

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 373 110**

51 Int. Cl.:  
**C12N 15/10** (2006.01)  
**C12Q 1/68** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **06013825 .2**  
96 Fecha de presentación: **14.01.1998**  
97 Número de publicación de la solicitud: **1712623**  
97 Fecha de publicación de la solicitud: **18.10.2006**

54 Título: **SELECCIÓN DE PROTEÍNAS USANDO FUSIONES DE ARN-PROTEÍNA.**

30 Prioridad:  
**21.01.1997 US 35963 P**  
**06.11.1997 US 64491 P**

45 Fecha de publicación de la mención BOPI:  
**31.01.2012**

45 Fecha de la publicación del folleto de la patente:  
**31.01.2012**

73 Titular/es:  
**THE GENERAL HOSPITAL CORPORATION**  
**55 FRUIT STREET**  
**BOSTON, MA 02114, US**

72 Inventor/es:  
**Szostak, Jack W.;**  
**Roberts, Richard W. y**  
**Liu, Rihe**

74 Agente: **Carpintero López, Mario**

ES 2 373 110 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

Selección de proteínas usando fusiones de ARN-proteína

Actualmente existen procedimientos para el aislamiento de moléculas de ARN y ADN basándose en sus funciones. Por ejemplo, los experimentos de Ellington y Szostak (Nature 346: 818 (1990); y Nature 355: 850 (1992)) y Tuerk y Gold (Science 249: 505 (1990); y J. Mol. Biol. 222: 739 (1991)) han demostrado que pueden aislarse moléculas de ácido nucleico muy poco frecuentes (es decir, menos de 1 en  $10^{13}$ ) con propiedades deseadas de combinaciones complejas de moléculas por rondas repetidas de selección y amplificación. Estos procedimientos ofrecen ventajas sobre las selecciones genéticas tradicionales en el sentido de que (i) pueden explorarse combinaciones de candidatos muy grandes ( $>10^{15}$ ), (ii) la viabilidad del huésped y las condiciones *in vivo* no suponen una preocupación y (iii) las selecciones pueden llevarse a cabo incluso si no existe una exploración genética *in vivo*. La potencia de la selección *in vitro* se ha demostrado en la definición de nuevas secuencias de ARN y ADN con funciones de unión a proteínas muy específicas (véase, por ejemplo, Tuerk y Gold, Science 249: 505 (1990); Irvine y col., J. Mol. Biol. 222: 739 (1991); Oliphant y col., Mol. Cell Biol. 9: 2944 (1989); Blackwell y col., Science 250: 1104 (1990); Pollock y Treisman, Nuc. Acids Res. 18: 6197 (1990); Thiesen y Bach, Nuc. Acids Res. 18: 3203 (1990); Bartel y col., Cell 57: 529 (1991); Stormo y Yoshioka, Proc. Natl. Acad. Sci. USA 88: 5699 (1991); y Bock y col., Nature 355: 564 (1992)), funciones de unión a moléculas pequeñas (Ellington y Szostak, Nature 346: 818 (1990); Ellington y Szostak, Nature 355: 850 (1992)) y funciones catalíticas (Green y col., Nature 347: 406 (1990); Robertson y Joyce, Nature 344: 467 (1990); Beaudry y Joyce, Science 257: 635 (1992); Bartel y Szostak, Science 261: 1411 (1993); Lorsch y Szostak, Nature 371: 31-36 (1994); Cuenoud y Szostak, Nature 375: 611-614 (1995); Chapman y Szostak, Chemistry and Biology 2: 325-333 (1995); y Lohse y Szostak, Nature 381: 442-444 (1996)). No se ha demostrado un esquema similar para la selección y amplificación de proteínas.

Se asume que una molécula de tipo virus que comprende un cofactor específico de una ribozima de replicación unida a su ARN genético ha sido un mediador en la coevolución de la replicación y la traducción. Esta molécula también se ha descrito como útil para la modificación por ingeniería genética molecular evolutiva usando el sistema de traducción sin células (Husimi y col., Progress in biophysics and molecular biology 65, supl.1, página 64, 1996).

Niemeyer y col. (Nucleic Acids Research 22(25): 5530-5539, 1994) describen moléculas híbridas en las que un ADN monocatenario corto se entrecruza con estreptavidina. Por lo tanto, estas moléculas híbridas semisintéticas comprenden un dominio de unión específico para ácidos nucleicos complementarios, así como los cuatro sitios de unión a biotina nativos de la estreptavidina. Esta bioespecificidad permite que las moléculas híbridas atraigan ácidos nucleicos complementarios y proteínas biotiniladas hacia un complejo y por lo tanto sirvan como conectores para la generación de bioconjugados supramoleculares y matrices macroscópicas.

**Sumario de la invención**

La invención se refiere a una microplaca que comprende una matriz de ácidos nucleicos monocatenarios inmovilizados, estando dichos ácidos nucleicos hibridados con fusiones de ARN-proteína, en la que cada una de dichas fusiones de ARN-proteína comprende un ARN unido covalentemente en el extremo 3' del ARN a una proteína a través de un aceptor peptídico, y en la que dicha proteína está codificada por dicho ARN.

La invención también se refiere a una biblioteca de moléculas, comprendiendo cada una un ácido ribonucleico unido covalentemente a una proteína codificada por dicho ácido ribonucleico.

La invención también se refiere a una molécula que comprende una porción de ácido ribonucleico (ARN) y a una porción de proteína codificada por dicha porción de ácido ribonucleico; en la que la porción de ácido ribonucleico está unida covalentemente a dicha porción de proteína a través de un aceptor peptídico, y en la que la porción de ácido ribonucleico está unida covalentemente en su extremo 3' a dicho aceptor peptídico a través de un sitio de pausa.

La invención también se refiere a una molécula que comprende un ácido ribonucleico unido covalentemente a un anticuerpo, en la que dicho anticuerpo está codificado en su totalidad por dicho ácido ribonucleico.

La invención también se refiere a un ácido ribonucleico que comprende una secuencia de inicio de la traducción y un codón de inicio unido operativamente a una secuencia codificante de proteína candidata, estando dicho ácido ribonucleico unido covalentemente a un aceptor peptídico en el extremo 3' de dicha secuencia codificante de proteína candidata, que comprende además una secuencia de ADN o análogo de ADN unida covalentemente al extremo 3' del ácido ribonucleico.

El propósito de la presente invención es permitir que los principios de la selección *in vitro* y de la evolución *in vitro* se apliquen a proteínas. La invención facilita el aislamiento de proteínas con propiedades deseadas a partir de grandes combinaciones de secuencias de aminoácidos parcialmente o totalmente aleatorias. Además, la invención resuelve el problema de recuperar y amplificar la información de secuencia proteica uniendo covalentemente la secuencia codificante de ARNm a la molécula de proteína.

En general, el método de la invención consiste en un protocolo de transcripción/traducción *in vitro* o *in situ* que

genera una proteína unida covalentemente al extremo 3' de su propio ARNm, es decir, una fusión de ARN-proteína. Esto se logra por síntesis y traducción *in vitro* o *in situ* de una molécula de ARNm con un aceptor peptídico unido a su extremo 3'. Un aceptor peptídico preferido es puromicina, un análogo nucleosídico que se añade al extremo C-terminal de una cadena peptídica en crecimiento y termina la traducción. En un diseño preferido, se incluye una

5 secuencia de ADN entre el extremo del mensaje y el aceptor peptídico que está diseñado para provocar que el ribosoma se detenga en el extremo de la fase de lectura abierta, proporcionando un tiempo adicional para que el aceptor peptídico (por ejemplo, puromicina) acepte la cadena peptídica naciente antes de la hidrólisis del enlace de peptidil-ARNt.

Si se desea, la fusión de ARN-proteína resultante permite rondas repetidas de selección y amplificación porque la información de secuencia proteica puede recuperarse por transcripción inversa y amplificación (por ejemplo, mediante amplificación por PCR, así como cualquier otra técnica de amplificación, incluyendo técnicas de amplificación basadas en ARN, tales como 3SR o TSA). Después, el ácido nucleico amplificado puede transcribirse, modificarse y traducirse *in vitro* o *in situ* para generar fusiones de ARNm-proteína para la siguiente ronda de selección. La capacidad para llevar a cabo múltiples rondas de selección y amplificación permite el enriquecimiento y

10 aislamiento de moléculas muy poco frecuentes, por ejemplo, una molécula deseada de una combinación de  $10^{15}$  miembros. Esto a su vez permite el aislamiento de proteínas nuevas o mejoradas que reconozcan específicamente prácticamente cualquier diana o que catalicen reacciones químicas deseadas.

Por consiguiente, una proteína deseada puede seleccionarse por un procedimiento que implique las etapas de: (a) proporcionar una población de moléculas de ARN candidatas, cada una de las cuales incluye una secuencia de inicio de la traducción y un codón de inicio unido operativamente a una secuencia codificante de proteína candidata, y cada una de las cuales está unida operativamente a un aceptor peptídico en el extremo 3' de la secuencia codificante de proteína candidata; (b) traducir *in vitro* o *in situ* las secuencias codificantes de proteínas candidatas para producir una población de fusiones de ARN-proteína candidatas; y (c) seleccionar una fusión de ARN-proteína deseada, seleccionando de este modo la proteína deseada.

20

Una molécula de ADN que codifica una proteína deseada puede seleccionarse mediante un procedimiento que implique las etapas de: (a) proporcionar una población de moléculas de ARN candidatas, cada una de las cuales incluye una secuencia de inicio de la traducción y un codón de inicio unido operativamente a una secuencia codificante de proteína candidata, y cada una de las cuales está unida operativamente a un aceptor peptídico en el extremo 3' de la secuencia codificante de proteína candidata; (b) traducir *in vitro* o *in situ* las secuencias codificantes de proteínas candidatas para producir una población de fusiones de ARN-proteína candidatas; (c) seleccionar una fusión de ARN-proteína deseada; y (d) generar a partir de la porción de ARN de la fusión una molécula de ADN que codifique la proteína deseada.

25

30

Una proteína que tenga una función alterada respecto a una proteína de referencia puede seleccionarse mediante un procedimiento que implique las etapas de: (a) producir una población de moléculas de ARN candidatas a partir de una población de moldes de ADN, teniendo cada uno de los moldes de ADN candidatos una secuencia codificante de proteína candidata que difiere de la secuencia codificante de la proteína de referencia, comprendiendo cada una de las moléculas de ARN una secuencia de inicio de la traducción y un codón de inicio unido operativamente a la secuencia codificante de proteína candidata, y estando cada una unida operativamente a un aceptor peptídico en el extremo 3'; (b) traducir *in vitro* o *in situ* las secuencias codificantes de proteínas candidatas para producir una población de fusiones de ARN-proteínas candidatas; y (c) seleccionar una fusión de ARN-proteína que tenga una función alterada, seleccionando de este modo la proteína que tiene la función alterada.

35

40

Una molécula de ADN que codifica una proteína que tiene una función alterada respecto a una proteína de referencia puede seleccionarse mediante un procedimiento que implique las etapas de: a) producir una población de moléculas de ARN candidatas a partir de una población de moldes de ADN candidatos, teniendo cada uno de los moldes de ADN candidatos una secuencia codificante de proteína candidata que difiere de la secuencia codificante de la proteína de referencia, comprendiendo cada una de las moléculas de ARN una secuencia de inicio de la traducción y un codón de inicio unido operativamente a la secuencia codificante de proteína candidata, y estando cada una unida operativamente a un aceptor peptídico en el extremo 3'; (b) traducir *in vitro* o *in situ* las secuencias codificantes de proteínas candidatas para producir una población de fusiones de ARN-proteína; y (c) seleccionar una fusión de ARN-proteína que tenga una función alterada; y (d) generar a partir de la porción de ARN de la fusión una molécula de ADN que codifique la proteína que tenga la función alterada.

45

50

Un ARN deseado puede seleccionarse mediante un procedimiento que implique las etapas de: (a) proporcionar una población de moléculas de ARN candidatas, cada una de las cuales incluye una secuencia de inicio de la traducción y un codón de inicio unido operativamente a una secuencia codificante de proteína candidata, y cada una de las cuales está unida operativamente a un aceptor peptídico en el extremo 3' de la secuencia codificante de proteína candidata; (b) traducir *in vitro* o *in situ* las secuencias codificantes de proteínas candidatas para producir una población de fusiones de ARN-proteínas candidatas; y (c) seleccionar una fusión de ARN-proteína deseada, seleccionando de este modo el ARN deseado.

55

En realizaciones preferidas de los procedimientos anteriores, el aceptor peptídico es puromicina; cada una de las moléculas de ARN candidatas incluye además una secuencia de pausa o incluye además una secuencia de ADN o

60

análogo de ADN unida covalentemente al extremo 3' del ARN; la población de moléculas de ARN candidatas incluye al menos  $10^9$ , preferentemente al menos  $10^{10}$ , más preferentemente al menos  $10^{11}$ ,  $10^{12}$  ó  $10^{13}$  y, más preferentemente, al menos  $10^{14}$  moléculas de ARN diferentes; la reacción de traducción *in vitro* se lleva a cabo en un lisado preparado a partir de una célula eucariota o porción de la misma (y por ejemplo se lleva a cabo en un lisado de reticulocitos o lisado de germen de trigo); la reacción de traducción *in vitro* se lleva a cabo en un extracto preparado a partir de una célula procariota (por ejemplo, *E. coli*) o porción de la misma; la etapa de selección implica la unión de la proteína deseada a un compañero de unión inmovilizado; la etapa de selección implica el ensayo para una actividad funcional de la proteína deseada; la molécula de ADN se amplifica; el procedimiento implica además repetir las etapas de los procedimientos de selección anteriores; el procedimiento implica además transcribir una molécula de ARN a partir de la molécula de ADN y repetir las etapas (a) a (d); después de la etapa de traducción *in vitro*, el procedimiento implica además una etapa de incubación llevada a cabo en presencia de  $Mg^{2+}$  50-100 mM; y la fusión de ARN-proteína incluye además una secuencia de ácido nucleico o de análogo de ácido nucleico situada proximal al aceptor peptídico que aumente la flexibilidad.

La invención presenta una fusión de ARN-proteína seleccionada por cualquiera de los procedimientos de la invención; un ácido ribonucleico unido covalentemente a través de un enlace amida a una secuencia de aminoácidos, estando la secuencia de aminoácidos codificada por el ácido ribonucleico; y un ácido ribonucleico que incluye una secuencia de inicio de la traducción y un codón de inicio unido operativamente a una secuencia codificante de proteína candidata, estando el ácido ribonucleico unido operativamente a un aceptor peptídico (por ejemplo, puromicina) en el extremo 3' de la secuencia codificante de proteína candidata.

Una proteína deseada o ARN deseado pueden seleccionarse por enriquecimiento de una combinación de secuencias. Este procedimiento implica las etapas de: (a) proporcionar una población de moléculas de ARN candidatas, cada una de las cuales incluye una secuencia de inicio de la traducción y un codón de inicio unido operativamente a una secuencia codificante de proteína candidata, y cada una de las cuales está unida operativamente a un aceptor peptídico en el extremo 3' de la secuencia codificante de proteína candidata; (b) traducir *in vitro* o *in situ* las secuencias codificantes de proteínas candidatas para producir una población de fusiones de ARN-proteínas candidatas; (c) poner en contacto la población de fusiones de ARN-proteínas con un compañero de unión específico para cualquiera de la porción de ARN o la porción de proteína de la fusión de ARN-proteína en condiciones que separen sustancialmente los complejos de compañero de unión-fusión de ARN-proteína de los miembros no unidos de la población; (d) liberar las fusiones de ARN-proteína unidas de los complejos; y (e) poner en contacto la población de fusiones de ARN-proteína de la etapa (d) con un compañero de unión específico para la porción de proteína de la fusión de ARN-proteína deseada en condiciones que separen sustancialmente el complejo de compañero de unión-fusión de ARN-proteína de los miembros no unidos de dicha población, seleccionando de este modo la proteína deseada y el ARN deseado.

En realizaciones preferidas, el procedimiento implica además repetir las etapas (a) a (e). Además, para estas etapas repetidas, pueden usarse los mismos o diferentes compañeros de unión, en cualquier orden, para el enriquecimiento selectivo de la fusión de ARN-proteína deseada. En otra realización preferida, la etapa (d) implica el uso de un compañero de unión (por ejemplo, un anticuerpo monoclonal) específico para la porción de proteína de la función deseada. Esta etapa se lleva a cabo preferentemente después de la transcripción inversa de la porción de ARN de la fusión para generar un ADN que codifique la proteína deseada. Si se desea, este ADN puede aislarse y/o amplificarse por PCR. Esta técnica de enriquecimiento puede usarse para seleccionar una proteína deseada o puede usarse para seleccionar una proteína que tenga una función alterada respecto a una proteína de referencia.

En otras realizaciones preferidas de los procedimientos de enriquecimiento, el aceptor peptídico es puromicina; cada una de las moléculas de ARN candidatas incluye además una secuencia de pausa o incluye además una secuencia de ADN o análogo de ADN unida covalentemente al extremo 3' del ARN; la población de moléculas de ARN candidatas incluye al menos  $10^9$ , preferentemente al menos  $10^{10}$ , más preferentemente al menos  $10^{11}$ ,  $10^{12}$  ó  $10^{13}$  y, más preferentemente al menos  $10^{14}$  moléculas de ARN diferentes; la reacción de traducción *in vitro* se lleva a cabo en un lisado preparado a partir de una célula eucariota o porción de la misma (y se lleva a cabo, por ejemplo, en un lisado de reticulocitos o lisado de germen de trigo); la reacción de traducción *in vitro* se lleva a cabo en un extracto preparado a partir de una célula procariota o porción de la misma (por ejemplo, *E. coli*); la molécula de ADN se amplifica; al menos uno de los compañeros de unión se inmoviliza en un soporte sólido; después de la etapa de traducción *in vitro*, el procedimiento implica además una etapa de incubación llevada a cabo en presencia de  $Mg^{2+}$  50-100 mM ; y la fusión de ARN-proteína incluye además una secuencia de ácido nucleico o análogo de ácido nucleico situada proximal al aceptor peptídico que aumente la flexibilidad.

Además, se describen en el presente documento kits para llevar a cabo cualquiera de los procedimientos de selección.

En un aspecto adicional, la invención presenta una microplaca que incluye una matriz de ácidos nucleicos monocatenarios inmovilizados, estando los ácidos nucleicos hibridados con fusiones de ARN-proteína. El componente proteico de la fusión de ARN-proteína está codificado por el ARN.

Como se usa en el presente documento, por una "población" se entiende más de una molécula (por ejemplo, más de una molécula de ARN, ADN o fusión de ARN-proteína). Debido a que los procedimientos de la invención facilitan

selecciones que comiencen, si se desea, con grandes cantidades de moléculas candidatas, una "población" de acuerdo con la invención significa preferentemente más de  $10^9$  moléculas, más preferentemente más de  $10^{11}$ ,  $10^{12}$  ó  $10^{13}$  moléculas y, más preferentemente, más de  $10^{13}$  moléculas.

5 Por "selección" se entiende separar sustancialmente una molécula de otras moléculas en una población. Como se usa en el presente documento, una etapa "de selección" proporciona al menos un enriquecimiento de 2 veces, preferentemente de 30 veces, más preferentemente de 100 veces y más preferentemente de 1000 veces de una molécula deseada respecto a moléculas no deseadas en una población después de la etapa de selección. Como se indica en el presente documento, una etapa de selección puede repetirse cualquier número de veces y pueden combinarse diferentes tipos de etapas de selección en una estrategia dada.

10 Por una "proteína" se entiende cualquiera de dos o más aminoácidos de origen natural o modificados unidos por uno o más enlaces peptídicos. Los términos "proteína" y "péptido" se usan indistintamente en el presente documento.

Por "ARN" se entiende una secuencia de dos o más ribonucleótidos de origen natural o modificados unidos covalentemente. Un ejemplo de un ARN modificado incluido dentro de este término es ARN de fosforotioato.

15 Por una "secuencia de inicio y la traducción" se entiende cualquier secuencia que sea capaz de proporcionar un sitio de entrada al ribosoma funcional. En sistemas bacterianos, estas regiones se denominan a veces secuencia de Shine-Dalgarno.

Por "codón de inicio" se entienden tres bases que señalizan el comienzo de una secuencia codificante de proteína. Generalmente estas bases son AUG (o ATG); sin embargo, cualquier otro triplete de bases capaz de utilizarse de esta forma puede sustituirse.

20 Por "unido covalentemente" a un aceptor peptídico se entiende que el aceptor peptídico se une a una "secuencia codificante de proteína" directamente a través de un enlace covalente o indirectamente a través de otra secuencia unida covalentemente (por ejemplo, ADN correspondiente a un sitio de pausa).

25 Por "aceptor peptídico" se entiende cualquier molécula capaz de añadirse al extremo C-terminal de una cadena proteica en crecimiento por la actividad catalítica de la función de la peptidil transferasa ribosómica. Típicamente dichas moléculas contienen (i) un nucleótido o resto de tipo nucleótido (por ejemplo, adenosina o un análogo de adenosina (la dimetilación en la posición N-6 amino es aceptable)), (ii) un aminoácido o resto de tipo aminoácido (por ejemplo, cualquiera de los 20 D- o L-aminoácidos o cualquier análogo de aminoácido de los mismos (por ejemplo, O-metil tirosina o cualquiera de los análogos descritos por Ellman y col., Meth. Enzymol. 202: 301, 1991), y (iii) un enlace entre los dos (por ejemplo, un enlace éster, amida o cetona en la posición 3' o, menos preferentemente, en la posición 2'); preferentemente, este enlace no altera significativamente el fruncido del anillo de la conformación de ribonucleótido natural. Los aceptores peptídicos también pueden poseer un nucleófilo, que puede ser, sin limitación, un grupo amino, un grupo hidroxilo o un grupo sulfhidrilo. Además, los aceptores peptídicos pueden estar compuestos por miméticos de nucleótidos, miméticos de aminoácidos o miméticos de la estructura de nucleótidos-aminoácidos combinada.

35 Por un aceptor peptídico que está situado "en el extremo 3'" de una secuencia codificante de proteína se entiende que la molécula aceptora peptídica está situada después del codón final de esa secuencia codificante de proteína. Este término incluye, sin limitación, una molécula aceptora peptídica que esté situada exactamente en el extremo 3' de la secuencia codificante de proteína, así como una que esté separada del codón final por una secuencia codificante o no codificante intermedia (por ejemplo, una secuencia correspondiente a un sitio de pausa). Este término también incluye construcciones en las que secuencias codificantes y no codificantes siguen a (es decir, están 3' respecto a) la molécula aceptora peptídica. Además, este término incluye, sin limitación, una molécula aceptora peptídica que esté unida covalentemente (directa o indirectamente a través de una secuencia de ácido nucleico intermedia) a la secuencia codificante de proteína, así como una que esté unida a la secuencia codificante de proteína por algunos medios no covalentes, por ejemplo, por hibridación usando una segunda secuencia de ácido nucleico que se una a o próxima al extremo 3' de la secuencia codificante de proteína y que por sí misma se una a una molécula aceptora peptídica.

Por una "función alterada" se entiende cualquier cambio cualitativo o cuantitativo en la función de una molécula.

Por "secuencia de pausa" se entiende una secuencia de ácido nucleico que provoca que un ribosoma ralentice o detenga su velocidad de traducción.

50 Por "compañero de unión", como se usa en el presente documento, se entiende cualquier molécula que tenga una afinidad covalente o no covalente específica por una porción de una fusión de ARN-proteína deseada. Los ejemplos de compañeros de unión incluyen, sin limitación, miembros de parejas de antígeno/anticuerpo, parejas de proteína/inhibidor, parejas de receptor/ligando (por ejemplo, parejas de receptor de superficie celular/ligando, tales como parejas de receptor de hormona/hormona peptídica), parejas de enzima/sustrato (por ejemplo, parejas de quinasa/sustrato), parejas de lectina/carbohidrato, agregados de proteínas oligoméricas o heterooligoméricos, parejas de proteína de unión ADN/sitio de unión a ADN, parejas de ARN/proteína y dúplex de ácido nucleico, heterodúplex o cadenas ligadas, así como cualquier otra molécula que sea capaz de formar uno o más enlaces

covalentes o no covalentes (por ejemplo, enlaces disulfuro) con cualquier porción de una fusión de ARN-proteína. Los compañeros de unión incluyen, sin limitación, cualquiera de los “motivos de selección” presentados en la Figura 2.

5 Por “soporte sólido” se entiende, sin limitación, cualquier columna (o material de columna), perla, tubo de ensayo, placa de microtitulación, partícula sólida (por ejemplo, agarosa o sepharose), microplaca (por ejemplo, microplaca de silicio, silicio-vidrio u oro) o membrana (por ejemplo, la membrana de un liposoma o vesícula) a la que puede unirse un complejo de afinidad, directa o indirectamente (por ejemplo, a través de otros intermedios compañeros de unión tales como otros anticuerpos o Proteína A) o en la que puede embeberse un complejo de afinidad (por ejemplo, a través de un receptor o canal).

10 Los procedimientos descritos en el presente documento proporcionan varias ventajas significativas. Para comenzar, es el primer ejemplo de este tipo de esquema para la selección y amplificación de proteínas. Esta técnica supera el punto muerto creado por la necesidad de recuperar secuencias de nucleótidos correspondientes a proteínas aisladas deseadas (puesto que sólo los ácidos nucleicos pueden replicarse). En particular, muchos procedimientos anteriores que permitían el aislamiento de proteínas a partir de combinaciones parcialmente o totalmente aleatorizadas lo hacían a través de una etapa *in vivo*. Los procedimientos de esta clase incluyen tecnología de anticuerpos monoclonales (Milstein, *Sci. Amer.* 243: 66 (1980); y Schultz y col., *J. Chem. Engng. News* 68: 26 (1990)), presentación en fago (Smith, *Science* 228: 1315 (1985); Parmley y Smith, *Gene* 73: 305 (1988); y McCafferty y col., *Nature* 348: 552 (1990)), fusiones de péptido-represor lac (Cull y col., *Proc. Natl. Acad. Sci. USA* 89: 1865 (1992)) y selecciones genéticas clásicas. A diferencia de la presente técnica, cada uno de estos procedimientos depende de un vínculo topológico entre la proteína y el ácido nucleico de modo que se retenga la información de la proteína y pueda recuperarse en forma legible de ácido nucleico.

25 Además, los procedimientos descritos en el presente documento proporcionan ventajas sobre el procedimiento de traducción parada (Tuerk y Gold, *Science* 249: 505 (1990); Irvine y col., *J. Mol. Biol* 222: 739 (1991); Korman y col., *Proc. Natl. Acad. Sci. USA* 79: 1844-1848 (1982); Mattheakis y col., *Proc. Natl. Acad. Sci. USA* 91: 9022-9026 (1994); Mattheakis y col., *Meth. Enzymol.* 267: 195 (1996); y Hanes y Pluckthun, *Proc. Natl. Acad. Sci. USA* 94:4937 (1997)), una técnica en la que la selección es por alguna propiedad de una cadena proteica naciente que todavía está formando un complejo con el ribosoma y su ARNm. A diferencia de la técnica de traducción parada, el presente procedimiento no depende del mantenimiento de la integridad de un complejo ternario de ARNm:ribosoma:cadena naciente, un complejo que es muy frágil y por lo tanto limitante con respecto a los tipos de selecciones que son técnicamente viables.

35 El presente procedimiento también proporciona ventajas sobre la estrategia de síntesis ramificada propuesta por Brenner y Lerner (*Proc. Natl. Acad. Sci. USA* 89: 5381-5383 (1992)), en la que se generan fusiones de ADN-péptido, y la información genética se recupera teóricamente después de una ronda de selección. A diferencia de la estrategia de síntesis ramificada, el presente procedimiento no requiere la regeneración de un péptido a partir de la porción de ADN de una fusión (que, en la estrategia de síntesis ramificada, se logra generalmente por rondas individuales de síntesis química). Por consiguiente, el presente procedimiento permite rondas repetidas de selección usando poblaciones de moléculas candidatas. Además, a diferencia de la técnica de síntesis ramificada, que generalmente está limitada a la selección de secuencias bastantes cortas, el presente procedimiento es aplicable a la selección de moléculas proteicas de una longitud considerable.

40 En otra ventaja más, la presente técnica de selección y evolución dirigida puede hacer uso de bibliotecas muy grandes y complejas de secuencias candidatas. Por el contrario, los procedimientos de selección de proteína existentes que dependen de una etapa *in vivo* se limitan típicamente a bibliotecas relativamente pequeñas de una complejidad algo limitada. Esta ventaja es particularmente importante cuando se seleccionan secuencias de proteínas funcionales considerando, por ejemplo, que existen  $10^{13}$  secuencias posibles para un péptido de sólo 10 aminoácidos de longitud. En técnicas genéticas clásicas, estrategias de fusión con represor lac y procedimientos de presentación en fago, las complejidades máximas generalmente caen a órdenes de magnitud por debajo de  $10^{13}$  miembros. Un gran tamaño de biblioteca también proporciona una ventaja para aplicaciones de evolución dirigidas, en el sentido de que puede explorarse el espacio de secuencia en una mayor profundidad alrededor de cualquier secuencia de partida dada.

50 La presente técnica también difiere de las técnicas anteriores en el sentido de que la etapa de selección es independiente del contexto. En muchos otros esquemas de selección, el contexto en el que, por ejemplo, está presente una proteína expresada puede influir profundamente en la naturaleza de la biblioteca generada. Por ejemplo, una proteína expresada puede no expresarse apropiadamente en un sistema particular o puede no presentarse apropiadamente (por ejemplo, en la superficie de una partícula de fago). Como alternativa, la expresión de una proteína puede en realidad interferir con una o más etapas críticas en un ciclo de selección, por ejemplo, la viabilidad o infectividad del fago o la unión de represor lac. Estos problemas pueden dar como resultado la pérdida de moléculas funcionales o limitaciones en la naturaleza de los procedimientos de selección que pueden aplicarse.

60 Por último, el presente procedimiento es ventajoso porque proporciona control sobre el repertorio de proteínas que pueden ensayarse. En ciertas técnicas (por ejemplo, selección de anticuerpos), existe escaso o ningún control sobre la naturaleza de la combinación de partida. En otras técnicas más (por ejemplo, fusiones con lac y presentación en

fago), la combinación candidata debe expresarse en el contexto de una proteína de fusión. Por el contrario, las construcciones de fusión de ARN-proteína proporcionan control sobre la naturaleza de las combinaciones de candidatos disponibles para exploración. Además, el tamaño de la combinación de candidatos tiene el potencial de ser tan elevado como combinaciones de ARN o ADN (~10<sup>15</sup> miembros), estando sólo limitado por el tamaño de la reacción de traducción *in vitro* realizada. Y la composición de la combinación de candidatos depende completamente del diseño experimental; pueden explorarse regiones aleatorias en aislamiento o dentro del contexto de una proteína de fusión deseada, y la mayoría, si no todas las posibles secuencias pueden expresarse en combinaciones de candidatos de fusiones de ARN-proteína.

Otras características y ventajas de la invención serán evidentes a partir de la descripción detallada siguiente y de las reivindicaciones.

### **Descripción detallada**

Los dibujos se describirán primero brevemente.

### **Breve descripción de los dibujos**

Las FIGURAS 1A-1C son representaciones esquemáticas de etapas implicadas en la producción de fusiones de ARN-proteína. La Figura 1A ilustra una construcción de ADN de muestra para la generación de una porción de ARN de una fusión. La Figura 1B ilustra la generación de un conjugado de ARN/puromicina. Y la Figura 1C ilustra la generación de una fusión de ARN-proteína.

La FIGURA 2 es una representación esquemática de un protocolo de selección generalizado de acuerdo con la invención.

La FIGURA 3 es una representación esquemática de un protocolo de síntesis para moldes de traducción mínima que contienen puromicina 3'. La etapa (A) muestra la adición de grupos protectores a los grupos funcionales reactivos en la puromicina (5'-OH y NH<sub>2</sub>); como modificados, estos grupos están convenientemente protegidos para su uso en la síntesis de oligonucleótidos basada en fosoramidita. La puromicina protegida estaba unida a un vidrio de tamaño de poro controlado (CPG) de aminoheptilo a través de el grupo OH 2' usando el protocolo convencional para unión de ADN a través de su OH 3' (Gait, *Oligonucleotide Synthesis, A Practical Approach*, The Practical Approach Series (IRL Press, Oxford, 1984)). En la etapa (B), se sintetizó un molde de traducción mínima (denominado "43-P"), que contenía 43 nucleótidos usando química de ARN y ADN convencional (Millipore, Bedford, MA), se desprotegió usando NH<sub>4</sub>OH y TBAF y se purificó en gel. El molde contenía 13 bases de ARN en el extremo 5', seguidas de 29 bases de ADN unidas a la puromicina 3' en su OH 5'. La secuencia de ARN contenía (i) una secuencia de consenso de Shine-Dalgarno complementaria a cinco bases de ARNr 16S (Stormo y col., *Nucleic Acids Research* 10: 2971-2996 (1982); Shine y Dalgarno, *Proc. Natl. Acad. Sci. USA* 71: 1342-1346 (1974); y Steitz y Jakes, *Proc. Natl. Acad. Sci. USA* 72: 4734-4738 (1975)), (ii) un espaciador de cinco bases y (iii) un solo codón de inicio AUG. La secuencia de ADN era dA<sub>27</sub>dCdCP, donde "P" es puromicina. La FIGURA 4 es una representación esquemática de un procedimiento preferido para la preparación de puromicina unida a CPG protegida.

La FIGURA 5 es una representación esquemática que muestra posibles modos de incorporación de metionina en un molde de la invención. Como se muestra en la reacción (A), el molde se une al ribosoma, permitiendo la formación del complejo de inicio 70S. El fmet ARNt se une al sitio P y sus bases se emparejan con el molde. La puromicina en el extremo 3' del molde entra en el sitio A de una forma intramolecular y forma un enlace amida con la N-formil metionina mediante el centro de peptidil transferasa, desacilando de este modo el ARNt. La extracción con fenol/cloroformo de la reacción produce el molde con metionina unida covalentemente. En la reacción (B) se muestra una reacción intramolecular no deseada del molde con oligonucleótidos que contienen puromicina. Como anteriormente, el molde mínimo estimula la formación del ribosoma 70S que contiene fmet ARNt unido al sitio P. Esto viene seguido por la entrada de un segundo molde en *trans* para dar una metionina unida covalentemente.

Las FIGURAS 6A-6H son fotografías que muestran la incorporación de <sup>35</sup>S metionina (<sup>35</sup>S met) en moldes de traducción. La Figura 6A demuestra la dependencia de magnesio (Mg<sup>2+</sup>) de la reacción. La Figura 6B demuestra la estabilidad a bases del producto; el cambio en la movilidad mostrado en esta figura corresponde a una pérdida de la secuencia de ARN 5' de 43-P (también denominado "molde Met") para producir la porción de ADN-puromicina, denominada 30-P. La retención del marcador después del tratamiento con bases concordaba con la formación de un enlace peptídico entre la <sup>35</sup>S metionina y la puromicina 3' del molde. La Figura 6C demuestra la inhibición de la formación de producto en presencia de inhibidores de peptidil transferasa. La Figura 6D demuestra la dependencia de la incorporación de <sup>35</sup>S metionina en una secuencia codificante de molde. La Figura 6E demuestra la dependencia de la longitud del molde de ADN de la incorporación de <sup>35</sup>S metionina. La Figura 6F ilustra la formación de producto *cis* frente a *trans* usando moldes 43-P y 25-P. La Figura 6G ilustra la formación de producto *cis* frente a *trans* usando moldes 43-P y 13-P. La Figura 6H ilustra la formación de producto *cis* frente a *trans* usando moldes 43-P y 30-P en un sistema de lisado de reticulocitos.

Las FIGURAS 7A-7C son ilustraciones esquemáticas de construcciones para ensayar la formación y selección de fusiones de péptidos. La Figura 7A muestra LP77 ("producto ligado", "77" nucleótidos de longitud) (también denominado "molde de myc corto") (SEC ID N°: 1). Esta secuencia contiene la etiqueta epitópica de anticuerpo monoclonal de c-myc EQKLISEEDL (SEC ID N°: 2) (Evan y col., *Mol. Cell Biol.* 5: 3610-3616 (1985)) flanqueada por un codón de inicio 5' y un engarce 3'. La región 5' contiene una secuencia de Shine-Dalgarno bacteriana

idéntica a la de 43-P. La secuencia codificante se optimizó para su traducción en sistemas bacterianos. En particular, las UTR 5' de 43-P y LP77 contenían una secuencia de Shine-Dalgarno complementaria a cinco bases del ARNr 16S (Steitz y Jakes, Proc. Natl. Acad. Sci. USA 72: 4734-4738 (1975)) y estaban separadas de forma similar a las secuencias proteicas ribosómicas (Stormo y col, Nucleic Acids Res. 10: 2971-2996 (1982)). La Figura 7B muestra LP154 (producto ligado, 154 nucleótidos de longitud) (también denominado "molde de myc largo") (SEC ID N°: 3). Esta secuencia contiene el código para la generación del péptido usado para aislar el anticuerpo de c-myc. El extremo 5' contiene una versión truncada de la secuencia cadena arriba de TMV (denominada "TE"). Esta UTR 5' contenía una secuencia de 22 nucleótidos derivada de la UTR 5' de TMV que incluía dos repeticiones directas ACAAUUAC (Gallie y col., Nucl. Acids Res. 16:883 (1988)). La Figura 7C muestra la Combinación n° 1 (SEC ID N°: 4), una secuencia ejemplar que se usará para selección de péptidos. Los siete aminoácidos finales del péptido myc original se incluyeron en el molde para servir como la región constante 3' necesaria para la amplificación por PCR del molde. Se sabe que esta secuencia no es parte del epítipo de unión a anticuerpo.

La FIGURA 8 es una fotografía que demuestra la síntesis de fusiones de ARN-proteína usando moldes 43-P, LP77 y LP154, y sistemas de traducción de reticulocitos ("Retic") y germen de trigo ("Trigo"). La mitad izquierda de la figura ilustra la incorporación de <sup>35</sup>S metionina en cada uno de los tres moldes. La mitad derecha de la figura ilustra los productos resultantes después del tratamiento con ARNasa A de cada uno de los tres moldes para eliminar la región codificante de ARN; se muestran fusiones de ADN-proteína marcadas con <sup>35</sup>S metionina. La porción de ADN de cada una era idéntica al oligonucleótido 30-P. Por lo tanto, las diferencias en la movilidad eran proporcionales a la longitud de las regiones codificantes, lo que concuerda con la existencia de proteínas de diferente longitud en cada caso.

La FIGURA 9 es una fotografía que demuestra la sensibilidad a proteasa de una fusión de ARN-proteína sintetizada a partir de LP154 y analizada por electroforesis en gel de poliacrilamida desnaturizante. El carril 1 contiene 30-P marcado con <sup>32</sup>P. Los carriles 2-4, 5-7 y 8-10 contienen los moldes de traducción marcados con <sup>35</sup>S recuperados de reacciones de lisado de reticulocitos sin tratamiento, con tratamiento con ARNasa A o con tratamiento con ARNasa A y proteinasa K, respectivamente.

La FIGURA 10 es una fotografía que muestra los resultados de reacciones de inmunoprecipitación usando una proteína con epítipo myc de 33 aminoácidos traducida *in vitro*. Los carriles 1 y 2 muestran los productos de traducción de la proteína con epítipo myc y moldes de β-globina, respectivamente. Los carriles 3-5 muestran los resultados de la inmunoprecipitación del péptido con epítipo myc usando un anticuerpo monoclonal de c-myc y los tampones de lavado PBS, DB y PBSTDS, respectivamente. Los carriles 6-8 muestran las mismas reacciones de inmunoprecipitación pero usando el producto de traducción de β-globina.

La FIGURA 11 es una fotografía que demuestra la inmunoprecipitación de una fusión de ARN-proteína a partir de una reacción de traducción *in vitro*. Se indican los picomoles de molde usados en la reacción. Los carriles 1-4 muestran ARN124 (la porción de ARN de la fusión LP154) y los carriles 5-7 muestran la fusión de ARN-proteína LP154. Después de la inmunoprecipitación usando un anticuerpo monoclonal de c-myc y proteína G-sepharose, las muestras se trataron con ARNasa A y polinucleótido quinasa de T4, después se cargaron en un gel de urea poliacrilamida desnaturizante para visualizar la fusión. En los carriles 1-4, con muestras que no contenían molde o que contenían solamente la porción de ARN del molde de myc largo (ARN124), no se observó fusión. En los carriles 5-7, se visualizaron claramente bandas que correspondían a la fusión. Se indica la posición de 30-P marcado con <sup>32</sup>P y la cantidad de molde introducido se indica en la parte superior de la figura.

La FIGURA 12 es una gráfica que muestra una cuantificación de material de fusión obtenido a partir de una reacción de traducción *in vitro*. La intensidad de las bandas de fusión mostradas en los carriles 5-7 de la Figura 11 y la banda 30-P (aislada de una forma paralela en dT<sub>25</sub>, no se muestra) se cuantificó en placas de Phosphorimager y se representó en función de la concentración de LP154 introducida. El 30-P modificado recuperado (eje y izquierdo) era linealmente proporcional al molde introducido (eje x), mientras que la fusión de engarce-péptido (eje y derecho) era constante. A partir de este análisis, se calculó que se formaban ~10<sup>12</sup> fusiones por ml de muestra de reacción de traducción.

La FIGURA 13 es una representación esquemática de tiopropil sepharose y dT<sub>25</sub> agarosa, y de la capacidad de estos sustratos para interactuar con las fusiones de ARN-proteína de la invención.

La FIGURA 14 es una fotografía que muestra los resultados del aislamiento secuencial de fusiones de la invención. El carril 1 contiene 30-P marcado con <sup>32</sup>P. Los carriles 2 y 3 muestran LP154 aislado de reacciones de traducción y tratado con ARNasa A. En el carril 2, LP154 se aisló de forma secuencial, usando tiopropil sepharose seguida de dT<sub>25</sub> agarosa. El carril 3 muestra el aislamiento usando solamente dT<sub>25</sub> agarosa. Los resultados indicaban que el producto contenía un tiol libre, probablemente la penúltima cisteína en la secuencia codificante del epítipo myc.

Las FIGURAS 15A y 15B son fotografías que muestran la formación de productos de fusión usando moldes de β-globina como se ensayó por SDS-tricina-PAGE (electroforesis en gel de poliacrilamida). La Figura 15A muestra la incorporación de <sup>35</sup>S sin usar molde (carril 1), usando un molde de syn-β-globina (carriles 2-4) o un molde de LP-β-globina (carriles 5-7). La Figura 15B (carriles marcados como en la Fig. 15A) muestra material marcado con <sup>35</sup>S aislado por cromatografía de afinidad de oligonucleótidos. No se aisló material en ausencia de una cola de 30-P (carriles 2-4).

Las FIGURAS 16A-16C son diagramas y fotografías que ilustran el enriquecimiento de ADNbc de myc frente a ADNbc de combinación por selección *in vitro*. La Figura 16A es un esquema del protocolo de selección. Se tradujeron cuatro mezclas de los moldes de myc y combinación *in vitro* y se aislaron en dT<sub>25</sub> agarosa seguida de



TP sepharose para purificar las fusiones de moldes a partir de moldes no modificados. Las fusiones de ARNm-péptido se sometieron después a transcripción inversa para suprimir cualquier estructura secundaria o terciaria presente en los moldes. Se retiraron alícuotas de cada mezcla tanto antes (Figura 16B) como después (Figura 16C) de la selección por afinidad, se amplificaron por PCR en presencia de un cebador marcado y se digirieron con una enzima de restricción que escindía solamente el ADN de myc. Las mezclas introducidas de moldes eran myc puro (carril 1) o myc:combinación de 1:20, 1:200 ó 1:2000 (carriles 2-4). El material no seleccionado se desvió de las proporciones introducidas debido a la traducción y transcripción inversa preferentes del molde de myc. El enriquecimiento del molde de myc durante la etapa selectiva se calculó a partir del cambio en la proporción de combinación:myc antes y después de la selección.

La FIGURA 17 es una fotografía que ilustra la traducción de moldes de ARN de myc. Se usaron los engarces siguientes: carriles 1-4, dA<sub>27</sub>dCdCP; carriles 5-8, dA<sub>27</sub>rCrCP; y carriles 9-12, dA<sub>21</sub>C<sub>9</sub>C<sub>9</sub>dAdCdCP. En cada carril, la concentración de molde de ARN era de 600 nM, y se usó <sup>35</sup>S-Met para el marcaje. Las condiciones de reacción eran las siguientes: carriles 1, 5 y 9, 30°C durante 1 hora; carriles 2, 6 y 10, 30°C durante 2 horas; carriles 3, 7 y 11, 30°C durante 1 hora, -20°C durante 16 horas; y carriles 4, 8 y 12, 30°C durante 1 hora, -20°C durante 16 horas con Mg<sup>2+</sup> 50 mM. En esta Figura, "A" representa péptido libre y "B" representa fusión de ARNm-péptido.

La FIGURA 18 es una fotografía que ilustra la traducción de moldes de ARN de myc marcados con <sup>32</sup>P. El engarce utilizado era dA<sub>21</sub>C<sub>9</sub>C<sub>9</sub>dAdCdCP. La traducción se realizó a 30°C durante 90 minutos y las incubaciones se llevaron a cabo a -20°C durante 2 días sin Mg<sup>2+</sup> adicional. Las concentraciones de molde de ARNm eran de 400 nM (carril 3), 200 nM (carril 4), 100 nM (carril 5) y 100 nM (carril 6). El carril 1 muestra la fusión de ARNm-péptido marcada con <sup>35</sup>S-Met. El carril 2 muestra ARNm marcado con <sup>32</sup>P. En el carril 6, la reacción se llevó a cabo en presencia de análogo de caperuza 0,5 mM.

La FIGURA 19 es una fotografía que ilustra la traducción de molde de ARN de myc usando un lisado obtenido de Ambion (carril 1), Novagen (carril 2) y Amersham (carril 3). El engarce utilizado era dA<sub>27</sub>dCdCP. La concentración del molde era de 600 nM y se usó <sup>35</sup>S-Met para el marcaje. Las traducciones se realizaron a 30°C durante 1 hora y las incubaciones se llevaron a cabo a -20°C durante una noche en presencia de Mg<sup>2+</sup> 50 mM.

En el presente documento se describe un procedimiento general para la selección de proteínas con funciones deseadas usando fusiones en las que estas proteínas se unen covalentemente a sus propios ARN mensajeros. Estas fusiones de ARN-proteína se sintetizan por traducción *in vitro* o *in situ* de combinaciones de ARNm que contienen un aceptor peptídico unido a sus extremos 3' (Figura 1B). En una realización preferida, después de la ultralectura de la fase de lectura abierta del mensaje, el ribosoma se detiene cuando alcanza el sitio de pausa diseñado, y el resto aceptor ocupa el sitio A ribosómico y acepta la cadena peptídica nascente del peptidil-ARNt en el sitio P para generar la fusión de ARN-proteína (Figura 1C). La unión covalente entre la proteína y el ARN (en forma de un enlace amida entre el extremo 3' del ARNm y el extremo C-terminal de la proteína que codifica) permite que se recupere y se amplifique la información genética en la proteína (por ejemplo, por PCR) después de la selección por transcripción inversa del ARN. Una vez que se genera la fusión, la selección o enriquecimiento se lleva a cabo basándose en las propiedades de la fusión de ARNm-proteína o, como alternativa, puede llevarse a cabo una transcripción inversa usando el molde de ARNm mientras está unido a la proteína para evitar cualquier efecto del ARN monocatenario sobre la selección. Cuando se usa la construcción de ARNm-proteína, las fusiones seleccionadas pueden ensayarse para determinar qué resto (la proteína, el ARN o ambos) proporciona la función deseada.

En una realización preferida, la puromicina (que se parece a la tirosil adenosina) actúa como aceptor para unir el péptido en crecimiento a su ARNm. La puromicina es un antibiótico que actúa terminando la elongación peptídica. Como un mimético de aminoacil-ARNt, actúa como inhibidor universal de la síntesis de proteína por unión al sitio A, aceptación de la cadena peptídica en crecimiento y desprendimiento del ribosoma (a una K<sub>d</sub> = 10<sup>4</sup> M) (Traut y Monro, J. Mol. Biol. 10: 63 (1964); Smith y col., J. Mol. Biol. 13: 617 (1965)). Una de las características más atractivas de la puromicina es el hecho de que forma un enlace amida estable con la cadena peptídica en crecimiento, permitiendo de este modo fusiones más estables que los aceptores potenciales que forman enlaces éster inestables. En particular, la molécula de peptidil-puromicina contiene un enlace amida estable entre el péptido y la porción O-metil tirosina de la puromicina. La O-metil tirosina se une a su vez mediante un enlace amida estable al grupo amino 3' de la porción de adenosina modificada de la puromicina.

Otras elecciones posibles de aceptores incluyen estructuras de tipo ARNt en el extremo 3' del ARNm, así como otros compuestos que actúan de una forma similar a la puromicina. Dichos compuestos incluyen, sin limitación, cualquier compuesto que posea un aminoácido unido a una adenina o un compuesto de tipo adenina, tal como los nucleótidos de aminoácidos fenilalanil-adenosina (A-Phe), tirosil adenosina (A-Tyr) y alanil adenosina (A-Ala), así como estructuras unidas por amida tales como fenilalanil 3' desoxi 3' amino adenosina, alanil 3' desoxi 3' amino adenosina y tirosil 3' desoxi 3' amino adenosina; en cualquiera de estos compuestos, puede utilizarse cualquiera de los L-aminoácidos de origen natural o sus análogos. Además, también puede usarse en la invención un conjugado de estructura 3' de tipo ARNt-puromicina combinado.

En la Figura 2 se muestra un esquema de selección preferido de acuerdo con la invención. Las etapas implicadas en esta selección se llevan a cabo generalmente de la forma siguiente.

**Etapa 1. Preparación del molde de ADN.** Como etapa hacia la generación de las fusiones de ARN-proteína de la invención, se sintetiza la porción de ARN de la fusión. Esto puede efectuarse por síntesis química directa de ARN o, más comúnmente, se efectúa por transcripción de un molde de ADN bicatenario apropiado.

Dichos moldes de ADN pueden crearse por cualquier técnica convencional (incluyendo cualquier técnica de tecnología de ADN recombinante, síntesis química o ambas). En principio, cualquier procedimiento que permita la producción de uno o más moldes que contienen una secuencia conocida, aleatoria, aleatorizada o mutada puede usarse para este fin. En una estrategia particular, se sintetiza un oligonucleótido (por ejemplo, que contiene bases aleatorias) y se amplifica (por ejemplo, por PCR) antes de la transcripción. También puede usarse síntesis química para producir un casete aleatorio que después se inserta en la mitad de una secuencia codificante de proteína conocida (véase, por ejemplo, el capítulo 8.2, Ausubel y col., *Current Protocols in Molecular Biology*, John Wiley & Sons y Greene Publishing Company, 1994). Esta última estrategia produce una alta densidad de mutaciones alrededor de un sitio específico de interés en la proteína.

Una alternativa a la aleatorización total de una secuencia de molde de ADN es la aleatorización parcial, y una combinación sintetizada de este modo se denomina generalmente combinación "adulterada". Un ejemplo de esta técnica, realizada en una secuencia de ARN, se describe, por ejemplo, por Eklund y col. (*Nucl. Acids Research* 23: 3231 (1995)). Puede realizarse una aleatorización parcial por derivación de las reacciones de síntesis de modo que cada mezcla de reacción de adición de bases contenga un exceso de una base y pequeñas cantidades de cada una de las otras; mediante el control cuidadoso de las concentraciones de bases, puede conseguirse una frecuencia de mutación deseada mediante esta estrategia. También pueden generarse combinaciones parcialmente aleatorizadas usando técnicas de PCR propensa a errores, por ejemplo, como se describe en Beaudry y Joyce (*Science* 257: 635 (1992)) y Bartel y Szostak (*Science* 261: 1411 (1993)).

También están disponibles numerosos procedimientos para generar una construcción de ADN que comience con una secuencia conocida y después crear una combinación de ADN mutado. Se describen ejemplos de dichas técnicas en Ausubel y col. (anteriormente, capítulo 8) y Sambrook y col. (*Molecular Cloning: A Laboratory Manual*, capítulo 15, Cold Spring Harbor Press, Nueva York, 2ª ed. (1989)). También pueden generarse secuencias aleatorias mediante la técnica de "transposición" resumida en Stemmer (*Nature* 370: 389 (1994)).

Para optimizar un esquema de selección de la invención, las secuencias y estructuras en los extremos 5' y 3' de un molde también pueden alterarse. Preferentemente, esto se lleva a cabo en dos selecciones separadas, implicando cada una la inserción de dominios aleatorios en el molde proximal al extremo apropiado, seguido de selección. Estas elecciones pueden servir (i) para maximizar la cantidad de fusión generada (y por lo tanto maximizar la complejidad de una biblioteca) o (ii) para proporcionar secuencias de traducción optimizadas. Además, el procedimiento puede ser generalmente aplicable, combinado con PCR mutagénica, a la optimización de moldes de traducción tanto en las regiones codificantes como no codificantes.

**Etapa 2. Generación de ARN.** Como se ha señalado anteriormente, la porción de ARN de una fusión de ARN-proteína puede sintetizarse químicamente usando técnicas convencionales de síntesis de oligonucleótidos. Como alternativa, y particularmente si se utilizan secuencias de ARN más largas, la porción de ARN se genera por transcripción *in vitro* de un molde de ADN. En una estrategia preferida, se usa polimerasa de T7 para generar enzimáticamente la cadena de ARN. Otras ARN polimerasas apropiadas para este uso incluyen, sin limitación, las ARN polimerasas de SP6, T3 y *E. coli* (descritas, por ejemplo, en Ausubel y col. (anteriormente, capítulo 3). Además, el ARN sintetizado puede ser, en su totalidad o en parte, ARN modificado. En un ejemplo particular, puede producirse ARN de fosforotioato (por ejemplo, por transcripción con T7) usando ribonucleótidos modificados y técnicas convencionales. Dicho ARN modificado proporciona la ventaja de ser estable a nucleasas.

**Etapa 3. Ligación de puromicina con el molde.** A continuación, la puromicina (o cualquier otro aceptor peptídico apropiado) se une covalentemente a la secuencia de molde. Esta etapa puede efectuarse usando ARN ligasa de T4 para unir la puromicina directamente a la secuencia de ARN, o preferentemente la puromicina puede unirse por medio de un "puente" de ADN usando ADN ligasa de T4 o cualquier otra enzima que sea capaz de unir entre sí dos secuencias de nucleótidos (véase la Figura 1B) (véase también, por ejemplo, Ausubel y col., anteriormente, capítulo 3, secciones 14 y 15). También pueden usarse ARNt sintetasas para unir compuestos de tipo puromicina a ARN. Por ejemplo, la fenilalanil ARNt sintetasa une la fenilalanina a moléculas de fenilalanil-ARNt que contienen un grupo amino 3', generando moléculas de ARN con extremos 3' de tipo puromicina (Fraser y Rich, *Proc. Natl. Acad. Sci. USA* 70: 2671 (1973)). Otros aceptores peptídicos que pueden usarse incluyen, sin limitación, cualquier compuesto que posea un aminoácido unido a una adenina o compuesto de tipo adenina, tal como los nucleótidos de aminoácidos, fenilalanil-adenosina (A-Phe), tirosil adenosina (A-Tyr) y alanil adenosina (A-Ala), así como estructuras ligadas a amida, tales como fenilalanil 3' desoxi 3' amino adenosina, alanil 3' desoxi 3' amino adenosina; y tirosil 3' desoxi 3' amino adenosina; en cualquiera de estos compuestos puede utilizarse cualquiera de los L-aminoácidos de origen natural o sus análogos. Se describen varios aceptores peptídicos, por ejemplo, en Krayevsky y Kukhanova, *Progress in Nucleic Acids Research and Molecular Biology* 23: 1 (1979).

**Etapa 4. Generación y recuperación de las fusiones de ARN-proteína.** Para generar fusiones de ARN-proteína puede utilizarse cualquier sistema de traducción *in vitro* o *in situ*. Como se muestra a continuación, se prefieren sistemas eucariotas, y dos sistemas particularmente preferidos incluyen los sistemas de germen de trigo y lisado de

reticulocitos. En principio, sin embargo, cualquier sistema de traducción que permita la formación de una fusión de ARN-proteína y que no degrade significativamente la porción de ARN de la fusión es útil en la invención. Además, para reducir la degradación de ARN en cualquiera de estos sistemas, pueden incluirse oligonucleótidos antisentido de bloqueo de la degradación en la mezcla de la reacción de traducción; dichos oligonucleótidos hibridan específicamente con y abarcan secuencias dentro de la porción de ARN de la molécula que desencadenan la degradación (véase, por ejemplo, Hanes y Pluckthun, Proc. Natl. Acad. Sci USA 94: 4937 (1997)).

Como se ha señalado anteriormente, están disponibles cualquier cantidad de sistemas de traducción eucariotas para su uso en la invención. Estos incluyen, sin limitación, lisados de levadura, ascitis, células tumorales (Leibowitz y col., Meth. Enzymol. 194: 536 (1991)), y huevos de ovocitos de *Xenopus*. Los sistemas de traducción *in vitro* útiles de sistemas bacterianos incluyen, sin limitación, los descritos en Zubay (Ann. Rev. Genet. 7: 267 (1973)); Chen y Zubay (Meth. Enzymol. 101: 44 (1983)); y Ellman (Meth. Enzymol. 202: 301 (1991)).

Además, pueden llevarse a cabo reacciones de traducción *in situ*. En un ejemplo particular, la traducción puede llevarse a cabo por inyección de ARNm en huevos de *Xenopus* usando técnicas convencionales.

Una vez generadas, las fusiones de ARN-proteína pueden recuperarse de la mezcla de reacción de traducción mediante cualquier técnica convencional de purificación de ARN o proteínas. Típicamente, se utilizan técnicas de purificación de proteínas. Como se muestra a continuación, por ejemplo, la purificación de una fusión puede facilitarse mediante el uso de reactivos cromatográficos adecuados tales como dT<sub>25</sub> agarosa o tiopropil sepharose. La purificación, sin embargo, también, o como alternativa, puede implicar una purificación basada en la porción de ARN de la fusión; se describen técnicas para dicha purificación, por ejemplo, en Ausubel y col. (anteriormente, capítulo 4).

Etapa 5. Selección de la fusión de ARN-proteína deseada. La selección de una fusión de ARN-proteína deseada puede efectuarse por cualquier medio disponible para separar o aislar selectivamente una fusión deseada de una población de fusiones candidatas. Los ejemplos de técnicas de aislamiento incluyen, sin limitación, unión selectiva, por ejemplo, a un compañero de unión que esté directa o indirectamente inmovilizado en una columna, perla, membrana u otro soporte sólido, e inmunoprecipitación usando un anticuerpo específico para el resto proteico de la fusión. La primera de estas técnicas hace uso de un motivo de selección inmovilizado que puede consistir en cualquier tipo de molécula con el que es posible la unión. Se presenta una lista de posibles moléculas de motivos de selección en la Figura 2. La selección también puede basarse en el uso de moléculas de sustrato unidas a un marcador de afinidad (por ejemplo, sustrato-biotina) que reaccione con una molécula candidata, o en cualquier otro tipo de interacción con una molécula de fusión. Además, pueden seleccionarse proteínas basándose en su actividad catalítica de una forma análoga a la descrita por Bartel y Szostak para el aislamiento de enzimas de ARN (anteriormente); de acuerdo con esa técnica particular, las moléculas deseadas se seleccionan basándose en su capacidad para unir una molécula diana consigo mismas, y las moléculas funcionales se aíslan después basándose en la presencia de esa diana. Se permiten mediante la presente invención esquemas de selección para aislar proteínas catalíticas nuevas o mejoradas usando esta misma estrategia o cualquier otra selección funcional.

Además, como se describe en el presente documento, la selección de una fusión de ARN-proteína deseada (o su copia de ADN) puede facilitarle por enriquecimiento a favor de esa fusión en una combinación de moléculas candidatas. Para llevar a cabo dicho enriquecimiento opcional, se pone en contacto una población de fusiones de ARN-proteína candidatas con un compañero de unión (por ejemplo, uno de los compañeros de unión descritos anteriormente) que es específico para la porción de ARN o la porción de proteína de la fusión, en condiciones que separen sustancialmente el complejo de compañero de unión-fusión de miembros no unidos en la muestra. Esta etapa puede repetirse y la técnica incluye preferentemente al menos dos etapas de enriquecimiento secuenciales, una en la que las fusiones se seleccionan usando un compañero de unión específico para la porción de ARN y otra en la que las fusiones se seleccionan usando un compañero de unión específico para la porción de proteína. Además, si se repiten las etapas de enriquecimiento que se dirigen a la misma porción de la fusión (por ejemplo, la porción de proteína) se utilizan preferentemente diferentes compañeros de unión. En un ejemplo particular descrito en el presente documento, una población de moléculas se enriquece para fusiones deseadas usando primero un compañero de unión específico para la función de ARN de la fusión y, después, en dos etapas secuenciales, usando dos compañeros de unión diferentes, siendo ambos específicos para la porción de proteína de la fusión. De nuevo, estos complejos pueden separarse de los componentes de muestra mediante cualquier técnica de separación convencional incluyendo, sin limitación, cromatografía de afinidad en columna, centrifugación o inmunoprecipitación.

Además, la elución de una fusión de ARN-proteína a partir de un complejo de enriquecimiento (o selección) puede efectuarse mediante varias estrategias. Por ejemplo, como se describe en el presente documento, se puede utilizar una etapa de elución química inespecífica o desnaturizante para aislar una fusión de ARN-proteína deseada. Dicha etapa facilita la liberación de componentes del complejo entre sí o de un soporte sólido asociado de una forma relativamente inespecífica por rotura de enlaces no covalentes entre los componentes y/o entre los componentes y el soporte sólido. Como se describe en el presente documento, un reactivo de elución química inespecífica o desnaturizante ejemplar es HOAC al 4%/H<sub>2</sub>O. Otros reactivos de elución química inespecífica o desnaturizante incluyen guanidina, urea, alta concentración salina, detergente o cualquier otro medio por el que puedan eliminarse generalmente aductos no covalentes. Como alternativa, se puede utilizar una estrategia de elución química específica, en la que se aproveche un agente químico que provoque la liberación específica de una molécula de

fusión. En un ejemplo particular, si el brazo de engarce de una proteína de fusión deseada contiene uno o más de enlaces disulfuro, los aptámeros de fusión unidos pueden eluirse por adición, por ejemplo, de DTT, dando como resultado la reducción del enlace disulfuro y la liberación de la diana unida.

5 Como alternativa, la elución puede efectuarse alterando específicamente complejos de afinidad; dichas técnicas liberan selectivamente componentes de complejo por adición de un exceso de un miembro del complejo. Por ejemplo, en una selección de unión a ATP, la elución se realiza por adición de ATP en exceso a la mezcla de incubación. Por último, se puede llevar a cabo una etapa de elución enzimática. Mediante esta estrategia, una molécula unida de por sí o una proteasa añadida exógenamente (u otra enzima hidrolítica apropiada) se escinde y libera la diana o la enzima. En un ejemplo particular, un sitio diana de proteasa puede incluirse en cualquiera de los componentes del complejo y las moléculas unidas eluirse por adición de la proteasa. Como alternativa, en una selección catalítica, puede usarse elución como etapa de selección para aislar moléculas capaces de liberarse (por ejemplo, escindirse) por sí mismas de un soporte sólido.

15 Etapa 6. Generación de una copia de ADN de la secuencia de ARN usando transcriptasa inversa. Si se desea, una copia de ADN de una secuencia de fusión de ARN seleccionada está fácilmente disponible por transcripción inversa de esa secuencia de ARN usando cualquier técnica convencional (por ejemplo, usando transcriptasa inversa Superscript). Esta etapa puede llevarse a cabo antes de la etapa de selección o enriquecimiento (por ejemplo, como se describe en la Figura 16) o después de esa etapa. Como alternativa, el procedimiento de transcripción inversa puede llevarse a cabo antes del aislamiento de la fusión de la mezcla de traducción *in vitro* o *in situ*.

20 A continuación, se amplifica el molde de ADN, como una secuencia bicatenaria parcial o de longitud completa. Preferentemente, en esta etapa, se generan moldes de ADN de longitud completa usando oligonucleótidos apropiados y amplificación por PCR.

Estas etapas, y los reactivos y técnicas para llevar a cabo estas etapas, se describen ahora en detalle usando ejemplos particulares. Estos ejemplos se proporcionan con el fin de ilustrar la invención y no deben interpretarse como limitantes.

## 25 **Generación de moldes para fusiones de ARN-proteína**

Como se muestra en las Figuras 1A y 2, el esquema de selección de los procedimientos descritos en el presente documento hace uso preferentemente de moldes de ADN bicatenarios que incluyen varios elementos de diseño. El primero de estos elementos es un promotor que se usará junto con una ARN polimerasa deseada para síntesis de ARNm. Como se muestra en la Figura 1A y se describe en el presente documento, se prefiere el promotor de T7, aunque puede usarse cualquier promotor capaz de dirigir la síntesis de un ADN bicatenario lineal.

30 El segundo elemento del molde mostrado en la Figura 1A se denomina la región no traducida 5' (o UTR 5') y corresponde al ARN cadena arriba del sitio de inicio de la traducción. En la Figura 1A se muestra una UTR 5' preferida (denominada "TE") que es una mutante de delección de la región no traducida 5' del Virus del Mosaico del Tabaco y, en particular, corresponde a las bases directamente 5' del inicio de traducción de TMV; la secuencia de esta UTR es la siguiente: rGrGrG rArCrA rArUrU rArCrU rArUrU rUrArC rArArU rUrArC rA (insertándose los primeros 3 nucleótidos G para aumentar la transcripción) (SEC ID N°: 5). Puede utilizarse cualquier otra UTR 5' apropiada (véase, por ejemplo, Kozak, Microbiol. Rev. 47: 1 (1983)).

40 El tercer elemento mostrado en la Figura 1A es el sitio de inicio de la traducción. En general, éste es un codón AUG. Sin embargo, hay ejemplos en los que se utilizan codones distintos de AUG en secuencias codificantes de origen natural, y estos codones también pueden usarse en el esquema de selección de la invención.

El cuarto elemento en la Figura 1A es la fase de lectura abierta de la proteína (denominada ORF), que codifica la secuencia proteica. Esta fase de lectura abierta puede codificar cualquier secuencia proteica de origen natural, aleatoria, aleatorizada, mutada o totalmente sintética.

45 El quinto elemento mostrado en la Figura 1A es la región constante 3'. Esta secuencia facilita la amplificación por PCR de las secuencias de combinación y ligación del oligonucleótido que contiene puromicina al ARNm. Si se desea, esta región también puede incluir un sitio de pausa, una secuencia que causa que el ribosoma se detenga y de este modo permite tiempo adicional para que un resto aceptor (por ejemplo, puromicina) acepte una cadena peptídica naciente del peptidil-ARNt; este sitio de pausa se analiza en más detalle a continuación.

50 Para desarrollar la presente metodología, se generaron inicialmente fusiones de ARN-proteína usando moldes de ARN altamente simplificados que contenían 1-2 codones. Esta estrategia se adoptó por dos razones. En primer lugar, podían generarse fácilmente moldes de este tamaño mediante síntesis química. Y en segundo lugar, una fase de lectura abierta pequeña permitía que se ensayaran fácilmente características críticas de la reacción, incluyendo eficacia de enlace, heterogeneidad de extremos, dependencia de molde y exactitud de traducción.

55 Diseño de construcción. Se usó una construcción básica para generar fusiones de ARN-proteína de ensayo. La molécula consistía en un ARNm que contenía una secuencia de Shine-Dalgarno (SD) para inicio de la traducción que contenía una delección de 3 bases de la secuencia SD de la proteína ribosómica L1 y que era complementaria a

5 bases del ARNr 16S (es decir, rGrGrA rGrGrA rCrGrA rA) (SEC ID N°: 6) (Stormo y col., Nucleic Acids Research 10: 2971-2996 (1982); Shine y Dalgarno, Proc. Natl. Acad. Sci. USA 71: 1342-1346 (1974); y Steitz y Jakes, Proc. Natl. Acad. Sci. USA 72: 4734-4738 (1975)), (ii) un codón de inicio AUG, (iii) un engarce de ADN para actuar como sitio de pausa (es decir, 5'-(dA)<sub>27</sub>, (iv) dCdC-3' y (v) una puromicina 3' (P). La secuencia poli dA se seleccionó porque se sabía que sirve mal de molde al ARNt en el sitio A (Morgan y col., J. Mol. Biol. 26: 477-497 (1967); Ricker y Kaji, Nucleic Acid Research 19: 6573-6578 (1991)) y estaba diseñada para actuar como un buen sitio de pausa. La longitud del engarce de oligo dA se seleccionó para abarcar la distancia de ~60-70 Å entre el sitio de descodificación y el centro de transferencia de peptidilo del ribosoma. El dCdCP mimetizaba el extremo CCA de un ARNt y estaba diseñado para facilitar la unión de la puromicina al sitio A del ribosoma.

10 Síntesis química de molde mínimo 43-P. Para sintetizar la construcción 43-P (mostrada en la Figura 3), la puromicina se unió primero a un soporte sólido de tal modo que sería compatible con química de síntesis de oligonucleótidos de fosforamidita convencional. El protocolo de síntesis para este oligonucleótido se resume esquemáticamente en la Figura 3 y se describe en más detalle a continuación. Para unir puromicina a un soporte sólido de vidrio de tamaño de poro controlado (CPG), el grupo amino se protegió con un grupo trifluoroacetilo como se ha descrito en el Boletín de Usuario de Applied Biosystems n° 49 para el modelo de sintetizador de ADN 380 (1988). A continuación, la protección del OH 5' se llevó a cabo usando una estrategia de DMT-CI convencional (Gait, Oligonucleotide Synthesis a practical approach The Practical Approach Series (IRL Press, Oxford, 1984)), y la unión a aminohexil CPG a través del OH 2' se efectuó exactamente de la misma forma que se usaría el OH 3' para la unión de un desoxinucleósido (véase la Fig. 3 y Gait, anteriormente, pág. 47). La puromicina protegida unida a 5' DMT-CPG era entonces adecuada para la extensión de la cadena con monómeros de fosforamidita. La síntesis del oligo avanzaba en la dirección 3' → 5' en el orden: (i) puromicina 3', (ii) pdCpdC, (iii) ~27 unidades de dA como engarce, (iv) AUG y (v) la secuencia de Shine-Dalgarno. La secuencia de la construcción 43-P se muestra a continuación.

25 Síntesis de CPG puromicina. La síntesis de CPG puromicina protegida siguió la ruta general usada para desoxinucleósidos como se ha resumido anteriormente (Gait, Oligonucleotide Synthesis, A Practical Approach, The Practical Approach Series (IRL Press, Oxford, 1984)). Las salidas principales incluían la selección de un grupo de N bloqueo apropiado, la unión en el OH 2' al soporte sólido y la reacción de enlace con el soporte sólido. En el caso de esta última, la reacción se llevó a cabo a concentraciones muy pequeñas de nucleótidos activados, ya que este material era significativamente más precioso que el soporte sólido. El rendimiento resultante (~20 µmol/g soporte sólido) era bastante satisfactorio considerando las condiciones de reacción diluidas.

35 Síntesis de N-trifluoroacetil puromicina. Se convirtieron primero 267 mg (0,490 mmol) de puromicina\*HCl en la forma de fase libre por disolución en agua, adición de tampón carbonato pH 11 y extracción (3X) en cloroformo. La fase orgánica se evaporó hasta sequedad y se pesó (242 mg, 0,513 mmol). La base libre se disolvió después en 11 ml de piridina seca y 11 ml de acetonitrilo seco, y se añadieron 139 µl (2,0 mmol) de trietilamina (TEA) y 139 µl (1,0 mmol) de anhídrido trifluoroacético (TFAA) con agitación. Después se añadió el TFAA a la solución turbia en alícuotas de 20 µl hasta que no quedaba nada de material de partida, según se ensayó por cromatografía en capa fina (ccf) (cloroformo/MeOH 93:7) (un total de 280 µl). La reacción se dejó que se desarrollara durante una hora. En ese punto, se pusieron de manifiesto dos bandas por cromatografía en capa fina, ambas de mayor movilidad que el material de partida. La preparación de la reacción con NH<sub>4</sub>OH y agua reducía el producto a una sola banda. La cromatografía en sílice (cloroformo/MeOH 93:7) producía 293 mg (0,515 mmol) del producto, N-TFA-Pur. El producto de esta reacción se muestra esquemáticamente en la Figura 4.

45 Síntesis de N-trifluoroacetil 5'-DMT puromicina. El producto de la reacción anterior se dividió en alícuotas y se coevaporó 2X con piridina seca para eliminar el agua. Se prepararon múltiples tubos para ensayar múltiples condiciones de reacción. En una reacción a pequeña escala, se disolvieron 27,4 mg (4,82 µmoles) de N-TFAF-Pur en 480 µl de piridina que contenía 0,05 eq de DMAP y 1,4 eq de TEA. A esta mezcla se añadieron 20,6 mg de cloruro de tritilo (60 µmol) y la reacción se dejó que se desarrollara hasta completarse con agitación. La reacción se detuvo por adición de un volumen equivalente de agua (aproximadamente 500 µl) a la solución. Debido a que esta reacción parecía tener éxito, se realizó una versión a gran escala. En particular, se disolvieron 262 mg (0,467 mmol) de N-TFA-Pur en 2,4 de piridina, seguido de adición de 1,4 eq de TEA, 0,05 eq de DMAP y 1,2 eq de cloruro de tritilo. Después de aproximadamente dos horas, se añadieron 50 mg adicionales (0,3 eq) de dimetoxitritil\*Cl (DMT\*Cl) y la reacción se dejó que se desarrollara durante 20 minutos adicionales. La reacción se detuvo por adición de 3 ml de agua y se coevaporó 3X con CH<sub>3</sub>CN. La reacción se purificó mediante cloroformo/MeOH 95:5 en una columna de sílice (seco) de 100 ml de 2 mm de diámetro. Debido a la purificación incompleta, se procesó una segunda columna idéntica con cloroformo/MeOH 97,5:2,5. El rendimiento total fue de 325 mg o 0,373 mmol (o un rendimiento del 72%). El producto de esta reacción se muestra esquemáticamente en la Figura 4.

60 Síntesis de N-trifluoroacetil 5'-DMT, 2'-succinil puromicina. En una reacción a pequeña escala, se combinaron 32 mg (37 µmol) del producto sintetizado anteriormente con 1,2 eq de DMAP disuelto en 350 µl de piridina. A esta solución, se añadieron 1,2 equivalentes de anhídrido succínico en 44 µl de CH<sub>3</sub>CN seco y se dejó agitando durante una noche. La cromatografía en capa fina puso de manifiesto que quedaba poco material de partida. En una reacción a gran escala, se combinaron 292 mg (336 µmol) del producto previo con 1,2 eq de DMAP en 3 ml de piridina. A esto, se añadieron 403 µl de anhídrido succínico 1 M en CH<sub>3</sub>CN seco y la mezcla se dejó agitando durante una noche. La

5 cromatografía en capa fina puso de manifiesto de nuevo que quedaba poco del material de partida. Las dos reacciones se combinaron, y se añadieron 0,2 eq adicionales de DMAP y succinato. El producto se coevaporó con tolueno 1X y se secó hasta una espuma amarilla en alto vacío. Se añadió CH<sub>2</sub>Cl<sub>2</sub> (20 ml) y esta solución se extrajo dos veces con 15 ml de ácido cítrico helado al 10% y después dos veces con agua pura. El producto se secó, se redisolvió en 2 ml de CH<sub>2</sub>Cl<sub>2</sub> y se precipitó por adición de 50 ml de hexano con agitación. Después, el producto se agitó vorticialmente y se centrifugó a 600 rpm durante 10 minutos en la centrifuga clínica. La mayoría del eluyente se retiró y el resto del producto se secó, primero a bajo vacío, después alto vacío en un desecador. El rendimiento de esta reacción era de aproximadamente 260 μmol para un rendimiento por etapas de ~70%.

10 Síntesis de N-trifluoroacetil 5'-DMT, 2'-succinilo CPG puomicina. El producto de la etapa previa se disolvió a continuación con 1 ml de dioxano seguido de 0,2 ml de dioxano/0,2 ml de piridina. A esta solución se añadieron 40 mg de p-nitrofenol y 140 mg de dicitclohexilcarbodiimida (DCC), y se dejó que la reacción se desarrollara durante 2 horas. La ciclohexil urea insoluble producida por la reacción se retiró por centrifugación y la solución de producto se añadió a 5 g de vidrio de tamaño controlado (CPG) de aminohehilo suspendido en 22 ml de DMF seco y se agitó durante una noche. La resina se lavó después con DMF, metanol y éter y se secó. La resina resultante se ensayó que contenía 22,6 μmol de trilito por g, bien dentro del intervalo aceptable para este tipo de soporte. Después el soporte se protegió terminalmente por incubación con 15 ml de piridina, 1 ml de anhídrido acético y 60 mg de DMAP durante 30 minutos. El material de columna resultante producía un ensayo de ninhidrina negativo (sin color), al contrario que los resultados obtenidos antes del bloqueo, en los que el material producía una reacción de color azul oscuro. El producto de esta reacción se muestra esquemáticamente en la Figura 4.

20 Síntesis de conjugado de ARNm puomicina. Como se ha analizado anteriormente, un oligonucleótido unido a puomicina puede usarse en cualquiera de dos formas para generar un conjugado de ARNm-puomicina que actúe como molde de traducción. Para fases de lectura abierta extremadamente cortas, típicamente el oligonucleótido de puomicina se extiende químicamente con monómeros de ARN o ADN para crear un molde totalmente sintético. Cuando se desean fases de lectura abierta más largas, el oligonucleótido de ARN o ADN se liga generalmente al extremo 3' de un ARNm usando un puente de ADN y ADN ligasa de T4 como se ha descrito por Moore y Sharp (Science 256: 992 (1992)).

#### Traducción y ensayo *in vitro* de fusiones de ARN-proteína

Los moldes generados anteriormente se tradujeron *in vitro* usando sistemas de traducción *in vitro* tanto bacterianos como eucariotas.

30 Traducción *in vitro* de moldes mínimos. Se añadieron 43-P y conjugados de ARN-puomicina relacionados a varios sistemas de traducción *in vitro* diferentes incluyendo: (i) el sistema S30 derivado de *E. coli* MRE600 (Zubay, Ann. Rev. Genet. 7: 267 (1973); Collins, Gene 6: 29 (1979); Chen y Zubay, Methods Enzymol, 101: 44 (1983); Pratt, en Transcription and Translation: A Practical Approach, B. D. Hammes, S. J. Higgins, Eds. (IRL Press, Oxford, 1984) págs. 179-209; y Ellman y col., Methods Enzymol. 202: 301 (1991)) preparado como se describe por Ellman y col. (Methods Enzymol. 202: 301 (1991)); (ii) la fracción ribosómica derivada de la misma cepa, preparada como se describe por Kudlicki y col. (Anal. Chem. 206: 389 (1992)); y (iii) el sistema S30 derivado de *E. coli* BL21, preparado como se describe por Lesley y col. (J. Biol. Chem. 266: 2632 (1991)). En cada caso, la premezcla usada era la de Lesley y col. (J. Biol. Chem. 266: 2632 (1991)) y las incubaciones eran de 30 minutos de duración.

40 Ensayo de la naturaleza de la fusión. El molde 43-P se ensayó primero usando extractos de traducción de S30 de *E. coli*. La Figura 5 (Reacción "A") demuestra la reacción intramolecular (*cis*) deseada en la que 43-P se une al ribosoma y actúa como molde y un aceptor de fMet al mismo tiempo. La incorporación de <sup>35</sup>S-metionina y su posición en el molde se ensayaron primero, y los resultados se muestran en las Figuras 6A y 6B. Después de la extracción de la mezcla de reacción de traducción *in vitro* con fenol/cloroformo y del análisis de los productos por SDS-PAGE, aparecía una banda marcada con <sup>35</sup>S con la misma movilidad que el molde 43-P. La cantidad de este material sintetizado dependía de la concentración de Mg<sup>2+</sup> (Figura 6A). La concentración óptima de Mg<sup>2+</sup> parecía estar entre 9 y 18 mM, lo que era similar al óptimo para la traducción en este sistema (Zubay, Ann. Rev. Genet. 7: 267 (1973); Collins, Gene 6: 29 (1979); Chen y Zubay, Methods Enzymol, 101: 44 (1983); Pratt, en Transcription and Translation: A Practical Approach, B. D. Hammes, S. J. Higgins, Eds. (IRL Press, Oxford, 1984) págs 179-209; Ellman y col., Methods Enzymol. 202: 301 (1991); Kudlicki y col., Anal. Chem. 206: 389 (1992); y Lesley y col., J. Biol. Chem. 266: 2632 (1991)). Además, el marcador incorporado era estable al tratamiento con NH<sub>4</sub>OH (Figura 6B), indicando que el marcador estaba localizado en la mitad 3' de la molécula (la porción de ADN estable a bases) y estaba unido mediante un enlace estable a bases, según se esperaba para un enlace amida entre puomicina y fMet.

55 Dependencia de ribosoma y molde. Para demostrar que la reacción observada anteriormente se producía en el ribosoma, se ensayaron los efectos de inhibidores específicos de la función peptidil transferasa del ribosoma (Figura 6C), y se examinó el efecto de cambiar la secuencia que codifica metionina (Figura 6D). La Figura 6C demuestra claramente que la reacción se inhibía fuertemente por los inhibidores de peptidil transferasa, virginiamicina, gougerotina y cloranfenicol (Monro y Vazquez, J. Mol. Biol. 28: 161-165 (1967); y Vazquez y Monro, Biochemica et Biophysical Acta 142: 155-173 (1967)). La Figura 6D demuestra que cambiar una sola base en el molde de A a C suprimía la incorporación de <sup>35</sup>S metionina a Mg<sup>2+</sup> 9 mM y la disminuía enormemente a 18 mM (lo que concuerda

con el hecho de que altos niveles de  $Mg^{2+}$  permiten una lectura errónea del mensaje). Estos experimentos demostraron que la reacción se producía en el ribosoma de una forma dependiente del molde.

Longitud de engarce. También se ensayó la dependencia de la reacción de la longitud del engarce (Figura 6E). El molde original estaba diseñado de modo que el engarce abarcaba la distancia desde el sitio de descodificación (ocupado por el AUG del molde) hasta el sitio aceptor (ocupado por el resto de puromicina), una distancia que era aproximadamente la misma longitud que la distancia entre el bucle anticodón y el tallo aceptor en una ARNt, o de aproximadamente 60-70 Å. El primer engarce ensayado era de 30 nucleótidos de longitud, basándose en un mínimo de 3,4 Å por base ( $\geq 102$  Å). En el intervalo entre 30 y 21 nucleótidos ( $n = 27-18$ ; longitud  $\geq 102-71$  Å) se observaban escasos cambios en la eficacia de la reacción. Por consiguiente, puede variarse la longitud del engarce. Aunque un engarce de entre 21 y 30 nucleótidos representa una longitud preferida, también pueden utilizarse en la invención engarces más cortos que de 80 nucleótidos y, preferentemente, más cortos que de 45 nucleótidos.

Reacciones intramoleculares frente a intermoleculares. Por último, los presentes inventores ensayaron si la reacción se producía de una forma intramolecular (Figura 5, Reacción "A") según se desea o intermolecularmente (Figura 5, Reacción "B"). Esto se ensayó por adición de oligonucleótidos con puromicina 3' pero sin secuencia de unión a ribosoma (es decir, los moldes 25-P, 13-P y 30-P) a las reacciones de traducción que contenían el molde 43-P (Figuras 6F, 6G y 6H). Si la reacción se producía mediante un mecanismo intermolecular, también se marcarían los oligonucleótidos más cortos. Como se muestra en las Figuras 6F-H, había escasa incorporación de  $^{35}S$  metionina en los tres oligonucleótidos más cortos, indicando que la reacción se producía principalmente de una forma intramolecular. Las secuencias de 25-P (SEC ID N°: 10), 13-P (SEC ID N°: 9) y 30-P (SEC ID N°: 8) se muestran a continuación.

Lisado de reticulocitos. La Figura 6H demuestra que la  $^{35}S$ -metionina puede incorporarse en el molde 43-P usando un lisado de reticulocitos de conejo (véase a continuación) para la traducción *in vitro*, además de los lisados de *E. coli* usados anteriormente. Esta reacción se producía principalmente en un mecanismo intramolecular, según se desee.

#### **Síntesis y ensayo de fusiones que contienen una etiqueta epitópica c-myc**

También se generaron fusiones ejemplares que contenían, dentro de la porción de proteína, la etiqueta epitópica para el anticuerpo monoclonal de c-myc 9E10 (Evan y col., Mol Cell Biol. 5: 3610 (1985)).

Diseño de moldes. Se diseñaron tres moldes con etiqueta epitópica iniciales (es decir, LP77, LP154 y Combinación N° 1) y se muestran en las Figuras 7A-C. Los primeros dos moldes contenían la secuencia de etiqueta epitópica c-myc EQKLISEEDL (SEC ID N°: 2) y el tercer molde era el diseño usado en la síntesis de una combinación de selección aleatoria. El LP77 codificaba una secuencia de 12 aminoácidos, con los codones optimizados para la traducción bacteriana. El LP 154 y sus derivados contenían una secuencia de ARNm de 33 aminoácidos en la que los codones estaban optimizados para la traducción eucariota. La secuencia de aminoácidos codificada de MAEEQKLISEEDLLRKRREQKLKHKLEQLRNSCA (SEC ID N°: 7) correspondía al péptido original usado para aislar el anticuerpo 9E10. La Combinación N° 1 contenía 27 codones de NNG/C (para generar péptidos aleatorios) seguidos por una secuencia correspondiente a los últimos siete aminoácidos del péptido myc (que no eran parte de la secuencia de epitopo myc). Estas secuencias se muestran a continuación.

Sistemas de traducción *in vitro* de reticulocitos frente a germen de trigo. Los moldes 43-P, LP77 y LP 154 se ensayaron en sistemas de traducción tanto de reticulocitos de conejo como de extracto de germen de trigo (Promega, Boehringer Mannheim) (Figura 8). Las traducciones se realizaron a 30°C durante 60 minutos. Los moldes se aislaron usando dT<sub>25</sub> agarosa a 4°C. Los moldes se eluyeron de la agarosa usando NaOH 15 mM, EDTA 1 mM, se neutralizaron con tampón NaOAc/HOAc, se precipitaron inmediatamente con etanol (2,5-3 vol), se lavaron (con etanol al 100%) y se secaron en un concentrador Speedvac. La Figura 8 muestra que la  $^{35}S$  metionina se incorporó en los tres moldes, en los sistemas tanto de germen de trigo como de reticulocitos. Se observó menos degradación del molde en las reacciones de fusión del sistema de reticulocitos y, por consiguiente, este sistema se prefiere para la generación de fusiones de ARN-proteína. Además, en general, se prefieren sistemas eucariotas sobre sistemas bacterianos. Debido a que las células eucariotas tienden a contener menores niveles de nucleasas, las vidas útiles del ARNm son generalmente 10-100 veces más prolongadas en estas células que en células bacterianas. En experimentos usando un sistema de traducción de *E. coli* particular, no se observó generación de fusiones usando un molde que codifica el epitopo c-myc; el marcaje del molde en diversos sitios demostró que esto se debía probablemente a la degradación de porciones tanto de ARN como de ADN del molde.

Para examinar la porción peptídica de estas fusiones, las muestras se trataron con ARNasa para eliminar las secuencias codificantes. Después de este tratamiento, el producto 43-P corría con una movilidad casi idéntica a la del oligonucleótido 30-P marcado con  $^{32}P$ , lo que concuerda con un péptido muy pequeño (quizá sólo metionina) añadido a 30-P. Para LP77, la eliminación de la secuencia codificante producía un producto con menor movilidad que el oligonucleótido 30-P, lo que concuerda con la noción de que se añadía un péptido de 12 aminoácidos a la puromicina. Por último, para LP154, la eliminación de la secuencia codificante producía un producto de todavía menor movilidad, lo que concuerda con una secuencia de 33 aminoácidos unida al oligonucleótido 30-P. No se observó oligonucleótido en el carril de reticulocitos de LP154 tratado con ARNasa debido a un error de carga. En la

Figura 9, se demostró que la movilidad de este producto era la misma que el producto generado en el extracto de germen de trigo. En suma, estos resultados indicaban que se añadieron productos resistentes a ARNasa a los extremos de los oligonucleótidos 30-P, que los tamaños de los productos eran proporcionales a la longitud de las secuencias codificantes, y que los productos eran bastante homogéneos en tamaño. Además, aunque ambos sistemas producían productos de fusión similares, el sistema de reticulocitos parecía superior debido a la mayor estabilidad del molde.

Sensibilidad a ARNasa A y proteinasa K. En la Figura 9, se ensayó la sensibilidad a ARNasa A y proteinasa K usando la fusión LP154. Como se muestra en los carriles 2-4, la incorporación de <sup>35</sup>S metionina se demostró para el molde LP154. Cuando este producto se trató con ARNasa A, disminuyó la movilidad de la fusión, pero todavía era significativamente superior a la del oligonucleótido 30-P marcado con <sup>32</sup>P, lo que concuerda con la adición de un péptido de 33 aminoácidos en el extremo 3'. Cuando este material se trató también con proteinasa K, la señal de <sup>35</sup>S desaparecía completamente, lo que concuerda de nuevo con la noción de que el marcador estaba presente en un péptido en el extremo 3' del fragmento 30-P. Se han obtenido resultados similares en experimentos equivalentes usando las fusiones 43-P y LP77.

Para confirmar que el marcaje de molde mediante <sup>35</sup>S Met era una consecuencia de la traducción y, más específicamente, que se obtenía como resultado de la actividad peptidil transferasa del ribosoma, se examinó el efecto de diversos inhibidores sobre la reacción de marcaje. Los inhibidores específicos de peptidil transferasa eucariota, anisomicina, gougerotina y esparsomicina (Vazquez, *Inhibitors of Protein Biosynthesis* (Springer Verlag, Nueva York), págs. 312 (1979)), así como los inhibidores de la translocación cicloheximida y emetina (Vazquez, *Inhibitors of Protein Biosynthesis* (Springer Verlag, Nueva York), pág. 312 (1979)) disminuían todos la formación de fusión de ARN-péptido en ~95% usando el molde de myc largo y un extracto de traducción de lisado de reticulocitos.

Experimentos de inmunoprecipitación. En un experimento diseñado para ilustrar la eficacia de inmunoprecipitar una fusión de ARNm-péptido, se realizó un intento para inmunoprecipitar un péptido c-myc libre generado por traducción *in vitro*. La Figura 10 muestra los resultados de estos experimentos ensayados en un gel de péptidos SDS-PAGE. Los carriles 1 y 2 muestran el material marcado de reacciones de traducción que contenían ARN124 (la porción de ARN de LP154) o ARNm de β-globina. Los carriles 3-8 muestran la inmunoprecipitación de estas muestras de reacción usando el anticuerpo monoclonal de c-myc 9E10 en varias condiciones de tampón diferentes (descritas a continuación). Los carriles 3-5 muestran que el péptido derivado de ARN124 se inmunoprecipitaba eficazmente, siendo el mejor caso el carril 4, en el que se aislaron ~83% de los recuentos precipitables por TCA totales. Los carriles 6-8 muestran poca de la proteína β-globina, indicando una purificación de >100 veces. Estos resultados indicaban que el péptido codificado por el ARN124 (y por LP154) puede aislarse cuantitativamente mediante este protocolo de inmunoprecipitación.

Inmunoprecipitación de la fusión. Los presentes inventores ensayaron a continuación la capacidad para inmunoprecipitar un producto de ARN-péptido químico usando una reacción de traducción de LP154 y el anticuerpo monoclonal c-myc 9E10 (Figura 11). Los productos de traducción de una reacción de reticulocitos se aislaron por inmunoprecipitación (como se describe en el presente documento) y se trataron con 1 μg de ARNasa A a temperatura ambiente durante 30 minutos para eliminar la secuencia codificante. Esto generó un OH 5' que ese marcó con <sup>32</sup>P con polinucleótido quinasa de T4 y se ensayó mediante PAGE desnaturalizante. La Figura 11 demuestra que se aislaba un producto con una movilidad similar a la observada para la fusión del epítipo c-myc con 30-P generado por tratamiento con ARNasa de la fusión LP154 (véase anteriormente), pero no se generaba ningún producto correspondiente cuando sólo se traducía la porción de ARN del molde (ARN124). En la Figura 12, se determinó la cantidad de proteína de fusión aislada y se representó frente a la cantidad de 30-P no modificado (no se muestra en esta figura). La cuantificación de la proporción de engarce no modificado respecto a la fusión de engarce-péptido myc muestra que el 0,2-0,7% del mensaje introducido se convertía en producto de fusión. Una mayor fracción del ARN introducido se convertía en producto de fusión en presencia de una mayor proporción de ribosoma/molde; por encima del intervalo de concentraciones de ARNm introducido que se ensayaron, se generaban aproximadamente  $0,8-1,0 \times 10^{12}$  moléculas de fusión por ml de extracto de traducción.

Además, los resultados de los presentes inventores indicaban que los péptidos unidos a las especies de ARN estaban codificados por ese ARNm, es decir, el péptido naciente no se transfería a la puomicina de algún otro ARNm. No se observó ningún indicio de transferencia cruzada cuando un engarce (30-P) se coincubaba con el molde de myc largo en extractos de traducción en proporciones tan elevadas como de 20:1, ni la presencia de engarce libre disminuía significativamente la cantidad de fusión con myc largo producida. De forma similar, la cotraducción de los moldes corto y largo, 43-P y LP154, producía solamente los productos de fusión observados cuando los moldes se traducían en solitario y no se observaron productos de movilidad intermedia, como se esperaba para la fusión del molde corto con el péptido de myc largo. Ambos resultados sugerían que la formación de fusión se producía principalmente entre un péptido naciente y ARNm unido al mismo ribosoma.

Aislamiento secuencial. Como confirmación adicional de la naturaleza del producto de molde LP154 traducido *in vitro*, los presentes inventores examinaron el comportamiento de este producto en dos tipos diferentes de medios de cromatografía. La tiopropil (TP) sepharose permite el aislamiento de un producto que contiene una cisteína libre (por ejemplo, el producto LP154 que tiene un resto de cisteína adyacente al extremo C-terminal) (Figura 13). De forma similar, la dT<sub>25</sub> agarosa permite el aislamiento de moldes que contienen una secuencia de poli dA (por ejemplo, 30-



P) (Figura 13). La Figura 14 demuestra que el aislamiento secuencial en TP sepharose seguido de dT<sub>25</sub> agarosa producía el mismo producto que el aislamiento en dT<sub>25</sub> agarosa solamente. El hecho de que el producto de traducción *in vitro* contuviera tanto un tracto de poli-A como un tiol libre indicaba claramente que el producto de traducción era la fusión de ARN-péptido deseada.

- 5 Los resultados anteriores concuerdan con la capacidad para sintetizar fusiones de ARNm-péptido y para recuperarlas intactas a partir de extractos de traducción *in vitro*. Las porciones peptídicas de fusiones así sintetizadas parecían tener las secuencias deseadas según se demostró por inmunoprecipitación y aislamiento usando técnicas cromatográficas apropiadas. De acuerdo con los resultados presentados anteriormente, las reacciones son intramoleculares y se producen de una forma dependiente de molde. Por último, incluso con una  
10 modificación de molde de menos del 1%, el presente sistema facilita selecciones basadas en complejidades de candidatos de aproximadamente 10<sup>13</sup> moléculas.

Selección de recuperación de epítipo c-myc. Para seleccionar epítopos c-myc adicionales, se genera una gran biblioteca de moldes de traducción (por ejemplo, 10<sup>11</sup> miembros) que contienen una región aleatorizada (véase la Figura 7C y a continuación). Esta biblioteca se usa para generar ~10<sup>12</sup>-10<sup>13</sup> fusiones (como se describe en el  
15 presente documento) que se tratan con el anticuerpo anti-c-myc (por ejemplo, por inmunoprecipitación o usando un anticuerpo inmovilizado en una columna u otro soporte sólido) para enriquecer para moldes que codifiquen c-myc en rondas repetidas de selección *in vitro*.

Modelos para formación de fusiones. Sin desear quedar ligado a teoría particular alguna, los presentes inventores proponen un modelo para el mecanismo de formación de fusiones en el que la traducción se inicia normalmente y la elongación avanza hasta el final de la fase de lectura abierta. Cuando el ribosoma alcanza la porción de ADN del molde, la traducción se detiene. En este punto, el complejo puede repartirse entre dos destinos: disociación del péptido naciente o transferencia del péptido naciente a la puromicina en el extremo 3' del molde. Es probable que la eficacia de la reacción de transferencia esté controlada por varios factores que influyen en la estabilidad del complejo de traducción detenido y la entrada del resto de puromicina 3' en el sitio A del centro de peptidil transferasa. Después de la reacción de transferencia, la fusión de ARNm-péptido probablemente permanece formando un complejo con el ribosoma puesto que los factores de liberación conocidos no pueden hidrolizar el enlace amida estable entre los dominios de ARN y péptido.  
20  
25

Tanto el modelo clásico para elongación (Watson, Bull. Soc Chim. Biol. 46:1399 (1964)) como el modelo de estados intermedios (Moazed y Noller, Nature 342:142 (1989)) requieren que el sitio A esté vacío para la entrada de la puromicina en el centro de peptidil transferasa. Para que la puromicina entre en el sitio A vacío, el engarce debe formar un bucle alrededor de la parte externa del ribosoma o pasar directamente desde el sitio de descodificación a través del sitio A hasta el centro de peptidil transferasa. Los datos descritos en el presente documento no distinguen claramente entre estas alternativas porque el engarce más corto ensayado (21 nucleótidos) es todavía lo bastante largo para pasar alrededor de la parte externa del ribosoma. En algunos modelos de estructura de ribosoma (Frank y col, Nature 376: 441 (1995)), el ARNm se enhebra a través de un canal que se extiende en cualquier lado del sitio de descodificación, en cuyo caso sería necesario el desenhebrado del engarce del canal para permitir que la puromicina alcance el centro de peptidil transferasa a través del sitio A.  
30  
35

La transferencia del péptido naciente a la puromicina parecía ser lenta respecto al procedimiento de elongación según se demostró por la homogeneidad y longitud del péptido unido al engarce. Si la puromicina competía eficazmente con los aminoacil ARNt durante la elongación, se esperaría que las fusiones de engarce-péptido presentes en los productos de fusión fueran de un tamaño heterogéneo. Además, el ribosoma no parecía leer en la región de engarce, según se indicaba por la similitud en las moviidades de gel entre la fusión de Met-molde y el engarce no modificado. dA<sub>3n</sub> debería codificar (lisina)<sub>n</sub>, que con toda certeza disminuiría la movilidad del engarce. La baja velocidad de desenhebrado del ARNm puede explicar la baja velocidad de formación de fusiones respecto a la velocidad de translocación. Los resultados preliminares sugieren que la cantidad de producto de fusión formado aumenta notablemente después de una incubación post-traducción prolongada a baja temperatura, quizá debido al mayor tiempo disponible para la transferencia del péptido naciente a la puromicina.  
40  
45

#### **Materiales y procedimientos detallados**

A continuación se describen materiales y procedimientos detallados en relación con la traducción *in vitro* y el ensayo de fusiones de ARN-proteína, incluyendo fusiones que tienen una etiqueta epitópica myc.  
50

Secuencias. Se usaron varios oligonucleótidos anteriormente para la generación de las fusiones de ARN-proteína. Estos oligonucleótidos tienen las secuencias siguientes.

NOMBRE	SECUENCIA
30-P	5'AAA AAA AAA AAA AAA AAA AAA AAA AAA CCP (SEC ID N°: 8)
13-P	5'AAA AAA AAA ACC P (SEC ID N°: 9)
25-P	5'CGC GGT TTT TAT TTT TTT TTT TCC P (SEC ID N°: 10)
43P	5'rGrGrA rGrGrA rCrGrA rArArU rGAA AAA AAA AAA ACC P (SEC ID N°: 11)
43-P [CUG]	5'rGrGrA rGrGrA rCrGrA rArCrU rGAA AAA AAA AAA ACC P (SEC ID N°: 12)
40-P	5'rGrGrA rGrGrA rCrGrA rArCrU rGAA AAA AAA AAA ACC P (SEC ID N°: 13)
37-P	5'rGrGrA rGrGrA rCrGrA rArCrU rGAA AAA AAA AAA ACC P (SEC ID N°: 14)
34P	5'rGrGrA rGrGrA rCrGrA rArCrU rGAA AAA AAA AAA ACC P (SEC ID N°: 15)
31-P	5'rGrGrA rGrGrA rCrGrA rArCrU rGAA AAA AAA AAA ACC P (SEC ID N°: 16)
LP77	5TGIGIG rArGrG rArCrG rArArA rUjGjG rArArC rArGrA rArArC rUrGrA rUrCrU rArArG rArArG rArArG rArArC rUrGrA rArC AAA AAA AAA AAA AAA CCP (SEC ID N°: 1)
LP154	5'rGrGrG rArCrA rArUrU rArCrU rArUrU rUrArC rArArU rUrArC rA rArUrG rGrArA rGrArA rGrArA rCrArG rArArA rCrUrG rArUrC rUrCrU rGrArA rGrArA rGrArC rCrUrG rCrUrG rCrUrG rArArA rCrUrG rCrUrG rCrUrG rCrUrG rArArC rCrUrG rArArA rCrArC rArArA rCrUrG rGrArA rCrArG rCrUrG rArArC rUrCrU rUrGrC rGrCrU AAA AAA AAA AAA AAA CCP (SEC ID N°: 3)
LP160	5'5'rGrGrG rArCrA rArUrU rArCrU rArUrU rUrArC rArArU rUrArC rA rArUrG rNfNfS rNfNfS rNfNfS rNfNfS rNfNfS rNfNfS rNfNfS rNfNfS rNfNfS rNfNfS rNfNfS rNfNfS rNfNfS rNfNfS rNfNfS rNfNfS rNfNfS rNfNfS rNfNfS rArArC rArArC rUrCrU rUrGrC rGrCrU AAA AAA AAA AAA AAA AAA AAA AAA AAA CCP (SEC ID N°: 17)

Todos los oligonucleótidos se enumeran en la dirección 5' a 3'. Las bases de ribonucleótidos se indican en minúsculas "r" antes de la designación de nucleótido; P es puromicina; rN indica cantidades equivalentes de rA, rG, rC y rU; rS indica cantidades equivalentes de rG y rC; y todas las demás designaciones de bases indican oligonucleótidos de ADN.

- 5 Productos químicos. La puromicina HCl, vidrio de tamaño de poro controlado de alquilamina de cadena larga, gougertina, cloranfenicol, virginiamicina, DMAP, cloruro de dimetiltrilito y anhídrido acético se obtuvieron en Sigma Chemical (St. Louis, MO). La piridina, dimetilformamida, tolueno, anhídrido succínico y para-nitrofenol se obtuvieron en Fluka Chemical (Ronkonkoma, NY). El ARNm de beta-globina se obtuvo en Novagen (Madison, WI). El ARN de TMV se obtuvo en Boehringer Mannheim (Indianapolis, IN).
- 10 Enzimas. La proteinasa K se obtuvo en Promega (Madison, WI). La ARNasa sin ADNasa se produjo mediante el protocolo de Sambrook y col (anteriormente) o se adquirió en Boehringer Mannheim. La polimerasa de T7 se preparó mediante el protocolo publicado de Grodberg y Dunn (J. Bacteriol. 170: 1245 (1988)) con las modificaciones de Zawadzki y Gross (Nucl. Acids. Res. 19:1948 (1991)). La ADN ligasa de T4 se obtuvo en New England Biolabs (Beverly, MA).
- 15 Cuantificación de la incorporación de radiomarcador. Para bandas de geles radiactivas, la cantidad de radiomarcador (<sup>35</sup>S o <sup>32</sup>P) presente en cada banda se determinó por cuantificación en un analizador de transferencia Betagen 603 (Betagen, Waltham, MA) o usando placas de Phosphorimager (Molecular Dynamics, Sunnyvale, CA). Para muestras líquidas y sólidas, la cantidad de radiomarcador (<sup>35</sup>S o <sup>32</sup>P) presente se determinó por recuento de escintilación (Beckman, Columbia, MD).
- 20 Imágenes de gel. Las imágenes de geles se obtuvieron por autorradiografía (usando película Kodak XAR) o usando placas de Phosphorimager (Molecular Dynamics).

Síntesis de CPG puromicina. Se han resumido anteriormente protocolos detallados para la síntesis de CPG-puromicina.

- 25 Reacciones enzimáticas. En general, la preparación de ácidos nucleicos para reacciones de quinasa, transcripción, PCR y traducción usando extractos de *E. coli* era la misma. Cada protocolo preparativo comenzaba con la extracción usando un volumen equivalente de fenol/cloroformo 1:1, seguida de centrifugación y aislamiento de la fase acuosa. Se añadió acetato sódico (pH 5,2) y espermidina a una concentración final de 300 mM y 1 mM, respectivamente, y la muestra se precipitó por adición de 3 volúmenes de etanol al 100% e incubación a -70°C durante 20 minutos. Las muestras se centrifugaron a >12.000 g, se retiró el sobrenadante y los sedimentos se lavaron con un exceso de etanol al 95% a 0°C. Los sedimentos resultantes se secaron después al vacío y se resuspendieron.
- 30

- Oligonucleótidos. Todos los ADN y ARN sintéticos se sintetizaron en un sintetizador Millipore Expedite usando química convencional para cada uno según se suministró por el fabricante (Milligen, Bedford, MA). Se sintetizaron oligonucleótidos que contenían puromicina 3' usando columnas de CPG puromicina envasadas con 30-50 mg de soporte sólido (~20 μmol puromicina/gramo). Se sintetizaron oligonucleótidos que contenían una biotina 3' usando 1 μmol de columnas de CPG Bioteg de Glen Research (Sterling, VA). Se sintetizaron oligonucleótidos que contenían una biotina 5' por adición de fosforamida Bioteg (Glen Research) como base 5'. Los oligonucleótidos a ligar con los extremos 3' de moléculas de ARN se fosforilaron químicamente en el extremo 5' (usando reactivo de fosforilación química de Glen Research) antes de la desprotección o se fosforilaron enzimáticamente usando ATP y polinucleótido quinasa de T4 (New England Biolabs) después de la desprotección. Las muestras que contenían solamente ADN (y puromicina 3' o biotina 3') se desprotegeron por adición de NH<sub>4</sub>OH al 25% seguido de incubación durante 12 horas a 55°C. Las muestras que contenían monómeros de ARN (por ejemplo, 43-P) se desprotegeron por adición de etanol (25% (v/v)) a la solución de NH<sub>4</sub>OH e incubación durante 12 horas a 55°C. El OH 2' se desprotegió usando TBAF 1 M en THF (Sigma) durante 48 horas a temperatura ambiente. El TBAF se eliminó usando una columna NAP-25 Sephadex (Farmacia, Piscataway, NJ).
- 35
- 40

- 45 Las muestras de ADN y ARN desprotegidas se purificaron después usando PAGE desnaturalizante, seguido de remojo o electroelución del gel usando un Elutrap (Schleicher y Schuell, Keene, NH) y desalinización usando una columna NAP-25 Sephadex o precipitación con etanol como se ha descrito anteriormente.

- Construcción de ADN de myc. Se construyeron dos moldes de ADN que contenían la etiqueta epitópica c-myc. El primer molde se generó a partir de una combinación de los oligonucleótidos 64.27 (5'-GTT CAG GTC TTC TTG AGA GAT CAG TTT CTG TTC CAT TTC GTC CTC CTT ATA GTG AGT CGT ATT A-3') (SEC ID N°: 18) y 18.109 (5'-TAA TAC GAC TCA CTA TAG-3') (SEC ID N°: 19). La transcripción usando este molde produjo el ARN 47.1 que codificaba el péptido MEQKLISEEDLN (SEC ID N°: 20). La ligación del ARN 47.1 con 30-P produjo el LP77 mostrado en la Figura 7A.
- 50

- 55 El segundo molde se generó primero como un solo oligonucleótido de 99 bases de longitud que tenía la designación RWR 99.6 y la secuencia 5'AGC GCA AGA GTT ACG CAG CTG TTC CAG TTT GTG TTT CAG CTG TTC ACG ACG TTT ACG CAG CAG GTC TTC TTC AGA GAT CAG TTT CTG TTC TTC AGC CAT-3' (SEC ID N°: 21). Se construyeron moldes de transcripción bicatenarios que contenían esta secuencia mediante PCR con los oligonucleótidos RWR 21.103 (5'-AGC GCA AGA GTT ACG CAG CTG-3') (SEC ID N°: 22) y RWR 63.26 (5'TAA TAC

GAC TCA CTA TAG GGA CAA TTA CTA TTT ACA ATT ACA ATG GCT GAA GAA CAG AAA CTG-3') (SEC ID N°: 23) de acuerdo con protocolos publicados (Ausubel y col., anteriormente, capítulo 15). La transcripción usando este molde producía un ARN denominado ARN124 que codificaba el péptido MAEEQKLISEEDLLRKRREQLKHKLEQLRNSCA (SEC ID N°: 24). Este péptido contenía la secuencia usada para generar el anticuerpo monoclonal 9E10 cuando se conjugaba con una proteína transportadora (Oncogene Science Technical Bulletin). El ARN124 tenía 124 nucleótidos de longitud, y la ligación de ARN124 con 30-P producía el LP154 mostrado en la Figura 7B. La secuencia del ARN 124 es la siguiente (SEC ID N°: 32): 5'-rGrGrG rArCrA rArUrU rArCrU rArUrU rUrArC rArArU rUrArC rArArUrG rGrCrU rGrArA rGrArA rCrArG rArArA rCrUrG rArUrC rUrCrU rGrArA rGrArA rGrArC rCrUrG rCrUrG rCrGrU rArArA rCrGrU rCrGrU rGrArA rCrArG rCrUrG rArArA rCrArC rArArA rCrUrG rGrArA rCrArG rCrUrG rCrGrU rArArC rUrCrU rUrGrC rGrCrU-3'.

Construcción de una combinación aleatorizada. La combinación aleatorizada se construyó con un solo oligonucleótido de 130 bases de longitud indicado como RWR130.1. Comenzando en el extremo 3', la secuencia era 3' CCCTGTTAATGATAAATGTTAATGTTAC (NNS)<sub>27</sub> GTC GAC GCA TTG AGA TAC CGA-5' (SEC ID N°: 25). N indica una posición aleatoria y esta secuencia se generó de acuerdo con el protocolo de sintetizador convencional. S indica una mezcla igual de bases dG y dC. Se realizó PCR con los oligonucleótidos 42.108 (5'-TAA TAC GAC TCA CTA TAG GGA CAA TTA CTA TTT ACA ATT ACA) (SEC ID N°: 26) y 21.103 (5'-AGC GCA AGA GTT ACG CAG CTG) (SEC ID N°: 27). La transcripción de este molde produjo un ARN indicado como combinación 130.1. La ligación de la combinación 130.1 con 30-P produjo la Combinación N° 1 (también denominada LP 160) mostrada en la Figura 7C.

Se realizaron siete ciclos de PCR de acuerdo con los protocolos publicados (Ausubel y col., anteriormente) con las excepciones siguientes: (i) la concentración de partida de RWR130.1 era de 30 nanomolar, (ii) cada cebador se usó a una concentración de 1,5  $\mu\text{M}$ , (iii) la concentración de dNTP era de 400  $\mu\text{M}$  para cada base y (iv) se usó la polimerasa Taq (Boehringer Mannheim) a 5 unidades por 100  $\mu\text{l}$ . El producto bicatenario se purificó sobre PAGE no desnaturizante y se aisló por electroelución. La cantidad de ADN se determinó tanto por absorbancia UV a 260 nm como por fluorescencia con bromuro de etidio en comparación con patrones conocidos.

Síntesis enzimática de ARN. Se realizaron reacciones de transcripción a partir de ADN de PCR bicatenario y oligonucleótidos sintéticos como se ha descrito anteriormente (Milligan y Uhlenbeck, Meth. Enzymol. 180: 51 (1989)). Se purificó ARN de longitud completa por PAGE desnaturizante, se electroeluyó y se desalinizó como se ha descrito anteriormente. La concentración de ARN de combinación se estimó usando un coeficiente de extinción de 1300 D.O./ $\mu\text{mol}$ ; ARN124, 1250 D.O./ $\mu\text{mol}$ ; ARN 47.1, 480 D.O./ $\mu\text{mol}$ . La transcripción del ADN de combinación bicatenario produjo ~90 nanomoles de ARN de combinación.

Síntesis enzimática de conjugados de ARN-puromicina. La ligación de las secuencias de ARN mensajero de myc y de combinación con el oligonucleótido que contenía puromicina se realizó usando un puente de ADN denominado 19.35 (5'-TTT TTT TTT TAG CGC AAG A) (SEC ID N°: 28) usando procedimiento análogo al descrito por Moore y Sharp (Science 250: 992 (1992)). La reacción consistía en ARNm, puente y oligonucleótido de puromicina (30-P, dA27dCdCP) en una proporción molar de 0,8:0,9:1,0 y 1-2,5 unidades de ADN ligasa por picomol de ARNm de combinación. Las reacciones se realizaron durante una hora a temperatura ambiente. Para la construcción de las fusiones de ARN de combinación, la concentración de ARNm era de ~6,6  $\mu\text{M}$ . Después de la ligación, el conjugado de ARN-puromicina se preparó como se ha descrito anteriormente para reacciones enzimáticas. El precipitado se resuspendió y se purificaron fusiones de longitud completa en PAGE desnaturizante y se aislaron por electroelución como se ha descrito anteriormente. La concentración de ARN de combinación se estimó usando un coeficiente de extinción de 1650 D.O./ $\mu\text{mol}$  y el molde de myc 1600 D.O./ $\mu\text{mol}$ . De esta forma, se generaron 2,5 nanomoles de conjugado.

Preparación de dT<sub>25</sub> estreptavidina agarosa. dT<sub>25</sub> que contenía una biotina 3' (sintetizada en columnas de fosforamida Bioteg (Glen Research)) se incubó a 1-10  $\mu\text{M}$  con una suspensión de estreptavidina agarosa (agarosa al 50% en volumen, Pierce, Rockford, IL) durante 1 hora a temperatura ambiente en TE (Tris Cloruro 10 mM pH 8,2, EDTA 1 mM) y se lavó. La capacidad de unión de la agarosa se estimó después ópticamente por la desaparición de biotina-dT<sub>25</sub> de la solución y/o por valoración de la resina con cantidades conocidas de oligonucleótido complementario.

Reacciones de traducción usando extractos y ribosomas derivados de *E. coli*. En general, se realizaron reacciones de traducción con kits adquiridos (por ejemplo, extracto de *E. coli* S30 para moldes lineales, Promega, Madison, WI). Sin embargo, también se usó *E. coli* MRE600 (obtenida en la ATCC, Rockville, MD) para generar extractos de S30 preparados de acuerdo con protocolos publicados (por ejemplo, Ellman y col., Meth. Enzymol. 202: 301(1991)), así como una fracción ribosómica preparada como se describe por Kudlicki y col. (Anal. Biochem. 206: 389 (1992)). La reacción patrón se realizó en un volumen de 50  $\mu\text{l}$  con 20-40  $\mu\text{Ci}$  de <sup>35</sup>S metionina como marcador. La mezcla de reacción consistía en extracto al 30% v/v, MgCl<sub>2</sub> 9-18 mM, premezcla sin metionina al 40% (Promega) v/v y 5  $\mu\text{M}$  de molde (por ejemplo, 43-P). Para experimentos de coincubación se añadieron los oligonucleótidos 13-P y 25-P a una concentración de 5  $\mu\text{M}$ . Para experimentos usando ribosomas, se añadieron 3  $\mu\text{l}$  de solución de ribosomas por reacción en lugar del lisado. Todas las reacciones se incubaron a 37°C durante 30 minutos. Los moldes se purificaron como se ha descrito anteriormente en reacciones enzimáticas.

Reacciones de traducción en germen de trigo. Las reacciones de traducción en la Figura 8 se realizaron usando kits adquiridos que carecían de metionina (Promega), de acuerdo con las recomendaciones del fabricante. Las concentraciones de molde eran de 4  $\mu\text{M}$  para 43-P y 0,8  $\mu\text{M}$  para LP77 y LP154. Las reacciones se realizaron a 25°C con 30  $\mu\text{Ci}$  de  $^{35}\text{S}$  metionina en un volumen total de 25  $\mu\text{l}$ .

5 Reacciones de traducción de reticulocitos. Se realizaron reacciones de traducción con kits adquiridos (Novagen, Madison, WI) o usando extracto preparado de acuerdo con protocolos publicados (Jackson y Hunt, Meth. Enzymol. 96: 50 (1983)). Se obtuvo sangre rica en reticulocitos en Pel-Freez Biologicals (Rogers, AK). En ambos casos, las condiciones de reacción eran las recomendadas para su uso con Red Nova Lysate (Novagen). Las reacciones consistían en KCl 100 mM, MgOAc 0,5 mM, DTT 2 mM, HEPES 20 mM, pH 7,6, fosfato de creatina 8 mM, 25  $\mu\text{M}$  de cada aminoácido (con la excepción de metionina si se usaba  $^{35}\text{S}$  Met) y 40% v/v de lisado. La incubación fue a 30 °C durante 1 hora. Las concentraciones de molde dependían del experimento pero generalmente variaban de 50 nM a 1  $\mu\text{M}$  con la excepción de 43-P (Figura 6H) que era de 4  $\mu\text{M}$ .

10 Para la generación de la combinación aleatorizada, se realizaron 10 ml de reacción de traducción en una concentración de molde de  $\sim 0,1$   $\mu\text{M}$  (1,25 nanomoles de molde). Además, se incluyó un molde marcado con  $^{32}\text{P}$  en la reacción para permitir la determinación de la cantidad de material presente en cada etapa del procedimiento de purificación y selección. Después de la traducción a 30 °C durante una hora, la reacción se enfrió en hielo durante 30-60 minutos.

15 Aislamiento de fusión con dT<sub>25</sub> estreptavidina agarosa. Después de la incubación, la reacción de traducción se diluyó aproximadamente 150 veces en tampón de aislamiento (NaCl 1,0 M, Tris cloruro 0,1 M pH 8,2, EDTA 10 mM, DTT 1 mM) que contenía más de un exceso molar de 10X de dT<sub>25</sub>-biotina-estreptavidina agarosa cuya concentración de dT<sub>25</sub> era  $\sim 10$   $\mu\text{M}$  (volumen de suspensión igual o superior al volumen de lisado) y se incubó con agitación a 4°C durante una hora. La agarosa se retiró después de la mezcla por filtración (Millipore ultrafree MC filters) o centrifugación y lavado con tampón de aislamiento frío 2-4 veces. Después, el molde se liberó de la dT<sub>25</sub> estreptavidina agarosa por lavado repetido con alícuotas de 50-100  $\mu\text{l}$  de NaOH 15 mM, EDTA 1 mM. El eluyente se neutralizó inmediatamente en NaOAc 3 M pH 5,2, espermidina 10 mM y se precipitó con etanol. Para la reacción de combinación, la reactividad total recuperada indicaba que se recuperaba aproximadamente el 50-70% del molde introducido.

20 Aislamiento de fusión con tiopropil sepharose. Pueden purificarse fusiones que contienen cisteína usando tiopropil sepharose 6B como en la Figura 13 (Pharmacia). En los experimentos descritos en el presente documento, el aislamiento se llevó a cabo directamente a partir de la reacción de traducción o después del aislamiento inicial de la fusión (por ejemplo, con estreptavidina agarosa). Para muestras purificadas directamente, se usó una proporción de lisado respecto a sepharose de 1:10 (v/v). Para la combinación, se usaron 0,5 ml de suspensión de sepharose para aislar todo el material de fusión a partir de 5 ml de mezcla de reacción. Las muestras se diluyeron en una suspensión de tiopropil sepharose 50:50 (v/v) en TE 1X 8,2 (Tris-Cl 10 mM, EDTA 1 mM, pH 8,2) que contenía ARNasa sin ADNasa (Boehringer Mannheim) y se incubaron con rotación durante 1-2 horas a 4°C para permitir una reacción completa. Se eliminó el exceso de líquido y la sepharose se lavó repetidamente con tampón de aislamiento que contenía DTT 20 mM y se recuperó por centrifugación o filtración. Las fusiones se eluyeron de la sepharose usando una solución de ditiotretol 25-30 mM (DTT) en Tris cloruro 10 mM pH 8,2, EDTA 1 mM. Después la fusión se concentró mediante una combinación de evaporación a alto vacío y precipitación con etanol como se ha descrito anteriormente. Para la reacción de combinación, la radiactividad total recuperada indicaba que aproximadamente el 1% del molde se convertía en fusión.

30 Para ciertas aplicaciones, se añadió dT<sub>25</sub> a este eluido y se hizo girar durante 1 hora a 4°C. La agarosa se aclaró tres veces con tampón de aislamiento frío, se aisló por filtración y el material unido se eluyó como anteriormente. Se añadió ARNt transportador y el producto de fusión se precipitó con etanol. La muestra se resuspendió en TE pH 8,2 que contenía ARNasa A sin ADNasa para eliminar la porción de ARN del molde.

35 Reacciones de inmunoprecipitación. Se realizaron inmunoprecipitaciones de péptidos de reacciones de traducción (Figura 10) mezclando 4  $\mu\text{l}$  de reacción de traducción de reticulocitos, 2  $\mu\text{l}$  de suero de ratón normal y 20  $\mu\text{l}$  de Proteína G + A agarosa (Calbiochem, La Jolla, CA) con 200  $\mu\text{l}$  de PBS (Na<sub>2</sub>HPO<sub>4</sub> 58 mM, NaH<sub>2</sub>PO<sub>4</sub> 17 mM, NaCl 68 mM), tampón de dilución (Tris cloruro 10 mM pH 8,2, NaCl 140 mM, Triton X-100 al 1% v/v) o PBSTDS (PBS + Triton X-100 al 1%, desoxicolato al 0,5%, SDS al 0,1%). Después las mezclas se hicieron girar durante una hora a 4°C, seguido de centrifugación a 2500 rpm durante 15 minutos. El eluyente se retiró y se añadieron 10  $\mu\text{l}$  de anticuerpo monoclonal de c-myc 9E10 (Calbiochem, La Jolla, CA) y 15  $\mu\text{l}$  de Proteína G + A agarosa y se hizo girar durante 2 horas a 4°C. Después, las muestras se lavaron con dos volúmenes de 1 ml de PBS, tampón de dilución o PBSTDS. Se añadieron 40  $\mu\text{l}$  de tampón de carga de gel (Calbiochem Product Bulletin) a la mezcla y se cargaron 20  $\mu\text{l}$  en un PAGE desnaturante como se describe por Schagger y von Jagow (Anal. Biochem. 166: 368 (1987)).

40 Se realizaron inmunoprecipitaciones de fusiones (como se muestra en la Figura 11) mezclando 8  $\mu\text{l}$  de reacción de traducción de reticulocitos con 300  $\mu\text{l}$  de tampón de dilución (Tris cloruro 10 mM pH 8,2, NaCl 140 mM, Tritón X-100 al 1% v/v), 15  $\mu\text{l}$  de proteína G sepharose (Sigma) y 10  $\mu\text{l}$  (1  $\mu\text{g}$ ) de anticuerpo de c-myc 9E10 (Calbiochem), seguido por rotación durante varias horas a 4°C. Después del aislamiento, las muestras se lavaron, se trataron con

ARNasa A sin ADNasa, se marcaron con polinucleótido quinasa y <sup>32</sup>P gamma ATP y se separaron mediante PAGE con urea desnaturante (Figura 11).

5 Transcripción inversa de combinación de fusión. Se realizaron reacciones de transcripción inversa de acuerdo con la recomendación de los fabricantes para Superscript II, excepto por que el molde, el agua y el cebador se incubaron a 70°C durante sólo dos minutos (Gibco BRL, Grand Island, NY). Para controlar la extensión, se incluyeron 50 µCi de alfa <sup>32</sup>P dCTP en algunas reacciones; en otras reacciones, se controló la transcripción inversa usando cebadores marcados con <sup>32</sup>P 5' que se prepararon usando <sup>32</sup>P αATP (New England Nuclear, Boston, MA) y polinucleótido quinasa T4 (New England Biolabs, Beverly, MA).

10 Preparación de sepharose con proteína g y anticuerpo. Se lavaron dos alícuotas de 50 µl de suspensión de Proteína G sepharose (50% de sólidos en volumen) (Sigma) con tampón de dilución (Tris cloruro 10 mM pH 8,2, NaCl 140 mM, NaN<sub>3</sub> al 0,025%, Triton X-100 al 1% v/v) y se aislaron por centrifugación. La primera alícuota se reservó para su uso como una precolumna antes de la matriz de selección. Después de la resuspensión de la segunda alícuota en tampón de dilución, se añadieron 40 µg de anticuerpo monoclonal de c-myc AB-1 (Oncogene Science) y la reacción se incubó durante una noche a 4°C con rotación. La sepharose con anticuerpo se purificó después por centrifugación durante 15 minutos a 1500-2500 rpm en una microcentrífuga y se lavó 1-2 veces con tampón de dilución.

15 Selección. Después del aislamiento de la fusión y de la síntesis de cadena complementaria se usó la reacción de transcriptasa inversa completa directamente en el procedimiento de selección. Se resumen en el presente documento dos protocolos. Para la ronda uno, la reacción de transcriptasa inversa se añadió directamente a la sepharose con anticuerpo preparada como se ha descrito anteriormente y se incubó 2 horas. Para rondas posteriores, la reacción se incubó ~2 horas con sepharose con proteína G lavada antes de la columna de anticuerpo para disminuir el número de agentes de unión que interaccionan con proteína G más que con el anticuerpo inmovilizado.

20 Para eluir la combinación de la matriz, pueden adoptarse varias estrategias. La primera es lavar la matriz de selección con ácido acético al 4%. Este procedimiento libera el péptido de la matriz. Como alternativa, puede usarse un lavado más riguroso (por ejemplo, usando urea u otro agente desnaturante) en lugar de o además de la estrategia de ácido acético.

25 PCR de fusiones seleccionadas. Las moléculas seleccionadas se amplifican por PCR usando protocolos convencionales como se han descrito anteriormente para la construcción de la combinación.

### Síntesis y ensayo de fusiones de beta-globina

30 Para sintetizar una construcción de fusión de β-globina, se generó ADNc de β-globina a partir de 2,5 µg de ARNm de globina por transcripción inversa con 200 pmoles de cebador 18.155 (5' GTG GTA TTT GTG AGC CAG) (SEC ID N°: 29) y transcriptasa inversa Superscript (Gibco BRL) de acuerdo con el protocolo del fabricante. La secuencia de cebador era complementaria a los 18 nucleótidos de β-globina 5' del codón de terminación. Para añadir un promotor de T7, se retiraron 20 µl de la reacción de transcripción inversa y se sometieron a 6 ciclos de PCR con los cebadores 35 18.155 y 40.54 (5' TAA TAC GAC TCA CTA TAG GGA CAC TTG CTT TTG ACA CAA C) (SEC ID N°: 30). El ARNm de "syn-β-globina" resultante se generó después por transcripción de salida con T7 de acuerdo con Milligan y Uhlenbeck (Methods Enzymol. 180: 51 (1989)) y el ARN se purificó en gel, se electroeluyó y se desalinizó como se describe en el presente documento. Después se generó "LP-β-globina" a partir de la construcción de syn-β-globina por ligación de esa construcción con 30-P de acuerdo con el procedimiento de Moore y Sharp (Science 256: 992 (1992)) usando el cebador 20.262 (5' TTT TTT TTT T GTG GTA TTT G) (SEC ID N°: 31) como puente. Después, el producto de la reacción de ligación se purificó en gel, se electroeluyó y se desalinizó como anteriormente. La concentración del producto final se determinó por absorbancia a 260 nm.

40 Estos moldes de β-globina se tradujeron después *in vitro* como se describe en la Tabla 1 en un volumen total de 25 µl cada uno. Se añadió Mg<sup>2+</sup> a partir de una solución madre 25mM. Todas las reacciones se incubaron a 30°C durante una hora y se pusieron a -20°C durante una noche. Después se determinaron los CPM precipitables en dT<sub>25</sub> dos veces usando 6 µl de lisado y se calculó el promedio sin el fondo.

TABLA 1

Reacciones de traducción con moldes de beta-globina

Reacción	Molde	Mg <sup>2+</sup> (mM)	<sup>35</sup> S Met (µl)	TCA CPM (2 µl)	dT <sub>25</sub> CPM (6 µl)
1	---	1,0	2,0 (20 µCi)	3312	0
2	2,5 µg de syn-β-globina	0,5	2,0 (20 µCi)	33860	36
3	2,5 µg de syn-β-globina	1,0	2,0 (20 µCi)	22470	82

(continuación)

## Reacciones de traducción con moldes de beta-globina

Reacción	Molde	Mg <sup>2+</sup> (mM)	<sup>35</sup> S Met (μl)	TCA CPM (2 μl)	dT <sub>25</sub> CPM (6 μl)
4	2,5 μg de syn-β-globina	2,0	2,0 (20 μCi)	15696	86
5	2,5 μg de LP-β-globina	0,5	2,0 (20 μCi)	32712	218
6	2,5 μg de LP-β-globina	1,0	2,0 (20 μCi)	24226	402
7	2,5 μg de LP-β-globina	2,0	2,0 (20 μCi)	15074	270

Para preparar las muestras para análisis en gel, se mezclaron 6 μl de cada reacción de traducción con 1000 μl de tampón de aislamiento (NaCl 1M, Tris-Cl 100 mM pH 8,2, EDTA 10 mM, DTT 0,1 mM), 1 μl de ARNasa A (sin ADNasa Boehringer Mannheim) y 20 μl de dT<sub>25</sub> estreptavidina agarosa 20 μM. Las muestras se incubaron a 4°C durante una hora con rotación. Se retiró el exceso de tampón de aislamiento y las muestras se añadieron a un filtro Millipore MC para eliminar cualquier tampón de aislamiento restante. Después, las muestras se lavaron cuatro veces con 50 μl de H<sub>2</sub>O y dos veces con 50 μl de NaOH 15 mM, EDTA 1 mM. La muestra (300 μl) se neutralizó con 100 μl de TE pH 6,8 (Tris-Cl 10 mM, EDTA 1 mM), se añadió 1 μl de ARNasa A 1 mg/ml (como anteriormente) y las muestras se incubaron a 37°C. Después se añadieron 10 μl de tampón de carga SDS 2X (Tris-Cl 125 mM pH 6,8, SDS al 2%, β-mercaptoetanol al 2%, glicerol al 20%, azul de bromofenol al 0,001%) y la muestra se liofilizó hasta sequedad y se resuspendió en 20 μl de H<sub>2</sub>O y β-mercaptoetanol al 1%. Después las muestras se cargaron sobre un gel de resolución de péptidos como se describe por Schagger y von Jagow (Analytical Biochemistry 166: 368 (1987)) y se visualizaron por autorradiografía.

Los resultados de estos experimentos se muestran en las Figuras 15A y 15B. Como se indica en la Figura 15A, se incorporó <sup>35</sup>S-metionina en la porción de proteína de las fusiones de syn-β-globina y LP-β-globina. La proteína era heterogénea, pero una intensa banda mostraba la movilidad esperada para el ARNm de β-globina. Además, como se muestra en la Figura 15B, después del aislamiento con dT<sub>25</sub> y de la digestión con ARNasa A, no quedaba material marcado con <sup>35</sup>S en los carriles de syn-β-globina (Figura 15B, carriles 2-4). Por el contrario, en los carriles de LP-β-globina se observó un producto marcado con <sup>35</sup>S de un tamaño homogéneo.

Estos resultados indicaban que, como anteriormente, se aisló un producto de fusión por cromatografía de afinidad con oligonucleótido sólo cuando el molde contenía una puromicina 3'. Esto se confirmó por recuento de escintilación (véase la Tabla 1). Se espera que el material obtenido contenga el engarce 30-P fusionado a alguna porción de la β-globina. El producto de fusión parecía bastante homogéneo en tamaño a juzgar por el análisis en gel. Sin embargo, puesto que el producto mostraba una movilidad muy similar a la β-globina natural (Figuras 15A y 15B, carriles de control) era difícil determinar la longitud exacta de la porción de proteína del producto de fusión.

**Optimización adicional de formación de fusiones de ARN-proteína**

Se ha descubierto que ciertos factores aumentan adicionalmente la eficacia de formación de fusiones de ARN-péptido. La formación de fusiones, es decir, la transferencia de la cadena de péptido naciente de su ARNt al resto de puromicina en el extremo 3' del ARNm es una reacción lenta que sigue a la traducción inicial relativamente rápida de la fase de lectura abierta para generar el péptido naciente. El grado de formación de fusiones puede aumentar sustancialmente mediante una incubación post-traducciona en condiciones elevadas de Mg<sup>2+</sup> (preferentemente, en un intervalo de 50-100 mM) y/o mediante el uso de un engarce más flexible entre el ARNm y el resto de puromicina. Además, incubaciones largas (12-48 horas) a bajas temperaturas (referentemente, -20°C) también dan como resultado rendimientos aumentados de fusiones con menos degradación de ARNm de la que se produce durante la incubación a 30°C. Por combinación de estos factores, hasta el 40% del ARNm introducido puede convertirse en productos de fusión de ARNm-péptido como se muestra a continuación.

Síntesis de conjugados de ARNm-puromicina. En estos experimentos de optimización, se ligaron oligonucleótidos de engarce que contenían puromicina con los extremos 3' de ARNm usando ADN ligasa de bacteriófago T4 en presencia de puentes de ADN complementario, generalmente como se ha descrito anteriormente. Puesto que la ADN ligasa de T4 prefiere un emparejamiento de bases exacto cerca de la unión de ligación y los productos de la transcripción de salida con ARN polimerasa T7, T3 o SP6 son con frecuencia heterogéneos en sus extremos 3' (Nucleic Acids Research 15: 8783 (1987)), sólo los ARN que contienen el nucleótido 3' terminal correcto se ligaron eficazmente. Cuando se usó un puente de ADN convencional, aproximadamente el 40% de los productos de la transcripción de salida se ligaron con el oligonucleótido puromicina. La cantidad de producto de ligación se aumentó usando ARN en exceso pero no se aumentó usando un exceso de oligonucleótido de puromicina. Sin desear quedar ligado a teoría particular alguna, parecía que el factor limitante para la ligación en la cantidad de ARN que era totalmente complementario a la región correspondiente del puente de ADN.

Para permitir la ligación de esos transcritos que terminaban con un nucleótido no moldeado en el extremo terminal 3' (denominados "productos N+1"), se usó una mezcla del puente de ADN convencional con un nuevo puente de ADN que contenía una base aleatoria adicional en la unión de ligación. La eficacia de la ligación aumentó hasta más del 70% para un molde de ARN de myc ejemplar (es decir, ARN124) en presencia de dicho puente de ADN mixto.

5 Además de esta estrategia de puente de ADN modificado, la eficacia de la formación de conjugado de ARNm-puromicina también se optimizó adicionalmente teniendo en cuenta los siguientes tres factores. En primer lugar, se diseñaron o utilizaron preferentemente ARNm que carecían de extremo terminal 3' que tuviera cualquier estructura secundaria estable significativa que interfiriese con la hibridación con un oligonucleótido de puente. Además, debido a que una alta concentración de sal a veces causaba el fracaso de la reacción de ligación, se incluía preferentemente una desalinización minuciosa de los oligonucleótidos usando columnas NAP-25 como etapa en el procedimiento. Por último, debido a que la reacción de ligación era relativamente rápida y generalmente completa en 40 minutos a temperatura ambiente, generalmente no se utilizaron periodos de incubación significativamente más prolongados y que con frecuencia daban como resultado una degradación innecesaria del ARN.

15 Usando las condiciones anteriores, se sintetizaron conjugados de ARNm-puromicina de la forma siguiente. La ligación de la secuencia de ARN de myc (ARN124) con el oligonucleótido que contiene puromicina se realizó usando un puente de ADN convencional (por ejemplo, 5'-TTTTTTTTTAGCGCAAGA) o un puente que contenía una base aleatoria (N) en la unión de ligación (por ejemplo, 5'-TTTTTTTTTNAGCGCAAGA). Las reacciones consistían en ARNm, el puente de ADN y el oligonucleótido de puromicina en una proporción molar de 1,0:1,5-2,0:1,0. Se calentó primero una mezcla de estos componentes a 94°C durante 1 minuto y después se enfrió en hielo durante 15 minutos. Las reacciones de ligación se realizaron durante una hora a temperatura ambiente en Tris-HCl 50 mM (pH 7,5), MgCl<sub>2</sub> 10 mM, DTT 10 mM, ATP 1 mM, BSA 25 µg/ml, oligonucleótido de puromicina 15 µM, ARNm 15 µM, puente de ADN 22,5-30 µM, inhibidor de ARNasina (Promega) a 1 U/µl, y 1,6 unidades de ADN ligasa de T4 por picomol de oligonucleótido de puromicina. Después de la incubación, se añadió EDTA a una concentración final de 30 mM y las mezclas de reacción se extrajeron con fenol/cloroformo. Los conjugados de longitud completa se purificaron por PAGE desnaturante y se aislaron por electroelución.

25 Condiciones de traducción de reticulocitos generales. Además de mejorar la síntesis del conjugado de ARNm-puromicina, las reacciones de traducción también se optimizaron adicionalmente de la forma siguiente. Las reacciones se realizaron en lisados de reticulocitos de conejo de diferentes fuentes comerciales (Novagen, Madison, WI; Amersham, Arlington Heights, IL; Boehringer Mannheim, Indianapolis, IN; Ambion, Austin, TX; y Promega, Madison, WI). Una mezcla de reacción típica (volumen final de 25 µl) consistía en HEPES 20 mM pH 7,6, DTT 2 mM, fosfato de creatina 8 mM, KCl 100 mM, Mg(OAc)<sub>2</sub> 0,75 mM, ATP 1 mM, GTP 0,2 mM, 25 µM de cada aminoácido (metionina 0,7 µM si se usaba <sup>35</sup>S-Met), ARNasina a un 1 U/µl y lisado al 60% (v/v). La concentración final de molde estaba en el intervalo de 50 nM a 800 nM. Para cada incubación, todos los componentes excepto el lisado se mezclaron cuidadosamente en hielo y el lisado congelado se descongeló inmediatamente antes del uso. Después de la adición de lisado, la mezcla de reacción se mezcló minuciosamente por pipeteo suave y se incubó a 30°C para iniciar la traducción. Las concentraciones óptimas de Mg<sup>2+</sup> y K<sup>+</sup> variaban dentro de los intervalos de 0,25 mM-2 mM y 75 mM-200 mM, respectivamente, para diferentes ARNm, y se determinaron preferentemente en experimentos preliminares. Particularmente para ARNm escasamente traducidos, las concentraciones de hemina, fosfato de creatina, ARNt y aminoácidos también se optimizaron a veces. Generalmente se prefirió cloruro de potasio sobre acetato de potasio para las reacciones de fusión, pero una mezcla de KCl y KOAc a veces producía mejores resultados.

35 Después de la traducción a 30°C durante 30 a 90 minutos, la reacción se enfrió en hielo durante 40 minutos y se añadió Mg<sup>2+</sup>. La concentración final de Mg<sup>2+</sup> añadida en esta etapa se optimizó también para diferentes moldes de ARNm pero generalmente estaba en el intervalo de 50 mM a 100 mM (usándose preferentemente 50 mM para combinaciones de moldes mixtos). La mezcla resultante se incubó a -20°C durante 16 o 28 horas. Para visualizar los productos de fusión marcados, se mezclaron 2 µl de la mezcla de reacción con 4 µl de tampón de carga y la mezcla se calentó a 75°C durante 3 minutos. La mezcla resultante se cargó después sobre un gel de SDS-poliacrilamida con glicina al 6% (para moldes marcados con <sup>32</sup>P) o un gel de SDS-poliacrilamida con tricina al 8% (para moldes marcados con <sup>35</sup>S-Met). Como alternativa a esta estrategia, los productos de fusión también pueden aislarse usando dT<sub>25</sub> estreptavidina agarosa o tiopropil sepharose (o ambas), generalmente como se describe en el presente documento.

45 Para eliminar la porción de ARN del conjugado de ARN-engarce-puromicina-péptido para su análisis posterior mediante SDS-PAGE, se añadió una cantidad apropiada de EDTA después de la incubación post-traduccional y la mezcla de reacción se desalinizó usando una columna microcon-10 (o microcon-30). Se mezclaron 2 µl de la mezcla resultante (aproximadamente 25 µl totales) con 18 µl de tampón de ARNasa H (Tris-HCl 30 mM, pH 7,8, (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> 30 mM, MgCl<sub>2</sub> 8 mM, β-mercaptoetanol 1,5 mM y una cantidad apropiada de puente de ADN complementario) y la mezcla se incubó a 4°C durante 45 minutos. Después se añadió ARNasa H y la digestión se realizó a 37°C durante 20 minutos.

55 Calidad del oligonucleótido de puromicina. La calidad del oligonucleótido de puromicina también era importante para la generación eficaz de productos de fusión. El acoplamiento de 5'-DMT, 2'-succinil, N-trifluoroacetil puromicina con

60



CPG no era tan eficaz como el acoplamiento de los nucleótidos convencionales. Como tal, la reacción de acoplamiento se controló cuidadosamente para evitar la formación de CPG con una concentración demasiado baja de puomicina acoplada, y los grupos amino sin reaccionar en el CPG se inactivaron totalmente para evitar la posterior síntesis de oligonucleótidos que carecieran de una puomicina 3'-terminal. También era importante evitar el uso de CPG que contuviera partículas de tamaño de malla muy fino, ya que éstas son capaces de causar problemas con obstrucción de válvulas durante etapas de síntesis de oligonucleótidos automáticas posteriores.

Además, el oligonucleótido de puomicina sintetizado se ensayó preferentemente antes de su uso a gran escala para asegurar la presencia de puomicina en el extremo 3'. En los experimentos de los presentes inventores, no se detectó fusión si la puomicina se sustituía con una desoxiadenosina que contenía un grupo amino primario en el extremo 3'. Para ensayar la presencia de grupos hidroxilo 3' (es decir, la síntesis no deseada de oligonucleótidos que carecían de una puomicina 3'-terminal), el oligonucleótido de puomicina puede radiomarcarse primero (por ejemplo, por fosforilación 5') y después usarse como un cebador para extensión con desoxinucleotidil transferasa terminal. En presencia de un resto de puomicina 3'-terminal, no debería observarse producto de extensión.

Transcurso de tiempo de la incubación de traducción y post-traduccional. La reacción de traducción era relativamente rápida y se completaba generalmente en 25 minutos a 30°C. La reacción de fusión, sin embargo, era más lenta. Cuando se usaba un engarce convencional (dA<sub>27</sub>dCdCP) a 30°C, la síntesis de fusión alcanzaba su nivel máximo en 45 minutos adicionales. La incubación post-traducciona podía llevarse a cabo a menores temperaturas, por ejemplo, temperatura ambiente, 0°C o -20°C. Se observaba menos degradación del molde de ARNm a -20°C y los mejores resultados de fusión se obtuvieron después de la incubación a -20°C durante 2 días.

El efecto de la concentración de Mg<sup>2+</sup>. Una alta concentración de Mg<sup>2+</sup> en la incubación post-traducciona estimulaba enormemente la formación de fusiones. Por ejemplo, para el molde de ARN de myc descrito anteriormente, se observó una estimulación de 3-4 veces de la formación de fusiones usando un engarce convencional (dA<sub>27</sub>dCdCP) en presencia de Mg<sup>2+</sup> 50 mM durante la incubación de 16 horas a -20°C (Figura 17, compárense los carriles 3 y 4). De forma similar, también se observó una formación de fusiones eficaz usando una incubación post-traducciona en presencia de una concentración de Mg<sup>2+</sup> 50-100 mM cuando las reacciones se llevaron a cabo a temperatura ambiente durante 30-45 minutos.

Secuencia y longitud del engarce. La dependencia de la reacción de fusión de la longitud del engarce también se examinó. En el intervalo de entre 21 y 30 nucleótidos (n = 18-27) se observaron pocos cambios en la eficacia de la reacción de fusión (como se ha descrito anteriormente). Engarces más cortos (por ejemplo, de 13 nucleótidos de longitud) dieron como resultado una menor fusión. Además, aunque engarces particulares de mayor longitud (es decir, de 45 nucleótidos y 54 nucleótidos) también daban como resultado eficacias de fusión algo menores, continúa siendo probable que también puedan usarse engarces todavía más largos para optimizar la eficacia de la reacción de fusión.

Con respecto a la secuencia de engarce, la sustitución de restos de desoxirribonucleótidos próximos al extremo 3' con restos de ribonucleótidos no cambiaba significativamente la eficacia de fusión. La secuencia de dCdCP (o rCrCP) en el extremo 3' del engarce era sin embargo importante para la formación de fusiones. La sustitución de dCdCP con dUdUP reducía la eficacia de formación de fusiones significativamente.

Flexibilidad del engarce. La dependencia de la reacción de fusión de la flexibilidad del engarce también se ensayó. En estos experimentos, se determinó que la eficacia de fusión era baja si la rigidez del engarce se aumentaba por hibridación con un oligonucleótido complementario próximo al extremo 3'. De forma similar, cuando se usaba un engarce más flexible (por ejemplo dA<sub>21</sub>C<sub>9</sub>C<sub>9</sub>dAdCdCP, en el que C<sub>9</sub> representa HO(CH<sub>2</sub>CH<sub>2</sub>O)<sub>3</sub>PO<sub>2</sub>), la eficacia de fusión se mejoraba significativamente. En comparación con el engarce convencional (dA<sub>27</sub>dCdCP), el uso del engarce más flexible (dA<sub>21</sub>C<sub>9</sub>C<sub>9</sub>dAdCdCP) mejoraba la eficacia de fusión para el ARN124 más de 4 veces (Figura 17, compárense los carriles 1 y 9). Además, al contrario que el molde con el engarce convencional cuya fusión post-traducciona apenas avanzaba en ausencia de una alta concentración de Mg<sup>2+</sup> (Figura 17, carril 3 y 4), el molde con el engarce flexible no requería Mg<sup>2+</sup> elevado para producir un buen rendimiento de producto de fusión en una incubación post-traducciona prolongada a -20°C (Figura 17, compárense los carriles 11 y 12). Este engarce, por lo tanto, era muy útil si no se deseaban adiciones post-traduccionales de altas concentraciones de Mg<sup>2+</sup>. Además, el engarce flexible también producía rendimientos de fusión óptimos en presencia de Mg<sup>2+</sup> elevado.

Cuantificación de la eficacia de fusión. La eficacia de fusión puede expresarse como la fracción de péptido traducido convertido en producto de fusión, o la fracción de molde introducido convertido en producto de fusión. Para determinar la fracción de péptido traducido convertido en producto de fusión, se utilizó marcaje con <sup>35</sup>S-Met del péptido traducido. En estos experimentos, cuando se usó un engarce dA<sub>27</sub>dCdCP o dA<sub>27</sub>rCrCP, aproximadamente el 3,5% del péptido traducido se fusionó a su ARNm después de una incubación de traducción de 1 hora a 30°C. Este valor aumentaba hasta el 12% después de una incubación durante 1 noche a -20°C. Cuando la incubación post-traducciona se llevaba a cabo en presencia de una alta concentración de Mg<sup>2+</sup>, más del 50% del péptido traducido se fusionaba con el molde.

Para un molde con un engarce flexible, aproximadamente el 25% del péptido traducido se fusionaba con el molde después de 1 hora de traducción a 30°C. Este valor aumentaba hasta más del 50% después de la incubación

durante 1 noche a -20°C, y hasta más del 75% si la incubación post-traducciona se realizaba en presencia de Mg<sup>2+</sup> 50 mM.

5 Para determinar el porcentaje del molde introducido convertido en producto de fusión, las traducciones se realizaron usando molde de ARNm-engarce marcado con <sup>32</sup>P. Cuando se usó el engarce flexible y la incubación post-traducciona se realizó a -20°C sin adición de Mg<sup>2+</sup>, aproximadamente el 20%, 40%, 40%, 35% y 20% del molde introducido se convirtió en fusión de ARNm-péptido cuando la concentración del molde de ARN introducido era de 800, 400, 200, 100 y 50 nM, respectivamente (Figura 18). Se obtuvieron resultados similares cuando la incubación post-traducciona se realizó en presencia de Mg<sup>2+</sup> 50 mM. Los mejores resultados se consiguieron usando lisados obtenidos en Novagen, Amersham o Ambion (Figura 19).

10 Las diferencias de movilidad entre ARNm y fusiones de ARNm-péptido según se midieron por SDS-PAGE pueden ser muy pequeñas si el molde de ARNm es largo. En tales casos, el molde puede marcarse en el extremo 5' del engarce con <sup>32</sup>P. La porción de ARN larga puede digerirse después con ARNasa H en presencia de un puente de ADN complementario después de la traducción/incubación, y la eficacia de fusión determinarse por cuantificación de la proporción de engarce no modificado respecto a fusión de engarce-péptido. En comparación con la digestión con ARNasa A, que produce 3'-P y 5'-OH, esta estrategia tiene la ventaja de que no se elimina el <sup>32</sup>P en el extremo 5' del engarce.

15 Fusión intramolecular frente a intermolecular durante la incubación post-traducciona. Además de los experimentos anteriores, los presentes inventores ensayaron si la reacción de fusión que se producía a -20°C en presencia de Mg<sup>2+</sup> era intra- o intermolecular por naturaleza. Se coincubó engarce libre (dA<sub>27</sub>dCdCP o dA<sub>21</sub>C<sub>9</sub>C<sub>9</sub>dAdCdCP, donde C<sub>9</sub> es - O(CH<sub>2</sub>CH<sub>2</sub>O)<sub>3</sub>PO<sub>2</sub>-) con un molde que contenía un engarce de ADN, pero sin puromicina en el extremo 3', en las condiciones de incubación de traducción y post-traducciona descritas anteriormente. En estos experimentos, no se incorporó una cantidad detectable (es decir, inferior al 2% del nivel normal) de <sup>35</sup>S-Met en el producto de engarce-péptido, sugiriendo que la fusión post-traducciona se producía principalmente entre el péptido naciente y el ARNm unido al mismo ribosoma.

25 Resultados de optimización. Como se ha ilustrado anteriormente, mediante el uso del engarce flexible y/o la realización de la incubación post-traducciona en presencia de una alta concentración de Mg<sup>2+</sup>, se aumentaron las eficacias de fusión hasta aproximadamente el 40% del ARNm introducido. Estos resultados indicaban que podían generarse tantas como 10<sup>14</sup> moléculas de fusión de ARNm-péptido por ml de mezcla de reacción de traducción *in vitro*, produciendo combinaciones de fusiones de ARN-péptido de una complejidad muy elevada para su uso en experimentos de selección *in vitro*.

### Enriquecimiento selectivo de fusiones de ARN-proteína

Los presentes inventores han demostrado la viabilidad del uso de fusiones de ARN-péptido en experimentos de selección y evolución por enriquecimiento de una fusión de ARN-péptido particular a partir de una combinación compleja de fusiones de secuencias aleatorias basándose en el péptido codificado. En particular, los presentes inventores prepararon una serie de mezclas en las que se combinaba una pequeña cantidad de secuencia conocida (en este caso, el molde de myc largo, LP154) con cierta cantidad de combinación de secuencia aleatoria (es decir, LP160). Estas mezclas se tradujeron y los productos de fusión de ARN-péptido se seleccionaron por cromatografía de afinidad de oligonucleótido y disulfuro como se describe en el presente documento. Las fusiones de molde de myc se inmunoprecipitaron selectivamente con anticuerpo monoclonal anti-myc (Figura 16A). Para medir el enriquecimiento obtenido en esta etapa selectiva, se amplificaron alícuotas de la mezcla de fusiones de ADNc/ARNm-péptido antes y después de la inmunoprecipitación por PCR en presencia de un cebador radiomarcado. El ADN amplificado se digirió con una endonucleasa de restricción que cortaba la secuencia de molde de myc pero no la combinación (Figuras 16B y 16C). La cuantificación de la proporción de ADN cortado y no cortado indicaba que la secuencia de myc estaba enriquecida en 20-40 veces respecto a la biblioteca aleatoria por inmunoprecipitación.

Estos experimentos se llevaron a cabo de la forma siguiente.

50 Reacciones de traducción. Se realizaron reacciones de traducción generalmente como se ha descrito anteriormente. En concreto, las reacciones se realizaron a 30°C durante una hora de acuerdo con las especificaciones del fabricante (Novagen) y se congelaron durante una noche a -20°C. Se prepararon dos versiones de seis muestras, una que contenía <sup>35</sup>S metionina y una que contenía metionina fría añadida a una concentración final de 52 μM. Las reacciones 1-6 contenían las cantidades de moldes descritas en la Tabla 2. Todos los números de la Tabla 2 representan picomoles de molde por 25 μl de mezcla de reacción.

TABLA 2

Proporciones de molde usadas en selección adulterada

Reacción	LP154	LP160
1	---	---
2	5	---
3	1	20
4	0,1	20
5	0,01	20
6	---	20

Preparación de dT<sub>25</sub> estreptavidina agarosa. La estreptavidina agarosa (Pierce) se lavó tres veces con TE 8,2 (Tris-Cl 10 mM pH 8,2, EDTA 1 mM) y se resuspendió como una suspensión 1:1 (v/v) en TE 8,2. Después se añadió 3' biotil T<sub>25</sub> sintetizado usando Bioteg CPG (Glen Research) a la concentración final deseada (generalmente de 10 ó 20 µM) y la incubación se llevó a cabo con agitación durante 1 hora. La dT<sub>25</sub> estreptavidina agarosa se lavó después tres veces con TE 8,2 y se almacenó a 4°C hasta el uso.

Purificación de moldes a partir de reacciones de traducción. Para purificar moldes a partir de reacciones de traducción, se retiraron 25 µl de cada reacción y se añadieron a 7,5 ml de tampón de aislamiento (NaCl 1 M; Tris-Cl 100 mM pH 8,2, EDTA 10 mM, DTT 0,1 mM) y 125 µl de dT<sub>25</sub> estreptavidina agarosa 20 µM. Esta solución se incubó a 4°C durante una hora con rotación. Los tubos se centrifugaron y se retiró el eluyente. Se añadió un ml de tampón de aislamiento, la suspensión se resuspendió y las mezclas se transfirieron a tubos de microcentrífuga de 1,5 ml. Después, las muestras se lavaron cuatro veces con alícuotas de 1 ml de tampón de aislamiento helado. Las muestras calientes y frías de reacciones idénticas se combinaron después en una unidad de filtro de Millipore MC filter y se eluyeron de la dT<sub>25</sub> agarosa por lavado con dos volúmenes de 100 µl de H<sub>2</sub>O, DTT 0,1 mM y 2 volúmenes de NaOH 15 mM, EDTA 1 mM.

A este eluyente se añadieron 40 µl de una suspensión al 50% de tiopropil sepharose lavada (Pharmacia) y la incubación se llevó a cabo a 4°C con rotación durante 1 hora. Después las muestras se lavaron con tres volúmenes de 1 ml de TE 8,2 y se retiró el eluyente. Se añadió un µl de DTT 1 M al sólido (volumen total de aproximadamente 20-30 µl) y la muestra se incubó durante varias horas, se retiró y se lavó cuatro veces con 20 µl de H<sub>2</sub>O (volumen total 90 µl). El eluyente contenía tiopiridona 2,5 mM a juzgar por la absorbancia UV. 50 µl de esta muestra se precipitaron con etanol por adición de 6 µl de NaOAc 3 M pH 5,2, espermina 10 mM, 1 µl de glucógeno (10 mg/ml, Boehringer Mannheim) y 170 µl de EtOH al 100%, incubación durante 30 minutos a -70°C y centrifugación durante 30 minutos a 13.000 rpm en una microcentrífuga.

Reacciones de transcriptasa inversa. Se realizaron reacciones de transcripción inversa en las muestras tanto precipitadas con etanol como de eluyente de tiopiridona de la forma siguiente. Para las muestras precipitadas con etanol, 30 µl de molde resuspendido, H<sub>2</sub>O a 48 µl y 200 picomoles de cebador 21.103 (SEC ID N°: 22) se hibridaron a 70°C durante 5 minutos y se enfriaron en hielo. A esta muestra, se añadieron 16 µl de tampón de primera cadena (Tris-Cl 250 mM pH 8,3, KCl 375 mM y MgCl<sub>2</sub> 15 mM; disponible en Gibco BRL, Grand Island, NY), se añadieron 8 µl de DTT 100 mM y 4 µl de NTP 10 mM y se equilibraron a 42°C y se añadieron 4 µl de transcriptasa inversa Superscript II (Gibco BRL, Grand Island, NY). Se añadió H<sub>2</sub>O (13 µl) al eluyente de TP sepharose (35 µl) y las reacciones se realizaron como anteriormente. Después de la incubación durante una hora, las muestras numeradas igual se combinaron (volumen total 160 µl). Se reservaron 10 µl de muestra para la PCR de cada muestra no seleccionada y se reservaron 150 µl de muestra para la inmunoprecipitación.

Inmunoprecipitación. Para llevar a cabo inmunoprecipitaciones se añadieron 170 µl de reacción de transcripción inversa a 1 ml de tampón de dilución (Tris-Cl 10 mM, pH 8,2, NaCl 140 mM, Triton X-100 al 1% v/v) y 20 µl de conjugado de Proteína G/A (Calbiochem, La Jolla, CA), y se preaclararon por incubación a 4°C con rotación durante 1 hora. El eluyente se retiró y se añadieron 20 µl de conjugado de G/A y 20 µl de anticuerpo monoclonal (2 µg, 12 picomoles) y la muestra se incubó con rotación durante dos horas a 4°C. El conjugado se precipitó por microcentrifugación a 2.500 rpm durante 5 minutos, se retiró el eluyente y el conjugado se lavó tres veces con alícuotas de 1 ml de tampón de dilución helado. La muestra se lavó después con 1 ml de Tris-Cl 10 mM helado, pH 8,2, NaCl 100 mM. Los fragmentos unidos se retiraron usando 3 volúmenes de HOAc al 4% congelado y las muestras se liofilizaron hasta sequedad.

PCR de muestras seleccionadas y no seleccionadas. Se llevaron a cabo reacciones de PCR por adición de 20 µl de NH<sub>4</sub>OH concentrado a 10 µl del material no seleccionado y la totalidad del material seleccionado, e incubación

5 durante 5 minutos cada una a 55°C, 70°C y 90°C para destruir cualquier ARN presente en la muestra. Después, las muestras se evaporaron hasta sequedad usando un Speedvac. Se añadieron a cada muestra 200 µl de mezcla de PCR (cebadores 21.103 y 42.108 1 µM, dNTP 200 µM en tampón de PCR más Mg<sup>2+</sup> (Boehringer Mannheim) y 2 µl de Taq polimerasa (Boehringer Mannheim)). Se realizaron 16 ciclos de PCR sobre la muestra no seleccionada número 2 y se realizaron 19 ciclos sobre todas las demás muestras.

Las muestras se amplificaron después en presencia de cebador 21.103 marcado con <sup>32</sup>P 5' de acuerdo con la Tabla 3 y se purificaron dos veces individualmente usando kits de purificación por PCR directa Wizard (Promega) para eliminar todos los cebadores y fragmentos más cortos.

TABLA 3

10 Amplificación de muestras de PCR seleccionadas y no seleccionadas

Muestra	Tipo	Volumen	Ciclos
1	no seleccionada	20 µl	5
2	no seleccionada	5 µl	4
3	no seleccionada	20 µl	5
4	no seleccionada	20 µl	5
5	no seleccionada	20 µl	5
6	no seleccionada	20 µl	5
1	seleccionada	20 µl	5
2	seleccionada	5 µl	4
3	seleccionada	20 µl	5
4	seleccionada	20 µl	7
5	seleccionada	20 µl	7
6	seleccionada	20 µl	7

15 Digestiones de restricción. ADN marcado con <sup>32</sup>P preparado a partir de cada una de las reacciones de PCR anteriores se añadió en cantidades iguales (por cpm de muestra) a reacciones de digestión de restricción de acuerdo con la Tabla 4. El volumen total de cada reacción era de 25 µl. Se añadieron 0,5 µl de AlwNI (5 unidades, New England Biolabs) a cada reacción. Las muestras se incubaron a 37°C durante 1 hora y la enzima se inactivó por calor mediante una incubación de 20 minutos a 65°C. Después, las muestras se mezclaron con 10 µl de tampón de carga desnaturante (1 ml de formamida Ultrapure (USB), 20 µl de EDTA 0,5 M y 20 µl de NaOH 1 M), se calentaron a 90°C durante 1 minuto, se enfriaron y se cargaron sobre un gel de poliacrilamida desnaturante al 12% que contenía urea 8 M. Después de la electroforesis, el gel se fijó con HOAc al 10% (v/v), MeOH al 10% (v/v), H<sub>2</sub>O.

20 TABLA 4

Condiciones de digestión de Restricción con AlwNI

Muestra	Tipo	Volumen de ADN añadido a la reacción	Volumen total
1	no seleccionada	20 µl	25 µl
2	no seleccionada	4 µl	25 µl
3	no seleccionada	20 µl	25 µl
4	no seleccionada	20 µl	25 µl
5	no seleccionada	4 µl	25 µl
6	no seleccionada	20 µl	25 µl

(continuación)

Muestra	Tipo	Volumen de ADN añadido a la reacción	Volumen total
1	seleccionada	20 µl	25 µl
2	seleccionada	8 µl	25 µl
3	seleccionada	12 µl	25 µl
4	seleccionada	12 µl	25 µl
5	seleccionada	20 µl	25 µl
6	seleccionada	20 µl	25 µl

5 Quantificación de producto de digestión. La cantidad de ADN de myc frente a combinación presente en una muestra se cuantificó usando un Phosphorimager (Molecular Dynamics). La cantidad de material presente en cada banda se determinó como el volumen integrado de rectángulos idénticos dibujados alrededor de las bandas de gel. Los cpm  
10 totales presentes en cada banda se calcularon como el volumen menos el fondo. Se usaron tres valores de fondo: (1) un promedio de cuadrados idénticos fuera del área en el que se produjeron los recuentos en el gel; (2) los cpm presentes en el carril de combinación no seleccionado donde debería aparecer la banda de myc (no aparece banda en esta posición en el gel); y (3) un valor normalizado que reproducía el valor más próximo a los aumentos de molde de 10 veces entre carriles no seleccionados. Los carriles 2, 3 y 4 en las Figuras 16B y 16C demuestran el enriquecimiento de la diana frente a la secuencia de combinación. El enriquecimiento demostrable en el carril 3 (no seleccionado/seleccionado) producía los mayores valores (17, 43 y 27 veces usando los procedimientos 1-3, respectivamente) debido a la optimización de la proporción de señal respecto a interferencia para esta muestra. Estos resultados se resumen en la Tabla 5.

TABLA 5

15 Enriquecimiento de molde de myc frente a combinación

Procedimiento	Carril 2 (20)	Carril 3 (200)	Carril 4 (2000)
1	7,0	16,6	5,7
2	10,4	43	39
3	8,7	27	10,2

En un segundo conjunto de experimentos, estos mismos productos de PCR se purificaron una vez usando kits de purificación por PCR directa Wizard y los productos de digestión se cuantificaron mediante el procedimiento (2) anterior. En estos experimentos se obtuvieron resultados similares; se midieron enriquecimientos de 10,7, 38 y 12 veces, respectivamente, para muestras equivalentes a las de los carriles 2, 3 y 4 anteriores.

## 20 **Uso de sistemas de selección de proteína**

Los sistemas de selección descritos en el presente documento tienen aplicaciones comerciales en cualquier área en la que se use tecnología de proteínas para resolver problemas terapéuticos, de diagnóstico o industriales. Esta tecnología de selección es útil para mejorar o alterar proteínas existentes, así como para aislar nuevas proteínas con funciones deseadas. Estas proteínas pueden ser secuencias de origen natural, pueden ser formas alteradas de  
25 secuencias de origen natural o pueden ser secuencias parcialmente o totalmente sintéticas.

Aislamiento de nuevo reactivo de unión. En una aplicación particular, la tecnología de fusión de ARN-proteína descrita en el presente documento es útil para el aislamiento de proteínas con propiedades de unión específicas (por ejemplo, unión a ligando). Las proteínas que presentan interacciones de unión altamente específicas pueden usarse como reactivos de reconocimiento distintos de anticuerpos que permiten que la tecnología de fusión de ARN-proteína evite la tecnología de anticuerpos monoclonales tradicional. Pueden usarse reactivos de tipo anticuerpo aislados mediante este procedimiento en cualquier área en el que se utilizan los anticuerpos tradicionales, incluyendo aplicaciones de diagnóstico y terapéuticas.  
30

Mejora de anticuerpos humanos. La tecnología de fusión de ARN-proteína descrita en el presente documento también puede usarse para mejorar anticuerpos humanos o humanizados para el tratamiento de cualquiera de  
35 varias enfermedades. En esta solicitud, se desarrollan bibliotecas de anticuerpos y se exploran *in vitro*, eliminando la necesidad de técnicas tales como fusión de células o presentación en fago. En una aplicación importante, la tecnología de fusión de ARN-proteína descrita en el presente documento es útil para mejorar bibliotecas de

anticuerpos de cadena sencilla (Ward y col., Nature 341: 544 (1989); y Goulot y col., J. Mol. Biol. 213: 617 (1990)). Para esta aplicación, la región variable puede construirse a partir de una fuente humana (para minimizar posibles reacciones inmunes adversas del destinatario) o puede contener un casete totalmente aleatorizado (para maximizar la complejidad de la biblioteca). Para explorar moléculas de anticuerpos mejoradas, se ensaya una combinación de moléculas candidatas para determinar la unión a una molécula diana (por ejemplo, un antígeno inmovilizado como se muestra en la Figura 2). Después, se aplican mayores niveles de rigurosidad a la etapa de unión a medida que avanza la selección de una ronda a la siguiente. Para aumentar la rigurosidad, se alteran condiciones tales como el número de etapas de lavado, la concentración de competidor en exceso, las condiciones de tampón, la duración del tiempo de reacción de unión y la elección de matriz de inmovilización.

Pueden usarse anticuerpos de cadena sencilla directamente para terapia o indirectamente para el diseño de anticuerpos convencionales. Dichos anticuerpos tienen varias aplicaciones potenciales, incluyendo el aislamiento de anticuerpos anti-autoinmunes, supresión inmune y en el desarrollo de vacunas para enfermedades víricas tales como SIDA.

Aislamiento de nuevos catalizadores. La tecnología de fusión de ARN-proteína descrita en el presente documento también puede usarse para seleccionar nuevas proteínas catalíticas. La selección y evolución *in vitro* se ha usado previamente para el aislamiento de nuevos ARN y ADN catalíticos y, en la presente invención, se usa para el aislamiento nuevas enzimas proteicas. En un ejemplo particular de esta estrategia, un catalizador puede aislarse indirectamente por selección para la unión a un análogo químico del estado de transición del catalizador. En otro ejemplo particular, el aislamiento directo puede llevarse a cabo por selección para la formación de enlaces covalentes con un sustrato (por ejemplo, usando un sustrato unido a una etiqueta de afinidad) o por escisión (por ejemplo, por selección para determinar la capacidad para romper un enlace específico y de este modo liberar miembros catalíticos de una biblioteca de un soporte sólido).

Esta estrategia para el aislamiento de nuevos catalizadores tiene al menos dos ventajas importantes sobre la tecnología de anticuerpos catalíticos (revisada en Schultz y col., J. Chem. Engng. News 68: 26 (1990)). En primer lugar, en la tecnología de anticuerpos catalíticos, la combinación inicial se limita generalmente al plegamiento de inmunoglobulina; por el contrario, la biblioteca de partida de fusiones de ARN-proteína puede ser completamente aleatoria o puede consistir, sin limitación, en variantes de estructuras enzimáticas o armazones proteicos conocidos. Además, el aislamiento de anticuerpos catalíticos generalmente depende de una selección inicial para la unión a análogos de estado de transición, seguida de una exploración laboriosa para anticuerpos activos; de nuevo, por el contrario, la selección directa para catálisis es posible usando una estrategia de biblioteca de fusión de ARN-proteína como se ha demostrado previamente usando bibliotecas de ARN. En una estrategia alternativa para aislar enzimas proteicas, pueden combinarse estrategias de análogo de estado de transición y de selección directa.

Las enzimas obtenidas mediante este procedimiento son altamente valiosas. Por ejemplo, actualmente existe una necesidad urgente de catalizadores industriales novedosos y eficaces que permitan que se desarrollen procedimientos químicos mejorados. Una ventaja principal de la tecnología de fusión de ARN-proteína descrita en el presente documento es que pueden llevarse a cabo selecciones en condiciones arbitrarias y no están limitadas, por ejemplo, a las condiciones *in vivo*. La tecnología facilita por lo tanto el aislamiento de nuevas enzimas o variantes mejoradas de enzimas existentes que puedan llevar a cabo transformaciones altamente específicas (y de este modo minimizar la formación de subproductos no deseados), al tiempo que funcionan en entornos predeterminados, por ejemplo, entornos de elevada temperatura, presión o concentración de disolvente.

Una trampa de interacción *in vitro*. La tecnología de fusión de ARN-proteína también es útil para explorar genotecas de ADNc y clonar nuevos genes basándose en las interacciones de proteína-proteína. Mediante este procedimiento, se genera una genoteca de ADNc a partir de una fuente deseada (por ejemplo, mediante el procedimiento de Ausubel y col., anteriormente, capítulo 5). A cada uno de los ADNc candidatos se liga un aceptor peptídico (por ejemplo, como una cola de puromicina) (por ejemplo, usando las técnicas descritas anteriormente para la generación de LP77, LP154 y LP160). Después se generan fusiones de ARN-proteína como se describe en el presente documento, y la capacidad de estas fusiones (o versiones mejoradas de las fusiones) para interaccionar con moléculas particulares se ensaya después como se ha descrito anteriormente. Si se desea, pueden evitarse codones de terminación y regiones UTR 3' en este procedimiento por (i) adición de ARNt supresor para permitir la ultralectura de las regiones de terminación, (ii) eliminación del factor de liberación de la reacción de traducción por inmunoprecipitación, (iii) una combinación de (i) y (ii), o (iv) eliminación de los codones de terminación y UTR 3' de las secuencias de ADN.

El hecho de que la etapa de interacción tenga lugar *in vitro* permite el control cuidadoso de la rigurosidad de la reacción, usando condiciones inespecíficas de competidor, temperatura e iónicas. La alteración de moléculas pequeñas normales con análogos no hidrolizables (por ejemplo, ATP frente a ATPgS) proporciona selecciones que discriminan entre diferentes conformeros de la misma molécula. Esta estrategia es útil tanto para la clonación como para la identificación funcional de muchas proteínas, puesto que la secuencia de ARN del compañero de unión seleccionado se une covalentemente y por lo tanto puede aislarse fácilmente. Además, la técnica es útil para identificar funciones de interacciones de los ~50-100.000 genes humanos, cuyas secuencias se están determinando actualmente mediante el Proyecto del Genoma Humano.

**Uso de fusiones de ARN-proteína en un formato de microplaca**

Las "microplacas de ADN" consisten en matrices definidas espacialmente de oligonucleótidos inmovilizados o fragmentos clonados de ADNc o ADN genómico y tienen aplicaciones tales como secuenciación rápida y generación de perfiles de transcritos. Por hibridación de una mezcla de fusiones de ARN-proteína (por ejemplo, generadas a partir de una combinación de ADN o ARN celular) con dicha microplaca de ADN es posible generar una "microplaca de presentación de proteína" en la que cada mancha correspondiente a una secuencia inmovilizada es capaz de hibridar con su secuencia de ARN correspondiente en la combinación de fusiones de ARN-proteína. Mediante esta estrategia, la proteína correspondiente se inmoviliza de una forma definida espacialmente debido a su enlace a su propio ARNm, y las microplacas que contienen conjuntos de secuencias de ADN presentan los conjuntos de proteínas correspondientes. Como alternativa, pueden presentarse fragmentos peptídicos de estas proteínas si la biblioteca de fusión se genera a partir de fragmentos más pequeños de ADNc o ADN genómicos.

Dichas presentaciones ordenadas de proteínas y péptidos tienen muchos usos. Por ejemplo, representan herramientas potentes para la identificación de interacciones de proteína-proteína previamente desconocidas. En un formato específico, una proteína de sonda está marcada de forma detectable (por ejemplo, con un colorante fluorescente) y la proteína marcada se incuba con una microplaca de presentación de proteínas. Mediante esta estrategia, la identidad de proteínas que son capaces de unirse a la proteína de sonda se determina a partir de la localización de las manchas en la microplaca que se marcan debido a la unión de la sonda. Otra aplicación es la determinación rápida de proteínas que estén químicamente modificadas a través de la acción de enzimas modificadoras (por ejemplo, proteínas quinasas, acil transferasas y metil transferasas). Por incubación de la microplaca de presentación de proteínas con la enzima de interés y un sustrato marcado radiactivamente, seguida de lavado y autorradiografía, la localización y por tanto la identidad de esas proteínas que son sustratos para la enzima modificadora pueden determinarse fácilmente. Además, el uso de esta estrategia con presentaciones ordenadas de pequeños péptidos permite la localización adicional de dichos sitios de modificación.

La tecnología de presentación de proteínas puede llevarse a cabo usando matrices de ácidos nucleicos (incluyendo ARN, pero preferentemente ADN) inmovilizados en cualquier soporte sólido apropiado. Los soportes sólidos ejemplares pueden estar hechos de materiales tales como vidrio (por ejemplo, placas de vidrio), silicio o silicio-vidrio (por ejemplo, microplacas) u oro (por ejemplo, placas de oro). Se conocen bien en la técnica procedimientos para unir ácidos nucleicos a regiones exactas en dichas superficies sólidas, por ejemplo, procedimientos fotolitográficos, y pueden usarse para generar soportes sólidos (tales como microplacas de ADN) para su uso en la invención. Los procedimientos ejemplares para este fin incluyen, sin limitación, Schena y col., *Science* 270: 467-470 (1995); Kozal y col., *Nature Medicine* 2: 753-759 (1996); Cheng y col., *Nucleic Acids Research* 24: 380-385 (1996); Lipshutz y col., *BioTechniques* 19: 442-447 (1995); Pease y col., *Proc. Natl. Acad. Sci. USA* 91: 5022-5026 (1994); Fodor y col., *Nature* 364: 555-556 (1993); Pirrung y col., *Patente de Estados Unidos* N° 5.143.854; y Fodor y col., documento WO 92/10092.

35

## REIVINDICACIONES

**Reivindicaciones para los Estados Contratantes siguientes: AT, BE, CH, DK, ES, FR, GB, IT, LI, LU, MC, NL, SE,**

- 5 1. Una microplaca que comprende una matriz de ácidos nucleicos monocatenarios inmovilizados, estando dichos ácidos nucleicos hibridados con fusiones de ARN-proteína, en la que cada una de dichas fusiones de ARN-proteína comprende un ARN unido covalentemente en el extremo 3' del ARN a una proteína a través de un aceptor peptídico, y en la que dicha proteína está codificada por dicho ARN.
2. La microplaca de la reivindicación 1, en la que dicho ARN está unido covalentemente a dicho aceptor peptídico a través de un sitio de pausa.
- 10 3. La microplaca de la reivindicación 1 ó 2, en la que dicho aceptor peptídico se selecciona de puromicina, conjugado de ARNt-puromicina, fenilalanina-adenosina, tirosil adenosina, alanil adenosina, fenilalanil 3' desoxi 3' amino adenosina, alanil 3' desoxi 3' amino adenosina y tirosil 3' desoxi 3' amino adenosina, y en la que dicha proteína está codificada por dicho ARN unido covalentemente.
- 15 4. Una biblioteca de moléculas comprendiendo cada una un ácido ribonucleico unido covalentemente a una proteína codificada por dicho ácido ribonucleico.
5. La biblioteca de la reivindicación 4, en la que dicho ácido ribonucleico está unido covalentemente en el extremo 3' a dicha proteína a través de un aceptor peptídico.
6. La biblioteca de la reivindicación 5, en la que dicho aceptor peptídico es puromicina.
- 20 7. La biblioteca de una cualquiera de las reivindicaciones 4-6, en la que dicho ácido ribonucleico está unido covalentemente a dicho aceptor peptídico a través de un sitio de pausa.
8. La biblioteca de una cualquiera de las reivindicaciones 4-7, en la que dicha biblioteca comprende al menos 10<sup>9</sup> moléculas diferentes.
- 25 9. Una molécula que comprende una porción de ácido ribonucleico (ARN) y una porción de proteína codificada por dicha porción de ácido ribonucleico; en la que la porción de ácido ribonucleico se une covalentemente a dicha porción de proteína a través de un aceptor peptídico, y en la que la porción de ácido ribonucleico se une covalentemente en su extremo 3' a dicho aceptor peptídico a través de un sitio de pausa.
10. La molécula de la reivindicación 9, en la que el sitio de pausa es ácido desoxirribonucleico.
11. La molécula de la reivindicación 10, en la que el ácido desoxirribonucleico comprende una secuencia de oligo dA.
- 30 12. La molécula de la reivindicación 9, en la que el sitio de pausa es una combinación de ácido desoxirribonucleico y un resto no nucleotídico.
13. La molécula de la reivindicación 12, en la que el resto no nucleotídico comprende uno o más restos HO(CH<sub>2</sub>CH<sub>2</sub>O)<sub>3</sub>PO<sub>2</sub>.
14. La molécula de una cualquiera de las reivindicaciones 9-13, en la que el aceptor peptídico es puromicina.
- 35 15. Una molécula que comprende un ácido ribonucleico unido covalentemente a un anticuerpo, en la que dicho anticuerpo está totalmente codificado por dicho ácido ribonucleico.
16. La molécula de la reivindicación 15, en la que dicho anticuerpo se selecciona en el grupo que consiste en un anticuerpo humano, un anticuerpo humanizado y un anticuerpo de cadena sencilla.
- 40 17. La molécula de la reivindicación 15 ó 16, que comprende además un efector peptídico situado entre dicho ácido ribonucleico y dicho anticuerpo.
18. La molécula de una cualquiera de las reivindicaciones 9-17, en la que dicha molécula está inmovilizada en un soporte sólido.
19. La molécula de la reivindicación 18, en la que dicho soporte sólido es una microplaca.
- 45 20. Un ácido ribonucleico que comprende una secuencia de inicio de la traducción y un codón de inicio unido operativamente a una secuencia codificante de proteína candidata, estando dicho ácido ribonucleico unido covalentemente a un aceptor peptídico en el extremo 3' de dicha secuencia codificante de proteína candidata, que comprende además una secuencia de ADN o análogo de ADN unida covalentemente al extremo 3' del ácido ribonucleico.



21. El ácido ribonucleico de la reivindicación 20, en el que el aceptor peptídico es puromicina.

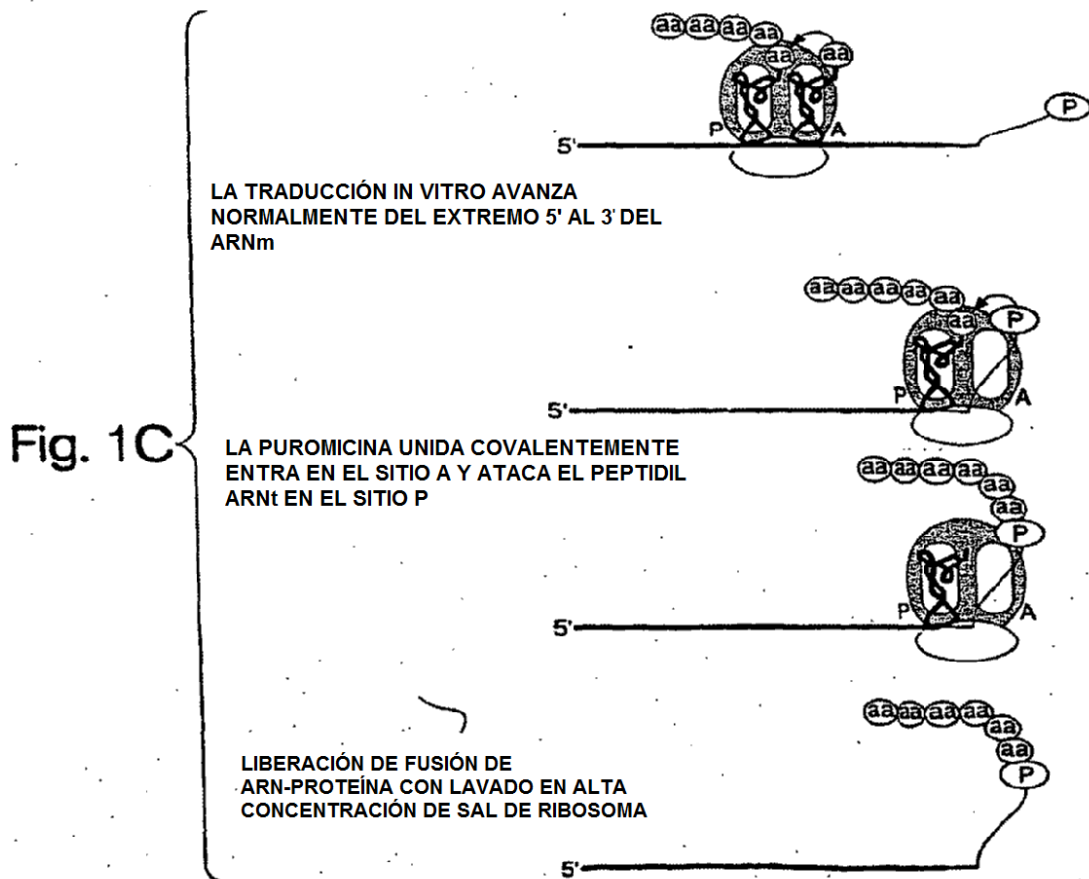
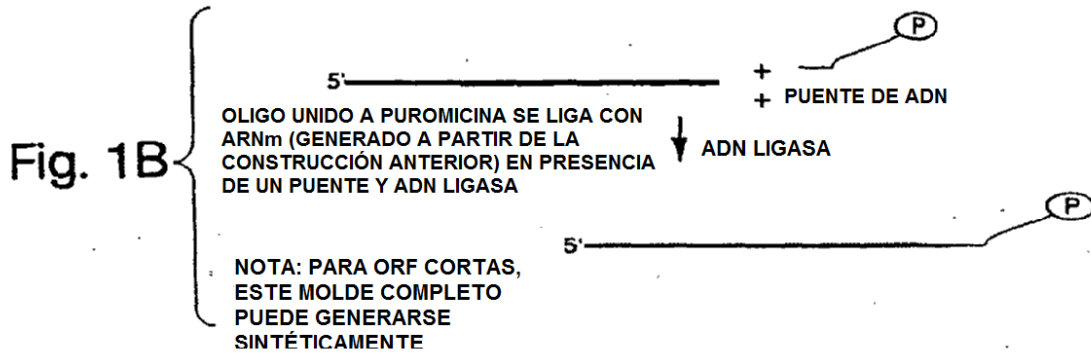
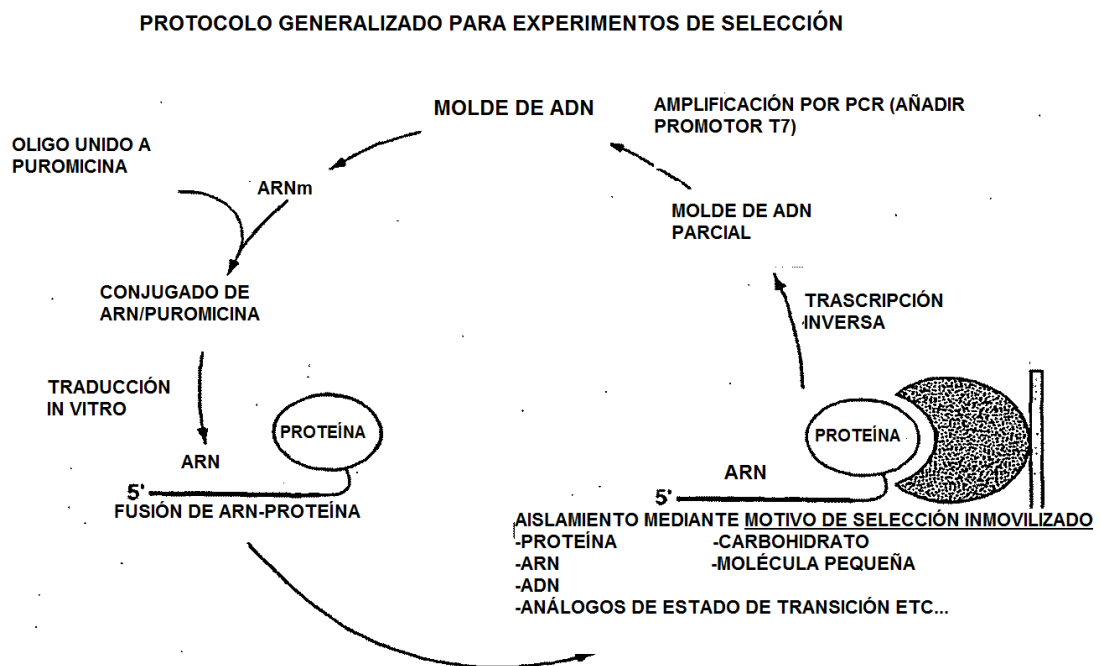


FIG. 2



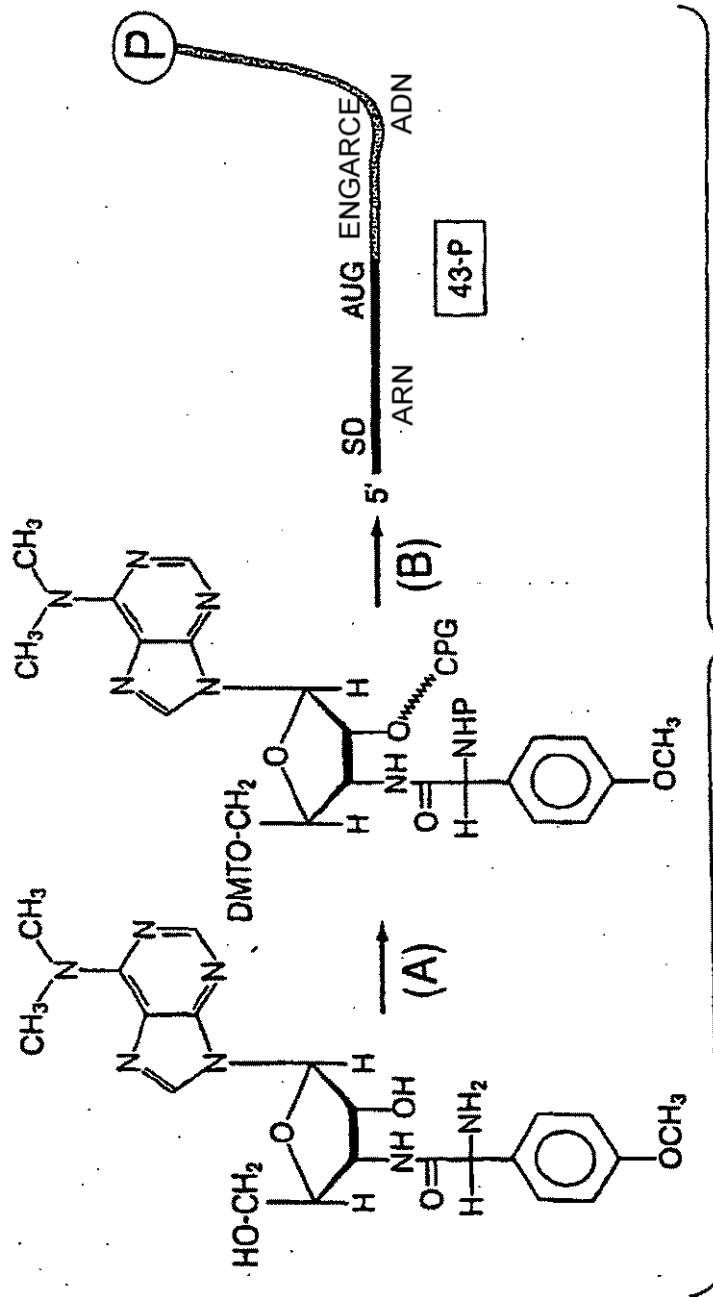


Fig. 3

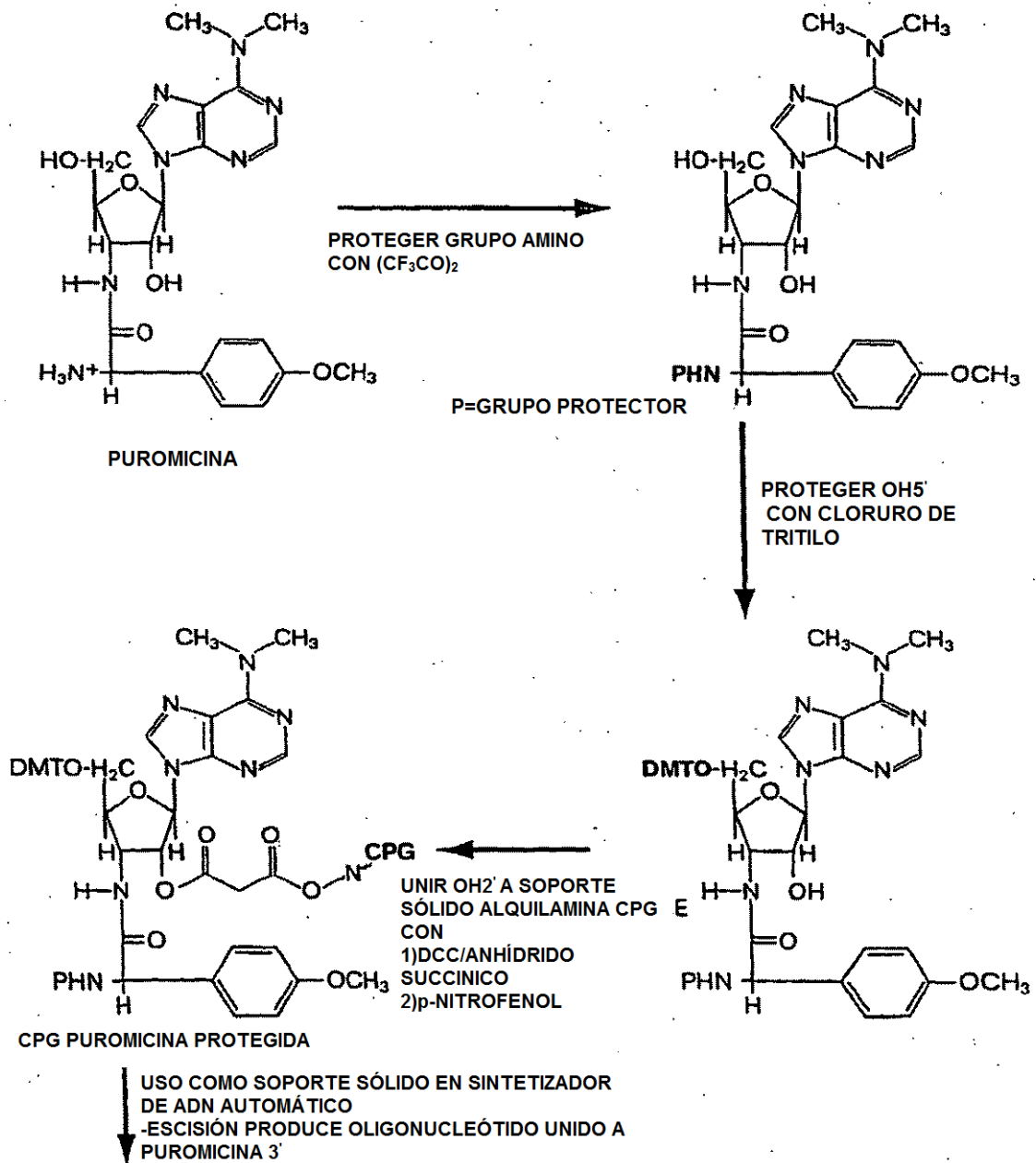


Fig. 4

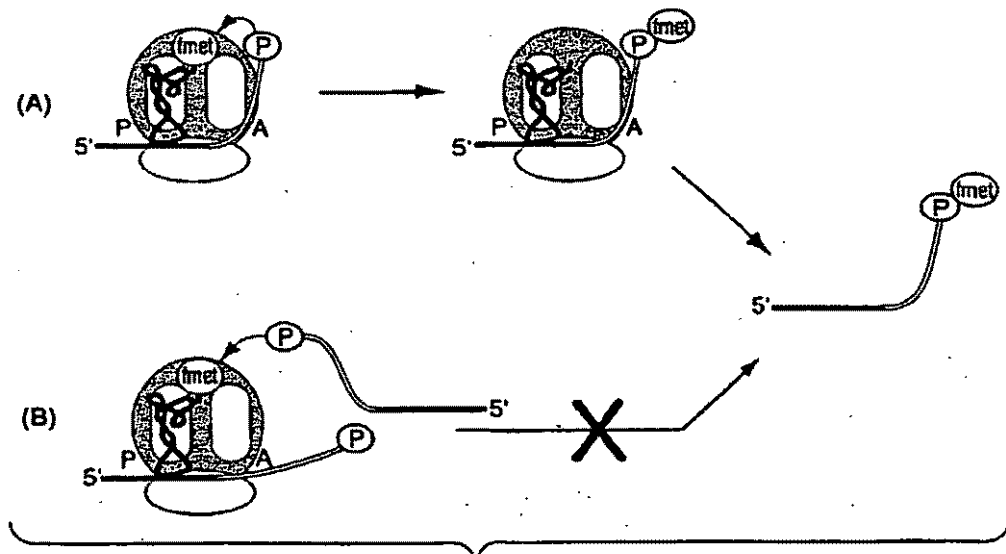


Fig. 5

Carril	1	2	3	4	5	6
Mg <sup>+2</sup> [mM]	6	9	12	15	18	

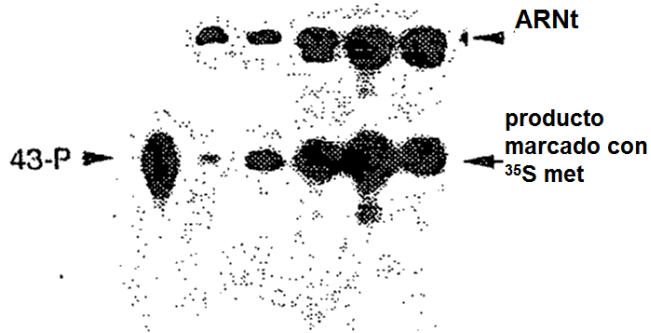


Fig. 6A

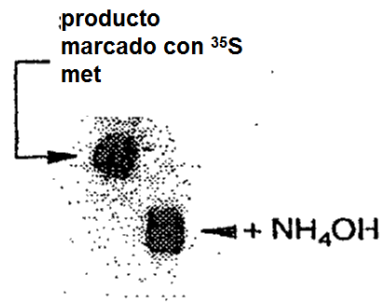


Fig. 6B

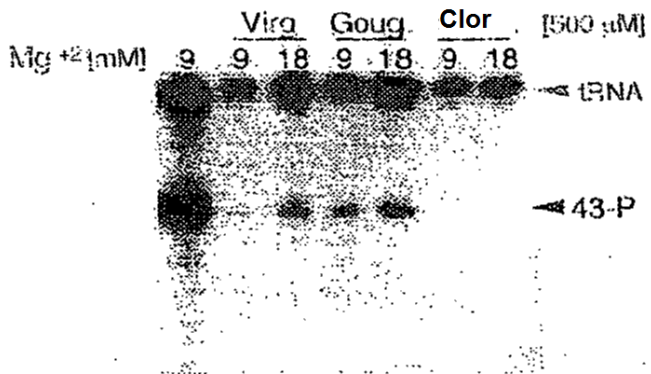


Fig. 6C

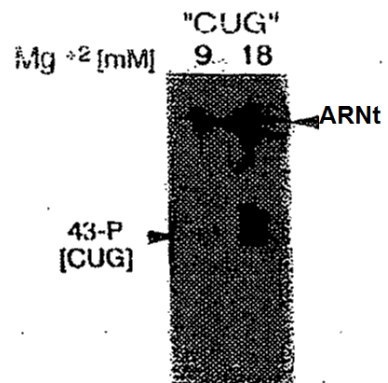


Fig. 6D

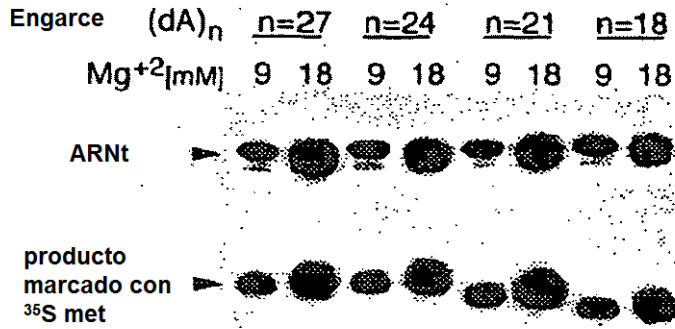


Fig. 6E

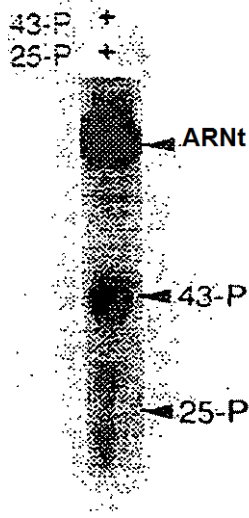


Fig. 6F

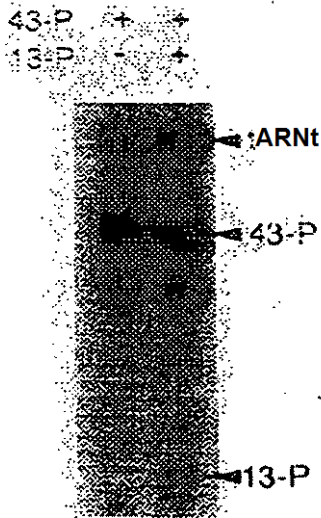


Fig. 6G

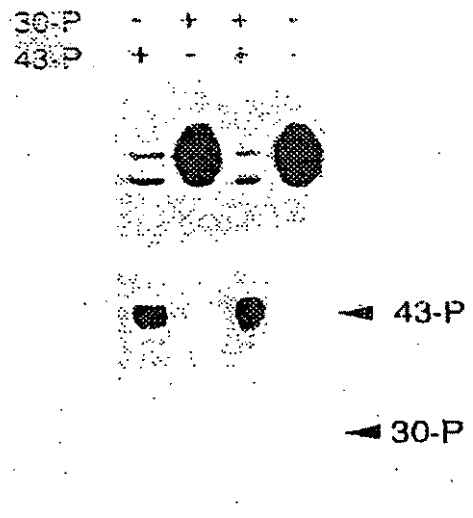


Fig. 6H



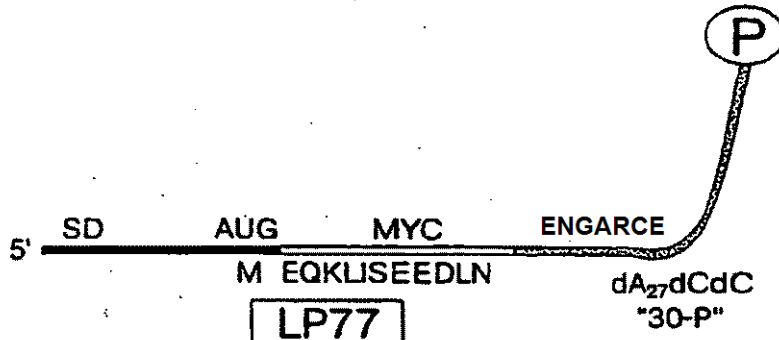


Fig. 7A

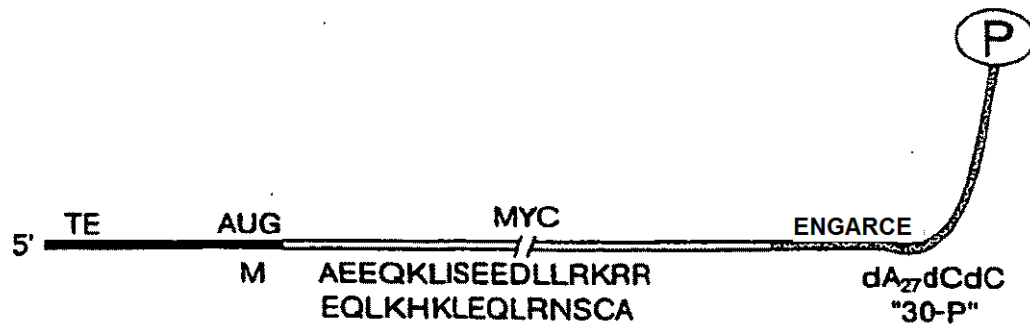


Fig. 7B

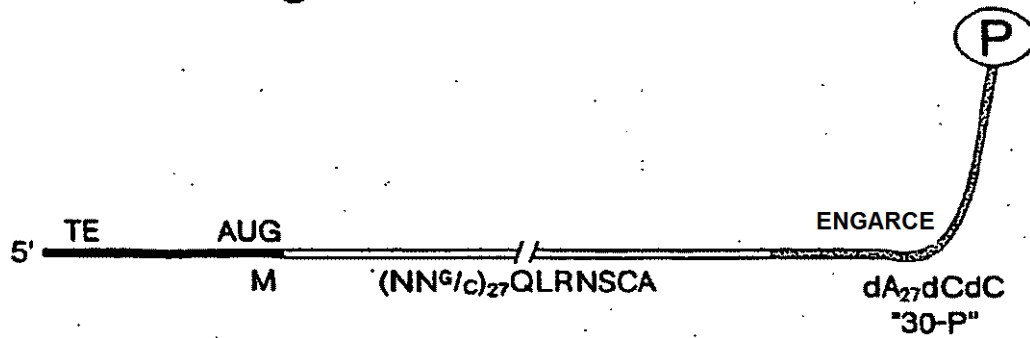


Fig. 7C

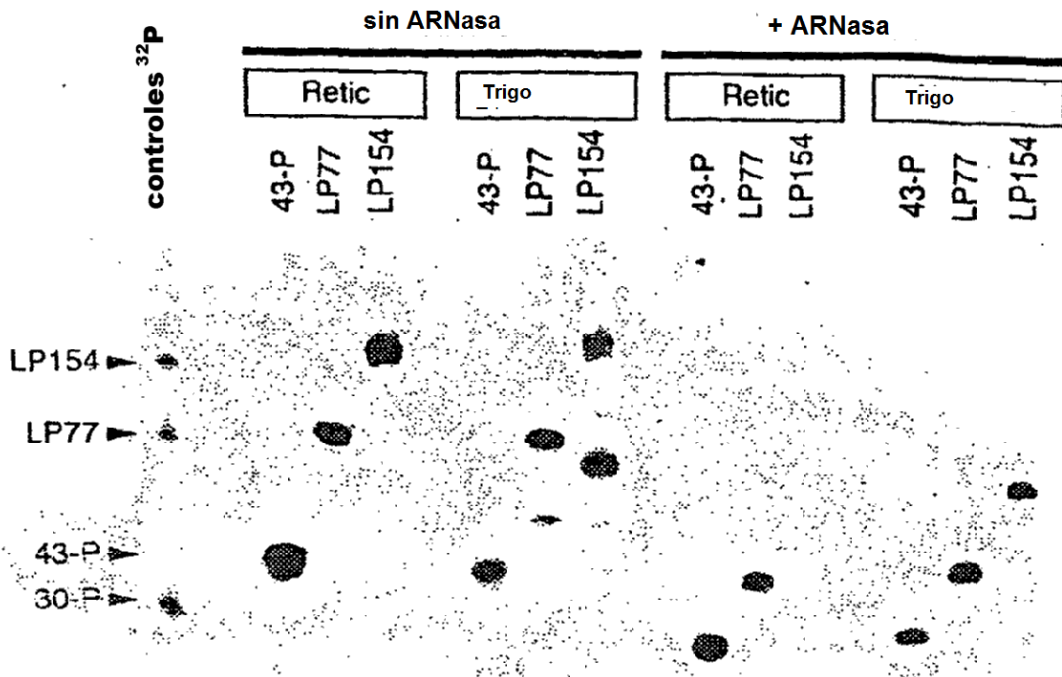


Fig. 8

Efectos de ARNasa A y Proteinas K sobre la fusión

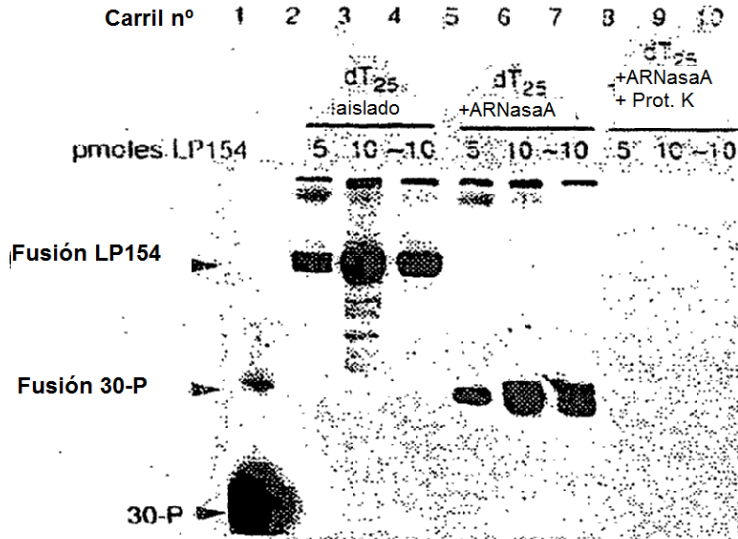
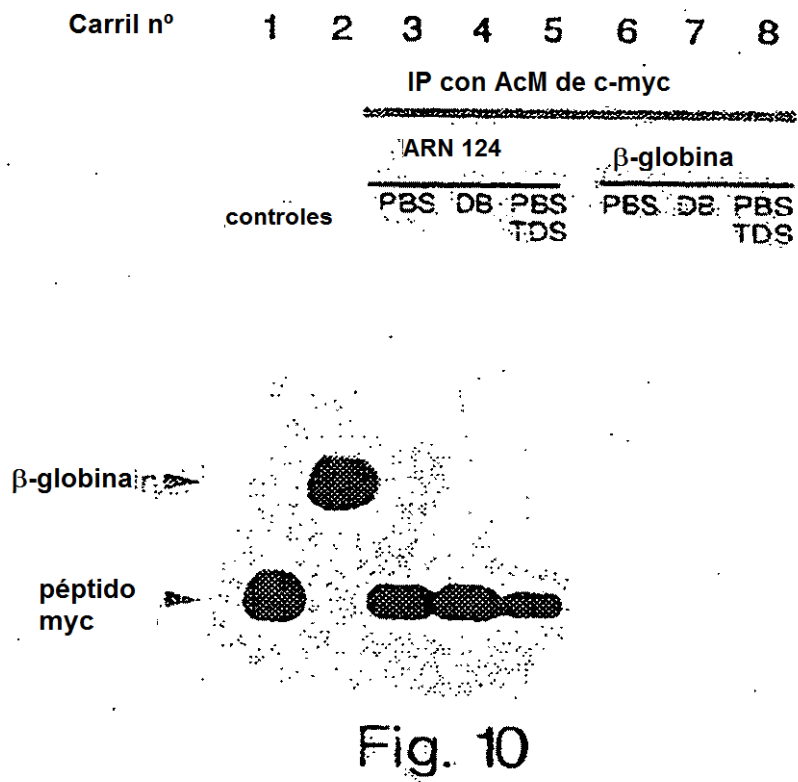


Fig. 9



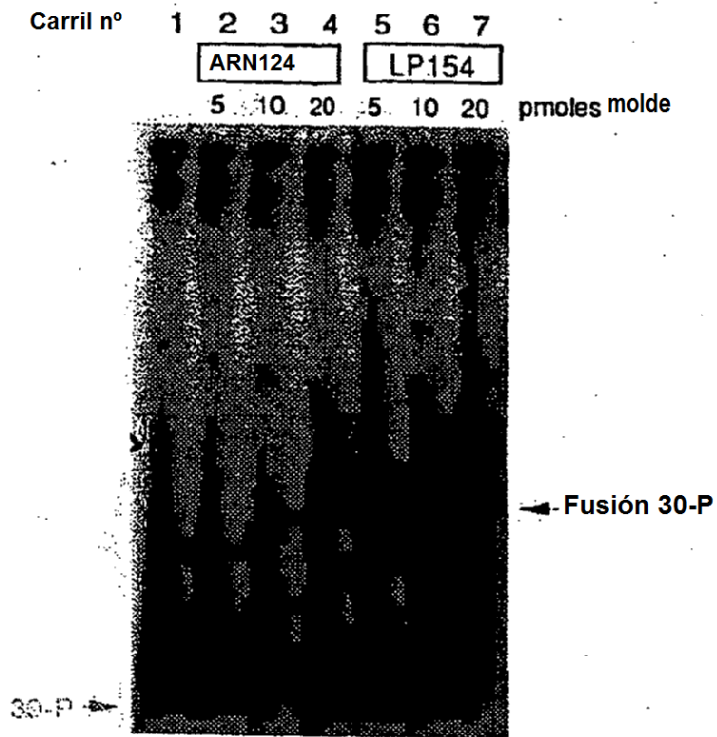


Fig. 11

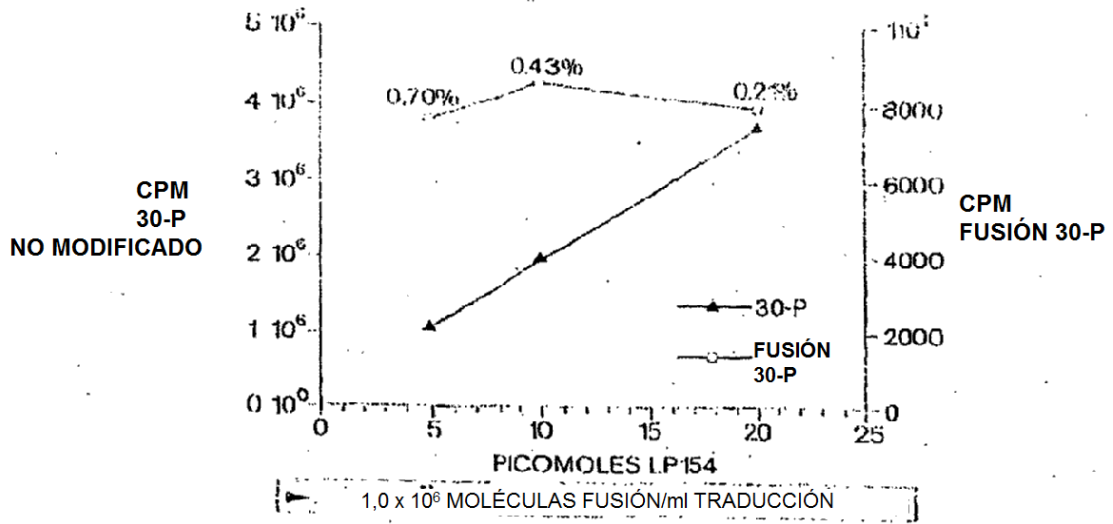


Fig. 12

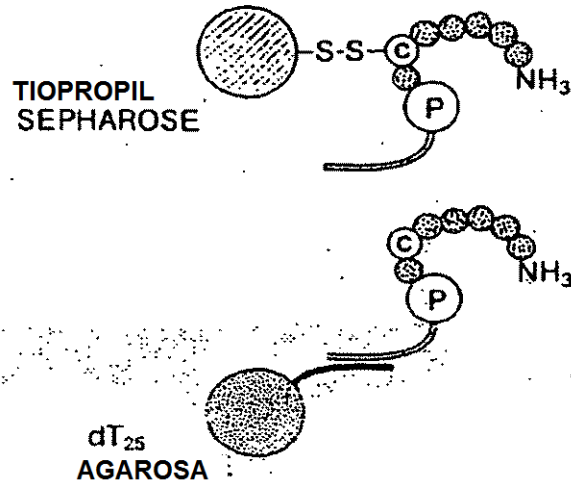


Fig. 13

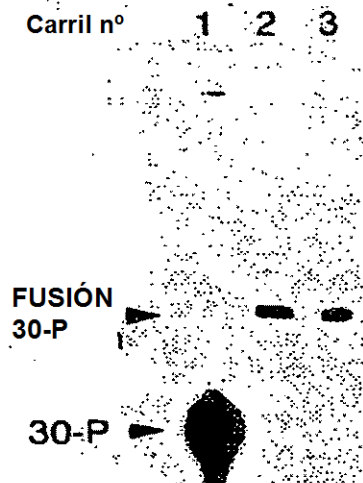


Fig. 14

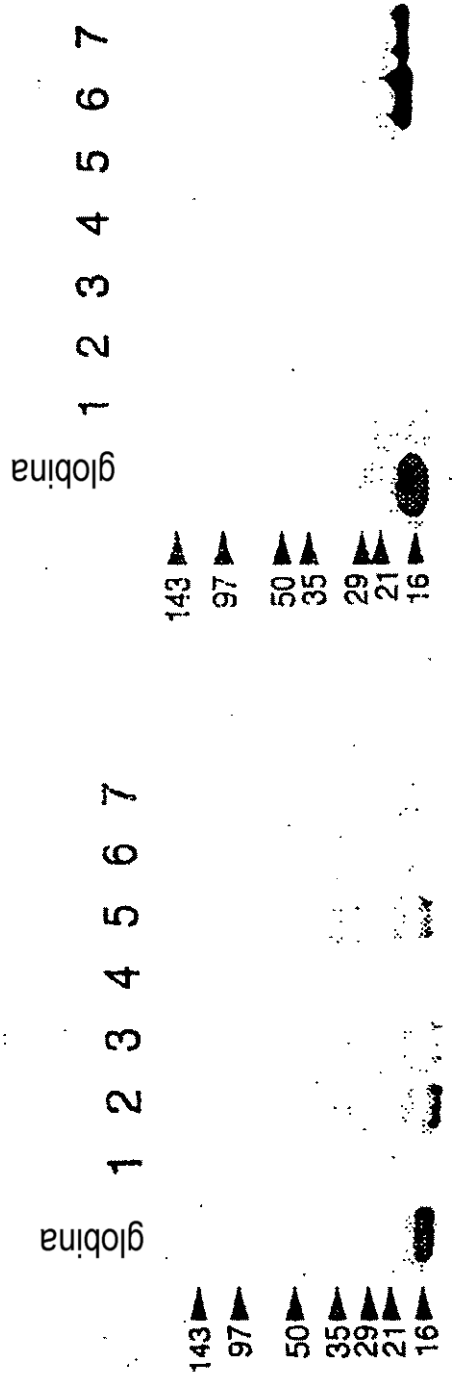


Fig. 15B

Fig. 15A

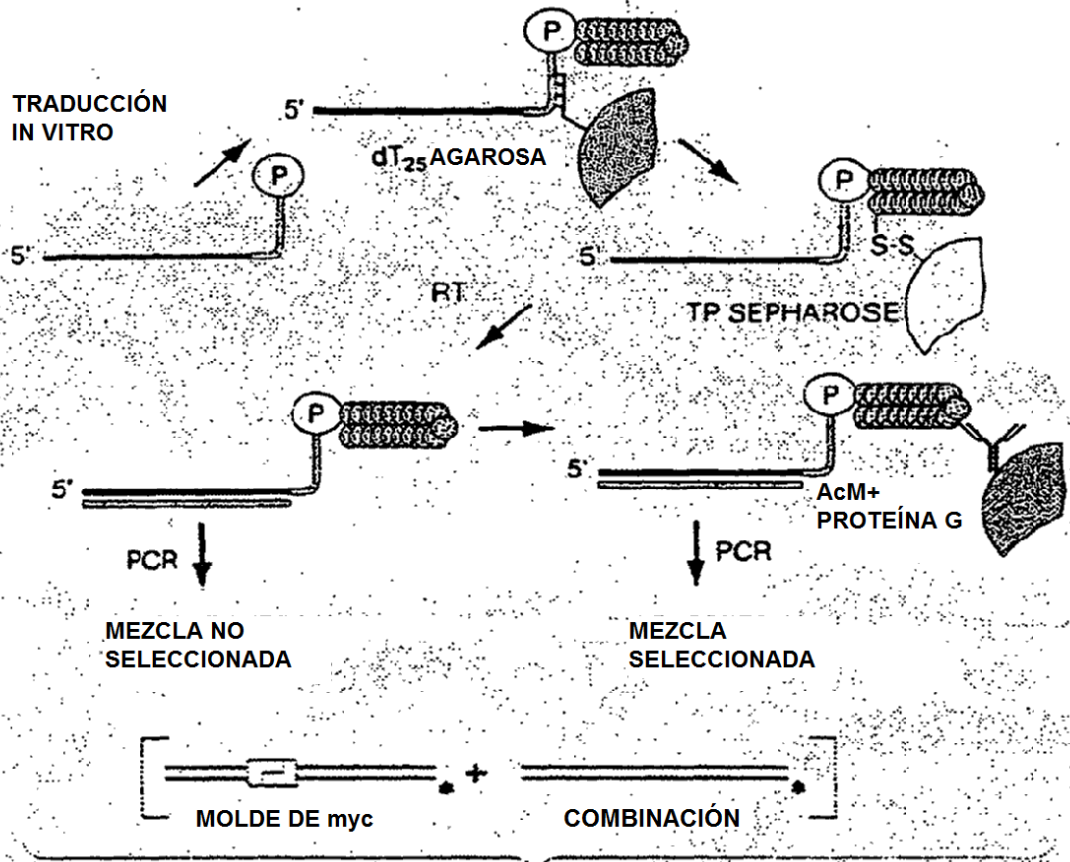


Fig. 16A

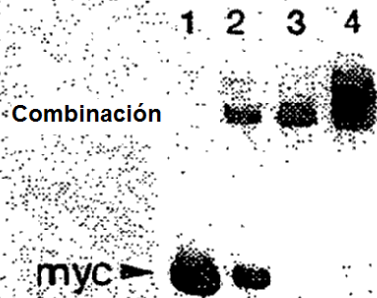


Fig. 16B



Fig. 16C

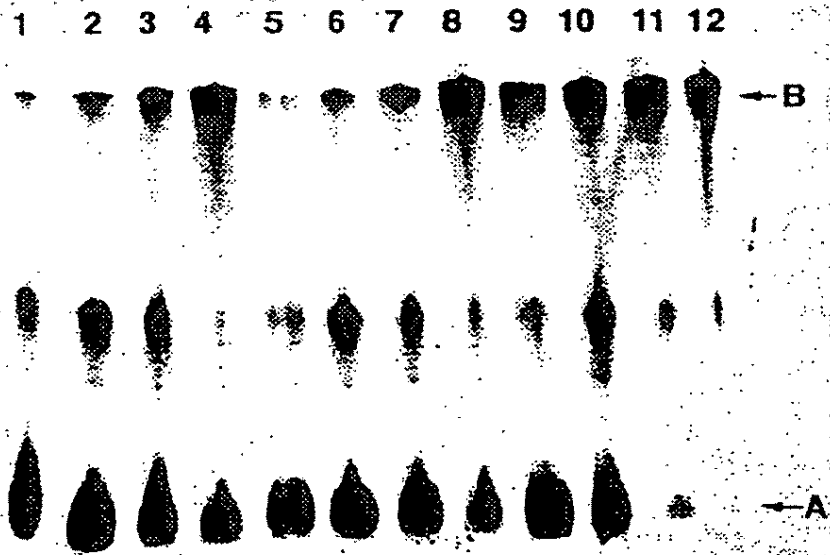


Fig. 17



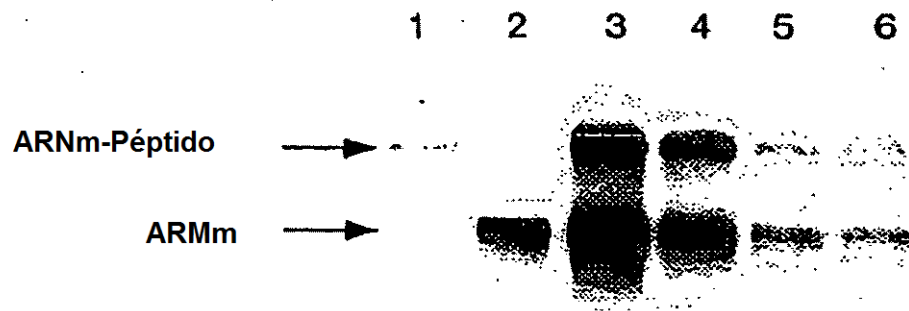


Fig. 18

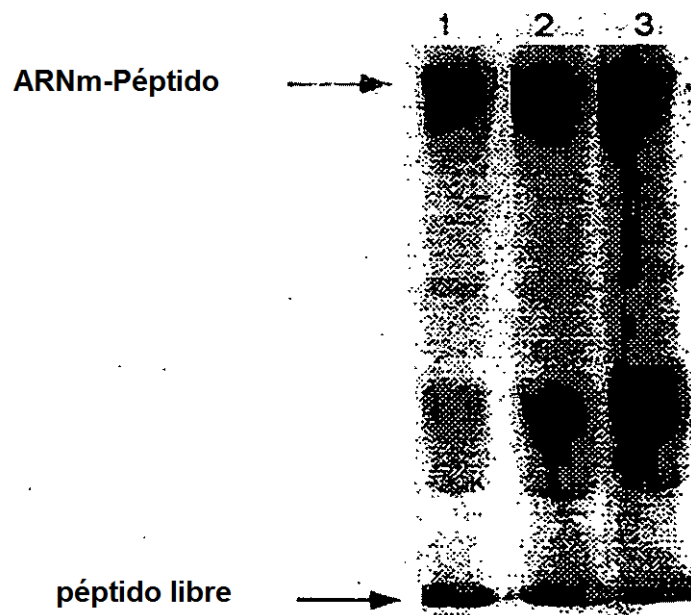


Fig. 19