

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 536 509**

51 Int. Cl.:

**C07K 16/12** (2006.01)

**C07K 16/42** (2006.01)

**C07K 14/195** (2006.01)

**C12N 15/10** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **04.12.2007 E 07872052 (1)**

97 Fecha y número de publicación de la concesión europea: **11.02.2015 EP 2099817**

54 Título: **OB-fold usado como estructura para diseño por ingeniería de nuevos agentes de unión específicos**

30 Prioridad:

**04.12.2006 EP 06291869**

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

**26.05.2015**

73 Titular/es:

**INSTITUT PASTEUR (50.0%)  
28 rue du Docteur Roux  
75015 Paris, FR y  
CENTRE NATIONAL DE LA RECHERCHE  
SCIENTIFIQUE (50.0%)**

72 Inventor/es:

**PECORARI, FRÉDÉRIC y  
ALZARI, PEDRO**

74 Agente/Representante:

**ISERN JARA, Jorge**

Observaciones :

**Véase nota informativa (Remarks) en el folleto original publicado por la Oficina Europea de Patentes**

**ES 2 536 509 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

OB-Fold usado como estructura para diseño por ingeniería de nuevos agentes de unión específicos

5 La presente invención se refiere al campo de la ingeniería de proteínas. Más específicamente, la invención proporciona un medio para obtener moléculas estables que se unan específicamente a una diana seleccionada entre una amplia diversidad de familias de ligandos.

10 La mayoría de las proteínas naturales no están adaptadas a usos terapéuticos o incluso biotecnológicos. El reto de la ingeniería de proteínas es definir modos eficaces y extensivos para el diseño de proteínas nuevas o mejoradas con las propiedades requeridas. Entre las propiedades abordadas, el reconocimiento de un ligando con alta especificidad y afinidad es de suprema importancia. Para muchas aplicaciones, se han usado anticuerpos a causa de extraordinaria adaptabilidad a la unión con muchos tipos diferentes de ligandos, tales como proteínas, péptidos, ácidos nucleicos, azúcares, etc. Sin embargo, los anticuerpos o sus fragmentos derivados a menudo son difíciles y caros de producir a causa de su complejidad molecular y/o su insuficiente estabilidad. Por estas razones, recientemente se han desarrollado alternativas a los anticuerpos usando proteínas estructurales modificadas por ingeniería mediante un enfoque de mutación combinatoria/selección. El objetivo fue mantener las propiedades favorables de los anticuerpos, eliminando al mismo tiempo sus desventajas. Se ha informado de las aplicaciones satisfactorias de esta estrategia (Binz et al., 2005; Mathonet y Fastrez 2004), siendo la inmensa mayoría de las dianas proteínas y, excepcionalmente, compuestos orgánicos pequeños.

25 La presente invención está dirigida a proporcionar herramientas para modificar por ingeniería moléculas bien adaptadas a usos médicos o biotecnológicos, y capaces de unirse específicamente y con alta afinidad a diversas dianas, seleccionadas entre ácidos nucleicos, proteínas, carbohidratos y cualquier otra molécula biológica. Para este objetivo, los inventores han ensayado la hipótesis de que una estructura proteica particular, llamada plegamiento de unión a oligonucleótido/oligosacárido (OB-fold), podría modificarse por aleatorización de los restos de su cara de unión, manteniendo al menos parcialmente al mismo tiempo las propiedades biofísicas favorables de la proteína precursora (es decir, alta estabilidad y eficacia de plegamiento).

30 El plegamiento de unión a oligonucleótido/oligosacárido (OB-fold), se describió por primera vez por Murzin (Murzin, 1993), se encuentra en todos los reinos y se clasificó como el 28º plegamiento más representado en un compendio de 20 genomas (Qian et al., 2001). Este plegamiento, un barril- $\beta$  de cinco hebras encapsulado con una  $\alpha$ -hélice anfífila, parece adecuado para una amplia diversidad de secuencias. El OB-fold presenta una cara de unión de lámina- $\beta$  que confiere diversidad remarcable en los tipos de compuestos que pueden reconocerse: ADNmc, ADNbc, ARN, oligosacáridos, proteínas, iones metálicos y sustratos catalíticos. Un estudio de varios OB-fold de diferentes proteínas mostró que el núcleo de unión está siempre localizado en la misma posición de la cara de unión (Arcus, 2002).

40 Sac7d tiene una topología OB-fold (figura 1a) (Agbaek et al., 1998; Gao et al., 1998; Robinson et al., 1998; Su et al., 2000). Pertenece a una clase de pequeñas proteínas cromosómicas de la archaea hipertermófila *Sulfolobus acidocaldarius*. Sac7d se une a ADN bicatenario sin ninguna preferencia de secuencia particular, induciendo al mismo tiempo una aguda torsión en el ADN (Robinson et al., 1998). Se cree que desempeña un papel en la estabilización de la hélice de ADN a alta temperatura de crecimiento de *Sulfolobus acidocaldarius* (óptima a 85°C). Sac7d es extremadamente estable. Es termoestable y se despliega con una  $T_m$  de 91°C, mantiene un plegamiento nativo entre pH 0 y 10, y su desnaturalización inducida por clorhidrato de guanidinio sucede de forma reversible con una concentración de punto medio de 2,8 M de desnaturalizante (McCrary et al., 1996). Contrario a los anticuerpos, su organización molecular es bastante simple. Sac7d es una proteína monomérica pequeña de 66 aminoácidos, con un dominio estructural (es decir, el OB-fold), y no tiene un enlace disulfuro (McAfee et al., 1995). Se han observado varias formas truncadas de Sac7d (McAfee et al., 1995). Pueden alcanzarse niveles de sobreproducción de hasta 50 10-15 mg de proteína soluble por litro de cultivo en matraz de *Escherichia coli* (Edmondson y Shriver, 2001). Estudios estructurales de Sac7d y su homólogo cercano Sso7d de *Sulfolobus solfataricus* han demostrado que los dos núcleos proteicos son superponibles y que la unión al surco menor del ADN sucede principalmente mediante un área girada formada por 16 restos (Agback et al., 1998; Gao et al., 1998; Robinson et al., 1998).

55 Los inventores han explorado la posibilidad de cambiar la especificidad de unión de Sac7d por evolución proteica dirigida *in vitro*, a través de la introducción de mutaciones aleatorias en varios restos implicados en la unión a ligando, seguido por una selección de variantes que se unen a una diana dada. De forma más precisa, los inventores han construido una gran biblioteca de aproximadamente  $3 \times 10^{12}$  variantes correspondientes a sustituciones aleatorias de 11 restos de la interfaz de unión de Sac7d. Esta biblioteca se ha usado para seleccionar variantes capaces de unirse a una diana proteica definida, por presentación en ribosoma. Se obtuvieron combinaciones de agentes de unión para tres dianas proteicas, con afinidades en el intervalo de cien nanomolar. En un segundo enfoque, los inventores ampliaron el área de unión potencial hasta 13 y 14 restos. Estas nuevas bibliotecas se usaron para la selección sobre una de las dianas previas (PulD-N), y se obtuvieron agentes de unión específica con afinidad en el intervalo picomolar. Por tanto, los inventores han demostrado que, sorprendentemente, 65 un agente de unión general a ADN (Sac7d) puede ir evolucionando hacia un agente de unión específico a proteína que tiene alta especificidad y afinidad. Esta demostración proporciona una prueba del principio de que los agentes

de unión para casi cualquier tipo de ligando pueden obtenerse de Sac7d u otras proteínas OB-fold, por ejemplo usando los procesos descritos en más detalle a continuación.

Un primer aspecto de la presente solicitud describe el uso de una proteína OB-fold, como molécula de partida para obtener, a través de un enfoque de mutación combinatoria/selección, una molécula que se une específicamente a una diana, especialmente a una diana a la cual no se une la proteína OB-fold de partida, es decir, a una diana diferente de la diana de la proteína OB-fold usada como la molécula de partida. A continuación, la diana de la proteína OB-fold de partida se llamará la "diana nativa", mientras que la diana usada en la etapa de selección se denominará como "diana de interés". En una realización preferida, la diana de interés es de una naturaleza química diferente de la de la diana nativa (por ejemplo, proteína frente a ácido nucleico, e ión metálico o azúcar frente a proteína). Un aspecto particular de la presente solicitud describe el uso de una proteína OB-fold que se une de forma natural a un ácido nucleico, como molécula de partida para obtener, a través de un enfoque de mutación combinatoria/selección, una molécula que se une específicamente a una diana diferente de un ácido nucleico (por ejemplo, una proteína). El uso de una proteína OB-fold que se une de forma natural a una proteína, como molécula de partida para obtener, a través de un enfoque de mutación combinatoria/selección, una molécula que se une específicamente a una diana no proteica (por ejemplo, un ácido nucleico) también es parte de la presente invención. Dichas proteínas OB-fold, de acuerdo con la invención, son proteínas Sac7d y proteínas de su familia como se describe a continuación.

"Proteína OB-fold" denomina cualquier polipéptido que comprenda o consista en un dominio que tenga topología OB-fold, como se describe por (Murzin, 1993) y (Arcus, 2002). Esta topología corresponde a una arquitectura que comprende un barril- $\beta$  de cinco hebras encapsulado en un extremo por una hélice- $\alpha$  anfilila. Con referencia a la clasificación de estructuras proteicas CATH (Pearl et al., 2003), la topología OB-fold corresponde a una familia de plegamiento 2.40.50 (base de datos CATH versión 3.0.0: Publicada en mayo de 2006). Dicha proteína OB-fold puede ser una proteína nativa (es decir, una proteína aislada, purificada o recombinante que tiene la misma secuencia que una proteína natural), o una proteína modificada por ingeniería (como, por ejemplo, un fragmento de una proteína nativa, o una proteína de fusión que comprende un dominio OB-fold de una primera proteína, y otro resto de otra proteína). Ejemplos no limitantes de proteínas OB-fold que pueden usarse de acuerdo con la invención son Sac7d, Sso7d, el dominio N-terminal de SEB (Papageorgiou et al., 1998), la cadena A de la toxina tipo Shiga Ile (PDB 2bosa), el péptido 2 de activación de neutrófilos humanos (NAP-2, PDB 1tvxA), la proteína de unión a molibdeno (modg) de *Azotobacter vinelandii* (PDB h9j), el dominio N-terminal de SPE-C (Roussel et al., 1997), la subunidad B<sub>5</sub> de la toxina tipo Shiga de *E. coli* Shiga (Kitov et al., 2000), Cdc13 (Mitton-Fry et al., 2002), el dominio de unión a ADN por choque frío de la proteína YB-1 de caja Y humana (Kloks et al., 2002), la pirofosfatasa inorgánica de *E. coli* EPPasa (Samyгина et al., 2001), o cualquiera de las proteínas enumeradas en la Tabla 3 del artículo de (Arcus, 2002), tales como 1krs (Lisil-ARNt sintetasa LysS, *E. coli*), 1c0aA (Asp-ARNt sintetasa, *E. coli*), 1b8aA (Asp-ARNt sintetasa, *P. kodakaraensis*), 1lylA (Lisil-ARNt sintetasa LysU, *E. coli*), 1quqA (proteína de replicación A, subunidad de 32kDa, ser humano), 1quqB (proteína de replicación A, subunidad de 14kDa, ser humano), 1jmcA (proteína de replicación A, fragmento de subunidad de 70kDa (RPA70), ser humano), 1otc (proteína de unión a extremo telomérico, *O. nova*), 3ulla (proteína de unión a ADNmc mitocondrial, ser humano), 1prtF (subunidad S5 de toxina pertussis, *B. pertussis*), 1bcpD (subunidad S5 de toxina pertussis (unida a ATP), *B. pertussis*), 3chbD (Toxina colérica, *V. cholerae*), 1tiiD (toxina inestable al calor, *E. coli*), 2bosA (Verotoxina-1/toxina Shiga, pentámero-B, *E. coli*), 1br9 (TIMP-2, ser humano), 1an8 (Superantígeno SPE-C, *S. pyogenes*), 3seb (Superantígeno SPE, *S. aureus*), 1aw7A (toxina de síndrome de choque tóxico, *S. aureus*), 1jmc (proteína principal de choque frío, *E. coli*), 1bbk (factor de inicio de la traducción 5a, *P. aerophilum*), 1sro (dominio de unión a ARN S1 de PNPasa, *E. coli*), 1d7qA (factor de inicio de la traducción 1, e1F1a, ser humano), 1ah9 (factor de inicio de la traducción 1, IF1, *E. coli*), 1b9mA (regulador transcripcional dependiente de ModE, *E. coli*), 1ckmA (ARN guaniltransferasa, virus Chlorella, PBCV-1), 1a0i (ADN ligasa dependiente de ATP, Bacteriófago T7), 1snc (nucleasa de estafilococos, *S. aureus*), 1hjp (subunidad RuvA de ADN helicasa, dominio N-terminal, *E. coli*), 1pfsA (proteína de gen V, bacteriófago pf3 de *Pseudomonas*), 1gvp (proteína de gen V, bacteriófago filamentoso (f1, M13)), 1gpc (núcleo de proteína de gen 32 (gp32), Bacteriófago T4), 1wgjA (pirofosfatasa inorgánica, *S. cerevisiae*), y 2prd (pirofosfatasa inorgánica, *T. thermophilus*). Puede apreciarse que los dominios OB-fold que se originan a partir de toxinas pueden usarse como moléculas de partida incluso para propósitos en los que debe evitarse la toxicidad, ya que mutaciones en su sitio de unión, y por tanto cambios en su especificidad de unión, pueden anular completamente su toxicidad.

Como se ejemplifica en más detalle en la siguiente parte experimental, el enfoque de mutación combinatoria/selección consiste en obtener una biblioteca combinatoria correspondiente a la aleatorización de varios restos elegidos de la proteína OB-fold de partida (especialmente, aleatorización de varios restos implicados en la unión de la proteína con su ligando nativo), seguido por una selección, en dicha biblioteca, de variantes que tienen las propiedades deseadas.

Un objeto de la presente invención es por tanto una biblioteca combinatoria correspondiente a la aleatorización de 5 a 32, preferiblemente de 8 a 20, y más preferiblemente de 11 a 16 restos de la interfaz de unión de la proteína OB-fold con su ligando nativo, posiblemente combinada con la delección de 1 a 4 restos y/o la inserción de 1 a 50 restos. Por supuesto, la "interfaz de unión de una proteína OB-fold" en este documento indica, incluso en casos de proteínas con múltiples dominios de unión a diferentes ligandos, la interfaz entre el dominio OB-fold y el ligando que

se une a este dominio. Los especialistas en la técnica encontrarán en la bibliografía científica la información necesaria para identificar los restos implicados en la unión de la proteína OB-fold con su ligando nativo, que a menudo están localizados en las láminas- $\beta$ ,  $\beta$ 3,  $\beta$ 4 y  $\beta$ 5 y en los bucles 1, 3 y 4 del OB-fold (figuras 1b y 2).

5 Las bibliotecas combinatorias de acuerdo con la presente invención se obtienen con Sac7d de *Sulfolobus acidocaldarius*, o con sus homólogos (figura 4) tales como Sso7d de *Sulfolobus solfataricus*. Por supuesto, las bibliotecas obtenidas con una forma truncada de Sac7d, tales como las descritas por (McAfee et al., 1995), también son parte de la presente invención.

10 Por superposición de varias secuencias y estructuras 3D de dominios OB-fold usando los sitios web (<http://www.ebi.ac.uk/blast2/index.html>) (Lopez et al., 2003), T-COFFEE (<http://www.ch.embnet.org/software/TCoffee.html>) (Notredame et al., 2000) y DALI lite (<http://www.ebi.ac.uk/DaliLite/>) (Holm y Park, 2000), los inventores han identificado las posiciones que tienen que modificarse para obtener las bibliotecas de acuerdo con la presente invención. Tomando como referencia la secuencia de Sac7d de la SEC ID N° 1, los restos que pueden aleatorizarse son los siguientes: V2, K3, K5, K7, Y8, K9, G10, E14, T17, K21, K22, W24, V26, G27, K28, M29, S31, T33, D36, N37, G38, K39, T40, R42, A44, S46, E47, K48, D49, A50 y P51.

Aún con Sac7d como referencia, los restos que pueden delecionarse son: A59, R60, A61 y E64.

20 Una superposición de estructuras 3D de 10 proteínas o dominios OB-fold (incluyendo Sac7d), usando el sitio web DALI (<http://www.ebi.ac.uk/dali/Interactive.html>) (Holm y Sander, 1998), reveló que este tipo de proteínas tiene bucles de diversos tamaños, que muy probablemente están implicados en la unión de ligandos (véase la figura 2). Esta observación es coherente con los datos publicados previamente (Arcus, 2002). Como Sac7d es una de las proteínas que tiene los bucles más cortos, pueden realizarse de forma ventajosa inserciones aleatorias en dichos bucles para obtener bibliotecas particularmente adaptadas para la unión de ciertas familias de ligandos. De forma ventajosa, pueden realizarse inserciones de 1 a 15 restos aminoacídicos en el bucle 3 (identificado en las figuras 1b y 2), por ejemplo en la región de los restos 25 a 30 de Sac7d, preferiblemente entre los restos 27 y 28, pueden realizarse inserciones de 1 a 15 restos aminoacídicos en el bucle 4 (identificado en las figuras 1b y 2), por ejemplo en la región de los restos 35 a 40 de Sac7d, preferiblemente entre los restos 37 y 38, y pueden realizarse inserciones de 1 a 20 restos aminoacídicos en el bucle 1 (identificado en las figuras 1b y 2), por ejemplo en la región de los restos 7 a 12 de Sac7d, preferiblemente entre los restos 9 y 10. Para evitar cualquier ambigüedad, en este documento se especifica que los bucles 3, 4 y 1 identificados en las figuras 1a y 2 corresponden respectivamente a los bucles 1, 2 y 4 identificados en el artículo de revisión de Arcus (*supra*).

35 De acuerdo con una realización particular de la invención, la biblioteca combinatoria corresponde a la aleatorización de 11, 12, 13, 14, 15, 16, 17 ó 18 restos de Sac7d seleccionados entre K7, Y8, K9, K21, K22, W24, V26, K28, M29, S31, T33, K39, T40, R42, A44, S46, E47 y K48, por ejemplo 11, 12, 13, 14, 15 ó 16 restos seleccionados entre K7, Y8, K9, K21, K22, W24, V26, K28, M29, S31, T33, K39, T40, R42, A44 y S46.

40 En una biblioteca combinatoria preferida de la realización anterior, los restos aleatorizados comprenden al menos los restos K7, Y8, K9, W24, V26, M29, S31, T33, R42, A44 y S46 de Sac7d. Una biblioteca específica de acuerdo con esta realización, y denominada en la siguiente parte experimental como "biblioteca 11", corresponde a la aleatorización de los restos K7, Y8, K9, W24, V26, M29, S31, T33, R42, A44 y S46 de Sac7d. En otras bibliotecas combinatorias de acuerdo con esta realización, también se aleatorizan 1, 2 ó 3 restos adicionales de Sac7d seleccionados entre K21, K22 y T40. Otras dos bibliotecas preferidas también descritas en la parte experimental, son la "biblioteca 13", que corresponde a la aleatorización de los restos K7, Y8, K9, K21, K22, W24, V26, M29, S31, T33, R42, A44 y S46 de Sac7d, y la "biblioteca 14", que corresponde a la aleatorización de los restos K7, Y8, K9, K21, K22, W24, V26, M29, S31, T33, T40, R42, A44 y S46 de Sac7d.

50 Por supuesto, una biblioteca combinatoria obtenida aleatorizando 11, 12, 13, 14, 15 ó 16 restos de Sso7d seleccionados entre los restos localizados en las posiciones que son equivalentes a las enumeradas anteriormente, también es parte de la presente invención. Dichas posiciones equivalentes pueden identificarse usando el archivo para Sso7d (cadena A) en el banco de datos de proteínas RSCB (Berman et al., 2000) (1bf4A), y comparando su estructura 3D con la de Sac7d (figura 7a):

55 Sac7d: 7 8 9 21 22 24 26 28 29 31 33 39 40 42 44 46  
1bf4A: 7 8 9 21 22 24 26 28 29 31 33 40 41 43 45 47

60 Otras proteínas homólogas a Sac7d también pueden usarse para obtener bibliotecas combinatorias como se ha descrito anteriormente. Ejemplos de dichas proteínas se describen en la siguiente Tabla.

Nombre	Fuente	Número de acceso principal	Número de acceso secundario
Sac7d: (proteína 7d de unión a ADN)	<i>Sulfolobus acidocaldarins</i> NCBI: AAA80315	P13123	Q4JCI7
Sac7e: (proteína 7e de unión a ADN)	<i>Sulfolobus acidocaldarius</i> NCBI: YP 255071	P13125:	Q4JBQ1
Proteína 7 de unión a ADN (DBP 7)	<i>Sulfolobus tokodaii</i> NCBI: Q96X56	Q96X56	
Ssh7b (proteína 7b de unión a ADN)	<i>Sulfolobus shibatae</i> NCBI: BAA28275	O59632	
Sso7d (proteína 7d de unión a ADN)	<i>Sidfolobus solfataricus</i> NCBI: P39476	P39476	P81550
Ssh7a (proteína 7a de unión a ADN)	<i>Sulfolobus shibatae</i> NCBI: BAA28274	P61990	O59631, P80170 Q9UWI8
p7ss (proteína 7a de unión a ADN)	<i>Sulfolobus solfataricus</i> NCBI: P61991	P61991	O59631, P80170, Q9UWI8

Tabla 1: Ejemplos de homólogos de Sac7d que pueden usarse de acuerdo con la presente invención

5 La alineación de los dominios OB-fold de los homólogos de Sac7d enumerados en la Tabla 1 se muestra en la figura 4.

10 Otras bibliotecas combinatorias particulares de acuerdo con la presente invención pueden obtenerse con la toxina Ile tipo Shiga (PDB 1r4pB), por ejemplo aleatorizando 11, 12, 13, 14, 15 ó 16 restos seleccionados entre los restos localizaos en las posiciones que son equivalentes a las enumeradas anteriormente para Sac7d, siendo dichas posiciones equivalentes (figura 3b) las siguientes:

15 Sac7d: 7 8 9 21 22 24 26 28 29 31 33 39 40 42 44 46  
1r4pB: 60 59 58 9 10 12 14 17 18 20 22 X X 27 29 31

20 También se describen bibliotecas combinatorias que pueden obtenerse aleatorizando 11, 12, 13, 14, 15 ó 16 restos del Péptido-2 de activación de neutrófilos (NAP-2, PDB 1tvxA), donde dichos restos se seleccionan entre los restos localizados en las posiciones que son equivalentes a las enumeradas anteriormente para Sac7d. Estas posiciones equivalentes, obtenidas por comparación de la estructura 3D de dicha proteína con la de Sac7d (figura 3c), son las siguientes:

25 Sac7d: 7 8 9 21 22 24 26 28 29 31 33 39 40 42 44 46  
1tvxA: 25 26 27 40 41 43 45 54 55 57 59 61 63 65 67 69

30 También se describen bibliotecas combinatorias que pueden obtenerse con la proteína de unión a molibdeno (modg) de *Azotobacter vinelandii* (PDB 1h9j), por ejemplo aleatorizando 11, 12, 13, 14, 15 ó 16 restos seleccionados entre los restos localizados en posiciones que son equivalentes a las enumeradas anteriormente para Sac7d, donde dichas posiciones equivalentes (figura 3d) son las siguientes:

35 Sac7d: 7 8 9 21 22 24 26 28 29 31 33 39 40 42 44 46  
1h9jA: 60 61 62 86 87 89 91 93 94 96 98 105 106 108 110 112

40 La presente invención también se refiere al uso de una biblioteca combinatoria como las descritas anteriormente, para modificar por ingeniería una molécula que se una específicamente a una diana. En el presente texto, una molécula se considera un agente de unión "específico" de una diana, si se une al mismo con una proporción de señal a ruido superior a 10. Una molécula de unión que se modifica por ingeniería usando una biblioteca combinatoria de acuerdo con la invención se une preferiblemente a su diana (también llamada "diana de interés") con una alta afinidad, es decir, con una afinidad mejor de 10 nM cuando la diana es un péptido o una proteína, y mejor de 1 µM cuando la diana es una molécula de carbohidrato, por ejemplo.

45 Otro objeto de la invención es un proceso para obtener una molécula que se una específicamente a una diana, que comprende una etapa de selección, en una biblioteca combinatoria como la descrita anteriormente, de aquellas variantes de dicha proteína OB-fold que se unen específicamente a dicha diana. Por "variante" en este documento se entiende una proteína que pertenece a la biblioteca, es decir, una proteína que se obtiene de la proteína OB-fold de partida por mutaciones en su sitio de unión.

Como se ha mencionado anteriormente, uno de los intereses de la presente invención es que posibilite modificar por ingeniería una molécula que se una específicamente a una diana de interés, casi sin ninguna limitación respecto a

su diana. Por ejemplo, la diana de interés puede ser un péptido, una proteína, un oligosacárido, un lípido, un lipopéptido, un carbohidrato (por ejemplo, un azúcar), un ADN monocatenario, un ADN bicatenario, un ARN. De forma interesante, la diana de interés puede ser una molécula pequeña natural o sintética, limitada a unos pocos átomos, o incluso a solamente un átomo. Ejemplos de dichas moléculas pequeñas incluyen cualquier tipo de hapteno (es decir, moléculas pequeñas que pueden provocar una respuesta inmune solamente cuando están unidas a un vehículo grande), vitaminas, o iones metálicos. La amplia diversidad de moléculas que puede usarse como dianas es especialmente interesante, ya que las proteínas OB-fold que se unen a estas dianas pueden usarse para las mismas aplicaciones que las ya conocidas para los anticuerpos, o para diferentes aplicaciones. Por ejemplo, las proteínas OB-fold obtenidas de acuerdo con la presente invención y que se unen a iones metálicos pueden usarse en procesos de biorremediación, por ejemplo para eliminar metales pesados de un material complejo tal como suelo o agua contaminada.

Los especialistas en la técnica pueden usar cualquier técnica de la técnica para realizar la etapa de selección del proceso de acuerdo con la invención y/o para realizar la etapa de selección para obtener una proteína con una propiedad deseada. Las técnicas de selección puede ser, por ejemplo, presentación en fagos (Smith, 1985), presentación de ARNm (Wilson et al., 2001), presentación bacteriana (Georgiou et al., 1997), presentación en levaduras (Boder y Wittrup, 1997) o presentación en ribosoma (Hanes y Pluckthun, 1997).

La presentación en ribosoma es particularmente ventajosa para realizar la etapa de selección, ya que se realiza completamente *in vitro* y por tanto esquivada muchas limitaciones de los sistemas *in vivo* (especialmente respecto al tamaño de la biblioteca) (He y Taussig, 2002; Schaffitzel et al., 1999). Los especialistas en la técnica encontrarán en la siguiente parte experimental, así como en los artículos de He and Taussig (2002), y Schaffitzel et al. (1999), o en otros artículos y manuales, protocolos para realizar presentación en ribosoma.

En una realización preferida del proceso de la invención, se realizan 2, 3, 4 ó 5 rondas de selección por presentación en ribosoma.

Como se describe en la parte experimental, el proceso de acuerdo con la invención también puede comprender una etapa adicional de aislamiento y caracterización de una o varias moléculas que se unen específicamente a la diana de interés. Por ejemplo, las etapas de aislamiento y caracterización pueden realizarse del siguiente modo:

(i) se transcriben de forma inversa ARNm de una combinación seleccionada por presentación en ribosoma y se amplifican por PCR, y los ADN resultantes se clonan en vectores de expresión;

(ii) se transforman bacterias (por ejemplo, DH5 $\alpha$  competentes) por dichos vectores de expresión (que comprenden dichos ADN) y se cultivan clones individuales; y

(iii) se ensayan proteínas extraídas de dichos clones bacterianos para su unión a la diana y/u otras propiedades biológicas (tales como estabilidad y similares).

Los clones que expresan las proteínas que tienen las mejores propiedades después pueden cultivarse para producir cantidades mayores de proteínas. La secuencia codificante de la proteína modificada por ingeniería puede determinarse secuenciando la parte pertinente del vector de expresión, y el gen que codifica esta proteína también puede usarse para modificar por ingeniería adicionalmente la proteína. De hecho, puede ser ventajoso usar un agente de unión obtenido a través de un proceso de acuerdo con la invención para construir proteínas multifuncionales que tengan al menos un resto dirigido y otro resto con actividad catalítica, fluorescente, enzimática o cualquier otro tipo de actividad. Esto puede hacerse, por ejemplo, a través de la construcción de una proteína de fusión que comprenda, como primer resto, la molécula aislada y caracterizada, y al menos un segundo resto. Este segundo resto puede seleccionarse ventajosamente entre enzimas, marcadores o proteínas indicadoras (tales como PhoA, GFP y similares), proteínas terapéuticas, etc. De acuerdo con una variante del proceso de la invención, se unen varios agentes de unión juntos (en una proteína de fusión, o a través de enlaces no covalentes), para construir fusiones o complejos con múltiples especificidades de unión.

Otros objetos de la invención son proteínas obtenidas a través de un proceso como el descrito anteriormente. En particular, se describen proteínas que se unen específicamente a PulD, y que tienen una secuencia seleccionada entre las SEC ID N° 2 a 8, así como el agente de unión a GarA de la SEC ID N° 47, que es un objeto de la invención.

La invención se ilustra adicionalmente mediante las siguientes figuras y ejemplos.

#### Leyendas de las figuras

Figura 1: a) representación esquemática de Sac7d de tipo silvestre en complejo con ADN bicatenario (código PDB 1azp). b) Restos que participan en la unión al ADN. El área cóncava se define con restos coloreados en gris claro, el área plana con aquellos coloreados en gris medio y oscuro. Los restos en el límite entre las regiones del área de unión están coloreados en gris medio.

**Figura 2:** Representación esquemática de los bucles 1, 3 y 4 de diversos dominios OB-fold. Los archivos de estructura PBD de RSCB usados para realizar esta superposición son los siguientes: \_1azp (es decir, Sac7d), 1tvxA, 1dokA, 1bqq, 1esp, 3seb, 1br9, 2bosa, 1bf4A y 1quqB. Por motivos de claridad de la figura, solamente se representa la estructura de Sac7d.

**Figura 3:** Representación esquemática y superposición con Sac7d de a) Sso7d (archivo de estructura: 1bf4); b) toxina Ile tipo Shiga (archivo de estructura: 2bosa); c) Péptido-2 de activación de neutrófilos humanos (NAP-2, archivo de estructura: 1tvxA); d) Proteína de unión a molibdeno (modg) de *Azotobacter vinelandii* (archivo de estructura: 1h9j). *Nota bene:* para proteínas que comprenden dominios adicionales diferentes del dominio OB-fold, se muestra solamente el dominio OB-fold en la figura. Las superposiciones de estructura se obtuvieron usando los servidores DALI y DALI Lite. La secuencia de proteínas, con su numeración en el archivo NCBI, se muestra en la siguiente Tabla 2.

Nombre	Archivo ref.	Fuente	Secuencia
Sac7d	1azp	NCBI: AAA80315	<i>MVKVKFKYKGEEKEVDTSKIKKVVWRVGMVS FTYDDNGKTGRGAVSEKDAPKELLMLLRAE REKK (SEC ID N° 1)</i>
Sso7d	1bf4	NCBI: P39476	<i>MATVKFKYKGEEKEVDISKIKKVVWRVGMKIS FTYDEGGGKTGRGAVSEKDAPKELLQMLEKQ KK (SEC ID N° 28)</i>
Toxina Ile	1r4pb	NCBI: CAA90631	<i>MKKMFMAVLFALVSVNAMAADCAKGIFFSK YNEDDTFTVKVDGKEYWTSRWNLQPLLQSAQ LTGMTVTIKSSTCESGSGFAEVQFNND (SEC ID N° 29)</i>
NAP-2	1tvxA	NCBI: CAG33086	<i>MSLRLDTPSCNSARPLHALQVLLLLLLLT ALASSTKGQTKRNLAKGKEESLDSLYAELR <b>CMCIKTTSGIHPKNIQSLEVIGKGTNCNQE VIATLKDGRKICLDPPAPRIKKIYQKLAGD ESAD (SEC ID N° 30)</b></i>
Modg	1h9j	NCBI: CAA90038	<i>MKISARNVFKGTVSALKEGAVNAEVDILLGG GDKLAAVVTLESARSLQAAGKEVVA <b>VVKAP WVLLMTDSSGYRLSARNILGTVKTITGAV NAEVTALQGGTEITSMVTKEAVAELGLKPG ASASAVIKASNVILG VPA (SEC ID N° 31)</b></i>

Tabla 2: Secuencias de proteínas que se han usado para obtener las superposiciones mostradas en la figura 3. Los fragmentos de secuencia en cursiva corresponden a las secuencias usadas de forma eficaz en el archivo de estructura usado para determinar las posiciones que podrían mutarse ventajosamente de forma aleatoria. Dichas posiciones de mutación están en negrita.

**Figura 4:** Alineación de los dominios OB-fold de varios homólogos de Sac7d.

**Figura 5:** Síntesis génica de las secuencias mutadas que codifican Sac7d. Las posiciones aleatorizadas para la biblioteca 11 están marcadas en negro. Las posiciones aleatorizadas adicionales para la biblioteca 14 están marcadas en gris. Los oligonucleótidos usados se representan con flechas delgadas (véase el texto para sus secuencias). Las flechas grandes corresponden a la secuencia codificante de Sac7d (sin el enlazador ToIA C-ter).

**Figura 6:** a) RIA (radioinmunoensayo) después de cuatro rondas de selección con la biblioteca 11 frente a DnaK, PulD-N y GarA. Las combinaciones de selecciones se tradujeron *in vitro* en presencia de metionina <sup>35</sup>S y se seleccionaron para la unión a DnaK, PulD-N o GarA inmovilizados en una placa ELISA. Después de lavado y elución con trietanolamina 0,1 M, se estimaron las cantidades de agentes de unión con un contador β. Se realizaron competiciones en paralelo con pre-incubación de combinaciones traducidas con proteínas libres a concentraciones de 1 μM y 10 μM. b) Selección de agentes de unión anti-PulD-N después de la ronda 4. c) Ensayo de las especificidades de unión para seis agentes de unión anti-PulD-N por ELISA. Se compararon las interacciones entre los agentes de unión y DnaK, PulD-N, GarA y BSA inmovilizados.

**Figura 7:** RIA después de cinco rondas de selección con las bibliotecas 11, 13 y 14 frente a PulD-N. Las combinaciones de selecciones se tradujeron *in vitro* en presencia de metionina <sup>35</sup>S y se ensayaron para la unión a DnaK, PulD-N, GarA o BSA inmovilizados en una placa ELISA. Después de lavado y elución con trietanolamina 0,1 M, se estimaron las cantidades de agentes de unión con un contador β. Se realizaron competiciones en paralelo con pre-incubación de combinaciones traducidas con PulD-N libre a 1 nM, 10 nM, 100 nM, 1000 nM. Se usó BSA como control negativo para la unión combinaciones (sin competición hecha).

**Figura 8:** Secuencias de ocho agentes de unión seleccionados frente a PulD-N. La secuencia de Sac7d de tipo silvestre se muestra en la parte superior de la figura. Los restos comunes a Sac7d de tipo silvestre y los agentes de

unión se representan por un punto. Las posiciones que se programaron en el esquema de sustitución están resaltadas en gris.

5 Figura 9: Análisis de expresión y purificación de ocho agentes de unión seleccionados. Las proteínas se cargaron en un SDS-PAGE al 15% teñido con azul brillante de Coomassie. a) Extracto crudo de *E. coli* después de expresión a 30°C. Las células se recogieron después de 19 horas de incubación con IPTG 0,5 mM y se lisaron en tampón de carga. b) Fracciones solubles de extractos crudos antes de purificación de los agentes de unión. c) Agentes de unión purificados después de purificación por IMAC (cromatografía de afinidad por iones metálicos inmovilizados) de una etapa de la fracción soluble del extracto crudo de *E. coli*. Todas las muestras cargadas fueron equivalentes a 13 µl de cultivo líquido.

15 Figura 10: Determinación de afinidad del clon 6 usando análisis SPR (resonancia de plasmón superficial). a) Se realizó la cinética de unión usando un BIAcore. Se inmovilizó PulD-N biotinilado (250 UR) en un chip de estreptavidina y se inyectó el clon 6 a 50 nM, 25 nM, 12,5 nM, 6,25 nM, 3,13 nM y 1,56 nM. Se usó una celda de flujo sin PulD-N inmovilizado para controlar que no sucedía unión específica y se sustrajo la señal de la celda a la obtenida de la celda de medición. Se usó el software BIAevaluation para analizar los datos con un procedimiento de ajuste global. b) También se midió la afinidad en equilibrio usando BIAcore de competición. Los parámetros determinados a partir de ambos enfoques se presentan en la Tabla 3.

20 Figura 11: Estabilidades térmicas de proteínas recombinantes Sac7d de tipo silvestre y variantes determinadas por calorimetría de exploración diferencial. La absorción de calor en exceso de wt Sac7d, clon 40, clon 33, y clon 6, se presentan a 1°C/min a una concentración de proteína de 200 µg/ml en tampón MES 50 mM pH 5,6 con NaCl 300 mM. Para cada muestra de proteína, la curva de capacidad calorífica en exceso experimental (línea gruesa) se ajustó mejor a un modelo no de dos estados (línea delgada) por regresión no lineal. Los parámetros termodinámicos resultantes se dan en la Tabla 4.

30 Figura 12: a) Transferencia de Western para la detección de PulD-N mezclado con extracto crudo de *E. coli* usando la fusión clon 6-fosfatasa alcalina. b) Purificación de una etapa de PulD-N con clon 6 inmovilizado en una columna. La fracción soluble de un extracto de *E. coli* que contenía PulD-N se inyectó en la columna equilibrada con tampón HBS pH 7,0. Después del lavado, el pH se disminuyó hasta 2,5 con un tampón glicina-HCl. Las fracciones se analizaron en un SDS-PAGE al 15% y se tiñeron con azul brillante de Coomassie. CE: extracto celular; PulD-N: proteína pura usada como marcador; Fracciones: fracciones eluidas de la columna después de cambio a pH ácido; FT: flujo continuo.

35 Figura 13: Unión de Sac7\*40, Sac7\*33, y Sac7\*6 a envueltas celulares aisladas y a dodecámeros de PulD (PulDD) y monómeros de PulD (PulDM). (A) Se incubaron cantidades crecientes de envuelta celular de la cepa PAP105 que producía PulD y PulS o PulD-CS y PulS con Sac7\*40-GFP, y después se analizó la fracción de membrana por SDS e inmunotransferencia con anticuerpos contra GFP. (B) Transferencia de Far-Western de las cantidades indicadas de proteínas de envuelta celular de las mismas cepas usando las tres quimeras Sac7\*-PhoA y anticuerpos contra PhoA.

40 Figura 14: Producción de quimeras Sac7-PhoA con y sin inducción por IPTG y sus efectos sobre la secreción y multimerización de PulD en cepa PAP5198 deficiente en proteasa de envuelta que porta pCHAP231. (A) Niveles de Sac7-PhoA detectados por inmunotransferencia (con anticuerpo contra PhoA) de la misma cantidad de extracto celular. (B) Niveles de secreción (%) y presencia de dodecámeros de PulD (PulDD) y monómeros de PulD (PulDM) detectados por inmunotransferencia de extractos celulares tratados con fenol y no tratados con anticuerpos contra PulD. Las flechas indican PulDM detectado sin tratamiento con fenol. S indica Sac7d-PhoA. (C) Como en B, pero sin inducción con IPTG.

50 Figura 15: PCR realizada sobre ADNc obtenido de ARNm, eluidos durante el ciclo nº 4 (véase la Figura 19, fase RT-PCR).

55 Figura 16: Radioinmunoensayos competitivos realizados después de cinco ciclos de selección anti-PKng y anti-lisozima.

Figura 17: Selección por ELISA de clones anti-lisozima (A), anti-PKng (B) y anti-GarA (C) obtenidos después de 5 ciclos de selección.

60 Figura 18: Secuencias del clon anti-lisozima obtenidas después de cinco ciclos de selección.

Figura 19: Alineación de un agente de unión a GarA con Sac7d.

Figura 20: ELISA de competición después de la 4ª ronda de selección frente al fragmento Fc.

65 Figura 21: Representación esquemática de una ronda de selección por presentación en ribosoma.



## Ejemplos

Los siguientes datos experimentales se han obtenido usando el material y los métodos descritos a continuación.

### 5 Materiales y métodos

#### Biología molecular general

10 Las enzimas y tampones fueron de New England Biolabs (EEUU) o Fermentas (Lituania). Los oligonucleótidos fueron de MWG Biotech (Alemania). Todas las PCR (reacciones en cadena de la polimerasa) se realizaron usando polimerasa Vent si no se indica en el texto. El vector de clonación y expresión para Sac7d de tipo silvestre y mutantes seleccionados fue pQE30 de Qiagen (Alemania).

#### Síntesis del ADN que codifica Sac7d de tipo silvestre

15 La secuencia de ADN de Sac7d de tipo silvestre se generó por PCR de ensamblaje usando los siguientes seis oligonucleótidos: SC1 (SEC ID N° 9: GAAACTCCTAGGTATTGTGCTGACGACCCCGATCGCGATCTCTAGCTTTGC GGTGAAAGTGAAATT), SC2 (SEC ID N° 10: GATCTTGCTGGTGTCCACTTCTTTTTCTTCGCGCAAAGCTAG), SC3 (SEC ID N° 11: GAAGTGGACACCAGCAAGATCAAGAAAGTTTGGCGTGTGGGCAAATGGTGAGCTTTACCT  
20 ACGACGACAACGGCAAG), SC4 (SEC ID N° 12: CTCTTTCCGGGCATCTTTCTCGTCCACGGCGCCACGGCCGGT CTTGCCGTTGTGTCGTA), SC5 (SEC ID N° 13: GAGAAAGATGCCCCGAAAGAGTTATTAGATATGTTAGCGCGT GCGGAAAGCTTCAACCA), SC6 (SEC ID N° 14: TGGTGGTTGAAGCTTTCCGCACG). El producto purificado de PCR sirvió como molde para una segunda amplificación por PCR usando los dos siguientes cebadores: SC07 (SEC ID N° 15: ATTAATGGTACCGGATCCGTGAAAGTAAATTTAAATATAAAG) y SC08 (SEC ID N° 16  
25 ATAATTGAGCTCTAAGCTTTTTTTCACGTTCCGCACGCGCTAACATATC). El producto de PCR se clonó en el vector de expresión pQe30 usando los sitios de restricción BamHI y HindIII. Se usó un clon con la secuencia esperada para posteriores expresiones.

#### Generación de bibliotecas combinatorias

30 El protocolo fue el mismo para la generación de las tres bibliotecas 11, 13, 14 en un formato compatible con la presentación en ribosoma. Las bibliotecas se construyeron principalmente por etapas de síntesis génica y ensamblaje por PCR. Con una PCR de una etapa usando una combinación de cuatro oligonucleótidos convencionales y tres degenerados que codifican tripletes NNS (donde N=A, C, T o G, y S=C o G), se obtuvo  
35 producto de ADN que incluía la región 5'-flanqueante necesaria para la presentación en ribosoma y el gen aleatorizado de Sac7d. Para la biblioteca 14, se usaron los siguientes oligonucleótidos: T7C (SEC ID N° 17: ATACGAAATTAATACGACTCACTATAGGGAGACCACAACGGTTTCCCTC), SDA\_MRGS (SEC ID N° 18: AGACCACAACGGTTTCCCTCTAGAAATAATTTTGTAACTTAAAGAAGGAGATATATCCATGAGAGGATCG), SClib1 (SEC ID N° 19: GGAGATATATCCATGAGAGGATCGCATCACCATCACCATCACGGATCCGTCAAGGTGAA  
40 ATTC), SClib2 (SEC ID N° 20: GGATCCGTCAAGGTGAAATTCNNSNNSNNSGGCGAAGAAAAAGAAGTGACT AGTAAGATC), SClib3 (SEC ID N° 21: CTTGCCGTTGTGTCGTASNNAASNNCACSNNTTTGCCSNNACGSNNAA CSNNSNNGATCTTACTAGTGTCCACTTC), SClib4 (SEC ID N° 22: TAATAACTCTTTCCGGGCATCTTTCTCSNNC ACSNNGCCSNNCCSNNCTTGCCGTTGTGTCGTA), SClib5 (SEC ID N° 23: CCATATAAAGCTTTTTCTCGCGTT CCGCACGCGCTAACATATCTAATAACTCTTTCCGGGCATC). Se usaron cebadores que codifican tripletes de tipo  
45 silvestre en lugar de tripletes NNS en las posiciones correspondientes a los restos 21, 22 y 40, o al resto 40 en Sac7d para la construcción de las bibliotecas 11 y 13, respectivamente.

Con el propósito de que la proteína presentada en el ribosoma sea accesible a ligandos potenciales, la proteína tiene que fusionarse a un enlazador. La secuencia del enlazador, correspondiente a una parte de la proteína de *E. coli* TolA, codificada en el vector pRDV plasmídico (Binz et al., 2004a), se amplificó por PCR usando los cebadores SClink (SEC ID N° 24: GCGGAACGCGAGAAAAAGCTTTATATGGCCTCGGGGGCC) y tolAk (SEC ID N° 25: CGGCACACCAGTAAGGTGTGCGGTTTTCAGTTGCCGCTTTCTTTCT) (codificando el último la región 3'-flanqueante  
50 necesaria para la presentación en ribosoma). Finalmente la biblioteca se ensambló con el enlazador tolA mediante ensamblaje por PCR usando los cebadores tolAk y T7B (SEC ID N° 25: ATACGAAATTAATACGACTCACTATAGGG AGACCACAACGG). El producto de ensamblaje final correspondía a una biblioteca de Sac7d con todas las regiones 5' y 3' necesarias para su uso para selecciones de presentación en ribosoma como se ha descrito previamente (Hanes et al., 1998; Schaffitzel et al., 1999).

#### Rondas de selección de presentación en ribosoma

60 Para experimentos de selección, se usaron proteínas diana biotiniladas. La biotinilación se realizó por incubación de una solución 10  $\mu$ M de DnaK, GarA o PulD-N con un exceso molar de 20 veces de sulfosuccinimidil-6-(biotinamido) hexanoato (Sulfo-NHS-LC-LC-Biotina, Pierce) en PBS en hielo durante 1 h. A las proteínas biotiniladas se les intercambió el tampón usando columnas de centrifugación de desalación de proteínas de Pierce equilibradas en  
65 TBS. El grado de biotinilación se determinó usando el ensayo HABA (Sigma) de 2 a 3 moléculas de biotina por molécula de proteína. Las proteínas diana biotiniladas se unieron a neutravidina inmovilizada en una placa Maxisorp

(Nunc) y se realizaron selecciones por presentación en ribosoma a 4°C esencialmente como se ha descrito (Binz et al., 2004b) con algunas modificaciones. En resumen, después de cada ronda de selección se transcribió de forma inversa el ARNm eluido a ADNc y se amplificó los cebadores SDA\_MRGS y SClib5. El producto de PCR se desaló en una columna NucleSpin ExtractII (Macherey-Nagel) y se usó para una ensamble por PCR con el fragmento de ADN enlazador ToIA (véase anteriormente). Esto generó una construcción de presentación en ribosoma de longitud completa con regiones 5'- y 3'-flanqueantes recién añadidas para minimizar la pérdida de clones debido a la degradación de los extremos de ARNm durante su manipulación. La cantidad de ciclos de RT-PCR fue 35 ciclos (ronda 1), 30 ciclos (ronda 2), 30 ciclos (ronda 3), 25 ciclos (ronda 4). En la ronda de selección se aumentó usando selección por disociación como se ha descrito (Jermutus et al., 2001). Para esta selección, se añadieron 10 nM de PulD-N biotinilado a la traducción detenida de la combinación de la ronda 4. La mezcla se equilibró 1 h a 4°C y se añadió PulD-N no biotinilado a una concentración final de 10 µM (exceso de 1000 veces sobre PulD-N biotinilado). Después de 2 h de incubación a 4°C con agitación, los complejos ternarios (ARNm-ribosoma-agente de unión) unidos a PulD-N biotinilado se capturaron con 30 µl de perlas recubiertas con estreptavidina magnéticas (Roche) 15 min a 4°C. Las perlas se lavaron y se aisló el ARNm como para las primeras cuatro rondas. La cantidad de ciclos de RT-PCR fue 35 para la 5ª ronda de selección. Los progresos de las selecciones se comprobaron controlando las cantidades de productos de RT-PCR de una ronda a otra.

#### Análisis de combinaciones seleccionadas y clones aislados

Después de cuatro o cinco rondas, las combinaciones seleccionadas se ensayaron por RIA (radioinmunoensayo) como se ha descrito (Hanes et al., 1998) usando directamente proteína diana recubierta en una placa Maxisorp y usando de 1 nM a 10 µM de proteína diana libre como competidor.

Los productos de RT-PCR de las combinaciones seleccionadas se clonaron en el vector pQE30 usando los sitios de restricción BamHI y HindIII y los ligamientos resultantes se usaron para transformar la cepa DH5α de *E. coli*. Los clones se picaron de la placa Petri para inocular una placa de pocillo profundo que contenía 1,5 ml de medio LB por pocillo (100 µg/ml de ampicilina, glucosa al 1%). Después de cultivo durante una noche a 37°C con agitación a 250 rpm, se usaron 0,2 ml de cada pocillo de esta placa maestra para inocular otra placa de pocillo profundo que contenía 1,3 ml de medio 2YT (100 µg/ml de ampicilina) por pocillo. La placa después se incubó a 37°C durante 1 h con agitación (250 rpm). La expresión se indujo con la adición de IPTG 0,5 mM e incubación a 30°C durante 4 h con agitación (250 rpm). Las células se sedimentaron con una etapa de centrifugación (2250 g) y se descartaron los sobrenadantes. Las proteínas se extrajeron con 50 µl de BugBuster (Novagen) por pocillo con agitación a 250 rpm durante 30 min, después se añadieron 250 µl de TBS pH 7,4 (Tris-HCl 20 mM, NaCl 150 mM). Los desechos celulares se sedimentaron con una etapa de centrifugación (2500 g). Para selección por ELISA (ensayo inmunoabsorbente ligado a enzimas), se usaron 100 µl de cada sobrenadante para ensayar la unión en proteínas diana recubiertas en una placa Maxisorp. La detección se realizó usando el conjugado de anticuerpo RGS His con HRP (Qiagen) que detecta solamente la marca RGS-(His)<sub>6</sub> de agentes de unión y no la marca (His)<sub>6</sub> de proteínas diana y sustrato BM-Blue de Roche. Todas las etapas de incubación se realizaron en TBS pH 7,4 con Tween 20 al 0,1%. Los clones positivos se secuenciaron por técnicas convencionales de secuenciación.

#### Producción de proteínas para selección por SPR (resonancia de plasmón superficial)

Para realizar la selección de agentes de unión en base a su tiempo de disociación de PulD-N, se expresaron clones en un cultivo a escala de 5 ml (cepa DH5α de *E. coli*). Se usaron quinientos microlitros de un precultivo durante una noche (medio LB, 100 µg/ml de ampicilina, glucosa al 1%, 37°C) en pocillo profundo para inocular 4,5 ml de cultivo (2YT, 100 µg/ml de ampicilina, 37°C) en pocillo profundo. La expresión se indujo a DO<sub>600</sub> = 1,0 mediante la adición de IPTG 0,5 mM y los cultivos se incubaron durante 19 h a 30°C (250 rpm). Las células se recogieron por centrifugación (2500 g) y las proteínas se extrajeron en pocillos profundos por resuspensión de células en 0,5 ml de TBS pH 7,4 que contenía imidazol 25 mM, BugBuster y Benzonase (Novagen). Después de 1 h de agitación a 4°C el pocillo profundo se centrifugó para sedimentar los desechos celulares. Los sobrenadantes se purificaron en microcolumnas de centrifugación que contenían 100 µl de Sepharose quelante Ni-Fast Flow (General Electric) equilibradas con TBS pH 7,4 que contenía imidazol 20 mM. La resina se lavó 4 veces con el tampón de carga y las proteínas purificadas se eluyeron con 400 µl de TBS pH 7,4 e imidazol 250 mM.

#### Producción de proteínas y purificación de agentes de unión y Sac7d de tipo silvestre

Para experimentos Biacore y de microcalorimetría, los agentes de unión se expresaron en la cepa DH5α de *E. coli* en una escala de 1 litro y se purificaron como se describe a continuación. Se usaron cincuenta mililitros de un precultivo durante una noche (medio LB, 100 µg/ml de ampicilina, glucosa al 1%, 37°C) para inocular 1 litro de cultivo (2YT, 100 µg/ml de ampicilina, 37°C). La expresión se indujo a DO<sub>600</sub>=1,0 mediante la adición de IPTG 0,5 mM y los cultivos se incubaron durante 19 h a 30°C (250 rpm). Las células se recogieron por centrifugación y se resuspendieron en 30 ml TBS de pH 7,4 a 30°C que contenía imidazol 25 mM. Las células se lisaron con un homogeneizador Avestin Emulsiflex y los desechos celulares se desecharon con una etapa de centrifugación. Las proteínas se purificaron en una columna quelante HiTrap de 5 ml equilibrada con TBS pH 7,4 que contenía imidazol 25 mM. Se realizó elución con TBS pH 7,4 e imidazol 250 mM. Las proteínas se purificaron adicionalmente por

5 cromatografía por exclusión de tamaño sobre una columna de filtración en gel Superdex 75 26/60 (GE Healthcare) equilibrada con HBS pH 7,0 (HEPES 20 mM, NaCl 150 mM). La filtración analítica en gel se hizo en un sistema SMART (GE Healthcare) usando una columna Superdex 75 3,2/20 equilibrada con HBS pH 7,0 (60  $\mu$ l/min, 50  $\mu$ l de muestra inyectados). Se usaron BPTI (6,5 kDa), ribonucleasa A (14,6 kDa), quimotripsinógeno A (20,3 kDa), ovalbúmina (46,7 kDa) y albúmina (62,9 kDa) como patrones de peso molecular de calibración de filtración en gel.

#### Resonancia de plasmón superficial

10 La SPR se midió usando un instrumento BIAcore 2000 a 25°C. Se preparó PulD biotinilado como para las selecciones por presentación en ribosoma y se inmovilizó en celdas de flujo de un SA-chip. Las densidades de PulD-N biotinilado inmovilizado para las mediciones cinéticas, y para las mediciones de inhibición fueron 200 UR y 800 UR (chip saturado), respectivamente. El tampón de procesamiento fue HBST pH 7,0 (HEPES 20 mM, NaCl 150 mM, Tween 20 al 0,05%).

15 La selección y la clasificación de los agentes de unión de acuerdo con su disociación se realizaron usando proteínas purificadas micro IMAC (véase anteriormente) diluidas 1/20 en tampón de procesamiento antes de las inyecciones para las mediciones cinéticas a un caudal de 60  $\mu$ l/min.

20 Las mediciones cinéticas para las determinaciones de afinidad se realizaron con proteínas purificadas por exclusión de tamaño inyectadas a una concentración que variaba de 1 nM a 50 nM. Las mediciones de inhibición se hicieron como se ha descrito (Ostermeier et al., 1995) a una concentración de agente de unión de 5 nM y diferentes concentraciones de PulD-N como competidor en una concentración que variaba de 3,2 nM a 20 nM a un caudal de 25  $\mu$ l/min. Se determinó la pendiente del sensograma para la fase de unión y se representó frente a la concentración de PulD-N. La evaluación de los datos se hizo usando software BIAeval (BIAcore).

#### Microcalorimetría

25 Se realizó calorimetría de exploración diferencial con un calorímetro MicroCal VP-DSC como se ha descrito previamente (Hible et al., 2005). Se dializaron soluciones madre de Sac7d recombinante de tipo silvestre y variantes durante una noche a 4°C frente a tampón MES 50 mM (pH 5,6 NaCl 300 mM). Las muestras de proteína después se diluyeron hasta 200  $\mu$ g/ml en la misma preparación de tampón, se desgasificaron al vacío durante 10 min con agitación suave antes de cargarlas en la celda del calorímetro (0,5 ml). La celda de referencia se llenó primero con el mismo tampón. Las muestras después se mantuvieron *in situ* a una presión externa constante de 172,32 kPa (25 psi) para evitar la formación de burbujas y evaporación hasta 130°C, se equilibró durante 25 min a 25°C, se calentó a una tasa de calentamiento constante de 1 grado/min, y los datos se recogieron con un filtro de 16 segundos. El análisis de los datos se hizo con el software Origin7™ (Plotnikov et al., 1997) proporcionado por el fabricante. La función de capacidad calorífica de exceso ( $C_p$ , exceso) se obtuvo después de sustracción de dos líneas basales a partir de la función de capacidad calorífica, el termograma de referencia de tampón y la línea basal química calculada después de normalizar a la concentración a partir del progreso de la transición de desplegamiento. Los parámetros termodinámicos de las transiciones de desplegamiento de cada muestra de proteína son los resultados de regresión no lineal de tres parámetros ( $T_m$ ,  $\Delta H_{cal}$ ,  $\Delta H_{vh}$ ) de la curva de capacidad calorífica de exceso asumiendo un modelo no de dos estados.

#### Construcción de fusiones con fosfatasa alcalina

30 Los genes que codifican agentes de unión se amplificaron por PCR con los cebadores SCPhoAF (SEC ID N° 26: ATTAATGGTACCGGATCCGTGAAGGTGAAATTC) y SCPhoAR (SEC ID N° 27: ATAATTGAGCTCTAAGCTTTTTTTT CACGCTCCGCAC) y la polimerasa Phusion para introducir KpnI y SacI en los extremos 5' y 3', respectivamente. Los productos de PCR se digirieron con KpnI y SacI y se clonaron en el vector pQUANTagen (Qbiogene). Por tanto, se fusionó fosfatasa alcalina al extremo C-terminal del agente de unión. La selección de clones activos de fosfatasa alcalina y extracciones periplásmicas se realizaron usando la cepa DH5 $\alpha$  de *E. coli* y las instrucciones del manual del fabricante. En resumen, se seleccionaron clones positivos en placas LB (100  $\mu$ g/ml de ampicilina, 4  $\mu$ g/ml de 5-bromo-4-cloro-3-indolil-fosfato) para su actividad fosfatasa alcalina. Se usaron ocho mililitros de un precultivo durante una noche (medio LB, 100  $\mu$ g/ml de ampicilina, 37°C) para inocular un cultivo de 400 ml (medio 2YT, 100  $\mu$ g/ml de ampicilina, 1 g de fosfato dipotásico pH 7,5, 37°C). La inducción del promotor tac se hizo con la adición de IPTG 0,5 mM cuando la DO<sub>600</sub> era aproximadamente 0,7 y el cultivo se continuó durante 4 h a 30°C. Las células se sedimentaron por centrifugación y se resuspendieron en 40 ml de tampón TSE (Tris-HCl 30 mM pH 8,0, sacarosa al 20%, EDTA 0,5 mM, 0,1 mg/ml de lisozima). El choque lisozímico se realizó por incubación de la suspensión durante 20 min a 4°C con agitación suave. Los desechos celulares se desecharon por centrifugación durante 30 min a 20000 g a 4°C y los sobrenadantes se filtraron a través de una membrana de 0,45  $\mu$ m antes de almacenamiento a -20°C.

#### Transferencias de Western usando fusiones con fosfatasa alcalina

65 Se centrifugó un cultivo de 10 ml durante una noche de la cepa DH5 $\alpha$  de *E. coli* (medio LB, 37°C) sin ningún vector de expresión para sedimentar las células. El sedimento celular se resuspendió en 1 ml de TBS (BugBuster 1X y 1  $\mu$ l

de Benzonase). La lisis celular sucedió a temperatura ambiente durante 30 min y la suspensión se centrifugó a 20000 g durante 5 min para retirar los desechos celulares. El sobrenadante después se usó para dilución en serie de PulD-N purificado. Las muestras se prepararon con un volumen constante de sobrenadante y una cantidad variable de PulD-N purificado correspondiente de 0,4 ng a 400 ng de PulD-N cuando se cargaron 5 µl de las muestras en un gel SDS. Después de la migración, las proteínas se transfirieron del gel SDS a una membrana de nitrocelulosa (Hybond-C extra, 0,45 µm, General Electric). La membrana se bloqueó con leche en polvo al 5% en TBST (Tris-HCl 20 mM, NaCl 150 mM, pH 7,5, Tween 20 al 0,1%). La membrana después se incubó con extracto periplásmico soluble de fusiones diluidas 1 a 15 en leche TBST durante 1 h con agitación suave. Después de lavar la membrana, se hizo la detección con sustrato de precipitación NBT/BCIP (kit de sustrato conjugado AP, Biorad) diluido en tampón de reacción (Tris-HCl 100 mM pH 9,5, NaCl 100 mM, MgCl<sub>2</sub> 5 mM).

#### Construcción y expresión de fusiones con GFP (proteína fluorescente verde)

Los genes que codifican agentes de unión se clonaron mediante sitios BamHI y HindIII en un plásmido derivado de pQE30 (pFP3000) que contenía el gen para eGFP y una región que codifica un enlazador peptídico flexible KLGSAAGSGEF (SEC ID N° 32). Esto produjo fusiones N-ter-agente de unión-enlazador-eGFP-C-ter con una marca MRGS (His)<sub>6</sub> en el extremo N-terminal. Las secuencias de los clones individuales se comprobaron por secuenciación de ADN.

La expresión y purificación de las fusiones agente de unión-GFP se hicieron en las mismas condiciones que para los agentes de unión en solitario (es decir, con una etapa IMAC y de cromatografía por exclusión de tamaño).

#### Aislamiento de fracciones de membrana y estudios de interacción con la quimera 40-eGFP

Se usaron las cepas de *E. coli* K-12 PAP 105 (Guilvout et al., 1999) que portan los plásmidos pCHAP3671 (pulD) y pCHAP580 (pulS) (Guilvout et al., 1999) o los plásmidos pCHAP371 (pulD-CS) (Guilvout et al., 2006) y pCHAP580 para el aislamiento de las vesículas de membrana. Los cultivos se hicieron crecer hasta fase exponencial (DO<sub>600</sub>: 0,9-1,1) en medio LB (Miller, 1992) que contenía antibióticos apropiados (100 µg/ml de ampicilina, 25 µg/ml de cloranfenicol) a 30°C con aireación vigorosa. Las fracciones de membrana se aislaron por centrifugación (180.000 x g durante 30 min) después de desintegración en prensa French de las células y se redisolviéron en Tris 50 mM pH 7,5, NaCl 150 mM a una concentración final de 450 µg/ml.

Se incubaron diferentes cantidades de las vesículas de membrana durante 1 hora a temperatura ambiente con 70 pmol de la proteína de fusión 40-eGFP purificada. Después de centrifugación (80.000 x g durante 20 min) los sedimentos se lavaron con el mismo tampón, se centrifugaron de nuevo y se resuspendieron en el mismo volumen. Las proteínas en cada muestra se separaron por SDS-PAGE y se transfirieron a nitrocelulosa. La detección de la proteína de fusión 40-eGFP se realizó por inmunotransferencia con anticuerpos primarios contra GFP.

#### Cromatografía de afinidad usando el agente de unión clon 6 inmovilizado

Se dializaron diecisiete miligramos de clon 6 purificado por IMAC frente a tampón carbonato 0,2 M pH 8,3 (NaCl 0,5 M) y se inyectaron en una columna HP activada por NHS de 1 ml HiTrap (General Electric) previamente enjuagada con 6 ml de HCl 1 mM. La inmovilización sucedió a temperatura ambiente durante 30 minutos. Después de lavar la columna y desactivar cualquier grupo activo restante con etanolamina 0,5 M (NaCl 0,5 M, pH 8,3) y acetato 0,1 M (NaCl 0,5 M, pH 8,3) durante 30 min a temperatura ambiente, la columna se equilibró con TBS pH 8,0 (Tris 20 mM, NaCl 500 mM) y estuvo lista para su uso para purificación.

La proteína PulD-N se expresó usando la cepa BL21(DE3) de *E. coli* transformada con el plásmido pCHAP3702 (Chami et al., 2005). Se usaron cincuenta mililitros de un precultivo durante una noche (medio LB, 100 µg/ml de ampicilina, glucosa al 1%, 37°C) para inocular 1 l de cultivo (2YT, 100 µg/ml de ampicilina, 37°C). La expresión se indujo a DO<sub>600</sub> = 1,0 mediante la adición de IPTG 1,0 mM y los cultivos se incubaron durante 4 h a 30°C (250 rpm). Las células se recogieron por centrifugación y se resuspendieron en 30 ml de TBS pH 8,0. Las células se lisaron con un homogeneizador Avestin Emulsiflex y los desechos celulares se desecharon con una etapa de centrifugación. Se inyectó un mililitro y medio de esta fracción soluble en la columna de NHS-clon 6 inmovilizado a un caudal de 0,5 ml/min. Las proteínas no específicas se retiraron por lavado con 40 ml de tampón de procesamiento y se eluyó PulD-N con un cambio de pH ácido usando un tampón glicina 100 mM (pH 2,5, NaCl 250 mM). La pureza de la proteína eluida se comprobó cargando las fracciones en un gel SDS.

#### Transferencia de Far Western

Se prepararon como anteriormente membranas externas de *E. coli* PAP 105+/- PulD (pCHAP3671; (Guilvout et al., 1999)) o PulD-CS (pCHAP3711; (Guilvout et al., 2006)) junto con PulS (pCHAP580; (Daefler et al., 1997)). Las membranas se resuspendieron y almacenaron en Tris-HCl 50 mM, pH 7,5, que contenía sacarosa al 10% y 0,1 mg/ml del inhibidor de proteasa Pefabloc (Interchim, Montluçon, Francia) y se sometieron a SDS/PAGE y se transfirieron a láminas de nitrocelulosa que después se bloquearon e incubaron con extractos periplásmicos (choque

osmótico) de cepas que producían quimeras Sac7-PhoA. Después de lavar, se detectó PhoA unido por anticuerpos contra PhoA, anticuerpos secundarios acoplados a peroxidasa de rábano rusticano, y quimioluminiscencia.

### Secreción

La cepa PAP7232 (Hardie et al., 1996) se transformó con el vector vacío o con plásmidos que codificaban quimeras Sac7-PhoA. Los transformantes se cultivaron en medio que contenía maltosa al 0,4% (para inducir producción de pululanasa y sus sistema de secreción, incluyendo PulD) e IPTG 1 mM para inducir producción de Sac7-PhoA. Los niveles de secreción se midieron como se describe en (d'Enfert et al., 1989) y se expresan como la cantidad de actividad enzimática detectada en células completas en comparación con la detectada en células lisadas (100%). Los extractos celulares también se examinaron por inmunotransferencia con anticuerpos contra PulD y PhoA.

Para analizar los efectos de las quimeras Sac7-PhoA a niveles superiores (codificados por plásmido) de producción de PulD, se amplificó un gen de resistencia a zeocina con sitios PstI flanqueantes y se insertaron en el sitio PstI único en el gen *blaM* de los plásmidos correspondientes. Los plásmidos recombinantes después se transformaron junto con pCHAP231 (d'Enfert et al., 1987) en la cepa deficiente en proteasa de envuelta PAP5198 (*degP*, *ompT*, *ptr*). Se analizaron la secreción de pululanasa y los niveles de PulD (con o sin tratamiento previo con fenol) como anteriormente con o sin inducción por IPTG.

### Resultados

Ejemplo 1: Diseño de la biblioteca de primera generación (biblioteca 11)

La primera etapa crítica para investigar la posibilidad de usar Sac7d como estructura para obtener agentes de unión para diversos ligados fue diseñar una biblioteca por aleatorización del área de unión potencial, manteniendo al mismo tiempo la estabilidad y la solubilidad de la proteína Sac7d precursora.

Se conocen varias estructuras tridimensionales de complejos Sac7d-ADN (Agback et al., 1998; Edmondson y Shriver, 2001; Gao et al., 1998; McAfee et al., 1995; McCrary et al., 1996; Robinson et al., 1998; Su et al., 2000) (figura 1a). De acuerdo con estas estructuras, la unión de Sac7d al surco menor del ADN implica un área superficial de unión que concuerda de forma remarcable (formas y cargas) con el ADN. Esta área de unión está compuesta por dieciséis restos (K7, Y8, K9, K21, K22, W24, V26, K28, M29, S31, T33, K39, T40, R42, A44, S46) (figura 1b). Una inspección visual de esta área de unión muestra que está girada y que comprende dos tipos de geometrías: una ligeramente cóncava (K7, Y8, K9, W24, V26, K28, M29, S31, T33, R42, A44, S46) y la otra esencialmente plana (K21, K22, W24, T33, K39, T40, R42). Los restos W24, T33, y R42 están compartidos por estas dos superficies. Aproximadamente un cuarto de la secuencia de Sac7d está dedicado a la unión del ADN. Aunque todos estos restos están expuestos en la superficie, una mutagénesis aleatoria masiva del 25% de la secuencia de Sac7d podría ser drástica para el plegamiento y la estabilidad de los mutantes obtenidos. En una primera etapa, los inventores decidieron excluir K28 y K39 del esquema de mutagénesis, ya que estos dos restos no están completamente orientados hacia la superficie de unión. La segunda razón para excluirlos fue que la generación de bibliotecas por una PCR de una etapa como se describe a continuación habría sido imposible debido a solapamientos reducidos disponibles para los cebadores que hibridan (véase a continuación). También se postuló que la geometría del lado cóncavo del área de unión podría concordar bien con la forma esférica de proteínas globulares o, al menos, que podría acomodar la unión de sus bucles expuestos. Por tanto, la estrategia de sustitución se centró en los once restos de esta región de Sac7d (K7, Y8, K9, W24, V26, M29, S31, T33, R42, A44, S46).

El gen que codifica Sac7d es bastante corto (aproximadamente 200 pares de bases). Por tanto, la biblioteca correspondiente de sustituciones aleatorias de los once restos podría obtenerse por una PCR de una etapa. Esto se hizo usando una mezcla de tres oligonucleótidos degenerados (esquema NNS) y tres oligonucleótidos convencionales. Los codones de las posiciones aleatorizadas estaban codificados por tripletes NNS que permiten la representación de todos los aminoácidos. Con esta estrategia de mutagénesis, la diversidad teórica es de aproximadamente  $3,2 \cdot 10^{16}$  ( $32^{11}$ ), que excede aproximadamente cuatro órdenes de magnitud la diversidad conseguida de forma experimental de nuestra biblioteca. De hecho, de acuerdo con la cantidad de producto de ensamblaje de PCR usada para generar la construcción de presentación en ribosoma, la estimación superior de la biblioteca fue de aproximadamente  $3,0 \cdot 10^{12}$  variantes. La secuenciación de cuarenta clones aleatorios confirmó que la frecuencia observada de restos fue similar a la predicha (datos no mostrados). Se descubrió que el porcentaje de clones correctos, sin ningún desplazamiento de fase o delección, era de aproximadamente el 50%. Por tanto, la diversidad "funcional" se consideró satisfactoria y la biblioteca se usó para las selecciones.

Ejemplo 2: Selecciones de presentación en ribosoma usando la biblioteca 11

Se eligieron tres proteínas como dianas: DnaK, GarA (tanto de *Mycobacterium tuberculosis* como el dominio N-terminal de PulD (de *Klebsiella oxytoca*)). PulD es una proteína de membrana externa que no puede mantenerse en solución sin detergentes iónicos. Por lo tanto, se usó un fragmento monomérico soluble. Este fragmento, llamado PulD-N, corresponde a la región N-terminal de PulD. Estas dianas se eligieron porque agentes de unión específicos y ávidos podrían ser herramientas útiles para estudiar DnaK, GarA y PulD. Además, estas proteínas son difíciles de

cristalizar para estudios estructurales. Los agentes de unión que reconocen estas proteínas podrían usarse para ensayos de co-cristalización del mismo modo que fragmentos de anticuerpo (Ostermeier et al., 1995).

5 Se usaron placas ELISA recubiertas con neutravidina para inmovilizar proteínas biotiniladas y para realizar selecciones. Después de cuatro rondas de selección, se observó enriquecimiento con agentes de unión específicos para las tres dianas (figura 6a). En todos los casos, más del 50% de la unión de combinaciones se inhibió con 10  $\mu\text{M}$  de dianas libres, de acuerdo con RIA.

10 Para evaluar combinaciones de agentes de unión adicionalmente por ELISA para la selección con PulD-N, se clonaron los productos de RT-PCR de combinaciones seleccionadas en el vector pQE30 (Qiagen) usando sitios de restricción BamHI y HindIII. Los ligamientos resultantes se usaron para transformar la cepa DH5 $\alpha$  de *E. coli*. Se picaron 96 clones aleatoriamente de la placa Petri para inocular una placa de pocillo profundo que contenía 1,5 ml de medio LB por pocillo (100  $\mu\text{g/ml}$  de ampicilina, glucosa al 1%). Después de cultivo durante una noche a 37°C con agitación a 250 rpm, se usaron 0,2 ml de cada pocillo de esta placa maestra para inocular otra placa de pocillo profundo que contenía 1,3 ml de medio 2YT (100  $\mu\text{g/ml}$  de ampicilina) por pocillo. La placa después se incubó a 37°C durante 1 h con agitación (250 rpm). La expresión se indujo con la adición de IPTG 0,5 mM e incubación a 30°C durante 4 h con agitación (250 rpm). Las células se sedimentaron con una etapa de centrifugación (2250 g) y los sobrenadantes se desecharon. Las proteínas se extrajeron con 50  $\mu\text{l}$  de BugBuster (Novagen) por pocillo con agitación a 250 rpm durante 30 min, después se añadieron 250  $\mu\text{l}$  de TBS pH 7,4 (Tris-HCl 20 mM, NaCl 150 mM). Los desechos celulares se sedimentaron con una etapa de centrifugación (2500 g). Para selección ELISA, se usaron 100  $\mu\text{l}$  de cada sobrenadante para ensayar la unión en PulD-N o BSA recubierto en una placa Maxisorp. La detección se realizó usando un conjugado de anticuerpo RGS His con HRP (Qiagen) que detecta solamente la marca RGS-(His)6 de agentes de unión y no la marca (His)6 de proteínas diana y sustrato BM-Blue de Roche. Todas las etapas de incubación se realizaron en TBS pH 7,4 con Tween 20 al 0,1%. Para todos los clones se midió la DO a 340 nm para la unión de PulD-N y BSA. Se calculó la proporción del valor obtenido para la unión de BSA sobre el valor para PulD-N para cada clon. Se observó una proporción de señal a ruido superior a 10 para aproximadamente el 92% de los clones desde la ronda 4 (figura 6b). Se ensayó la unión de 6 clones purificados de la selección PulD-N (ronda 4) para DnaK, GarA, PulD-N y BSA por ELISA. Como se muestra en la figura 6c, la unión fue específica para PulD-N. La secuenciación de 28 clones de la ronda 4 reveló una elevada diversidad de secuencias; de hecho ninguna secuencia se encontró más de una vez, lo que indica que podría haberse usado una presión de selección mayor. Además, no pudo encontrarse ningún resto de tipo silvestre en ningunas de las once posiciones aleatorizadas (datos no mostrados). El rendimiento promedio de la expresión estimada después de purificación por IMAC (cromatografía de afinidad por iones metálicos inmovilizados) de nueve clones fue de aproximadamente 70 mg/litro de cultivo en matraz (4 h de inducción a 30°C). Los agentes de unión podían concentrarse al menos hasta 45 mg/ml. En conjunto, estas observaciones sugirieron que el área de unión de Sac7d es muy tolerante a sustitución y que las variantes seleccionadas retenían propiedades biofísicas satisfactorias. Sin embargo, un análisis más detallado por RIA mostró que la afinidad promedio para agentes de unión en la combinación de la ronda 4 era de aproximadamente 100 nM y una quinta ronda realizada con una selección de disociación más suave (2 h) o más rigurosa (17 h) no logró aumentar la afinidad promedio de la combinación anti PulD-N (figura 7). Dicha baja afinidad podría ser limitante para aplicaciones donde se requieren bajas constantes de disociación nanomolares. Por tanto, los inventores decidieron explorar otros enfoques para obtener mayor afinidad.

#### Ejemplo 3: Diseño de las bibliotecas de segunda generación (biblioteca 13 y 14)

45 Un modo intuitivo de mejorar la afinidad es ampliar el área de unión y por tanto aumentar la cantidad potencial de interacciones entre los agentes de unión y sus ligandos. El área de unión potencial correspondiente a la aleatorización de 11 restos tenía una superficie accesible a disolvente de 890  $\text{\AA}^2$ . Siendo Sac7d muy tolerante a 11 sustituciones, se decidió aleatorizar dos o tres posiciones adicionales: K21 y K22 (biblioteca 13) o K21, K22 y T40 (biblioteca 14) que corresponden a 1130  $\text{\AA}^2$  y 1200  $\text{\AA}^2$ , respectivamente. Otra explicación podría ser que el área de unión cóncava no es muy adecuada para unirse a proteína y, por lo tanto, el uso de una superficie más plana podría mejorar las afinidades finales. Estas dos nuevas bibliotecas (cada una de 3,0  $10^{12}$  variantes) se construyeron del mismo modo que la biblioteca once. La secuenciación de setenta clones aleatorios de la biblioteca 14 confirmó que la frecuencia observada de restos era similar a la predicha, con una ligera sub-representación de restos aminoacídicos S, L, y R, mientras que los restos aminoacídicos P, N, Q, y H estaban un poco sobre-representados (datos no mostrados). Se descubrió que el porcentaje de clones correctos, sin ningún desplazamiento de fase o deleción, era de aproximadamente el 65%. Por tanto, la diversidad "funcional" se consideró satisfactoria y la biblioteca se usó para selecciones.

#### Ejemplo 4: Selecciones de presentación en ribosoma con PulD-N como proteína diana, usando la biblioteca 13 y 14

60 Se realizaron cuatro rondas de selecciones usando estas bibliotecas y PulD-N como proteína diana. Se realizó una quinta ronda en paralelo con las bibliotecas 11, 13, 14 con una selección de disociación suave durante dos horas para enriquecer las combinaciones para agentes de unión con disociaciones más lentas. Un radioinmunoensayo para las combinaciones después de la quinta ronda (figura 7) indicó un enriquecimiento para agentes de unión a PulD-N específicos con las dos nuevas bibliotecas. La unión no pudo detectarse en BSA o neutravidina inmovilizada.

Además, este RIA mostró que 10 nM de PulD-N libre fue suficiente para inhibir aproximadamente el 50% de la señal de unión con la biblioteca 13 en comparación con 100 nM de PulD-N para el mismo nivel de inhibición con la biblioteca 11. La biblioteca 14 se comportó incluso mejor, ya que 1 nM de competidor fue suficiente para conseguir inhibición significativa de aproximadamente el 20% de la señal de unión. Por tanto, se adquirió al menos un orden de magnitud para la afinidad promedio con el diseño usado para la biblioteca 14.

Ejemplo 5: Caracterización de agentes de unión anti-PulD-N de la biblioteca 14

### 5.1. Análisis de secuencia, expresión y purificación de agentes de unión seleccionados

Como las afinidades más prometedoras pudieron encontrarse en la combinación obtenida de la biblioteca 14 (ronda 5), se analizó la combinación de agentes de unión de esta biblioteca. La combinación enriquecida se clonó en el plásmido de expresión pQE30, y se usó la cepa DH5 $\alpha$  de *E. coli* para la producción de proteína. Se ensayaron cuarenta y ocho clones individuales por un procedimiento ELISA usando PulD-N o BSA inmovilizado y extractos crudos de *E. coli*. Para todos los clones, se detectó una unión significativa y específica sobre el fondo. Por tanto, todos los agentes de unión se secuenciaron para análisis adicional.

Como se observa para la biblioteca 11, permaneció una gran diversidad de agentes de unión incluso después de la quinta ronda de selección (figura 8). Sin embargo, la mayoría de los clones pudo clasificarse en seis familias de acuerdo con sus homologías (datos no mostrados). Además, se encontraron dos clones idénticos en dos casos, siendo todos los demás clones únicos, los que sugiere una convergencia de selección. Los residuos encontrados en posiciones aleatorizadas fueron de naturaleza muy diferente, con cadenas laterales alifáticas y aromáticas así como cadenas laterales cargadas o hidrófilas. Pudo observarse una clara preferencia para un resto particular en algunas posiciones. Por ejemplo, la posición R42 estaba ocupada por una tirosina en aproximadamente la mitad de los clones, reflejando probablemente su importancia para la unión de PulD-N. Entre todas las mutaciones programadas, no pudo observarse conservación estricta del resto nativo, incluyendo las tres nuevas posiciones abordadas (K21, K22 y T40) en comparación con la biblioteca 11. Finalmente, se retuvieron solamente siete restos nativos de los 560 restos en los cuarenta clones secuenciados (40 x 14). Por tanto, estas observaciones sugieren que las 14 posiciones abordadas toleran bien sustituciones aleatorias.

Los agentes de unión se acumularon en grandes cantidades en el citoplasma de *E. coli* a 30°C después de cultivo durante una noche y pudieron purificarse hasta homogeneidad en una IMAC de una etapa con rendimiento de hasta 200 mg a partir de un cultivo en matraz de agitación de un litro (figura 9). Las proteínas corrieron en un gel SDS-PAGE de acrilamida al 15% en la posición esperada para sus masas moleculares calculadas. Los agentes de unión purificados pudieron concentrarse hasta 60 mg/ml en un tampón TBS convencional con ausencia de signos de precipitación y permanecieron solubles durante varios meses a 4°C.

La cromatografía por exclusión de tamaño mostró que los agentes de unión clones 6, 39, 40 y 41 estudiados eran monoméricos. Como Sac7d, todos los agentes de unión parecían más grandes de lo predicho en este ensayo (11,5 kDa determinados en lugar de 9,1 kDa calculados), probablemente debido a una desviación desde la forma verdaderamente globular causada por la presencia de una marca de nueve restos N-terminal.

### 5.2. Afinidades de agentes de unión a PulD-N

Para identificar clones con las mayores afinidades, se usaron cuarenta y ocho clones para micro expresiones y purificaciones por IMAC. Estas proteínas purificadas después se seleccionaron por SPR usando PulD-N biotinilado inmovilizado en un chip recubierto con estreptavidina. Para todos los agentes de unión ensayados, no se observó unión significativa en la superficie blanco recubierta solamente con estreptavidina, lo que apoya la idea de que la unión era específica para PulD-N. De acuerdo con el análisis de las fases de disociación, se eligieron cinco clones con las disociaciones más lentas para determinación detallada de la afinidad de las proteínas monovalentes por SPR.

Estos cinco clones, 6, 33, 39, 40 y 41, se purificaron adicionalmente por filtración en gel. Se realizaron análisis cinéticos a diferentes concentraciones y se analizaron con un ajuste cinético global. Se descubrió que todas las constantes de disociación de todos los agentes de unión a PulD-N estaban en un intervalo de picomolar o bajo nanomolar (figura 10a y tabla 3). Los clones 6 y 33 tuvieron las afinidades más elevadas ( $K_D$ = 130 pM y 190 pM, respectivamente). Las cinéticas de asociación estaban ligeramente reducidas (factor 1,3) en condiciones de elevada fuerza iónica (NaCl 300 mM), lo que indica que estas asociaciones no estaban asistidas electrostáticamente (datos no mostrados). Los valores de  $K_D$  de estos cinco agentes de unión se confirmaron (figura 10b y tabla 3) por análisis SPR de competición (Nieba et al., 1996). Un experimento de control mostró que no sucedía unión cuando se ensayaba Sac7d de tipo silvestre para la unión a PulD-N inmovilizado, lo que indica que la propiedad de unión observada para los clones seleccionados era el resultado de una función recién introducida y no el resultado de una afinidad pre-existente de Sac7d por PulD-N.

Clon	$k_{on}$ [ $M^{-1}s^{-1}$ ]	$K_{off}$ [ $s^{-1}$ ]	$K_D$ [nM] <sup>a</sup>	$K_D$ [nM] <sup>b</sup>	% de actividad <sup>c</sup>
6	$1,9 \cdot 10^7$	$2,5 \cdot 10^{-3}$	0,13	$0,14 \pm 0,03$	90-100
33	$1,7 \cdot 10^7$	$3,2 \cdot 10^{-3}$	0,19	$0,17 \pm 0,03$	90-100
39	$1,7 \cdot 10^7$	$8,6 \cdot 10^{-3}$	0,52	$1,1 \pm 0,2$	90-100
40	$2,3 \cdot 10^7$	$1,4 \cdot 10^{-2}$	0,60	$0,9 \pm 0,2$	90-100
41	$2,0 \cdot 10^7$	$1,2 \cdot 10^{-2}$	0,61	$1,1 \pm 0,1$	90-100

Tabla 3: Resumen de determinaciones de constante de disociación por resonancia de plasmón superficial.

<sup>a</sup>: Valores de  $K_D$  obtenidos del análisis cinético. <sup>b</sup>: Valores de  $K_D$  obtenidos de experimentos SPR de competición. <sup>c</sup>: de análisis experimental SPR de competición.

El análisis de secuencias de estos agentes de unión reveló diferentes situaciones (figura 8). Las posiciones aleatorizadas en el clon 6 y 33 eran diferentes, a pesar del hecho de que sus respectivos valores de  $K_D$  estaban en el mismo intervalo (130-140 pM). En contraste, el clon 40 tenía 11/14 restos aleatorizados en común con aquellos del clon 41 y tenía valores de  $K_D$  similares ( $\approx 1$  nM). El clon 39 compartía 10/14 restos aleatorizados con el clon 6 pero tenía una afinidad 7 veces menor. De forma interesante, se descubrió que el clon 39 estaba truncado por 4 restos en su extremo C-terminal pero permanecía capaz de unirse a PulD-N. Esto no es sorprendente, ya que también existen varias formas truncadas de Sac7d (McAfee et al., 1995).

### 5.3. Estabilidad de agentes de unión a PulD-N

Sac7d es una proteína hipertermoestable que se despliega con una  $T_m$  de 91°C a pH 7,0, disminuyendo la estabilidad térmica de la proteína a pH inferiores (McCrory et al., 1996). Las estabilidades térmicas de los clones 6, 33 y 40 se compararon con la de Sac7d de tipo silvestre por calorimetría de exploración diferencial. Para asegurar el desplegamiento completo de las proteínas por debajo del alto límite de temperatura de DSC (125°C), se usó a pH 5,6 para todas las exploraciones. Se descubrió que los clones eran termoestables con temperaturas de desnaturalización entre 68°C y 83°C (figura 11 y tabla 4). El clon 40 era remarcablemente estable con un valor de  $T_m$  inferior al de tipo silvestre (89,7°C) en solamente 5,8°C, seguido del clon 33 ( $\Delta T_m = -10,5^\circ\text{C}$ ), siendo el clon menos estable el clon 6 ( $\Delta T_m = -22,0^\circ\text{C}$ ). Sin embargo, la  $T_m$  del clon 6 aún sigue estando 8°C por encima del valor medio de  $T_m$  de proteínas del Banco de Datos de Proteínas (Freire, 2001). Las exploraciones DSC fueron características de desplegamiento cooperativo lo que indica que, aunque con una elevada carga de mutación (hasta 14 restos mutados, es decir 21% de Sac7d mutado), las variantes de Sac7d eran capaces de adoptar una estructura plagada como para el tipo silvestre. La entalpía calorimétrica (el campo de calor por mol,  $\Delta H_{cal}$ ) fue inferior con respecto al tipo silvestre (42,5 kcal.mol<sup>-1</sup>) en solamente 5,8 y 4,8 kcal.mol<sup>-1</sup> para el clon 40 y el clon 53, respectivamente (tabla 4). Además, tanto para el clon 40 como el clon 33,  $\beta$ , el valor de proporción de la entalpía de van't Hoff (el cambio de calor por unidad de desplegamiento cooperativo) a la entalpía calorimétrica era cercano a uno, y cercano al valor  $\beta$  de Sac7d de tipo silvestre. Esta observación indicó firmemente que el desplegamiento del clon 40 y el clon 33 correspondían al del monómero proteico plegado (tabla 4). El clon 6 presentó una entalpía calorimétrica inferior con respecto al tipo silvestre (en 18,10 kcal.mol<sup>-1</sup>) y una mayor proporción  $\beta$  (tabla 4). Estas observaciones pueden estar relacionadas con el desplegamiento parcial de la proteína clon 6 a pH ácidos donde la proteína es menos estable ( $T_m$  del clon 6 es 8°C mayor a pH 7; datos no mostrados). Tomadas en conjunto, estas observaciones muestran que los clones 6, 33 y 40 retenían en gran medida la estabilidad térmica favorable de la proteína de tipo silvestre.

Clon	$T_m$ [°C] <sup>a</sup>	$\Delta H_{cal}$ [kcal.mol <sup>-1</sup> ] <sup>b</sup>	$\Delta H_{vH}$ [kcal.mol <sup>-1</sup> ] <sup>b</sup>	$\beta^c$
6	67,7	24,4	46,3	1,9
33	79,2	37,7	45,2	1,2
40	83,9	36,7	47,7	1,3
Sac7d wt	89,7	42,5	55,2	1,3

Tabla 4: Parámetros de desplegamiento térmico para Sac7d wt y variantes

<sup>a</sup>:  $T_m$  es  $\pm 0,1^\circ\text{C}$ ; <sup>b</sup>:  $\Delta H_{cal}$ , es  $\pm 0,3$  kcal.mol<sup>-1</sup>,  $\Delta H_{vH}$  es  $\pm 0,5$  kcal.mol<sup>-1</sup>; <sup>c</sup>:  $\beta = \Delta H_{vH}/\Delta H_{cal}$

### 5.4. Especificidad de agentes de unión a PulD-N

¿Cómo de específicos son estos agentes de unión a PulD-N, y cómo pueden usarse para aplicaciones biotecnológicas en las que es crucial la especificidad?

Para responder a estas cuestiones, se fusionaron los agentes de unión 6, 33 y 40 al extremo N-terminal de fosfatasa alcalina de *E. coli* (PhoA). Se produjeron las quimeras (Sac7\*-PhoA) como proteínas periplásmicas en la cepa DH5 $\alpha$  usando el vector pQUANTagen que codifica el péptido señal PhoA. Después, las quimeras se extrajeron del periplasma por choque osmótico y se ensayó su funcionalidad por ELISA con pocillos recubiertos con PulD-N o BSA. Las fusiones unidades se cuantificaron con sustrato p-nitrofenilfosfato cromogénico y espectrofotometría (datos no mostrados). Se observó una alta señal sobre el fondo para las tres quimeras (proporción señal/fondo > 10). Esto demostró que las quimeras 1) se exportaban al periplasma, 2) podían seguir reconociendo específicamente PulD-N y 3) eran catalíticamente activas. Por tanto, estas fusiones podían usarse como reactivo de detección en ELISA de una etapa.



Para evaluar si estas fusiones son capaces de discriminar la proteína PulD-N en una mezcla completa de proteínas, tal como en un extracto crudo de *E. coli*, se usaron como reactivo de detección para inmunotransferencias. En este experimento, se recogió y lisó un cultivo durante una noche culture de DH5 $\alpha$  libre de plásmido. El extracto crudo se dividió en alícuotas y se añadieron cantidades decrecientes de PulD-N purificado a cada tubo. Después de SDS-PAGE y transferencia a una membrana de nitrocelulosa, se detectó PulD-N con las quimeras agente de unión-PhoA en presencia del sustrato cromogénico de precipitación (NBT/BCIP) en cantidades tan bajas como 1,5-3 ng (figura 12a) sin reactividad cruzada. Esto apoyó la doble funcionalidad de las fusiones observada con ELISA (figura 12a). La detección de bajas cantidades de PulD-N se hizo mientras estaban presentes cantidades mucho mayores de proteínas endógenas, lo que demuestra que los agentes de unión 6, 33, y 40 eran capaces de discriminar PulD-N entre miles de proteínas y por lo tanto eran altamente específicos.

Para evaluar la especificidad adicionalmente, el agente de unión 6 se inmovilizó covalentemente mediante una reacción de acoplamiento de amina en una columna activada de NHS-agarosa de un mililitro. La fracción soluble de un extracto crudo de *E. coli* correspondiente a 40 ml de un cultivo que produce PulD-N se inyectó después en la columna. Las fracciones se recogieron durante las etapas de carga, lavado y elución (cambio a pH ácido). El análisis por SDS-PAGE de estas fracciones demostró que la columna era capaz de atrapar el PulD-N presente en el extracto crudo y que esto se hacía con alta especificidad, ya que la única banda visible en el gel SDS-PAGE teñido correspondía a PulD-N (figura 12b). Por tanto, se confirmó de nuevo la capacidad de los agentes de unión de discriminar PulD-N entre miles de proteínas.

#### Ejemplo 6: Unión de agentes de unión a PulD-N a PulD dodecamérico

A continuación, los inventores investigaron si los tres derivados seleccionados de Sac7d con las mayores afinidades por PulD-N (agentes de unión 6, 33 y 40) eran capaces de reconocer la proteína PulD de longitud completa integrada en la membrana externa de *E. coli*. PulD de longitud completa forma dodecámeros en la membrana (mientras que PulD-N es monomérico) (Chami et al., 2005) y no podía excluirse que el epítipo reconocido por cada uno de esos agentes de unión estuviera afectado por la multimerización de PulD, evitando de este modo su unión a PulD nativo. Cuando se mezclaron cantidades crecientes de membranas que contenían PulD de *E. coli* PAP105 (pCHAP3671 pCHAP580) con cantidades saturantes de agentes de unión 40 o 33 marcados con GFP, la cantidad de agente de unión que permanecía en los sedimentos después de centrifugación estaba correspondientemente aumentada (Fig. 13A). Los agentes de unión marcados con GFP no se sedimentaron con las membranas de PAP 105 (pCHAP3711 pCHAP580) que contenían una variante de PulD que carecía del dominio N (Guilvout et al., 2006) (Fig. 13A). Por tanto, la unión de estas quimeras GFP-agente de unión a las membranas es específica de PulD-N, y se unen al complejo nativo dodecamérico de secretina a pesar de la presencia de la marca GFP. El agente de unión 6 se unió solamente de forma muy débil a las membranas que contenían PulD (datos no mostrados).

Se usó una inmunotransferencia de Far-Western para validar la unión de los derivados de Sac7d a dodecámeros de PulD. Las tres quimeras Sac7\*-PhoA se unieron específicamente a PulD tanto monomérico como dodecamérico pero no a PulD-CS (Fig. 13B). Aunque estas tres quimeras se unieron igual de bien a PulD disociado en fenol (monomérico; véase (Hardie et al., 1996)) (datos no mostrados), mostraron de forma coherente diferentes afinidades aparentes por PulD dodecamérico, que variaban de alta (agente de unión 40), a baja (agente de unión 6).

#### Ejemplo 7: Agentes de unión a PulD-N inhiben la secreción de pululanasa y evitan la multimerización de PulD

Las quimeras Sac7\*-PhoA, en las que Sac7d o sus derivados están empareados entre el péptido señal PhoA y la parte catalítica de PhoA, se exportaron de forma eficaz al periplasma, controlado por la elevada actividad PhoA y la liberación tras choque periplásmico. Se transformaron plásmidos que codificaban las quimeras Sac7\*-PhoA en la cepa PAP7232 de *E. coli*, en los que los genes *pul* están integrados en el cromosoma. Las tres quimeras se produjeron en cantidades similares después de inducción con IPTG e inhibieron la secreción de pululanasa completamente, mientras que Sac7d-PhoA exportada no tuvo efecto. Además, no se detectaron ni dodecámeros ni monómeros de PulD en las cepas que producían alguna de las quimeras Sac7\*-PhoA (datos no mostrados). Para obtener información más precisa sobre estos fenómenos y para estudiar el destino de PulD en cepas que producen las quimeras, se produjeron en la cepa deficiente en proteasa de envuelta PAP5198 que porta pCHAP231 (para aumentar la producción de T2SS; véase Guilvout et al., 2006). Los niveles inducidos por IPTG de producción de Sac7\*-PhoA inhibieron la secreción de pululanasa en >90% (Fig. 14A). PulD dodecamérico fue mucho menos abundante, y los monómeros de PulD eran correspondientemente más abundantes (flechas en la Fig. 14B) que en controles sin quimeras o con Sac7d-PhoA, lo que demuestra que las quimeras evitan la multimerización de PulD y causan degradación de monómero PulD por proteasas de la envuelta. Se obtuvieron resultados similares con niveles no inducidos de Sac7\*33-PhoA y Sac7\*40-PhoA, pero sucedió secreción sustancial de pululanasa (>50%) y multimerización de PulD cuando Sac7\*6-PhoA estaba presente a niveles no inducidos (Fig. 14C).

#### Ejemplo 8: Selección de variantes de Sac7d que se unen a otras dianas

Se usó la mejor biblioteca (biblioteca 14) para obtener agentes de unión específicos para otras dianas, usando la misma tecnología que la descrita anteriormente.

Se realizaron selecciones por presentación en ribosoma anti-proteína centrada en las siguientes proteínas: PulDN1 (26,8 kDa), NGF (13 kDa), PknG (81,6 kDa), GarA (12 kDa) y lisozima (14,3kDa) hasta el ciclo N° 5. Durante el cuarto ciclo, el proceso de selección se realizó en paralelo, con y sin la diana. La Figura 15 muestra que, después de 4 ciclos de selección, se obtiene un producto de PCR solamente cuando está presente la diana (carriles marcados como "+"). Esto sugiere que las combinaciones de selección estaban enriquecidas con clones específicos para cada una de las dianas ensayadas.

Para evaluar adicionalmente el éxito del proceso de selección, se realizó un radioinmunoensayo (RIA) competitivo usando las proteínas diana a diversas concentraciones inhibitorias. Estos RIA muestran claramente que existe un marcado enriquecimiento para selecciones anti-PknG, anti-lisozima y anti-PulDN1 mientras que en el caso de NGF y GarA, los resultados sugieren que la selección no ha convergido aún suficientemente. La Figura 16 muestra que la afinidad anticipada promedio para las variantes anti-PknG y anti-lisozima es del orden de 1 a 10 nM.

Las combinaciones de selección anti-lisozima, anti-PknG y anti-GarA usadas en el ciclo N° 5 se seleccionaron de acuerdo con el método ELISA preparando micro-cultivos de 96 pocillos y usando sobrenadantes de lisis bacteriana después de inducir expresión del clon con IPTG durante 4 horas a 30°C. Los resultados presentados en la Figura 17 muestran claramente que hay clones específicos para cada una de las dianas. Aunque el RIA para GarA no fue alentador, no obstante se detectó una proporción significativa de clones positivos con el método ELISA. Se aisló y secuenció un clon particular que se une a GarA. La Figura 19 muestra la alineación de la secuencia de este agente de unión (GSVKVKFLYLGEKEVDTSKIWFVMRAGKrTVYFQYDDNGKYGIGWVREKDPKELLDMLARAEREKKL, SEC ID N° 47) con Sac7d.

Respecto a PknG, una proporción significativa de clones pareció unirse a BSA, lo que sugiere que se requiere un ciclo adicional para eliminarlos. Se realizaron selecciones anti-PulDN1 y anti-NGF.

Los clones anti-lisozima positivos se seleccionaron y clasificaron por Biacore para las mejores afinidades anticipadas (tiempo de disociación de complejo largo ( $k_{off}$ )), preparando microcultivos de 96 pocillos. Los 24 mejores clones seleccionados después se secuenciaron (Figura 18). Es claramente evidente una secuencia preferente, tal como la del clon Lys\_B11, ya que está representada varias veces (aunque codificada para ciertas posiciones por diferentes codones para el mismo aminoácido). La selección del ciclo N° 4 reveló una mayor diversidad (no mostrado).

Se produjeron tres clones de diferentes secuencias (Lys\_B3, Lys\_H4 y Lys\_H8) en *E. coli*, se purificaron a gran escala por IMAC y después se pasaron a través de un tamiz molecular. Los niveles de producción obtenidos fueron aproximadamente 120, 40 y 25 mg/l de cultivo, respectivamente. Las afinidades determinadas por Biacore se están obteniendo y procesando. Los datos actuales (preliminares) indican afinidades del orden de 15 nM, 3 nM y 25 nM (Lys\_B3, Lys\_H4 y Lys\_H8, respectivamente).

Los tres clones anti-lisozima (Lys\_B3, Lys\_H4 y Lys\_H8) se están caracterizando actualmente por microcalorimetría para determinar sus termoestabilidades y afinidades en solución.

Los clones obtenidos de otras selecciones también se seleccionarán por Biacore y se secuenciarán, y se determinarán las afinidades de ciertos clones de forma más precisa por Biacore.

Ejemplo 9: Selección de variantes de Sac7d que se unen a un fragmento Fc de IgG humana

#### 9.1. Selección de agentes de unión a Fc

Se biotiniló químicamente un fragmento Fc de IgG humana (PM = 50 kDa) proporcionado por Bio-Rad usando sulfosuccinimidil-6-(biotinamido) hexanoato. El grado de biotinilación, determinado por un ensayo HABA, fue de aproximadamente 2 a 3 moléculas de biotina por molécula de proteína. El Fc biotinilado se unió a neutravidina o estreptavidina inmovilizada en una placa Maxisorp (Nunc) y las secciones por presentación en ribosoma se realizaron a 4°C usando la biblioteca derivada de Sac7d (como se ha descrito anteriormente). Se realizaron cuatro rondas de selección para aislar agentes de unión. Se usaron neutravidina y estreptavidina alternativamente de una ronda a otra para evitar agentes de unión no específicos.

#### 9.2. Evaluación de la combinación de secuencias seleccionadas

Después de una ronda de selección, se obtuvo una combinación de secuencias (supuestamente agentes de unión) que portan una marca RGS-His6. Esta combinación se tradujo *in vitro* usando un extracto de *E. coli* S30, y se ensayó su actividad de unión por ELISA usando un conjugado anti-RGS-marca His6-HRP. Se detectó una actividad de unión en la combinación de selección frente al fragmento Fc inmovilizado de forma pasiva, y no se observó señal de unión frente a BSA, Neutravidina y Estreptavidina (no mostrado). Este resultado sugiere que la combinación de agentes de unión es probablemente específica del fragmento Fc y que no hay contribución significativa de Neutravidina o Estreptavidina a la actividad de unión.

También se realizó un ELISA de competición con esta combinación traducida para evaluar de forma aproximada la afinidad promedio de los agentes de unión seleccionados. En este experimento, la traducción se pre-incubó con varias concentraciones de fragmento Fc antes de la incubación en la placa ELISA en la que estaba inmovilizado el fragmento Fc. Como se muestra en la figura 20, pudo inhibirse el 20% y el 70% de la señal con 10 nM y 100 nM de Fc, respectivamente. Este resultado sugiere que la afinidad promedio esperada es de aproximadamente docenas de nM. Este intervalo en la afinidad es similar a lo observado para la selección con PulD-N.

#### Ejemplo 10: Automatización de la tecnología

Para aumentar el potencial de la tecnología descrita anteriormente, se transfiere el protocolo de selección por presentación en ribosoma (figura 21) a una plataforma automatizada para el posterior ensayo simultáneo y rápido de numerosas dianas (aún usando la biblioteca 14, o cualquier otra biblioteca, por ejemplo partiendo de una proteína OB-fold diferente de Sac7d).

Se ha equipado completamente un robot Tecan Gemini 150 para este propósito. Esta estación comprende: varios bloques de calentamiento/refrigeración/agitación, un termociclador MJ research para estación automatizada con tapa motorizada, todos los accesorios necesarios para realizar las configuraciones de reacción, limpiezas de ARN y ADN así como para medir las concentraciones de ácidos nucleicos con un lector de microplaca integrado. El objetivo es conseguir una ronda de selección sin intervención manual en absoluto. La intervención manual sería necesaria solamente entre las rondas de selección ya que tienen que remplazarse muchos residuos plásticos al final de un ciclo. Las selecciones manuales por presentación en ribosoma (figura 21) son repetitivas (cinco rondas; necesarias para enriquecer los agentes de unión), exigentes (muchas etapas de transcripción, traducción, PCR, RT-PCR) y por lo tanto largas para muchas dianas, o incluso no manejables para más de cuatro dianas por un técnico de experimentación. Por tanto, existe una gran ventaja práctica en realizar selecciones usando un robot. Después de la validación del protocolo de selección en la estación robótica (con diana PulD), es posible seleccionar agentes de unión frente hasta 48 dianas en una escala de tiempo de una o dos semanas.

#### Discusión

Los inventores exploraron los beneficios potenciales de lo que la evolución ha conseguido con OB-fold durante millones de años para adaptar este plegamiento a la unión a una amplia diversidad de ligandos: iones metálicos, azúcares, ácidos nucleicos (ARN, ADN mono y bicatenario) y proteínas. Fueron capaces de mostrar que un miembro de la familia OB-fold puede de hecho convertirse de reconocimiento de ADN en reconocimiento de proteína por evolución *in vitro*. Esto condujo satisfactoriamente a la generación de agentes de unión de alta afinidad, alta especificidad para una diana dada.

#### *Diseño de bibliotecas de Sac7d*

La observación de que la arquitectura OB-fold está polarizada, que significa que la cara de unión es siempre la misma en todas las proteínas OB-fold (Arcus, 2002), posibilitó el diseño de bibliotecas de variantes de Sac7d que retenían las propiedades biofísicas favorables de la proteína de tipo silvestre. Diferentes proteínas OB-fold pueden mostrar variación sustancial en las longitudes de los bucles 1, 3 y 4 (figura 1b y 2), que a menudo están implicadas en la unión al ligando. Aunque la mayoría de las estructuras usadas hasta ahora se basan en la aleatorización de bucles flexibles de forma similar a los de los anticuerpos (Binz et al., 2005), los inventores decidieron mantener las longitudes originales de los bucles de la proteína Sac7d nativa. Primero postularon que la parte ligeramente cóncava del área de unión se adaptaría mejor a la forma globular de proteínas usadas como dianas. Sin embargo, como se ha presentado anteriormente, la aleatorización de los restos de esta área no permitió obtener agentes de unión mejores de alto nanomolar. Esta superficie de unión potencial corresponde a un área superficial accesible a disolvente de aproximadamente  $890 \text{ \AA}^2$  y está en el intervalo típico de asociaciones anticuerpo-proteína ( $777 \pm 135 \text{ \AA}^2$ ) (Jones y Thornton, 1996). Por tanto, esta superficie debería ser suficiente para proporcionar suficiente energía de interacción para obtener agentes de unión monoméricos con alta afinidad. Entonces se ensayó la extensión del área de unión potencial con tres sustituciones más.

#### *Selección con bibliotecas de Sac7d*

Se realizaron satisfactoriamente selecciones de agentes de unión frente a diferentes dianas usando presentación en ribosoma. En contraste a las selecciones con bibliotecas de scFv vírgenes completamente sintéticos, para las que eran necesarias de 5 a 6 rondas para el enriquecimiento (Hanes et al., 2000), se observó un enriquecimiento significativo casi después de la tercera ronda con bibliotecas de Sac7d. Esto podría explicarse por un plegamiento más eficaz de Sac7d en comparación con anticuerpos, ya que no requiere ni formación de enlaces disulfuro para alcanzar su estado nativo ni la asociación correcta de dos dominios estructurales ligados (como para fragmentos scFv) para ser funcional. La elevada tasa de enriquecimiento es similar a la descrita para la estructura DARPin (Binz et al., 2004a). Finalmente, los resultados de las selecciones frente a PulD-N demuestran que las bibliotecas contienen suficiente diversidad funcional para proporcionar diferentes agentes de unión capaces de reconocer una misma proteína diana. Esto validó el diseño de bibliotecas de Sac7d de acuerdo con la invención.

*Propiedades de agentes de unión seleccionados*

La caracterización de los agentes de unión monoméricos seleccionados frente a PulD-N mostró que pueden obtenerse afinidades picomolares con bibliotecas de Sac7d. Estas afinidades están entre las mayores obtenidas con una proteína estructural sin la necesidad de una etapa de maduración de afinidad (para una revisión, véase (Hosse et al., 2006) y referencias en el mismo). Las cinéticas de las uniones también fueron remarcables, con tasas de asociación de aproximadamente  $2 \cdot 10^7 \text{ M}^{-1} \cdot \text{s}^{-1}$  que clasifican los agentes de unión para PulD-N en el intervalo superior de tasas de asociación limitadas por difusión para interacciones proteína-proteína (Gabdoulline y Wade, 1997). Las tasas de asociación no parecen estar electrostáticamente asistidas. Por tanto, las elevadas tasas de asociación podrían explicarse por la forma "preformada" del agente de unión que concuerda geométricamente con la diana sin la necesidad de cambios conformacionales.

Las altas afinidades obtenidas están asociadas con una especificidad muy elevada. Los inventores han demostrado que cualquiera que sea el contexto (extracto crudo o membranas de *E. coli*) y el enfoque (inmunotransferencia o cromatografía de afinidad), todos los agentes de unión ensayados fueron capaces de discriminar la diana entre miles de proteínas diferentes. Esta rigurosa especificidad podría estar relacionada con la rigidez de la estructura que evita ligeras adaptaciones a diferentes dianas. Además, las docenas de agentes de unión seleccionados demostraron ser específicos de diana en ensayos ELISA, lo que indica que la presión de selección aplicada era suficiente para deshacerse de las moléculas adhesivas.

Los rendimientos de proteínas recombinante para las variantes (hasta 200 mg/l de cultivo de *E. coli*) fueron mucho mayores que los presentados para Sac7d (aproximadamente 10-15 mg/l). Esta diferencia puede explicarse por la lisis que sucede pocas horas después de la inducción de la expresión del gen *sac7d* de tipo silvestre. Esta toxicidad se debe probablemente a perturbaciones de las vías de regulación inducidas por la unión de Sac7d al genoma de *E. coli*, ya que Sac7d es una proteína de unión a ADN general. Esta limitación no se observa para variantes, lo que sugiere que las variantes habían perdido de hecho su propiedad de unión a ADN.

La estabilidad termodinámica observada para tres variantes es muy similares a otras varias proteínas de termófilos (Kahsai et al., 2005) y la estabilidad térmica de los agentes de unión permaneció cercana a la de Sac7d de tipo silvestre. El clon 6, el clon menos estable analizado, siguió siendo una proteína termoestable con una  $T_m$  de aproximadamente 68°C a pH 5,6, una  $T_m$  en 8°C inferior en comparación con la de pH 7,0. Por tanto, parece que a pesar de los valores variables de  $T_m$  observados de un clon a otro, globalmente el diseño de biblioteca descrito anteriormente, combinado con el uso de una proteína de un origen extremófilo, condujo a la generación de proteínas muy estables con las propiedades deseadas de reconocimiento.

*Unión in vivo e inhibición intracelular de secreción de pululanasa*

Las tres quimeras Sac7\*-PhoA que se ensayaron se unieron a PulD de longitud completa monomérico. Sac7\*40-PhoA y Sac7\*33-PhoA se unieron bien a PulD dodecamérico, lo que indica que sus epítomos permanecen accesibles. En contraste, Sac7\*6-PhoA se unió sólo débilmente a dodecámeros de PulD pero tenía la mayor afinidad por PulD-N *in vitro*, lo que indica que su epítomo está parcialmente enmascarado tras la multimerización y es diferente de los reconocidos por Sac7\*40 y Sac7\*33. Los experimentos de competición ITC indicaron que los epítomos reconocidos por Sac7\*6 y Sac7\*40 son idénticos o solapan (datos no mostrados). Las diferencias en los efectos *in vivo* de estos dos agentes de unión, cuando están fusionados a PhoA, sugieren que su epítomos se solapan.

Las tres variantes Sac7\*-PhoA evitaron la multimerización de PulD y abordaron monómeros de PulD para la degradación por proteasas de la envuelta, bloqueando de este modo la secreción de pululanasa. Evidencias previas indicaron que el dominio N no influye en la multimerización de PulD (Guilvout et al., 2006). La dimerización de PhoA en Sac7\*-PhoA podría causar impedancia estérica y consecuente mal posicionamiento de los monómeros de PulD. Los niveles de secreción en las cepas que producen Sac7\*-PhoA permanecieron muy bajos cuando el nivel de PulD producido se aumentaba y se inactivaban las proteasas de la envuelta (Fig. 14B). La baja secreción podría deberse a la presencia de solamente unos pocos dodecámeros de PulD, oclusión de los canales, o enmascaramiento de un sitio esencial de interacción con el sustrato (Shevchik et al., 1997) u otro componente de secreción (Possot et al., 2000) por las quimeras unidas.

La reducción del nivel de Sac7\*33 o Sac7\*40 (eliminando la inducción por IPTG) no disminuyó su efecto sobre la secreción o multimerización de PulD, pero Sac7\*6-PhoA estaba casi sin efecto en estas condiciones (Fig. 14C), aunque era al menos tan abundante como las otras quimeras (Fig. 14A). Dos diferentes escenarios podrían explicar esta observación. Primero, el Sac7\*6-PhoA altamente abundante se une a casi todos los monómeros de PulD y evita su multimerización. Sin embargo, cuando los niveles de Sac7\*6-PhoA son inferiores, no puede competir de forma eficaz con las interacciones monómero-monómero de PulD, y suficientes multímeros se ensamblan para permitir una secreción eficaz. La afinidad aparentemente inferior de Sac7\*6-PhoA por PulD dodecamérico es insuficiente para evitar la secreción. Segundo, la unión de la chaperona PulS a monómeros de PulD, un pre-requisito para su correcto direccionamiento a la membrana externa (Guilvout et al., 2006; Hardie et al., 1996), evita la unión de Sac7\*6-PhoA y permite la correcta multimerización. Es este escenario, el resultado depende de la proteína que se una primero a

PulD, PulS o Sac7\*6-PhoA. Ambos escenarios están en acuerdo con el hecho de que el epítipo reconocido por Sac7\*6 está fuertemente enmascarado tras la multimerización de PulD, lo que sugiere que está en una interfaz entre dos monómeros.

5 *Aplicaciones biotecnológicas potenciales*

Estos agentes de unión retenían la mayoría de las propiedades biofísicas muy favorables en una estructura alternativa a los anticuerpos. Sus propiedades son similares a las alternativas previamente propuestas a los anticuerpos tales como aficuerpos (Nord et al., 1997), fibronectina (Xu et al., 2002), o anquirinas (Binz et al., 2004a). De hecho, se expresan muy bien en *E. coli* (en el citoplasma o el periplasma), son estables, solubles, son capaces de reconocer una diana proteica con alta afinidad y alta especificidad, y pueden fusionarse de forma funcional con diferentes proteínas indicadoras (PhoA y GFP). En otras palabras, son baratos de producir, fáciles de purificar y manipular. Esto abre la puerta a varias aplicaciones biotecnológicas.

15 Además, su tamaño muy pequeño, tres o diecinueve veces más pequeño que un scFv o una IgG, respectivamente, los hace candidatos ideales para el diseño de proteínas quiméricas en las que son necesarios módulos de unión manteniendo al mismo tiempo un tamaño razonable. También es concebible ligar varios agentes de unión con diferentes especificidades y, por tanto, construir fusiones con múltiples especificidades. Otra ventaja de los agentes de unión pequeños tales como Sac7d es su capacidad potencial de unirse a epítipos enterrados que son estéricamente inaccesibles a los anticuerpos naturales o sus fragmentos.

Puede apreciarse que la capacidad de los agentes de unión de unirse a proteínas diana no está limitada a dianas con un alto peso molecular. De hecho, los resultados ya obtenidos, además de los referentes a PulD, muestran que el banco desarrollado a partir de Sac7d ha permitido el aislamiento de clones que se unen específicamente a tres nuevas dianas, que cubren un amplio intervalo de pesos moleculares (Lisozima = 14,3 kDa, GarA = 17,3 kDa, PknG = 81,6 kDa). La expresión de estos agentes de unión en *E. coli* es muy superior a los niveles observados para fragmentos recombinantes de anticuerpo. Las afinidades obtenidas son del orden de nM sin selección basada en valores largos de  $k_{off}$  (selección por disociación), debido a la escasez de tiempo. Esta etapa de selección se necesita para obtener afinidades sub-nanomolares, y se está realizando actualmente.

Para abordar cómo de específicos podrían ser los agentes de unión de PulD, los inventores ya han demostrado que podrían usarse para el desarrollo de reactivos de detección. Por ejemplo, se pueden conseguir ELISA de una etapa o transferencias de Western por fusión de agentes de unión a fosfatasa alcalina. Otro ejemplo es el uso de fusiones con GFP para la detección *in vitro* que allana el camino a la localización intracelular *in vivo*. También se demostró que puede realizarse cromatografía de afinidad con estos agentes de unión y que puede purificarse una proteína hasta homogeneidad por purificación de una etapa. Otras aplicaciones podrían usar estos agentes de unión, tales como series de chip de proteína, biosensores o knock out *in vivo* conmutable por ejemplo.

40 *Conclusiones*

La estrategia de usar una proteína OB-fold para obtener agentes de unión capaces de reconocer proteínas diana con alta afinidad y especificidad muy fina se validó satisfactoriamente por los resultados anteriores. El punto de partida fue Sac7d, una proteína de unión a ADN general. Sac7d y sus derivados son ahora capaces de reconocer dos familias estructuralmente no relacionadas de ligandos: ADN y proteínas. Por tanto, los inventores han reproducido *in vitro* la adaptación de la unión de un OB-fold para una diana dada, tal como ya ha conseguido la naturaleza. Por lo tanto es concebible, que Sac7d también pueda adaptarse a otros ligandos conocidos de proteínas OB-fold, tales como ADN monocatenario, ARN o azúcares. Además de su uso potencial en experimentos de knockout funcional, estos agentes de unión podrían usarse para cromatografía de afinidad, detección, experimentos de localización *in vivo*, etc. De hecho, las propiedades biofísicas favorables de estos agentes de unión, junto con su fácil fusión a diferentes proteínas indicadoras y su pequeño tamaño (3 y 19 veces más pequeños que scFv e IgG, respectivamente), podrían facilitar estas aplicaciones.

Referencias

55 Agback, P., Baumann, H., Knapp, S., Ladenstein, R, y Hard, T. (1998). Architecture of nonspecific protein-DNA interactions in the Sso7d-DNA complex. *Nat Struct Biol* 5, 579-584.

Arcus, V. (2002). OB-fold domains: a snapshot of the evolution of sequence, structure and function. *Curr Opin Struct Biol* 12, 794-801.

60 Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., y Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Res* 28, 235-242.

Binz, H. K., Amstutz, P., Kohl, A., Stumpp, M. T., Briand, C, Forrer, P., Grutter, M. G., y Pluckthun, A. (2004a). High-affinity binders selected from designed ankyrin repeat protein libraries. *Nat Biotechnol* 22, 575-582.

- Binz, H. K., Amstutz, P., Kohl, A., Stumpp, M. T., Briand, C., Forrer, P., Grutter, M. G., y Pluckthun, A. (2004b). High-affinity binders selected from designed ankyrin repeat protein libraries. *Nat Biotechnol* 22, 575-582. Epub 2004 Abr 2018.
- 5 Binz, H. K., Amstutz, P., y Pluckthun, A. (2005). Engineering novel binding proteins from nonimmunoglobulin domains. *Nat Biotechnol* 23,1257-1268.
- Boder, E. T., y Wittrup, K. D. (1997). Yeast surface display for screening combinatorial polypeptide libraries. *Nat Biotechnol* 15, 553-557.
- 10 Chami, M., Guilvout, I., Gregorini, M., Remigy, H. W., Muller, S. A., Valerio, M., Engel, A., Pugsley, A. P., y Bayan, N. (2005). Structural insights into the secretin PulD and its trypsin-resistant core. *J Biol Chem* 280, 37732-37741.
- 15 Daefler, S., Guilvout, I., Hardie, K. R., Pugsley, A. P., y Russel, M. (1997). The C-terminal domain of the secretin PulD contains the binding site for its cognate chaperone, PulS, and confers PulS dependence On pIVf function. *Mol Microbiol* 24, 465-475.
- d'Enfert, C., Chapon, C, y Pugsley, A. P. (1987). Export and secretion of the lipoprotein pullulanase by *Klebsiella pneumoniae*. *Mol Microbiol* 1, 107-116.
- 20 d'Enfert, C, Reyss, I., Wandersman, C, y Pugsley, A. P. (1989). Protein secretion by gram-negative bacteria. Characterization of two membrane proteins required for pullulanase secretion by *Escherichia coli* K-12. *J Biol Chem* 264, 17462-17468.
- 25 Edmondson, S. P., y Shriver, J. W. (2001). DNA binding proteins Sac7d and Sso7d from *Sulfolobus*. *Methods Enzymol* 334,129-145.
- Freire, E. (2001). The thermodynamic linkage between protein structure, stability, and function. *Methods Mol Biol* 168, 37-68.
- 30 Gabdoulline, R. R., y Wade, R. C. (1997). Simulation of the diffusional association of bamase and barstar. *Biophys J* 72, 1917-1929.
- 35 Gao, Y. G., Su, S. Y, Robinson, H., Padmanabhan, S., Lim, L., McCrary, B. S., Edmondson, S. P., Shriver, J. W., y Wang, A. H. (1998). The crystal structure of the hyperthermophile chromosomal protein Sso7d bound to DNA. *Nat Struct Biol* 5, 782-786.
- Georgiou, G., Stathopoulos, C, Daugherty, P. S., Nayak, A. R., Iverson, B. L., y Curtiss, R., 3rd (1997). Display of heterologous proteins on the surface of microorganisms: from the screening of combinatorial libraries to live recombinant vaccines. *Nat Biotechnol* 75, 29-34.
- 40 Guilvout, I., Chami, M., Engel, A., Pugsley, A. P., y Bayan, N. (2006). Bacterial outer membrane secretin PulD assembles and inserts into the inner membrane in the absence of its pilotin. *Embo J* 25, 5241-5249.
- 45 Guilvout, I., Hardie, K. R., Sauvonnnet, N., y Pugsley, A. P. (1999). Genetic dissection of the outer membrane secretin PulD: are there distinct domains for multimerization and secretion specificity? *J Bacteriol* 181, 7212-7220.
- Hanes, J., Jermutus, L., Weber-Bornhauser, S., Bosshard, H. R., and Pluckthun, A. (1998). Ribosome display efficiently selects and evolves high-affinity antibodies in vitro from immune libraries. *Proc Natl Acad Sci U S A* 95, 14130-14135.
- 50 Hanes, J., y Pluckthun, A. (1997). In vitro selection and evolution of functional proteins by using ribosome display. *Proc Natl Acad Sci U S A* 94, 4937-4942.
- Hanes, J., Schaffitzel, C, Knappik, A., y Pluckthun, A. (2000). Picomolar affinity antibodies from a fully synthetic naive library selected and evolved by ribosome display. *Nat Biotechnol* 18, 1287-1292.
- 55 Hardie, K. R., Lory, S., y Pugsley, A. P. (1996). Insertion of an outer membrane protein in *Escherichia coli* requires a chaperone-like protein. *Embo J* 15, 978-988.
- 60 He, M., y Taussig, M. J. (2002). Ribosome display: cell-free protein display technology. *Brief Funct Genomic Proteomic* 1, 204-212.
- Hible, G., Renault, L., Schaeffer, F., Christova, P., Zoe Radulescu, A., Evrin, C, Gilles, A. M., y Cherfils, J. (2005). Calorimetric and crystallographic analysis of the oligomeric structure of *Escherichia coli* GMP kinase. *J Mol Biol* 352, 1044-1059.
- 65

- 5 Holm, L., y Park, J. (2000). DaliLite workbench for protein structure comparison. *Bioinformatics* 16,566-567.
- Holm, L., y Sander, C. (1998). Touring protein fold space with Dali/FSSP. *Nucleic Acids Res* 26, 316-319.
- 5 Hosse, R. J., Rothe, A., y Power, B. E. (2006). A new generation of protein display scaffolds for molecular recognition. *Protein Sci* 15, 14-27.
- Jermutus, L., Honegger, A., Schwesinger, F., Hanes, J., y Pluckthun, A. (2001). Tailoring in vitro evolution for protein affinity or stability. *Proc Natl Acad Sci U S A* 98, 75-80.
- 10 Jones, S., y Thornton, J. M. (1996). Principles of protein-protein interactions. *Proc Natl Acad Sci U S A* 93, 13-20.
- Kahsai, M. A., Martin, E., Edmondson, S. P., y Shriver, J. W. (2005). Stability and flexibility in the structure of the hyperthermophile DNA-binding protein Sac7d. *Biochemistry* 44, 13500-13509.
- 15 Kitov, P. I., Sadowska, J. M., Mulvey, G., Armstrong, G. D., Ling, H., Pannu, N. S., Read, R. J., y Bundle, D. R. (2000). Shiga-like toxins are neutralized by tailored multivalent carbohydrate ligands. *Nature* 403, 669-672.
- Kloks, C. P., Spronk, C. A., Lasonder, E., Hoffmann, A., Vuister, G. W., Grzesiek, S., y Hilbers, C. W. (2002). The solution structure and DNA-binding properties of the cold-shock domain of the human Y-box protein YB-1. *J Mol Biol* 316, 317-326.
- 20 Lopez, R., Silventoinen, V., Robinson, S., Kibria, A., y Gish, W. (2003). WU-Blast2 server at the European Bioinformatics Institute. *Nucleic Acids Res* 31, 3795-3798.
- 25 Mathonet, P., y Fastrez, J. (2004). Engineering of non-natural receptors. *Curr Opin Struct Biol* 14, 505-511.
- McAfee, J. G., Edmondson, S. P., Datta, P. K., Shriver, J. W., y Gupta, R. (1995). Gene cloning, expression, and characterization of the Sac7 proteins from the hyperthermophile *Sulfolobus acidocaldarius*. *Biochemistry* 34, 10063-10077.
- 30 McCrary, B. S., Edmondson, S. P., y Shriver, J. W. (1996). Hyperthermophile protein folding thermodynamics: differential scanning calorimetry and chemical denaturation of Sac7d. *J Mol Biol* 264, 784-805.
- 35 Mitton-Fry, R. M., Anderson, E. M., Hughes, T. R., Lundblad, V., y Wuttke, D. S. (2002). Conserved structure for single-stranded telomeric DNA recognition. *Science* 296, 145-147.
- Murzin, A. G. (1993). OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *Embo J* 12, 861-867.
- 40 Nieba, L., Krebber, A., y Pluckthun, A. (1996). Competition BIAcore for measuring true affinities: large differences from values determined from binding kinetics. *Anal Biochem* 234, 155-165.
- Nord, K., Gunneriusson, E., Ringdahl, J., Stahl, S., Uhlen, M., y Nygren, P. A. (1997). Binding proteins selected from combinatorial libraries of an alpha-helical bacterial receptor domain. *Nat Biotechnol* 15,112-111.
- 45 Notredame, C., Higgins, D. G., y Heringa, J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* 302, 205-217.
- 50 Ostermeier, C., Iwata, S., Ludwig, B., y Michel, H. (1995). Fv fragment-mediated crystallization of the membrane protein bacterial cytochrome c oxidase. *Nat Struct Biol* 2, 842-846.
- Papageorgiou, A. C., Tranter, H. S., y Acharya, K. R. (1998). Crystal structure of microbial superantigen staphylococcal enterotoxin B at 1.5 Å resolution: implications for superantigen recognition by MHC class II molecules and T-cell receptors. *J Mol Biol* 277, 61-79.
- 55 Pearl, F. M., Bennett, C. F., Bray, J. E., Harrison, A. P., Martin, N., Shepherd, A., Sillitoe, I., Thornton, J., y Orengo, C. A. (2003). The CATH database: an extended protein family resource for structural and functional genomics. *Nucleic Acids Res* 31,452-455.
- 60 Plotnikov, V. V., Brandts, J. M., Lin, L. N., y Brandts, J. F. (1997). A new ultrasensitive scanning calorimeter. *Anal Biochem* 250, 237-244.
- 65 Possot, O. M., Vignon, G., Bomchil, N., Ebel, F., y Pugsley, A. P. (2000). Multiple interactions between pullulanase secreton components involved in stabilization and cytoplasmic membrane association of PulE. *J Bacterid* 182, 2142-2152.

- Qian, J., Stenger, B., Wilson, C. A., Lin, J., Jansen, R., Teichmann, S. A., Park, J., Krebs, W. G., Yu, H., Alexandrov, V., *et al.* (2001). PartsList: a web-based system for dynamically ranking protein folds based on disparate attributes, including whole-genome expression and interaction information. *Nucleic Acids Res* 29, 1750-1764.
- 5 Robinson, H., Gao, Y. G., McCrary, B. S., Edmondson, S. P., Shriver, J. W., y Wang, A. H. (1998). The hyperthermophile chromosomal protein Sac7d sharply kinks DNA. *Nature* 392, 202-205.
- Roussel, A., Anderson, B. F., Baker, H. M., Fraser, J. D., y Baker, E. N. (1997). Crystal structure of the streptococcal superantigen SPE-C: dimerization and zinc binding suggest a novel mode of interaction with MHC class II molecules. *Nat Struct Biol* 4, 635-643.
- 10 Samygina, V. R., Popov, A. N., Rodina, E. V., Vorobyeva, N. N., Lamzin, V. S., Polyakov, K. M., Kurilova, S. A., Nazarova, T. I., y Avaeva, S. M. (2001). The structures of *Escherichia coli* inorganic pyrophosphatase complexed with Ca(2+) or CaPP(i) at atomic resolution and their mechanistic implications. *J Mol Biol* 314,633-645.
- 15 Schaffitzel, C., Hanes, J., Jermutus, L., y Pluckthun, A. (1999). Ribosome display: an in vitro method for selection and evolution of antibodies from libraries. *J Immunol Methods* 231,119-135.
- Shevchik, V. E., Robert-Baudouy, J., y Condemine, G. (1997). Specific interaction between OutD, an *Erwinia chrysanthemi* outer membrane protein of the general secretory pathway, and secreted proteins. *Embo J* 16, 3007-3016.
- 20 Smith, G. P. (1985). Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* 228, 1315-1317.
- 25 Su, S., Gao, Y. G., Robinson, H., Liaw, Y. C., Edmondson, S. P., Shriver, J. W., y Wang, A. H. (2000). Crystal structures of the chromosomal proteins Sso7d/Sac7d bound to DNA containing T-G mismatched base-pairs. *J Mol Biol* 303,395-403.
- 30 Wilson, D. S., Keefe, A. D., y Szostak, J. W. (2001). The use of mRNA display to select high-affinity protein-binding peptides. *Proc Natl Acad Sci U S A* 98, 3750-3755.
- Xu, L., Aha, P., Gu, K., Kuimelis, R. G., Kurz, M., Lam, T., Lim, A. C., Liu, H., Lohse, P. A., Sun, L., *et al.* (2002). Directed evolution of high-affinity antibody mimics using mRNA display. *Chem Biol* 9, 933-942.
- 35

LISTADO DE SECUENCIAS

- <110> INSTITUT PASTEUR  
CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE  
40 PECORARI, Frederic  
ALZARI, Pedro
- <120> OB-FOLD USADO COMO ESTRUCTURA PARA DISEÑO POR INGENIERÍA DE NUEVOS AGENTES DE UNIÓN ESPECÍFICOS
- 45 <130> VMA/ah-F226/133
- <160> 47
- 50 <170> Patentin versión 3.3
- <210> 1  
<211> 66  
<212> PRT  
55 <213> *Sulfolobus sp.*
- <400> 1



ES 2 536 509 T3

Met Val Lys Val Lys Phe Lys Tyr Lys Gly Glu Glu Lys Glu Val Asp  
 1 5 10 15

Thr Ser Lys Ile Lys Lys Val Trp Arg Val Gly Lys Met Val Ser Phe  
 20 25 30

Thr Tyr Asp Asp Asn Gly Lys Thr Gly Arg Gly Ala Val Ser Glu Lys  
 35 40 45

Asp Ala Pro Lys Glu Leu Leu Asp Met Leu Ala Arg Ala Glu Arg Glu  
 50 55 60

Lys Lys  
 65

5 <210> 2  
 <211> 79  
 <212> PRT  
 <213> Artificial

10 <220>  
 <223> Clon 6

15 <220>  
 <221> DOMINIO  
 <222> (1)..(12)  
 <223> Marca MRGS (His)6

<400> 2

Met Arg Gly Ser His His His His His His Gly Ser Val Lys Val Lys  
 1 5 10 15

Phe Val Leu Gly Gly Glu Glu Lys Glu Val Asp Thr Ser Lys Ile Arg  
 20 25 30

His Val Tyr Arg Tyr Gly Lys His Val Thr Phe Ser Tyr Asp Asp Asn  
 35 40 45

Gly Lys Leu Gly Leu Gly Leu Val Lys Glu Lys Asp Ala Pro Lys Glu  
 50 55 60

Leu Leu Asp Met Leu Ala Arg Ala Glu Arg Glu Lys Lys Leu Asn  
 65 70 75

20 <210> 3  
 <211> 79  
 <212> PRT  
 <213> Artificial

25 <220>  
 <223> Clon 33

30 <220>  
 <221> DOMINIO

ES 2 536 509 T3

<222> (1)..(12)  
 <223> Marca MRGS (His)6

<400> 3

5

```

Met Arg Gly Ser His His His His His His Gly Ser Val Lys Val Lys
1           5           10           15

Phe Thr Trp Tyr Gly Glu Glu Lys Glu Val Asp Thr Ser Lys Ile Met
          20           25           30

Ser Val Cys Arg Trp Gly Lys Arg Val Ala Phe Ser Tyr Asp Asp Asn
          35           40           45

Gly Lys Ile Asp Tyr Gly Leu Val Asp Glu Lys Asp Ala Pro Lys Glu
50           55           60

Leu Leu Asp Met Leu Ala Arg Ala Glu Arg Glu Lys Lys Leu Asn
65           70           75
    
```

<210> 4  
 <211> 73  
 <212> PRT  
 <213> Artificial

10

<220>  
 <223> Clon 39

15

<220>  
 <221> DOMINIO  
 <222> (1)..(12)  
 <223> Marca MRGS (His)6

20

<400> 4

```

Met Arg Gly Ser His His His His His His Gly Ser Val Lys Val Lys
1           5           10           15

Phe Cys Trp Glu Gly Glu Glu Lys Glu Val Asp Thr Ser Lys Ile Arg
          20           25           30

His Val Tyr Arg Tyr Gly Lys Tyr Val Thr Phe Ser Tyr Asp Asp Asn
          35           40           45

Gly Lys Leu Gly Leu Gly Leu Val Asn Glu Lys Asp Ala Pro Lys Glu
50           55           60

Leu Leu Asp Met Leu Ala Arg Ala Ala
65           70
    
```

25

<210> 5  
 <211> 79  
 <212> PRT  
 <213> Artificial

ES 2 536 509 T3

<220>  
 <223> Clon 40

5 <220>  
 <221> DOMINIO  
 <222> (1) .. (12)  
 <223> Marca MRGS (His)6

10 <400> 5

Met Arg Gly Ser His His His His His His Gly Ser Val Lys Val Lys  
 1 5 10 15

Phe Gly Glu His Gly Glu Glu Lys Glu Val Asp Thr Ser Lys Ile Met  
 20 25 30

Ser Val Val Arg Phe Gly Lys Trp Val Cys Phe Arg Tyr Asp Asp Asn  
 35 40 45

Gly Lys Ala Gly Tyr Gly Ala Val Leu Glu Lys Asp Ala Pro Lys Glu  
 50 55 60

Leu Leu Asp Met Leu Ala Arg Ala Glu Arg Glu Lys Lys Leu Asn  
 65 70 75

15 <210> 6  
 <211> 79  
 <212> PRT  
 <213> Artificial

20 <220>  
 <223> Clon 41

25 <220>  
 <221> DOMINIO  
 <222> (1)..(12)  
 <223> Marca MRGS (His)6

<400> 6

ES 2 536 509 T3

Met Arg Gly Ser His His His His His His Gly Ser Val Lys Val Lys  
1 5 10 15

Phe Val Met Ser Gly Glu Glu Lys Glu Val Asp Thr Ser Lys Ile Met  
20 25 30

Ser Val Val Arg Phe Gly Lys Trp Val Cys Phe Arg Tyr Asp Asp Asn  
35 40 45

Gly Lys Ala Gly Tyr Gly Ala Val Leu Glu Lys Asp Ala Pro Lys Glu  
50 55 60

Leu Leu Asp Met Leu Ala Arg Ala Glu Arg Glu Lys Lys Leu Asn  
65 70 75

- 5 <210> 7
- <211> 548
- <212> PRT
- <213> Artificial
- 10 <220>
- <223> Fusión de PhoA con clon 6
- 15 <220>
- <221> SEÑAL
- <222> (1)..(21)
- <223> Péptido señal
- 20 <220>
- <221> CADENA
- <222> (29)..(33)
- <223> enlazador
- 25 <220>
- <221> UNIÓN
- <222> (34).. (98)
- <223> PalD-N
- 30 <220>
- <221> CADENA
- <222> (99)..(104)
- <223> enlazador
- 35 <220>
- <221> SITIO ACT
- <222> (105).. (548)
- <223> PhoA
- <400> 7

ES 2 536 509 T3

Met Lys Gln Ser Thr Ile Ala Leu Ala Leu Leu Pro Leu Leu Phe Thr  
1 5 10 15

Pro Val Thr Lys Ala Arg Thr Pro Glu Met Pro Val Asp Gly Thr Gly  
20 25 30

Ser Val Lys Val Lys Phe Val Leu Gly Gly Glu Glu Lys Glu Val Asp  
35 40 45

Thr Ser Lys Ile Arg His Val Tyr Arg Tyr Gly Lys His Val Thr Phe  
50 55 60

Ser Tyr Asp Asp Asn Gly Lys Leu Gly Leu Gly Leu Val Lys Glu Lys  
65 70 75 80

Asp Ala Pro Lys Glu Leu Leu Asp Met Leu Ala Arg Ala Glu Arg Glu  
85 90 95

Lys Lys Leu Arg Ala Gln Pro Gly Val Leu Glu Asn Arg Ala Ala Gln  
100 105 110

Gly Asp Ile Thr Ala Pro Gly Gly Ala Arg Arg Leu Thr Gly Asp Gln  
115 120 125

Thr Ala Ala Leu Arg Asp Ser Leu Ser Asp Lys Pro Ala Lys Asn Ile  
130 135 140

Ile Leu Leu Ile Gly Asp Gly Met Gly Asp Ser Glu Ile Thr Ala Ala  
145 150 155 160

Arg Asn Tyr Ala Glu Gly Ala Gly Gly Phe Phe Lys Gly Ile Asp Ala  
165 170 175

Leu Pro Leu Thr Gly Gln Tyr Thr His Tyr Ala Leu Asn Lys Lys Thr  
180 185 190

ES 2 536 509 T3

Gly Lys Pro Asp Tyr Val Thr Asp Ser Ala Ala Ser Ala Thr Ala Trp  
 195 200 205

Ser Thr Gly Val Lys Thr Tyr Asn Gly Ala Leu Gly Val Asp Ile His  
 210 215 220

Glu Lys Asp His Pro Thr Ile Leu Glu Met Ala Lys Ala Ala Gly Leu  
 225 230 235 240

Ala Thr Gly Asn Val Ser Thr Ala Glu Leu Gln Gly Ala Thr Pro Ala  
 245 250 255

Ala Leu Val Ala His Val Thr Ser Arg Lys Cys Tyr Gly Pro Ser Ala  
 260 265 270

Thr Ser Glu Lys Cys Pro Gly Asn Ala Leu Glu Lys Gly Gly Lys Gly  
 275 280 285

Ser Ile Thr Glu Gln Leu Leu Asn Ala Arg Ala Asp Val Thr Leu Gly  
 290 295 300

Gly Gly Ala Lys Thr Phe Ala Glu Thr Ala Thr Ala Gly Glu Trp Gln  
 305 310 315 320

Gly Lys Thr Leu Arg Glu Gln Ala Gln Ala Arg Gly Tyr Gln Leu Val  
 325 330 335

Ser Asp Ala Ala Ser Leu Asn Ser Val Thr Glu Ala Asn Gln Gln Lys  
 340 345 350

Pro Leu Leu Gly Leu Phe Ala Asp Gly Asn Met Pro Val Arg Trp Leu  
 355 360 365

Gly Pro Lys Ala Thr Tyr His Gly Asn Ile Asp Lys Pro Ala Val Thr  
 370 375 380

Cys Thr Pro Asn Pro Gln Arg Asn Asp Ser Val Pro Thr Leu Ala Gln  
 385 390 395 400

Met Thr Asp Lys Ala Ile Glu Leu Leu Ser Lys Asn Glu Lys Gly Phe  
 405 410 415

Phe Leu Gln Val Glu Gly Ala Ser Ile Asp Lys Gln Gly His Ala Ala  
 420 425 430

ES 2 536 509 T3

Asn Pro Cys Gly Gln Ile Gly Glu Thr Val Asp Leu Asp Glu Ala Val  
 435 440 445

Gln Arg Ala Leu Glu Phe Ala Lys Lys Glu Gly Asn Thr Leu Val Ile  
 450 455 460

Val Thr Ala Asp His Ala His Ala Ser Gln Ile Val Ala Pro Asp Thr  
 465 470 475 480

Lys Ala Pro Gly Leu Thr Gln Ala Leu Asn Thr Lys Asp Gly Ala Val  
 485 490 495

Met Val Met Ser Tyr Gly Asn Ser Glu Glu Asp Ser Gln Glu His Thr  
 500 505 510

Gly Ser Gln Leu Arg Ile Ala Ala Tyr Gly Pro His Ala Ala Asn Val  
 515 520 525

Val Gly Leu Thr Asp Gln Thr Asp Leu Phe Tyr Thr Met Lys Ala Ala  
 530 535 540

Leu Gly Leu Lys  
 545

5 <210> 8  
 <211> 330  
 <212> PRT  
 <213> Artificial

10 <220>  
 <223> Fusión de GFP con clon 6

15 <220>  
 <221> UNIÓN  
 <222> (13)..(77)  
 <223> PalD-N

20 <220>  
 <221> CADENA  
 <222> (78)..(87)  
 <223> enlazador

25 <220>  
 <221> SITIO ACT  
 <222> (88)..(329)  
 <223> GFP

<400> 8

Met Arg Gly Ser His His His His His His Gly Ser Val Lys Val Lys  
 1 5 10 15

ES 2 536 509 T3

Phe Val Leu Gly Gly Glu Glu Lys Glu Val Asp Thr Ser Lys Ile Arg  
20 25 30

His Val Tyr Arg Tyr Gly Lys His Val Thr Phe Ser Tyr Asp Asp Asn  
35 40 45

Gly Lys Leu Gly Leu Gly Leu Val Lys Glu Lys Asp Ala Pro Lys Glu  
50 55 60

Leu Leu Asp Met Leu Ala Arg Ala Glu Arg Glu Lys Lys Leu Gly Ser  
65 70 75 80

Ala Gly Ser Ala Ala Gly Ser Gly Glu Phe Ser Lys Gly Glu Glu Leu  
85 90 95

Phe Thr Gly Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn  
100 105 110

Gly His Lys Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr  
115 120 125

Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val  
130 135 140

Pro Trp Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe  
145 150 155 160

Ser Arg Tyr Pro Asp His Met Lys Arg His Asp Phe Phe Lys Ser Ala  
165 170 175

Met Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile Ser Phe Lys Asp Asp  
180 185 190

Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu  
195 200 205

Val Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn  
210 215 220

Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr  
225 230 235 240

Ile Thr Ala Asp Lys Gln Lys Asn Gly Ile Lys Ala Asn Phe Lys Ile  
245 250 255



ES 2 536 509 T3

Arg His Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln  
 260 265 270

Gln Asn Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His  
 275 280 285

Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg  
 290 295 300

Asp His Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr His  
 305 310 315 320

Gly Met Asp Glu Leu Tyr Lys Thr Gly Asn  
 325 330

5 <210> 9  
 <211> 66  
 <212> ADN  
 <213> Artificial

10 <220>  
 <223> Cebador SCI  
 <400> 9

gaaactccta ggtattgtgc tgacgacccc gatcgcgata tctagctttg cggtgaaagt 60  
 gaaatt 66

15 <210> 10  
 <211> 68  
 <212> ADN  
 <213> Artificial

20 <220>  
 <223> Cebador SC2  
 <400> 10

gatcttgctg gtgtccactt cttttcttc gcctttatat ttaaatttca ctttcaccgc 60  
 aaagctag 68

30 <210> 11  
 <211> 78  
 <212> ADN  
 <213> Artificial

<220>  
 <223> Cebador SC3  
 <400> 11

gaagtggaca ccagcaagat caagaaagtt tggcgtgtgg gcaaaatggt gagctttacc 60  
 tacgacgaca acggcaag 78

<210> 12

ES 2 536 509 T3

<211> 60  
<212> ADN  
<213> Artificial

5 <220>  
<223> Cebador SC4

<400> 12  
ctctttcggg gcatctttct cgctcacggc gccacggcgc gtcttgccgt tgcgctcgta 60

10 <210> 13  
<211> 59  
<212> ADN  
<213> Artificial

15 <220>  
<223> Cebador SC5

<400> 13  
20 gagaaagatg ccccgaaaga gttattagat atgtagcgc gtgcggaaag ctcaacca 59

<210> 14  
<211> 23  
<212> ADN  
25 <213> Artificial

<220>  
<223> Cebador SC6

30 <400> 14  
tggtggttga agctttccgc acg 23

<210> 15  
<211> 43  
35 <212> ADN  
<213> Artificial

<220>  
<223> Cebador SC07

40 <400> 15  
attaatgga ccggatccgt gaaagtgaaa tttaatata aag 43

<210> 16  
<211> 49  
45 <212> ADN  
<213> Artificial

<220>  
50 <223> Cebador SC08

<400> 16  
ataahtgagc tctaagcttt tttcacggt cgcacgcgc taacatatc 49

55 <210> 17  
<211> 49  
<212> ADN  
<213> Artificial

60 <220>  
<223> Cebador T7C

<400> 17  
65 atacgaaatt aatagcactc actatagga gaccacaacg gtttcctc 49

<210> 18

ES 2 536 509 T3

<211> 72  
 <212> ADN  
 <213> Artificial  
 5 <220>  
 <223> Cebador SDA\_MRGS  
 <400> 18  
 agaccacaac ggtttccctc tagaataat tttgtttaac ttttaagaagg agatatatcc 60  
 10 atgagaggat cg 72  
 <210> 19  
 <211> 63  
 <212> ADN  
 15 <213> Artificial  
 <220>  
 <223> Cebador SClib1  
 20 <400> 19  
 ggagatatat ccatgagagg atcgcatcac catcaccatc acggatccgt caaggtgaaa 60  
 ttc 63  
 <210> 20  
 <211> 63  
 <212> ADN  
 <213> Artificial  
 25 <220>  
 <223> Cebador SClib2  
 <220>  
 <221> misc\_feature  
 <222> (22)..(23)  
 35 <223> n es a, c, g o t  
 <220>  
 <221> misc\_feature  
 <222> (24)..(24)  
 40 <223> s es g o c  
 <220>  
 <221> misc\_feature  
 <222> (25)..(26)  
 45 <223> n es a, c, g o t  
 <220>  
 <221> misc feature  
 <222> (27)..(27)  
 50 <223> s es g o c  
 <220>  
 <221> misc\_feature  
 <222> (28)..(29)  
 55 <223> n es a, c, g o t  
 <220>  
 <221> misc\_feature  
 <222> (30)..(30)  
 60 <223> s es g o c

<400> 20

**ggatccgtca aggtgaaatt cnnnsnnsnsnns ggcgaagaaa aagaagtgga cactagtaag**

**60**

**atc**

**63**

- 5 <210> 21  
<211> 78  
<212> ADN  
<213> Artificial
- 10 <220>  
<223> Cebador SCLib3
- 15 <220>  
<221> misc\_feature  
<222> (19)..(19)  
<223> s es g o c
- 20 <220>  
<221> misc\_feature  
<222> (20)..(21)  
<223> n es a, c, g o t
- 25 <220>  
<221> misc\_feature  
<222> (25)..(25)  
<223> s es g o c
- 30 <220>  
<221> misc\_feature  
<222> (26)..(27)  
<223> n es a, c, g o t
- 35 <220>  
<221> misc\_feature  
<222> (31)..(31)  
<223> s es g o c
- 40 <220>  
<221> misc\_feature  
<222> (32)..(33)  
<223> n es a, c, g o t
- 45 <220>  
<221> misc\_feature  
<222> (40)..(40)  
<223> s es g o c
- 50 <220>  
<221> misc\_feature  
<222> (41)..(42)  
<223> n es a, c, g o t
- 55 <220>  
<221> misc\_feature  
<222> (46)..(46)  
<223> s es g o c
- 60 <220>  
<221> misc\_feature  
<222> (47)..(48)  
<223> n es a, c, g o t
- <220>  
<221> misc\_feature

<222> (52)..(52)  
 <223> s es g o c  
  
 <220>  
 5 <221> misc\_feature  
 <222> (53)..(54)  
 <223> n es a, c, g o t  
  
 <220>  
 10 <221> misc\_feature  
 <222> (55)..(55)  
 <223> s es g o c  
  
 <220>  
 15 <221> misc\_feature  
 <222> (56)..(57)  
 <223> n es a, c, g o t  
  
 <400> 21  
 20  
     cttgccgttg tcgtcgtasn naaasnncac snnttgccs nnacgsnnaa csmsnngat      60  
     cttactagtg tccacttc      78  
  
 <210> 22  
 <211> 66  
 25 <212> ADN  
 <213> Artificial  
  
 <220>  
 30 <223> Cebador SCLib4  
  
 <220>  
 <221> misc\_feature  
 <222> (28)..(28)  
 <223> s es g o c  
 35  
 <220>  
 <221> misc\_feature  
 <222> (29)..(30)  
 <223> n es a, c, g o t  
 40  
 <220>  
 <221> misc\_feature  
 <222> (34)..(34)  
 <223> s es g o c  
 45  
 <220>  
 <221> misc\_feature  
 <222> (35)..(36)  
 <223> n es a, c, g o t  
 50  
 <220>  
 <221> misc\_feature  
 <222> (40)..(40)  
 <223> s es g o c  
 55  
 <220>  
 <221> misc\_feature  
 <222> (41)..(42)  
 <223> n es a, c, g o t  
 60  
 <220>  
 <221> misc\_feature  
 <222> (46)..(46)

ES 2 536 509 T3

<223> s e s g o c

<220>  
 <221> misc\_feature  
 <222> (47)..(48)  
 <223> n e s a, c, g o t

<400> 22

**taataactct ttccggggcat ctttctcsnn caacsngccs nngccsnct tgcggttgc 60**

**gtcgtgta 66**

<210> 23  
 <211> 64  
 <212> ADN  
 <213> Artificial

<220>  
 <223> Cebador SClib5

<400> 23

**ccatataaag ctttttctcg cgttccgcac gcgctaacat atctaataac tctttcgggg 60**

**catc 64**

<210> 24  
 <211> 39  
 <212> ADN  
 <213> Artificial

<220>  
 <223> Cebador SClink

<400> 24  
 gcggaacgcg agaaaaagct ttatatggcc tcgggggcc 39

<210> 25  
 <211> 45  
 <212> ADN  
 <213> Artificial

<220>  
 <223> Cebador T7B

<400> 25  
 ccgcacacca gtaaggtgtg cggttcagt tgccgcttc ttct 45

<210> 26  
 <211> 33  
 <212> ADN  
 <213> Artificial

<220>  
 <223> Cebador SCPhoAF

<400> 26  
 attaatggta ccgatccgt gaagtgaaa ttc 33

<210> 27  
 <211> 36  
 <212> ADN  
 <213> Artificial

ES 2 536 509 T3

<220>

<223> Cebador SCPHoAR

<400> 27

5 ataattgagc tctaagcttt ttttcacgct cgcac 36

<210> 28

<211> 64

<212> PRT

10 <213> *Sulfolobus* sp.

<400> 28

Met Ala Thr Val Lys Phe Lys Tyr Lys Gly Glu Glu Lys Glu Val Asp  
1 5 10 15

Ile Ser Lys Ile Lys Lys Val Trp Arg Val Gly Lys Met Ile Ser Phe  
20 25 30

Thr Tyr Asp Glu Gly Gly Gly Lys Thr Gly Arg Gly Ala Val Ser Glu  
35 40 45

Lys Asp Ala Pro Lys Glu Leu Leu Gln Met Leu Glu Lys Gln Lys Lys  
50 55 60

15

<210> 29

<211> 89

<212> PRT

20 <213> *Escherichia coli*

<400> 29

Met Lys Lys Met Phe Met Ala Val Leu Phe Ala Leu Val Ser Val Asn  
1 5 10 15

Ala Met Ala Ala Asp Cys Ala Lys Gly Lys Ile Glu Phe Ser Lys Tyr  
20 25 30

Asn Glu Asp Asp Thr Phe Thr Val Lys Val Asp Gly Lys Glu Tyr Trp  
35 40 45

Thr Ser Arg Trp Asn Leu Gln Pro Leu Leu Gln Ser Ala Gln Leu Thr  
50 55 60

Gly Met Thr Val Thr Ile Lys Ser Ser Thr Cys Glu Ser Gly Ser Gly  
65 70 75 80

Phe Ala Glu Val Gln Phe Asn Asn Asp  
85

25

<210> 30

<211> 128

<212> PRT

30 <213> *Homo sapiens*

<400> 30

ES 2 536 509 T3

Met Ser Leu Arg Leu Asp Thr Thr Pro Ser Cys Asn Ser Ala Arg Pro  
 1 5 10 15

Leu His Ala Leu Gln Val Leu Leu Leu Leu Ser Leu Leu Leu Thr Ala  
 20 25 30

Leu Ala Ser Ser Thr Lys Gly Gln Thr Lys Arg Asn Leu Ala Lys Gly  
 35 40 45

Lys Glu Glu Ser Leu Asp Ser Asp Leu Tyr Ala Glu Leu Arg Cys Met  
 50 55 60

Cys Ile Lys Thr Thr Ser Gly Ile His Pro Lys Asn Ile Gln Ser Leu  
 65 70 75 80

Glu Val Ile Gly Lys Gly Thr His Cys Asn Gln Val Glu Val Ile Ala  
 85 90 95

Thr Leu Lys Asp Gly Arg Lys Ile Cys Leu Asp Pro Asp Ala Pro Arg  
 100 105 110

Ile Lys Lys Ile Val Gln Lys Lys Leu Ala Gly Asp Glu Ser Ala Asp  
 115 120 125

5 <210> 31  
 <211> 142  
 <212> PRT  
 <213> *Azotobacter vinelandii*

10 <400> 31



ES 2 536 509 T3

Met Lys Ile Ser Ala Arg Asn Val Phe Lys Gly Thr Val Ser Ala Leu  
 1 5 10 15

Lys Glu Gly Ala Val Asn Ala Glu Val Asp Ile Leu Leu Gly Gly Gly  
 20 25 30

Asp Lys Leu Ala Ala Val Val Thr Leu Glu Ser Ala Arg Ser Leu Gln  
 35 40 45

Leu Ala Ala Gly Lys Glu Val Val Ala Val Val Lys Ala Pro Trp Val  
 50 55 60

Leu Leu Met Thr Asp Ser Ser Gly Tyr Arg Leu Ser Ala Arg Asn Ile  
 65 70 75 80

Leu Thr Gly Thr Val Lys Thr Ile Glu Thr Gly Ala Val Asn Ala Glu  
 85 90 95

Val Thr Leu Ala Leu Gln Gly Gly Thr Glu Ile Thr Ser Met Val Thr  
 100 105 110

Lys Glu Ala Val Ala Glu Leu Gly Leu Lys Pro Gly Ala Ser Ala Ser  
 115 120 125

Ala Val Ile Lys Ala Ser Asn Val Ile Leu Gly Val Pro Ala  
 130 135 140

5 <210> 32  
 <211> 14  
 <212> PRT  
 <213> Artificial

10 <220>  
 <223> Enlazador peptídico  
 <400> 32

Lys Leu Gly Ser Ala Gly Ser Ala Ala Gly Ser Gly Glu Phe  
 1 5 10

15 <210> 33  
 <211> 65  
 <212> PRT  
 <213> *Sulfolobus* sp.

20 <400> 33

Val Lys Val Lys Phe Lys Tyr Lys Gly Glu Glu Lys Glu Val Asp Thr  
 1 5 10 15

Ser Lys Ile Lys Lys Val Trp Arg Val Gly Lys Met Val Ser Phe Thr  
 20 25 30

Tyr Asp Asp Asn Gly Lys Thr Gly Arg Gly Ala Val Ser Glu Lys Asp  
 35 40 45

Ala Pro Lys Glu Leu Leu Asp Met Leu Ala Arg Ala Glu Arg Glu Lys  
 50 55 60

Lys  
 65

<210> 34  
 <211> 63  
 <212> PRT  
 <213> *Sulfolobus acidocaldarius*

5

<400> 34

Lys Val Arg Phe Lys Tyr Lys Gly Glu Glu Lys Glu Val Asp Thr Ser  
 1 5 10 15

Lys Ile Lys Lys Val Trp Arg Val Gly Lys Met Val Ser Phe Thr Tyr  
 20 25 30

Asp Asp Asn Gly Lys Thr Gly Arg Gly Ala Val Ser Glu Lys Asp Ala  
 35 40 45

Pro Lys Glu Leu Met Asp Met Leu Ala Arg Ala Glu Lys Lys Lys  
 50 55 60

10

<210> 35  
 <211> 63  
 <212> PRT  
 <213> *Sulfolobus tokodaii*

15

<400> 35

Val Thr Val Lys Phe Lys Tyr Lys Gly Glu Glu Lys Glu Val Asp Ile  
 1 5 10 15

Ser Lys Ile Lys Lys Val Trp Arg Val Gly Lys Met Ile Ser Phe Thr  
 20 25 30

Tyr Asp Asp Asn Gly Lys Thr Gly Arg Gly Ala Val Ser Glu Lys Asp  
 35 40 45

Ala Pro Lys Glu Leu Leu Gln Met Leu Glu Lys Ser Gly Lys Lys  
 50 55 60

20

ES 2 536 509 T3

<210> 36  
 <211> 62  
 <212> PRT  
 <213> *Sulfolobus shibatae*

5 <400> 36

Val Thr Val Lys Phe Lys Tyr Lys Gly Glu Glu Lys Glu Val Asp Thr  
 1 5 10 15  
 Ser Lys Ile Lys Lys Val Trp Arg Val Gly Lys Met Ile Ser Phe Thr  
 20 25 30  
 Tyr Asp Glu Gly Gly Lys Thr Gly Arg Gly Ala Val Ser Glu Lys Asp  
 35 40 45  
 Ala Pro Lys Glu Leu Leu Gln Met Leu Glu Lys Gln Lys Lys  
 50 55 60

10 <210> 37  
 <211> 60  
 <212> PRT  
 <213> *Sulfolobus solfataricus*

15 <400> 37

Val Lys Phe Lys Tyr Lys Gly Glu Glu Lys Glu Val Asp Ile Ser Lys  
 1 5 10 15  
 Ile Lys Lys Val Trp Arg Val Gly Lys Met Ile Ser Phe Thr Tyr Asp  
 20 25 30  
 Glu Gly Gly Lys Thr Gly Arg Gly Ala Val Ser Glu Lys Asp Ala Pro  
 35 40 45  
 Lys Glu Leu Leu Gln Met Leu Glu Lys Gln Lys Lys  
 50 55 60

20 <210> 38  
 <211> 60  
 <212> PRT  
 <213> *Sulfolobus shibatae*

25 <400> 38

ES 2 536 509 T3

Val Lys Phe Lys Tyr Lys Gly Glu Glu Lys Gln Val Asp Ile Ser Lys  
 1 5 10 15

Ile Lys Lys Val Trp Arg Val Gly Lys Met Ile Ser Phe Thr Tyr Asp  
 20 25 30

Glu Gly Gly Lys Thr Gly Arg Gly Ala Val Ser Glu Lys Asp Ala Pro  
 35 40 45

Lys Glu Leu Leu Gln Met Leu Glu Lys Gln Lys Lys  
 50 55 60

5 <210> 39  
 <211> 65  
 <212> PRT  
 <213> Artificial

10 <220>  
 <223> Aglutinante

<400> 39

Val Lys Val Lys Phe Val Leu Gly Gly Glu Glu Lys Glu Val Asp Thr  
 1 5 10 15

Ser Lys Ile Arg His Val Tyr Arg Tyr Gly Lys His Val Thr Phe Ser  
 20 25 30

Tyr Asp Asp Asn Gly Lys Leu Gly Leu Gly Leu Val Lys Glu Lys Asp  
 35 40 45

Ala Pro Lys Glu Leu Leu Asp Met Leu Ala Arg Ala Glu Arg Glu Lys  
 50 55 60

Lys  
 65

15 <210> 40  
 <211> 65  
 <212> PRT  
 <213> Artificial

20 <220>  
 <223> Aglutinante

<400> 40

ES 2 536 509 T3

Val Lys Val Lys Phe Trp Gly Ser Gly Glu Glu Lys Glu Val Asp Thr  
1 5 10 15

Ser Lys Ile Arg Ser Val Gly Arg Met Gly Lys Leu Val Ala Phe Arg  
20 25 30

Tyr Asp Asp Asn Gly Lys Asn Gly Tyr Gly Val Val Asp Glu Lys Asp  
35 40 45

Ala Pro Lys Glu Leu Leu Asp Met Leu Ala Arg Ala Glu Arg Glu Lys  
50 55 60

Lys  
65

5 <210> 41  
<211> 65  
<212> PRT  
<213> Artificial

10 <220>  
<223> Aglutinante

<400> 41

Val Lys Val Lys Phe Thr Val Leu Gly Glu Glu Lys Glu Val Asp Thr  
1 5 10 15

Ser Lys Ile Leu Arg Val Ala Arg Tyr Gly Lys Ala Val Cys Phe Arg  
20 25 30

Tyr Asp Asp Asn Gly Lys Ile Gly Tyr Gly Leu Val Ile Glu Lys Asp  
35 40 45

Ala Pro Lys Glu Leu Leu Asp Met Leu Ala Arg Ala Glu Arg Glu Lys  
50 55 60

Lys  
65

15 <210> 42  
<211> 65  
<212> PRT  
<213> Artificial

20 <220>  
<223> Aglutinante

<400> 42

ES 2 536 509 T3

Val Lys Val Lys Phe Thr Trp Tyr Gly Glu Glu Lys Glu Val Asp Thr  
 1 5 10 15

Ser Lys Ile Met Ser Val Cys Arg Trp Gly Lys Arg Val Ala Phe Ser  
 20 25 30

Tyr Asp Asp Asn Gly Lys Ile Asp Tyr Gly Leu Val Asp Glu Lys Asp  
 35 40 45

Ala Pro Lys Glu Leu Leu Asp Met Leu Ala Arg Ala Glu Arg Glu Lys  
 50 55 60

Lys  
 65

5 <210> 43  
 <211> 61  
 <212> PRT  
 <213> Artificial

10 <220>  
 <223> Aglutinante

<400> 43

Val Lys Val Lys Phe Cys Trp Glu Gly Glu Glu Lys Glu Val Asp Thr  
 1 5 10 15

Ser Lys Ile Arg His Val Tyr Arg Tyr Gly Lys Tyr Val Thr Phe Ser  
 20 25 30

Tyr Asp Asp Asn Gly Lys Leu Gly Leu Gly Leu Val Asn Glu Lys Asp  
 35 40 45

Ala Pro Lys Glu Leu Leu Asp Met Leu Ala Arg Ala Ala  
 50 55 60

15 <210> 44  
 <211> 65  
 <212> PRT  
 <213> Artificial

20 <220>  
 <223> Aglutinante

<400> 44

ES 2 536 509 T3

Val Lys Val Lys Phe Gly Glu His Gly Glu Glu Lys Glu Val Asp Thr  
 1 5 10 15

Ser Lys Ile Met Ser Val Val Arg Phe Gly Lys Trp Val Cys Phe Arg  
 20 25 30

Tyr Asp Asp Asn Gly Lys Ala Gly Tyr Gly Ala Val Leu Glu Lys Asp  
 35 40 45

Ala Pro Lys Glu Leu Leu Asp Met Leu Ala Arg Ala Glu Arg Glu Lys  
 50 55 60

Lys  
 65

5 <210> 45  
 <211> 65  
 <212> PRT  
 <213> Artificial

10 <220>  
 <223> Aglutinante  
 <400> 45

Val Lys Val Lys Phe Val Met Ser Gly Glu Glu Lys Glu Val Asp Thr  
 1 5 10 15

Ser Lys Ile Met Ser Val Val Arg Phe Gly Lys Trp Val Cys Phe Arg  
 20 25 30

Tyr Asp Asp Asn Gly Lys Ala Gly Tyr Gly Ala Val Leu Glu Lys Asp  
 35 40 45

Ala Pro Lys Glu Leu Leu Asp Met Leu Ala Arg Ala Glu Arg Glu Lys  
 50 55 60

Lys  
 65

15 <210> 46  
 <211> 65  
 <212> PRT  
 <213> Artificial

20 <220>  
 <223> Aglutinante  
 <400> 46

ES 2 536 509 T3

Val Lys Val Lys Phe Glu Ile Ser Gly Glu Glu Lys Glu Val Asp Thr  
1 5 10 15

Ser Lys Ile Ser Ser Val Ala Arg Leu Gly Lys Leu Val Ala Phe Arg  
20 25 30

Tyr Asp Asp Asn Gly Lys Thr Gly Tyr Gly Val Val Pro Glu Lys Asp  
35 40 45

Ala Pro Lys Glu Leu Leu Asp Met Leu Ala Arg Ala Glu Arg Glu Lys  
50 55 60

Lys  
65

<210> 47  
<211> 68  
<212> PRT  
<213> Artificial

5

<220>  
<223> D5\_GarA

10

<400> 47

Gly Ser Val Lys Val Lys Phe Leu Tyr Leu Gly Glu Glu Lys Glu Val  
1 5 10 15

Asp Thr Ser Lys Ile Trp Phe Val Met Arg Ala Gly Lys His Val Tyr  
20 25 30

Phe Gln Tyr Asp Asp Asn Gly Lys Tyr Gly Ile Gly Trp Val Arg Glu  
35 40 45

Lys Asp Ala Pro Lys Glu Leu Leu Asp Met Leu Ala Arg Ala Glu Arg  
50 55 60

Glu Lys Lys Leu  
65

15



**REIVINDICACIONES**

- 5 1. Una biblioteca combinatoria de variantes de una proteína OB-fold, en la que de 5 a 32, preferiblemente de 8 a 20, y más preferiblemente de 11 a 16 restos de la interfaz de unión de una proteína OB-fold con su ligando nativo se han sustituido aleatoriamente, en la que dicha proteína OB-fold es Sac7d o Sac7e de *Sulfolobus acidocaldarius*, o Sso7d, de *Sulfolobus solfataricus*, o DBP 7 de *Sulfolobus tokodaii*, o Ssh7b de *Sulfolobus shibatae*, o Ssh7a de *Sulfolobus shibatae*, o p7ss de *Sulfolobus solfataricus*, en la que dichos restos que se aleatorizan en dichas variantes se seleccionan entre los restos correspondientes a V2, K3, K5, K7, Y8, K9, G10, E14, T17, K21, K22, W24, V26, G27, K28, M29, S31, T33, D36, N37, G38, K39, T40, R42, A44, S46, E47, K48, D49, A50 y P51 de Sac7d.
- 10 2. La biblioteca combinatoria de la reivindicación 1, en la que dichas variantes comprenden adicionalmente una inserción de 1 a 15 restos aminoacídicos aleatorios en el bucle 3 y/o una inserción de 1 a 15 restos aminoacídicos aleatorios en el bucle 4, y/o una inserción de 1 a 20 restos aminoacídicos aleatorios en el bucle 1.
- 15 3. La biblioteca combinatoria de la reivindicación 1 o 2, en la que 11, 12, 13, 14, 15, 16, 17 o 18 restos se aleatorizan, elegidos entre los restos correspondientes a K7, Y8, K9, K21, K22, W24, V26, K28, M29, S31, T33, K39, T40, R42, A44, S46, E47 y K48 de Sac7d.
- 20 4. La biblioteca combinatoria de la reivindicación 3, en la que 11, 12, 13, 14, 15 o 16 restos se aleatorizan, elegidos entre los restos correspondientes a K7, Y8, K9, K21, K22, W24, V26, K28, M29, S31, T33, K39, T40, R42, A44 y S46 de Sac7d.
- 25 5. La biblioteca combinatoria de la reivindicación 4, en la que los restos aleatorizados en dichas variantes comprenden al menos los restos correspondientes a K7, Y8, K9, W24, V26, M29, S31, T33, R42, A44 y S46 de Sac7d.
- 30 6. La biblioteca combinatoria de la reivindicación 5, en la que solamente los restos correspondientes a K7, Y8, K9, W24, V26, M29, S31, T33, R42, A44 y S46 de Sac7d se aleatorizan en dichas variantes.
- 35 7. La biblioteca combinatoria de la reivindicación 6, en la que uno, dos o tres restos adicionales seleccionados entre los correspondientes a K21, K22 y T40 de Sac7d se aleatorizan también en dichas variantes.
- 40 8. La biblioteca combinatoria de la reivindicación 7, en la que los restos correspondientes a K7, Y8, K9, K21, K22, W24, V26, M29, S31, T33, R42, A44 y S46 de Sac7d se aleatorizan en dichas variantes.
- 45 9. La biblioteca combinatoria de cualquiera de las reivindicaciones 1 a 8, en la que de 1 a 15 restos aminoacídicos aleatorios se insertan en la región correspondiente a los aminoácidos en la posición 25 a 30 de Sac7d, preferiblemente entre G27 y K28, y/o de 1 a 15 restos aminoacídicos aleatorios se insertan en la región correspondiente a los aminoácidos en la posición 35 a 40 de Sac7d, preferiblemente entre N37 y G38, y/o de 1 a 20 restos aminoacídicos aleatorios se insertan en la región correspondiente a los aminoácidos en la posición 7 a 12 de Sac7d, preferiblemente entre K9 y G10 en dichas variantes.
- 50 10. La biblioteca combinatoria de cualquiera de las reivindicaciones 1 a 8, en la que de 1 a 4 restos aminoacídicos seleccionados entre los restos correspondientes a A59, R60, A61 y E64 de Sac7d se delecionan en dichas variantes.
- 55 11. Un proceso para obtener una molécula que se une específicamente a una diana de interés, que comprende una etapa de selección y aislamiento, en una biblioteca combinatoria de acuerdo con cualquiera de las reivindicaciones 1 a 10, de una o varias variantes de dicha proteína OB-fold que se une a dicha diana de interés.
- 60 12. El proceso de la reivindicación 11, en el que dicha diana es un péptido, una proteína, un oligosacárido, un ADN monocatenario, un ADN bicatenario, un ARN, un ión metálico, o un carbohidrato.
- 65 13. El proceso de la reivindicación 11 o 12, en el que dicha etapa de selección se realiza por presentación en ribosoma.
14. El proceso de la reivindicación 13, en el que se realizan 2, 3, 4 o 5 rondas de selección por presentación en ribosoma.
15. El proceso de cualquiera de las reivindicaciones 11 a 14, que comprende adicionalmente una etapa de caracterización de una molécula que se une específicamente a dicha diana.
16. El proceso de la reivindicación 15, en el que dicha etapa de aislamiento y caracterización comprende las siguientes etapas:

- (i) se transcriben de forma inversa los ARNm de una combinación seleccionada por presentación en ribosoma y se amplifican por PCR, y los ADN resultantes se clonan en vectores de expresión;
- (ii) se transforman bacterias por dichos vectores de expresión y se cultivan clones individuales; y
- (iii) se ensayan las proteínas extraídas de dichos clones bacterianos para su unión a la diana.

- 5
17. Una proteína, caracterizada porque se une específicamente a GarA, y porque su secuencia es la SEC ID N° 47.
18. Un complejo que consiste en una variante de una proteína OB-fold de partida en el que de 5 a 32 restos en la interfaz de unión de dicha proteína OB-fold de partida con su ligando nativo se han sustituido, en el que dicha variante se une a una diana a la que no se une la proteína OB-fold de partida, en el que dicha proteína OB-fold de partida se elige entre Sac7d o Sac7e de *Sulfolobus acidocaldarius*, Sso7d, de *Sulfolobus solfataricus*, DBP 7 de *Sulfolobus tokodaii*, Ssh7b de *Sulfolobus shibatae*, Ssh7a de *Sulfolobus shibatae*, y p7ss de *Sulfolobus solfataricus*, en el que dichos restos sustituidos corresponden a los restos de Sac7d seleccionados entre V2, K3, K5, K7, Y8, K9, G10, E14, T17, K21, K22, W24, V26, G27, K28, M29, S31, T33, D36, N37, G38, K39, T40, R42, A44, S46, E47, K48, D49, A50 y P51.
19. El complejo de la reivindicación 18, en el que dicha variante comprende adicionalmente la inserción de 1 a 15 restos aminoacídicos aleatorios en la región correspondiente a los aminoácidos en la posición 25 a 30 de Sac7d, preferiblemente entre G27 y K28, y/o al inserción de 1 a 15 restos aminoacídicos aleatorios en la región correspondiente a los aminoácidos en la posición 35 a 40 de Sac7d, preferiblemente entre N37 y G38, y/o la inserción de 1 a 20 restos aminoacídicos aleatorios en la región correspondiente a los aminoácidos en la posición 7 a 12 de Sac7d, preferiblemente entre K9 y G10.
20. El complejo de cualquiera de las reivindicaciones 18 o 19, en el que 11, 12, 13, 14, 15, 16, 17 o 18 restos seleccionados entre los restos correspondientes a K7, Y8, K9, K21, K22, W24, V26, K28, M29, S31, T33, K39, T40, R42, A44, S46, E47 y K48 de Sac7d están sustituidos en dicha variante respecto a la proteína OB-fold de partida.
- 10
- 15
- 20
- 25

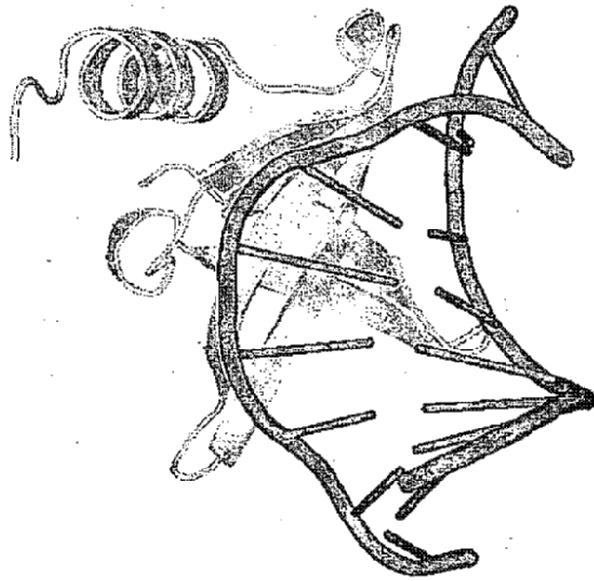


Figura 1a

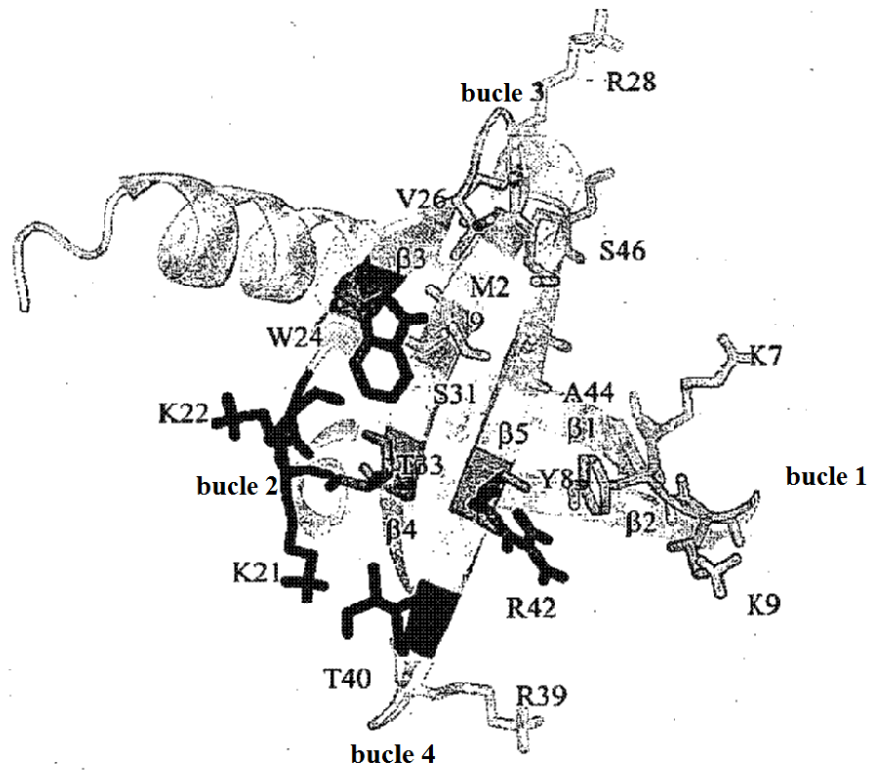


Figura 1b

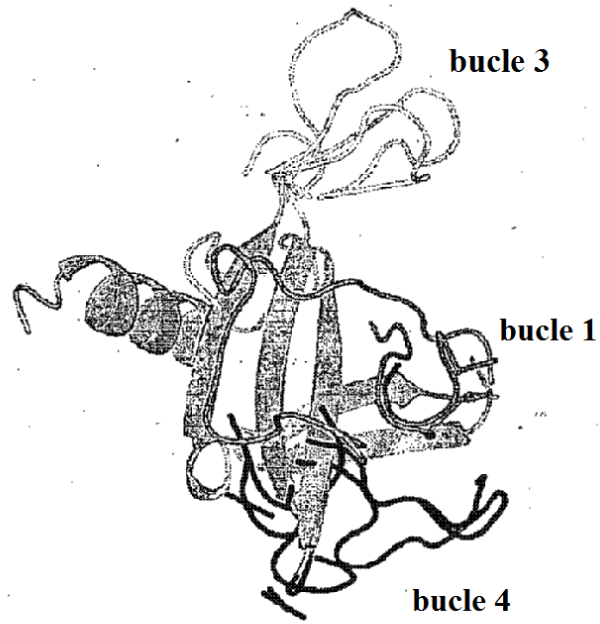


Figura 2

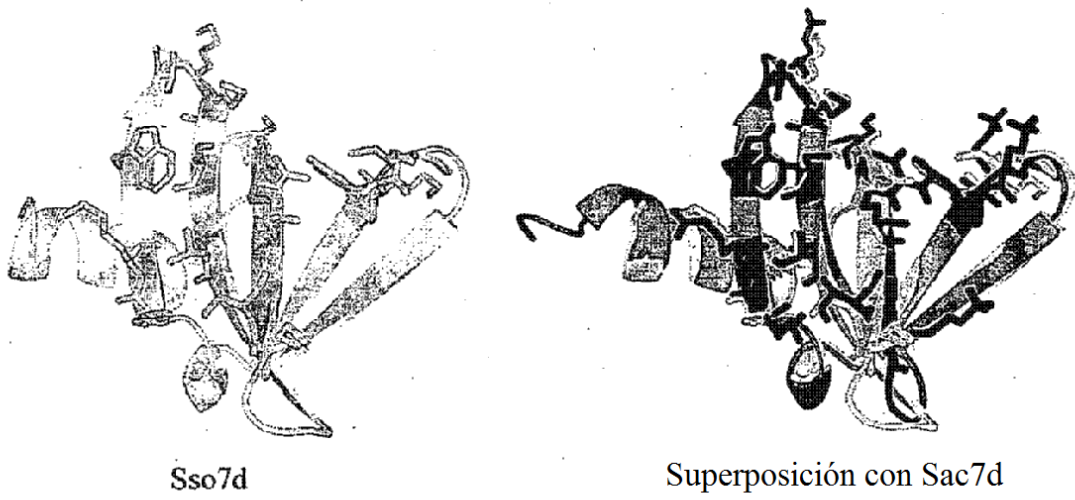
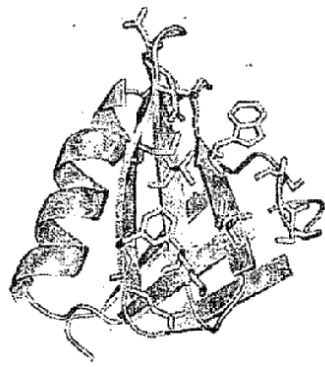


Figura 3a



Toxina IIe tipo Shiga



Superposición con Sac7d

Figura 3b

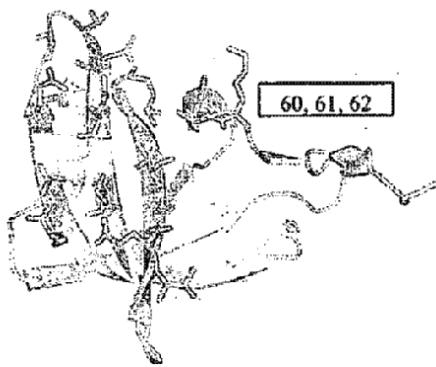


NAP-2

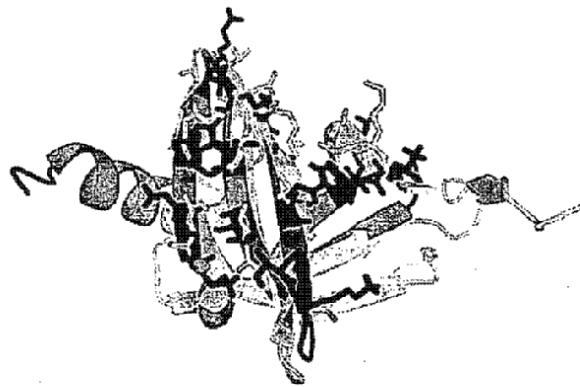


Superposición con Sac7d

Figura 3c



modg



Superposición con Sac7d

Figura 3d

Sac7d	VKVKFKYKGEKEVDTSKIRKVKVRVGRKMSFTYDDNCKTGRGAVSEKDAPKELLDMLARAERE--KK
Sac7e	--R.....M.....K--
DBP 7	.T.....I.....I.....Q....KSG..
Ssh7b	.T.....I.....I.....EG....Q....KQ--
Sso7d	--.....I.....I.....EG....Q....KQ--
Ssh7a	--.....Q.I.....I.....EG....Q....KQ--

Figura 4

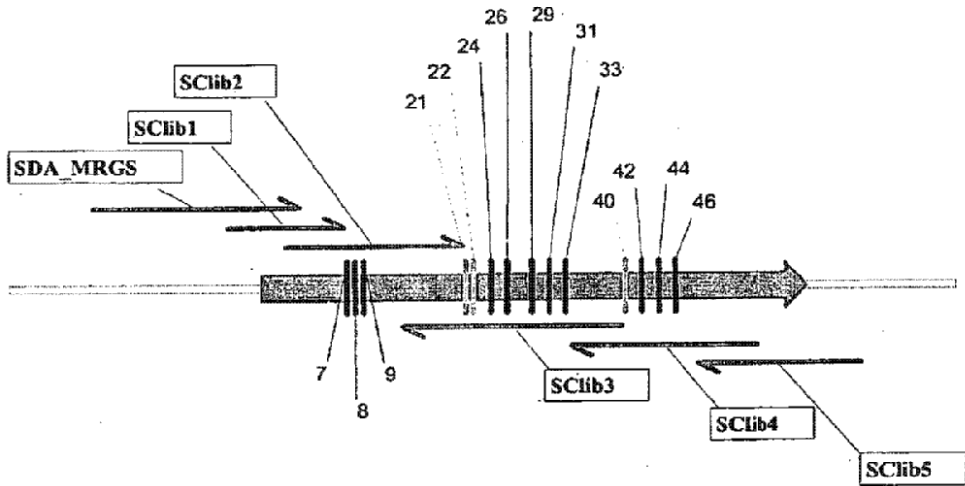


Figura 5

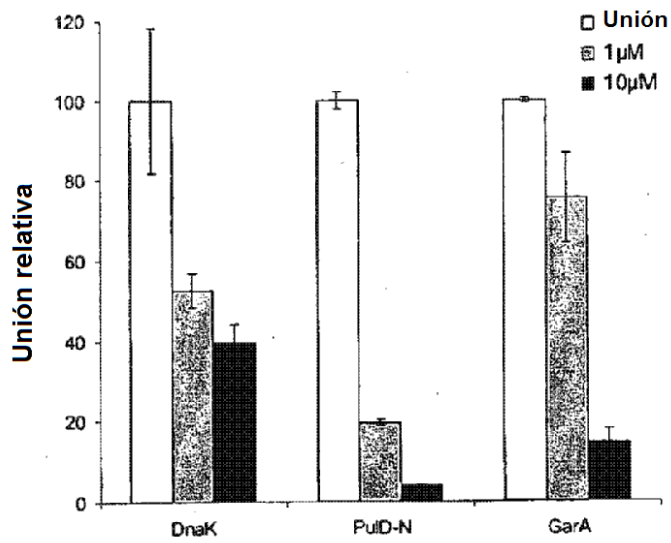


Figura 6a

Selección de agentes de unión anti-PuID-N después de ronda 4

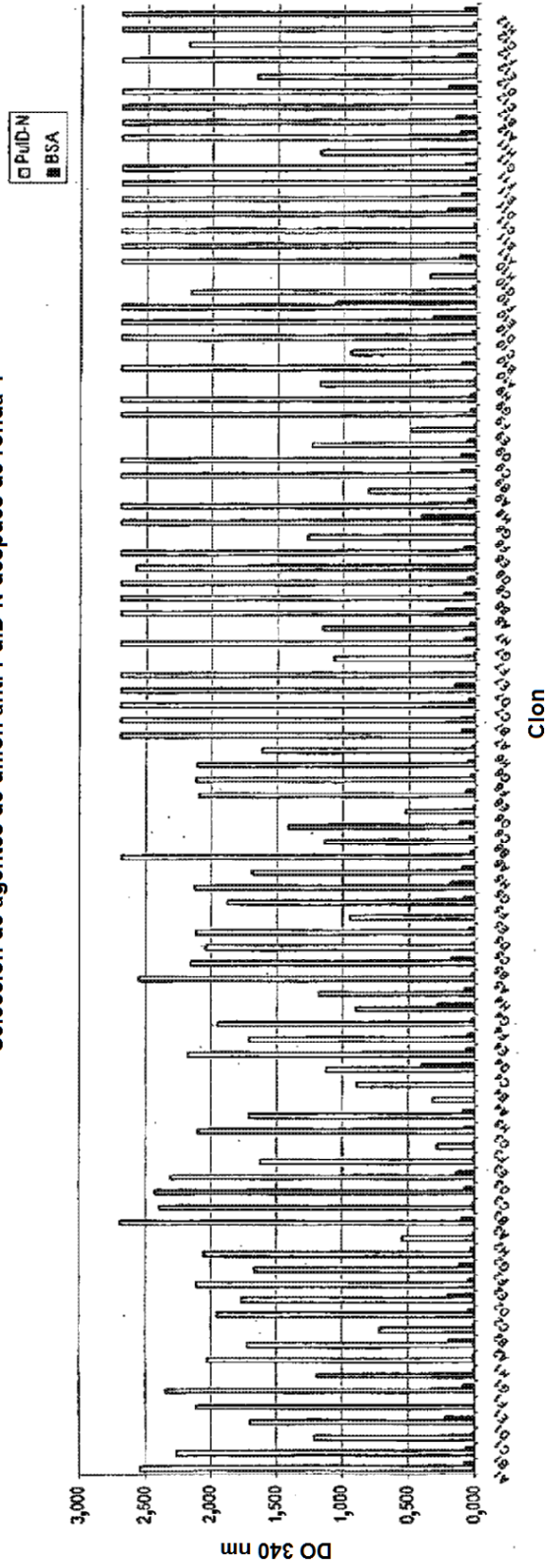


Figura 6b

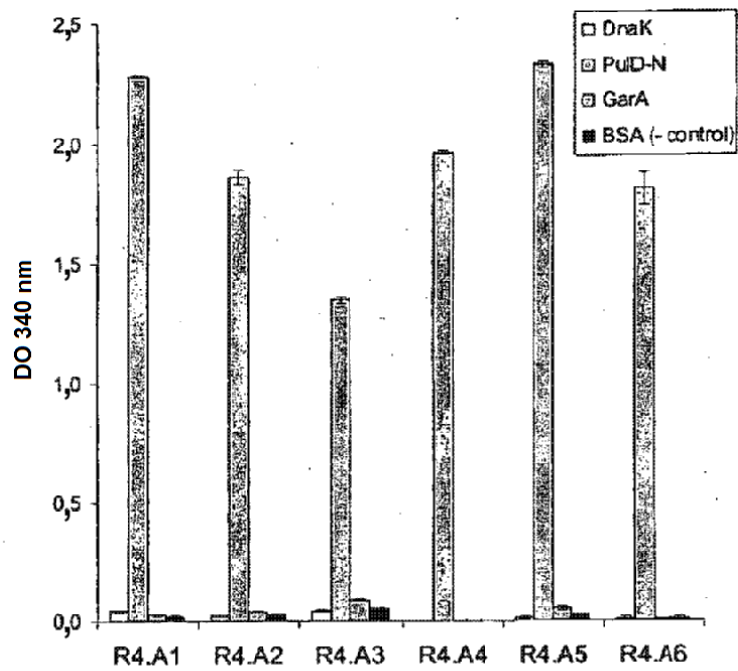


Figura 6c

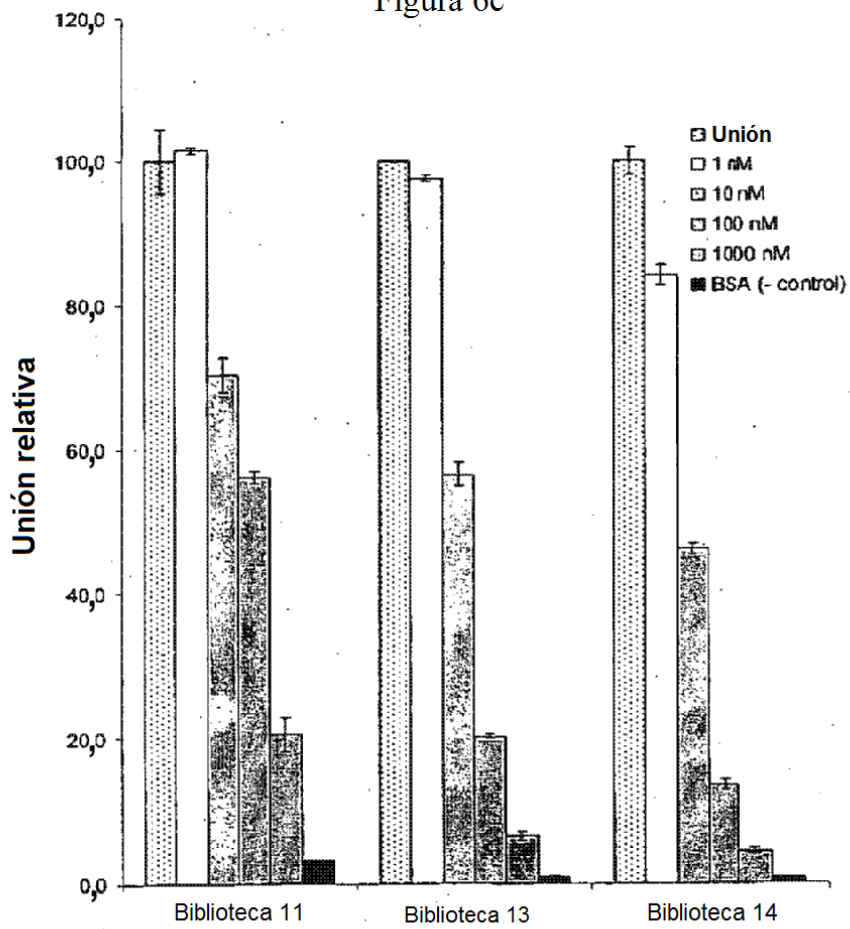


Figura 7





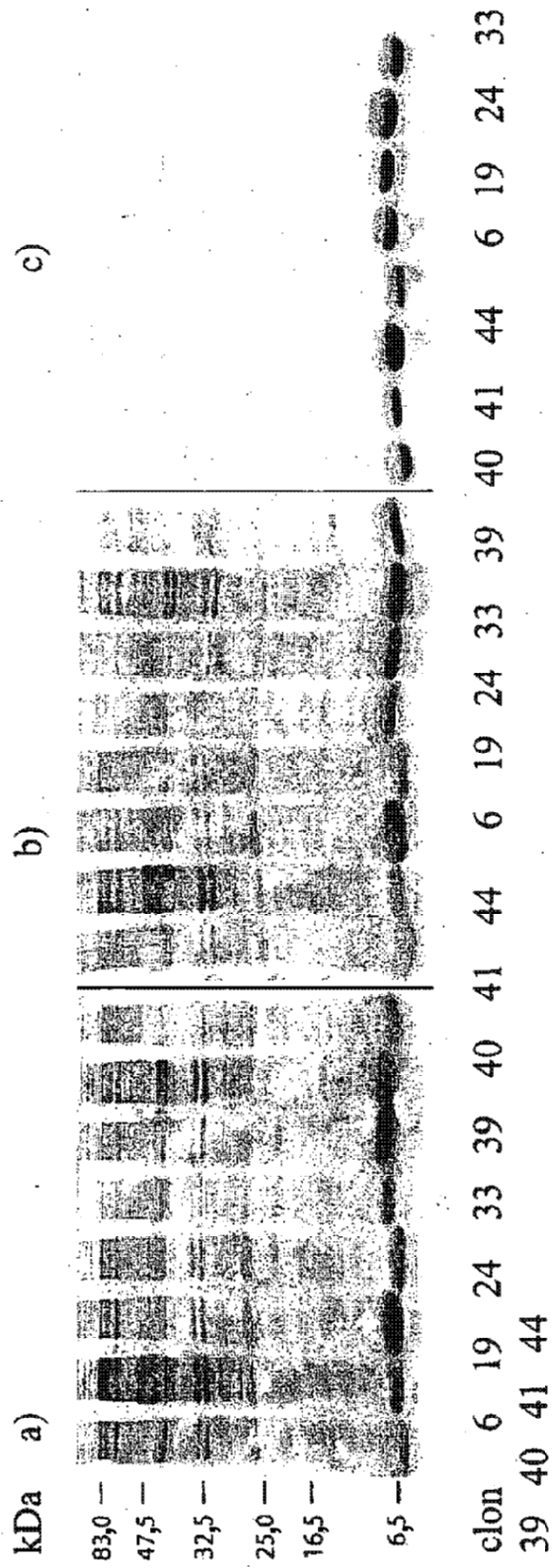


Figura 9

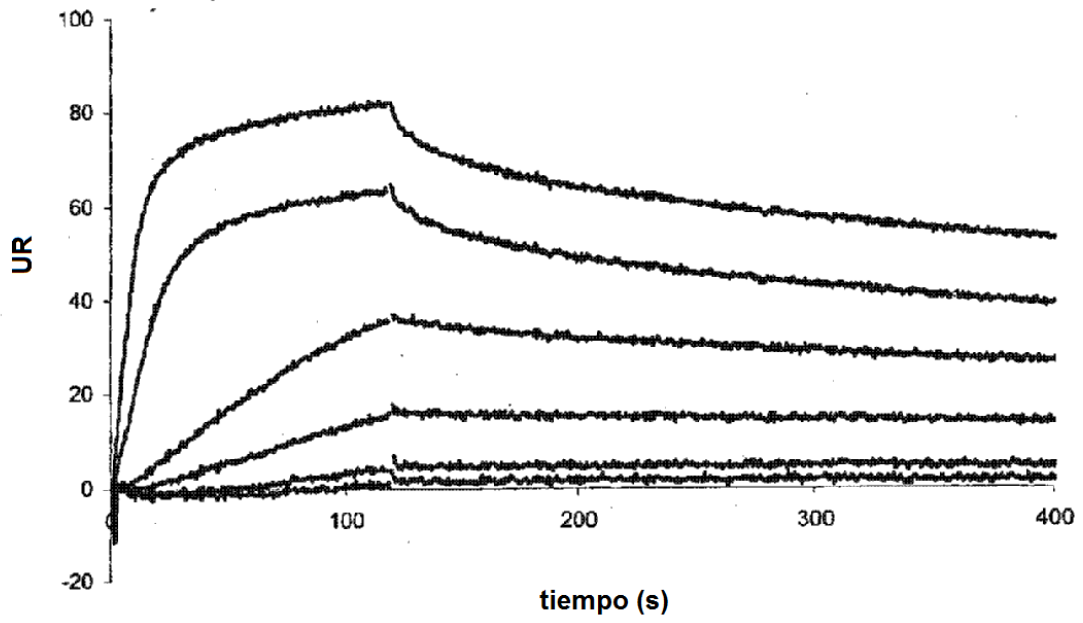


Figura 10a

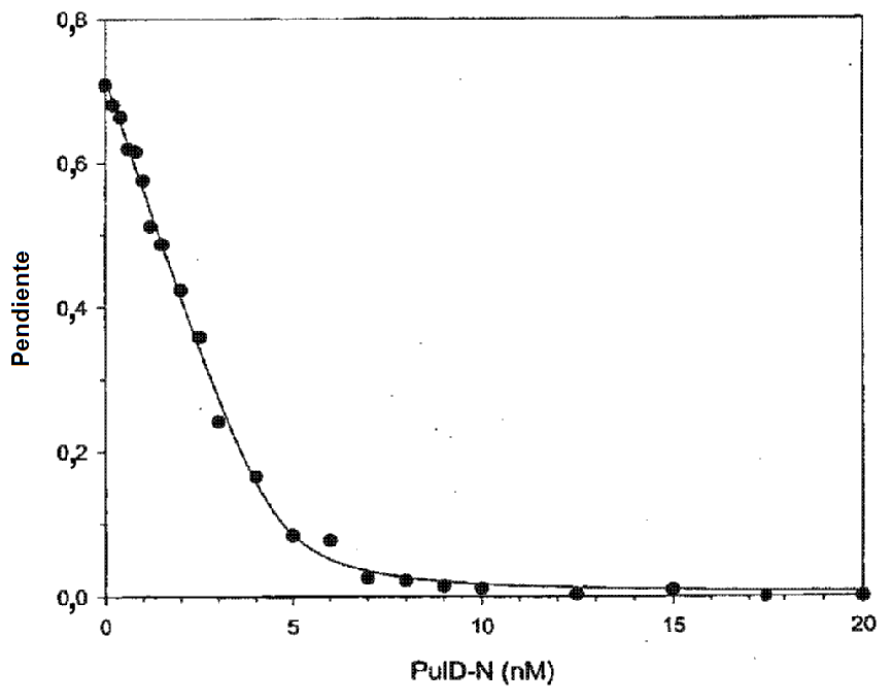


Figura 10b

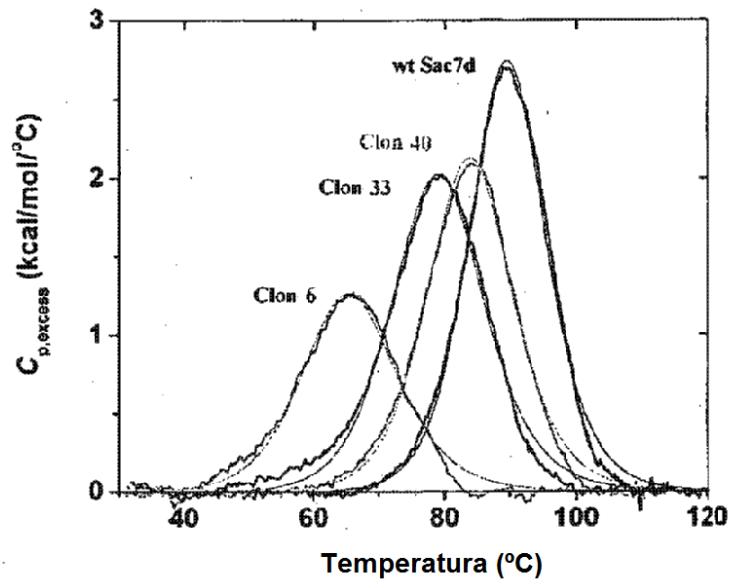


Figura 11

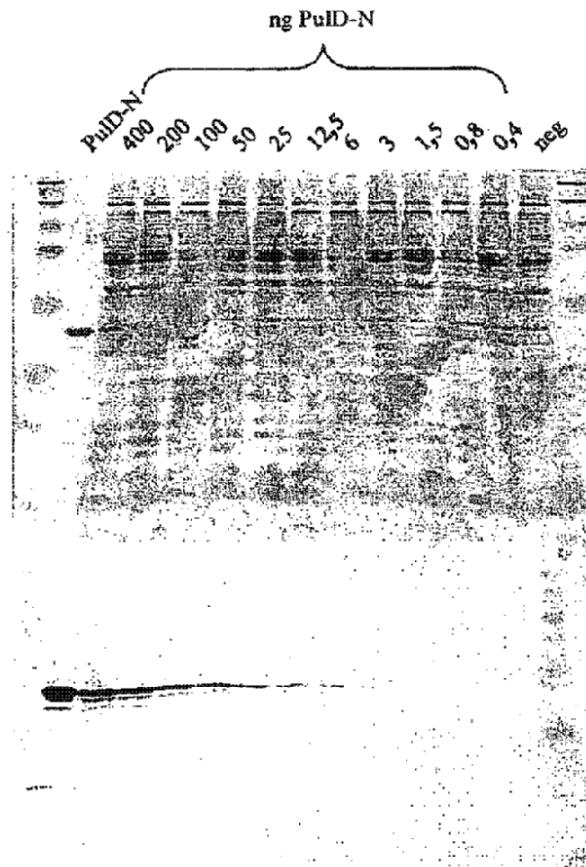


Figura 12a

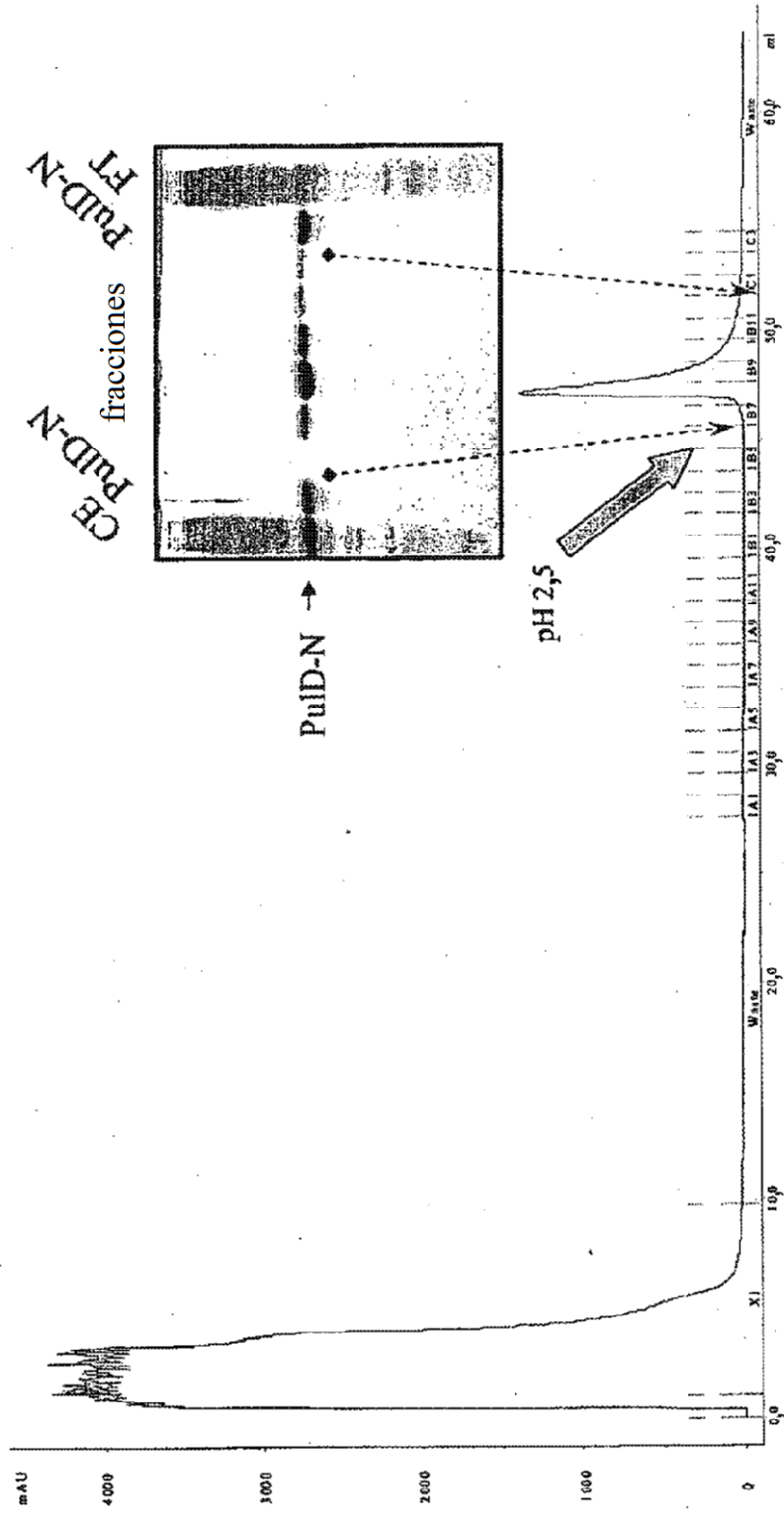


Figura 12b

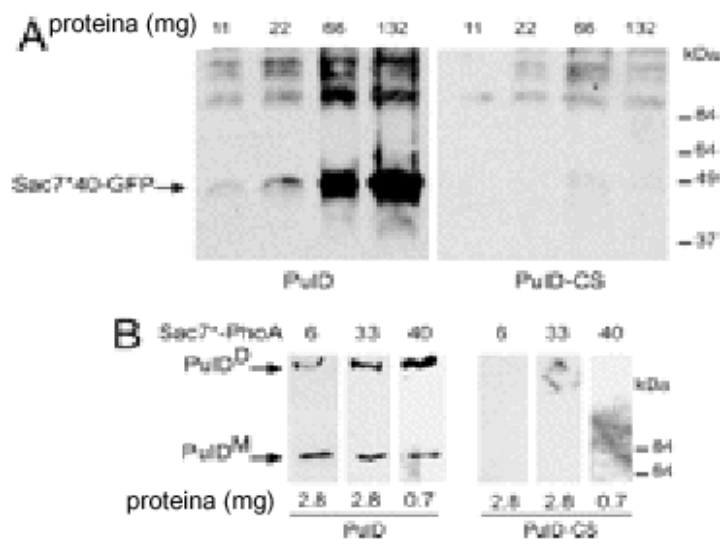


Fig 13

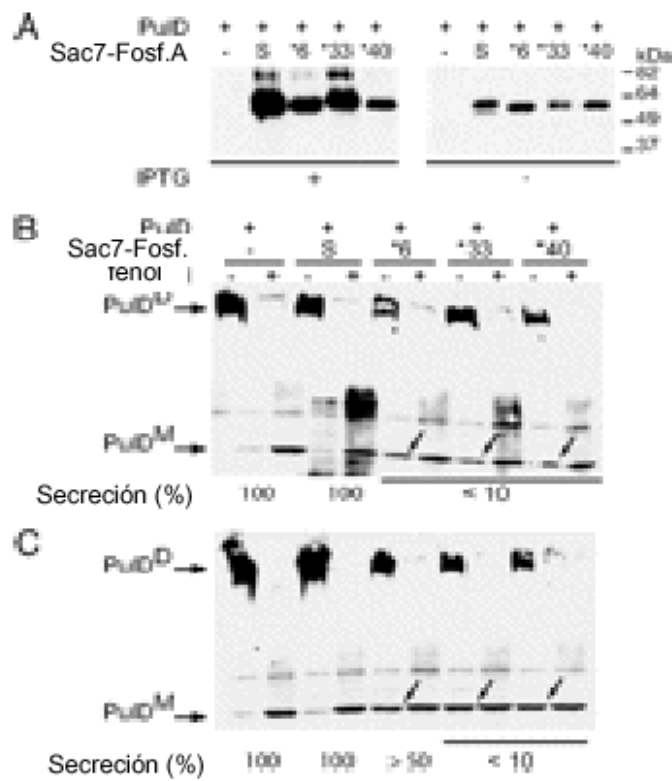


Fig 14

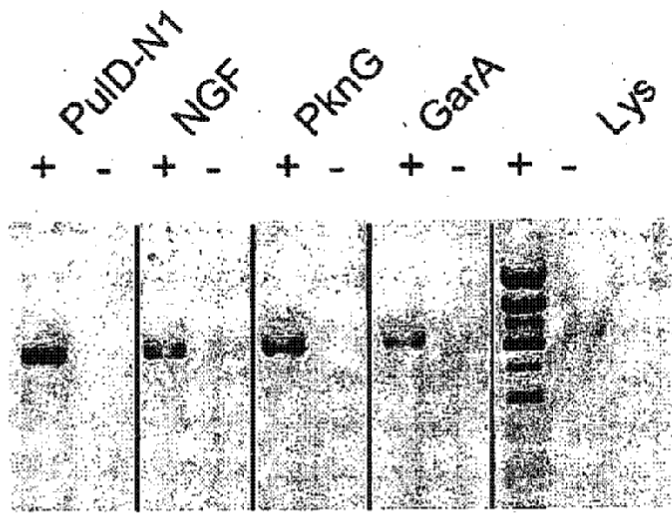


Figura 15

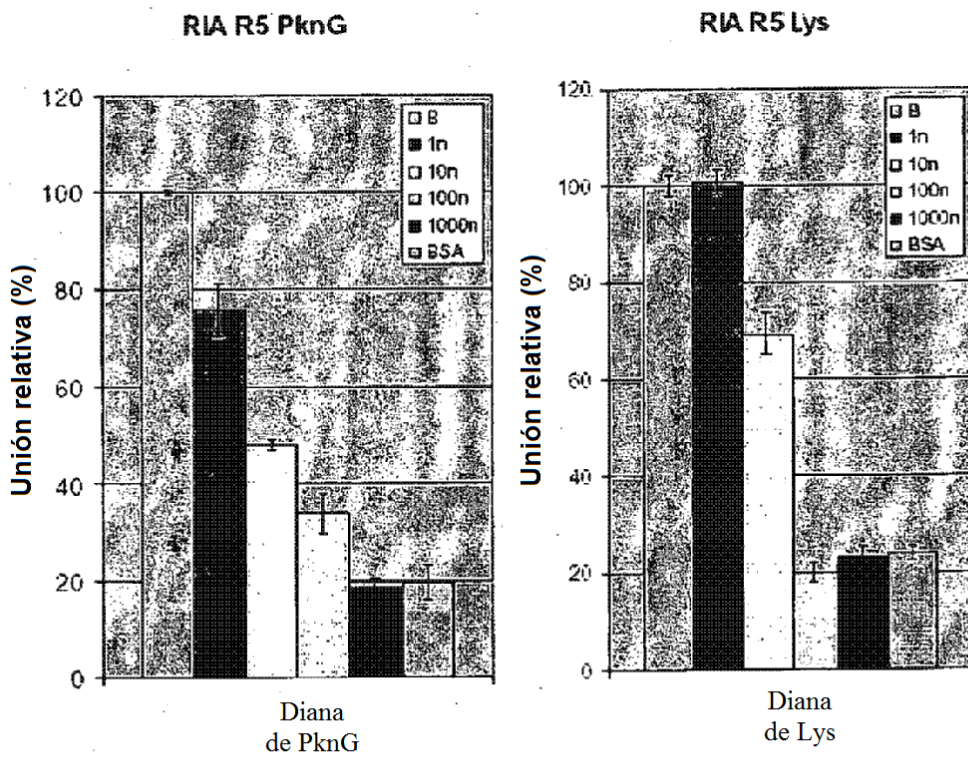


Figura 16

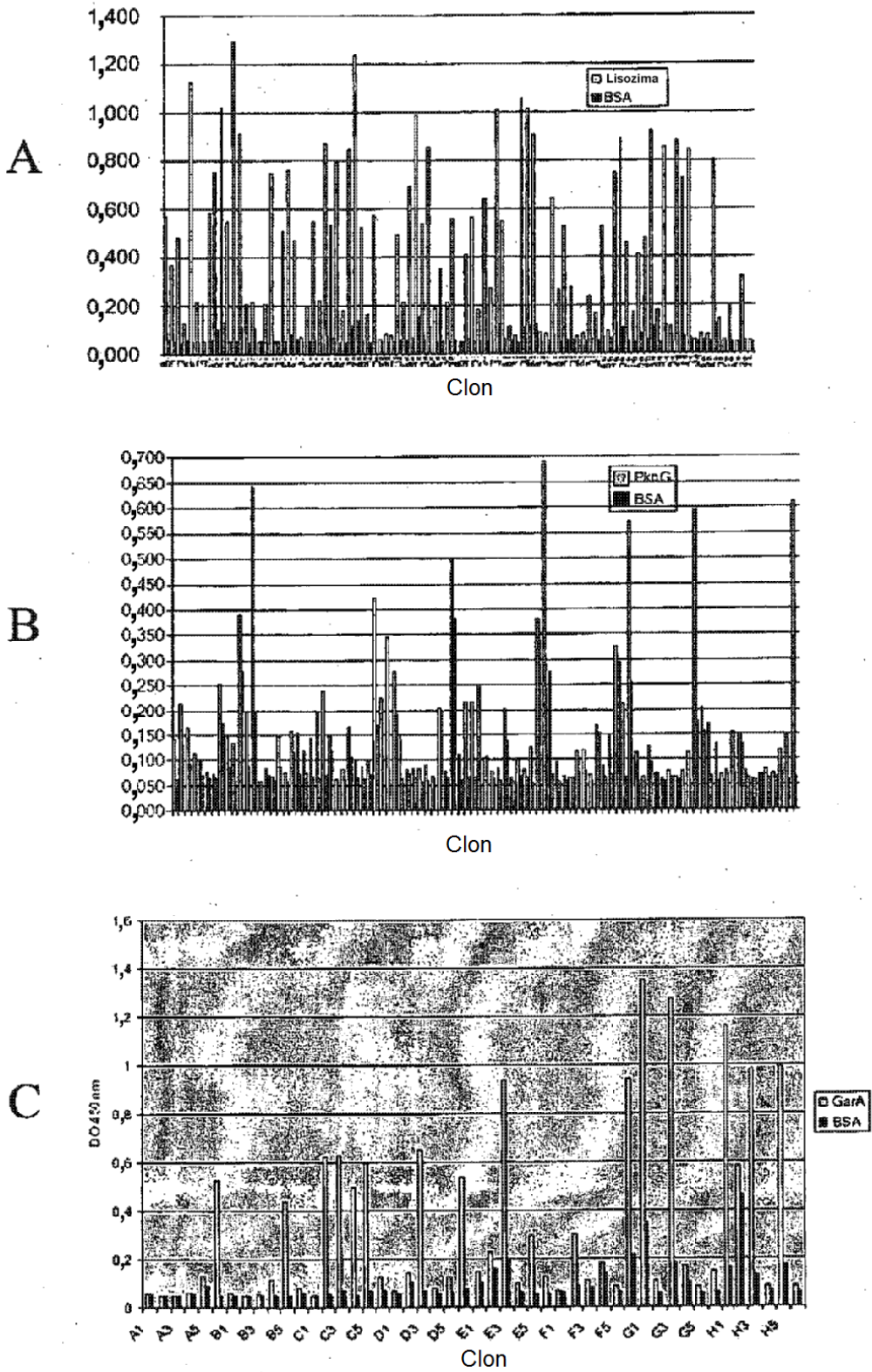


Figura 17

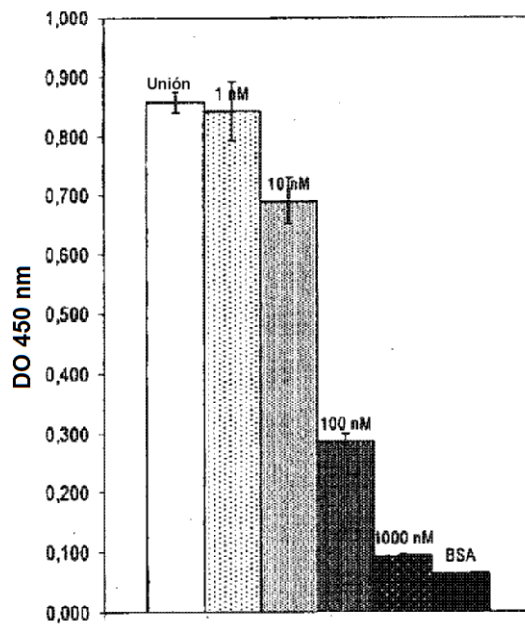


	1	10	20	30	40	50	60	68	
Sac_BK_1	GSVKVKFXXXGEEKEVDTSKIDGVXRXGKQVMEKYDDNGKKGXGVXKEKDPKELLQHLARREREKKL								
Lys_B10_1	GVD.....VM.K.A.C.L.I.....N.Y.D.T								
Lys_B11_1	IVM.....VM.K.A.C.L.I.....N.Y.D.T								
Lys_B2_1	IVM.....VM.K.A.C.L.I.....N.Y.D.T								
Lys_C4_1	IVM.....VM.K.A.C.L.I.....N.Y.D.T								
Lys_D11_1	IVM.....VM.K.A.C.L.I.....N.Y.D.T								
Lys_H10_1	IVM.....VM.K.A.C.L.I.....N.Y.D.T								
Lys_E2_1	IVM.....VM.K.A.C.L.I.....N.Y.D.T								
Lys_D6_1	IVM.....VM.K.A.C.L.I.....N.Y.D.T								
Lys_D4_1	IVMD.....VM.K.A.C.L.I.....N.Y.D.T								
Lys_H4_1	FMN.....VM.K.A.C.L.I.....N.Y.D.T								
Lys_H8_1	ITK.T.E.D.L.....N.Y.T.Y								
Lys_B3_1	DLC.....VY.A.D.T.L.N.....S.R.E.S								

Figura 18

	1	10	20	30	40	50	60	68	
Sac7_MT	GSVKVKFKYKGEKEVDTSKIKKVVHRVGGKHYSEFYDDNGKTRGRGVSEKDPKELLQHLARREREKKL								
D5_GarA	L.L.....M.F.N.A.H.Y.Q.....Y.T.H.R								
Consenso	L.L.....k.k.n.a.h.s.g.....t.r.a.r								

Figura 19



Cuarta ronda Fc  
Combinación  
Figura 20

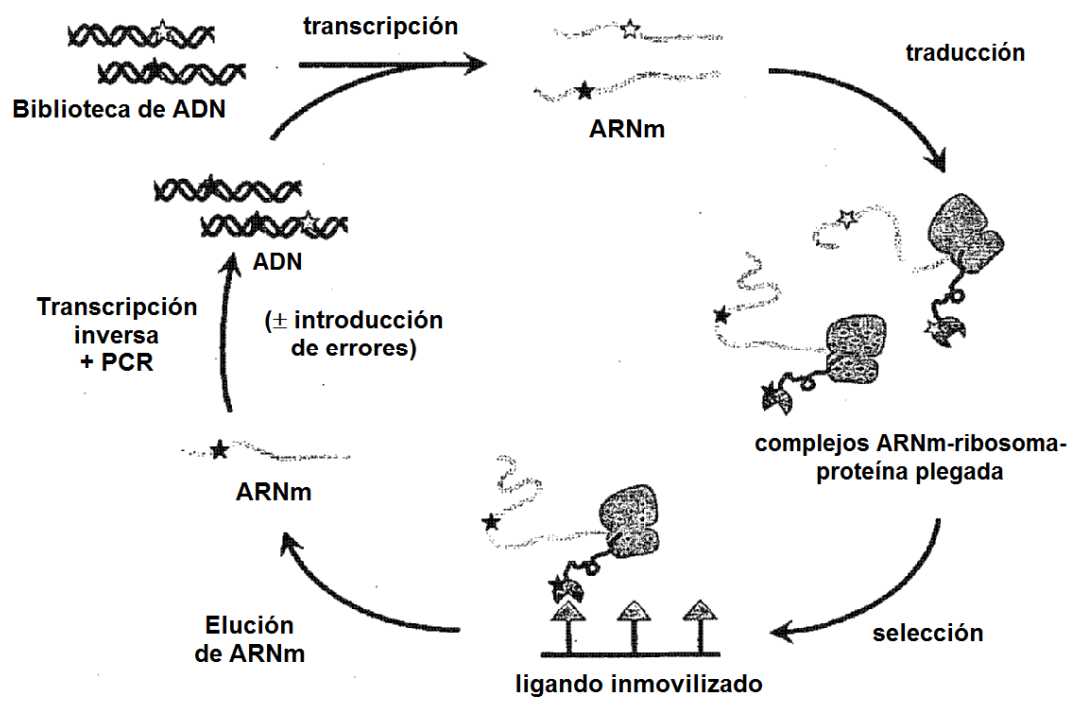


Figura 21