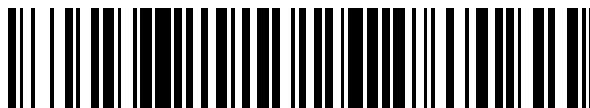


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 458 015**

21 Número de solicitud: 201231635

51 Int. Cl.:

C12Q 1/10 (2006.01)

C12R 1/19 (2006.01)

12

SOLICITUD DE PATENTE

A1

22 Fecha de presentación:

24.10.2012

43 Fecha de publicación de la solicitud:

29.04.2014

71 Solicitantes:

HIPSITEC, S.A. (100.0%)
C/ Isidoro Chamorro Pérez nº 7 - 1º A
33008 Oviedo (Asturias) ES

72 Inventor/es:

PELEGRY CUESTA, Sergio;
COBIÁN FERNÁNDEZ, Natalia;
LOMBÓ BRUGOS, Felipe;
ÁLVAREZ SAN MILLÁN, María;
FERNÁNDEZ FERNÁNDEZ, Javier;
VILLAMIZAR RODRÍGUEZ, German;
FERRERO MARTÍN, Francisco Javier;
CAMPO RODRÍGUEZ, Juan Carlos;
VALLEDOR LLOPIS, Marta y
GONZÁLEZ ALONSO, Pablo

74 Agente/Representante:

CARVAJAL Y URQUIJO, Isabel

54 Título: **Uso de las proteínas GFP-Hadrurina, GFP-gp12 y GFP-Colicina-S4 para la detección de Escherichia coli y procedimiento de detección**

57 Resumen:

Uso de una mezcla de las proteínas GFP-Hadrurina, GFP-gp12 y GFP-Colicina-S4 (y procedimiento de obtención de dichas proteínas) para la detección de Escherichia coli en un líquido, por ejemplo: aguas de consumo, bebidas embotelladas, homogeneizado procedente de un alimento, aguas residuales destinadas a consumo agrícola y líquidos de la industria láctea y conservera. La invención también se refiere a un procedimiento de detección de Escherichia coli en un líquido que comprende: filtrar dicho líquido a través de un filtro, añadir las proteínas GFP-Hadrurina, GFP-gp12 y GFP-Colicina-S4 sobre dicho filtro, detectar la señal de fluorescencia emitida por las moléculas fluorescentes presentes en el filtro y cuantificar el número de células de E. coli presentes en dicho líquido a partir del valor de fluorescencia detectado en la etapa anterior.

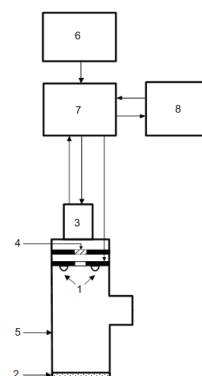


Fig. 1

ES 2 458 015 A1

DESCRIPCIÓN

Uso de las proteínas GFP-Hadrurina, GFP-gp12 y GFP-Colicina-S4 para la detección de *Escherichia coli* y procedimiento de detección

CAMPO DE LA INVENCION

5 La presente invención está relacionada con la detección de contaminación microbiológica en tiempo real. En la presente invención, se detecta *Escherichia coli* utilizando proteínas con afinidad específica para componentes de la pared celular y/o membrana celular de *E. coli*. Estas proteínas son Hadrurina (procedente del alacrán *Hadrurus aztecus*), Colicina S4 (bacteriocina procedente de *E. coli*) y proteína de cápside gp12 (procedente del fago T4), unidas a la proteína verde fluorescente GFP.

ANTECEDENTES DE LA INVENCION

En el estado de la técnica se han utilizado con anterioridad diversos métodos para detectar la presencia de *Escherichia coli*. Podemos citar entre otros la hibridación de ADN, PCR o qRT-PCR, inmunoensayos, separación inmunomagnética, separación inmune por flujo lateral y detección por nanopartículas.

15 Los métodos previos del estado de la técnica presentan ciertos inconvenientes. Así, por ejemplo, los métodos basados en el uso de anticuerpos necesitan mantener las soluciones de anticuerpos refrigerados y en unas condiciones óptimas de conservación y conllevan normalmente múltiples pasos de lavado. Otros métodos requieren muchas copias de la molécula diana para dar una señal medible por lo que es necesario amplificar esta señal.

20 La presente invención propone un sistema alternativo de detección de *E. coli*, que permite detectar la presencia de unas pocas células de *E. coli* en una muestra de agua, sin necesidad de amplificación de la señal.

DESCRIPCION DE LA INVENCION

Una realización de la invención es el uso, en adelante uso de la invención, de una mezcla de las proteínas GFP-Hadrurina, GFP-gp12 y GFP-Colicina-S4, identificadas por las secuencias SEQ ID NO: 3, 6 y 9 respectivamente, para la detección de *Escherichia coli* en un líquido.

25 La Hadrurina es un péptido antimicrobiano de 41 aminoácidos de longitud, que forma parte del cóctel químico presente en el veneno del alacrán centroamericano *Hadrurus aztecus*, y que es capaz de unirse a la membrana externa de *E. coli*.

30 La estructura tridimensional de este péptido es una hélice alfa corta, que físicamente se intercala entre los constituyentes de esta membrana externa, produciendo poros en la misma que desorganizan las envueltas celulares y posteriormente causan la lisis de la célula bacteriana.

La unión del fago T4 a *E. coli* se realiza mediante las espículas o fibras cortas de la cola denominadas gp12. Los receptores de gp12 son los lipopolisacáridos y las proteínas OmpC de *E. coli*.

35 La Colicina-S4 es un tipo de bacteriocina específica para *E.coli*. Las bacteriocinas son sustancias de naturaleza proteica que actúan como bactericidas y son sintetizadas por algunas cepas bacterianas. El mecanismo de acción es diverso (bloques metabólicos, cambios de permeabilidad en la membrana celular, daño o degradación de DNA, etc.). La Colicina-S4 forma poros en la membrana citoplasmática, rompiéndola por gradiente electroquímico. Los receptores de esta bacteriocina son OmpW, F y Tol System.

Las tres proteínas quiméricas comprenden un dominio catalítico y un péptido de unión específico para una proteína diana situada en la pared de *E. coli*.

40 En la presente invención se entiende por "quimeras" o "proteínas quiméricas" a proteínas que comprenden secuencias procedentes de diferentes organismos, y que han sido por tanto diseñadas, sintetizadas y aisladas por los inventores.

Las tres quimeras son GFP-Hadrurina, GFP-gp12 y GFP-Colicina S4. El uso de las tres quimeras consigue una mayor unión a la pared de *E.coli* y una mayor emisión de fluorescencia.

45 Las tres proteínas quiméricas comprenden:

- Dominio catalítico: GFP
- Espaciador: Secuencia de 101 aminoácidos
- Péptido de unión específica a *E. coli*: Hadrurina, gp12 o Colicina S4

Adicionalmente, dichas proteínas quiméricas contienen también una cola de histidinas, útil para purificar en columnas cromatográficas de Ni-sefarosa.

Como dominio catalítico se ha utilizado GFP (Proteína Verde Fluorescente). El gen de la GFP ha sido aislado de la medusa *Aequorea victoria* y genera una proteína de 238 aminoácidos que adopta una conformación tridimensional de barril, capaz de emitir luz fluorescente (rango de absorción-emisión: 395nm-508nm) una vez excitada con un haz de luz ultravioleta. Su detección se realiza por medio de un equipo capaz de medir fluorescencia.

Una realización es el uso de la invención, donde dicho líquido se selecciona del grupo compuesto por aguas de consumo, bebidas embotelladas, homogeneizado procedente de un alimento, aguas residuales destinadas a consumo agrícola, líquidos de la industria láctea y líquidos de la industria conservera.

Una realización de la invención es un procedimiento de detección de *Escherichia coli* en un líquido, en adelante procedimiento de detección de la invención, que comprende las siguientes etapas:

(a) filtrar dicho líquido a través de un filtro,

(b) añadir las proteínas GFP-Hadrurina, GFP-gp12 y GFP-Colicina-S4, identificadas por las secuencias SEQ ID NO: 3, 6 y 9 respectivamente, sobre dicho filtro,

(c) detectar la señal de fluorescencia emitida por las moléculas fluorescentes presentes en el filtro y

(d) cuantificar el número de células de *E. coli* presentes en dicho líquido a partir del valor de fluorescencia detectado en la etapa anterior.

La filtración de la etapa (a) se puede realizar utilizando un filtro de tamaño de poro 0,45 micras donde quedan retenidos todos los microorganismos.

Después de la etapa (b) se puede realizar un lavado para que difundan las proteínas que no se hayan unido a la pared de *E.coli*.

El procedimiento de detección de la invención garantiza la posibilidad de detectar la presencia de una única célula bacteriana de *E. coli*. Este procedimiento permite determinar si la muestra testada cumple con las exigencias sanitarias recogidas en la legislación vigente en tiempo real.

Una realización es el procedimiento de detección de la invención, donde dicho líquido se selecciona del grupo compuesto por aguas de consumo, bebidas embotelladas, homogeneizado procedente de un alimento, aguas residuales destinadas a consumo agrícola, líquidos de la industria láctea y líquidos de la industria conservera.

TEXTO LIBRE DE LA LISTA DE SECUENCIAS

A continuación se aporta una traducción al español del texto libre que aparece en la lista de secuencias.

SEQ ID NO: 1. Quimera GFP-Hadrurina

SEQ ID NO: 2. Quimera GFP-Hadrurina con cola de His

1..7 Cola de His; 8..244 Proteína GFP; 245..344 Espaciador; 345..386 Hadrurina

SEQ ID NO: 3. Quimera GFP-Hadrurina sin cola de His

1..236 proteína GFP; 237..336 Espaciador; 337..379 Hadrurina

SEQ ID NO: 4. Quimera GFP-gp12

SEQ ID NO: 5. Quimera GFP-gp12 con cola de His

1..7 Cola de His; 8..244 Proteína GFP; 245..344 Espaciador; 345..386 Proteína del fago T4 gp12

SEQ ID NO: 6. Quimera GFP-gp12 sin cola de His

1..236 Proteína GFP; 237..336 Espaciador; 337..977 Proteína del fago T4 gp12

SEQ ID NO: 7. Quimera GFP-colicina-S4

SEQ ID NO: 8. Quimera GFP-colicina-S4 con cola de His

1..7 Cola de His; 8..244 Proteína GFP; 245..344 Espaciador; 345..843 Proteína del fago T4 gp12

SEQ ID NO: 9. Quimera GFP-colicina-S4 sin cola de His

1..236 Proteína GFP; 237..336 Espaciador; 337..836 Proteína del fago T4 gp12

BREVE DESCRIPCIÓN DE LAS FIGURAS

Figura 1. La figura muestra el diagrama de bloques del sistema de medida utilizado para la detección de *E.coli*. Una fuente de luz (1), constituida por dos LED UV de 308 nm, que excitan la proteína fluorescente que recubre cada célula de *E.coli*, retenida en el filtro de papel (2). La fluorescencia emitida por la proteína es recogida por un módulo fotosensor de tipo PMT (3), previamente filtrada por el filtro óptico paso-banda centrado en 508 nm (4). Todo ello va situado sobre un tubo de PVC (5). El resto de bloques son la fuente de alimentación (6), el circuito de medida (7) y el circuito de control (8). La elección del filtro óptico y del número de LEDs de excitación se escogió en base a la cantidad de luz reflejada por el filtro de papel y las dimensiones del tubo de PVC.

Figura 2. Esquema de la estructura del plásmido pNCF68. "GFP-gp12" en esta figura corresponde a la SEQ ID NO: 4.

Figura 3. Esquema de la estructura del plásmido pNCF69. "GFPHD" en esta figura corresponde a la SEQ ID NO: 1.

Figura 4. Esquema de la estructura del plásmido pNCF72. "GFPS4" en esta figura corresponde a la SEQ ID NO: 7.

MODOS DE REALIZACIÓN PREFERENTE

Ejemplo 1. Expresión y purificación de las quimeras GFP-Hadrurina, GFP-gp12 y GFP-Colicina S4.

Los genes sintéticos de las tres quimeras se clonaron en el vector plasmídico pMAS49 integrativo en *Streptomyces*, especie utilizada como factoria microbiana para la producción y aislamiento de las quimeras, dando lugar a los plásmidos pNCF68 (GFP-gp12), pNCF69 (GFP-Hadrurina) y pNCF72 (GFP-Colicina-S4). En la expresión de las quimeras se han utilizado las especies *S. coelicolor* y *S. albus*.

El plásmido pMAS49 es un vector replicativo en *E. coli* e integrativo en *Streptomyces*, de 6,5 kb, que contiene como marcador de resistencia para ambas especies el gen de resistencia a apramicina *aac(3)IV*. El plásmido pMAS49 se integra de forma estable en el cromosoma de bacterias del género *Streptomyces* por el sitio *attP* proveniente del fago Φ C31. Este plásmido pMAS49 contiene como promotor P_{rplM} , de alta eficiencia en *Streptomyces*, derivado de un operón ribosómico, y que está flanqueado en su extremo 5' por las secuencias de reconocimiento para las enzimas de restricción *PstI*, *EcoRV*, *XbaI*, *SphI*, *SpeI*, *BamHI* y *BglII*.

El plásmido pNCF68 contiene, clonada en los sitios de restricción *SpeI* y *BamHI* la banda de 2,9 kb que codifica la proteína quimera GFP-gp12. La Figura 2 muestra un esquema de la estructura del plásmido.

El plásmido pNCF69 contiene, clonada en los sitios de restricción *BamHI* y *BglII* la banda de 1,2 kb que codifica la proteína quimera GFP-hadrurina. La Figura 3 muestra un esquema de la estructura del plásmido.

El plásmido pNCF72 contiene, clonada en los sitios de restricción *PstI* y *SphI*, la banda de 2,5 kb que codifica la proteína quimera GFP-colicina-S4. La Figura 4 muestra un esquema de la estructura del plásmido.

Se obtuvieron cultivos semipreparativos (250 ml) de las cepas que contienen los genes para las tres quimeras integrados en su cromosoma, y estos cultivos se lisaron mediante French Press y se purificaron las tres proteínas quiméricas (identificadas por las secuencias SEQ ID NO: 2, 5 Y 8) mediante Fast protein liquid chromatography (FPLC) utilizando columnas de His-Trap con níquel.

Ejemplo 2. Detección de *E. coli* en aguas destinadas al consumo humano.

Se filtró un volumen de 200 ml de agua a través de un filtro de 20 micras para retener sólidos en suspensión. Después se pasó el agua a través de un filtro de acetato de celulosa de 0,45 micras de tamaño de poro donde quedarán retenidos los microorganismos. A continuación se añadió 1 ml del reactivo que contiene las tres quimeras y se realizó un lavado. El reactivo contenía las tres quimeras a una concentración 45 µg/ml en una disolución 5 mM Tris-HCl, 5 mM EDTA, pH 8.0.

Se han realizado ensayos con muestras de agua no contaminadas y contaminadas con *E.coli* a distintas concentraciones. Se midió la fluorescencia emitida por las tres quimeras con un fluorímetro. El fluorímetro transforma la señal física recibida en una señal eléctrica (0-5V). La diferencia de voltaje con respecto al blanco indica la presencia de *E. coli* en la muestra.

En la Tabla 1 se muestran los resultados obtenidos.

Tabla 1

Blanco [V]	10UFC	Voltaje [V]
5,53	6,81	0,64
6,11	7,22	1,11
6,16	7,46	1,3
6,11	7,37	1,26
6,02	7,35	1,33
6,08	7,39	1,31
5,98	6,85	0,87
5,96	6,96	1
5,96	6,97	1,01
	Promedio	1,09

Blanco [V]	1*10 ⁸ UFC	Voltaje [V]
6,01	6,66	0,65
6,15	8,33	2,18
6,1	8,18	2,08
6,01	7,86	1,85
6,17	7,7	1,53
6,01	7,7	1,76
5,92	7,31	1,39
6,16	7,27	1,11
6,16	7,68	1,52
6,1	7,66	1,56
	Promedio	1,56

5

La media de todas las muestras analizadas de agua no contaminada (Blanco) es de 6.03(V).

La media de muestras con una cantidad de *E.coli* de 10 UFC (Unidades formadoras de colonias) es de 7,15 (V).

La media de muestras con una cantidad de *E.coli* de 1*10⁸ UFC es de 7,63 (V).

10

La figura 1 muestra un sistema utilizado para realizar la detección de *E. coli* según el procedimiento de la invención. En la cámara de detección se acopla un filtro de manera hermética. En el filtro quedan retenidas las células de *E. coli* tras filtrar el agua y se unen las proteínas quiméricas que se detectan mediante técnicas de fluorescencia. El equipo de detección de señal está compuesto por un fluorímetro y una lámpara UV. El fluorímetro se ha diseñado y construido para optimizar la señal emitida y recibida de las quimeras, con un espectro de absorción-emisión de 395-510 nm.

15

Ejemplo 3. Detección de *E. coli* en alimentos.

5 Se mezclaron 10 g de una muestra (un producto lácteo, concretamente yogur) con 190 g de Agua de Peptona Tamponada con un homogenizador. La mezcla resultante se filtró a través de un filtro de 0,5 mm y luego a través de otro de 20 micras para retener sólidos en suspensión. Después la mezcla se pasó a través de un filtro de acetato de celulosa de 0,45 micras de tamaño de poro donde quedaron retenidos los microorganismos. A continuación se añadió 1 ml del reactivo que contiene las tres quimeras y se realizó un lavado. El reactivo contenía las tres quimeras a una concentración 45 µg/ml en una disolución 5 mM Tris-HCl, 5 mM EDTA, pH 8.0.

Se midió la fluorescencia emitida por las tres quimeras con un fluorímetro.

REIVINDICACIONES

1. Uso de una mezcla de las proteínas GFP-Hadrurina, GFP-gp12 y GFP-Colicina-S4, identificadas por las secuencias SEQ ID NO: 3, 6 y 9 respectivamente, para la detección de *Escherichia coli* en un líquido.
- 5 2. Uso según la reivindicación 1, caracterizado por que dicho líquido se selecciona del grupo compuesto por aguas de consumo, bebidas embotelladas, homogeneizado procedente de un alimento, aguas residuales destinadas a consumo agrícola, líquidos de la industria láctea y líquidos de la industria conservera.
3. Un procedimiento de detección de *Escherichia coli* en un líquido, caracterizado por que comprende las siguientes etapas:
 - 10 (a) filtrar dicho líquido a través de un filtro,
 - (b) añadir las proteínas GFP-Hadrurina, GFP-gp12 y GFP-Colicina-S4, identificadas por las secuencias SEQ ID NO: 3, 6 y 9 respectivamente, sobre dicho filtro,
 - (c) detectar la señal de fluorescencia emitida por las moléculas fluorescentes presentes en el filtro y
 - 15 (d) cuantificar el número de células de *E. coli* presentes en dicho líquido a partir del valor de fluorescencia detectado en la etapa anterior.
4. El procedimiento según la reivindicación 3, caracterizado por que dicho líquido se selecciona del grupo compuesto por aguas de consumo, bebidas embotelladas, homogeneizado procedente de un alimento, aguas residuales destinadas a consumo agrícola, líquidos de la industria láctea y líquidos de la industria conservera.

20

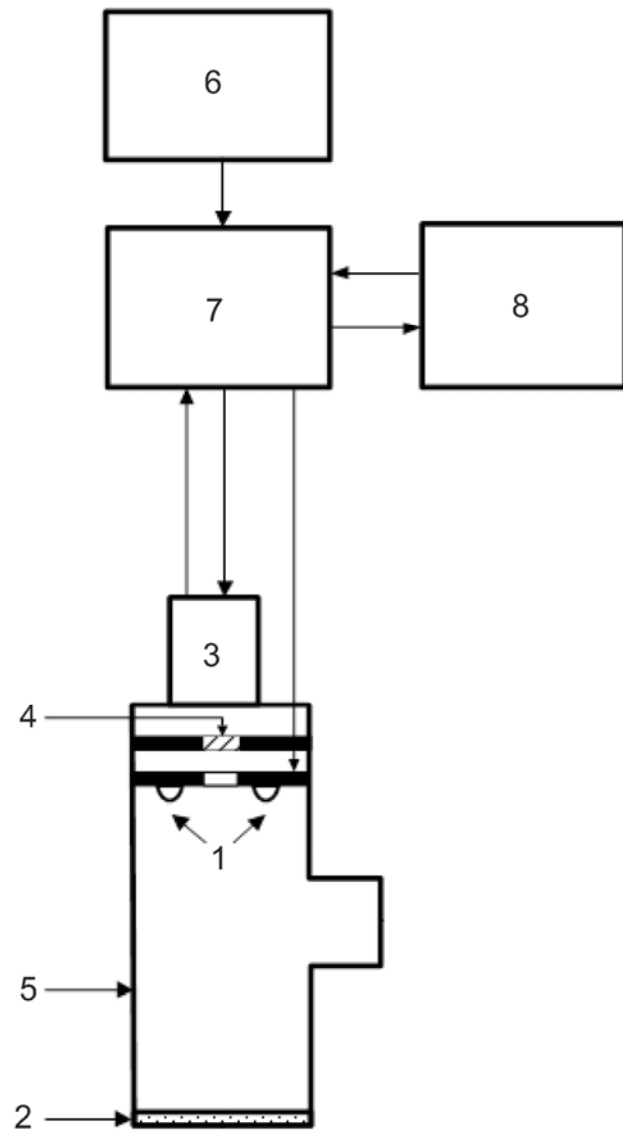


Fig. 1

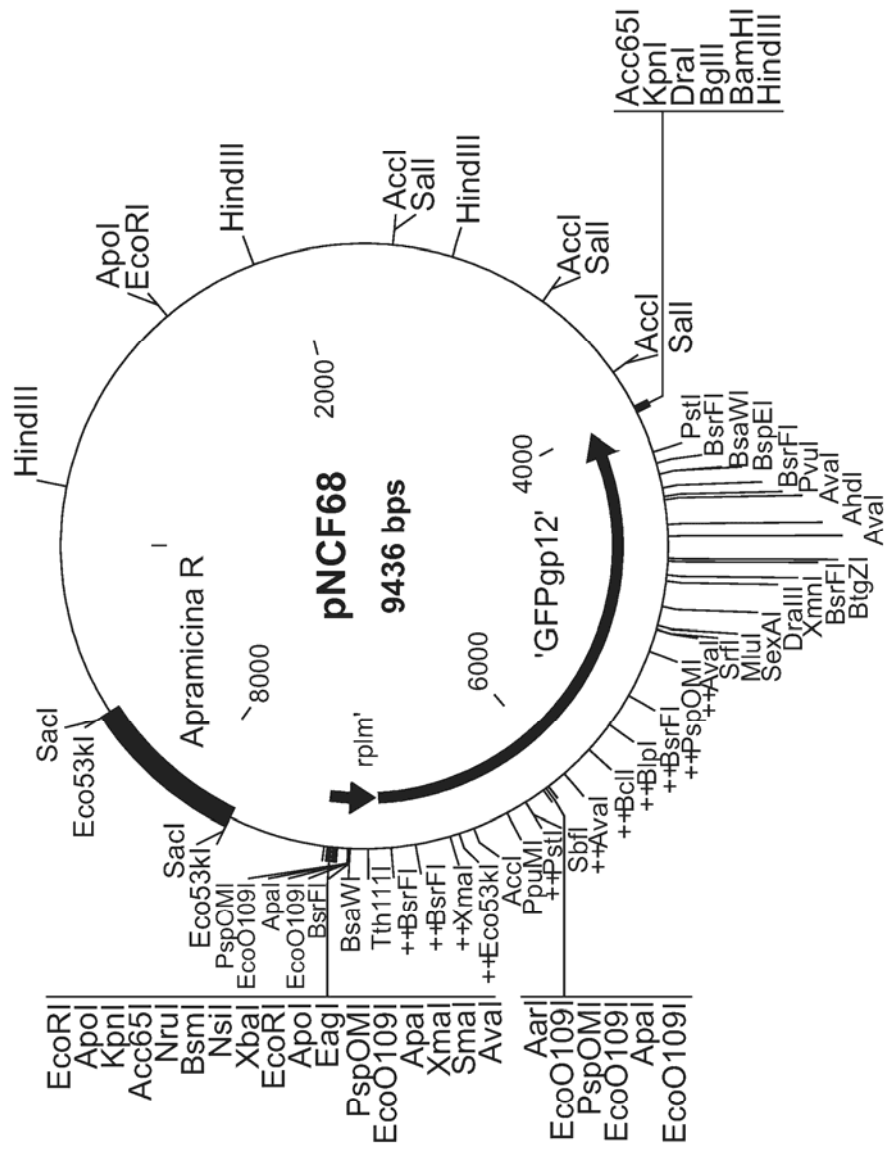


Fig. 2

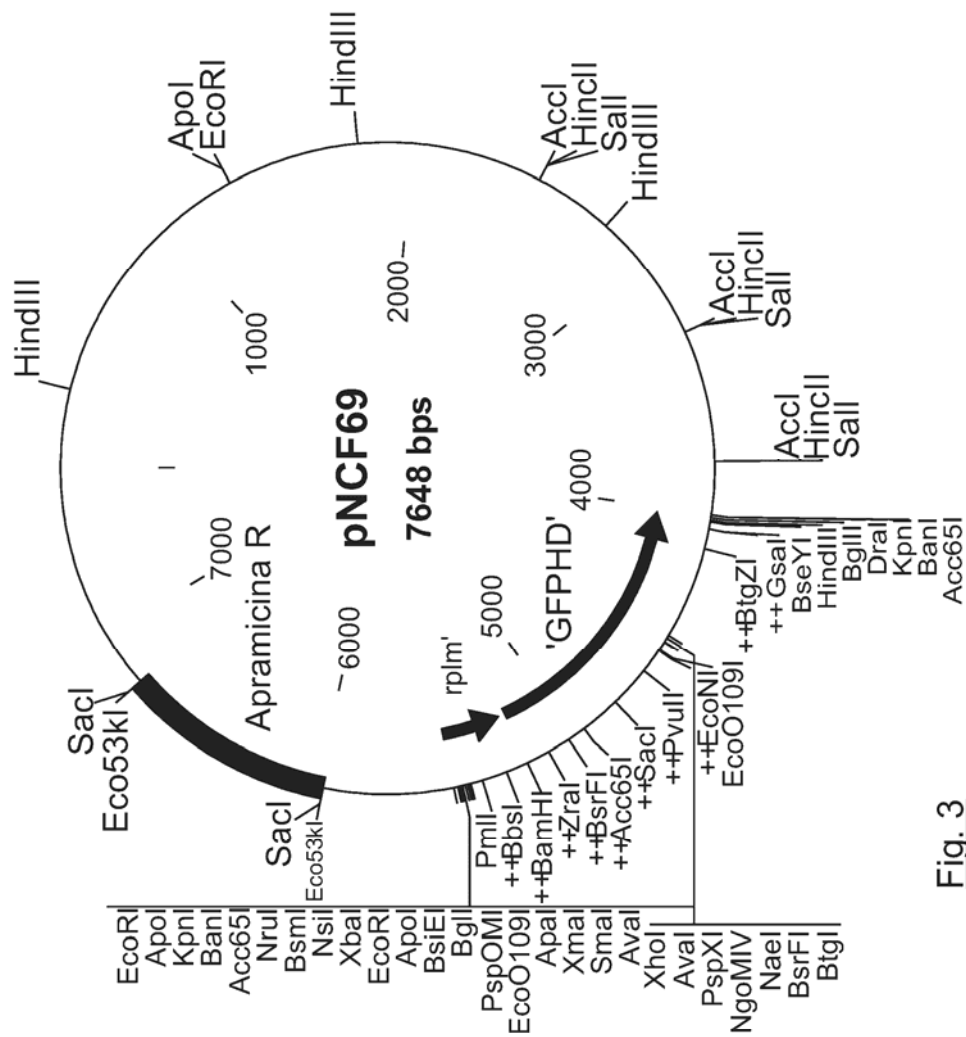


Fig. 3

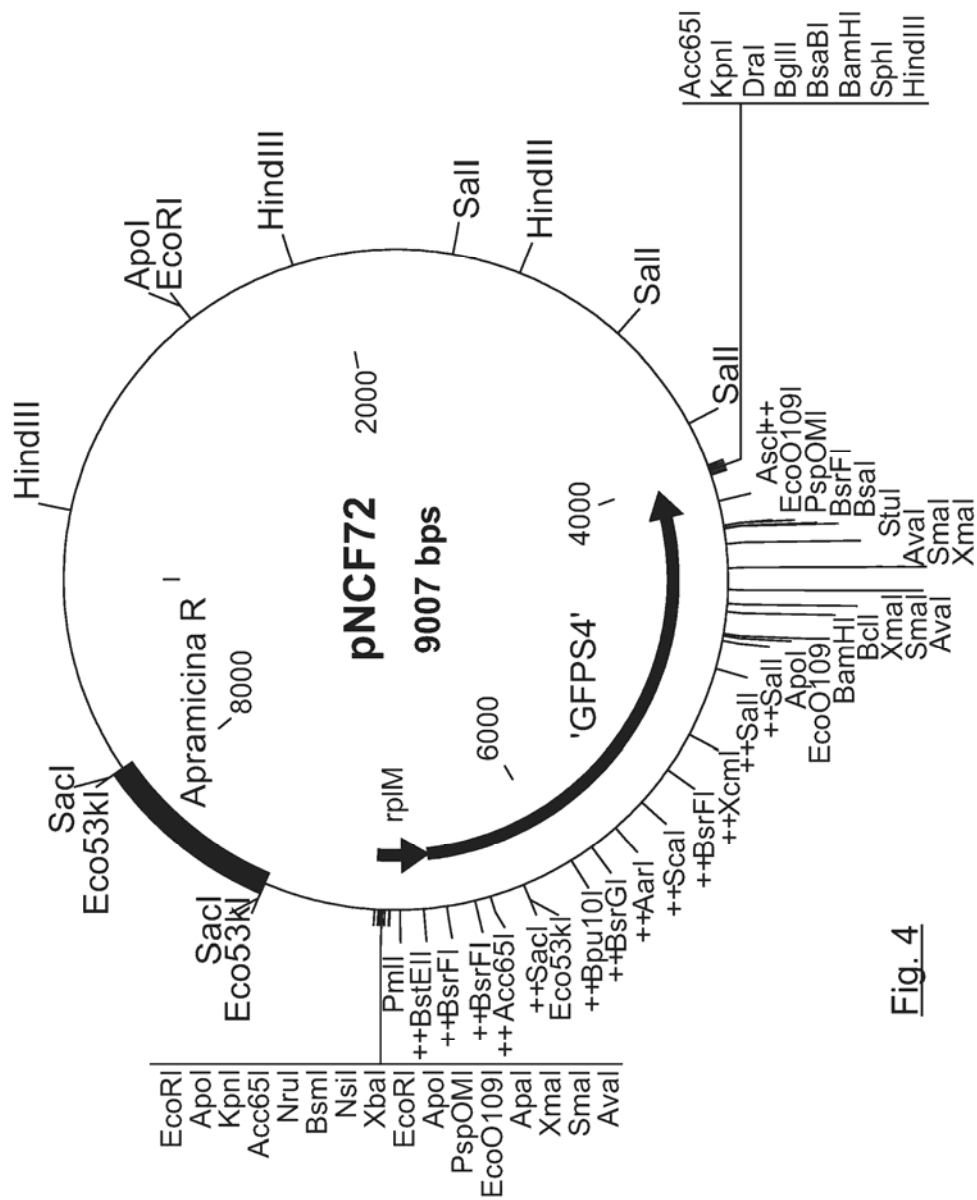


Fig. 4

LISTA DE SECUENCIAS

<110> HIPSITEC, S. A.

<120> USO DE LAS PROTEÍNAS GFP-HADRURINA, GFP-gp12 Y GFP-COLICINA-S4 PARA LA DETECCIÓN DE ESCHERICHIA COLI Y PROCEDIMIENTOS RELACIONADOS

<130> P3479/2012

<160> 9

<170> BiSSAP 1.0

<210> 1

<211> 1173

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<222> 1..1173

<223> /mol_type="DNA"
/note="GFP-Hadrurine chimera"
/organism="Artificial Sequence"

<220>

<221> CDS

<222> 7..1164

<223> /transl_table=1
/translation="MHHHHHHSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLT
FICTTGKLPVPWPTLVTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFYKDDGNYKSRA
EVKFEGDTLVNRIELKGIDFKEDGNILGHKMEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGS
VQLADHYQQNTPIGDGPVLLPDNHYLSTQSALS KDPNEKRDHMILLEFVTAAGITHGMDELYKEL
FLRSLQDKYSLNLFHVP TPLDPHTHV KQIDKYDLSNLHFHLP GHKAPDNKGPYTPKKGDPPKGLD
TGKPTGHNQRGHYPLLNNPEATNAGKYEYWPSHMGILDTIKSIASKVWNSKTVQDLKRKGINWVA
NKLGVSPQAA"

<220>

<221> CDS

<222> 28..1164

<223> /transl_table=1
/translation="SKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLT LKFICTTGK
LPVPWPTLVTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFYKDDGNYKSRAEVKFEGD
TLVNRIELKGIDFKEDGNILGHKMEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHY
QQNTPIGDGPVLLPDNHYLSTQSALS KDPNEKRDHMILLEFVTAAGITHGMDELYKELFLRSLQD
KYSNLHFHVP TPLDPHTHV KQIDKYDLSNLHFHLP GHKAPDNKGPYTPKKGDPPKGLDTGKPTGH
NQRGHYPLLNNPEATNAGKYEYWPSHMGILDTIKSIASKVWNSKTVQDLKRKGINWVANKLGVSP
QAA"

<400> 1

tctagaatgc accaccacca ccaccactcc aagggcgagg agctgttcac cggcgctcgtg	60
cccacccctgg tcgagctgga cggcgacgtc aacggccaca agttctccgt ctccggcgag	120
ggcgagggcg acgccaccta cggcaagctg accctcaagt tcacctgcac caccggcaag	180
ctgcccgtcc cctggccac cctcgtgacc accttctcct acggcgtgca gtgcttctcg	240

ES 2 458 015 A1

```

cggtagcccg accacatgaa gcagcacgac ttcttcaagt ccgccatgcc cgagggctac      300
gtccaggagc gcaccatctt ctacaaggac gacggcaact acaagagccg ggcggaggtc      360
aagttcgagg gcgacaccct ggtgaaccgc atcgagctca agggcatcga cttcaaggaa      420
gacggcaaca tcctggggcca caagatggag tacaactaca actcgcacaa cgtctacatc      480
atggccgaca agcagaagaa cggcatcaag gtgaacttca agatccgcca caacatcgag      540
gacggctccg tccagctggc cgaccactac cagcagaaca ccccgatcgg cgacggcccg      600
gtgctgctcc cggacaacca ctacctgagc acccagtccg ccctctcgaa ggacccgaac      660
gagaagcggg accacatgat cctgctcgag ttcgtcaccg ccgccggcat caccacgggc      720
atggacgagc tgtacaagga gctgttcctc cgctcgctgc aggacaagta cagcaacctc      780
cacttccacg tccccacccc gctggacccc cacacccacg tgaagcagat cgacaagtac      840
gacctgtcca acctgcactt ccacctgccg ggccacaagg ccccgacaaa caagggcccc      900
tacaccccga agaagggcga cccgcccaag ggccctggaca ccggcaagcc caccggccac      960
aaccagcggg gccactaccc cctgctcaac aacccgagg ccaccaacgc gggcaagtac     1020
gagtactggc cgagccatat gggcatcctg gacaccatca agagcatcgc gtccaaggtc     1080
tggaactcca agaccgtgca ggacctcaag cgcaagggca tcaactgggt ggcgaacaag     1140
ctgggcgtgt ccccgcaggc ggcgtgaaag ctt                                     1173

```

```

<210> 2
<211> 386
<212> PRT
<213> Artificial Sequence

```

```

<220>
<221> SOURCE
<222> 1..386
<223> /mol_type="protein"
      /note="[CDS]:7..1164 from SEQ ID NO 1"
      /note="1..7 His tag; 8..244 GFP protein; 245..344 Linker; 345..386
      Hadrurine"
      /organism="Artificial Sequence"

```

```

<400> 2
Met His His His His His Ser Lys Gly Glu Glu Leu Phe Thr Gly
1          5          10          15
Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys
20          25          30
Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu
35          40          45
Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro
50          55          60
Thr Leu Val Thr Thr Phe Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr
65          70          75          80

```

ES 2 458 015 A1

```

Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu
      85          90          95
Gly Tyr Val Gln Glu Arg Thr Ile Phe Tyr Lys Asp Asp Gly Asn Tyr
      100        105        110
Lys Ser Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg
      115        120        125
Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly
      130        135        140
His Lys Met Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala
      145        150        155        160
Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn
      165        170        175
Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr
      180        185        190
Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser
      195        200        205
Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met
      210        215        220
Ile Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr His Gly Met Asp
      225        230        235        240
Glu Leu Tyr Lys Glu Leu Phe Leu Arg Ser Leu Gln Asp Lys Tyr Ser
      245        250        255
Asn Leu His Phe His Val Pro Thr Pro Leu Asp Pro His Thr His Val
      260        265        270
Lys Gln Ile Asp Lys Tyr Asp Leu Ser Asn Leu His Phe His Leu Pro
      275        280        285
Gly His Lys Ala Pro Asp Asn Lys Gly Pro Tyr Thr Pro Lys Lys Gly
      290        295        300
Asp Pro Pro Lys Gly Leu Asp Thr Gly Lys Pro Thr Gly His Asn Gln
      305        310        315        320
Arg Gly His Tyr Pro Leu Leu Asn Asn Pro Glu Ala Thr Asn Ala Gly
      325        330        335
Lys Tyr Glu Tyr Trp Pro Ser His Met Gly Ile Leu Asp Thr Ile Lys
      340        345        350
Ser Ile Ala Ser Lys Val Trp Asn Ser Lys Thr Val Gln Asp Leu Lys
      355        360        365
Arg Lys Gly Ile Asn Trp Val Ala Asn Lys Leu Gly Val Ser Pro Gln
      370        375        380
Ala Ala
      385

```

<210> 3

<211> 379

<212> PRT

<213> Artificial Sequence

<220>

<221> SOURCE

<222> 1..379

<223> /mol_type="protein"

/note="[CDS]:28..1164 from SEQ ID NO 1"

/note="1..236 GFP protein; 237..336 Linker; 337..379 Hadrurine"

/organism="Artificial Sequence"

<400> 3

```

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

```

```

Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu Gly

```

ES 2 458 015 A1

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

<210> 4

<211> 2967

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<222> 1..2967

<223> /mol_type="DNA"

/note="GFP-gp12 chimera"

/organism="Artificial Sequence"

ES 2 458 015 A1

```
<220>
<221> CDS
<222> 7..2958
<223> /transl_table=1
      /translation="MHHHHHHSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLK
      FICTTGKLPVPWPTLVTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFYKDDGNYKSRA
      EVKFEGDTLVNRIELKGIDFKEDGNILGHKMEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGS
      VQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMILLEFVTAAGITHGMDELYKEL
      FLRSLQDKYSNLHFHVPTPLDPHPTHVKQIDKYDLSNLHFHLPGHKAPDNKGPYTPKKGDPPKGLD
      TGKPTGHNQRGHYPLLNNPEATNAGKYEYWP SHMSFFAGKLNNKSILSLRRGSGGDTNQHINPDS
      QTIFHSDMSHVIITETHSTGLRLDQGAGDYYWSEMPSRVTQLHNNDPNRVVLTEIEFSDGSRHML
      SGMSMGVGAAYGIINPQIMSQGGLKTQITASADLSLDVGYFNTGTSGTIPQKLRDGTGCQHMFG
      AFSGRRGFASSAMYLGAALYKSAWSGSGYVADAGTLTIPSDYVRHPGARNFGFNAIYVRGRSC
      NRVLYGMEGPNYTTGGAVQGASSGALNFTYNPSPNPESPKYSVG FARADPTNYAYWESMGDPNDS
      ANGPIGIYSEHLGIYPSKITWYVTNLVYNGSGYNIDGGLFNGNDIKLSPREFIIKGVNVNNTSWK
      FINFIEKNFNVGNRADFRDVGCNLSKDSPSTGISGIATFGLPTTESNNAPSIKGGNVGGLHANVV
      SIYNFLPSTSWYVSSNPPKIGNNYGDVWSENLLPLRLLGGSGSTILSGNIVFQNGSVHVGTVGL
      DLNSSRNGAIVCTMEFIDDTWLSAGGIGCFNPTEMLSQGAIEYGDSRFRIGGNTINKKLHQILSLP
      AGEYVPFFTICKGTVVNACKLQAAAYNPTPYWVWSGLPGSVGQTGYTTLTYMRNDGNNNISIWLDSS
      SMSNIIIGMKACLPNIKLIQRLT"
```

```
<220>
<221> CDS
<222> 28..2958
<223> /transl_table=1
      /translation="SKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICTTGK
      LPVPWPTLVTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFYKDDGNYKSRAEVKFEGD
      TLVNRIELKGIDFKEDGNILGHKMEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSSVQLADHY
      QQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMILLEFVTAAGITHGMDELYKELFLRSLQD
      KYSNLHFHVPTPLDPHPTHVKQIDKYDLSNLHFHLPGHKAPDNKGPYTPKKGDPPKGLDGTGKPTGH
      NQRGHYPLLNNPEATNAGKYEYWP SHMSFFAGKLNNKSILSLRRGSGGDTNQHINPDSQTIFHSD
      MSHVIITETHSTGLRLDQGAGDYYWSEMPSRVTQLHNNDPNRVVLTEIEFSDGSRHMLSGMSMGV
      GAKAYGIINPQIMSQGGLKTQITASADLSLDVGYFNTGTSGTIPQKLRDGTGCQHMFGAFSGRRG
      FASSAMYLGAALYKSAWSGSGYVADAGTLTIPSDYVRHPGARNFGFNAIYVRGRSCNRVLYGM
      EGPNYTTGGAVQGASSGALNFTYNPSPNPESPKYSVG FARADPTNYAYWESMGDPNDSANGPIGI
      YSEHLGIYPSKITWYVTNLVYNGSGYNIDGGLFNGNDIKLSPREFIIKGVNVNNTSWKFINFIEK
      NFNVGNRADFRDVGCNLSKDSPSTGISGIATFGLPTTESNNAPSIKGGNVGGLHANVVSIIYNFLP
      STSWYVSSNPPKIGNNYGDVWSENLLPLRLLGGSGSTILSGNIVFQNGSVHVGTVGLDLNSSRN
      GAIVCTMEFIDDTWLSAGGIGCFNPTEMLSQGAIEYGDSRFRIGGNTINKKLHQILSLPAGEYVPF
      FTICKGTVVNACKLQAAAYNPTPYWVWSGLPGSVGQTGYTTLTYMRNDGNNNISIWLDSSSMSNII
      G MKACLPNIKLIQRLT"
```

```
<400> 4
tctagaatgc accaccacca ccaccactcg aagggcgagg agctgttcac cggcgctcgtg      60
cccatcctcg tggagctgga cggcgacgtc aacggccaca agttctccgt ctccggcgag      120
ggcgagggcg acgccaccta cggcaagctg accctcaagt tcatctgcac caccggcaag      180
ctgcccgtcc cctggcccac cctcgtgacc accttctcct acggcgtgca gtgcttctcc      240
cggtagcccg accacatgaa gcagcacgac ttcttcaagt ccgcatgcc ggagggtac      300
gtccaggagc gcaccatctt ctacaaggac gacggcaact acaagtcccg ggccgaggtc      360
aagttcgagg gcgacaccct ggtgaaccgc atcgagctca agggcatcga cttcaaggaa      420
gacggcaaca tcctgggcca caagatggag tacaactaca actcgcacaa cgtctacatc      480
```

ES 2 458 015 A1

atggcggaaca agcagaagaa cggcatcaag gtgaacttca agatccgcca caacatcgag	540
gacgggtccg tccagctggc cgaccactac cagcagaaca ccccatcgg cgacggcccc	600
gtgctgctcc ccgacaacca ctacctgtcc acccagtcgg cctcagcaa ggaccccaac	660
gagaagcggg accacatgat cctgctcgag ttctgtaccg ccgcggcat caccacggc	720
atggacgagc tgtacaagga gctgttcctc cgctccctgc aggacaagta ctgaaacctc	780
cacttccacg tccccacccc gctggacccc cacaccacg tgaagcagat cgacaagtac	840
gacctgtcca acctgcactt ccacctgccc ggccacaagg ccccggaaca caagggcccc	900
tacaccccga agaagggcga cccgcccagg ggcttgga cgggaagcc caccggccac	960
aaccagcggg gccactaccc gctgctcaac aaccccgagg cgaccaacgc cggcaagtac	1020
gagtactggc cctcccatat gtctgttctc gccggcaagc tgaacaacaa gagcatcctg	1080
tccctccgtc gcggctcggg cggcgacacc aaccagcaca tcaaccggga cagccagacc	1140
atcttccact cggacatgag ccacgtgatc atcaccgaga cccactccac cggcctgcgg	1200
ctcgaccagg gcgcggcga ctactactgg tccgagatgc ccagccgct caccagctg	1260
cacaacaacg acccgaaccg ggtcgtgctc accgagatcg agttctccga cggctcgcgc	1320
cacatgctga gcggcatgtc catgggcgtg ggccccaagg cctacggcat catcaaccgc	1380
cagatcatgt cccagggcgg cctgaagacc cagatcaccg cctccgccga cctgtcgtc	1440
gacgtcggct acttcaacac cggcacctcc ggaccatcc cgcagaagct gcgggacggc	1500
accggctgcc agcacatgtt cggcgcttc tccggccgtc gcggcttcgc gtcctccgcc	1560
atgtacctgg gcggcgccgc cctctacaag agcgcttgt cgggtccgg ctacgtggtg	1620
gccgacgccg gcacctgac catccctcc gactacgtcc gccacccggg cggccggaac	1680
ttcggcttca acgccatcta cgtccggggc cggctctgca accgggtgct gtacggcatg	1740
gagggcccga actacaccac cggcggcgcg gtgcagggcg ccagctcctc gggcgccctg	1800
aacttcacct acaaccccag caaccggag tccccgaagt actcgtggg cttcgcccg	1860
gccgaccca ccaactacgc gtactgggag tcgatgggcg accccaacga cagcgccaac	1920
ggcccatcg gcattactc cgagcacctg ggcattacc cctcgaagat cacctggtac	1980
gtcaccaacc tcgtgtacaa cggctccggc tacaacatcg acggtggcct gttcaacggc	2040
aacgacatca agctctcccc gcgcgagttc atcatcaagg gcgtcaacgt gaacaacacc	2100
tcgtggaagt tcatcaactt catcgagaag aacttcaacg tgggcaaccg ggccgacttc	2160
cgggacgtgg gctgcaacct gtccaaggac tccccctcca ccggcatctc cggcatcgcc	2220

ES 2 458 015 A1

```

accttcggcc tccccaccac cgagagcaac aacgccccct ccatcaaggg cggaacgtc 2280
ggcggcctgc acgccaacgt ggtgtccatc tacaacttcc tcccagacac ctcttggtac 2340
gtcagctcca acccgcccaa gatcggaac aactacggcg acgtctggtc cgagaacctg 2400
ctccccctgc ggctgctcgg cggctccggc agcaccatcc tcagcggcaa catcgtcttc 2460
cagggcaacg gctccgtcca cgtgggcacc gtgggcctgg acctcaactc gagccgcaac 2520
ggcgcgatcg tctgcaccat ggagttcatc gacgacacct ggctgtccgc cgggtggcatc 2580
ggctgcttca accccaccga gatgctgtcc cagggcgccg agtacggcga ctcccgcttc 2640
cggatcggcg gcaacaccat caacaagaag ctgcaccaga tcctgtccct ccccgccggc 2700
gagtacgtcc ctttcttcac catcaagggc accgtggtga acgctgcaa gctgcaggcc 2760
gccgcctaca accccacccc ctactgggtc tccggcctgc ccggcagcgt gggccagacc 2820
ggctactaca ccctcaccta ctacatgcgc aacgacggca acaacaacat ctcgatctgg 2880
ctggactcct cgatgagcaa catcatcggc atgaaggcgt gcctccccaa catcaagctc 2940
atcatccagc ggctcacctg aaagctt 2967

```

```

<210> 5
<211> 984
<212> PRT
<213> Artificial Sequence

```

```

<220>
<221> SOURCE
<222> 1..984
<223> /mol_type="protein"
      /note="[CDS]:7..2958 from SEQ ID NO 4"
      /note="1..7 His tag; 8..244 GFP protein; 245..344 Linker; 345..984 gp12 T4 phage protein"
      /organism="Artificial Sequence"

```

```

<400> 5
Met His His His His His Ser Lys Gly Glu Glu Leu Phe Thr Gly
1      5      10      15
Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys
20     25     30
Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu
35     40     45
Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro
50     55     60
Thr Leu Val Thr Thr Phe Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr
65     70     75     80
Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu
85     90     95
Gly Tyr Val Gln Glu Arg Thr Ile Phe Tyr Lys Asp Asp Gly Asn Tyr
100    105    110
Lys Ser Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg

```

ES 2 458 015 A1

		115					120					125			
Ile	Glu	Leu	Lys	Gly	Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly
	130					135					140				
His	Lys	Met	Glu	Tyr	Asn	Tyr	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala
145					150					155					160
Asp	Lys	Gln	Lys	Asn	Gly	Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn
				165					170					175	
Ile	Glu	Asp	Gly	Ser	Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr
			180					185					190		
Pro	Ile	Gly	Asp	Gly	Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser
		195					200					205			
Thr	Gln	Ser	Ala	Leu	Ser	Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met
	210					215					220				
Ile	Leu	Leu	Glu	Phe	Val	Thr	Ala	Ala	Gly	Ile	Thr	His	Gly	Met	Asp
225					230					235					240
Glu	Leu	Tyr	Lys	Glu	Leu	Phe	Leu	Arg	Ser	Leu	Gln	Asp	Lys	Tyr	Ser
				245					250					255	
Asn	Leu	His	Phe	His	Val	Pro	Thr	Pro	Leu	Asp	Pro	His	Thr	His	Val
			260					265					270		
Lys	Gln	Ile	Asp	Lys	Tyr	Asp	Leu	Ser	Asn	Leu	His	Phe	His	Leu	Pro
		275					280					285			
Gly	His	Lys	Ala	Pro	Asp	Asn	Lys	Gly	Pro	Tyr	Thr	Pro	Lys	Lys	Gly
	290					295					300				
Asp	Pro	Pro	Lys	Gly	Leu	Asp	Thr	Gly	Lys	Pro	Thr	Gly	His	Asn	Gln
305					310					315					320
Arg	Gly	His	Tyr	Pro	Leu	Leu	Asn	Asn	Pro	Glu	Ala	Thr	Asn	Ala	Gly
				325					330					335	
Lys	Tyr	Glu	Tyr	Trp	Pro	Ser	His	Met	Ser	Phe	Phe	Ala	Gly	Lys	Leu
			340					345					350		
Asn	Asn	Lys	Ser	Ile	Leu	Ser	Leu	Arg	Arg	Gly	Ser	Gly	Gly	Asp	Thr
		355					360					365			
Asn	Gln	His	Ile	Asn	Pro	Asp	Ser	Gln	Thr	Ile	Phe	His	Ser	Asp	Met
	370					375					380				
Ser	His	Val	Ile	Ile	Thr	Glu	Thr	His	Ser	Thr	Gly	Leu	Arg	Leu	Asp
385					390					395					400
Gln	Gly	Ala	Gly	Asp	Tyr	Tyr	Trp	Ser	Glu	Met	Pro	Ser	Arg	Val	Thr
				405					410					415	
Gln	Leu	His	Asn	Asn	Asp	Pro	Asn	Arg	Val	Val	Leu	Thr	Glu	Ile	Glu
			420					425					430		
Phe	Ser	Asp	Gly	Ser	Arg	His	Met	Leu	Ser	Gly	Met	Ser	Met	Gly	Val
		435					440					445			
Gly	Ala	Lys	Ala	Tyr	Gly	Ile	Ile	Asn	Pro	Gln	Ile	Met	Ser	Gln	Gly
	450					455					460				
Gly	Leu	Lys	Thr	Gln	Ile	Thr	Ala	Ser	Ala	Asp	Leu	Ser	Leu	Asp	Val
465															

ES 2 458 015 A1

			580					585				590					
Ser	Ser	Ser	Gly	Ala	Leu	Asn	Phe	Thr	Tyr	Asn	Pro	Ser	Asn	Pro	Glu		
			595					600				605					
Ser	Pro	Lys	Tyr	Ser	Val	Gly	Phe	Ala	Arg	Ala	Asp	Pro	Thr	Asn	Tyr		
			610					615				620					
Ala	Tyr	Trp	Glu	Ser	Met	Gly	Asp	Pro	Asn	Asp	Ser	Ala	Asn	Gly	Pro		
625						630				635					640		
Ile	Gly	Ile	Tyr	Ser	Glu	His	Leu	Gly	Ile	Tyr	Pro	Ser	Lys	Ile	Thr		
				645						650					655		
Trp	Tyr	Val	Thr	Asn	Leu	Val	Tyr	Asn	Gly	Ser	Gly	Tyr	Asn	Ile	Asp		
			660						665						670		
Gly	Gly	Leu	Phe	Asn	Gly	Asn	Asp	Ile	Lys	Leu	Ser	Pro	Arg	Glu	Phe		
			675					680				685					
Ile	Ile	Lys	Gly	Val	Asn	Val	Asn	Asn	Thr	Ser	Trp	Lys	Phe	Ile	Asn		
690						695					700						
Phe	Ile	Glu	Lys	Asn	Phe	Asn	Val	Gly	Asn	Arg	Ala	Asp	Phe	Arg	Asp		
705					710					715					720		
Val	Gly	Cys	Asn	Leu	Ser	Lys	Asp	Ser	Pro	Ser	Thr	Gly	Ile	Ser	Gly		
				725						730					735		
Ile	Ala	Thr	Phe	Gly	Leu	Pro	Thr	Thr	Glu	Ser	Asn	Asn	Ala	Pro	Ser		
			740						745					750			
Ile	Lys	Gly	Gly	Asn	Val	Gly	Gly	Leu	His	Ala	Asn	Val	Val	Ser	Ile		
		755				760						765					
Tyr	Asn	Phe	Leu	Pro	Ser	Thr	Ser	Trp	Tyr	Val	Ser	Ser	Asn	Pro	Pro		
770						775					780						
Lys	Ile	Gly	Asn	Asn	Tyr	Gly	Asp	Val	Trp	Ser	Glu	Asn	Leu	Leu	Pro		
785					790					795					800		
Leu	Arg	Leu	Leu	Gly	Gly	Ser	Gly	Ser	Thr	Ile	Leu	Ser	Gly	Asn	Ile		
				805						810					815		
Val	Phe	Gln	Gly	Asn	Gly	Ser	Val	His	Val	Gly	Thr	Val	Gly	Leu	Asp		
			820						825					830			
Leu	Asn	Ser	Ser	Arg	Asn	Gly	Ala	Ile	Val	Cys	Thr	Met	Glu	Phe	Ile		
		835					840					845					
Asp	Asp	Thr	Trp	Leu	Ser	Ala	Gly	Gly	Ile	Gly	Cys	Phe	Asn	Pro	Thr		
850						855					860						
Glu	Met	Leu	Ser	Gln	Gly	Ala	Glu	Tyr	Gly	Asp	Ser	Arg	Phe	Arg	Ile		
865					870					875					880		
Gly	Gly	Asn	Thr	Ile	Asn	Lys	Lys	Leu	His	Gln	Ile	Leu	Ser	Leu	Pro		
				885						890					895		
Ala	Gly	Glu	Tyr	Val	Pro	Phe	Phe	Thr	Ile	Lys	Gly	Thr	Val	Val	Asn		
			900						905					910			
Ala	Cys	Lys	Leu	Gln	Ala	Ala	Ala	Tyr	Asn	Pro	Thr	Pro	Tyr	Trp	Val		
		915					920					925					
Ser	Gly	Leu	Pro	Gly	Ser	Val	Gly	Gln	Thr	Gly	Tyr	Tyr	Thr	Leu	Thr		
930						935					940						
Tyr	Tyr	Met	Arg	Asn	Asp	Gly	Asn	Asn	Asn	Ile	Ser	Ile	Trp	Leu	Asp		
945					950					955					960		
Ser	Ser	Met	Ser	Asn	Ile	Ile	Gly	Met	Lys	Ala	Cys	Leu	Pro	Asn	Ile		
				965						970					975		
Lys	Leu	Ile	Ile	Gln	Arg	Leu	Thr										
			980														

<210> 6

<211> 977

<212> PRT

<213> Artificial Sequence

<220>

ES 2 458 015 A1

<221> SOURCE
 <222> 1..977
 <223> /mol_type="protein"
 /note="1..236 GFP protein; 237..336 Linker; 337..977 gp12 T4 phage protein"
 /note="[CDS]:28..2958 from SEQ ID NO 4"
 /organism="Artificial Sequence"

<400> 6

```

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

```

ES 2 458 015 A1

385					390					395				400	
Trp	Ser	Glu	Met	Pro	Ser	Arg	Val	Thr	Gln	Leu	His	Asn	Asn	Asp	Pro
				405					410					415	
Asn	Arg	Val	Val	Leu	Thr	Glu	Ile	Glu	Phe	Ser	Asp	Gly	Ser	Arg	His
			420					425					430		
Met	Leu	Ser	Gly	Met	Ser	Met	Gly	Val	Gly	Ala	Lys	Ala	Tyr	Gly	Ile
		435					440					445			
Ile	Asn	Pro	Gln	Ile	Met	Ser	Gln	Gly	Gly	Leu	Lys	Thr	Gln	Ile	Thr
	450					455					460				
Ala	Ser	Ala	Asp	Leu	Ser	Leu	Asp	Val	Gly	Tyr	Phe	Asn	Thr	Gly	Thr
465						470				475					480
Ser	Gly	Thr	Ile	Pro	Gln	Lys	Leu	Arg	Asp	Gly	Thr	Gly	Cys	Gln	His
				485					490					495	
Met	Phe	Gly	Ala	Phe	Ser	Gly	Arg	Arg	Gly	Phe	Ala	Ser	Ser	Ala	Met
			500					505					510		
Tyr	Leu	Gly	Gly	Ala	Ala	Leu	Tyr	Lys	Ser	Ala	Trp	Ser	Gly	Ser	Gly
	515						520					525			
Tyr	Val	Val	Ala	Asp	Ala	Gly	Thr	Leu	Thr	Ile	Pro	Ser	Asp	Tyr	Val
	530					535					540				
Arg	His	Pro	Gly	Ala	Arg	Asn	Phe	Gly	Phe	Asn	Ala	Ile	Tyr	Val	Arg
545						550				555					560
Gly	Arg	Ser	Cys	Asn	Arg	Val	Leu	Tyr	Gly	Met	Glu	Gly	Pro	Asn	Tyr
				565					570					575	
Thr	Thr	Gly	Gly	Ala	Val	Gln	Gly	Ala	Ser	Ser	Ser	Gly	Ala	Leu	Asn
			580					585					590		
Phe	Thr	Tyr	Asn	Pro	Ser	Asn	Pro	Glu	Ser	Pro	Lys	Tyr	Ser	Val	Gly
		595					600					605			
Phe	Ala	Arg	Ala	Asp	Pro	Thr	Asn	Tyr	Ala	Tyr	Trp	Glu	Ser	Met	Gly
	610					615					620				
Asp	Pro	Asn	Asp	Ser	Ala	Asn	Gly	Pro	Ile	Gly	Ile	Tyr	Ser	Glu	His
625						630				635					640
Leu	Gly	Ile	Tyr	Pro	Ser	Lys	Ile	Thr	Trp	Tyr	Val	Thr	Asn	Leu	Val
				645					650					655	
Tyr	Asn	Gly	Ser	Gly	Tyr	Asn	Ile	Asp	Gly	Gly	Leu	Phe	Asn	Gly	Asn
			660					665					670		
Asp	Ile	Lys	Leu	Ser	Pro	Arg	Glu	Phe	Ile	Ile	Lys	Gly	Val	Asn	Val
		675					680					685			
Asn	Asn	Thr	Ser	Trp	Lys	Phe	Ile	Asn	Phe	Ile	Glu	Lys	Asn	Phe	Asn
	690					695					700				
Val	Gly	Asn	Arg	Ala	Asp	Phe	Arg	Asp	Val	Gly	Cys	Asn	Leu	Ser	Lys
705						710				715					720
Asp	Ser	Pro	Ser	Thr	Gly	Ile	Ser	Gly	Ile	Ala	Thr	Phe	Gly	Leu	Pro
				725					730					735	
Thr	Thr	Glu	Ser	Asn	Asn	Ala	Pro	Ser	Ile	Lys	Gly	Gly	Asn	Val	Gly
			740					745					750		
Gly	Leu	His	Ala	Asn	Val	Val	Ser	Ile	Tyr	Asn	Phe	Leu	Pro	Ser	Thr
		755						760				765			
Ser	Trp	Tyr	Val	Ser	Ser	Asn	Pro	Pro	Lys	Ile	Gly	Asn	Asn	Tyr	Gly
	770					775					780				
Asp	Val	Trp	Ser	Glu	Asn	Leu	Leu	Pro	Leu	Arg	Leu	Leu	Gly	Gly	Ser
785						790				795					800
Gly	Ser	Thr	Ile	Leu	Ser	Gly	Asn	Ile	Val	Phe	Gln	Gly	Asn	Gly	Ser
				805					810					815	
Val	His	Val	Gly	Thr	Val	Gly	Leu	Asp	Leu	Asn	Ser	Ser	Arg	Asn	Gly
			820					825					830		
Ala	Ile	Val	Cys	Thr	Met	Glu	Phe	Ile	Asp	Asp	Thr	Trp	Leu	Ser	Ala
		835						840				845			
Gly	Gly	Ile	Gly	Cys	Phe	Asn	Pro	Thr	Glu	Met	Leu	Ser	Gln	Gly	Ala

ES 2 458 015 A1

850 855 860
 Glu Tyr Gly Asp Ser Arg Phe Arg Ile Gly Gly Asn Thr Ile Asn Lys
 865 870 875 880
 Lys Leu His Gln Ile Leu Ser Leu Pro Ala Gly Glu Tyr Val Pro Phe
 885 890 895
 Phe Thr Ile Lys Gly Thr Val Val Asn Ala Cys Lys Leu Gln Ala Ala
 900 905 910
 Ala Tyr Asn Pro Thr Pro Tyr Trp Val Ser Gly Leu Pro Gly Ser Val
 915 920 925
 Gly Gln Thr Gly Tyr Tyr Thr Leu Thr Tyr Tyr Met Arg Asn Asp Gly
 930 935 940
 Asn Asn Asn Ile Ser Ile Trp Leu Asp Ser Ser Met Ser Asn Ile Ile
 945 950 955 960
 Gly Met Lys Ala Cys Leu Pro Asn Ile Lys Leu Ile Ile Gln Arg Leu
 965 970 975
 Thr

<210> 7
 <211> 2544
 <212> DNA
 <213> Artificial Sequence

<220>
 <221> source
 <222> 1..2544
 <223> /mol_type="DNA"
 /note="GFP-Colicin-S4 chimera"
 /organism="Artificial Sequence"

<220>
 <221> CDS
 <222> 7..2535
 <223> /transl_table=1
 /translation="MHHHHHHSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLT
 FICTTGKLPVPWPTLVTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFYKDDGNYKSRA
 EVKFEGDTLVNRIELKGIDFKEDGNILGHKMEYNYNSHNVIYIMADKQKNGIKVNFKIRHNIEDGS
 VQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMILLEFVTAAGITHGMDELYKEL
 FLRSLQDKYSLNLFHFVPTPLDPHTHVKQIDKYDLSNLFHFLPGHKAPDNKGPYTPKKGDPPKGLD
 TGKPTGHNQRGHYPLLNNPEATNAGKYEYWP SHMAKELSVYGPTAGESMGGTGANLNQQGGNNNS
 NSGVHWGGGSGSGNGGREHGSQTGWGWSKTNNPDVPPYVDDNGQVRITITNGLVKTPVYGVPGAG
 GNSDVQGGYIPENPNDEVARKWDKNNLPREIDVSDIGFKYRVTLNDNGRAIGILRTGVRPYVGSE
 KAKAGIMEKINHKTPEEIYEALGFNKNDESQRQEKAKQQAEDAWDRLPNVRKFDVDVEQFHYLVV
 LDDYGNVLSVTRTGVRPYVGSEKAKAGIMDKVDHKTPEEIYEALGFNNEEPQRQNQAKKAAVDVF
 YFSMNRDRIQSDVLNKAEEVISDIGNKVGDYLGDAYKSLAREIADDVKNFQGKTIRSYDDAMAS
 LNKVLSNPGFKFNRADSDALANVWRSIDAQDMANKLGNISKAFKFADVMMKVEKVKREKSIEGYET
 GNWGPLMLEVESWVLSGIAVAVALGVFSATLGAYALSLGAPAIAGVIGILLAAVVGALLDDKFA
 DALNKEIIPAH"

<220>
 <221> CDS
 <222> 28..2535
 <223> /transl_table=1
 /translation="SKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLT
 LKFICTTGKLPVPWPTLVTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFYKDDGNYKSRA
 EVKFEGDTLVNRIELKGIDFKEDGNILGHKMEYNYNSHNVIYIMADKQKNGIKVNFKIRHNIEDGS
 VQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMILLEFVTAAGITHGMDELYKEL
 FLRSLQDKYSLNLFHFVPTPLDPHTHVKQIDKYDLSNLFHFLPGHKAPDNKGPYTPKKGDPPKGLD
 TGKPTGHNQRGHYPLLNNPEATNAGKYEYWP SHMAKELSVYGPTAGESMGGTGANLNQQGGNNNS

ES 2 458 015 A1

NQRGHYPLLNNPEATNAGKYEYWPSHMAKELSVYGPTAGESMGGTGANLNQQGGNNNSNSGVHWG
 GSGSGNGGREHGSQTGWGWSKTNNPDVPPYVDDNGQVRITITNGLVKTPVYGVPGAGGNSDVQG
 GYIPENPNDEVARKWDKNNLPREIDVSIDGFKYRVTLNDNGRAIGILRTGVRPYVGSEKAKAGIM
 EKINHKTPEEIIYEALGFNKDESQRQEKAKQQAEDAWDRLPNVRKFDVDVEQFHLYLVLLDDYGNV
 LSVTRTGVRPYVGSEKAKAGIMDKVDHKTPEEIIYEALGFNNNEEPQRQNAKKAAYDVFYFSMNR
 DRIQSDVLNKA AEVISDIGNKVG DYLGDAYKSLAREIADDVKNFQGKTIRSYDDAMASLNKVL SN
 PGFKFN RADSDALANVWRSIDAQDMANKLGNISKAFKFADVMMKVEK VREKSIEGYETGNWGPLM
 LEVESWVLSG IASAVALGVFSATLGAYALSLGAPAI AVGIVGILLAAVVGALLDDKFADALNKEI
 IKPAH"

<400> 7

tctagaatgc accaccacca ccaccactcc aagggcgagg agctgttcac cggcgctcgtg	60
cccatcctgg tcgagctgga cggcgacgtc aacggccaca agttctccgt ctccggcgag	120
ggcgagggcg acgccaccta cggcaagctg accctcaagt tcctctgcac caccggcaag	180
ctgcccgtcc cctggcccac cctcgtgacc accttctcct acggcgtgca gtgcttctcg	240
cggtagcccg accacatgaa gcagcacgac ttcttcaagt ccgcatgcc cgagggctac	300
gtccaggagc gcaccatctt ctacaaggac gacggcaact acaagagccg ggcgagggtc	360
aagttcgagg gcgacacctt ggtgaaccgc atcgagctca agggcatcga cttcaaggaa	420
gacggcaaca tcctgggcca caagatggag tacaactaca actcgcacaa cgtctacatc	480
atggccgaca agcagaagaa cggcatcaag gtgaacttca agatccgcca caacatcgag	540
gacggctccg tccagctggc cgaccactac cagcagaaca ccccgatcgg cgacggcccg	600
gtgctgctcc cggacaacca ctacctgagc acccagtcgg ccctctcgaa ggacccgaac	660
gagaagcggg accacatgat cctgctcgag ttcgtcaccg ccgccggcat caccacggc	720
atggacgagc tgtacaagga gctgttcctc cgctcgctgc aggacaagta cagcaacctc	780
cacttccacg tccccacccc gctggacccc cacaccacg tgaagcagat cgacaagtac	840
gacctgtcca acctgcactt ccacctgccg ggccacaagg ccccgacaa caagggcccc	900
tacacccgga agaagggcga cccgcccaag ggcttgaca ccggcaagcc caccggccac	960
aaccagcggg gccactacc cctgctcaac aaccggagg ccaccaacgc gggcaagtac	1020
gagtactggc cgagccatat ggccaaggag ctgagcgtct acggccccac cgccggcgag	1080
tccatgggcg gcaccggcgc caacctcaac cagcagggcg gcaacaacaa ctccaactcc	1140
ggcgtccact ggggcgggcg cagcggctcc ggcaacggtg gccgcgagca cggctcccag	1200
accggctggg gctggtcgaa gaccaacaac ccggacgtcc cgccctacgt ggacgacaac	1260
ggccagggtc ggatcaccat caccaacggc ctggtcaaga ccccgctcta cggcgtgccg	1320
ggcgccggtg gcaactccga cgtgcagggc ggctacatcc ccgagaaccc caacgacgag	1380
gtggcccgca agtgggacaa gaacaacctg ccccgggaga tcgacgtgtc gatcgacggc	1440

ES 2 458 015 A1

```

ttcaagtacc gggtcaccct gaacgacaac ggccggggcca tcggcatcct ccgcaccggc 1500
gtccggccgt acgtgggctc cgagaaggcc aaggcgggca tcatggagaa gatcaaccac 1560
aagacccccg aggagatcta cgaggccctg ggcttcaaca aggacgagtc gcagcgccag 1620
gagaaggcga agcagcaggc cgaggacgcc tgggaccgcc tgccgccc aa cgtccggaag 1680
ttcgacgtcg acgtggagca gttccactac ctggtcgtgc tcgacgacta cggcaacgtg 1740
ctgtccgtga cccggaccgg cgtgcgggccg tacgtgggct cggagaaggc gaaggccggc 1800
atcatggaca aggtcgacca caagacccccg gaagaaattt atgaggccct gggttttaac 1860
aacgaggagc cccagcgcca gaaccaggcg aagaaggccg cgtacgacgt cttctactcg 1920
ttcagcatga accgcgaccg gatccagtcg gacgtcctga acaaggccgc ggaggtgatc 1980
agcgacatcg gcaacaaggt cggcgactac ctgggcgacg cctacaagtc gtcgcccgg 2040
gagatcgccg acgacgtcaa gaacttcag ggcaagacca tccggagcta cgacgacgcc 2100
atggcgtcgc tgaacaaggt gtcagcaac ccgggcttca agttcaaccg ggccgactcc 2160
gacgccctgg ccaacgtctg gcggtcgatc gacgcccgagg acatggcgaa caagctcggc 2220
aacatctcca aggccctcaa gttcgggac gtcgtgatga aggtcgagaa ggtgcgag 2280
aagtcgatcg agggctacga gaccggcaac tggggcccc tgatgctgga ggtcgagagc 2340
tgggtgctgt ccggcatcgc ctccgccgtg gccctgggcg tgttcagcgc caccctcggc 2400
gcgtacgcc tgctccctcg cgcgcccgcc atcgccgtgg gcacgtggg catcctgctc 2460
gccgcggtcg tgggcgccct gctggacgac aagttcgccg acgccctcaa caaggagatc 2520
atcaagcccc cccactgaaa gctt 2544

```

<210> 8

<211> 843

<212> PRT

<213> Artificial Sequence

<220>

<221> SOURCE

<222> 1..843

<223> /mol_type="protein"

/note="[CDS]:7..2535 from SEQ ID NO 7"

/note="1..7 His tag; 8..244 GFP protein; 245..344 Linker; 345..843 Colicin-S4"

/organism="Artificial Sequence"

<400> 8

```

Met His His His His His His Ser Lys Gly Glu Glu Leu Phe Thr Gly
1           5           10           15
Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys

```

ES 2 458 015 A1

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

ES 2 458 015 A1

				485					490					495		
Thr	Gly	Val	Arg	Pro	Tyr	Val	Gly	Ser	Glu	Lys	Ala	Lys	Ala	Gly	Ile	
			500					505					510			
Met	Glu	Lys	Ile	Asn	His	Lys	Thr	Pro	Glu	Glu	Ile	Tyr	Glu	Ala	Leu	
		515					520					525				
Gly	Phe	Asn	Lys	Asp	Glu	Ser	Gln	Arg	Gln	Glu	Lys	Ala	Lys	Gln	Gln	
	530					535					540					
Ala	Glu	Asp	Ala	Trp	Asp	Arg	Leu	Pro	Pro	Asn	Val	Arg	Lys	Phe	Asp	
545					550					555					560	
Val	Asp	Val	Glu	Gln	Phe	His	Tyr	Leu	Val	Val	Leu	Asp	Asp	Tyr	Gly	
			565						570					575		
Asn	Val	Leu	Ser	Val	Thr	Arg	Thr	Gly	Val	Arg	Pro	Tyr	Val	Gly	Ser	
			580					585					590			
Glu	Lys	Ala	Lys	Ala	Gly	Ile	Met	Asp	Lys	Val	Asp	His	Lys	Thr	Pro	
		595					600					605				
Glu	Glu	Ile	Tyr	Glu	Ala	Leu	Gly	Phe	Asn	Asn	Glu	Glu	Pro	Gln	Arg	
	610					615					620					
Gln	Asn	Gln	Ala	Lys	Lys	Ala	Ala	Tyr	Asp	Val	Phe	Tyr	Ser	Phe	Ser	
625					630					635					640	
Met	Asn	Arg	Asp	Arg	Ile	Gln	Ser	Asp	Val	Leu	Asn	Lys	Ala	Ala	Glu	
			645						650					655		
Val	Ile	Ser	Asp	Ile	Gly	Asn	Lys	Val	Gly	Asp	Tyr	Leu	Gly	Asp	Ala	
			660					665					670			
Tyr	Lys	Ser	Leu	Ala	Arg	Glu	Ile	Ala	Asp	Asp	Val	Lys	Asn	Phe	Gln	
		675					680					685				
Gly	Lys	Thr	Ile	Arg	Ser	Tyr	Asp	Asp	Ala	Met	Ala	Ser	Leu	Asn	Lys	
	690					695					700					
Val	Leu	Ser	Asn	Pro	Gly	Phe	Lys	Phe	Asn	Arg	Ala	Asp	Ser	Asp	Ala	
705					710					715					720	
Leu	Ala	Asn	Val	Trp	Arg	Ser	Ile	Asp	Ala	Gln	Asp	Met	Ala	Asn	Lys	
			725						730					735		
Leu	Gly	Asn	Ile	Ser	Lys	Ala	Phe	Lys	Phe	Ala	Asp	Val	Val	Met	Lys	
			740					745					750			
Val	Glu	Lys	Val	Arg	Glu	Lys	Ser	Ile	Glu	Gly	Tyr	Glu	Thr	Gly	Asn	
		755					760					765				
Trp	Gly	Pro	Leu	Met	Leu	Glu	Val	Glu	Ser	Trp	Val	Leu	Ser	Gly	Ile	
	770					775					780					
Ala	Ser	Ala	Val	Ala	Leu	Gly	Val	Phe	Ser	Ala	Thr	Leu	Gly	Ala	Tyr	
785					790					795					800	
Ala	Leu	Ser	Leu	Gly	Ala	Pro	Ala	Ile	Ala	Val	Gly	Ile	Val	Gly	Ile	
			805						810							

```
<210> 9
<211> 836
<212> PRT
<213> Artificial Sequence

<220>
<221> SOURCE
<222> 1..836
<223> /mol_type="protein"
      /note="[CDS]:28..2535 from SEQ ID NO 7"
      /note="1..236 GFP protein; 237..336 Linker; 337..836 Colicin-S4"
      /organism="Artificial Sequence"
```

ES 2 458 015 A1

<400> 9

```

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

```

ES 2 458 015 A1

Asn	Asp	Glu	Val	Ala	Arg	Lys	Trp	Asp	Lys	Asn	Asn	Leu	Pro	Arg	Glu
450						455				460					
Ile	Asp	Val	Ser	Ile	Asp	Gly	Phe	Lys	Tyr	Arg	Val	Thr	Leu	Asn	Asp
465					470					475					480
Asn	Gly	Arg	Ala	Ile	Gly	Ile	Leu	Arg	Thr	Gly	Val	Arg	Pro	Tyr	Val
				485					490					495	
Gly	Ser	Glu	Lys	Ala	Lys	Ala	Gly	Ile	Met	Glu	Lys	Ile	Asn	His	Lys
			500					505					510		
Thr	Pro	Glu	Glu	Ile	Tyr	Glu	Ala	Leu	Gly	Phe	Asn	Lys	Asp	Glu	Ser
		515					520					525			
Gln	Arg	Gln	Glu	Lys	Ala	Lys	Gln	Gln	Ala	Glu	Asp	Ala	Trp	Asp	Arg
	530					535					540				
Leu	Pro	Pro	Asn	Val	Arg	Lys	Phe	Asp	Val	Asp	Val	Glu	Gln	Phe	His
545					550					555					560
Tyr	Leu	Val	Val	Leu	Asp	Asp	Tyr	Gly	Asn	Val	Leu	Ser	Val	Thr	Arg
				565					570					575	
Thr	Gly	Val	Arg	Pro	Tyr	Val	Gly	Ser	Glu	Lys	Ala	Lys	Ala	Gly	Ile
			580					585					590		
Met	Asp	Lys	Val	Asp	His	Lys	Thr	Pro	Glu	Glu	Ile	Tyr	Glu	Ala	Leu
	595						600					605			
Gly	Phe	Asn	Asn	Glu	Glu	Pro	Gln	Arg	Gln	Asn	Gln	Ala	Lys	Lys	Ala
	610					615					620				
Ala	Tyr	Asp	Val	Phe	Tyr	Ser	Phe	Ser	Met	Asn	Arg	Asp	Arg	Ile	Gln
625					630					635					640
Ser	Asp	Val	Leu	Asn	Lys	Ala	Ala	Glu	Val	Ile	Ser	Asp	Ile	Gly	Asn
				645					650					655	
Lys	Val	Gly	Asp	Tyr	Leu	Gly	Asp	Ala	Tyr	Lys	Ser	Leu	Ala	Arg	Glu
			660					665					670		
Ile	Ala	Asp	Asp	Val	Lys	Asn	Phe	Gln	Gly	Lys	Thr	Ile	Arg	Ser	Tyr
		675					680					685			
Asp	Asp	Ala	Met	Ala	Ser	Leu	Asn	Lys	Val	Leu	Ser	Asn	Pro	Gly	Phe
	690					695					700				
Lys	Phe	Asn	Arg	Ala	Asp	Ser	Asp	Ala	Leu	Ala	Asn	Val	Trp	Arg	Ser
705					710					715					720
Ile	Asp	Ala	Gln	Asp	Met	Ala	Asn	Lys	Leu	Gly	Asn	Ile	Ser	Lys	Ala
			725						730					735	
Phe	Lys	Phe	Ala	Asp	Val	Val	Met	Lys	Val	Glu	Lys	Val	Arg	Glu	Lys
			740					745					750		
Ser	Ile	Glu	Gly	Tyr	Glu	Thr	Gly	Asn	Trp	Gly	Pro	Leu	Met	Leu	Glu
		755					760					765			
Val	Glu	Ser	Trp	Val	Leu	Ser	Gly	Ile	Ala	Ser	Ala	Val	Ala	Leu	Gly
	770					775					780				
Val	Phe	Ser	Ala	Thr	Leu	Gly	Ala	Tyr	Ala	Leu	Ser	Leu	Gly	Ala	Pro
785					790					795					800
Ala	Ile	Ala	Val	Gly	Ile	Val	Gly	Ile	Leu	Leu	Ala	Ala	Val	Val	Gly
			805						810					815	
Ala	Leu	Leu	Asp	Asp	Lys	Phe	Ala	Asp	Ala	Leu	Asn	Lys	Glu	Ile	Ile
			820					825					830		
Lys	Pro	Ala	His												
		835													



- ②① N.º solicitud: 201231635
②② Fecha de presentación de la solicitud: 24.10.2012
③② Fecha de prioridad:

INFORME SOBRE EL ESTADO DE LA TÉCNICA

⑤① Int. Cl.: **C12Q1/10** (2006.01)
C12R1/19 (2006.01)

DOCUMENTOS RELEVANTES

Categoría	⑤⑥ Documentos citados	Reivindicaciones afectadas
A	TANJI Y et al. " <i>Escherichia coli</i> detection by GFP-labeled lysozyme-inactivated T4 bacteriophage". Journal of Biotechnology, 20041019 ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, NL 19.10.2004 VOL: 114 No: 1-2 Págs: 11-20 ISSN 0168-1656 Doi: doi:10.1016/j.jbiotec.2004.05.011; todo el documento.	1-4
A	MIYANAGA K et al. "Detection of <i>Escherichia coli</i> in the sewage influent by fluorescent labeled T4 phage". BIOCHEMICAL ENGINEERING JOURNAL, 20060401 ELSEVIER, AMSTERDAM, NL 01.04.2006 VOL: 29 No: 1-2 Págs: 119-124 ISSN 1369-703X Doi: doi:10.1016/j.bej.2005.03.016; todo el documento.	1-4
A	BIRMELE M et al. "Characterization and validation of a bioluminescent bioreporter for the direct detection of <i>Escherichia coli</i> ". JOURNAL OF MICROBIOLOGICAL METHODS, 20081001 ELSEVIER, AMSTERDAM, NL 01.10.2008 VOL: 75 No: 2 Págs: 354-356 ISSN 0167-7012 Doi: doi:10.1016/j.mimet.2008.06.003; todo el documento.	1-4
A	US 6821770 B1 (HOGAN JAMES J) 23.11.2004, todo el documento.	1-4
A	US 2006094034 A1 (BROUSSEAU ROLAND et al.) 04.05.2006, todo el documento.	1-4

Categoría de los documentos citados

X: de particular relevancia

Y: de particular relevancia combinado con otro/s de la misma categoría

A: refleja el estado de la técnica

O: referido a divulgación no escrita

P: publicado entre la fecha de prioridad y la de presentación de la solicitud

E: documento anterior, pero publicado después de la fecha de presentación de la solicitud

El presente informe ha sido realizado

☒ para todas las reivindicaciones

☐ para las reivindicaciones nº:

Fecha de realización del informe
26.02.2014

Examinador
M. Á. García Coca

Página
1/4

Documentación mínima buscada (sistema de clasificación seguido de los símbolos de clasificación)

C12Q, C12R

Bases de datos electrónicas consultadas durante la búsqueda (nombre de la base de datos y, si es posible, términos de búsqueda utilizados)

INVENES, EPODOC, WPI, BIOSIS, MEDLINE/NLM, EMBASE/ ELSEVIER, XPESP2/ELSEVIER, Bases de datos de texto completo TXT

Fecha de Realización de la Opinión Escrita: 26.02.2014

Declaración

Novedad (Art. 6.1 LP 11/1986)	Reivindicaciones 1-4	SI
	Reivindicaciones	NO
Actividad inventiva (Art. 8.1 LP11/1986)	Reivindicaciones 1-4	SI
	Reivindicaciones	NO

Se considera que la solicitud cumple con el requisito de aplicación industrial. Este requisito fue evaluado durante la fase de examen formal y técnico de la solicitud (Artículo 31.2 Ley 11/1986).

Base de la Opinión.-

La presente opinión se ha realizado sobre la base de la solicitud de patente tal y como se publica.

1. Documentos considerados.-

A continuación se relacionan los documentos pertenecientes al estado de la técnica tomados en consideración para la realización de esta opinión.

Documento	Número Publicación o Identificación	Fecha Publicación
D01	TANJI Y et al. " <i>Escherichia coli</i> detection by GFP-labeled lysozyme-inactivated T4 bacteriophage". Journal of Biotechnology, 20041019 ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, NL 19.10.2004 VOL: 114 No: 1-2 Págs: 11-20 ISSN 0168-1656 Doi: doi:10.1016/j.jbiotec.2004.05.011.	19.10.2004
D02	MIYANAGA K et al. "Detection of <i>Escherichia coli</i> in the sewage influent by fluorescent labeled T4 phage". BIOCHEMICAL ENGINEERING JOURNAL, 20060401 ELSEVIER, AMSTERDAM, NL 01.04.2006 VOL: 29 No: 1-2 Págs: 119-124 ISSN 1369-703X Doi: doi:10.1016/j.bej.2005.03.016.	01.04.2006
D03	BIRMELE M et al. "Characterization and validation of a bioluminescent bioreporter for the direct detection of <i>Escherichia coli</i> ". JOURNAL OF MICROBIOLOGICAL METHODS, 20081001 ELSEVIER, AMSTERDAM, NL 01.10.2008 VOL: 75 No: 2 Págs: 354-356 ISSN 0167-7012 Doi: doi:10.1016/j.mimet.2008.06.003.	01.10.2008
D04	US 6821770 B1 (HOGAN JAMES J)	23.11.2004
D05	US 2006094034 A1 (BROUSSEAU ROLAND et al.)	04.05.2006

2. Declaración motivada según los artículos 29.6 y 29.7 del Reglamento de ejecución de la Ley 11/1986, de 20 de marzo, de Patentes sobre la novedad y la actividad inventiva; citas y explicaciones en apoyo de esta declaración

El objeto de la invención, tal y como se recoge en las reivindicaciones 1-4, es el uso de una mezcla de las proteínas GFP-hadrurina, GFP-gp12 y GFP-colicina-S4 para la detección de *Escherichia coli* en un líquido.

Novedad (art. 6.1 de la Ley 11/1986 de Patentes) y Actividad Inventiva (art. 8.1 de la Ley 11/1986 de Patentes).

Los documentos D01 y D02 se consideran los más cercanos del estado de la técnica. Estos documentos divulgan un método para la detección de *E. coli* basado en la utilización de la proteína SOC (small outer capsid) de la cápsida a la que se le une la proteína verde fluorescente (GFP). La muestra a analizar, se pone en contacto con el fago T4 e-/GFP, que además de llevar la proteína GFP unida a SOC (como proteína fusión), tiene inactivada la lisozima responsable de la lisis bacteriana. En el caso de que la muestra presente *E. coli*, ésta es infectada por el fago, pero no le provoca la lisis celular, con lo que la intensidad de la fluorescencia sólo es debida a la presencia de la bacteria.

El documento D03 divulga un sistema de detección de *E. coli*, que consiste en el fago lambda que tiene integrado el gen luxI y una célula que tiene un agente bioluminiscente basado en lux.

Los documentos D04 y D05 divulgan métodos para la detección de bacterias (entre ellas *E. coli*), mediante hibridación de ácidos nucleicos de una muestra, purificados y amplificados, con sondas presentes en un microarray.

Ninguno de los documentos citados, tomados solos o en combinación, revelan la invención definida en las reivindicaciones 1-4. Además, en los documentos citados no hay sugerencias que dirijan al experto en la materia hacia la invención definida por dichas reivindicaciones. Por lo tanto, el objeto de la invención, tal y como se recoge en las reivindicaciones 1-4, es nuevo en el sentido del art. 6.1 LP y se considera que implica actividad inventiva en el sentido del art. 8.1 LP.