



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 361 971**

51 Int. Cl.:
C12Q 1/68 (2006.01)
C07K 14/435 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **99966135 .8**
96 Fecha de presentación : **10.12.1999**
97 Número de publicación de la solicitud: **1135532**
97 Fecha de publicación de la solicitud: **26.09.2001**

54 Título: **Proteínas fluorescentes de especies sin bioluminiscencia de la clase de los antozoos, genes codificantes de tales proteínas y sus usos.**

30 Prioridad: **11.12.1998 US 210330**

45 Fecha de publicación de la mención BOPI:
24.06.2011

45 Fecha de la publicación del folleto de la patente:
24.06.2011

73 Titular/es: **CLONTECH LABORATORIES Inc.**
1020 East Meadow Drive
Palo Alto, California 94303, US

72 Inventor/es: **Lukyanoy, Sergey Anatolievich;**
Fradkov, Arcady Fedorovich;
Labas, Yulii Aleksandrovich y
Matz, Mikhail Vladimirovich

74 Agente: **Ponti Sales, Adelaida**

ES 2 361 971 T3

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

DESCRIPCIÓN

Proteínas fluorescentes de especies sin bioluminiscencia de la clase de los antozoos, genes codificantes de tales proteínas y sus usos.

5

ANTECEDENTES DE LA INVENCION

Campo de la invención

10 **[0001]** Esta invención se refiere al campo de la biología molecular. Más específicamente, esta invención se refiere a nuevas proteínas fluorescentes de especies sin bioluminiscencia de la clase de los antozoos, a procedimientos para la identificación de las secuencias de ADN que codifican las proteínas, a los genes que codifican dichas proteínas y a sus usos.

15 Descripción de la técnica relacionada

[0002] El marcaje fluorescente es una herramienta particularmente útil para marcar una proteína, célula u organismo de interés. Tradicionalmente, una proteína de interés se purifica y después se conjuga covalentemente con un derivado fluoróforo. Para los estudios *in vivo*, el complejo proteína-colorante se introduce después en las células de interés mediante micropipeteado o por un procedimiento de permeabilización reversible. Sin embargo, las etapas de unión y de introducción del colorante hacen que el procedimiento sea laborioso y difícil de controlar. Un procedimiento alternativo para el marcaje de proteínas de interés es la concatenación o fusión del gen que expresa la proteína de interés con un gen que expresa un marcador y la expresión posterior del producto de fusión. Los marcadores típicos para este procedimiento de marcaje de proteínas incluyen la β -galactosidasa, la luciferasa de luciérnaga y la luciferasa bacteriana. Sin embargo, estos marcadores requieren sustratos exógenos o cofactores y, por lo tanto, son de uso limitado para estudios *in vivo*.

20 **[0003]** Un marcador que no requiere un cofactor o sustrato exógeno es la proteína verde fluorescente (GFP) de la medusa *Aequorea victoria*, una proteína con un máximo de excitación a 395 nm, un segundo pico de excitación a 475 nm y un máximo de emisión a 510 nm. La GFP es una proteína de 238 aminoácidos, en que los aminoácidos 65-67 están implicados en la formación del cromóforo.

25 **[0004]** Los usos de la GFP para el estudio de la expresión génica y la localización de proteínas se discuten en detalle por Chalfie y col. en Science 263 (1994), 802-805 y Heim y col. en Proc. Natl. Acad. Sci. 91 (1994), 12501-12504. Adicionalmente, Rizzuto y col. discuten en Curr. Biology 5 (1995), 635-642, el uso de la GFP natural como herramienta de visualización de orgánulos subcelulares en células, mientras que Kaether y Gerdes describen en Febs Letters 369 (1995), 267-271, la visualización del transporte de proteínas a lo largo de la ruta de secreción mediante la GFP natural. La expresión de la GFP en células vegetales se discute por Hu y Cheng en Febs Letters 369 (1995), 331-334, mientras que la expresión de la GFP en embriones de *Drosophila* se describe por Davis y col. en Dev. Biology 170 (1995), 726-729.

30 **[0005]** Las estructuras cristalográficas de la GFP natural y de la GFP mutante S65T revelan que la estructura terciaria de la GFP se asemeja a un barril (Ormö y col., Science 273 (1996), 1392-1395; Yang y col., Nature Biotechnol. 14 (1996), 1246-1251). El barril consta de láminas β en una estructura compacta, en cuyo centro una hélice α que contiene el cromóforo queda protegida por el barril. La estructura compacta confiere gran estabilidad a la GFP en condiciones diversas y/o duras, como el tratamiento con proteasa, lo que hace que la GFP sea un indicador de gran utilidad, en general. Sin embargo, por su estabilidad, la GFP no es óptima para la determinación de sucesos a corto plazo o repetitivos.

35 **[0006]** Se está llevando a cabo gran cantidad de investigación para mejorar las propiedades de la GFP y para producir reactivos de GFP útiles y optimizados para diversos fines de investigación. Se han desarrollado nuevas versiones de la GFP, por ejemplo un ADN de GFP "humanizado" que produce una proteína cuya síntesis se incrementa en células de mamíferos (Haas y col., Current Biology 8 (1996), 315-324; Yang y col., Nucleic Acids Research 24 (1996), 4592-4593). Una proteína humanizada tal es la "proteína verde fluorescente mejorada" (EGFP). Otras mutaciones de la GFP han resultado en versiones que emiten luz azul, cian y verde amarillenta. Sin embargo, a pesar de la gran utilidad de la GFP, otras proteínas fluorescentes con propiedades similares o diferentes a las de la GFP serían de utilidad en la técnica. Nuevas proteínas fluorescentes resultan en la posibilidad de nuevos colores o la producción de fluorescencia dependiente del pH. Otros beneficios de proteínas fluorescentes nuevas incluyen las posibilidades de transferencia de energía por resonancia de fluorescencia (FERT) basadas en nuevos espectros

50

55

y una mejor adecuación para mayor excitación.

[0007] En su tesis de diploma “Zur Biologie der Ökotypen von *Anemonia sulcata* (Pennant)”, 1996, Jörg Wiedenmann describe sus intentos de clasificar *Anemonia sulcata* según su entorno. Al hacer esto, el autor examina el hábitat, el color de los tentáculos, el genotipo y la agresividad del comportamiento. Al examinar el color de los tentáculos, el autor determina que este se debe en parte a zooxantelas y en parte a los colores propios de la anémona, que el autor describe como proteínas fluorescentes fijadas genéticamente.

[0008] La técnica anterior es deficiente en proteínas fluorescentes nuevas de las que se conozcan las secuencias codificantes de ADN. La presente invención satisface esta necesidad existente desde hace mucho tiempo en la técnica.

RESUMEN DE LA INVENCION

[0009] La presente invención se dirige a un ácido nucleico en un entorno distinto de su entorno natural, cuya secuencia codifica una proteína fluorescente, en que dicha proteína fluorescente es una proteína de un organismo antozoo sin bioluminiscencia seleccionado de los géneros *Anemonia*, *Clavularia*, *Zoanthus* y *Discosoma* y en que:

(a) dicha proteína fluorescente es de *Anemonia majano* y dicho ácido nucleico puede amplificarse con un par de cebadores, el primero de los cuales consta de la secuencia de SEQ ID NO:36 y el segundo de los cuales consta de la secuencia de SEQ ID NO:37; o

(b) dicha proteína fluorescente es de *Clavularia sp.* y dicho ácido nucleico puede amplificarse con un par de cebadores compuesto de un primer cebador que consta de la secuencia de SEQ ID NO:38 y un segundo cebador que consta de la secuencia de SEQ ID NO:40; o

(c) dicha proteína fluorescente es de *Zoanthus sp.* y dicho ácido nucleico puede amplificarse con un par de cebadores compuesto de un primer cebador que consta de la secuencia de SEQ ID NO:41 y un segundo cebador que consta de la secuencia de SEQ ID NO:42; o

(d) dicha proteína fluorescente es de *Discosoma sp.* “roja” y dicho ácido nucleico puede amplificarse con un par de cebadores compuesto de un primer cebador que consta de la secuencia de SEQ ID NO:43 y un segundo cebador que consta de la secuencia de SEQ ID NO:44; o

(e) dicha proteína fluorescente es de *Discosoma striata* y dicho ácido nucleico puede amplificarse con un par de cebadores compuesto de un primer cebador que consta de la secuencia de SEQ ID NO:45 y un segundo cebador que consta de la secuencia de SEQ ID NO:46.

[0010] Se describe un procedimiento para la identificación de una secuencia de ADN codificante de una proteína fluorescente que comprende la etapa de cribado en busca de la existencia de una secuencia de ácido nucleico en una muestra, en que la secuencia de ácido nucleico codifica un péptido con una secuencia seleccionada del grupo que consta de SEQ ID NO:3, 5, 8, 11, 12 y 14. La existencia de la secuencia de ácido nucleico identifica la secuencia de ADN codificante de la proteína fluorescente.

[0011] También se describe un procedimiento para la identificación de una secuencia de ADN codificante de una proteína fluorescente que comprende la etapa de cribado en busca de la existencia de una secuencia de ácido nucleico en una muestra, en que la secuencia de ácido nucleico hibrida con un cebador seleccionado del grupo que consta de SEQ ID NO:4, 6, 7, 9, 10, 13, 15 y 16. La existencia de la secuencia de ácido nucleico identifica la secuencia de ADN codificante de la proteína fluorescente.

[0012] También se describe un procedimiento para el análisis de una proteína fluorescente en una célula que comprende las etapas de la expresión en la célula de una secuencia de ácido nucleico codificante de una proteína fluorescente con una secuencia de aminoácidos seleccionada del grupo que consta de SEQ ID NO:55-63; y la medida de una señal fluorescente procedente de la proteína. Este procedimiento comprende además una etapa de clasificación de la célula según la señal. Preferentemente, la célula se clasifica por clasificación de células activadas por fluorescencia. Con mayor preferencia, la secuencia de ácido nucleico comprende un gen de interés codificante de una proteína de interés fusionada con la proteína fluorescente, en que la proteína de interés es distinta de la proteína fluorescente. La señal fluorescente detectada indica la presencia en la célula del gen de interés y además de la proteína de interés. Al identificar una localización intracelular de la proteína fluorescente también se identifica una localización intracelular de la proteína de interés.

[0013] Otros aspectos, características y ventajas adicionales de la presente invención se harán evidentes a partir de la siguiente descripción de las realizaciones de la invención preferidas actualmente que se indican con el fin de su desvelado.

BREVE DESCRIPCIÓN DE LOS DIBUJOS**[0014]**

La figura 1 muestra la estrategia de RACE 3' modificada usada para aislar los fragmentos diana. Las secuencias de los oligonucleótidos usados se muestran en la tabla 2. Dp1 y Dp2 son los cebadores degenerados usados en la primera y en la segunda reacción de PCR, respectivamente (véanse las secuencias de los cebadores degenerados en las tablas 3 y 4).

La figura 2A muestra un alineamiento múltiple de proteínas fluorescentes nuevas. La numeración se basa en la proteína fluorescente verde (GFP) de *Aequorea victoria*. Se comparan entre sí dos proteínas de *Zoanthus* y cuatro de *Discosoma*: los restos idénticos a los correspondientes en la primera proteína de la serie se representan por guiones. Los huecos introducidos se representan por puntos. En la secuencia de la GFP de *A. victoria* se subrayan los tramos que forman láminas β ; los restos cuyas cadenas laterales forman el interior del barril β están sombreados (según Yang y col., Nature Biotechnol. 14, 1246-1251 (1996)). La figura 2B muestra la parte N-terminal de cFP484, que no tiene homología con las otras proteínas. El supuesto péptido señal está subrayado.

La figura 3 muestra el espectro de excitación y emisión de la proteína fluorescente nueva de *Anemonia majano* amFP486.

La figura 4 muestra el espectro de excitación y emisión de la proteína fluorescente nueva de *Clavularia* cFP484.

La figura 5 muestra el espectro de excitación y emisión de la proteína fluorescente nueva de *Zoanthus* zFP506.

La figura 6 muestra el espectro de excitación y emisión de la proteína fluorescente nueva de *Zoanthus* zFP538.

La figura 7 muestra el espectro de excitación y emisión de la proteína fluorescente nueva de *Discosoma striata* dsFP483.

La figura 8 muestra el espectro de excitación y emisión de la proteína fluorescente nueva de *Discosoma* drFP583.

La figura 9 muestra el espectro de excitación y emisión de la proteína fluorescente nueva de *Anemonia sulcata* asFP600. Esta proteína no es realización de la presente invención.

La figura 10 muestra el espectro de excitación y emisión de la proteína fluorescente nueva de *Discosoma* dgFP512. Esta proteína no es realización de la presente invención.

La figura 11 muestra el espectro de excitación y emisión de la proteína fluorescente nueva de *Discosoma* dmFP592. Esta proteína no es realización de la presente invención.

DESCRIPCIÓN DETALLADA DE LA INVENCION

Según se usa en este documento, el término "GFP" se refiere a la proteína fluorescente verde básica de *Aequorea victoria*, incluyendo versiones anteriores de la GFP de la técnica, manipuladas para producir mayor fluorescencia o fluorescencia de colores diferentes. La secuencia de la GFP de *Aequorea victoria* (SEQ ID NO:54) ha sido desvelada en Prasher y col., Gene 111 (1992), 229-33.

Según se usa en este documento, el término "EGFP" se refiere a una variante mutante de la GFP que tiene dos sustituciones de aminoácidos: F64L y S65T (Heim y col., Nature 373 (1995), 663-664). El término "humanizada" se refiere a cambios en la secuencia del ácido nucleico de la GFP para optimizar los codones para la expresión de la proteína en células humanas (Yang y col., Nucleic Acids Research 24 (1996), 4592-4593).

De acuerdo con la presente invención pueden emplearse técnicas convencionales de biología molecular, microbiología y ADN recombinante dentro de la experiencia de la técnica. Estas técnicas se explican detalladamente en la bibliografía. Véase, por ejemplo, Maniatis, Fritsch y Sambrook, "Molecular Cloning: A Laboratory Manual (1982); "DNA cloning: A Practical Approach, volúmenes I y II (D. N. Glover ed. 1985); "Oligonucleotide Synthesis (M. J. Gait ed. 1984); "Nucleic Acid Hybridization" (B. D. Hames y S. J. Higgins eds. 1985); "Transcription and Translation" (B. D. Hames y S. J. Higgins eds. 1984); "Animal Cell Culture" (R. I. Freshney ed. 1986); "Immobilized Cells and Enzymes" (IRL Press 1986); B. Perbal, "A Practical Guide to Molecular Cloning" (1984).

Un vector es un replicón, como un plásmido, fago o cósmido al que puede unirse otro segmento de ADN para conseguir la replicación del segmento unido.

Una molécula de ADN se refiere a la forma polimérica de los desoxirribonucleótidos (adenina, guanina, timina o citosina), bien en forma monocatenaria o como hélice bicatenaria. Este término se refiere solamente a la

estructura primaria y secundaria de la molécula y no la limita a ninguna forma terciaria concreta. Por lo tanto, este término incluye el ADN bicatenario que se encuentra, entre otros, en moléculas lineales de ADN (por ejemplo, fragmentos de restricción), virus, plásmidos y cromosomas.

5 **[0020]** Una “secuencia codificante” de ADN es una secuencia de ADN que se transcribe y se traduce en un polipéptido *in vivo* al colocarla bajo el control de las secuencias de regulación apropiadas. Los límites de la secuencia codificante están determinados por un codón de inicio en el extremo 5' (amino) y un codón de parada de la traducción en el extremo 3' (carboxilo). Una secuencia codificante puede incluir, pero no se limita a secuencias procarionóticas, ADNc de ARNm eucariótico, secuencias de ADN genómico de ADN eucariótico (por ejemplo, de mamíferos) y secuencias de ADN sintéticas. Una señal de poliadenilación y una secuencia de terminación de la transcripción pueden localizarse en el lado 3' de la secuencia codificante.

10 **[0021]** Según se usa en este documento, el término “hibridación” se refiere al proceso de asociación de dos hebras de ácido nucleico para formar un dúplex antiparalelo estabilizado por medio de puentes de hidrógeno entre restos de las hebras opuestas de ácido nucleico.

15 **[0022]** El término “oligonucleótido” se refiere a una molécula de poca longitud (menos de 100 bases) de ácido nucleico.

20 **[0023]** Las “secuencias de regulación de ADN” según se usan en este documento son secuencias de control transcripcional y traduccional, como promotores, potenciadores, señales de poliadenilación, terminadores y similares, que producen y/o regulan la expresión de una secuencia codificante en una célula hospedadora.

25 **[0024]** Una “secuencia promotora” es una región de regulación de ADN capaz de unirse a la polimerasa de ARN en una célula e iniciar la transcripción de una secuencia codificante situada secuencia abajo (en la dirección 3'). Para el fin de la definición de la presente invención, la secuencia promotora está limitada en su extremo 3' por el sitio de inicio de la transcripción y se extiende secuencia arriba (en la dirección 5') para incluir el mínimo número de bases o elementos necesarios para iniciar la transcripción a niveles detectables por encima del nivel de fondo. Dentro de la secuencia promotora se encontrará un sitio de inicio de la transcripción, así como dominios de unión a proteínas responsables de la unión a la polimerasa de ARN. Los promotores eucarióticos contendrán frecuentemente, pero no siempre, cajas “TATA” y cajas “CAT”. Para el funcionamiento de los diversos vectores de la presente invención pueden usarse diversos promotores, incluyendo promotores inducibles.

30 **[0025]** Según se usa en este documento, los términos “endonucleasas de restricción” y “enzimas de restricción” se refieren a enzimas bacterianas, cada una de las cuales corta ADN bicatenario en una secuencia específica o cerca de esta.

35 **[0026]** Una célula ha sido “transformada” o “transfectada” por un ADN exógeno o heterólogo cuando dicho ADN ha sido introducido dentro de la célula. El ADN transformante puede integrarse o no (ligado covalentemente) en el genoma de la célula. En procarionotas, levaduras y células de mamíferos, por ejemplo, el ADN transformante puede mantenerse en un elemento episómico como un plásmido. Con respecto a las células eucariotas, una célula transformada de manera estable es aquella en la que el ADN transformante se ha integrado en un cromosoma, de modo que se hereda por las células hijas a través de la replicación cromosómica. Esta estabilidad se demuestra por la capacidad de la célula eucariota de establecer líneas celulares o clones compuestos por una población de células hijas que contienen el ADN transformante. Un “clon” es una población de células derivadas de una única célula o antepasado común por mitosis. Una “línea celular” es un clon de una célula primaria que es capaz de un crecimiento estable *in vitro* durante numerosas generaciones.

40 **[0027]** Una región “heteróloga” de la construcción de ADN es un segmento identificable de ADN dentro de una molécula de ADN de mayor tamaño, que no se encuentra asociado con la molécula de mayor tamaño en la naturaleza. Por lo tanto, cuando la región heteróloga codifica un gen de mamífero, el gen estará flanqueado habitualmente por ADN que no flanquea el ADN genómico de mamífero en el genoma del organismo fuente. En otro ejemplo, el ADN heterólogo incluye una secuencia codificante en una construcción en la que la se han unido entre sí partes de genes de dos fuentes diferentes para producir como producto una proteína de fusión. Las variaciones alélicas de mutaciones de origen natural no dan lugar a una región heteróloga de ADN según se define en este documento.

55 **[0028]** Según se usa en este documento, el término “gen indicador” se refiere a una secuencia codificante unida a un promotor heterólogo o a elementos potenciadores y cuyo producto puede analizarse fácilmente y de

manera cuantificable al introducir la construcción en tejidos o células.

[0029] Se prefiere que los aminoácidos descritos en este documento estén en la forma isomérica "L". Las secuencias de aminoácidos se indican en el código de una letra (A: alanina; C: cisteína; D: ácido aspártico; E: ácido glutámico; F: fenilalanina; G: glicina; H: histidina; I: isoleucina; K: lisina; L: leucina; M: metionina; N: asparragina; P: prolina; Q: glutamina; R: arginina; S: serina; T: treonina; V: valina; W: triptofano; Y: tirosina; X: cualquier resto). NH₂ se refiere al grupo amino libre presente en el extremo amino de un polipéptido. COOH se refiere al grupo carboxilo libre presente en el extremo carboxilo de un polipéptido. Se usa J. Biol. Chem. 243 (1969) 3552-59 para atenerse a la nomenclatura estándar de polipéptidos.

[0030] La presente invención se dirige a una proteína fluorescente aislada y purificada, seleccionada del grupo que consta de amFP486, cFP484, zFP506, zFP538, dsFP483 y drFP583.

[0031] En este documento se describe un procedimiento para la identificación de una secuencia de ADN codificante de una proteína fluorescente que comprende la etapa del cribado en busca de la existencia de una secuencia de ácido nucleico en una muestra, en que la secuencia de ácido nucleico codifica un péptido que tiene una secuencia seleccionada del grupo que consta de SEQ ID NO:3, 5, 8, 11, 12 y 14. La existencia de la secuencia de ácido nucleico identifica la secuencia de ADN codificante de la proteína fluorescente.

[0032] También se describe un procedimiento para la identificación de una secuencia de ADN codificante de una proteína fluorescente que comprende la etapa del cribado en busca de la existencia de una secuencia de ácido nucleico en una muestra, en que la secuencia de ácido nucleico hibrida con un cebador seleccionado del grupo que consta de SEQ ID NO:4, 6, 7, 9, 10, 13, 15 y 16. La existencia de la secuencia de ácido nucleico identifica la secuencia de ADN codificante de la proteína fluorescente.

[0033] Además se describe un procedimiento para el análisis de una proteína fluorescente en una célula que comprende las etapas de la expresión en la célula de una secuencia de ácido nucleico codificante de una proteína fluorescente con una secuencia de aminoácidos seleccionada del grupo compuesto de SEQ ID NO:55-63; y la medida de una señal fluorescente procedente de la proteína. Este procedimiento comprende además una etapa de clasificación de la célula según la señal. Preferentemente, la célula se clasifica por clasificación de células activadas por fluorescencia. Con mayor preferencia, la secuencia de ácido nucleico comprende un gen de interés codificante de una proteína de interés fusionada con la proteína fluorescente, en que la proteína de interés es distinta de la proteína fluorescente. La señal fluorescente detectada indica la presencia en la célula del gen de interés y además de la proteína de interés. Al identificar una localización intracelular de la proteína fluorescente también se identifica una localización intracelular de la proteína de interés.

[0034] Los ejemplos siguientes se incluyen con el fin de ilustrar diversas realizaciones de la invención y no se pretende que sean limitantes de la presente invención en ningún modo.

Material biológico

[0035] Se identificaron nuevas proteínas fluorescentes de varios géneros de antozoos que no presentan ninguna bioluminiscencia pero tienen color fluorescente al observarlos con la luz blanca habitual o con luz ultravioleta. Se eligieron seis especies (véase la tabla 1).

TABLA 1

Especies de antozoos usadas en este estudio		
Especie	Área de origen	Color fluorescente
<i>Anemonia majano</i>	Pacífico occidental	puntas de los tentáculos verde brillante
<i>Clavularia sp</i>	Pacífico occidental	tentáculos y disco oral verde brillante
<i>Zoanthus sp.</i>	Pacífico occidental	tentáculos y disco oral verde amarillento
<i>Discosoma sp. "roja"</i>	Pacífico occidental	manchas rojo anaranjadas en el disco oral
<i>Discosoma striata</i>	Pacífico occidental	bandas verde azuladas en el disco oral
<i>Discosoma sp. "magenta"</i>	Pacífico occidental	disco oral ligeramente púrpura
<i>Discosoma sp. "verde"</i>	Pacífico occidental	manchas verdes en el disco oral
<i>Anemonia sulcata</i>	Mediterráneo	puntas de los tentáculos púrpura

EJEMPLO 2Preparación de ADNc

5 **[0036]** El ARN total de la especie de interés se aisló según el protocolo de Chomczynski y Sacchi (Chomczynski P. y col., Anal. Biochem. 162 (1987), 156-159). La primera hebra de ADNc se sintetizó con 1-3 µg de ARN total mediante el kit de síntesis de ADNc SMART PCR (CLONTECH) según el protocolo suministrado, con la única alteración de que el “cebador para síntesis de ADNc” suministrado en el kit se sustituyó por el cebador TN3 (5'-CGCAGTCGACCG(T)₁₃, SEQ ID NO:1) (tabla 2). Después se prepararon muestras de ADNc amplificado según se describe en el protocolo suministrado, con la excepción de que los dos cebadores usados para PCR fueron el cebador TS (5'-AAGCAGTGGTATCAACGCAGAGT, SEQ ID NO:2) (tabla 2) y el cebador TN3 (tabla 2), los dos en la concentración de 0,1 µM. Se llevaron a cabo de veinte a veinticinco ciclos de PCR para la amplificación de una muestra de ADNc. El ADNc amplificado se diluyó veinte veces en agua y en los procedimientos siguientes se usó 1 µl de esta dilución.

15

TABLA 2

Oligonucleótidos usados para la síntesis de ADNc y RACE

TN3:	5'-CGCAGTCGACCG(T) ₁₃ (SEQ ID No. 1)
T7-TN3 :	5'-GTAATACGACTCACTATAGGGCCGCAGTCGACCG(T) ₁₃ (SEQ ID No. 17)
Cebador TS:	5'-AAGCAGTGGTATCAACGCAGAGT (SEQ ID No. 2)
T7-TS:	5'-GTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGT (SEQ ID No. 18)
T7:	5'-GTAATACGACTCACTATAGGGC (SEQ ID No. 19)
Oligonucleótido TS	5'-AAGCAGTGGTATCAACGCAGAGTACGCrGrGrG (SEQ ID No. 53)

20

EJEMPLO 3Diseño de oligonucleótidos

25 **[0037]** Para aislar fragmentos de ADNc de proteínas fluorescentes nuevas se llevaron a cabo reacciones de PCR con cebadores degenerados. Los cebadores degenerados se diseñaron para coincidir con la secuencia del ARNm de las regiones predichas como de máxima invariabilidad en la familia de las proteínas fluorescentes. Se eligieron cuatro de estos tramos (tabla 3) y se diseñaron variantes de cebadores degenerados. Todos estos cebadores se dirigieron al extremo 3' del ARNm. Todos los oligonucleótidos se purificaron en gel antes de su uso. La tabla 2 muestra los oligonucleótidos usados para la síntesis de ADNc y RACE.

30

TABLA 3

Tramos de aminoácidos clave y cebadores degenerados correspondientes usados para el aislamiento de proteínas fluorescentes		
Posición del tramo según la GFP de <i>A. victoria</i> (7)	Secuencia de aminoácidos del tramo clave	Nombre y secuencia del cebador degenerado
20-25	GXVNGH (SEQ ID No. 3)	NGH: 5'- GA(C,T) GGC TGC GT(A,T,G,C) AA(T,C) GG(A,T,G) CA (SEQ ID No. 4)
31-35	GEGEG (SEQ ID No. 5) GEGNG (SEQ ID No. 8)	GEGa: 5'- GTT ACA GGT GA(A,G) GG(A,C) GA(A,G) GG (SEQ ID No. 6) GEGb: 5'- GTT ACA GGT GA(A,G) GG(T,G) GA(A,G) GG (SEQ ID No. 7) GNGa: 5'- GTT ACA GGT GA(A,G) GG(A,C) AA(C,T) GG (SEQ ID No. 9) GNGb: 5'- GTT ACA GGT GA(A,G) GG(T,G) AA(C,T) GG (SEQ ID No. 10)
127-131	GMNFP (SEQ ID No. 11) GVNFP (SEQ ID No. 12)	NFP: 5' TTC CA(C,T) GGT (G,A)TG AA(C,T) TT(C,T) CC (SEQ ID NO. 13)
134-137	GPVM (SEQ ID No. 14)	PVMa: 5' CCT GCC (G,A)A(C,T) GGT CC(A,T,G,C) GT(A,C) ATG (SEQ ID NO. 15) PVMb: 5' CCT GCC (G,A)A(C,T) GGT CC(A,T,G,C) GT(G,T) ATG (SEQ ID NO. 16)

5 EJEMPLO 4Aislamiento de fragmentos 3' de ADNc de las nFP

[0038] Para el aislamiento de los fragmentos diana se usó la estrategia de RACE 3' modificada (véase la figura 1). La estrategia de RACE requirió dos etapas de PCR consecutivas. En la primera etapa de PCR se usaron un primer cebador degenerado (tabla 4) y el cebador T7-TN3 (SEQ ID NO:17) que tiene una parte 3' idéntica al cebador TN3 usado para la síntesis de ADNc (véase la secuencia de T7-TN3 en la tabla 2). La razón de la sustitución por el cebador de mayor longitud T7-TN3 en esta etapa de PCR fue la supresión efectiva de la amplificación de fondo que tiene lugar al usar el cebador de menor longitud TN3, en particular cuando el cebador T7-TN3 se usa a baja concentración (0,1 μ M) (Frohman y col., (1998) PNAS USA 85, 8998-9002). En la segunda etapa de PCR se usaron el cebador TN3 (SEQ ID NO:1, tabla 2) y un segundo cebador degenerado (tabla 4).

TABLA 4

Combinaciones de cebadores degenerados para la primera y segunda PCR que resultan en la amplificación específica de fragmentos 3' de ADNc de las nFP

Especie	Primer cebador degenerado	Segundo cebador degenerado
Anemonia majano	NGH (SEQ ID No. 4)	GNGb (SEQ ID No. 10)
Clavularia sp.	NGH (SEQ ID No. 4)	GEGa (SEQ ID No. 6)
Zoanthus sp.	NGH (SEQ ID No. 4)	GEGa (SEQ ID No. 6)
Discosoma sp. "roja"	NGH (SEQ ID No. 4)	GEGa (SEQ ID No. 6), NFP (SEQ ID No. 13) or PVMb (SEQ ID No. 16)
Discosoma striata	NGH (SEQ ID No. 4)	NFP (SEQ ID No. 13)
Anemonia sulcata	NGH (SEQ ID No. 4)	GEGa (SEQ ID No. 6) or NFP (SEQ ID No. 13)

5 **[0039]** La primera reacción de PCR se realizó de la manera siguiente: 1 µl de la muestra de ADNc amplificado diluida 20 veces se añadió a la mezcla de reacción que contenía la mezcla de la polimerasa Advantage KlenTaq 1x con el tampón suministrado (CLONTECH), dNTP 200 µM, el primer cebador degenerado a una concentración de 0,3 µM (tabla 4) y el cebador T7-TN3 (SEQ ID NO:17) 0,1 µM en un volumen total de 20 µl. Los ciclos realizados fueron (termociclador Hybaid OmniGene, modo de control de tubo): un ciclo de 95°C, 10 s; 55°C, 1 min; 72°C, 40 s; 24 ciclos de 95°C, 10 s; 62°C, 30 s; 72°C, 40 s. La reacción se diluyó después 20 veces en agua y 1 µl de esta dilución se añadió a una segunda reacción de PCR que contenía la mezcla de la polimerasa Advantage KlenTaq 1x con el tampón suministrado por el fabricante (CLONTECH), dNTP 200 µM, el segundo cebador degenerado a una concentración de 0,3 µM (tabla 4) y el cebador TN3 0,1 µM. Los ciclos realizados fueron (termociclador Hybaid OmniGene, modo de control de tubo): un ciclo de 95°C, 10 s; 55°C (para GEG/GNG o PVM) o 52°C (para NFP), 1 min; 72°C, 40 s; 13 ciclos de 95°C, 10 s; 62°C (para GEG/GNG o PVM) o 58°C (para NFP), 30 s; 72°C, 40 s. El producto de la reacción de PCR se clonó en el vector PCR-Script (Stratagene) según el protocolo del fabricante.

20 **[0040]** Se probaron diferentes combinaciones de cebadores degenerados en las reacciones primera y segunda de PCR con el ADN de cada especie hasta encontrar una combinación que resultó en una amplificación específica, lo que indica que se detectó una banda intensa del tamaño esperado (aproximadamente 650-800 pb para NGH y GEG/GNG y 350-500 pb para NFP y PVM, a veces acompañadas de bandas menores) en un gel de agarosa después de las dos reacciones de PCR. Las combinaciones de cebadores elegidas para las diferentes especies de la clase de los antozoos se listan en la tabla 4. Algunas otras combinaciones de cebadores también resultaron en la amplificación de fragmentos del tamaño correcto, pero la secuencia de estos fragmentos no mostró ninguna homología con las otras proteínas fluorescentes identificadas o con la GFP de *Aequorea victoria*.

EJEMPLO 5

Obtención de copias de ADNc de longitud completa

30 **[0041]** Después de la secuenciación de los fragmentos 3' obtenidos de los ADNc de las nuevas proteínas fluorescentes se sintetizaron dos cebadores anidados dirigidos al extremo 5' por cada ADNc (tabla 5) y los extremos 5' de los ADNc se amplificaron entonces mediante dos reacciones de PCR consecutivas. En la siguiente reacción de PCR se usó la nueva estrategia de "step-out PCR" para suprimir la amplificación de fondo. La mezcla de la reacción de "step-out" contenía la mezcla de la polimerasa Advantage KlenTaq 1x con el tampón suministrado por el fabricante (CLONTECH), dNTP 200 µM, el primer cebador específico del gen (véase la tabla 5) 0,2 µM, el cebador T7-TS (SEQ ID NO:18) 0,02 µM, el cebador T7 (SEQ ID NO:19) 0,1 µM y 1 µl de la muestra de ADNc amplificado diluida 20 veces en un volumen total de 20 µl. Los ciclos realizados fueron (termociclador Hybaid OmniGene, modo

de control de tubo): 23-27 ciclos de 95°C, 10 s; 60°C, 30 s; 72°C, 40 s. El producto de la amplificación se diluyó 50 veces en agua y 1 µl de esta dilución se añadió a la segunda reacción de PCR (anidada). La reacción contenía la mezcla de la polimerasa Advantage KlenTaq 1x con el tampón suministrado (CLONTECH), dNTP 200 µM, el segundo cebador específico del gen 0,2 µM y el cebador TS (SEQ ID NO:2) 0,1 µM en un volumen total de 20 µl. Los ciclos realizados fueron (termociclador Hybaid OmniGene, modo de control de tubo): 12 ciclos de 95°C, 10 s; 60°C, 30 s; 72°C, 40 s. Después el producto de la amplificación se clonó en el vector pAtlas (CLONTECH) según el protocolo del fabricante.

[0042] Se señala que las tres últimas líneas de la tabla 5 no se refieren a realizaciones de la presente invención.

TABLA 5

Cebadores específicos del gen usados para RACE 5'		
Especie	Primer cebador	Segundo cebador (anidado)
Anemonia majano	5'-GAAATAGTCAGGCATACTGGT (SEQ ID No. 20)	5'-GTCAGGCATAC TGGTAGGAT (SEQ ID No. 21)
Clavularia sp.	5'-CTTGAAATAGTCTGCTATATC (SEQ ID No. 22)	5'-TCTGCTATATC GTCTGGGT (SEQ ID No. 23)
Zoanthus sp.	5'- GTTCTTGAAATAGTCTACTATGT (SEQ ID No. 24)	5'-GTCTACTATGTCTT GAGGAT (SEQ ID No. 25)
Discosoma sp. "roja"	5'-CAAGCAAATGGCAAAGGTC (SEQ ID No. 26)	5'-CGGTATTGTGGCC TTCGTA (SEQ ID No. 27)
Discosoma striata	5'-TTGTCTTCTTCTGCACAAC (SEQ ID No. 28)	5'-CTGCACAACGG GTCCAT (SEQ ID No. 29)
Anemonia sulcata	5'-CCTCTATCTTCATTTCCCTGC (SEQ ID No. 30)	5'-TATCTTCATTTCCCT GCGTAC (SEQ ID No. 31)
Discosoma sp. "magenta"	5'-TTCAGCACCCCATCACGAG (SEQ ID No. 32)	5'-ACGCTCAGAGCTG GGTTCC (SEQ ID No. 33)
Discosoma sp. "verde"	5'-CCCTCAGCAATCCATCACGTTTC (SEQ ID No. 34)	5'-ATTATCTCAGTGGA TGGTTC (SEQ ID No. 35)

EJEMPLO 6Expresión de las nFP en *E. coli*

- 5 **[0043]** Para preparar una construcción de ADN para la expresión de las nuevas proteínas fluorescentes se sintetizaron dos cebadores para cada ADNc: un cebador “secuencia abajo” dirigido al extremo 5’ con el sitio de hibridación localizado en la región UTR 3’ del ADNc y un cebador “secuencia arriba” dirigido al extremo 3’ correspondiente al sitio de inicio de la traducción (sin incluir el primer codón ATG) (tabla 6). Los dos cebadores tenían extensiones 5’ codificantes de un sitio para una endonucleasa de restricción; además, el cebador secuencia arriba se diseñó para permitir la clonación del producto de PCR en el vector pQE30 (Quiagen), de tal manera que resultó en la fusión de los marcos de lectura de la etiqueta de 6xHis codificada por el vector y la proteína. La reacción de PCR se realizó de la manera siguiente: 1 µl de la muestra de ADNc amplificado diluida 20 veces se añadió a una mezcla que contenía la mezcla de la polimerasa Advantage KlenTaq 1x con el tampón suministrado por el fabricante (CLONTECH), dNTP 200 µM, el cebador secuencia arriba 0,2 µM y el cebador secuencia abajo 0,2 µM en un volumen final total de 20 µl. Los ciclos realizados fueron (termociclador Hybaid OmniGene, modo de control de tubo): 23-27 ciclos de 95°C, 10 s; 60°C, 30 s; 72°C, 40 s. El producto de esta etapa de amplificación se purificó por extracción con fenol-cloroformo y precipitación con etanol y después se clonó en el vector pQE30 mediante las endonucleasas de restricción correspondientes a la secuencia de los cebadores según protocolos estándar.
- 10
- 15
- 20 **[0044]** Todos los plásmidos se amplificaron en *E. coli* XL-1 Blue y se purificaron mediante kits de minipreparación de ADN plasmídico (CLONTECH). Los clones recombinantes se seleccionaron por el color de las colonias y se cultivaron en 3 ml de medio LB (suplementado con 100 µg/ml de ampicilina) a 37°C durante la noche. Un volumen de 100 µl del cultivo incubado durante la noche se transfirió a 200 ml de medio LB recién preparado que contenía 100 µg/ml de ampicilina y se incubó a 37°C, 200 rpm hasta una DO₆₀₀ de 0,6-0,7. Después se añadió IPTG 1 mM al cultivo y se continuó la incubación a 37°C durante otras 16 horas. Las células se recogieron y la proteína recombinante, que incorporaba etiquetas de 6xHis en el extremo N se purificó por medio de la resina de afinidad metálica TALON™ según el protocolo del fabricante (CLONTECH).
- 25
- 30 **[0045]** Se señala que las últimas tres líneas de la tabla 6 no se refieren a realizaciones de la presente invención.

TABLA 6

Cebadores usados para la obtención de la región codificante de longitud completa de las nFP para su clonación en una construcción de expresión	
Especie	Cebador secuencia arriba Cebador secuencia abajo
Aneimonia majano	5' -acatggatccgctcttcaaca agtttate (SEQ ID No. 36) BamHI 5' -tagtactcggagcttattcgtgta tttcagtgaaatc (SEQ ID No. 37) XhoI
Clavulania sp.	L: 5' -acatggatccaacattttttga gaaacg (SEQ ID No. 38) BamHI S: 5' -acatggatccaagctetaacc accatg (SEQ ID No. 39) BamHI 5' -tagtactcggagcaacacaa acctcagacaa (SEQ ID No. 40) XhoI
Zoanthus sp.	5' - acatggatccgctcagtcgaag cacggt (SEQ ID No. 41) BamHI 5' -tagtactcggaggttgaactacat tattata (SEQ ID No. 42) XhoI
Discosoma sp. "roja"	5' - acatggatccagggtcttccaagaat gttate (SEQ ID No. 43) BamHI 5' -tagtactcggagggagccaagttc agcctta (SEQ ID No. 44) XhoI
Discosoma striata	5' - acatggatccagttggtccaagagtggtg (SEQ ID No. 45) BamHI 5' -tagcggagcctctatcgcctc gtcacct (SEQ ID No. 46) SacI
Aneimonia sulcata	5' - acatggatccgctctttaaagaagact (SEQ ID No. 47) BamHI 5' -tagtactcggagtccttgggagc ggcttg (SEQ ID No. 48) XhoI

Cebadores usados para la obtención de la región codificante de longitud completa de las nFP para su clonación en una construcción de expresión

Especie	Cebador secuencia arriba	Cebador secuencia abajo
Discosoma sp. "magenta"	5'- <u>aca</u> tgga <u>lcc</u> agttgttccaaagaatgtgat (SEQ ID No. 49) BamHI	5'-tagtactc <u>gag</u> ggccattacg ctaatc (SEQ ID No. 50) XhoI
Discosoma sp. "verde"	5'-aca <u>tgga</u> lccagtgcaacttaagaagaatg (SEQ ID No. 51)	5'-tagtactc <u>gag</u> gattcggtttaat gccttg (SEQ ID No. 52)

EJEMPLO 7Nuevas proteínas fluorescentes y ADNc codificantes de estas proteínas

- 5 **[0046]** Se encontraron nueve ADNc de longitud completa codificantes de proteínas fluorescentes y se produjeron nueve proteínas fluorescentes nuevas (SEQ ID NO:55-63). Las propiedades espectrales de las nuevas proteínas fluorescentes aisladas se muestran en la tabla 7 y los espectros de emisión y de excitación de las nuevas proteínas se muestran en las figuras 3-11.
- 10 **[0047]** Se señala que las proteínas asFP600, dgFP512 y dmFP592 (SEQ ID NO:61-63) no son realizaciones de la presente invención.

TABLA 7

Propiedades espectrales de las nFP aisladas						
Especie	Nombre de la proteína nFP	Máximo de absorbancia nm	Máximo de emisión nm	Coefficiente de extinción máximo	Rendimiento cuántico relativo*	Luminosidad relativa**
Anemonia majano	amFP486	458	486	40.000	0,3	0,43
Clavularia sp.	cFP484	456	484	35.300	0,6	0,77
Zoanthus sp.	zFP506	496	506	35.600	0,79	1,02
Zoanthus sp.	zFP538	528	538	20.200	0,52	0,38
Discosoma sp. "roja"	drFP583	558	583	22.500	0,29	0,24
Discosoma striata	dsFP483	443	483	23.900	0,57	0,50
Anemonia sulcata	asFP600	572	596	56.200	<0,001	
Discosoma sp. "verde"	dgFP512	502	512	20.360	0,3	0,21
Discosoma sp. "magenta"	dmFP592	573	593	21.800	0,11	0,09

* El rendimiento cuántico relativo se determinó por comparación con el rendimiento cuántico de la GFP de *A. victoria*.

** La luminosidad relativa es el coeficiente de extinción multiplicado por el rendimiento cuántico y dividido por el mismo valor para la GFP de *A. victoria*.

- 15
- 20 **[0048]** El alineamiento múltiple de las proteínas fluorescentes se muestra en la figura 2A. La numeración se basa en la proteína verde fluorescente de *Aequorea victoria* (GFP, SEQ ID NO:54). Las secuencias de aminoácidos de las proteínas fluorescentes nuevas se marcan como SEQ ID NO:55-63. Se comparan entre sí dos proteínas de *Zoanthus* y cuatro de *Discosoma*: los restos idénticos a los correspondientes en la primera proteína de la serie se representan por guiones. Los huecos introducidos se representan por puntos. En la secuencia de la GFP de *A. victoria* se subrayan los tramos que forman láminas β ; los restos cuyas cadenas laterales forman el interior del barril β están sombreados. La figura 2B muestra la parte N-terminal de cFP484, que no tiene homología con otras proteínas. El supuesto péptido señal está subrayado.

- 25 **[0049]** En este documento se han citado las referencias siguientes.
1. Ormo y col., (1996) Science 273: 1392-1395.
 2. Yang, F. y col., (1996) Nature Biotech 14: 1246-1251.

3. Cormack y col., (1996) Gene 173, 33-38.
4. Haas y col., (1996) Current Biology 6, 315-324.
5. Yang y col., (1996) Nucleic Acids Research 24, 4592-4593.
6. Ghoda y col., (1990) J. Biol. Chem. 265: 11823-11826.
7. Prasher D.C. y col., (1992) Gene 111:229-33.
8. Kain y col., (1995) Biotechniques 19(4):650-55.
9. Chomczynski P. y col., (1987) Anal. Biochem. 162, 156-159.
10. Frohman y col., (1998) PNAS USA, 85, 8998-9002.

10 **Lista de secuencias**

[0050]

- <110> Lukyanov, Sergey A.
 Labas, Yulii A.
 15 Matz, Mikhail V.
 Fradkov, Arcady F.
 <120> Proteínas fluorescentes de especies sin bioluminiscencia de la clase de los antozoos, genes codificantes de tales proteínas y sus usos
 <130> D6196PCT
 20 <141> 10-12-1999
 <150> 09/210.330
 <151> 11-12-1998
 <160> 63
 <210> 1
 25 <211> 25
 <212> ADN
 <213> secuencia artificial
 <220>
 <221> unión del cebador
 30 <223> cebador TN3 usado para la síntesis de ADNc y RACE
 <400> 1
 cgcagtcgaccgttttttttttt 25
 <210> 2
 <211> 23
 35 <212> ADN
 <213> secuencia artificial
 <220>
 <221> unión del cebador
 <223> cebador TS usado para la síntesis de ADNc y RACE
 40 <400> 2
 aagcagtggtatcaacgcagagt 23
 <210> 3
 <211> 6
 <212> Proteína
 45 <213> *Aequorea victoria*
 <220>
 <222> 21
 <223> secuencia de aminoácidos de un tramo clave en el que se basa el cebador NGH; Xaa en la posición 21
 representa desconocido
 50 <400> 3

Gly Xaa Val Asn Gly His

5

- <210> 4
 <211> 20
 <212> ADN
 55 <213> secuencia artificial
 <220>

<221> unión del cebador
 <222> 12
 <223> cebador NGH usado para el aislamiento de proteínas fluorescentes; n en la posición 12 representa cualquiera de las cuatro bases
 5 <400> 4
 gayggctgctnaayggdca 20
 <210> 5
 <211> 5
 <212> Proteína
 10 <213> *Aequorea victoria*
 <220>
 <222> 31...35
 <223> secuencia de aminoácidos de un tramo clave en el que se basan los cebadores GEGa y GEGb
 <400> 5

Gly Glu Gly Glu Gly

5

15 <210> 6
 <211> 20
 <212> ADN
 <213> secuencia artificial
 20 <220>
 <221> unión del cebador
 <223> cebador GEGa usado para el aislamiento de proteínas fluorescentes
 <400> 6
 gttacaggtgarggmargg 20
 25 <210> 7
 <211> 20
 <212> ADN
 <213> secuencia artificial
 <220>
 30 <221> unión del cebador
 <223> cebador GEGb usado para el aislamiento de proteínas fluorescentes
 <400> 7
 gttacaggtgarggkgargg 20
 35 <210> 8
 <211> 5
 <212> Proteína
 <213> *Aequorea victoria*
 <220>
 <222> 31...35
 40 <223> secuencia de aminoácidos de un tramo clave en el que se basan los cebadores GNGa y GNGb
 <400> 8

Gly Glu Gly Asn Gly

5

<210> 9
 <211> 20
 45 <212> ADN
 <213> secuencia artificial
 <220>
 <221> unión del cebador
 <223> cebador GNGa usado para el aislamiento de proteínas fluorescentes
 50 <400> 9
 gttacaggtgarggmaaygg 20
 <210> 10
 <211> 20
 <212> ADN

<213> secuencia artificial
 <220>
 <221> unión del cebador
 <223> cebador GNGb usado para el aislamiento de proteínas fluorescentes
 5 <400> 10
 gttacaggtgarggkaaygg 20
 <210> 11
 <211> 5
 <212> Proteína
 10 <213> *Aequorea victoria*
 <220>
 <222> 127...131
 <223> secuencia de aminoácidos de un tramo clave en el que se basa el cebador NFP
 <400> 11

Gly Met Asn Phe Pro

 15 5
 <210> 12
 <211> 5
 <212> Proteína
 <213> *Aequorea victoria*
 20 <220>
 <222> 127...131
 <223> secuencia de aminoácidos de un tramo clave en el que se basa el cebador NFP
 <400> 12

Gly Val Asn Phe Pro

 25 5
 <210> 13
 <211> 20
 <212> ADN
 <213> secuencia artificial
 <220>
 30 <221> unión del cebador
 <223> cebador NFP usado para el aislamiento de proteínas fluorescentes
 <400> 13
 ttccaygtrtgaaytlycc 20
 <210> 14
 35 <211> 4
 <212> Proteína
 <213> *Aequorea victoria*
 <220>
 <222> 134...137
 40 <223> secuencia de aminoácidos de un tramo clave en el que se basan los cebadores PVMa y PVMb
 <400> 14
 Gly Pro Val Met
 <210> 15
 <211> 21
 45 <212> ADN
 <213> secuencia artificial
 <220>
 <221> unión del cebador
 <222> 15
 50 <223> cebador PVMa usado para el aislamiento de proteínas fluorescentes; n en la posición 15 representa
 cualquiera de las cuatro bases
 <400> 15
 cctgccrayggtccngtmatg 21

<210> 16
 <211> 21
 <212> ADN
 <213> secuencia artificial
 5 <220>
 <221> unión del cebador
 <222> 15
 <223> cebador PVMb usado para el aislamiento de proteínas fluorescentes; n en la posición 15 representa cualquiera de las cuatro bases
 10 <400> 16
 cctgccrayggtccngtkatg 21
 <210> 17
 <211> 47
 <212> ADN
 15 <213> secuencia artificial
 <220>
 <221> unión del cebador
 <223> cebador T7-TN3 usado para la síntesis de ADNc y RACE
 <400> 17
 20 gtaatcgcactactataggccgcagtcgaccgtttttttttt 47
 <210> 18
 <211> 45
 <212> ADN
 <213> secuencia artificial
 25 <220>
 <221> unión del cebador
 <223> cebador T7-TS usado para la síntesis de ADNc y RACE
 <400> 18
 gtaatcgcactactataggccaagcagtggtatcaacgcagagt 45
 30 <210> 19
 <211> 22
 <212> ADN
 <213> secuencia artificial
 <220>
 35 <221> unión del cebador
 <223> cebador T7 usado para la síntesis de ADNc y RACE
 <400> 19
 gtaatcgcactactataggc 22
 40 <210> 20
 <211> 21
 <212> ADN
 <213> secuencia artificial
 <220>
 <221> unión del cebador
 45 <223> cebador específico del gen usado para RACE 5' para *Anemonia majano*
 <400> 20
 gaaatagtcaggcactactggt 21
 <210> 21
 <211> 20
 50 <212> ADN
 <213> secuencia artificial
 <220>
 <221> unión del cebador
 <223> cebador específico del gen usado para RACE 5' para *Anemonia majano*
 55 <400> 21
 gtcaggcactactgtaggat 20
 <210> 22
 <211> 21
 <212> ADN

<213> secuencia artificial
 <220>
 <221> unión del cebador
 <223> cebador específico del gen usado para RACE 5' para *Clavularia sp.*
 5 <400> 22
 ctgaaatagtctgtatatc 21
 <210> 23
 <211> 19
 <212> ADN
 10 <213> secuencia artificial
 <220>
 <221> unión del cebador
 <223> cebador específico del gen usado para RACE 5' para *Clavularia sp.*
 <400> 23
 15 tctgtatatcgtctgggt 19
 <210> 24
 <211> 23
 <212> ADN
 <213> secuencia artificial
 20 <220>
 <221> unión del cebador
 <223> cebador específico del gen usado para RACE 5' para *Zoanthus sp.*
 <400> 24
 25 gttcttgaatagtctactatgt 23
 <210> 25
 <211> 20
 <212> ADN
 <213> secuencia artificial
 <220>
 30 <221> unión del cebador
 <223> cebador específico del gen usado para RACE 5' para *Zoanthus sp.*
 <400> 25
 gtctactatgtcttgaggat 20
 <210> 26
 35 <211> 19
 <212> ADN
 <213> secuencia artificial
 <220>
 <221> unión del cebador
 40 <223> cebador específico del gen usado para RACE 5' para *Discosoma sp.* "roja"
 <400> 26
 caagcaaattgcaaaggctc 19
 <210> 27
 <211> 19
 45 <212> ADN
 <213> secuencia artificial
 <220>
 <221> unión del cebador
 <223> cebador específico del gen usado para RACE 5' para *Discosoma sp.* "roja"
 50 <400> 27
 cggattgtggccttcgta 19
 <210> 28
 <211> 19
 <212> ADN
 55 <213> secuencia artificial
 <220>
 <221> unión del cebador
 <223> cebador específico del gen usado para RACE 5' para *Discosoma striata*
 <400> 28

ttgtcttctgcacaac 19
 <210> 29
 <211> 17
 <212> ADN
 5 <213> secuencia artificial
 <220>
 <221> unión del cebador
 <223> cebador específico del gen usado para RACE 5' para *Discosoma striata*
 <400> 29
 10 ctgcacaacgggtccat 17
 <210> 30
 <211> 20
 <212> ADN
 <213> secuencia artificial
 15 <220>
 <221> unión del cebador
 <223> cebador específico del gen usado para RACE 5' para *Anemonia sulcata*
 <400> 30
 20 cctctatcttcattcctgc 20
 <210> 31
 <211> 20
 <212> ADN
 <213> secuencia artificial
 <220>
 25 <221> unión del cebador
 <223> cebador específico del gen usado para RACE 5' para *Anemonia sulcata*
 <400> 31
 tatcttcatttcctgcgtac 20
 <210> 32
 30 <211> 19
 <212> ADN
 <213> secuencia artificial
 <220>
 <221> unión del cebador
 35 <223> cebador específico del gen usado para RACE 5' para *Discosoma sp.* "magenta"
 <400> 32
 ttcagcaccatcacgag 19
 <210> 33
 <211> 19
 40 <212> ADN
 <213> secuencia artificial
 <220>
 <221> unión del cebador
 <223> cebador específico del gen usado para RACE 5' para *Discosoma sp.* "magenta"
 45 <400> 33
 acgctcagagctgggtcc 19
 <210> 34
 <211> 22
 <212> ADN
 50 <213> secuencia artificial
 <220>
 <221> unión del cebador
 <223> cebador específico del gen usado para RACE 5' para *Discosoma sp.* "verde"
 <400> 34
 55 ccctcagcaatccatcagttc 22
 <210> 35
 <211> 20
 <212> ADN
 <213> secuencia artificial

<220>
 <221> unión del cebador
 <223> cebador específico del gen usado para RACE 5' para *Discosoma sp.* "verde"
 <400> 35
 5 attatctcagtgatgggttc 20
 <210> 36
 <211> 31
 <212> ADN
 <213> secuencia artificial
 10 <220>
 <221> unión del cebador
 <223> cebador secuencia arriba usado para obtener la región codificante de longitud completa de las nFP de *Anemonia majano*
 <400> 36
 15 acatggatccgctctttcaacaagtattac 31
 <210> 37
 <211> 34
 <212> ADN
 <213> secuencia artificial
 20 <220>
 <221> unión del cebador
 <223> cebador secuencia abajo usado para obtener la región codificante de longitud completa de las nFP de *Anemonia majano*
 <400> 37
 25 tagtactcgagcttattcgtatttcagtgaaatc 34
 <210> 38
 <211> 29
 <212> ADN
 <213> secuencia artificial
 30 <220>
 <221> unión del cebador
 <223> cebador secuencia arriba usado para obtener la región codificante de longitud completa de las nFP de *Clavularia sp.*
 <400> 38
 35 acatggatccaacatttttgagaaacg 29
 <210> 39
 <211> 28
 <212> ADN
 <213> secuencia artificial
 40 <220>
 <221> unión del cebador
 <223> cebador secuencia arriba usado para obtener la región codificante de longitud completa de las nFP de *Clavularia sp.*
 <400> 39
 45 acatggatcaaagctctaaccaccatg 28
 <210> 40
 <211> 31
 <212> ADN
 <213> secuencia artificial
 50 <220>
 <221> unión del cebador
 <223> cebador secuencia abajo usado para obtener la región codificante de longitud completa de las nFP de *Clavularia sp.*
 <400> 40
 55 tagtactcgagcaacacaaaccctcagacaa 31
 <210> 41
 <211> 28
 <212> ADN
 <213> secuencia artificial

<220>
 <221> unión del cebador
 <223> cebador secuencia arriba usado para obtener la región codificante de longitud completa de las nFP de *Zoanthus sp.*
 5 <400> 41
 acatggatccgctcagtcaaagcacggt 28
 <210> 42
 <211> 32
 <212> ADN
 10 <213> secuencia artificial
 <220>
 <221> unión del cebador
 <223> cebador secuencia abajo usado para obtener la región codificante de longitud completa de las nFP de *Zoanthus sp.*
 15 <400> 42
 tagtactcgaggttgaactacattctatca 32
 <210> 43
 <211> 31
 <212> ADN
 20 <213> secuencia artificial
 <220>
 <221> unión del cebador
 <223> cebador secuencia arriba usado para obtener la región codificante de longitud completa de las nFP de *Discosoma sp.* "roja"
 25 <400> 43
 acatggatccaggtctccaagaatgtatc 31
 <210> 44
 <211> 29
 <212> ADN
 30 <213> secuencia artificial
 <220>
 <221> unión del cebador
 <223> cebador secuencia abajo usado para obtener la región codificante de longitud completa de las nFP de *Discosoma sp.* "roja"
 35 <400> 44
 tagtactcgaggagccaagtcagcctta 29
 <210> 45
 <211> 28
 <212> ADN
 40 <213> secuencia artificial
 <220>
 <221> unión del cebador
 <223> cebador secuencia arriba usado para obtener la región codificante de longitud completa de las nFP de *Discosoma striata*
 45 <400> 45
 acatggatccagttgtccaagagtgtg 28
 <210> 46
 <211> 28
 <212> ADN
 50 <213> secuencia artificial
 <220>
 <221> unión del cebador
 <223> cebador secuencia abajo usado para obtener la región codificante de longitud completa de las nFP de *Discosoma striata*
 55 <400> 46
 tagcgagctctatcatgcctcgtcacct 28
 <210> 47
 <211> 31
 <212> ADN

<213> secuencia artificial
 <220>
 <221> unión del cebador
 <223> cebador secuencia arriba usado para obtener la región codificante de longitud completa de las nFP de
 5 *Anemonia sulcata*
 <400> 47
 acatggatccgcttccttttaagaagact 31
 <210> 48
 <211> 28
 10 <212> ADN
 <213> secuencia artificial
 <220>
 <221> unión del cebador
 <223> cebador secuencia abajo usado para obtener la región codificante de longitud completa de las nFP de
 15 *Anemonia sulcata*
 <400> 48
 tagtactcgagtccttgggagcggcttg 28
 <210> 49
 <211> 30
 20 <212> ADN
 <213> secuencia artificial
 <220>
 <221> unión del cebador
 <223> cebador secuencia arriba usado para obtener la región codificante de longitud completa de las nFP de
 25 *Discosoma sp. "magenta"*
 <400> 49
 acatggatccagttgtccaagaatgtgat 30
 <210> 50
 <211> 26
 30 <212> ADN
 <213> secuencia artificial
 <220>
 <221> unión del cebador
 <223> cebador secuencia abajo usado para obtener la región codificante de longitud completa de las nFP de
 35 *Discosoma sp. "magenta"*
 <400> 50
 tagtactcgaggccattacgctaac 26
 <210> 51
 <211> 31
 40 <212> ADN
 <213> secuencia artificial
 <220>
 <221> unión del cebador
 <223> cebador secuencia arriba usado para obtener la región codificante de longitud completa de las nFP de
 45 *Discosoma sp. "verde"*
 <400> 51
 acatggatccagtcacttaagaagaaatg 31
 <210> 52
 <211> 29
 50 <212> ADN
 <213> secuencia artificial
 <220>
 <221> unión del cebador
 <223> cebador secuencia abajo usado para obtener la región codificante de longitud completa de las nFP de
 55 *Discosoma sp. "verde"*
 <400> 52
 tagtactcgagattcggtttaatgccttg 29
 <210> 53
 <211> 33

<212> ADN
 <213> secuencia artificial
 <220>
 <221> unión del cebador
 5 <223> oligonucleótido TS usado para la síntesis de ADNc y RACE
 <400> 53
 aagcagtggatcaacgcagagtacgcrgrg 33
 <210> 54
 <211> 238
 10 <212> Proteína
 <213> *Aequorea victoria*
 <220>
 <223> secuencia de aminoácidos de la GFP
 <400> 54

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

15

Val	Thr	Thr	Phe	Ser	Tyr	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro
				65					70					75
Asp	His	Met	Lys	Gln	His	Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu
				80					85					90
Gly	Tyr	Val	Gln	Glu	Arg	Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn
				95					100					105
Tyr	Lys	Thr	Arg	Ala	Glu	Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val
				110					115					120
Asn	Arg	Ile	Glu	Leu	Lys	Gly	Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn
				125					130					135
Ile	Leu	Gly	His	Lys	Leu	Glu	Tyr	Asn	Tyr	Asn	Ser	His	Asn	Val
				140					145					150
Tyr	Ile	Met	Ala	Asp	Lys	Gln	Lys	Asn	Gly	Ile	Lys	Val	Asn	Phe
				155					160					165
Lys	Ile	Arg	His	Asn	Ile	Glu	Asp	Gly	Ser	Val	Gln	Leu	Ala	Asp
				170					175					180
His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	Gly	Asp	Gly	Pro	Val	Leu	Leu
				185					190					195
Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	Ser	Ala	Leu	Ser	Lys	Asp
				200					205					210
Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu	Leu	Glu	Phe	Val	Thr
				215					220					225
Ala	Ala	Gly	Ile	Thr	His	Gly	Met	Asp	Glu	Leu	Tyr	Lys		
				230					235					

<210> 55

<211> 229

<212> Proteína

5 <213> *Anemonia majano*

<220>

<223> secuencia de aminoácidos de amFP486

<400> 55

Met	Ala	Leu	Ser	Asn	Lys	Phe	Ile	Gly	Asp	Asp	Met	Lys	Met	Thr
				5					10					15
Tyr	His	Met	Asp	Gly	Cys	Val	Asn	Gly	His	Tyr	Phe	Thr	Val	Lys
				20					25					30
Gly	Glu	Gly	Asn	Gly	Lys	Pro	Tyr	Glu	Gly	Thr	Gln	Thr	Ser	Thr
				35					40					45
Phe	Lys	Val	Thr	Met	Ala	Asn	Gly	Gly	Pro	Leu	Ala	Phe	Ser	Phe
				50					55					60
Asp	Ile	Leu	Ser	Thr	Val	Phe	Lys	Tyr	Gly	Asn	Arg	Cys	Phe	Thr
				65					70					75
Ala	Tyr	Pro	Thr	Ser	Met	Pro	Asp	Tyr	Phe	Lys	Gln	Ala	Phe	Pro
				80					85					90
Asp	Gly	Met	Ser	Tyr	Glu	Arg	Thr	Phe	Thr	Tyr	Glu	Asp	Gly	Gly
				95					100					105
Val	Ala	Thr	Ala	Ser	Trp	Glu	Ile	Ser	Leu	Lys	Gly	Asn	Cys	Phe
				110					115					120
Glu	His	Lys	Ser	Thr	Phe	His	Gly	Val	Asn	Phe	Pro	Ala	Asp	Gly
				125					130					135
Pro	Val	Met	Ala	Lys	Lys	Thr	Thr	Gly	Trp	Asp	Pro	Ser	Phe	Glu
				140					145					150
Lys	Met	Thr	Val	Cys	Asp	Gly	Ile	Leu	Lys	Gly	Asp	Val	Thr	Ala
				155					160					165
Phe	Leu	Met	Leu	Gln	Gly	Gly	Gly	Asn	Tyr	Arg	Cys	Gln	Phe	His
				170					175					180
Thr	Ser	Tyr	Lys	Thr	Lys	Lys	Pro	Val	Thr	Met	Pro	Pro	Asn	His
				185					190					195
Val	Val	Glu	His	Arg	Ile	Ala	Arg	Thr	Asp	Leu	Asp	Lys	Gly	Gly
				200					205					210
Asn	Ser	Val	Gln	Leu	Thr	Glu	His	Ala	Val	Ala	His	Ile	Thr	Ser
				215					220					225
Val	Val	Pro	Phe											

<210> 56
 <211> 266
 <212> Proteína

<213> *Clavularia* sp.

<220>

<223> secuencia de aminoácidos de cFP484

<400> 56

Met Lys Cys Lys Phe Val Phe Cys Leu Ser Phe Leu Val Leu Ala

5

10

15

Ile Thr Asn Ala Asn Ile Phe Leu Arg Asn Glu Ala Asp Phe Glu

5

	20		25		30
Glu Lys Thr Phe Arg Ile Pro Lys Ala Leu Thr Thr Met Gly Val					
	35		40		45
Ile Lys Pro Asp Met Lys Ile Lys Leu Lys Met Glu Gly Asn Val					
	50		55		60
Asn Gly His Ala Phe Val Ile Glu Gly Glu Gly Glu Gly Lys Pro					
	65		70		75
Tyr Asp Gly Thr His Thr Leu Asn Leu Glu Val Lys Glu Gly Ala					
	80		85		90
Pro Leu Pro Phe Ser Tyr Asp Ile Leu Ser Asn Ala Phe Gln Tyr					
	95		100		105
Gly Asn Arg Ala Leu Thr Lys Tyr Pro Asp Asp Ile Ala Asp Tyr					
	110		115		120
Phe Lys Gln Ser Phe Pro Glu Gly Tyr Ser Trp Glu Arg Thr Met					
	125		130		135
Thr Phe Glu Asp Lys Gly Ile Val Lys Val Lys Ser Asp Ile Ser					
	140		145		150
Met Glu Glu Asp Ser Phe Ile Tyr Glu Ile Arg Phe Asp Gly Met					
	155		160		165
Asp Phe Pro Pro Asn Gly Pro Val Met Gln Lys Lys Thr Leu Lys					
	170		175		180
Trp Glu Pro Ser Thr Glu Ile Met Tyr Val Arg Asp Gly Val Leu					
	185		190		195
Val Gly Asp Ile Ser His Ser Leu Leu Leu Glu Gly Gly Gly His					
	200		205		210
Tyr Arg Cys Asp Phe Lys Ser Ile Tyr Lys Ala Lys Lys Val Val					
	215		220		225
Lys Leu Pro Asp Tyr His Phe Val Asp His Arg Ile Glu Ile Leu					
	230		235		240
Asn His Asp Lys Asp Tyr Asn Lys Val Thr Leu Tyr Glu Asn Ala					
	245		250		255
Val Ala Arg Tyr Ser Leu Leu Pro Ser Gln Ala					
	260		265		

<211> 230
<212> Proteína
<213> *Zoanthus sp.*
<220>

5 <223> secuencia de aminoácidos de zFP506
<400> 57

Ala	Gln	Ser	Lys	His	Gly	Leu	Thr	Lys	Glu	Met	Thr	Met	Lys	Tyr
				5					10					15
Arg	Met	Glu	Gly	Cys	Val	Asp	Gly	His	Lys	Phe	Val	Ile	Thr	Gly
				20					25					30
Glu	Gly	Ile	Gly	Tyr	Pro	Phe	Lys	Gly	Lys	Gln	Ala	Ile	Asn	Leu
				35					40					45
Cys	Val	Val	Glu	Gly	Gly	Pro	Leu	Pro	Phe	Ala	Glu	Asp	Ile	Leu
				50					55					60
Ser	Ala	Ala	Phe	Asn	Tyr	Gly	Asn	Arg	Val	Phe	Thr	Glu	Tyr	Pro
				65					70					75
Gln	Asp	Ile	Val	Asp	Tyr	Phe	Lys	Asn	Ser	Cys	Pro	Ala	Gly	Tyr
				80					85					90
Thr	Trp	Asp	Arg	Ser	Phe	Leu	Phe	Glu	Asp	Gly	Ala	Val	Cys	Ile
				95					100					105
Cys	Asn	Ala	Asp	Ile	Thr	Val	Ser	Val	Glu	Glu	Asn	Cys	Met	Tyr
				110					115					120
His	Glu	Ser	Lys	Phe	Tyr	Gly	Val	Asn	Phe	Pro	Ala	Asp	Gly	Pro
				125					130					135
Val	Met	Lys	Lys	Met	Thr	Asp	Asn	Trp	Glu	Pro	Ser	Cys	Glu	Lys
				140					145					150
Ile	Ile	Pro	Val	Pro	Lys	Gln	Gly	Ile	Leu	Lys	Gly	Asp	Val	Ser
				155					160					165
Met	Tyr	Leu	Leu	Leu	Lys	Asp	Gly	Gly	Arg	Leu	Arg	Cys	Gln	Phe
				170					175					180
Asp	Thr	Val	Tyr	Lys	Ala	Lys	Ser	Val	Pro	Arg	Lys	Met	Pro	Asp
				185					190					195
Trp	His	Phe	Ile	Gln	His	Lys	Leu	Thr	Arg	Glu	Asp	Arg	Ser	Asp
				200					205					210
Ala	Lys	Asn	Gln	Lys	Trp	His	Leu	Thr	Glu	His	Ala	Ile	Ala	Ser
				215					220					225
Gly	Ser	Ala	Leu	Pro										
				230										

<210> 58
 <211> 230
 <212> Proteína

<213> *Zoanthus sp.*

<220>

<223> secuencia de aminoácidos de zFP538

<400> 58

Met	Ala	His	Ser	Lys	His	Gly	Leu	Lys	Glu	Glu	Met	Thr	Met	Lys
				5					10					15
Tyr	His	Met	Glu	Gly	Cys	Val	Asn	Gly	His	Lys	Phe	Val	Ile	Thr
				20					25					30
Gly	Glu	Gly	Ile	Gly	Tyr	Pro	Phe	Lys	Gly	Lys	Gln	Thr	Ile	Asn
				35					40					45
Leu	Cys	Val	Ile	Glu	Gly	Gly	Pro	Leu	Pro	Phe	Ser	Glu	Asp	Ile
				50					55					60
Leu	Ser	Ala	Gly	Phe	Lys	Tyr	Gly	Asp	Arg	Ile	Phe	Thr	Glu	Tyr
				65					70					75
Pro	Gln	Asp	Ile	Val	Asp	Tyr	Phe	Lys	Asn	Ser	Cys	Pro	Ala	Gly
				80					85					90
Tyr	Thr	Trp	Gly	Ser	Phe	Leu	Phe	Glu	Asp	Gly	Ala	Val	Cys	Ile
				95					100					105
Cys	Asn	Val	Asp	Ile	Thr	Val	Ser	Val	Lys	Glu	Asn	Cys	Ile	Tyr
				110					115					120
His	Lys	Ser	Ile	Phe	Asn	Gly	Met	Asn	Phe	Pro	Ala	Asp	Gly	Pro
				125					130					135
Val	Met	Lys	Lys	Met	Thr	Thr	Asn	Trp	Glu	Ala	Ser	Cys	Glu	Lys
				140					145					150
Ile	Met	Pro	Val	Pro	Lys	Gln	Gly	Ile	Leu	Lys	Gly	Asp	Val	Ser
				155					160					165
Met	Tyr	Leu	Leu	Leu	Lys	Asp	Gly	Gly	Arg	Tyr	Arg	Cys	Gln	Phe
				170					175					180
Asp	Thr	Val	Tyr	Lys	Ala	Lys	Ser	Val	Pro	Ser	Lys	Met	Pro	Glu
				185					190					195
Trp	His	Phe	Ile	Gln	His	Lys	Leu	Leu	Arg	Glu	Asp	Arg	Ser	Asp
				200					205					210
Ala	Lys	Asn	Gln	Lys	Trp	Gln	Leu	Thr	Glu	His	Ala	Ile	Ala	Phe
				215					220					225

Pro Ser Ala Leu Ala

230

<210> 59
 <211> 232
 <212> Proteína
 <213> *Discosoma striata*

5

<220>
 <223> secuencia de aminoácidos de dsFP483
 <400> 59

Met	Ser	Cys	Ser	Lys	Ser	Val	Ile	Lys	Glu	Glu	Met	Leu	Ile	Asp
				5					10					15
Leu	His	Leu	Glu	Gly	Thr	Phe	Asn	Gly	His	Tyr	Phe	Glu	Ile	Lys
				20					25					30
Gly	Lys	Gly	Lys	Gly	Gln	Pro	Asn	Glu	Gly	Thr	Asn	Thr	Val	Thr
				35					40					45
Leu	Glu	Val	Thr	Lys	Gly	Gly	Pro	Leu	Pro	Phe	Gly	Trp	His	Ile
				50					55					60
Leu	Cys	Pro	Gln	Phe	Gln	Tyr	Gly	Asn	Lys	Ala	Phe	Val	His	His
				65					70					75
Pro	Asp	Asn	Ile	His	Asp	Tyr	Leu	Lys	Leu	Ser	Phe	Pro	Glu	Gly
				80					85					90
Tyr	Thr	Trp	Glu	Arg	Ser	Met	His	Phe	Glu	Asp	Gly	Gly	Leu	Cys
				95					100					105
Cys	Ile	Thr	Asn	Asp	Ile	Ser	Leu	Thr	Gly	Asn	Cys	Phe	Tyr	Tyr
				110					115					120
Asp	Ile	Lys	Phe	Thr	Gly	Leu	Asn	Phe	Pro	Pro	Asn	Gly	Pro	Val
				125					130					135
Val	Gln	Lys	Lys	Thr	Thr	Gly	Trp	Glu	Pro	Ser	Thr	Glu	Arg	Leu
				140					145					150
Tyr	Pro	Arg	Asp	Gly	Val	Leu	Ile	Gly	Asp	Ile	His	His	Ala	Leu
				155					160					165
Thr	Val	Glu	Gly	Gly	Gly	His	Tyr	Ala	Cys	Asp	Ile	Lys	Thr	Val
				170					175					180
Tyr	Arg	Ala	Lys	Lys	Ala	Ala	Leu	Lys	Met	Pro	Gly	Tyr	His	Tyr
				185					190					195

```

Val Asp Thr Lys Leu Val Ile Trp Asn Asn Asp Lys Glu Phe Met
                200                205                210
Lys Val Glu Glu His Glu Ile Ala Val Ala Arg His His Pro Phe
                215                220                225
Tyr Glu Pro Lys Lys Asp Lys
                230

```

<210> 60

<211> 225

<212> Proteína

5 <213> *Discosoma sp.* "roja"

<220>

<223> secuencia de aminoácidos de drFP583

<400> 60

Met	Arg	Ser	Ser	Lys	Asn	Val	Ile	Lys	Glu	Phe	Met	Arg	Phe	Lys
				5					10					15
Val	Arg	Met	Glu	Gly	Thr	Val	Asn	Gly	His	Glu	Phe	Glu	Ile	Glu
				20					25					30
Gly	Glu	Gly	Glu	Gly	Arg	Pro	Tyr	Glu	Gly	His	Asn	Thr	Val	Lys
				35					40					45
Leu	Lys	Val	Thr	Lys	Gly	Gly	Pro	Leu	Pro	Phe	Ala	Trp	Asp	Ile
				50					55					60
Leu	Ser	Pro	Gln	Phe	Gln	Tyr	Gly	Ser	Lys	Val	Tyr	Val	Lys	His
				65					70					75
Pro	Ala	Asp	Ile	Pro	Asp	Tyr	Lys	Lys	Leu	Ser	Phe	Pro	Glu	Gly
				80					85					90
Phe	Lys	Trp	Glu	Arg	Val	Met	Asn	Phe	Glu	Asp	Gly	Gly	Val	Val
				95					100					105
Thr	Val	Thr	Gln	Asp	Ser	Ser	Leu	Gln	Asp	Gly	Cys	Phe	Ile	Tyr
				110					115					120
Lys	Val	Lys	Phe	Ile	Gly	Val	Asn	Phe	Pro	Ser	Asp	Gly	Pro	Val
				125					130					135
Met	Gln	Lys	Lys	Thr	Met	Gly	Trp	Glu	Ala	Ser	Thr	Glu	Arg	Leu
				140					145					150
Tyr	Pro	Arg	Asp	Gly	Val	Leu	Lys	Gly	Glu	Ile	His	Lys	Ala	Leu
				155					160					165
Lys	Leu	Lys	Asp	Gly	Gly	His	Tyr	Leu	Val	Glu	Phe	Lys	Ser	Ile
				170					175					180
Tyr	Met	Ala	Lys	Lys	Pro	Val	Gln	Leu	Pro	Gly	Tyr	Tyr	Tyr	Val
				185					190					195
Asp	Ser	Lys	Leu	Asp	Ile	Thr	Ser	His	Asn	Glu	Asp	Tyr	Thr	Ile
				200					205					210
Val	Glu	Gln	Tyr	Glu	Arg	Thr	Glu	Gly	Arg	His	His	Leu	Phe	Leu
				215					220					225

<210> 61

<211> 232

5 <212> Proteína

<213> *Anemonia sulcata*

<220>

<223> secuencia de aminoácidos de asFP600

<400> 61

Met	Ala	Ser	Phe	Leu	Lys	Lys	Thr	Met	Pro	Phe	Lys	Thr	Thr	Ile
				5					10					15
Glu	Gly	Thr	Val	Asn	Gly	His	Tyr	Phe	Lys	Cys	Thr	Gly	Lys	Gly
				20					25					30
Glu	Gly	Asn	Pro	Phe	Glu	Gly	Thr	Gln	Glu	Met	Lys	Ile	Glu	Val
				35					40					45
Ile	Glu	Gly	Gly	Pro	Leu	Pro	Phe	Ala	Phe	His	Ile	Leu	Ser	Thr
				50					55					60
Ser	Cys	Met	Tyr	Gly	Ser	Lys	Thr	Phe	Ile	Lys	Tyr	Val	Ser	Gly
				65					70					75
Ile	Pro	Asp	Tyr	Phe	Lys	Gln	Ser	Phe	Pro	Glu	Gly	Phe	Thr	Trp
				80					85					90
Glu	Arg	Thr	Thr	Thr	Tyr	Glu	Asp	Gly	Gly	Phe	Leu	Thr	Ala	His
				95					100					105
Gln	Asp	Thr	Ser	Leu	Asp	Gly	Asp	Cys	Leu	Val	Tyr	Lys	Val	Lys
				110					115					120
Ile	Leu	Gly	Asn	Asn	Phe	Pro	Ala	Asp	Gly	Pro	Val	Met	Gln	Asn
				125					130					135
Lys	Ala	Gly	Arg	Trp	Glu	Pro	Ala	Thr	Glu	Ile	Val	Tyr	Glu	Val
				140					145					150
Asp	Gly	Val	Leu	Arg	Gly	Gln	Ser	Leu	Met	Ala	Leu	Lys	Cys	Pro
				155					160					165
Gly	Gly	Arg	His	Leu	Thr	Cys	His	Leu	His	Thr	Thr	Tyr	Arg	Ser
				170					175					180
Lys	Lys	Pro	Ala	Ser	Ala	Leu	Lys	Met	Pro	Gly	Phe	His	Phe	Glu
				185					190					195
Asp	His	Arg	Ile	Glu	Ile	Met	Glu	Glu	Val	Glu	Lys	Gly	Lys	Cys
				200					205					210
Tyr	Lys	Gln	Tyr	Glu	Ala	Ala	Val	Gly	Arg	Tyr	Cys	Asp	Ala	Ala
				215					220					225
Pro	Ser	Lys	Leu	Gly	His	Asn								
				230										

<210> 62

5 <211> 231

<212> Proteína
 <213> *Discosoma sp.* "verde"
 <220>
 <223> secuencia de aminoácidos de dgFP512
 5 <400> 62

Met	Ser	Ala	Leu	Lys	Glu	Glu	Met	Lys	Ile	Asn	Leu	Thr	Met	Glu
				5					10					15
Gly	Val	Val	Asn	Gly	Leu	Pro	Phe	Lys	Ile	Arg	Gly	Asp	Gly	Lys
				20					25					30
Gly	Lys	Pro	Tyr	Gln	Gly	Ser	Gln	Glu	Leu	Thr	Leu	Thr	Val	Val
				35					40					45
Lys	Gly	Gly	Pro	Leu	Pro	Phe	Ser	Tyr	Asp	Ile	Leu	Thr	Thr	Met
				50					55					60
Phe	Gln	Tyr	Gly	Asn	Arg	Ala	Phe	Val	Asn	Tyr	Pro	Glu	Asp	Ile
				65					70					75
Pro	Asp	Ile	Phe	Lys	Gln	Thr	Cys	Ser	Gly	Pro	Asn	Gly	Gly	Tyr
				80					85					90
Ser	Trp	Gln	Arg	Thr	Met	Thr	Tyr	Glu	Asp	Gly	Gly	Val	Cys	Thr
				95					100					105
Ala	Thr	Ser	Asn	Ile	Ser	Val	Val	Gly	Asp	Thr	Phe	Asn	Tyr	Asp
				110					115					120

Ile His Phe Met Gly Ala Asn Phe Pro Leu Asp Gly Pro Val Met
 125 130 135
 Gln Lys Arg Thr Met Lys Trp Glu Pro Ser Thr Glu Ile Met Phe
 140 145 150
 Glu Arg Asp Gly Met Leu Arg Gly Asp Ile Ala Met Ser Leu Leu
 155 160 165
 Leu Lys Gly Gly Gly His Tyr Arg Cys Asp Phe Glu Thr Ile Tyr
 170 175 180
 Lys Pro Asn Lys Val Val Lys Met Pro Asp Tyr His Phe Val Asp
 185 190 195
 His Cys Ile Glu Ile Thr Ser Gln Gln Asp Tyr Tyr Asn Val Val
 200 205 210
 Glu Leu Thr Glu Val Ala Glu Ala Arg Tyr Ser Ser Leu Glu Lys
 215 220 225
 Ile Gly Lys Ser Lys Ala
 230

- <210> 63
- <211> 235
- <212> Proteína
- 5 <213> *Discosoma sp.* "magenta"
- <220>
- <223> secuencia de aminoácidos de dmFP592
- <400> 63

Met Ser Cys Ser Lys Asn Val Ile Lys Glu Phe Met Arg Phe Lys
 5 10 15
 Val Arg Met Glu Gly Thr Val Asn Gly His Glu Phe Glu Ile Lys
 20 25 30
 Gly Glu Gly Glu Gly Arg Pro Tyr Glu Gly His Cys Ser Val Lys
 35 40 45

Leu	Met	Val	Thr	Lys	Gly	Gly	Pro	Leu	Pro	Phe	Ala	Phe	Asp	Ile
				50					55					60
Leu	Ser	Pro	Gln	Phe	Gln	Tyr	Gly	Ser	Lys	Val	Tyr	Val	Lys	His
				65					70					75
Pro	Ala	Asp	Ile	Pro	Asp	Tyr	Lys	Lys	Leu	Ser	Phe	Pro	Glu	Gly
				80					85					90
Phe	Lys	Trp	Glu	Arg	Val	Met	Asn	Phe	Glu	Asp	Gly	Gly	Val	Val
				100					105					110
Thr	Val	Ser	Gln	Asp	Ser	Ser	Leu	Lys	Asp	Gly	Cys	Phe	Ile	Tyr
				115					120					125
Glu	Val	Lys	Phe	Ile	Gly	Val	Asn	Phe	Pro	Ser	Asp	Gly	Pro	Val
				130					135					140
Met	Gln	Arg	Arg	Thr	Arg	Gly	Trp	Glu	Ala	Ser	Ser	Glu	Arg	Leu
				145					150					155
Tyr	Pro	Arg	Asp	Gly	Val	Leu	Lys	Gly	Asp	Ile	His	Met	Ala	Leu
				160					165					170
Arg	Leu	Glu	Gly	Gly	Gly	His	Tyr	Leu	Val	Glu	Phe	Lys	Ser	Ile
				175					180					185
Tyr	Met	Val	Lys	Lys	Pro	Ser	Val	Gln	Leu	Pro	Gly	Tyr	Tyr	Tyr
				190					195					200
Val	Asp	Ser	Lys	Leu	Asp	Met	Thr	Ser	His	Asn	Glu	Asp	Tyr	Thr
				205					210					215
Val	Val	Glu	Gln	Tyr	Glu	Lys	Thr	Gln	Gly	Arg	His	His	Pro	Phe
				220					225					230
Ile	Lys	Pro	Leu	Gln										
				235										

REIVINDICACIONES

1. Un ácido nucleico en un entorno distinto de su entorno natural cuya secuencia codifica una proteína fluorescente de un organismo antozoo sin bioluminiscencia seleccionado de los géneros *Anemonia*, *Clavularia* y *Discosoma*, en que:
- 5 (a) dicha proteína fluorescente es de *Anemonia majano* y dicho ácido nucleico puede amplificarse con un par de cebadores, el primero de los cuales consta de la secuencia de SEQ ID NO:36 y el segundo de los cuales consta de la secuencia de SEQ ID NO:37; o
- 10 (b) dicha proteína fluorescente es de *Clavularia sp.* y dicho ácido nucleico puede amplificarse con un par de cebadores compuesto de un primer cebador que consta de la secuencia de SEQ ID NO:38 y un segundo cebador que consta de la secuencia de SEQ ID NO:40; o
- (c) dicha proteína fluorescente es de *Zoanthus sp.* y dicho ácido nucleico puede amplificarse con un par de cebadores compuesto de un primer cebador que consta de la secuencia de SEQ ID NO:41 y un segundo cebador que consta de la secuencia de SEQ ID NO:42; o
- 15 (d) dicha proteína fluorescente es de *Discosoma sp.* "roja" y dicho ácido nucleico puede amplificarse con un par de cebadores compuesto de un primer cebador que consta de la secuencia de SEQ ID NO:43 y un segundo cebador que consta de la secuencia de SEQ ID NO:44; o
- (e) dicha proteína fluorescente es de *Discosoma striata* y dicho ácido nucleico puede amplificarse con un par de cebadores compuesto de un primer cebador que consta de la secuencia de SEQ ID NO:45 y un segundo cebador que consta de la secuencia de SEQ ID NO:46.
- 20 2. Un ácido nucleico según la reivindicación 1, en que dicha proteína fluorescente es de *Anemonia majano* y dicho ácido nucleico puede amplificarse con un par de cebadores compuesto de un primer cebador que consta de la secuencia de SEQ ID NO:36 y un segundo cebador que consta de la secuencia de SEQ ID NO:37.
- 25 3. Un ácido nucleico según la reivindicación 1, en que dicha proteína fluorescente es de *Clavularia sp.* y dicho ácido nucleico puede amplificarse con un par de cebadores compuesto de un primer cebador que consta de la secuencia de SEQ ID NO:38 y un segundo cebador que consta de la secuencia de SEQ ID NO:40.
- 30 4. Un ácido nucleico según la reivindicación 1, en que dicha proteína fluorescente es de *Zoanthus sp.* y dicho ácido nucleico puede amplificarse con un par de cebadores compuesto de un primer cebador que consta de la secuencia de SEQ ID NO:41 y un segundo cebador que consta de la secuencia de SEQ ID NO:42.
- 35 5. Un ácido nucleico según la reivindicación 1, en que dicha proteína fluorescente es de *Discosoma sp.* "roja" y dicho ácido nucleico puede amplificarse con un par de cebadores compuesto de un primer cebador que consta de la secuencia de SEQ ID NO:43 y un segundo cebador que consta de la secuencia de SEQ ID NO:44.
- 40 6. Un ácido nucleico según la reivindicación 1, en que dicha proteína fluorescente es de *Discosoma striata* y dicho ácido nucleico puede amplificarse con un par de cebadores compuesto de un primer cebador que consta de la secuencia de SEQ ID NO:45 y un segundo cebador que consta de la secuencia de SEQ ID NO:46.
7. El ácido nucleico según una cualquiera de las reivindicaciones 1 a 6, en que dicho ácido nucleico está aislado.
- 45 8. Una construcción que comprende un vector y un ácido nucleico según una cualquiera de las reivindicaciones 1 a 6.
9. Una célula transformada con un ácido nucleico según una cualquiera de las reivindicaciones 1 a 6.
- 50 10. Un ácido nucleico según una cualquiera de las reivindicaciones 1 a 6, en que la secuencia del ácido nucleico puede amplificarse con un cebador seleccionado del grupo que consta de SEQ ID NO:4, 6, 7, 9, 10, 13, 15 y 16.
- 55 11. Una proteína fluorescente codificada por el ácido nucleico según una cualquiera de las reivindicaciones 1 a 6 presente en un entorno distinto de su entorno natural.

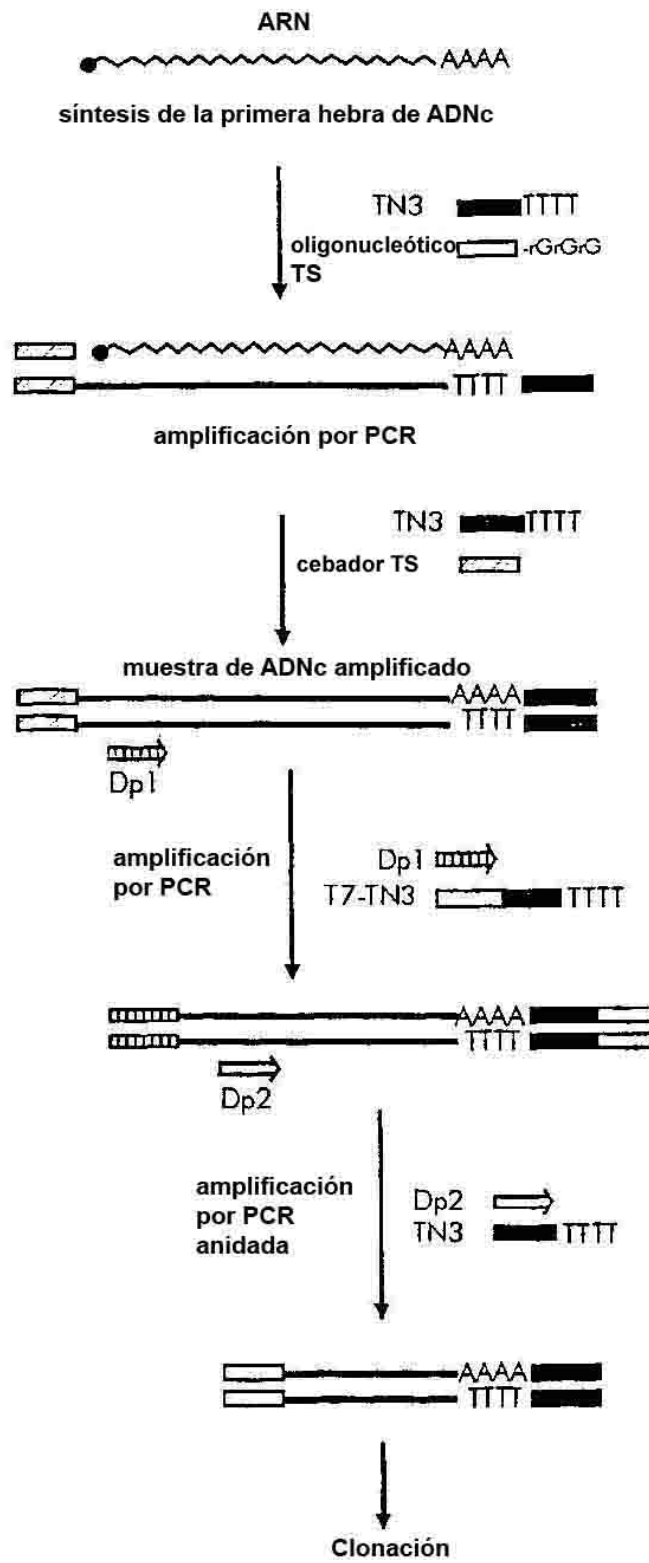


FIG. 1

10	20	30	40	50	SEQ ID#	
MSKGEELFTG.VVPILVELDGDVNGHKFSVSGEGEDATYGLKILKFICTT.GKLPVP..W GFP					54	
MAQSKHGLTK.EMTMKYRMEGCVDGHKFVITGEGIGYPFKGKQAINLCVV..EGGPLPFAE					zFP506 57	
--H----KE.-----H----N-----T-----I.-----S-					zFP538 58	
MSWSKSVIKE.EMLIDLHLEGTFNHGYFEIKGKGGKPNEGTNTVTLEVT..KGGPLPFGW					dsFP483 59	
...M-AL--Y-K-N-TM--VV--LP-K-R-D-----YQ-SQEL--T-V.-----SY					dgFP512 62	
-RS--N----F-RFKVRM--V---E---E-E-R-Y--H---K-K-.-----A-					drFP583 60	
M-C--N----F-RFKVRM--V---E---E-E-R-Y--HCS-K-M--.-----AF					dmFP592 63	
...MASFLKK.TMPFKTTIEGTVNGHYFKCTGKGEKNPFEGTQEMKIEVI..EGGPLPFAE					asFP600 61	
MALSINKFIGD.DMKMTYHMDGCVNGHYFTVKGEGNGKPYEGTQSTFKVTHANGGPLAFSE					amFP486 55	
KALTTMGVIKPDMKIKLKMEGNVNGHAEVIEGEGEGKPYDGTHTLNLEVKMAEGAPLPFSY					cFP484 56	
60	70	80	90	100	110	
PTLVTTFSYGVQCFSRYPDHHKQHDFFKSAM.:PEGYVQERTIFFKDOGNYKTRAEVKFEGD..					GFP	
DILSAAFNYGNRVFTEYPODIV..DYFKNSC...PAGYTWDRSFLFEDGAVCICNADITVSVEEN					zFP506	
-----G-K--D-I-----G-----V-----K--					zFP538	
HILCPQFQYGNKAFVHHFDDIP..DYLKLSF...PEGYTWERSMHFEDGGLCCITNDISLTGN..					dsFP483	
D--TTM-----R---HY-E---...IF-QTCSGPNG--S-Q-T-TY---V-TA-SN--VV-D..					dgFP512	
D--S-----S-VY-K--A---...K-----FK--V-N-----VTV-Q-S--QDG..					drFP583	
D--S-----S-VY-K--A---...K-----FK--V-N-----VTVSQ-S--KOG..					dmFP592	
HILSTSCHYGSKTFIKYVSGIP..DYFKQSF...PEGFTWERTTTYEDGGFLTAHQDTSLDGD..					asFP600	
DILSTVFKYGNRCFTAYPTSMP..DYFKQAF...PDGMSYERTFTYE DGGVATASWEISLKN..					amFP486	
DILSNAFQYGNRALTKYPDDIA..DYFKQSF...PEGYSWERTMTFEDKGIVKVKSDISMEED..					cFP484	
120	130	140	150	160	170	
TLVNRIELKGLDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNEKIRHNIEDGVSQVQL					GFP	
CMYHESKEYGVNFPADGPVM.KKMTDNWEPSCERII PVPKQGI LKGOVSMYLLIKDGGRLR					zFP506	
-I--K-I-N-M-----T--A-----M-----Y-					zFP538	
CENYDIKFTGLNFPPNGPVV.QKKTGWEPSTERLYP..RDGVLIGDIHHAHALTVEGGGHYV					dsFP483	
T-----H-M-A---LD--MM---R-MK-----IMFE ---L-R-D-AMS-LLK-----R					dgFP512	
--I-KV--I-V---SD---M.---M---A-----K-E--K--KLKD---L					drFP583	
--I-EV--I-V---SD---M--RR-R-----S-----K---M--RL-----L					dmFP592	
CLVYKVKILGNFPADGPVM.QNKAGRNEPATEIVYE..VDGVLRGQSLMALKCPGGRHLT					asFP600	
CFEHKSTFHGVNFPADGPVM.AKKTGWDPSEKMTV;.CDGILKGDVTAFLMLQGGGNYR					amFP486	
SFIYEIRFDGMNFPNGPVM.QKKTWKWEPSTEIMYV..RDGVLVGDISHSLLLEGGGHYR					cFP484	
180	190	200	210	220	230	
ADHYQONTPIGDG.PVLLPDNHVYLSSTOSALS KDPNEKRDHMLLEFVTAAGITHGMDELYK					GFP	
CQFDTVYKAKSV..PRKMPDWHFIQHKLTREDRSDAKNQKWHLTEHAIASGSALP					zFP506	
-----S--E-----L-----Q-----FP---A					zFP538	
CDIKTVYQAKK..PVKMPGYHYVDTKLVIRSNOKEFM.KVEEHEIAVARHHHPLOSQ					dsFP483	
--FE-I-KPN- V----D--F--HYIE-T-QQNYN V--LT-V-E--YSS-EKIGKSKA					dgFP512	
VEF-SI-M---...-QL---Y---S--D-T-HNEDYT.I--QY-RTEG---LEL					drFP583	
VEF-SI-MV-- PS-QL---Y---S--DMT-HNEDYT V--QY-KTQ-----FIKPLQ					dmFP592	
CHLHTTYRSKKPASALKMPGFHFEDHRIEIMEEVEKKGK.CYKQYEAAVGRYCDAAAPSKLGHN					asFP600	
CQFHTSYKTKK...PVTMPPNHVVEHRIARTDLKGGN.SVQLTEHAVAHITSVFPF					amFP486	
CDFKSIYKAKK...VVKLPDYHFVDHRIEILNHDKDYN.KVTLYENAVARYSLLPSQA					cFP484	

FIG. 2A



MKCKFVFLSFLVLAITNANI FLRNEADLEEKTLRIP

FIG. 2B

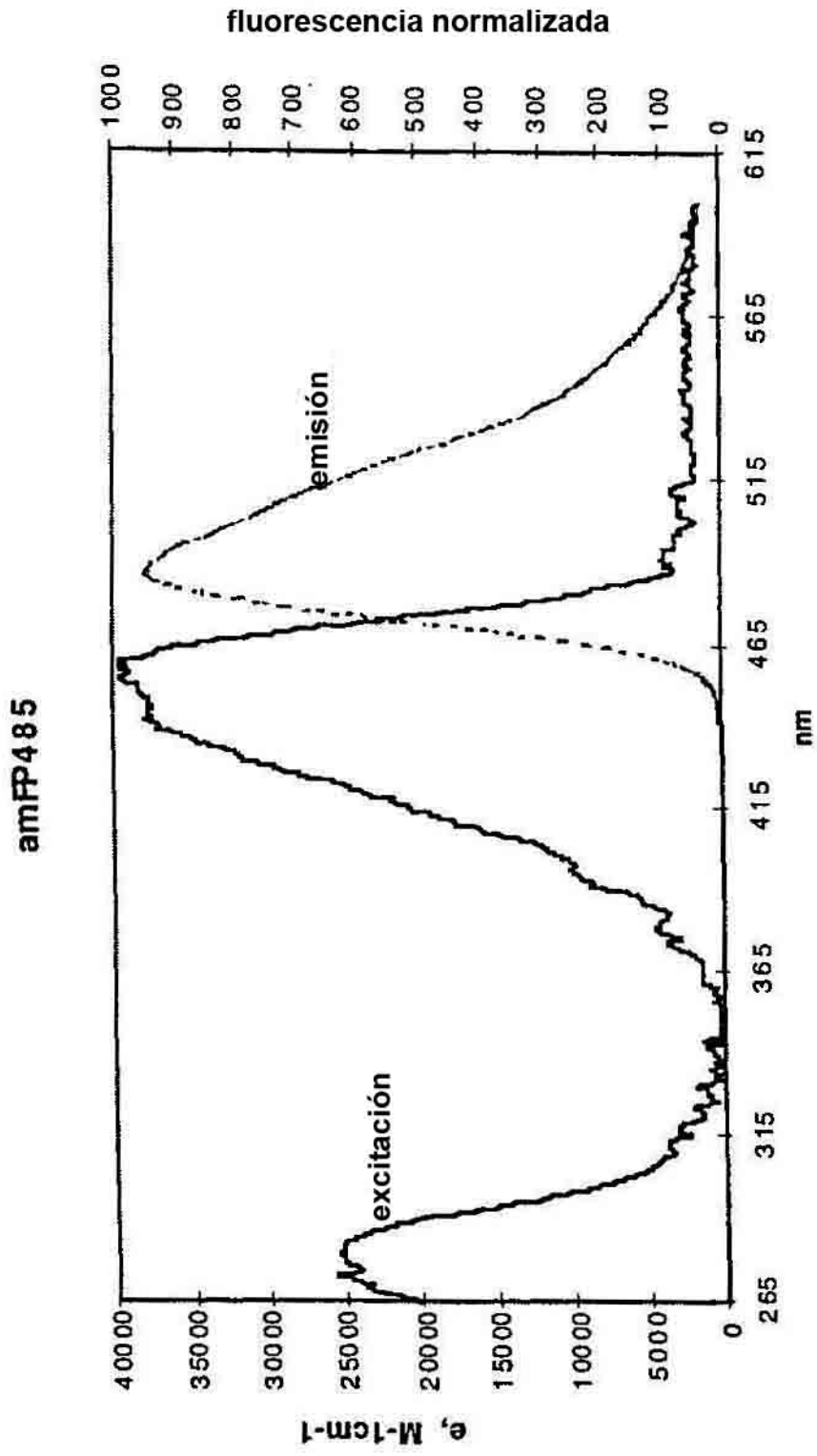


FIG. 3

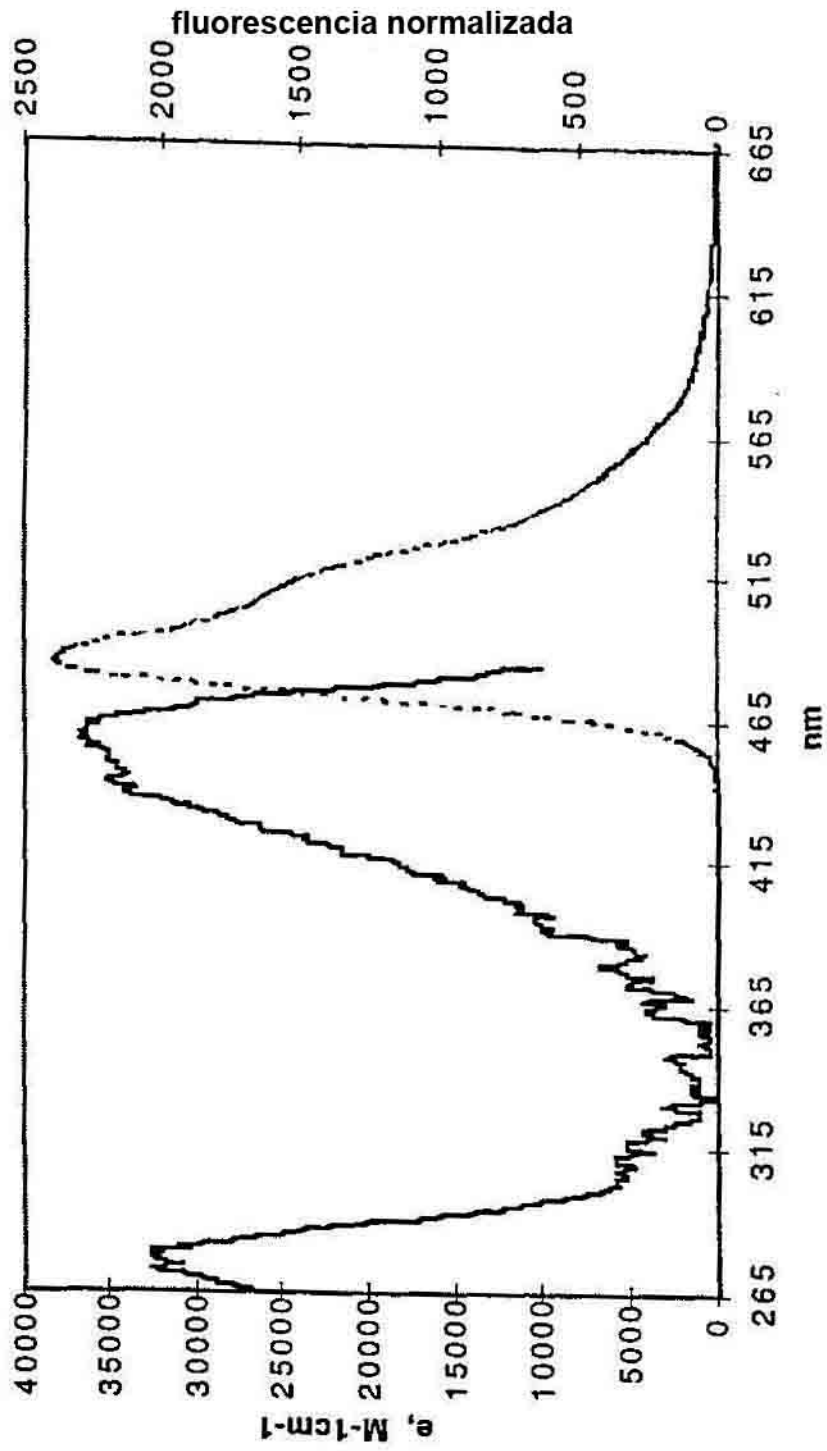


FIG. 4

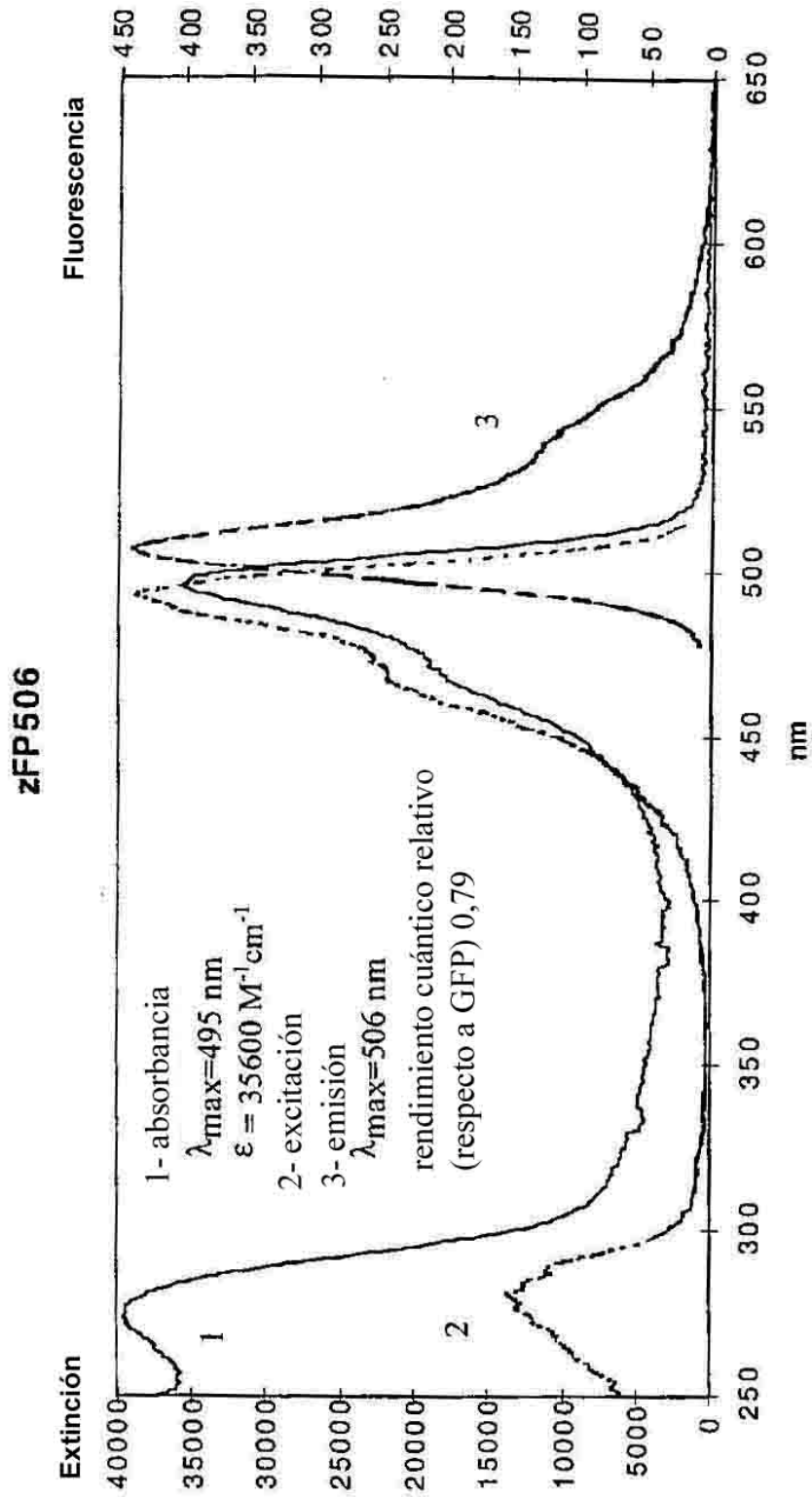


FIG. 5

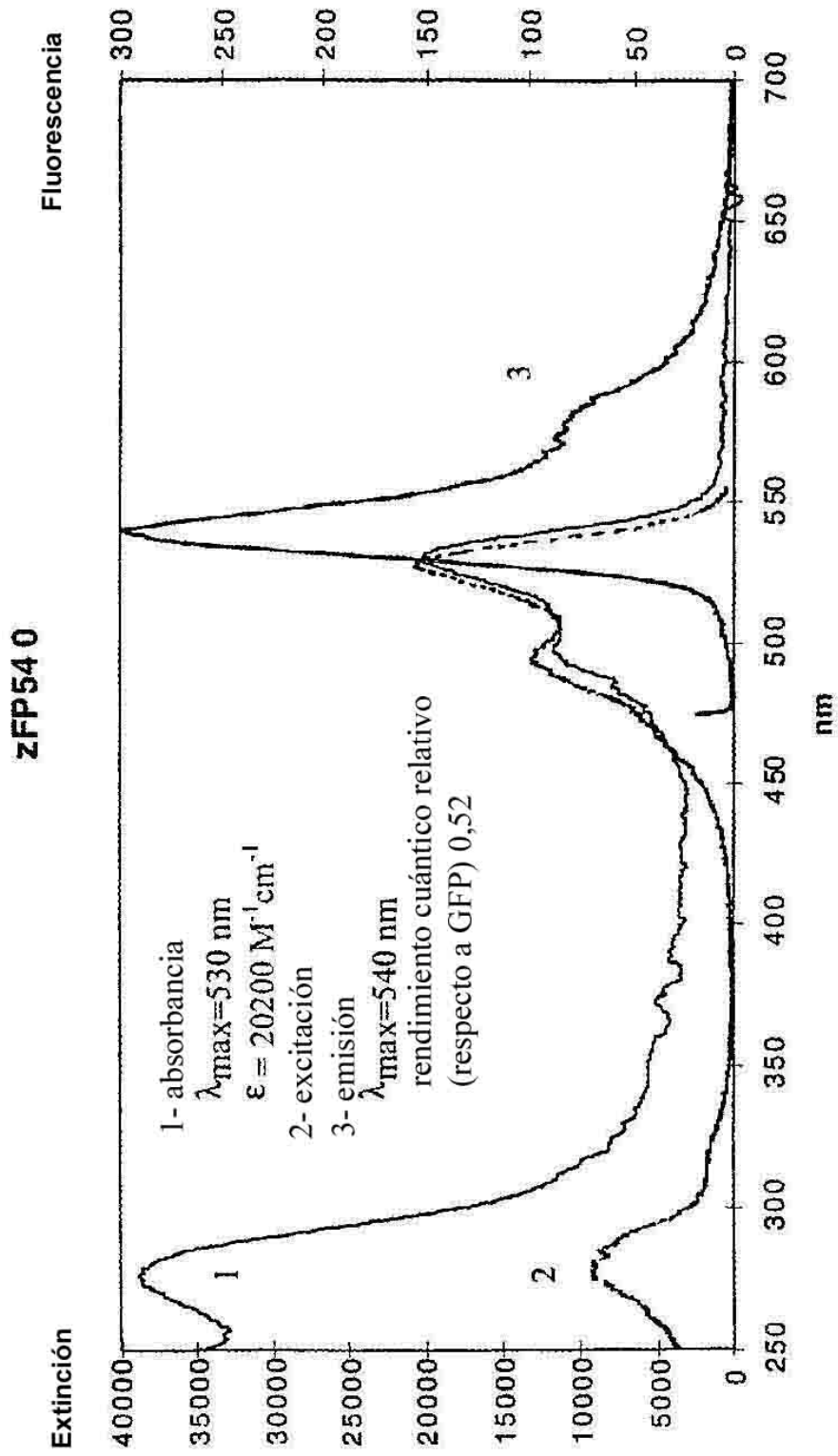


FIG. 6

dsFP484

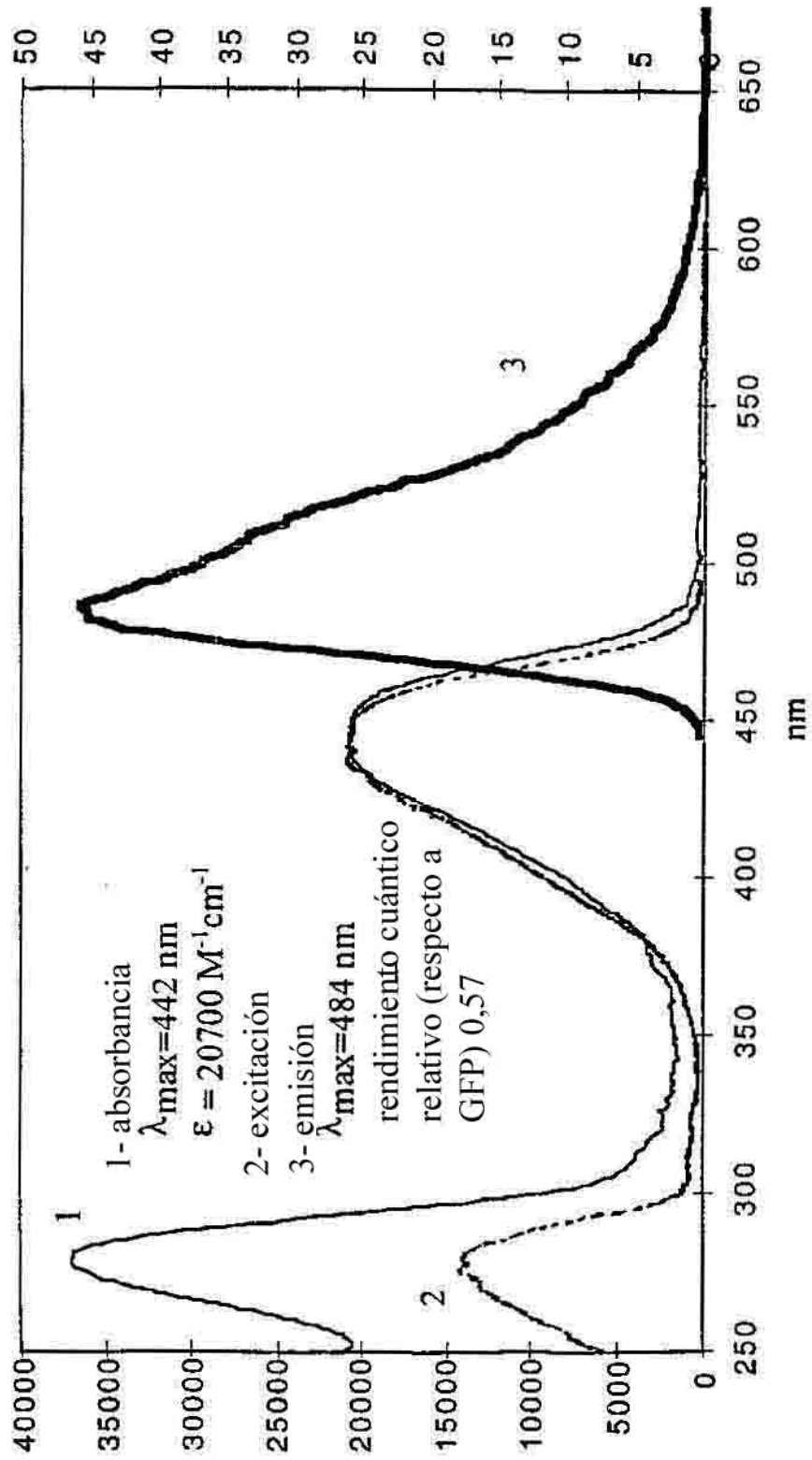


FIG. 7

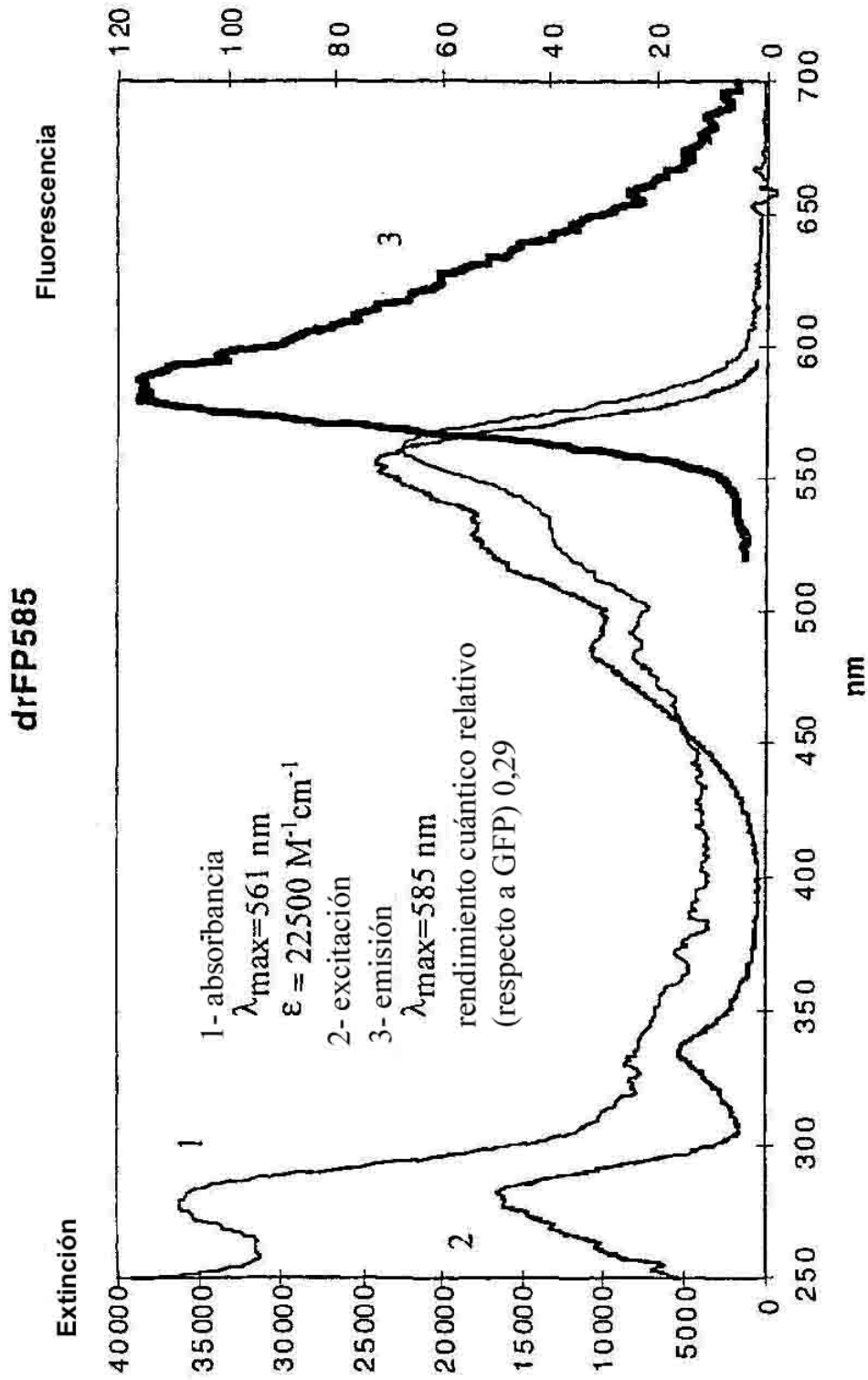


FIG. 8

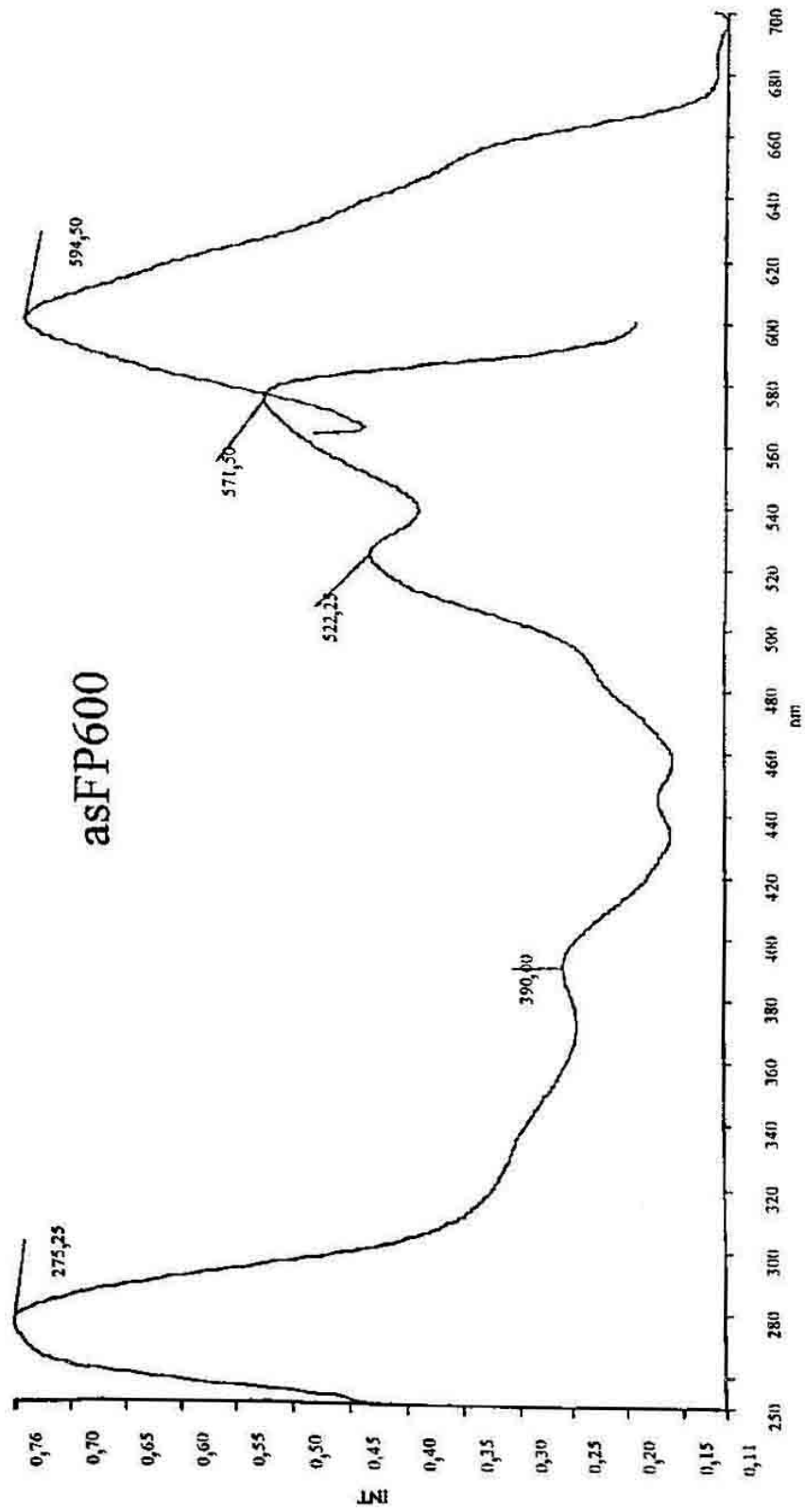


FIG. 9

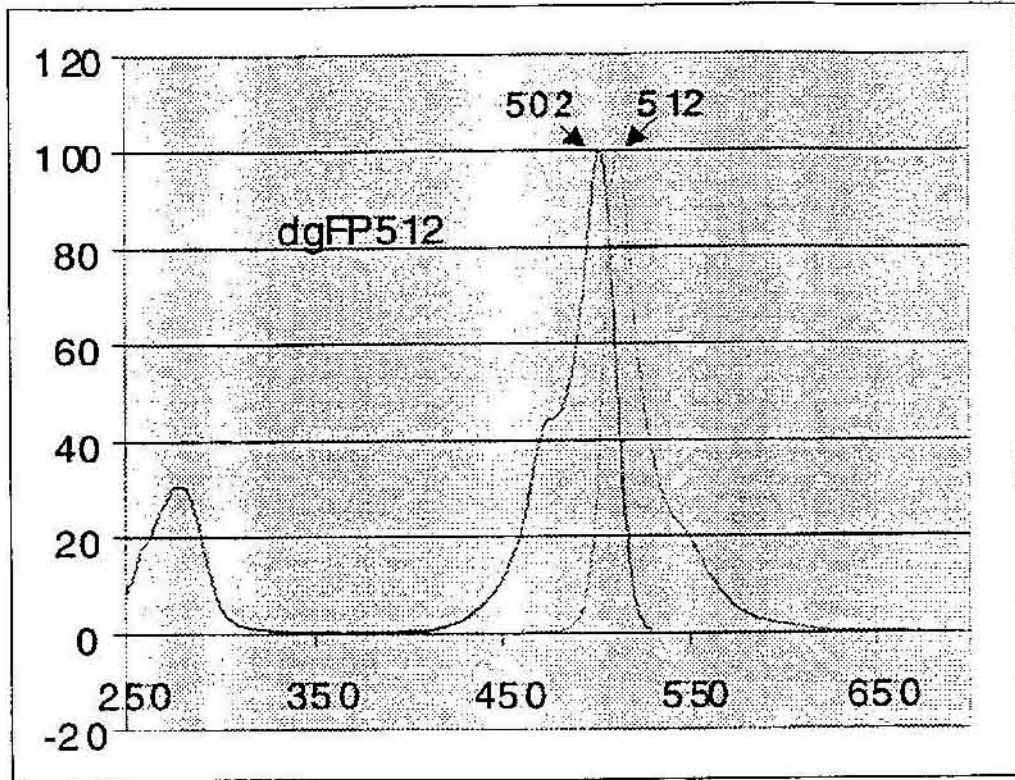


Fig. 10

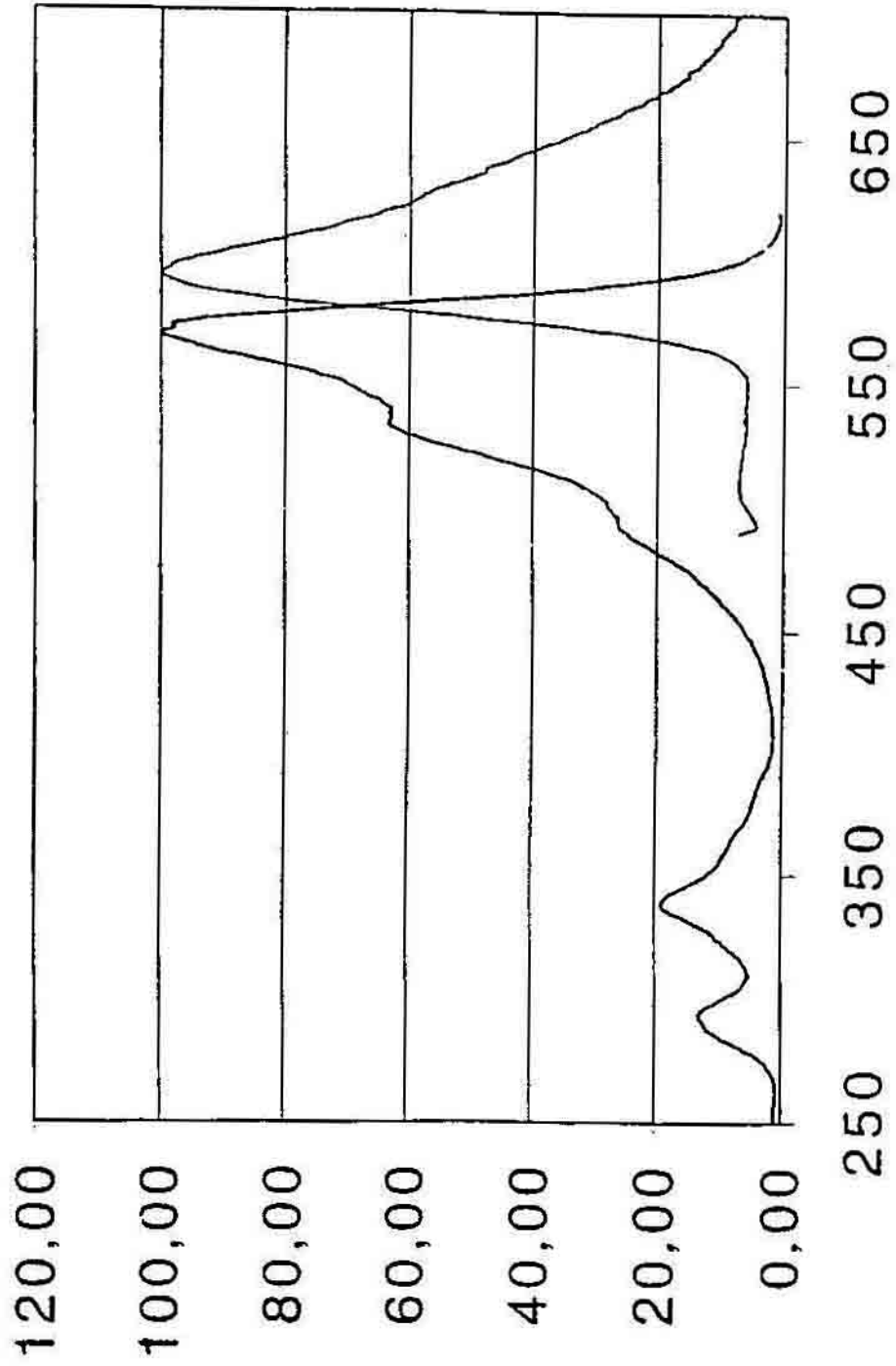


FIG. 11