(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
20 October 2011 (20.10.2011)

PCT

(10) International Publication Number
WO 2011/130343 A1

(51) International Patent Classification:
C12N 15/00 (2006.01) *F21V 9/16* (2006.01)

(21) International Application Number:
PCT/US2011/032214

(22) International Filing Date:
13 April 2011 (13.04.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

61/323,698	13 April 2010 (13.04.2010)	US
61/323,719	13 April 2010 (13.04.2010)	US
61/323,702	13 April 2010 (13.04.2010)	US
61/367,017	23 July 2010 (23.07.2010)	US
61/390,668	7 October 2010 (07.10.2010)	US
61/408,856	1 November 2010 (01.11.2010)	US
61/431,957	12 January 2011 (12.01.2011)	US

(71) Applicant *(for all designated States except US)*: **SIGMA-ALDRICH CO.** [US/US]; 3050 Spruce Street, St. Louis, Missouri 63103 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **FETTER, John** [US/US]; 3050 Spruce Street, St. Louis, Missouri 63103 (US). **MALKOV, Dmitry** [US/US]; 3050 Spruce Street, St. Louis, Missouri 63103 (US). **ZENSER, Nathan** [US/US]; 3050 Spruce Street, St. Louis, Missouri 63103 (US). **SONG, Keming** [US/US]; 3050 Spruce Street, St. Louis, Missouri 63103 (US).

(74) Agents: **BISSEN, Shirley T.** et al.; Polsinelli Shughart PC, 100 South Fourth Street, Suite 1000, St. Louis, Missouri 63102-1825 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

WO 2011/130343 A1

(54) Title: BIOSENSORS COMPRISING PROTEIN-BINDING DOMAINS AND FLUORESCENT PROTEINS

(57) Abstract: The present invention provides compositions and methods for monitoring cellular signal transduction events. In particular, signal transduction events may be monitored by the use of biosensors comprising protein-binding domains, which bind signal-transduction proteins, and fluorescent proteins. Provided herein, therefore, methods for detecting activation of signal transduction proteins or screening for agents that modulate the activity of signal transduction proteins. Also provided are cells comprising the biosensors; lentiviral particles comprising biosensor coding sequence; and kits comprising the lentiviral particles.

BIOSENSORS COMPRISING PROTEIN-BINDING DOMAINS AND FLUORESCENT PROTEINS

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the priority of US provisional application number 61/323,702, filed April 13, 2010, US provisional application number 61/323,719, filed April 13, 2010, US provisional application number 61/323,698, filed April 13, 2010, US provisional application number 61/367,017, filed July 23, 2010, US provisional application number 61/390,668, filed October 7, 2010, US provisional application number 61/408,856, filed November 1, 2010, and US provisional application number 61/431,957, filed January 12, 2011, each of which is hereby incorporated by reference in its entirety.

FIELD OF THE INVENTION

[0002] The present invention generally provides compositions and methods for monitoring cellular signal transduction events. More particularly, the present invention provides biosensors comprising protein-binding domains and fluorescent proteins and methods of using the biosensors for detecting changes in the activity of signal transduction proteins.

BACKGROUND OF THE INVENTION

[0003] Activation of protein tyrosine kinases (PTKs) plays an important role in modulating a wide variety of cellular events, including differentiation, proliferation, migration, metabolism and apoptosis. Within the human genome, 90 PTK genes have been identified. Among all PTKs identified, 58 are receptor tyrosine kinases (RTKs). Activation of RTKs initiates a series of cellular responses to growth factors and hormones that reside in the cellular microenvironment, and therefore, are an important class of drug targets. For example, Epidermal Growth Factor Receptor (EGFR) is a RTK which has been implicated in human breast cancer development, and thus EGFR is one of the most studied RTKs.

[0004] RTKs are activated by the binding of specific ligands, which leads to dimerization of the receptors and autophosphorylation of specific tyrosine residues. These phosphotyrosine residues deliver signals to downstream molecules via their interactions with proteins containing Src homology 2 (SH2) domains. The

protein-binding domain SH2 is approximately 100 amino acids long and recognizes phosphotyrosine residues within specific sequence contexts. SH2 domains are found in many intracellular proteins that generally function as adaptor proteins. For example, Grb2 (Growth factor binding protein 2) is a modular protein containing one SH2 domain that is constitutively associated with downstream regulators. Upon activation of EGFR, the SH2 domain of Grb2 binds a specific phosphorylated tyrosine residue in EGFR, thereby activating downstream kinases and intracellular events. More than 100 SH2 domains have been identified in proteins encoded by the human genome. The binding specificity of these domains plays a critical role in signal transduction within the cell, mediating the interaction of receptor proteins and other intermediary signaling molecules in response to changes in tyrosine phosphorylation states.

[0005] Several methods have been used to detect RTK activation or auto-phosphorylation upon ligand treatment. One of the most commonly used methods is antibody-based immunostaining that uses paired RTK-specific antibodies, one for detection of the total amount of RTK and the other for detection of phospho-RTK. Although widely used, the immunostaining-based method requires cell fixation and thus is not suitable for live cell imaging. The other commonly used method involves over-expression of an RTK-GFP fusion protein and tracks the distribution pattern of the fusion protein in live cells. This method, although useful in some cases, has two major drawbacks: 1) over-expression of an artificial target protein may compete against the endogenous protein and may also interfere with the physiological condition of the cell; and 2) tagging a fluorescent protein (>30KD) to the target RTK protein could inactivate or alter the function of the RTK.

[0006] There is a need, therefore, for improved methods to monitor changes in RTK activities in real time in live cells with minimum disturbance to the cells. In addition, there is a need for rapid, high-throughput screening methods to identity pharmaceutical agents that affect RTK activities.

SUMMARY OF THE INVENTION

[0007] Briefly, therefore, one aspect of the present disclosure provides a cell comprising a biosensor that is able to detect a change in activity of a signal transduction protein. The biosensor comprises at least one protein-binding domain

and a fluorescent protein. Moreover, the cell expresses the biosensor at a level that is substantially similar to the level of expression of the signal transduction protein.

[0008] Another aspect of the disclosure encompasses a method for detecting activation of a signal transduction protein in real time. The method comprises providing a cell comprising a biosensor, wherein the biosensor comprises at least one protein-binding domain and a first fluorescent protein, and the protein-binding domain of the biosensor is able to bind to the signal transduction protein when the signal transduction protein is either activated or inactivated. Additionally, the cell expresses a biosensor at a level that is substantially similar to the level of the signal transduction protein. The method further comprises contacting the cell with a signal that activates the signal transduction protein such that binding between the protein-binding domain of the biosensor and the signal transduction protein changes. Lastly, the method comprises monitoring the biosensor in the cell, wherein a change in the location of the biosensor indicates activation of the signal transduction protein.

[0009] Still another aspect of the disclosure provides a method for determining whether an agent modulates the activity of a signal transduction protein in real time. The method comprises providing a cell comprising a biosensor, wherein the biosensor comprises at least one protein-binding domain and a first fluorescent protein, and the protein-binding domain of the biosensor is able to bind to the signal transduction protein when the signal transduction protein is either activated or inactivated. Additionally, the cell expresses a biosensor at a level that is substantially similar to the level of the signal transduction protein. The method further comprises contacting the cell with the agents. The next step of the method comprises contacting the cell with a signal that activates the signal transduction protein such that binding between the protein-binding domain of the biosensor and the signal transduction protein changes. The final step of the method comprises monitoring the biosensor in the cell relative to a control cell, wherein a change in the location of the biosensor in the cell contacted with the agent relative to the control cell indicates that the agent modulates the activity of the signal transduction protein.

[0010] A further aspect of the present disclosure provides a lentiviral particle comprising a nucleic acid encoding a biosensor. The biosensor comprises at least one protein-binding domain and a fluorescent protein.

[0011] Yet another aspect encompasses a kit for generating a cell comprising a biosensor. The kit comprises a plurality of cells and a plurality of lentiviral particles. Each lentiviral particle comprises a nucleic acid encoding the biosensor, which comprises at least one protein-binding domain and a fluorescent protein.

[0012] Other aspects and iterations of the disclosure are described in more detail below.

REFERENCE TO COLOR FIGURES

[0013] The application file contains at least one photograph executed in color. Copies of this patent application publication with color photographs will be provided by the Office upon request and payment of the necessary fee.

DESCRIPTION OF THE FIGURES

[0014] **FIG. 1** illustrates the generic structure of biosensors that comprise a target-specific binding domain (i.e., "binder") and a fluorescent protein (i.e., "FP").

[0015] **FIG. 2** presents the structure of two EGFR biosensors. Each biosensor comprises one or two $SH2_{Grb2}$ domains and a green fluorescent protein tag (TurboGFP or TagGFP).

[0016] **FIG. 3** illustrates that a biosensor comprising two $SH2_{Grb2}$ domains has improved performance. Presented are fluorescence microscopy images of cells comprising a tagGFP-1x $SH2_{Grb2}$ biosensor (top panels) and cells comprising a tagGFP-2x $SH2_{Grb2}$ biosensor (bottom panels) in the absence or presence of 100 ng/mL of EGF.

[0017] **FIG. 4** depicts development of a stable cell line expressing a tagGFP-2x$SH2_{Grb2}$ biosensor. Single cell clones were selected and assayed for biosensor activity with EGF.

[0018] **FIG. 5** depicts the TagGFP:2x$SH2_{Grb2}$ biosensor translocation to the plasma membrane followed by internalization (granule formation) after treatment with 100 ng/mL EGF.

[0019] **FIG. 6** documents the increasing biosensor internalization over time in individual cells. Individual cells are numbered in the left-most image in the

panel of images. The number of granules per cell as a function of time is plotted at the bottom.

[0020] **FIG. 7** demonstrates that tyrphostin AG 1478 inhibits relocalization of the TagGFP:2x$SH2_{Grb2}$ biosensor. Presented are fluorescence microscopy images of cells treated with AG 1478 (upper panels) or control untreated cells (lower panels) in the absence or presence of 1 μg/mL of EGF. The bar represents 10 μm.

[0021] **FIG. 8** illustrates the specificity of EGFR signaling pathway as monitored by relocalization of the 2x$SH2_{Grb2}$ biosensors. Top panels show cells treated with 1μg/mL of heregulin-β1. Bottom panels show cells exposed to 100 ng/mL of HGF.

[0022] **FIG. 9** demonstrates ratiometric imaging. (**A**) presents the structure of a construct in which a biosensor coding sequence is fused with a GFP sequence through the 2A sequence, and depicts the two separate proteins that are made during translation. (**B**) presents fluorescent images showing the location of each protein and the fluorescence ratio between the two in the absence and presence of 100 ng/mL of EGF. The bar represents 10 μm.

[0023] **FIG. 10** depicts integration of a SH2 biosensor into the *TUBA1B* locus. Presented are schematics of the *TUBA1B* locus, site of integration, design of the SH2 biosensor, and the proteins expressed after successful integration.

[0024] **FIG. 11** presents differential interference contrast (DIC) and fluorescence microscopy images of individual isolated cell clones expressing the SH2 biosensor. Fluorescent images show a time course of SH2 biosensor translocation after exposure to 100 ng/mL of EGF.

[0025] **FIG. 12** depicts fluorescence microscopy images of individual isolated cell clones expressing the SH2 biosensor (upper panels) and RFP-actin (lower panels). Presented is a time course after exposure to 100 ng/mL of EGF.

DETAILED DESCRIPTION OF THE INVENTION

[0026] The present disclosure provides biosensors comprising protein-binding domains and a fluorescent protein, wherein the biosensors are able to detect or “sense” changes in the activity of specific signal transduction proteins. In particular, the protein-binding domain(s) of the biosensor is able to bind a signal

transduction protein that is either activated or inactivated, such that upon activation of the signal transduction protein, the cellular location of the biosensor changes. Additionally, the fluorescent protein of the biosensor provides means to track the change in location of the biosensor in response to activation of the signal transduction protein. Also provided herein are cells engineered to express the biosensor. Advantageously, it has been discovered that relocalization of the biosensor may be visualized in living cells if the cells express the biosensor at a level that is substantially similar to the level of the target signal transduction protein. The disclosure also provides methods for using the biosensor to detect activation of signal transduction proteins or screen for agents that modulate the activity of signal transduction proteins.

(I) Biosensors

[0027] One aspect of the present disclosure encompasses a biosensor comprising at least one protein-binding domain and a fluorescent protein. Stated another way, the biosensor is a fusion protein comprising at least one protein-binding domain and a fluorescent protein. The protein-binding domain of the biosensor binds to a target protein (e.g., a signal transduction protein) under different conditions of activation, and the fluorescent protein of the biosensor provides means to detect and monitor the changes in the localization of the biosensor, which reflects changes in activity of the target protein. The biosensor, therefore, detects and monitors changes in the activity of the target signal transduction protein.

(a) protein-binding domain

[0028] In general, a protein-binding domain is a structurally conserved region of a protein that recognizes and binds a specific site, a specific amino acid sequence, or a specific three-dimensional configuration in another protein. In the context of the present invention, a protein-binding domain is a structurally conserved region of a protein that recognizes and binds a specific site, a specific amino acid sequence, or a specific three dimensional configuration in a signal transduction protein. Protein-binding domains mediate many protein-protein interactions.

[0029] Numerous protein-binding domains and families of protein-binding domains and their target proteins have been identified and investigated.

Many relational databases designed to store protein interactions mediated by protein-binding domains (or protein recognition modules) in human, mouse, rat, yeast, plant, and other organisms are available. For example, interactions mediated by as many as 150 different domain families are stored in DOMINO (http://mint.bio.uniroma2.it/domino/); the largest database of human protein-protein interactions mediated by phosphoprotein binding domains (PPBDs) is the PepCyber :P~Pep, which currently includes 11,269 records of interactions between 387 PPBD proteins and 1,471 substrate proteins, curated from 4,852 published studies (http://pepcyber.biolead.org/PPEP/).

[0030] Preferred protein-binding domains include SH3, SH2, 14-3-3, PDZ, PTB, WW, EVH, VHS, FHA, EH, FF, BRCT, Bromo, Chromo, GYF, C2, MH2, PBD, WD40, and variants thereof. In one embodiment, the preferred protein-binding domain may be a 14-3-3 domain, which binds to phosphoserine-containing proteins. In another embodiment, the protein-binding domain may be BRCT (BRCA1 C Terminus domain), which binds other BRCT modules to form homo/hetero BRCT multimers. In yet another embodiment, the protein-binding domain may be Bromo, a domain found in many chromatin-associated proteins and in nuclear histone acetyltransferases. Bromo interacts specifically with an acetylated lysine of a protein. In an additional embodiment, the protein-binding domain may be a C2 domain, which mostly are Ca2+-dependent membrane-targeting modules that bind phospholipids, inositol polyphosphates, and intracellular proteins. In another embodiment, the protein-binding domain may be Chromo (Chromatin Organization Modifier domain), which often binds to methylated histone tails. Another embodiment provides EH (Eps15 homology domain) as the protein-binding domain that recognizes proteins containing Asn-Pro-Phe (NPF) sequences. In one iteration, the protein-binding domain may be EVH (enabled/VASP homology) domain, which binds G protein-coupled receptor proteins. In another iteration, the protein-binding domain may be FF domain, which binds the hyperphosphorylated C-terminal repeat domain of RNA polymerase II. In an alternate embodiment, the protein-binding domain may be Forkhead associated domain, FHA, which binds phosphothreonine, phosphoserine and sometimes phosphotyrosine. In still another embodiment, the protein-binding domain may be GYF that contains conserved Gly-Tyr-Phe residues and is a proline-binding domain. In another embodiment, the protein-binding domain

may be MAD homology 2 domain, MH2, which binds TGF-beta receptor kinase. In yet another embodiment, the protein-binding domain may be PBD (P21-Rho-binding domain), also known as CRIB (Cdc42/Rac interactive binding domain) that binds small GTPases. Alternatively, the protein-binding domain may be DHR (Dlg homologous region) or GLGF (named after a conserved sequence motif). Through target phosphorylation, many PDZ domains bind C-terminal polypeptides and even to lipids, and PDZ-PDZ domain interactions lead to heterodimerization. In another embodiment, the protein-binding domain may PTB (Phosphotyrosine-binding domain), which binds phosphotyrosine with specificity conferred by residues located amino-terminal to the phosphotyrosine, as compared to SH2 domain's specificity conferred by the adjacent carboxy-terminal residue of phosphotyrosine. In other embodiments, the protein-binding domain may be SH3 or VHS. SH3 domains bind to peptides that are rich in proline, mostly containing the motif PxxP. VHS domains are involved in general membrane targeting/cargo recognition in vesicular trafficking. In another embodiment, the protein-binding domain may be WD40. The 40 residues of a WD40 domain form a propeller-like structure to which proteins can bind either stably or reversibly, and the protein-protein interactions carry out functions in signal transduction, pre-mRNA processing and cytoskeleton assembly. In still another embodiment, the protein-binding domain may be a WW domain, which is a protein module with two highly conserved tryptophans that binds proline-rich peptide motif.

[0031] In a preferred embodiment, the protein-binding domain may be SH2, the Src homology 2 domain. SH2 is a phosphotyrosine binding domain of about 100 amino-acid residues first identified as a region with conserved sequence between the oncoproteins Src and Fps. Similar sequences have been found in many other intracellular signal-transducing proteins. SH2 domains are found in a wide variety of proteins, including phospholipase Cgamma (PLCG1) and the cellular non-receptor protein tyrosine kinases; structural proteins such as tensin (TNS1); a group of small adaptor molecules, i.e CRK and NCK1. Sometimes the SH2 domains are found as repeats in a single protein sequence and will then often bind both mono- and di-phosphorylated substrates. SH2 domains function as regulatory modules in intracellular signaling cascades and some other pathways, such as kinases, adaptors, phosphatases, and so on. In one embodiment, the SH2 domain may be from a natural protein or peptide. More specifically, in one preferred

embodiment, the protein binding domain may be the SH2 domain contained in Grb2 (Growth factor binding protein 2). Grb2 is a modular protein that specifically binds to a phosphotyrosine residue in activated EGFR (Epidermal Growth Factor Receptor), via its SH2 domain, and is constitutively associated with other downstream regulators.

[0032] In some embodiments, the protein-binding domain may be from a natural protein. In other embodiments, the protein-binding domain may be a variant of a natural protein-binding domain, which includes but is not limited to truncated versions, mutated versions, modified versions, and a version with at least one conservative amino acids substitution. Conservative amino acid substitutions, as used herein, include amino acid residues that may be substituted with another amino acid residue having a similar side chain without affecting the function of the protein-binding domain. For example, a group of amino acids having aliphatic side chains is glycine, alanine, valine, leucine, and isoleucine; a group of amino acids having aliphatic-hydroxyl side chains is serine and threonine; a group of amino acids having amide-containing side chains is asparagine and glutamine; a group of amino acids having aromatic side chains is phenylalanine, tyrosine, and tryptophan; a group of amino acids having basic side chains is lysine, arginine, and histidine; and a group of amino acids having sulfur-containing side chains is cysteine and methionine. Preferred conservative amino acid substitution groups include: valine-leucine-isoleucine, phenylalanine-tyrosine, lysine-arginine, alanine-valine, and asparagine-glutamine.

[0033] In other embodiments, the protein-binding domain may be an artificial protein-binding domain that was designed to bind a specific region of the target protein of interest. In yet another embodiment, the protein-binding domain may be a hybrid of two or more protein-binding domains.

[0034] In general, the protein-binding domain or variant thereof may be derived from a human protein, a mouse protein, a rat protein, a mammalian protein, a vertebrate protein, an invertebrate protein, a microbial protein, a bacterial protein, or a viral protein. In an exemplary embodiment, the protein-binding domain or variant thereof may be from a human protein.

[0035] The number of protein-binding domains in a biosensor can and will vary. In one embodiment, the biosensor may comprise one protein-binding

domain. In other embodiment, the biosensor may comprise two protein-binding domains. In a further embodiment, the biosensor may comprise three or more protein-binding domains. In embodiments in which the biosensor comprises more than one protein-binding domain, the protein-binding domains may be identical, the protein-binding domains may be from the same family, or the protein-binding domains may be from different families.

(b) fluorescent protein

[0036] The biosensor also comprises a fluorescent protein. Those of skill in the art appreciate that a variety of fluorescent proteins are suitable for inclusion in the biosensor.

[0037] Non limiting examples of suitable fluorescent proteins include green fluorescent proteins (e.g., GFP, GFP-2, tagGFP, turboGFP, EGFP, Emerald, Azami Green, Monomeric Azami Green, CopGFP, AceGFP, ZsGreen1), yellow fluorescent proteins (e.g. YFP, EYFP, Citrine, Venus, YPet, PhiYFP, ZsYellow1,), blue fluorescent proteins (e.g. EBFP, EBFP2, Azurite, mKalama1, GFPuv, Sapphire, T-sapphire,), cyan fluorescent proteins (e.g. ECFP, Cerulean, CyPet, AmCyan1, Midoriishi-Cyan), red fluorescent proteins (mKate, mKate2, mPlum, DsRed monomer, mCherry, mRFP1, DsRed-Express, DsRed2, DsRed-Monomer, HcRed-Tandem, HcRed1, AsRed2, eqFP611, mRasberry, mStrawberry, Jred), and orange fluorescent proteins (mOrange, mKO, Kusabira-Orange, Monomeric Kusabira-Orange, mTangerine, tdTomato) or any other suitable fluorescent protein.

[0038] Non-limiting examples also include circular permutations of green fluorescent proteins, in which the amino and carboxyl portions are interchanged and rejoined with a short spacer connecting the original termini, while still being fluorescent. These circular permutations of fluorescent protein have altered pKa values and orientations of the chromophore. Furthermore, certain locations within some fluorescent protein tolerate insertion of entire proteins, and conformational changes in the insert can have profound effects on the fluorescence, such as enhancement or changed colors. For example, insertions of calmodulin or a zinc finger domain in place of Tyr-145 of a yellow mutant (EYFP, enhanced yellow fluorescent protein) of GFP result in indicator proteins whose fluorescence can be enhanced several fold upon metal binding. The calmodulin graft into enhanced

yellow fluorescent protein can monitor cytosolic Ca^{2+} in single mammalian cells. In one possible embodiment, a protein-binding domain of a biosensor is inserted in fluorescent protein coding sequence.

(c) biosensor configurations

[0039] The orientation of the protein-binding domain(s) to the fluorescent protein of the biosensor can and will vary. As diagrammed in **FIG.1**, the fluorescent protein may be located at the C-terminal end of the biosensor, or the fluorescent protein may be located at the N-terminal end of the biosensor.

[0040] In embodiments in which the biosensor comprises two or more protein-binding domains, the protein-binding domains may be arranged in tandem at either the N-terminal end or the C-terminal end of the biosensor. Alternatively, the two more protein-binding domains may flank the fluorescent protein. Stated another way, the fluorescent protein may be internal with protein-binding domains at the N-terminal and C-terminal ends of the biosensor.

(d) exemplary biosensors

[0041] In a preferred embodiment, the at least one protein-binding domain of the biosensor may be a SH2 domain. In another preferred embodiment, the biosensor may comprise two SH2 domains. In an iteration of this embodiment, the SH2 domains may be from a Grb2 protein. In an exemplary embodiment, the biosensor may comprise two SH2 domains derived from human Grb2. Exemplary fluorescent proteins include turboGFP and tagGFP. In one exemplary embodiment, the biosensor comprises two SH2 domains derived from human Grb2 linked to turboGFP or tagGFP, with the fluorescent protein at the N-terminal end of the biosensor.

(II) Nucleic acids and vectors comprising nucleic acids encoding the biosensor

(a) nucleic acids

[0042] Another aspect of the disclosure encompasses nucleic acids encoding the biosensors, which are detailed above in section (I). The nucleic acids encoding the biosensors may be deoxyribonucleic acids or ribonucleic acids. In

some embodiments, the nucleic acids encoding the biosensors may be linear nucleic acid molecules. In other embodiments, the nucleic acids encoding the biosensors may be part of a vector, as detailed below.

(i) expression control sequences

[0043] In some embodiments, the nucleic acid encoding the biosensor may be operably linked to an expression control sequence. Although expression is regulated by many sequence elements, the term "promoter" is used below for ease of discussion. In some embodiments it may be desirable to use a constitutive promoter that is active in many types of cells or under different conditions. Non-limiting examples of suitable constitutive promoters include CMV, PGK, SV40, MMTV, adenovirus Ela, immediate early, immunoglobulin heavy chain, and RSV-LTR promoters. In other embodiments, it may be desirable to use an inducible promoter such that expression of the biosensor may be easily regulated. Examples of suitable inducible promoters include tetracycline-inducible promoters and those regulated by metal ions (e.g., metallothionein-1 promoters), steroid hormones, small molecules, heat shock, and the like.

[0044] Those of skill in the art will appreciate that the nucleic acid encoding the biosensor may also be operably linked to other transcriptional and translational control elements. Suitable transcription or translation control sequences include but are not limited to upstream control elements, enhancer elements, TATA boxes, cis regulatory regions, activator binding regions, repressor binding regions, transcription initiation sites, polyadenylation control elements, transcription termination sites, ribosome binding sites, translation initiation sites, and translation termination sites.

(ii) *optional sequence encoding a second fluorescent protein*

[0045] In certain embodiments, the nucleic acid encoding the biosensor may be linked to a nucleic acid encoding a second fluorescent protein that differs from the fluorescent protein in the biosensor. Suitable fluorescent proteins are detailed above in section (I)(b). The nucleic acid encoding the biosensor and the nucleic acid encoding the second fluorescent proteins may be linked by a sequence encoding a 2A peptide, such that two separate proteins are made during translation.

Alternatively, the nucleic acid encoding the biosensor and the nucleic acid encoding the second fluorescent proteins may be separated by an internal ribosome binding or entry site. Regardless of how the nucleic acid is constructed, the end result is that two separate proteins are made during translation (i.e., the biosensor and the second fluorescent protein), wherein the second fluorescent protein may be used as an internal standard or may be used for ratio imaging as detailed below.

[0046] As mentioned above, the nucleic acid encoding the biosensor and the nucleic acid encoding the second fluorescent proteins may be linked by a 2A coding sequence. As used herein, the term "2A peptide" refers to any 2A peptide or fragment thereof, any 2A-like peptide or fragment thereof, or an artificial peptide comprising the requisite amino acids. The 2A peptide was originally characterized in positive-strand RNA viruses, which produce a polyprotein that is "cleaved" during translation into mature individual proteins. More specifically, the 2A peptide region (~20 amino acids) mediates "cleavage" at its own C-terminus to release itself from the 2B region of the polyprotein. 2A peptide sequences terminate with a glycine and a proline residue. During translation of a 2A peptide, the ribosome pauses after the glycine residue, resulting in release of the nascent polypeptide chain. Translation resumes, with the proline residue of the 2A sequence becoming the first amino acid of the downstream protein.

[0047] In some embodiments, the 2A peptide coding sequence that links the biosensor coding sequence with the second fluorescent protein coding sequence may code for a full length 2A peptide. In other embodiments, the 2A peptide coding sequence may code for a C-terminal fragment of a 2A peptide. The C-terminal fragment may comprise about 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, or 4 amino acid residues of the C-terminal end of the 2A peptide.

[0048] The 2A peptide coding sequence may be linked to the 5' end or 3' end of the nucleic acid encoding the biosensor. Accordingly, the resultant nucleic acid may have the following configurations: 5'-(biosensor sequence)-(2A peptide sequence)-(2^{nd} fluorescent protein sequence)-3', or 5'-(2^{nd} fluorescent protein sequence)-(2A peptide sequence)-(biosensor sequence)-3'.

(b) vectors

[0049] A further aspect of the disclosure provides vectors comprising the nucleic acid encoding the biosensor as detailed above in section (I). In some embodiments, the vector may comprise the nucleic acid encoding the biosensor. In other embodiments, the vector may comprise the nucleic acid encoding the biosensor that is operably linked to expression control sequences. In further embodiments, the vector may comprise the nucleic acid encoding the biosensor, which is operably linked to expression control sequences, as well as a nucleic acid encoding a second fluorescent protein.

[0050] The vector may be used to introduce the nucleic acid encoding the biosensor into a cell of interest and/or regulate expression of the biosensor in the cell of interest. Suitable vectors include plasmids, phagemids, cosmids, BACS, and viral vectors, such as, e.g., adenoviruses, adeno-associated viruses, herpes viruses, lentiviruses, retroviruses, and so forth.

[0051] In some embodiments, the vector comprising the nucleic acid encoding the biosensor and any optional accessory sequences may be an integrating vector. Integrating vectors not only introduce the nucleic acid of interest into a cell, but also integrate the nucleic acid of interest into a location in a chromosome of the cell. Typically, the nucleic acid of interest is integrated randomly into a site in the chromosome. Viral vectors are examples of integrating vectors.

[0052] In other embodiments, the vector comprising the nucleic acid encoding the biosensor and any optional accessory sequences may be an expression vector. Expression vectors generally contain origins of replication such that they remain extrachromosomal and regulate expression of the nucleic acid of interest. Expression vectors suitable for expression in eukaryotic cells including yeast, avian, and mammalian cells are known in the art. One example of an expression vector is pcDNA3 (Invitrogen, San Diego, Calif.), in which transcription is driven by the cytomegalovirus (CMV) early promoter/enhancer.

[0053] In a preferred embodiment, the vector comprising the nucleic acid encoding the biosensor and optional accessory sequences may be a lentiviral vector. An advantage of lentiviral vectors is that they are able to replicate in non-dividing cells. In an exemplary embodiment, the nucleic acid encoding the biosensor and any optional accessory sequences may be within a lentiviral particle. Lentiviral

particles are ready to deliver and integrate the nucleic acid(s) of interest to a cell of interest. In some embodiments, the lentiviral particles may be highly purified. Methods of making lentiviral particles are well known in the art.

(III) Kits for generating cells comprising a biosensor

[0054] A further aspect of the disclosure provides kits for generating cells comprising the disclosed biosensors. The kits comprise a plurality of vectors comprising the nucleic acid encoding the biosensor (and optional accessory sequences) and a plurality of cells to be transfected.

(a) vectors

[0055] Suitable vectors comprising the nucleic acid encoding the biosensor and any optional accessory sequences are detailed above in section (II)(b). In an exemplary embodiment, the kits comprise lentiviral particles comprising the nucleic acid encoding the biosensor and any optional accessory sequences.

(b) cells

[0056] A variety of cells are suitable for inclusion in the kits. Typically the cells provided in the kits will be cultured cells. Suitable cultured cells include human cell lines, mammalian cell lines, and non-mammalian cell lines. The cell line may be adherent or non-adherent, or the cell line may be grown under conditions that encourage adherent, non-adherent or organotypic growth using standard techniques known to individuals skilled in the art. Non-limiting examples of suitable mammalian cell lines include Chinese hamster ovary (CHO) cells, monkey kidney CVI line transformed by SV40 (COS-7), human embryonic kidney line 293, baby hamster kidney cells (BHK), mouse sertoli cells (TM4), monkey kidney cells (CVI-76), African green monkey kidney cells (VERO), human cervical carcinoma cells (HeLa), canine kidney cells (MDCK), buffalo rat liver cells (BRL 3A), human lung cells (W138), human liver cells (Hep G2), mouse mammary tumor cells (MMT), rat hepatoma cells (HTC), HIH/3T3 cells, the human U2-OS osteosarcoma cell line, the human A549 cell line, the human A-431 cell line, the human K562 cell line, the human HEK293 cell lines, the human HEK293T cell line, and TRI cells. For an extensive list of mammalian cell lines, those of ordinary skill in the art may refer to

the American Type Culture Collection catalog (ATCC®, Mamassas, VA). In an exemplary embodiment, the cells may be A549, A-431, HeLa or COS-7.

(c) additional components

[0057] In other embodiments, the kits may comprise targeting endonucleases or nucleic acids encoding targeting endonucleases such that the sequence encoding the biosensor may be targeted to a specific chromosomal sequence. T he targeting endonuclease may be a naturally-occurring protein or an engineered protein. In some embodiments, the targeting endonuclease may be a meganuclease or a homing endonuclease. In other embodiments, the targeting endonuclease may be a transcription activator-like effector (TALE)-nuclease. In preferred embodiments, the targeting endonuclease may be a zinc finger nuclease. Typically, a zinc finger nuclease comprises a DNA binding domain (i.e., zinc finger) and a cleavage domain (i.e., nuclease).

[0058] Such kits generally also include donor polynucleotides comprising the nucleic acids encoding the biosensor and any optional accessory sequences. The nucleic acids encoding the biosensor and any optional accessory sequences are flanked by upstream and downstream sequences that share substantial sequence identity with sequences at either side of the site of integration in the targeted chromosomal sequence. For more detail regarding ZFNs and methods of using to edit chromosomal regions see PCT/US2010/43167, the disclosure of which is incorporated by reference herein in its entirety.

[0059] In some embodiments, the kit may further comprise at least one additional component. Suitable components include transfection reagents, agents to enhance vector delivery, culture media for growing the cells, control vectors, dilution reagents, and the like. Typically, the kits also include instructions for transfecting the cells with the vectors comprising the nucleic acid encoding the biosensor and any optional accessory sequences.

(IV) Cells comprising the biosensor

[0060] Yet another aspect of the disclosure encompasses cells comprising the disclosed biosensor. As detailed above in section (I), the biosensor is able to detect changes in the activity of a target signal transduction protein.

Additionally, the cells disclosed herein express the biosensor at a level that is substantially similar to the level of the target signal transduction protein.

(a) signal transduction proteins

[0061] The cells comprising the biosensor express the signal transduction protein of interest. Preferentially, the target signal transduction protein is an endogenous protein that is expressed from an endogenous chromosomal location. Alternatively, the target signal transduction protein may be an exogenously-introduced protein that is expressed from a chromosomally integrated nucleic acid sequence or an extrachromosomal nucleic acid sequence. The exogenous nucleic acid sequence may be overexpressed, however.

[0062] The identity of the target signal transduction protein can and will vary. Non-limiting examples of suitable signal transduction proteins include G protein-coupled receptors, transmembrane receptors, ligand-gated ion channels, voltage-gated ion channels, cytoplasmic protein kinases, serine-threonine kinases, protein phosphatases, phosphatidylinositol kinases, phospholipases and receptor tyrosine kinases, which may belong to the same or different groups for classification purposes. Classifying by compartmentation, there are intracellular receptors and cell-surface receptors.

[0063] In one embodiment, the target signal transduction protein may be an intracellular receptor. Intracellular receptors are soluble proteins localized within the nucleoplasm or the cytoplasm. Non-limiting examples include: cytoplasmic protein kinases, serine-threonine kinases, protein phosphatases, phosphatidylinositol kinases, phospholipases and non-receptor tyrosine kinases. Protein kinases transfer a phosphate group and covalently attach it to a serine, threonine, tyrosine, or histidine residue. At least 125 of the 500+ human protein kinases are serine/threonine kinases. With most kinases acting on both serine and threonine, others act on tyrosine, yet some others act on all three. While the catalytic domain of these kinases is highly conserved, the sequence variation in the kinase encoding genes provides for specificity in recognition of distinct substrate. Phosphorylation of the substrate by protein kinase results in a functional change of the substrate, which may be changing enzyme activity, cellular location, or association with other proteins.

[0064] In another embodiment, the target signal transduction protein may be a cell-surface receptor with or without a transmembrane structure. Non-limiting examples of cell-surface receptors include G-protein coupled receptors, integrins, toll-like receptors, ligand-gated ion channel receptors, and receptor tyrosine kinases.

(i) G protein-coupled receptors

[0065] Also known as GPCR, G protein-coupled receptors encompass a large group of transmembrane eukaryotic receptors that sense molecules outside the cell and activate inside signal transduction pathways. In one embodiment, the targeted signal transduction protein may be a GPCR. The G protein-coupled receptor is activated by a ligand or other signal mediator, which creates a conformational change in the receptor that causes activation of a G protein. The type of G protein activates more specific downstream effects. Non-limiting examples of the ligands include sensory signal mediators (e.g., light and olfactory stimulatory molecules); adenosine, bombesin, bradykinin, endothelin, γ-aminobutyric acid (GABA), hepatocyte growth factor (HGF), melanocortins, neuropeptide Y, opioid peptides, opsins, somatostatin, tachykinins, members of the vasoactive intestinal peptide family, and vasopressin; biogenic amines (e.g., dopamine, epinephrine, norepinephrine, histamine, glutamate (metabotropic effect), glucagon, acetylcholine (muscarinic effect), and serotonin); chemokines; lipid mediators of inflammation (e.g., prostaglandins, prostanoids, platelet-activating factor, and leukotrienes); and peptide hormones (e.g., calcitonin, C5a anaphylatoxin, follicle-stimulating hormone (FSH), gonadotropin-releasing hormone (GnRH), neurokinin, thyrotropin-releasing hormone (TRH), and oxytocin). GPCRs that act as receptors for stimuli that have not yet been identified are known as orphan receptors. G protein-coupled receptors are involved in many diseases, and are also the target of approximately 30% of modern therapeutic agents.

(ii) Integrins

[0066] During cell-cell or cell-ECM (extracellular matrix) signal transduction, an integrin receptor receives signals through its extracellular domain and induces a conformational change within itself and proteins at the cell surface to

initiate signal transduction. Integrin-mediated signal transduction is achieved through a variety of intracellular protein kinases and adaptor molecules, which include integrin-linked kinase (ILK), Src-family kinases, and GTPases. In one embodiment, the targeted signal transduction protein may be an integrin.

(iii) Toll-like receptors

[0067] Toll-like receptors (TLR) are important components of the immune system in animals, plants and bacteria. All TLR members of have a common TIR (Toll-IL-1 receptor) domain. One subgroup of TLRs includes receptors for interleukins that are produced by macrophages, monocytes and dendritic cells and all have extracellular Immunoglobulin (Ig) domains. Another subgroup of TLRs bind directly or indirectly to molecules of microbial origin. A third subgroup of proteins containing TIR domains are exclusively cytosolic and mediate signaling from proteins of subgroups 1 and 2.

(iv) Ligand-gated ion channel receptors

[0068] Ligand-gated ion channel receptors (LGIC) are a group of transmembrane ion channels that are opened or closed in response to the binding of a chemical messenger, such as a neurotransmitter. For example, the nicotinic acetylcholine receptor is a LGIC activated through binding acetylcholine. Receptor's configuration alters upon ligand binding, which allowing Na^+ ions to flow into the cell through the open channel until the cell membrane is depolarized and an action potential is initiated. In one embodiment, the targeted signal transduction protein is a LGIC.

(v) Receptor tyrosine kinases

[0069] In preferred embodiments, the signal transduction protein may be a receptor tyrosine kinase (RTK). RTKs are transmembrane receptors that, upon ligand binding, undergo dimerization and autophosphorylation of specific tyrosine residues. Adaptor proteins (e.g., those containing SH2 domains) recognize and bind the phosphorylated tyrosine residues, thereby activating downstream signaling processes.

[0070] RTKs have been classified into several families. One family is the epidermal growth factor receptor (EGFR) family (or ErbB protein family), which contains four structurally related receptor tyrosine kinases: ErbB-1 (or EGFR), ErbB-2, ErbB-3 and ErbB-4. Insufficient ErbB signaling in humans is associated with the development of neurodegenerative diseases, such as multiple sclerosis and Alzheimer's disease. In mice, loss of signaling by any member of the ErbB family results in embryonic lethality with defects in organs including the lungs, skin, heart, and brain. Excessive ErbB signaling is associated with the development of a wide variety of types of solid tumor. ErbB-1 and ErbB-2 are found in many human cancers and their excessive signaling may be critical factors in the development and malignancy of these tumors. In one embodiment, the target signal transduction protein may be selected from ErbB-1, ErbB-2, ErbB-3, ErbB-4 or their homologs in other organisms.

[0071] Another family of RTKs is the fibroblast growth factor receptor (FGFR) family. There are four fibroblast growth factor receptor (FGFR) genes in this family, which produce over 48 different isoforms of FGFR through alternate splicing during gene expression. In another embodiment, the target signal transduction protein may be selected from the FGFR isoforms or their homologs in other organisms.

[0072] RTKs also include members of vascular endothelial growth factor receptor (VEGFR) family. Members include VEGFR-1 (Flt-1), VEGFR-2 (KDR/Flk-1), and VEGFR-3. VEGFR-2 appears to mediate almost all of the known cellular responses to VEGF; VEGFR-1 modulates VEGFR-2 signaling; and VEGFR-3 mediates lymphangiogenesis in response to VEGF-C and VEGF-D. In still another embodiment, the target signal transduction protein may be selected from VEGFR-1, VEGFR-2, VEGFR-3, and variant thereof.

[0073] Another RTK family includes those activated by glial cell line-derived neurotrophic factor (GDNF), a family of extracellular signaling molecules or ligands. GDNF-activated RTKs include different isoforms of protein as the products of alternate splicing of the *RET*. *RET* loss of function mutations are associated with the development of Hirschsprung's disease, while gain of function mutations are associated with the development of various types of human cancer, including medullary thyroid carcinoma, multiple endocrine neoplasias type 2A and 2B. In an

alternate embodiment, the target signal transduction protein may be a member of the RET receptor family or variant thereof.

[0074] Still another family of RTKs is the Eph receptor family, which comprises the largest sub-family of RTKs. There are at least 16 Known Eph receptors that can be activated by at least one of the 9 known ephrin ligands. Eph/ephrin families form a principle cell guidance system during vertebrate and invertebrate development, particularly in cell positioning and cell morphology. In one embodiment, the target signal transduction protein may be selected from the Eph receptor family.

(b) types of cells

[0075] The type of cell comprising the biosensor described above can and will vary. In general, the cell will be a eukaryotic cell. In some instances, the cell may be a primary cell, a cultured cell, or immortal cell line cell. Suitable cells include fungi or yeast, such as *Pichia, Saccharomyces,* or *Schizosaccharomyces*; insect cells, such as SF9 cells from *Spodoptera frugiperda* or S2 cells from *Drosophila melanogaster;* and animal cells, such as mouse, rat, hamster, non-human primate, or human cells. Exemplary cells are mammalian. The mammalian cells may be primary cells. In general, any primary cell that is sensitive to double strand breaks may be used. The cells may be of a variety of cell types, e.g., fibroblast, myoblast, T or B cell, macrophage, epithelial cell, and so forth.

[0076] When mammalian cell lines are used, the cell line may be any established cell line or a primary cell line that is not yet described. Suitable mammalian cell lines are presented above in section (III)(b).

[0077] In still other embodiments, the cell may be a stem cell. Suitable stem cells include without limit embryonic stem cells, ES-like stem cells, fetal stem cells, adult stem cells, pluripotent stem cells, induced pluripotent stem cells, multipotent stem cells, oligopotent stem cells, and unipotent stem cells.

[0078] In further embodiments, the cell may be a one-cell embryo. The embryo may be a vertebrate or an invertebrate. Suitable vertebrates include mammals, birds, reptiles, amphibians, and fish. Examples of suitable mammals include without limit rodents, companion animals, livestock, and non-primates. Non-limiting examples of rodents include mice, rats, hamsters, gerbils, and guinea pigs.

Suitable companion animals include but are not limited to cats, dogs, rabbits, hedgehogs, and ferrets. Non-limiting examples of livestock include horses, goats, sheep, swine, cattle, llamas, and alpacas. Suitable non-primates include but are not limited to capuchin monkeys, chimpanzees, lemurs, macaques, marmosets, tamarins, spider monkeys, squirrel monkeys, and vervet monkeys. Non-limiting examples of birds include chickens, turkeys, ducks, and geese. Alternatively, the animal may be an invertebrate such as an insect, a nematode, and the like. Non-limiting examples of insects include *Drosophila* and mosquitoes.

(c) biosensor and its expression

[0079] The cells of the invention may comprise a variety of different biosensors, which are detailed above in section (I). The biosensor, however, is chosen such that it recognizes and binds a specific site or region in the target signal transduction protein. For example, the protein-binding domain of the biosensor may be a SH2 domain such that the biosensor may recognize and bind an activated RTK. Alternatively, the protein-binding domain of the biosensor may be a SH3 domain such that the biosensor may recognize and bind PI3 kinase or phospholipase. Those of skill in the art will appreciate that numerous combinations are possible.

[0080] In some embodiments, the biosensor may be expressed from an extrachromosomal nucleic acid. That is, the nucleic acid encoding the biosensor and associated expression control sequences may be within an extrachromosomal vector (e.g., a plasmid vector), such that expression is regulated by the operably linked expression control sequence (i.e., promoter). Such cells may also be engineered to express a second fluorescent protein.

[0081] In preferred embodiments, the biosensor may be expressed from a chromosomally integrated nucleic acid. The nucleic acid encoding the biosensor and any optional associated sequences may be randomly inserted into a chromosome, such that expression is regulated by the operably linked promoter. Random integration of the biosensor may be accomplished through the use of viral vectors, such as, e.g., lentiviral particles. Alternatively, the nucleic acid encoding the biosensor may be inserted into a targeted location in a chromosome such that expression of the biosensor is regulated by an endogenous promoter. Targeted

insertion of the biosensor may be accomplished by zinc finger nucleases and other targeting endonucleases.

[0082] Regardless of the mode of integration, the cells of the invention express the biosensor at substantially the same level as the signal transduction protein of interest. The phrase "at substantially the same level" as used herein means that the level of expression of the two proteins differs by less than about two-fold, or preferably less than about 50%. Stated another way, the cells of the invention do not overexpress the biosensor relative to the target signal transduction protein. Methods for measuring the level of protein expression are well known in the art. Cells that express the biosensor at substantially the same level as the target signal transduction protein are selected for and propagated using techniques well known in the art.

(d) exemplary cells

[0083] In preferred embodiments, the cell may be a mammalian cell comprising a biosensor that recognizes or "senses" a RTK. Generally, the signal transduction protein of interest is endogenous to the mammalian cell. In one embodiment, the cell comprises a biosensor comprising at least one SH2 domain. Preferentially, the SH2 domain is from a Grb2 protein. Thus, the SH2 domain recognizes and binds phosphotyrosine residues in activated EGFRs. In an exemplary embodiment, the cell comprises a biosensor comprising two $SH2_{Grb2}$ domains. Moreover, this exemplary cell expresses the SH2 biosensor and EGFR at substantially similar levels. The fluorescent protein of the biosensor may be a green fluorescent protein, but other fluorescent proteins also may be used.

(V) Methods for detecting activation of a signal transduction protein

[0084] A further aspect of the disclosure provides methods for detecting activation of a signal transduction protein in real time. The method comprises providing a cell comprising a biosensor, as detailed above, contacting the cell with a signal, and monitoring the location of the biosensor in the cell. As detailed in the Examples, upon activation of a signal transduction protein, the protein-binding domain of the biosensor binds to the signal transduction protein, and, as a consequence, the cellular location of the biosensor changes. By monitoring the

cellular translocation of fluorescent biosensor, one may monitor the activation and movement of the target signal transduction protein in live cells in real time.

(a) providing a cell comprising the biosensor

[0085] The first step of the method comprises providing a cell expressing the biosensor at substantially the same level as the target signal transduction protein. Suitable cells comprising the biosensor are detailed above in section (IV) and suitable biosensors are detailed above in section (I). In some embodiments, the cell may be an isolated cell and the method for detecting activation of a signal transduction protein may be carried out *in vitro*. In other embodiments, the cell may be within an organ, tissue, or organism and the method for detecting activation of a signal transduction protein may be carried out *in situ* or *in vivo*.

[0086] In general, the protein-binding domain of the biosensor recognizes and binds a specific site or region in the target signal transduction protein, wherein the binding changes upon activation or inactivation of the signal transduction protein. Accordingly, the cellular location of the biosensor changes spatially or temporally when the activity of the signal transduction protein changes. For example, the protein-binding domain of the biosensor may bind the target signal transduction protein when it is activated. Thus, upon activation of the signal transduction protein, the biosensor will translocate and bind to the signal transduction protein, thereby changing the cellular location of the biosensor. Alternatively, the protein-binding domain of the biosensor may bind the target signal transduction protein when it is an inactivate state. Thus, upon activation of the signal transduction protein, the biosensor will unbind the signal transduction protein and translocate to another location in the cell.

(b) contacting the cell with a signal

[0087] The method further comprises contacting the cell comprising the biosensor with a signal that activates the signal transduction protein. Suitable signal transduction proteins are detailed above in section (IV)(a). The signal is chosen such that it activates the target signal transduction protein. Suitable signals include growth factors, hormones, neurotransmitters, small molecule ligands, cytokines,

ions, and lipids. Exemplary signals include EGF, FGF, VEGF, GDNFs, TGFs, ephrins, BMPs, insulin, insulin-like growth factors, HGF, interleukins, epinephrine, norepinephrine, dopamine, serotonin, gastrin, cholecystokinin, thyroid hormone, erythropoietin, and so forth. In an exemplary embodiment, the signal may be EGF.

[0088] Depending on the nature of the signal used to activate the signal transduction protein, the cell is typically directly contacted with the signal. For example, the cell may be directly contacted with the signal by suspending the cells in a culture medium containing the signal. Alternatively, the cell may be directly contacted with the signal by administration of the signal to the organism comprising the cell. The administration may be via oral means or via parenteral means (e.g., injection).

[0089] The amount of signal contacted with the cell comprising the biosensor can and will vary. Those of skill in the art are familiar with means for determining the appropriate concentration of the signal to activate the target signal transduction protein.

[0090] The period of time the signal is contacted with the cell comprising the biosensor may vary. Typically, the signal transduction protein is immediately activated upon contact with the signal.

(c) monitoring the biosensor

[0091] The method further comprises monitoring the fluorescent biosensor in the cell, wherein a change in the cellular location of the biosensor indicates activation of the target signal transduction protein.

[0092] The monitoring step may be conducted in live cells or fixed cells. Preferentially, the monitoring is conducted in live cells in real time. Live-cell imaging techniques include a wide spectrum of imaging modalities, including fluorescence microscopy, widefield fluorescence, confocal, multiphoton, total internal reflection, FRET, lifetime imaging, superresolution, and transmitted light microscopy. In a preferred embodiment, the cell comprising the biosensor is imaged under low light fluorescence microscopy over time, such that the relocaliztion of the biosensor may be followed over the course of minutes or hours.

[0093] Upon activation of the target signal transduction protein, the binding of the protein-binding domain of the biosensor changes such that the

location of the biosensor changes in the cell. For example, the biosensor may be uniformly distributed throughout the cytoplasm of the cell in the absence of the signal. Upon activation of the signal transduction protein, the biosensor may be translocated to the cell periphery, to the plasma membrane, to the cell nucleus, to endocytotic vesicles, to a cytoskeletal component, and the like. Alternatively, in the absence of the signal, the biosensor may be located at the plasma membrane but be relocated to other cellular compartments upon activation of the signal transduction protein.

[0094] In an exemplary embodiment, the signal transduction protein may be EGFR, the biosensor may comprise two SH2 protein-binding domains, and the signal may be EGF. Prior to contact with the signal, the SH2 biosensor may be uniformly distributed throughout the cytoplasm of the cell. Upon contact of the cell with EGF, the SH2 biosensor initially may translocate to the plasma membrane (where it binds the phosphorylated tyrosine residue in the EGFR) and then the biosensor may translocate to endocytotic vesicles (as the EGFR undergoes endocytosis). See **FIG. 5** for a time course of the relocalization of the SH2 biosensor upon contact with EGF.

[0095] In some embodiments, the cell may further express a second fluorescent protein, as detailed above. The second fluorescent protein differs from the first fluorescent protein in the biosensor, such that the second fluorescent protein may be used an internal standard. For example, the ratio of the fluorescence signal of the fluorescent biosensor to the fluorescence signal of the second fluorescent protein may be used to account for differences in cell thickness and/or to improve the sensitivity of the fluorescence signal.

(VI) Methods for screening agents that may modulate the activity of a signal transduction protein

[0096] Still another aspect of the disclosure encompasses methods for determining whether agents modulate the activity of signal transduction proteins. The method comprises providing a cell comprising a biosensor, contacting the cell with the agent, contacting the treated cell with a signal, and monitoring the biosensor in the cell relative to a cell not contacted with the agent, wherein a change in the location of the biosensor between the two cells indicates that the agent modulates

the activity of the signal transduction protein. Examples 4 and 5 detail use of the method to identify agents that inhibit EGFR activity.

[0097] The method comprises providing a cell comprising the biosensor. Suitable cells comprising the biosensor are detailed above in section (V)(a).

[0098] The method further comprises contacting the cell comprising the biosensor with the agent. A variety of agents may be screened by the disclosed method. For example, agents that inhibit the activity of a signal transduction protein may be screened and identified. The types of agents screened will depend upon the identity of the signal transduction protein and the biosensor. For example, numerous small molecule compound libraries have been generated and are available commercially or from various sources.

[0099] The amount of agent contacted with the cell comprising the biosensor can and will vary. Those of skill in the art are familiar with means for determining the appropriate concentration of the agent.

[0100] The period of time the agent is contacted with the cell comprising the biosensor can and will vary. In general, the agent may be contacted with the cell comprising the biosensor for about 1 hour, about 2 hours, about 4 hours, about 8 hours, about 16 hours, or about 24 hours.

[0101] The method further comprises contacting the cells that were contacted with the agent with a signal, as detailed above in section (V)(b).

[0102] The last step of the method comprises monitoring the biosensor in the cell, in a manner similar to that detailed above in section (V)(c). The method for screening agents, however, differs from the previously described method in that the location of the biosensor in the agent-contacted cell is compared to the location of the biosensor in a control cell. The control cell is identical to the agent-contacted cell in all respects except it was not contacted with the agent. The agent may inhibit the activation of the signal transduction protein such that contact with the signal does change the localization of the biosensor. (See **FIG. 7**).

[0103] As various changes could be made in the above compounds, products and methods without departing from the scope of the invention, it is intended that all matter contained in the above description and in the examples given below, shall be interpreted as illustrative and not in a limiting sense.

DEFINITIONS

[0104] Unless defined otherwise, all technical and scientific terms used herein have the meaning commonly understood by a person skilled in the art to which this invention belongs. The following references provide one of skill with a general definition of many of the terms used in this invention: Singleton et al., Dictionary of Microbiology and Molecular Biology (2nd ed. 1994); The Cambridge Dictionary of Science and Technology (Walker ed., 1988); The Glossary of Genetics, 5th Ed., R. Rieger et al. (eds.), Springer Verlag (1991); and Hale & Marham, The Harper Collins Dictionary of Biology (1991). As used herein, the following terms have the meanings ascribed to them unless specified otherwise.

[0105] When introducing elements of the present disclosure or the preferred embodiments(s) thereof, the articles "a", "an", "the" and "said" are intended to mean that there are one or more of the elements. The terms "comprising", "including" and "having" are intended to be inclusive and mean that there may be additional elements other than the listed elements.

[0106] A "gene," as used herein, refers to a DNA region (including exons and introns) encoding a gene product, as well as all DNA regions which regulate the production of the gene product, whether or not such regulatory sequences are adjacent to coding and/or transcribed sequences. Accordingly, a gene includes, but is not necessarily limited to, promoter sequences, terminators, translational regulatory sequences such as ribosome binding sites and internal ribosome entry sites, enhancers, silencers, insulators, boundary elements, replication origins, matrix attachment sites, and locus control regions.

[0107] A “heterologous protein” is a protein that is not native (i.e., foreign) to the cell or organism of interest.

[0108] The terms “nucleic acid” and “polynucleotide” refer to a deoxyribonucleotide or ribonucleotide polymer, in linear or circular conformation, and in either single- or double-stranded form. For the purposes of the present disclosure, these terms are not to be construed as limiting with respect to the length of a polymer. The terms can encompass known analogs of natural nucleotides, as well as nucleotides that are modified in the base, sugar and/or phosphate moieties (e.g.,

phosphorothioate backbones). In general, an analog of a particular nucleotide has the same base-pairing specificity; i.e., an analog of A will base-pair with T.

[0109] The terms “polypeptide” and “protein” are used interchangeably to refer to a polymer of amino acid residues.

[0110] The term “recombination” refers to a process of exchange of genetic information between two polynucleotides. For the purposes of this disclosure, “homologous recombination” refers to the specialized form of such exchange that takes place, for example, during repair of double-strand breaks in cells. This process requires sequence similarity between the two polynucleotides, uses a “donor” or “exchange” molecule to template repair of a “target” molecule (i.e., the one that experienced the double-strand break), and is variously known as “non-crossover gene conversion” or “short tract gene conversion,” because it leads to the transfer of genetic information from the donor to the target. Without being bound by any particular theory, such transfer can involve mismatch correction of heteroduplex DNA that forms between the broken target and the donor, and/or “synthesis-dependent strand annealing,” in which the donor is used to resynthesize genetic information that will become part of the target, and/or related processes. Such specialized homologous recombination often results in an alteration of the sequence of the target molecule such that part or all of the sequence of the donor polynucleotide is incorporated into the target polynucleotide.

[0111] The term “sequence identity” refers to the extent in which two nucleotide sequences are invariant, i.e., the two sequences have the same nucleotide at the same position. Sequence identity is generally expressed as a percentage. Two nucleotide sequences that are identical in sequence and length have 100% sequence identity.

[0112] As used herein, the terms "target site" or "target sequence" refer to a nucleic acid sequence that defines a portion of a chromosomal sequence to be edited and to which a zinc finger nuclease is engineered to recognize and bind, provided sufficient conditions for binding exist.

[0113] Techniques for determining nucleic acid and amino acid sequence identity are known in the art. Typically, such techniques include determining the nucleotide sequence of the mRNA for a gene and/or determining the amino acid sequence encoded thereby, and comparing these sequences to a

second nucleotide or amino acid sequence. Genomic sequences can also be determined and compared in this fashion. In general, identity refers to an exact nucleotide-to-nucleotide or amino acid-to-amino acid correspondence of two polynucleotides or polypeptide sequences, respectively. Two or more sequences (polynucleotide or amino acid) can be compared by determining their percent identity. The percent identity of two sequences, whether nucleic acid or amino acid sequences, is the number of exact matches between two aligned sequences divided by the length of the shorter sequences and multiplied by 100. An approximate alignment for nucleic acid sequences is provided by the local homology algorithm of Smith and Waterman, Advances in Applied Mathematics 2:482-489 (1981). This algorithm can be applied to amino acid sequences by using the scoring matrix developed by Dayhoff, Atlas of Protein Sequences and Structure, M. O. Dayhoff ed., 5 suppl. 3:353-358, National Biomedical Research Foundation, Washington, D.C., USA, and normalized by Gribskov, Nucl. Acids Res. 14(6):6745-6763 (1986). An exemplary implementation of this algorithm to determine percent identity of a sequence is provided by the Genetics Computer Group (Madison, Wis.) in the "BestFit" utility application. Other suitable programs for calculating the percent identity or similarity between sequences are generally known in the art, for example, another alignment program is BLAST, used with default parameters. For example, BLASTN and BLASTP can be used using the following default parameters: genetic code=standard; filter=none; strand=both; cutoff=60; expect=10; Matrix=BLOSUM62; Descriptions=50 sequences; sort by=HIGH SCORE; Databases=non-redundant, GenBank+EMBL+DDBJ+PDB+GenBank CDS translations+Swiss protein+Spupdate+PIR. Details of these programs can be found on the GenBank website. With respect to sequences described herein, the range of desired degrees of sequence identity is approximately 80% to 100% and any integer value therebetween. Typically the percent identities between sequences are at least 70-75%, preferably 80-82%, more preferably 85-90%, even more preferably 92%, still more preferably 95%, and most preferably 98% sequence identity.

[0114] Alternatively, the degree of sequence similarity between polynucleotides can be determined by hybridization of polynucleotides under conditions that allow formation of stable duplexes between regions that share a degree of sequence identity, followed by digestion with single-stranded-specific

nuclease(s), and size determination of the digested fragments. Two nucleic acid, or two polypeptide sequences are substantially similar to each other when the sequences exhibit at least about 70%-75%, preferably 80%-82%, more-preferably 85%-90%, even more preferably 92%, still more preferably 95%, and most preferably 98% sequence identity over a defined length of the molecules, as determined using the methods above. As used herein, substantially similar also refers to sequences showing complete identity to a specified DNA or polypeptide sequence. DNA sequences that are substantially similar can be identified in a Southern hybridization experiment under, for example, stringent conditions, as defined for that particular system. Defining appropriate hybridization conditions is within the skill of the art. See, e.g., Sambrook et al., supra; Nucleic Acid Hybridization: A Practical Approach, editors B. D. Hames and S. J. Higgins, (1985) Oxford; Washington, D.C.; IRL Press).

[0115] Selective hybridization of two nucleic acid fragments can be determined as follows. The degree of sequence identity between two nucleic acid molecules affects the efficiency and strength of hybridization events between such molecules. A partially identical nucleic acid sequence will at least partially inhibit the hybridization of a completely identical sequence to a target molecule. Inhibition of hybridization of the completely identical sequence can be assessed using hybridization assays that are well known in the art (e.g., Southern (DNA) blot, Northern (RNA) blot, solution hybridization, or the like, see Sambrook, et al., Molecular Cloning: A Laboratory Manual, Second Edition, (1989) Cold Spring Harbor, N.Y.). Such assays can be conducted using varying degrees of selectivity, for example, using conditions varying from low to high stringency. If conditions of low stringency are employed, the absence of non-specific binding can be assessed using a secondary probe that lacks even a partial degree of sequence identity (for example, a probe having less than about 30% sequence identity with the target molecule), such that, in the absence of non-specific binding events, the secondary probe will not hybridize to the target.

[0116] When utilizing a hybridization-based detection system, a nucleic acid probe is chosen that is complementary to a reference nucleic acid sequence, and then by selection of appropriate conditions the probe and the reference sequence selectively hybridize, or bind, to each other to form a duplex molecule. A nucleic acid molecule that is capable of hybridizing selectively to a reference

sequence under moderately stringent hybridization conditions typically hybridizes under conditions that allow detection of a target nucleic acid sequence of at least about 10-14 nucleotides in length having at least approximately 70% sequence identity with the sequence of the selected nucleic acid probe. Stringent hybridization conditions typically allow detection of target nucleic acid sequences of at least about 10-14 nucleotides in length having a sequence identity of greater than about 90-95% with the sequence of the selected nucleic acid probe. Hybridization conditions useful for probe/reference sequence hybridization, where the probe and reference sequence have a specific degree of sequence identity, can be determined as is known in the art (see, for example, Nucleic Acid Hybridization: A Practical Approach, editors B. D. Hames and S. J. Higgins, (1985) Oxford; Washington, D.C.; IRL Press). Conditions for hybridization are well-known to those of skill in the art.

[0117] Hybridization stringency refers to the degree to which hybridization conditions disfavor the formation of hybrids containing mismatched nucleotides, with higher stringency correlated with a lower tolerance for mismatched hybrids. Factors that affect the stringency of hybridization are well-known to those of skill in the art and include, but are not limited to, temperature, pH, ionic strength, and concentration of organic solvents such as, for example, formamide and dimethylsulfoxide. As is known to those of skill in the art, hybridization stringency is increased by higher temperatures, lower ionic strength and lower solvent concentrations. With respect to stringency conditions for hybridization, it is well known in the art that numerous equivalent conditions can be employed to establish a particular stringency by varying, for example, the following factors: the length and nature of the sequences, base composition of the various sequences, concentrations of salts and other hybridization solution components, the presence or absence of blocking agents in the hybridization solutions (e.g., dextran sulfate, and polyethylene glycol), hybridization reaction temperature and time parameters, as well as, varying wash conditions. A particular set of hybridization conditions may be selected following standard methods in the art (see, for example, Sambrook, et al., Molecular Cloning: A Laboratory Manual, Second Edition, (1989) Cold Spring Harbor, N.Y.).

EXAMPLES

[0118] The following examples illustrate certain aspects of the invention.

Example 1- EGFR biosensors comprising one or two SH2 domains

[0119] The generic structure of domain-based biosensors is illustrated in **FIG. 1**. The general structure of domain-based biosensors comprises two components: a protein-binding domain (or "binder") that recognizes and binds a target protein and a fluorescent protein that permits monitoring of the cellular location of the biosensor. The protein-binding domain(s) may be located at the N-terminal end or the C-terminal end of the biosensor.

[0120] An EGFR biosensor comprising one SH2 domain, as shown in **FIG. 2** upper panel, was constructed. The chosen protein-binding domain was the Scr homology 2 (SH2) of the Grb2 (growth factor binding 2) protein, which specifically binds EGF-activated RTKs. The fluorescent protein portion of the biosensor was either TurboGFP or TagGFP, each of which is an improved green fluorescent protein. A second biosensor comprising two $SH2_{GRB2}$ domains was also constructed. See **FIG. 2**, lower panel. In this particular construct, the two SH2 domains were positioned side by side, with TurboGFP or TagGFP at the N-terminal.

[0121] A549 cells, which endogenously express EGFR, were transiently transfected with either the tagGFP-1xSH2 domain biosensor or the tagGFP-2xSH2 domain biosensor. Cells expressing the biosensors were exposed to 100 ng/mL of EGF and the location of the biosensors was monitored over time. As shown in **FIG. 3**, upper panel, cells expressing the 1xSH2 biosensor had much less pronounced relocalization of the biosensor (i.e., no granules were visible 45 min after exposure to EGF). In contrast, cells expressing low levels of the 2xSH2 biosensor displayed many fluorescent granules after 45 min (see **FIG. 3**, lower panel). The arrows point to cells expressing high levels of the 2xSH2 biosensor. Because of the high background of the unbound biosensor, relocalization of the biosensor was difficult to observe. In the right-most image in the lower panel of **FIG. 3**, fluorescent granules are slightly visible in the overexpressing cells (arrow). This suggests that it is important to match the expression levels of the biosensor and the target receptor for optimal signal sensitivity.

Example 2- Development of a stable cell line

[0122] The nucleic acid encoding TagGFP-2x$(SH2)_{Grb2}$ was cloned into a lentiviral vector. Lentiviral particles were prepared and used to stably transfect A549 cells with the TagGFP:2x$(SH2)_{Grb2}$ biosensor. Single cell clones were selected and assayed for activity with treatment of EGF. A stable cell line homogenous for biosensor expression and EGF activity was selected. (see **FIG. 4**).

Example 3- TagGFP:2X$(SH2)_{Grb2}$ biosensor relocalization upon EGF treatment

[0123] A549 cells stably transfected with the TagGFP:2x$(SH2)_{Grb2}$ biosensor were treated with 100 ng/ml EGF and the spatial and temporal relocalization of the biosensor was monitored in living cells over time. The cell nuclei were labeled with 1 μM DRAQ5. As shown in **FIG. 5**, EGF treatment led to biosensor translocation to the plasma membrane (compare 0 min to 3 min) followed by increasing internalization into endocytotic vesicles (see discrete punctuate granules at 17 and 69 min).

[0124] The increased internalization of the biosensor over time was quantitated using granularity analysis in Metamorph. The number of granules was counted in individual cells. The cells were numbered as shown in **FIG. 6**, the first image in the top panel. The number in the left hand corner of each of the four images refers to the granule count for the selected cell: No.4. Granule counts for each of the five cells over time are plotted in **FIG. 6**, bottom panel.

Example 4- Inhibitor of EGFR blocked the TagGFP:2X$(SH2)_{Grb2}$ relocalization

[0125] To determine the specificity of the biosensor relocalization, cells expressing the SH2 biosensor were treated with tyrphostin AG 1478, which is a selective inhibitor or ErbB1, an EGFR family member. The cells were first treated with 1 μM AG 1478, and then treated with 1 μg/mL EGF. The control cells were only treated with 1μg/mL EGF. The relocalization of the biosensor was monitored over time. As shown in **FIG. 7**, the translocation of TagGFP:2X$(SH2)_{Grb2}$ to the plasma membrane was blocked in cells treated with AG 1478 in the images taken at 3 min following EGF treatment. Further, the internalization of TagGFP:2X$(SH2)_{Grb2}$ was also blocked in the cells treated with AG 1478, shown in the images taken at 50 min

following EGF treatment. The results showed that the cells comprising the SH2 biosensor can be used to evaluate the inhibitors of signal transduction proteins.

Example 5- Identification of other activating ligands

[0126] A variety of other ligands were tested for their ability to activate EGFR, as monitored by the SH2 biosensor. The tested ligands included: EGF, TGF-α, HGF, PDGF-AB, insulin, IGF-1, NGF-Beta, FGF-acidic, angiopoietin, MSP, Gas6, VEGF, and FLT-3. Most of the ligands were tested for activity with the TagGFP:2X$(SH2)_{Grb2}$ biosensor in live cells. Heregulin-β1 (EGF domain 176-246), however, was tested for activity with a TurboGFP:2X$(SH2)_{Grb2}$ biosensor as compared to EGF, which was expressed as a second fluorescent protein. **Table 1** presents the results. Activity (i.e., relocalization of the SH2 biosensor) was seen with EGF and TGF-α treatment, both of which are ligands specific for EGFR (ErbB-1). Heregulin- β1, a ligand specific for ErbB-3 and ErbB-4, did not show activity at 1 µg/mL (**Table 1** and **FIG. 8**, top panel). HGF, a growth factor that binds HGFR, however, showed activity at 100 ng/mL, but much less than 100 mg/mL of EGF (**Table 1** and **FIG. 8**, bottom panel).

Table 1:

Ligand	Activity
EGF	+++
TGF-α	+++
HGF	+
Heregulin- β1	-
PDGF-AB	-
insulin	-
IGF-1	-
NGF-Beta	-
FGF-acidic	-
angiopoietin	-
MSP	-
Gas6	-
VEGF	-
FLT-3	-

Example 6- Ratiometric imaging improves detection

[0127] A nucleic acid construct was generated in which a sequence

encoding a RFP:2x$(SH2)_{Grb2}$ biosensor linked to a sequence encoding GFP, wherein the two sequences were linked by sequence encoding a 2A peptide. See **FIG. 9A**. During translation, the 2A peptide region (~20 amino acids) mediates "cleavage" at its own C-terminus, and produces two separate proteins: a RFP-2x$(SH2)_{Grb2}$-2A biosensor, and a separate GFP that serves as an internal marker or volume marker.

[0128] Cells expressing both proteins were exposed to 100 ng/mL of EGF. **FIG. 9B** presents images of the localization of the biosensor (red fluorescence) and the GFP volume marker (green fluorescence) over time, as well as the ratiometric image of the two (bottom panels). For ratiometric imaging, the fluorescence signal from the RFP biosensor was divided by the GFP signal in Metamorph. Ratiometric imaging improved detection of the biosensor because it removed the effect of changes in cell thickness on the fluorescence intensity of the biosensor and improved the sensitivity of detection at the thinner periphery of the cells.

Example 7: Targeted Integration into* TUBA1B *Locus Using ZFN Technology

[0129] The following example details a targeted insertion of a SH2 biosensor in the *TUBA1B* locus, such that expression of the biosensor was regulated by the endogenous tubulin promoter. For this, a pair of zinc finger nucleases (ZFNs) was designed to target a location in the *TUBA1B* coding region. One was designed to bind a sequence upstream and the second was designed to bind a sequence downstream of the target site. Upon binding, the ZFN pair will introduce a double-stranded break between the two ZFN binding sequences (see the top of **FIG. 10**). For more details regarding ZFNs and methods of using to edit chromosomal regions see PCT/US2010/43167, the disclosure of which is incorporated by reference herein in its entirely.

[0130] A donor plasmid was constructed that carried the sequence of the biosensor (i.e., GFP linked to two $SH2_{Grb2}$ domains), which was linked at the 3' end with a sequence encoding a 2A peptide. The sequence encoding the biosensor and 2A peptide (i.e., GFP-2x$SH2_{Grb2}$-2A) in the donor plasmid was flanked by upstream and downstream sequences of the ZFN cleavage site in the *TUBA1B* locus. The ZFNs and the donor plasmid were designed such that the SH2 biosensor coding sequence would be inserted in-frame with the endogenous tubulin sequence

just downstream of the start codon. Upon activation of the tubulin promoter, one transcript comprising the tubulin start codon, the biosensor sequence, the 2A peptide sequence, and tubulin sequence would be made. During translation, however, two separate proteins would be made (i.e., the biosensor and tubulin), as depicted in the bottom of **FIG. 10**.

[0131] The donor plasmid and nucleic acids encoding the pair of ZFNs were transfected into A549 cells. The transfected cells were then cultured under standard conditions. Analysis of individual cell clones revealed uniform cytoplasmic GFP fluorescence, indicating the expression of the SH2 biosensor.

[0132] To examine activity of the SH2 biosensor, the cells were exposed to 100 ng/ml of EGF and imaged over time. **FIG. 11** presents a time course of the translocation of the SH2 biosensor. Initially, the SH2 biosensor was translocated to the plasma membrane and then the SH2 biosensor was internalized by endocytosis.

Example 8: Targeted Integration into* ACTB *Locus Using ZFN Technology

[0133] To determine whether the level of expression of the biosensor affected its ability to track the relocalization of its target protein, a SH2 biosensor was inserted into a sequence under the control of a strong promoter. For this, the *ACTB* locus, which encodes β-actin, was chosen. ZFNs were designed to target the *ACTB* locus and introduce a cut site just downstream of the start codon. A donor plasmid was designed to provide the $SH2_{Grb2}$ biosensor sequence, as well as tag the endogenously produced β-actin (i.e., GFP-2x$SH2_{Grb2}$-2A-RFP). The nucleic acids were introduced into A549 cells, and two fluorescent proteins were made (i.e., biosensor and RFP-actin).

[0134] Cells were exposed to 100 ng/ml of EGF and imaged. **FIG. 12** presents a time course of the translocation of the GFP-biosensor and the location of RFP-actin. In many cells it was difficult to track the relocalization of the biosensor to the plasma membrane and endocytotic vesicles (see top panel in **FIG. 12**). In fact, GFP granules were visible only in cells with low levels of fluorescence (see arrow in **FIG. 12**). This example revealed that when high levels of the biosensor were produced, it was difficult to monitor the relocalization of the biosensor because of the high levels of unbound or “free” biosensor.

CLAIMS

What is claimed is:

1. A cell comprising a biosensor, the biosensor being able to detect a change in activity of a signal transduction protein, the biosensor comprising at least one protein-binding domain and a first fluorescent protein, the cell expressing the biosensor at a level that is substantially similar to the level of the signal transduction protein.

2. The cell of claim 1, wherein the signal transduction protein is chosen from a receptor tyrosine kinase, a G protein-coupled receptor, a transmembrane receptor, a ligand-gated ion channel, a voltage-gated ion channel, a cytoplasmic protein kinase, a serine-threonine kinase, a protein phosphatase, a phosphatidylinositol kinase, and a phospholipase.

3. The cell of claim 2, wherein the receptor tyrosine kinase is chosen from an EGFR, a FGFR, a VEGFR, a RET receptor, and an Eph/ephrin receptor.

4. The cell of claim 1, wherein the protein-binding domain is chosen from SH3, SH2, 14-3-3, PDZ, PTB, WW, EVH, VHS, FHA, EH, FF, BRCT, Bromo, Chromo, GYF, C2, MH2, WD40, and variants thereof.

5. The cell of claim 1, wherein the first fluorescent protein is chosen from a green fluorescent protein, a blue fluorescent protein, a cyan fluorescent protein, a yellow fluorescent protein, an orange fluorescent protein, and a red fluorescent protein.

6. The cell of claim 1, wherein the biosensor is expressed from a chromosomally integrated nucleic acid or an extrachromosomal nucleic acid.

7. The cell of claim 6, wherein the chromosomally integrated nucleic acid was integrated randomly using a lentivirus.

8. The cell of claim 6, wherein the chromosomally integrated nucleic acid was integrated in a targeted location using a targeting endonuclease.

9. The cell of claim 1, wherein the cell also expresses a second fluorescent protein at the same level as the biosensor, the second fluorescent protein being different from the first fluorescent protein in the biosensor.

10. The cell of claim 1, wherein the cell is a human cell or a mammalian cell.

11. The cell of claim 1, wherein the signal transduction protein is EGFR and the protein-binding domain is a SH2 domain.

12. The cell of claim 11, wherein the SH2 domain is from a Grb2 protein.1

13. A method for detecting activation of a signal transduction protein in real time, the method comprising:
    (a) providing a cell comprising a biosensor, the biosensor comprising at least one protein-binding domain and a first fluorescent protein, the cell expressing the biosensor at a level that is substantially similar to the level of the signal transduction protein, the protein-binding domain of the biosensor being able to bind to the signal transduction protein when the signal transduction protein is either activated or inactivated;
    (b) contacting the cell with a signal that activates the signal transduction protein such that binding between the protein-binding domain of the biosensor and the signal transduction protein changes; and
    (c) monitoring the biosensor in the cell over time, wherein a change in the location of the biosensor indicates activation of the signal transduction protein.

14. The method of claim 13, wherein the signal transduction protein is chosen from a receptor tyrosine kinase, a G protein-coupled receptor, a transmembrane receptor, a ligand-gated ion channel, a voltage-gated ion channel, a cytoplasmic protein kinase, a serine-threonine kinase, a protein phosphatase, a phosphatidylinositol kinase, and a phospholipase.

15. The method of claim 14, wherein the receptor tyrosine kinase is chosen from an EGFR, a FGFR, a VEGFR, a RET receptor, and an Eph/ephrin receptor.

16. The method of claim 13, wherein the signal is chosen from a growth factor, a small molecule ligand, a hormone, a cytokine, an ion, and a lipid.

17. The method of claim 13, wherein the protein-binding domain is chosen from SH3, SH2, 14-3-3, PDZ, PTB, WW, EVH, VHS, FHA, EH, FF, BRCT, Bromo, Chromo, GYF, C2, MH2, WD40, and variants thereof.

18. The method of claim 13, wherein the first fluorescent protein is chosen from a green fluorescent protein, a blue fluorescent protein, a cyan fluorescent protein, a yellow fluorescent protein, an orange fluorescent protein, and a red fluorescent protein.

19. The method of claim 13, wherein the cell also expresses a second fluorescent protein at the same level as the biosensor, the second fluorescent protein being different from the first fluorescent protein in the biosensor.

20. The method of claim 19, wherein the monitoring step further comprises determining the fluorescence ratio of the first fluorescent protein to the second fluorescent protein.

21. The method of claim 13, wherein the cell is a human cell or a mammalian cell.

22. The method of claim 13, wherein the cell is *in vitro*, *in situ*, or *in vivo*.

23. The method of claim 13, wherein signal transduction protein is EGFR, the signal is EGF, and the protein-binding domain is a SH2 domain.

24. The method of claim 23, wherein the SH2 domain is from a Grb2 protein.

25. A method for determining whether an agent modulates the activity of a signal transduction protein in real time, the method comprising:

(a) providing a cell comprising a biosensor, the biosensor comprising at least one protein-binding domain and a first fluorescent protein, the cell expressing the biosensor at a level that is substantially similar to the level of the signal transduction protein, the protein-binding domain of the biosensor being able to bind to the signal transduction protein when the signal transduction protein is either activated or inactivated;

(b) contacting the cell with the agent;

(c) contacting the cell with a signal that normally activates the signal transduction protein; and

(d) monitoring the biosensor in the cell relative to a control cell, wherein a change in the location of the biosensor in the cell contacted with the agent relative to the control cell indicates that the agent modulates the activity of the signal transduction protein.

26. The method of claim 25, wherein the control cell is as in (a) and (c) but is not contacted with the agent as in (b).

27. The method of claim 25, wherein the agent inhibits the activity of the signal transduction protein.

28. The method of claim 25, wherein the signal transduction protein is chosen from a receptor tyrosine kinase, a G protein-coupled receptor, a transmembrane receptor, a ligand-gated ion channel, a voltage-gated ion channel, a cytoplasmic protein kinase, a serine-threonine kinase, a protein phosphatase, a phosphatidylinositol kinase, and a phospholipase.

29. The method of claim 28, wherein the receptor tyrosine kinase is chosen from an EGFR, a FGFR, a VEGFR, a RET receptor, and an Eph/ephrin receptor.

30. The method of claim 25, wherein the signal is chosen from a growth factor, a small molecule ligand, a hormone, a cytokine, an ion, and a lipid.

31. The method of claim 25, wherein the protein-binding domain is chosen from SH3, SH2, 14-3-3, PDZ, PTB, WW, EVH, VHS, FHA, EH, FF, BRCT, Bromo, Chromo, GYF, C2, MH2, WD40, and variants thereof.

32. The method of claim 25, wherein the first fluorescent protein is chosen from a green fluorescent protein, a blue fluorescent protein, a cyan fluorescent protein, a yellow fluorescent protein, an orange fluorescent protein, and a red fluorescent protein.

33. The method of claim 25, wherein the cell also expresses a second fluorescent protein at the same level as the biosensor, the second fluorescent protein being different from the first fluorescent protein in the biosensor.

34. The method of claim 38, wherein the monitoring step further comprises determining the fluorescence ratio of the first fluorescent protein in the biosensor to the second fluorescent protein.

35. The method of claim 25, wherein the cell is a human cell or a mammalian cell.

36. The method of claim 25, wherein the cell is *in vitro*, *in situ*, or *in vivo*.

37. The method of claim 25, wherein signal transduction protein is EGFR, the signal is EGF, and the protein-binding domain is a SH2 domain.

38. The method of claim 37, wherein the SH2 domain is from a Grb2 protein.

39. A lentiviral particle, the lentiviral particle comprising a nucleic acid encoding a biosensor comprising at least one protein-binding domain and a first fluorescent protein.

40. The lentiviral particle of claim 39, wherein the protein-binding domain is chosen from SH3, SH2, 14-3-3, PDZ, PTB, WW, EVH, VHS, FHA, EH, FF, BRCT, Bromo, Chromo, GYF, C2, MH2, WD40, and variants thereof.

41. The lentiviral particle of claim 39, wherein the first fluorescent protein is chosen from a green fluorescent protein, a blue fluorescent protein, a cyan fluorescent protein, a yellow fluorescent protein, an orange fluorescent protein, and a red fluorescent protein.

42. The lentiviral particle of claim 39, wherein the nucleic acid encoding the biosensor is operably linked to an expression control sequence.

43. The lentiviral particle of claim 42, wherein the expression control sequence is chosen from a CMV promoter, a tetracycline-inducible promoter, a SV40 promoter, a PGK promoter, a MMTV promoter, a metallothionein-1 promoter, an adenovirus Ela promoter, an immediate early promoter, an immunoglobulin heavy chain promoter, and a RSV-LTR promoter.

44. The lentiviral particle of claim 39, wherein the nucleic acid encoding the biosensor is linked to a second nucleic acid encoding a second fluorescent protein by a sequence encoding a 2A peptide, wherein the second fluorescent protein differs from the first fluorescent protein.

45. The lentiviral particle of claim 44, wherein the second fluorescent protein is chosen from a green fluorescent protein, a blue fluorescent protein, a cyan fluorescent protein, a yellow fluorescent protein, an orange fluorescent protein, and a red fluorescent protein.

46. The lentiviral particle of claim 39, wherein the protein-binding domain is a SH2 domain.

47. The lentiviral particle of claim 46, wherein the SH2 domain is from a Grb2 protein.

48. A kit for generating a cell comprising a biosensor, the kit comprising (a) a plurality of cells and (b) a plurality of lentiviral particles, each lentiviral particle comprising a nucleic acid encoding the biosensor, the biosensor comprising at least one protein-binding domain and a first fluorescent protein.

49. The kit of claim 48, wherein the biosensor is able to detect a change in activity of a signal transduction protein.

50. The kit of claim 49, wherein the signal transduction protein is chosen from a receptor tyrosine kinase, a G protein-coupled receptor, a transmembrane receptor, a ligand-gated ion channel, a voltage-gated ion channel, a cytoplasmic protein kinase, a serine-threonine kinase, a protein phosphatase, a phosphatidylinositol kinase, and a phospholipase.

51. The kit of claim 48, wherein the protein-binding domain is chosen from SH3, SH2, 14-3-3, PDZ, PTB, WW, EVH, VHS, FHA, EH, FF, BRCT, Bromo, Chromo, GYF, C2, MH2, WD40, and variants thereof.

52. The kit of claim 48, wherein the first fluorescent protein is chosen from a green fluorescent protein, a blue fluorescent protein, a cyan fluorescent protein, a yellow fluorescent protein, an orange fluorescent protein, and a red fluorescent protein.

53. The kit of claim 48, wherein the nucleic acid encoding the biosensor is operably linked to an expression control sequence.

54. The kit of claim 53, wherein the expression control sequence is chosen from a CMV promoter, a PGK promoter, a tetracycline-inducible promoter, a SV40 promoter, a MMTV promoter, a metallothionein-1 promoter, an adenovirus Ela promoter, an immediate early promoter, an immunoglobulin heavy chain promoter, and a RSV-LTR promoter.

55. The kit of claim 48, wherein each lentiviral particle further comprises a second nucleic acid encoding a second fluorescent protein that differs from the first fluorescent protein in the biosensor, the second nucleic acid encoding the second fluorescent protein being linked to the nucleic acid encoding the biosensor by a sequence encoding a 2A peptide.

56. The kit of claim 48, wherein the plurality of cells is a mammalian cell line cell or a non-mammalian cell line cell.

57. The kit of claim 48, wherein the protein-binding domain is a SH2 domain.

58. The kit of claim 57, wherein the SH2 domain is from a Grb2 protein.

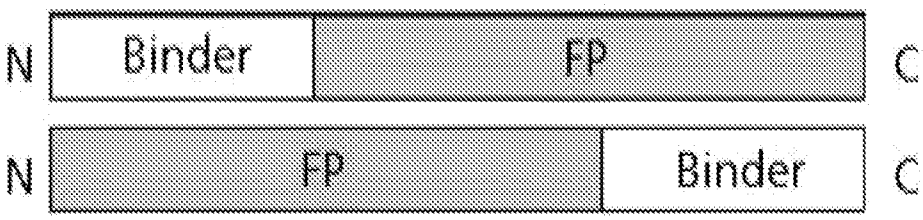
N
Binder
FP
C
N
FP
Binder
C

FIG. 1

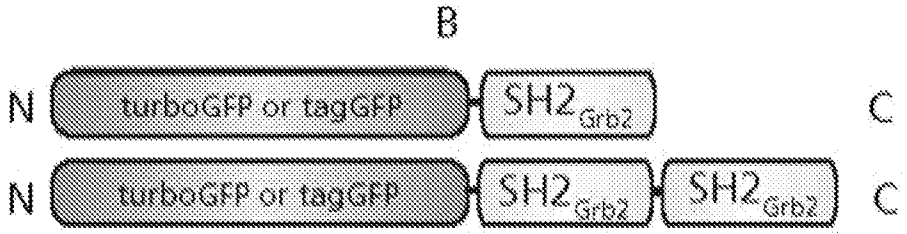
B
N
turboGFP or tagGFP
SH2Grb2
C
N
turboGFP or tagGFP
SH2Grb2
SH2Grb2
C

FIG. 2

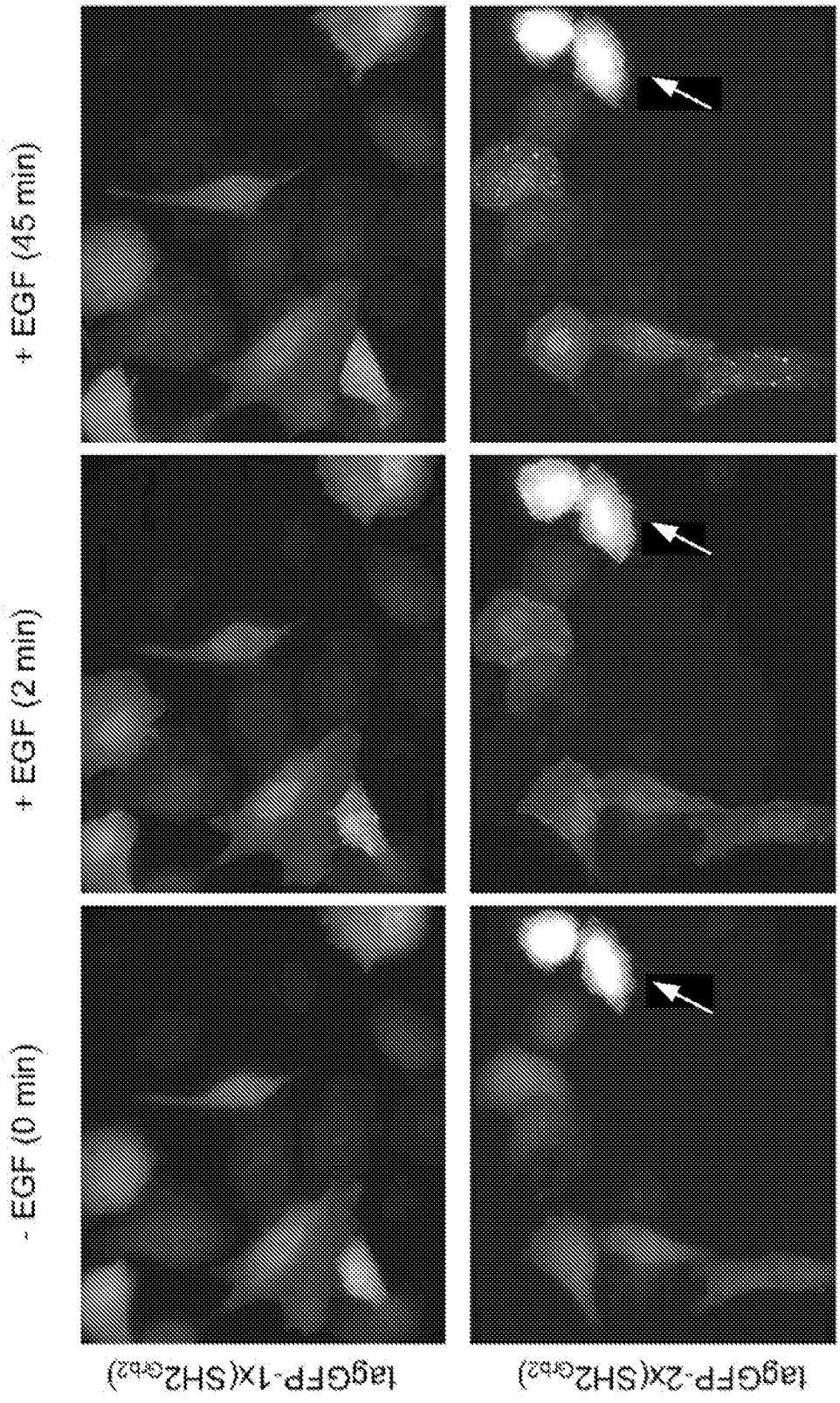
- EGF (0 min)
+ EGF (2 min)
+ EGF (45 min)
tagGFP-1x(SH2Grb2)
tagGFP-2x(SH2Grb2)

FIG. 3

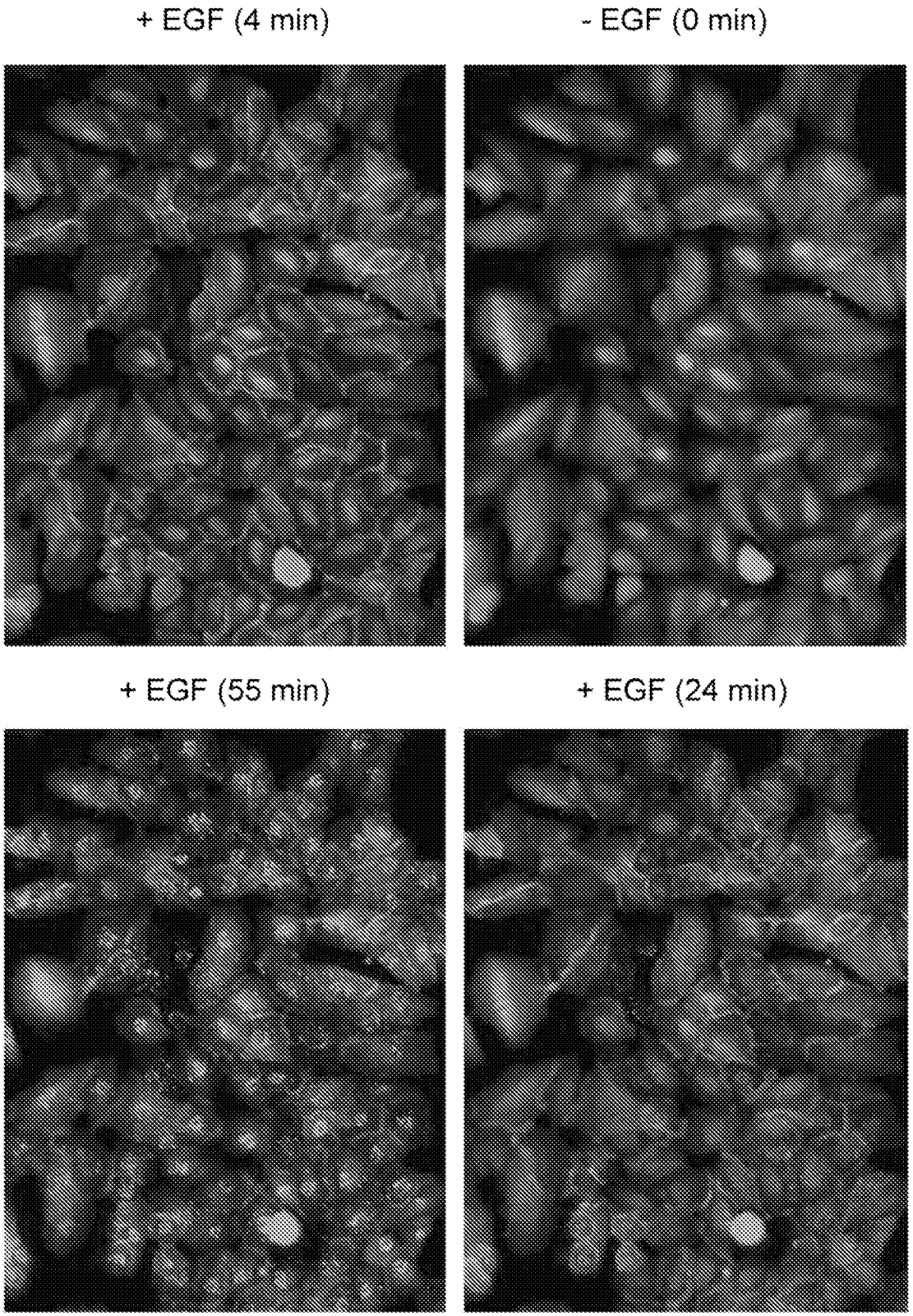

FIG. 4

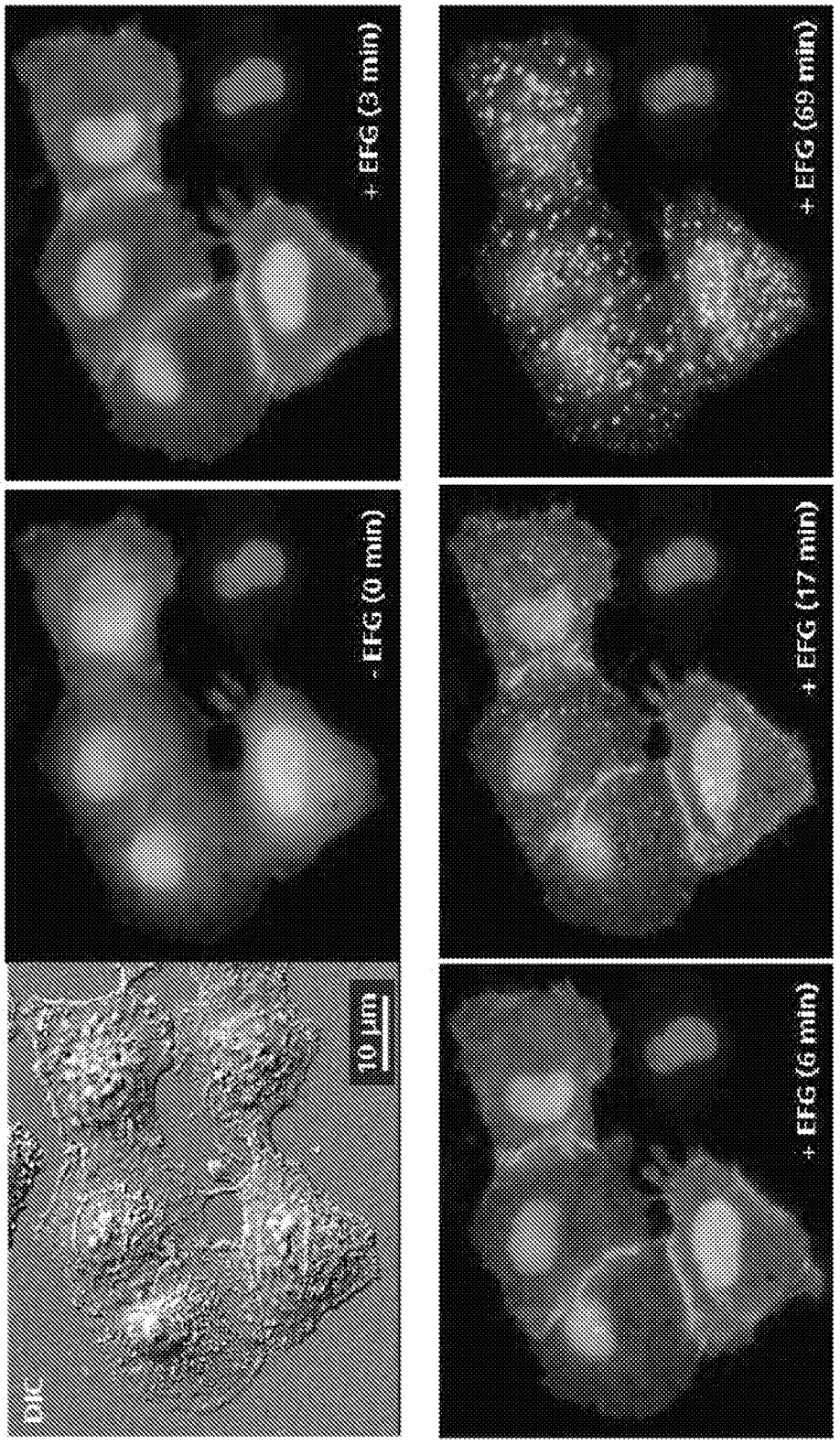

FIG. 5

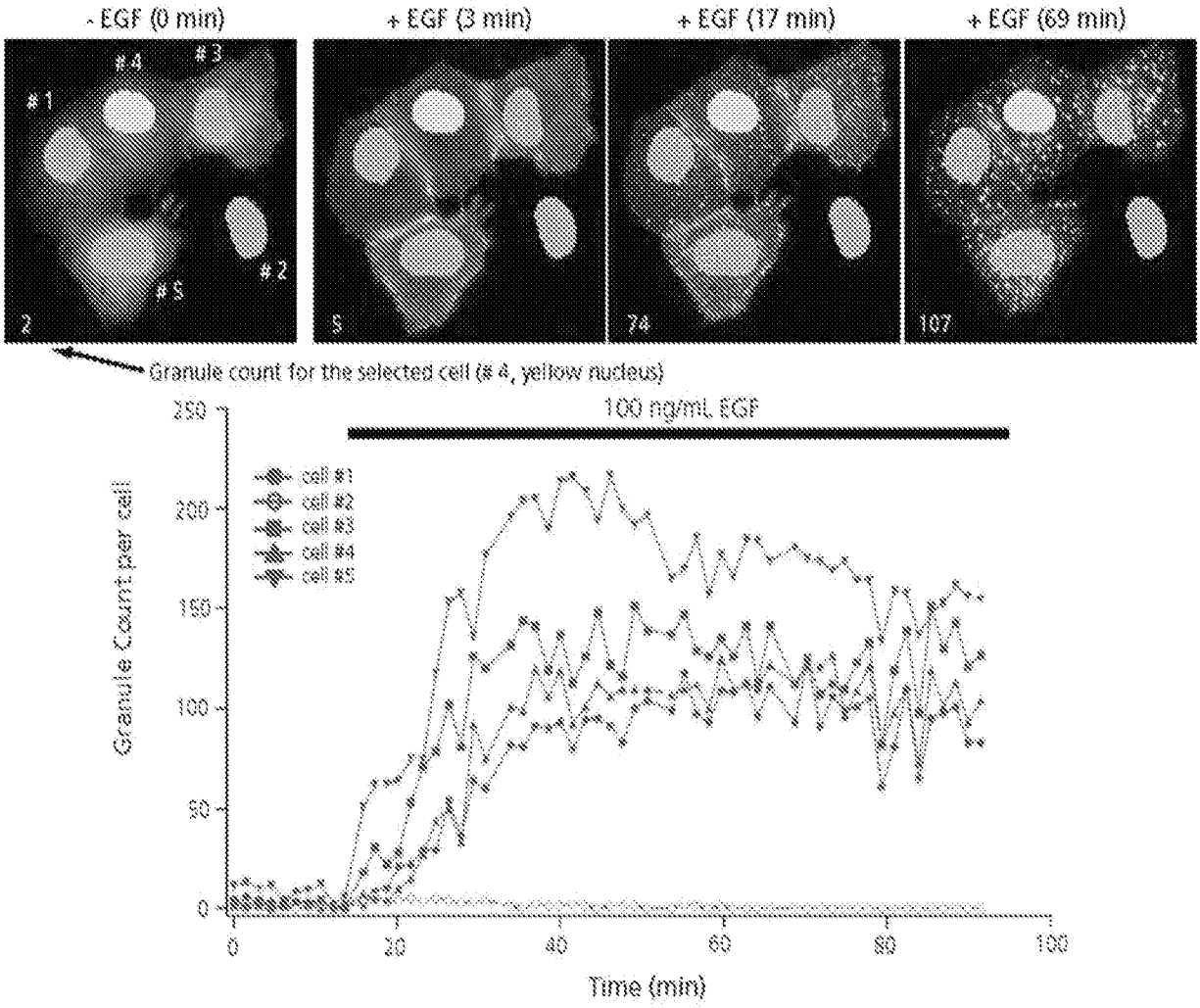
- EGF (0 min)
+ EGF (3 min)
+ EGF (17 min)
+ EGF (69 min)
#1
#2
#3
#4
#5
2
5
74
107
Granule count for the selected cell (# 4, yellow nucleus)
100 ng/mL EGF
cell #1
cell #2
cell #3
cell #4
cell #5
Granule Count per cell
Time (min)
0
50
100
150
200
250
0
20
40
60
80
100

FIG. 6

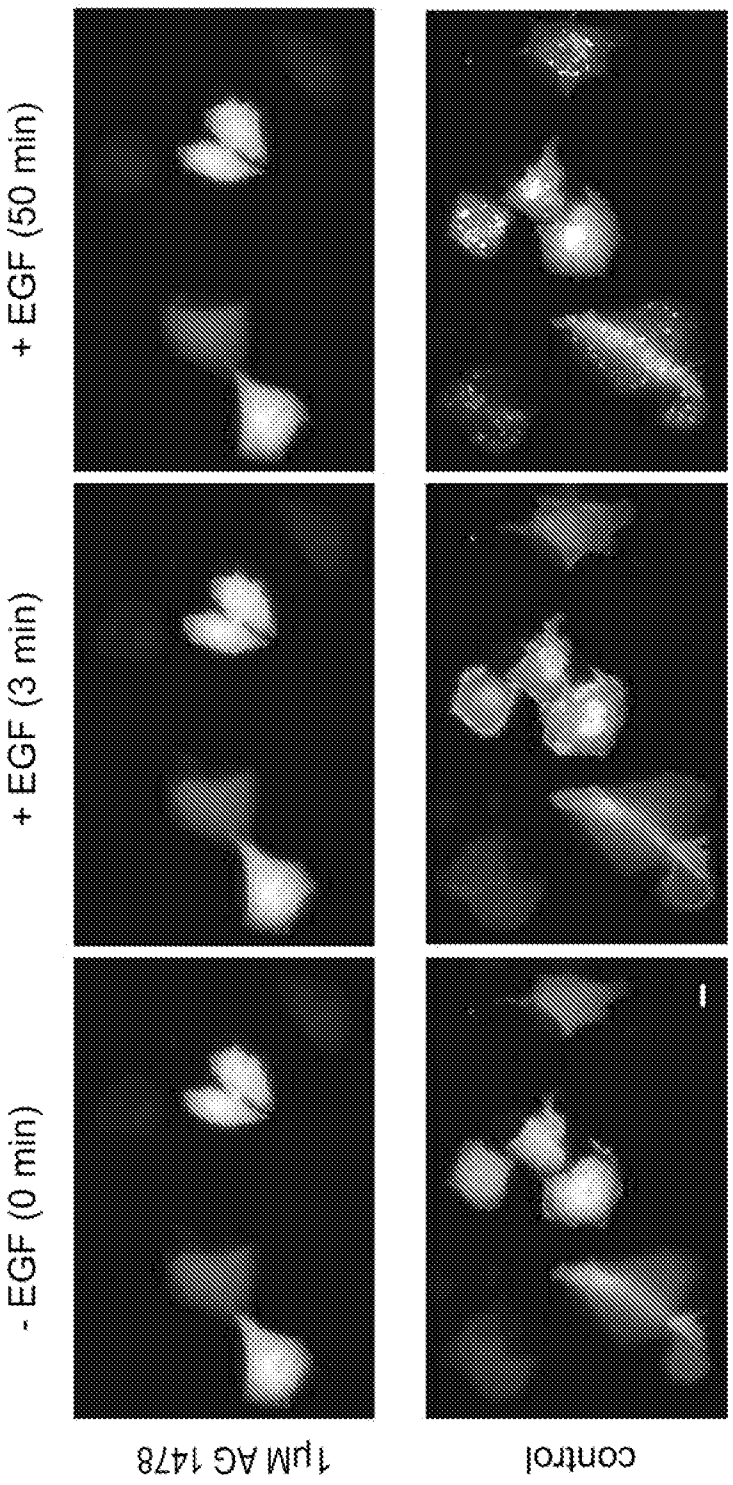

FIG. 7

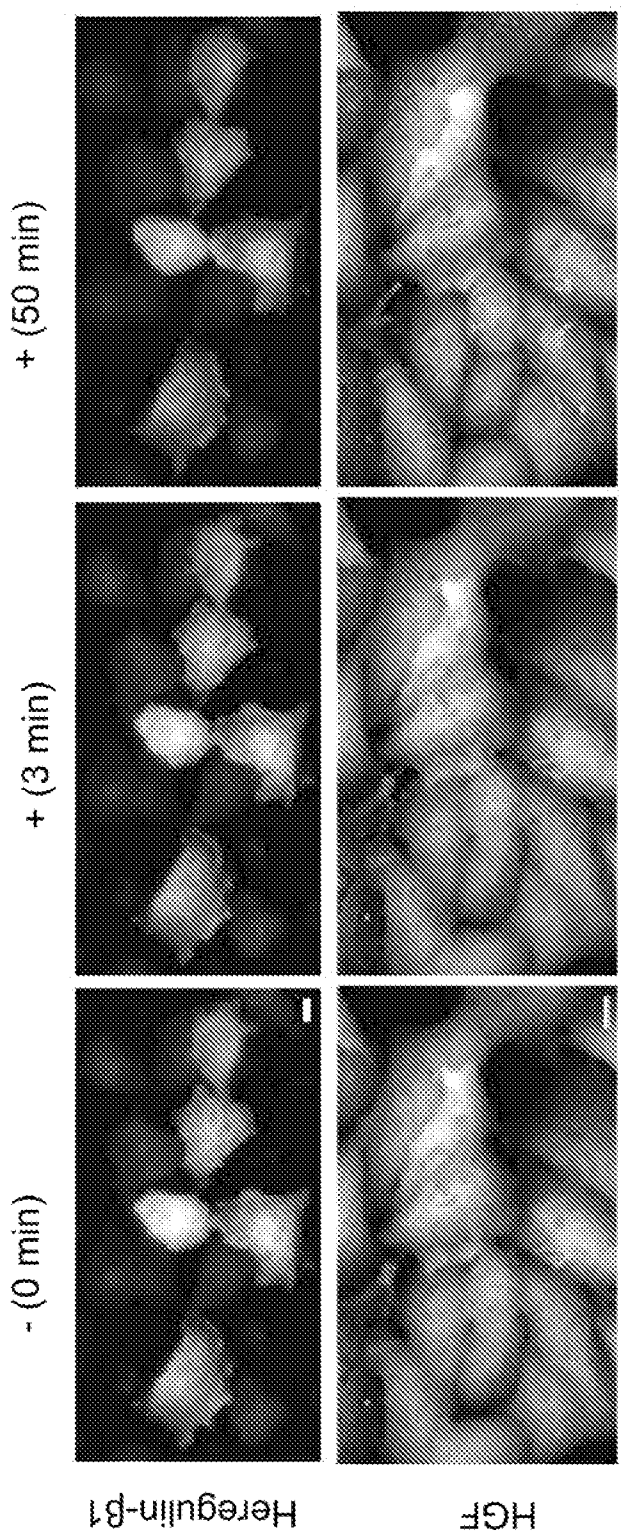

FIG. 8

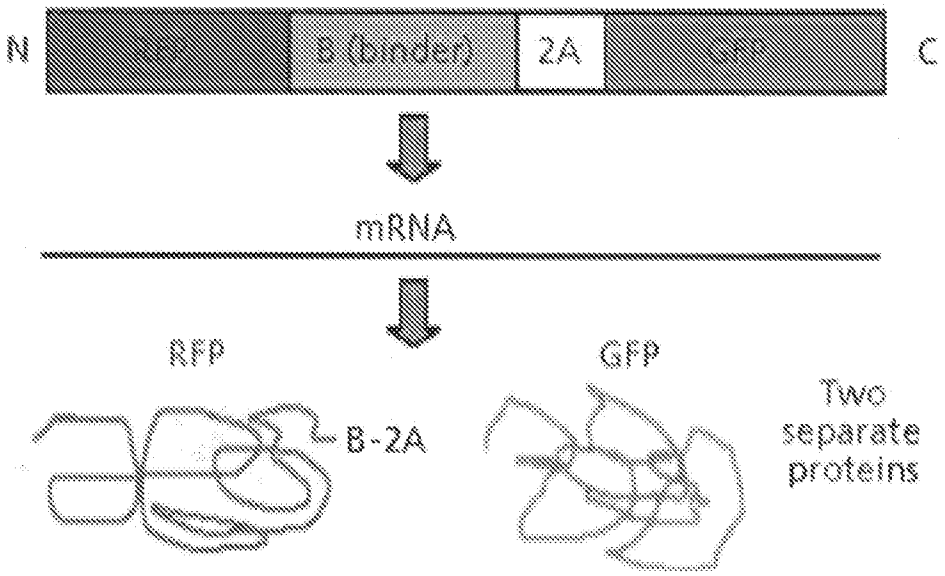
N
B (linker)
2A
C
mRNA
RFP
GFP
B-2A
Two separate proteins

FIG. 9A

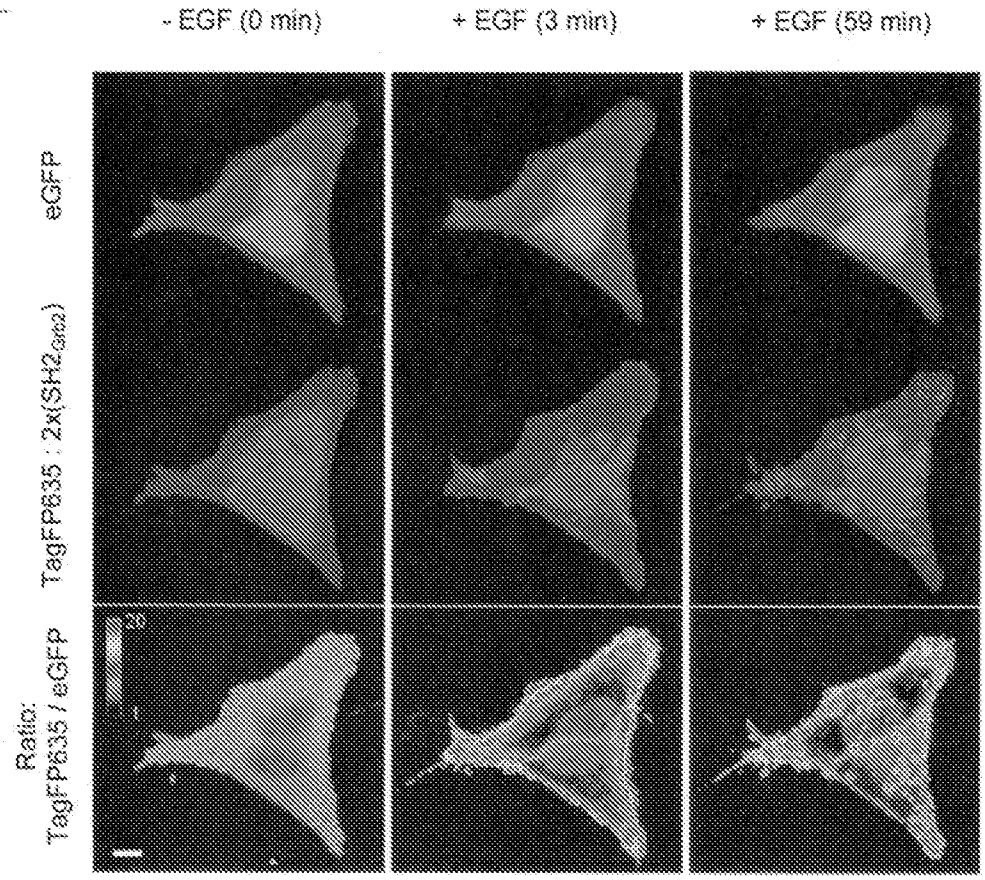
- EGF (0 min)
+ EGF (3 min)
+ EGF (59 min)
eGFP
TagFP635 : 2x(SH2$_{Grb2}$)
Ratio:
TagFP635 / eGFP

FIG. 9B

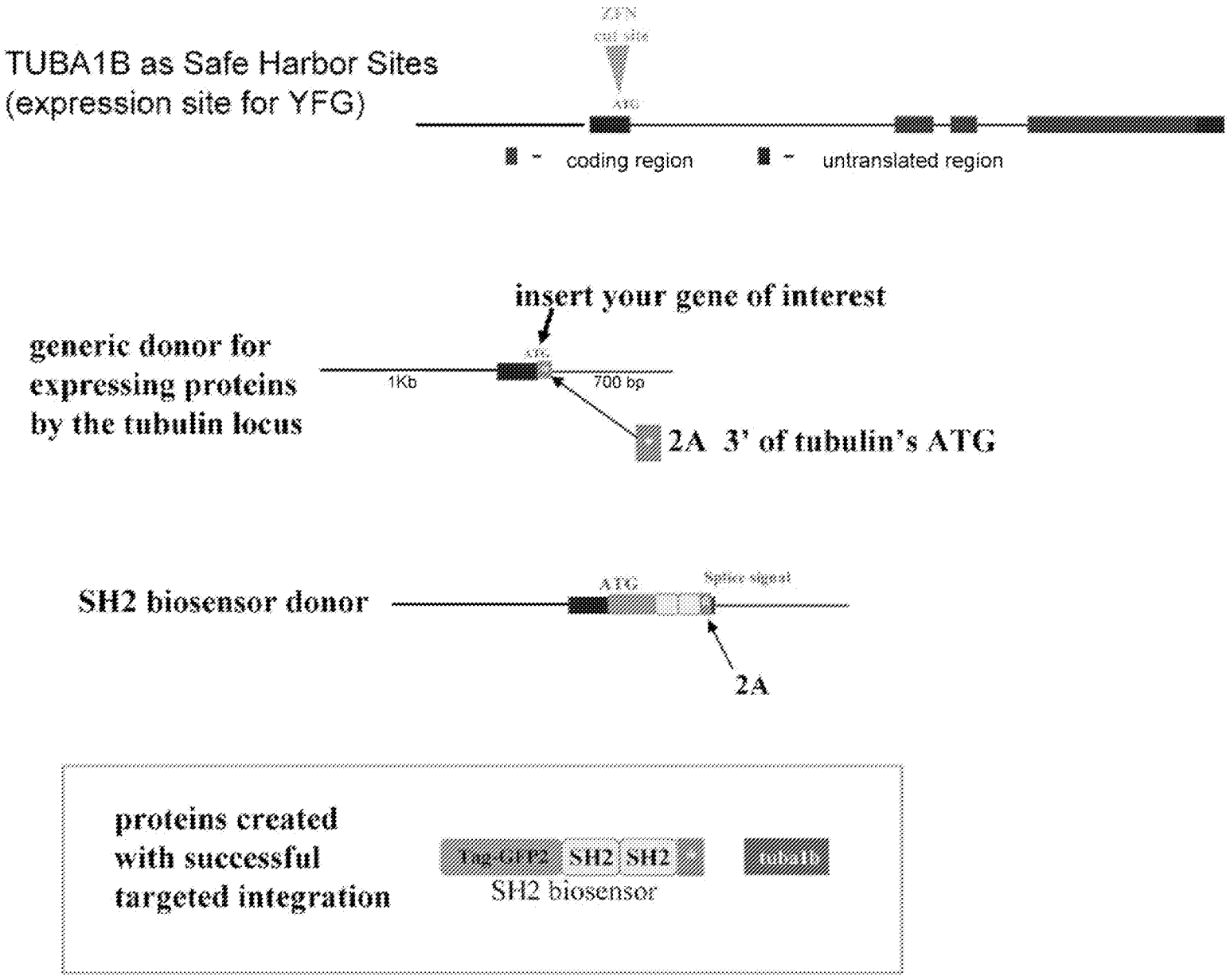
TUBA1B as Safe Harbor Sites
(expression site for YFG)
ZFN cut site
ATG
coding region
untranslated region
insert your gene of interest
generic donor for
expressing proteins
by the tubulin locus
ATG
1Kb
700 bp
2A 3' of tubulin's ATG
SH2 biosensor donor
ATG
Splice signal
2A
proteins created
with successful
targeted integration
Tag-GFP2
SH2
SH2
mKate
SH2 biosensor

FIG. 10

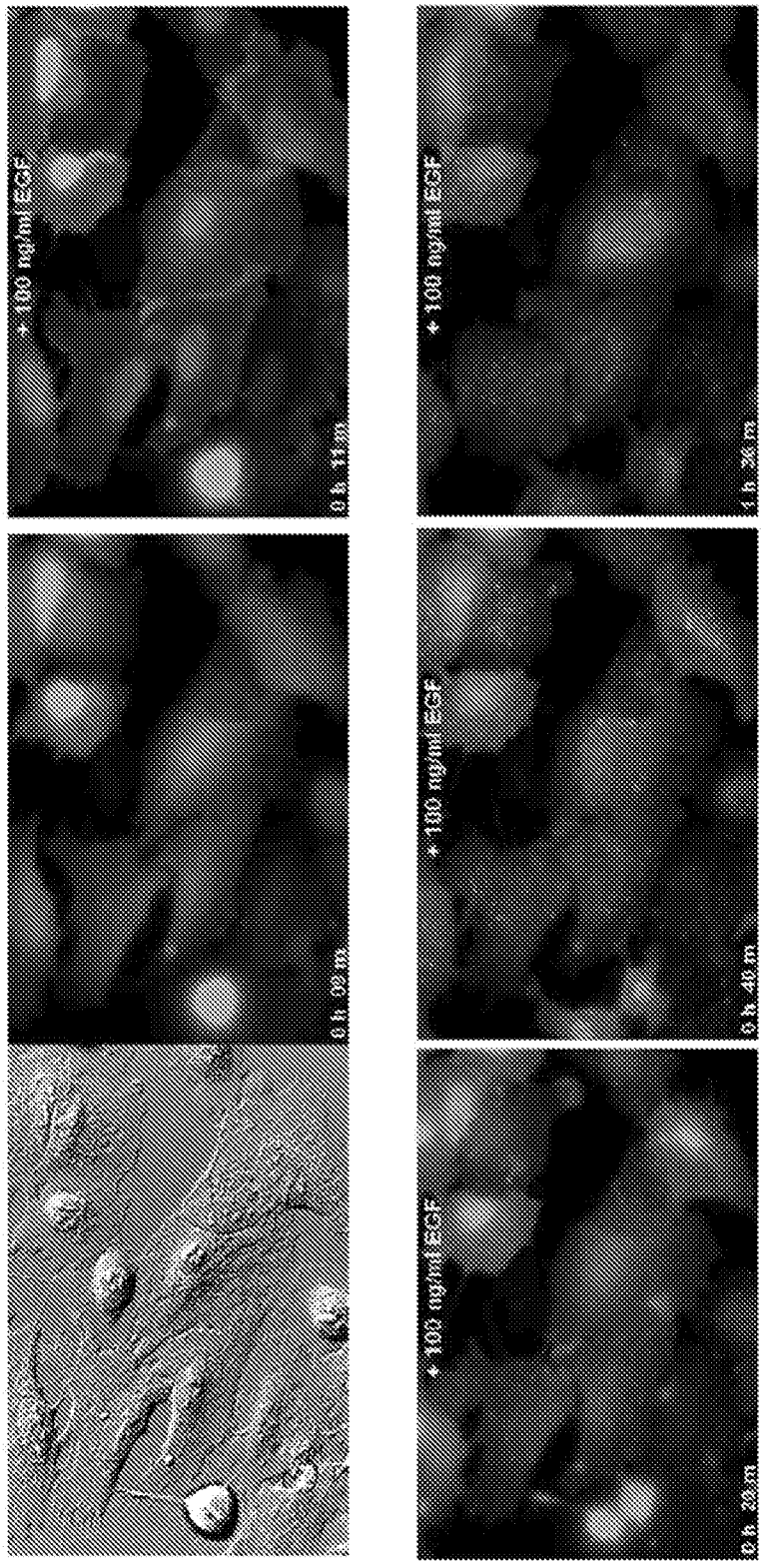
+ 100 ng/ml EGF
0 h 11 m
0 h 00 m
+ 100 ng/ml EGF
0 h 26 m
+ 100 ng/ml EGF
0 h 40 m
+ 100 ng/ml EGF
0 h 30 m

FIG. 11

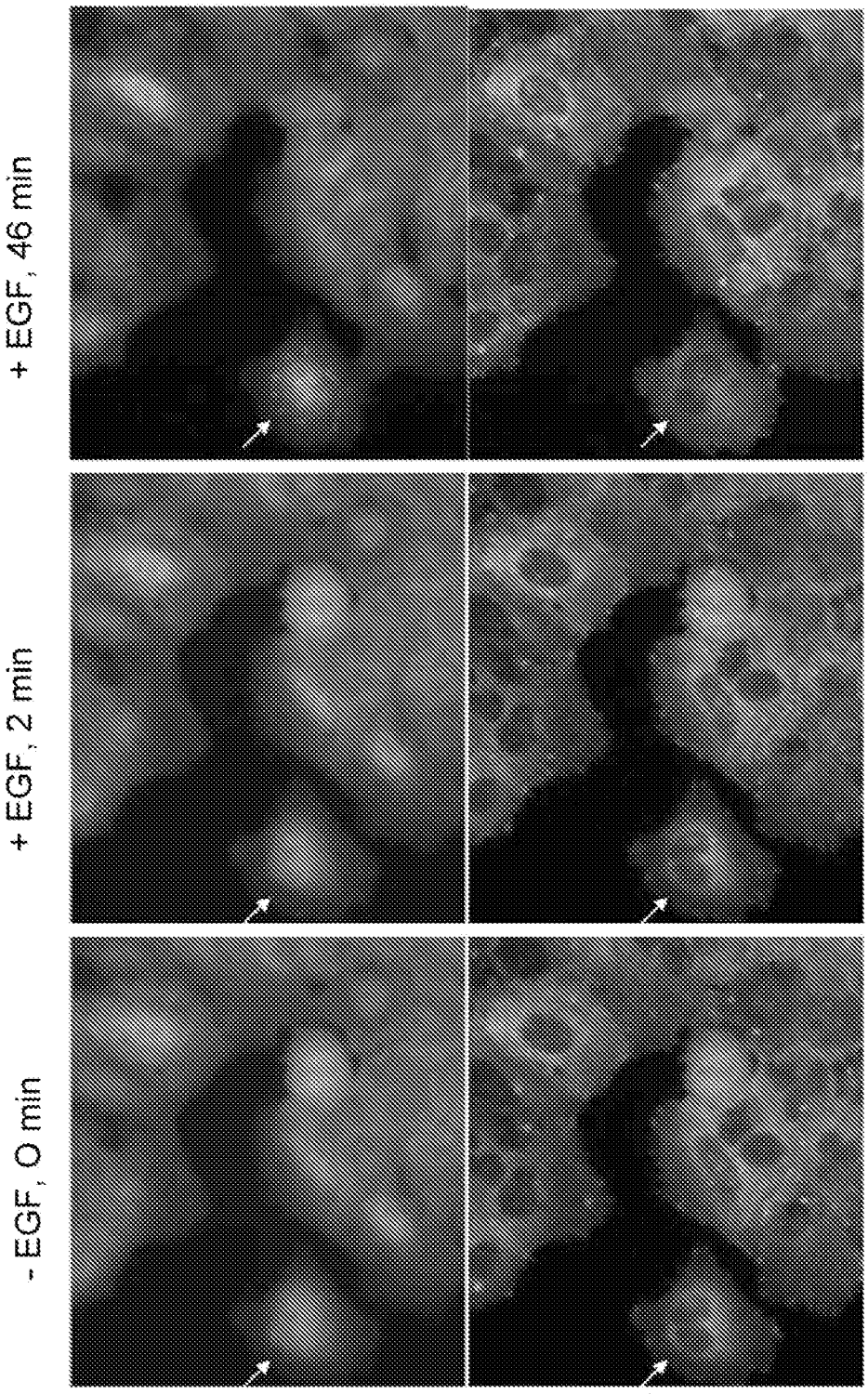
- EGF, O min
+ EGF, 2 min
+ EGF, 46 min
tagGFP2-2xSH2Grb2
tagRFP-β actin

FIG. 12

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 11/32214

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - C12N 15/00; F21V 9/16 (2011.01)
USPC - 435/320.1; 250/458.1
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
USPC: 435/320.1; 250/458.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PubWEST (USPT, PGPB, EPAB, JPAB), Google Patents/Scholar: GFP-Grb2, biosensors, EGFR, SH2, fluorescent protein, binding domain, lentiviral, Cre, loxP

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X -- Y	US 2006/0199226 A1 (Schiffer) 07 September 2006 (07.09.2006) para [0118], [0121]-[0123], [0126], Fig 3, 4	1-6, 9-12 ---------- 7-8
Y	US 2008/0200663 A1 (Yee et al.) 21 August 2008 (21.08.2008) para [0009]-[0010], [0014], Fig 1	7-8

☐ Further documents are listed in the continuation of Box C. ☐

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
13 July 2011 (13.07.2011)	20 JUL 2011
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer: Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (July 2009)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 11/32214

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
Group I+: claims 1-12, drawn to a cell comprising a biosensor, the biosensor being able to detect a change in activity of a signal transduction protein, the biosensor comprising at least one protein-binding domain and a first fluorescent protein, the cell expressing the biosensor at a level that is substantially similar to the level of the signal transduction protein. The first invention is restricted to EGFR and SH3. Should an additional fee(s) be paid, Applicant is invited to elect an additional specific signal transduction protein(s) and/or protein-binding domain(s) to be searched. The exact claims searched will depend on Applicant's election.
[NOTE: the first invention is restricted to EGFR and SH2 not to EGFR and SH3, as was indicated, because it is SH2, but not SH3, that binds to phosphorylated tyrosine of EGFR per para [0079] of the specification.]

Group II, claims 13-38, drawn to a method for detecting activation of a signal transduction protein in real time.

Group III, claims 39-58, drawn to a lentiviral particle, the lentiviral particle comprising a nucleic acid encoding a biosensor comprising at least one protein-binding domain and a first fluorescent protein.

- Please see extra sheet for continuation -

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-12 restricted to EGFR and SH2

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

********** Supplemental Box **********

Continuation of: Box NO III. Observations where unity of invention is lacking

The inventions listed as Groups I+ through III do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The inventions of Groups I+ and II do not include the inventive concept of a lentiviral particle, the lentiviral particle comprising a nucleic acid encoding a biosensor comprising at least one protein-binding domain and a first fluorescent protein, as required by Group III.

The inventions of Group I+ and II share the technical feature of a cell comprising a biosensor, the biosensor being able to detect a change in activity of a signal transduction protein, the biosensor comprising at least one protein-binding domain and a first fluorescent protein, the cell expressing the biosensor at a level that is substantially similar to the level of the signal transduction protein. However, this shared technical feature does not represent a contribution over prior art as being anticipated by a paper titled "Studies of Signal Transduction Events Using Chimeras To Green Fluorescent Protein" (Methods in Enzymology 2000, 327:500-513) by Meyer et al. (hereinafter "Meyer") that teaches a cell comprising a biosensor, the biosensor being able to detect a change in activity of a signal transduction protein (Title and pg 501, para 1; "monitoring the localization and translocation of signaling proteins in the intact cellular environment"), the biosensor comprising:

at least one protein-binding domain and a first fluorescent protein (pg 501, para 2, "GFP-tagged signaling proteins and GFP-tagged signaling domains can be used to understand signal transduction processes...GFP-tagged signaling proteins can be used to explore the translocation of signaling proteins and evaluate when GFP-tagged individual signaling domains are useful to understand specific signaling events such as phosphorylation"; also see pg 506, Fig 1 as an example, C1-GFP translocating to the PAF receptor, where C1 is the C1 domain of the y isoform of protein kinase C (PKC-y); pg 505, para 3);

the cell expressing the biosensor at a level that is substantially similar to the level of the signal transduction protein (pg 507, para 2, "docking partners of the GFP-tagged protein must be overexpressed in the same cell in order to resolve the translocation event. This enables the GFP-tagged protein to have a similar concentration as its binding partners"). As said cell was known at the time of the invention, this cannot be considered a special technical feature that would otherwise unify the groups.

Another special technical feature of the inventions listed as Group I+ is the specific signal transduction protein recited therein. As said signal transduction protein were known in the art at the time of the invention, the inventions do not share a special technical feature. Without a shared special technical feature, the inventions lack unity with one another.

Finally, another special technical feature of the inventions listed as Group I+ is the specific protein-binding domain. As said protein-binding domains were known in the art at the time of the invention, the inventions do not share a special technical feature. Without a shared special technical feature, the inventions lack unity with one another.

Groups I+ through III therefore lack unity under PCT Rule 13 because they do not share a same or corresponding special technical feature.