

(51) International Patent Classification:
C07D 401/06 (2006.01)(21) International Application Number:
PCT/US2015/010269(22) International Filing Date:
6 January 2015 (06.01.2015)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/923,937 6 January 2014 (06.01.2014) US

(71) Applicant: THE UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL [US/US]; The Office of Technology Development, Suite 430, 100 Europa Drive, Chapel Hill, North Carolina 27517 (US).

(72) Inventors: HAHN, Klaus M.; 1502 Michaux Road, Chapel Hill, North Carolina 27514 (US). MACNEVIN, Christopher J.; 1 Huntscroft Court, Durham, North Carolina 27713 (US).

(74) Agent: SCHWARTZMAN, Robert, A.; Myers Bigel Sibley & Sajovec, P.A., P.O. Box 37428, Raleigh, North Carolina 27627 (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, BN, 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, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, 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.

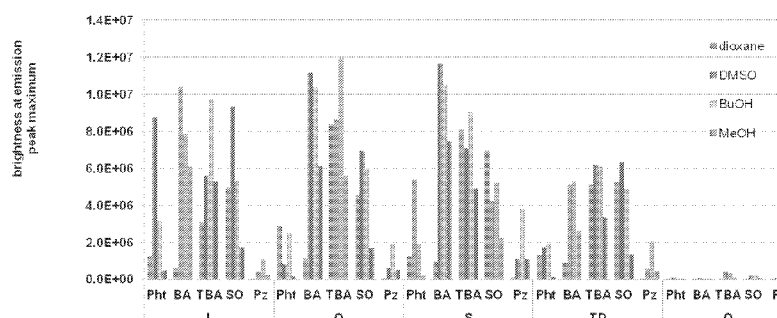
(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, 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, KM, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report (Rule 48.2(g))

(54) Title: IMPROVED ENVIRONMENT SENSING CYANINE AND MEROCYANINE DYES

FIG. 1



(57) Abstract: The present invention provides dyes having improved characteristics for imaging applications, biosensors comprising such dyes, and methods of use thereof, including methods for detecting target molecules in a sample and for live-cell imaging. The biosensors can include a binding member, including a biomolecule or fragments thereof, which can interact with target molecules of interest and can be specific to a given conformational state or covalent modification of the target molecule. The presently disclosed dyes can be used for detecting changes in the binding, conformational change, or posttranslational modification of the target molecule.

); Nalb

IMPROVED ENVIRONMENT SENSING CYANINE AND MEROCYANINE DYES

STATEMENT OF PRIORITY

[0001] This application claims the benefit of U.S. Provisional Application Serial No. 61/923,937, filed January 6, 2014, the entire contents of which are incorporated by reference herein.

STATEMENT OF FEDERAL SUPPORT

[0002] This invention was made with government support under Grant No. GM090317 awarded by the National Institutes of Health. The government has certain rights in this invention.

FIELD OF THE INVENTION

[0003] The present invention relates to fluorescent dyes and biosensors comprising said fluorescent dyes and methods of their use for detecting a target molecule and an activity or location of a target molecule within a living cell.

BACKGROUND OF THE INVENTION

[0004] Visualizing protein activation dynamics *in vivo* has proven essential to understanding how signaling networks are transiently configured to generate different cell phenotypes (Machacek *et al.*, *Nature* 461:99-103 (2009); Cai *et al.*, *X.*, *Mol. Cell. Biol.* 28:201-214 (2008); Gardiner *et al.*, *Curr. Biol.* 12:2029-2034 (2002)). Fluorescent biosensors have become a valuable tool for quantifying the spatio-temporal dynamics of cellular signaling (Hodgson *et al.*, *Meth. Cell Biol.* 85:63-81 (2008); Lawrence *et al.*, *ChemBioChem* 8:373-378 (2007); Zhang *et al.*, *Mol. BioSyst.* 3:759-765 (2007); Chen *et al.*, *C.*, *Biochim. Biophys. Acta, Proteins Proteomics* 1697:39-51 (2004)). Biosensors based on environment sensing fluorescent dyes can provide important advantages: 1) they can be brighter than fluorescent proteins, 2) unlike biosensors based on fluorescence resonance energy transfer (FRET), their fluorescence is the product of direct rather than indirect excitation, and 3) in some biosensor designs they report the conformational changes of endogenous, unmodified proteins (Hahn *et al.*, *K.*, *Curr. Opin. Cell Biol.* 14:167-172 (2002); Touthkine *et al.*, *J. Am. Chem. Soc.* 125:4132-4145 (2003); vant *et al.*, *Science* 305:1615-1619

(2004); Gulyani *et al.*, *Nat. Chem. Biol.* 7:437-444 (2011)). Nonetheless, currently available environment-sensing dyes for use in live cell imaging applications often suffer from one or more limitations, including less than adequate changes in fluorescence response on target binding, insufficient brightness, poor photostability, and/or short fluorescence wavelengths that are toxic to cells and overlap with autofluorescence. Elegant studies have resulted in several fluorophores that are sensitive to their environment and have been used to report protein activity *in vitro* (Goguen *et al.*, *Bioorg. Med. Chem. Lett.* 21:5058-5061 (2011); Loving *et al.*, *Trends Biotechnol.* 28:73-83 (2010)). For applications *in vivo*, a dye must be capable of providing sufficient signal/noise to report protein activities at biosensor concentrations that minimally perturb normal physiology. This is a product of both the fluorescence change in response to protein activity, and the brightness of the dye. Dyes used *in vitro* sometimes undergo impressive changes in fluorescence intensity, but are too dim to be used in living systems. Further, the membrane permeability of dyes is often insufficient for use in cells.

[0005] Cyanine and merocyanine dyes are characterized by electron donor and acceptor components linked by conjugation, usually a system of double bonds. This results in a ground state that may be represented as a resonance hybrid of charged and uncharged forms (Shirinian *et al.*, *Top. Heterocycl. Chem.* 14:75-105 (2008); Kulinich *et al.*, *Russ. Chem. Rev.* 78:141-164 (2009)). The potential for resonance delocalization across the polyene system renders these dyes especially sensitive to the effects of hydrogen bonding from the surrounding solvent (Buncel *et al.*, *Acc. Chem. Res.* 23:226-231 (1990)). Changes in the local solvation environment can result in large changes in fluorescence intensity, as well as shifts in excitation and emission maxima. These changes can be used in quantitative live cell imaging applications to measure the activity of a target protein, usually by attaching the dyes to an 'affinity reagent' that binds only to the activated state of the targeted protein. When the affinity reagent binds the activated, endogenous target, the dye is placed in a different solvent environment, leading to fluorescence changes (Nalbant *et al.*, *Science* 305:1615-1619 (2004); Garrett *et al.*, *Biochemistry* 47:986-996 (2008)). Dyes can also be directly attached to the protein of interest to report conformational changes.

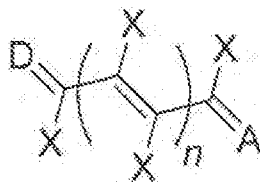
[0006] Pioneering work by Brooker demonstrated that absorbance properties of the merocyanines are strongly affected by varying the terminal donor and acceptor groups (Brooker *et al.*, *J. Am. Chem. Soc.* 67:1869-1874 (1945); Brooker *et al.*, *J. Am. Chem. Soc.* 67:1875-1889 (1945); Brooker *et al.*, *J. Am. Chem. Soc.* 73:5332-5350 (1951); Brooker *et al.*, *J. Am. Chem. Soc.* 73:5326-5332 (1951)), and later studies confirmed the profound effect on fluorescence properties that arises from the use of different donor and acceptor components (Marder *et al.*, *Science* 276:1233-1236 (1997); Bublitz *et al.*, *J. Am. Chem. Soc.* 119:2311-2312 (1997); Bublitz *et al.*, *J. Am. Chem. Soc.* 119:3365-3376 (1997); Bublitz *et al.*, *Annu. Rev. Phys. Chem.* 48:213-242 (1997); Marder *et al.*, *J. Am. Chem. Soc.* 115:3006-3007 (1993)). Nonetheless, a relatively small number of specific donor-acceptor combinations have been synthesized and fully characterized with respect to the photophysical properties relevant to live cell imaging.

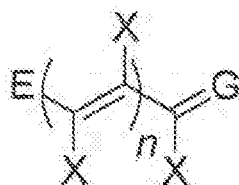
[0007] The presently disclosed subject matter addresses fluorophores with the requisite combination of brightness and fluorescence responsiveness for use in living cells, as well as increased membrane permeability.

SUMMARY

[0008] The present invention provides cyanine and merocyanine dyes having improved properties for use in imaging, including useful combinations of brightness, photostability, and solvent-dependent fluorescence. The invention further provides cyanine and merocyanine dyes having increased cell membrane permeability for enhanced effectiveness in living cells. These dyes retain brightness and solvent sensitivity while readily crossing cell membranes and showing minimal staining of intracellular membranes.

[0009] Thus, one aspect of the invention relates to a compound of Formula I or Formula II





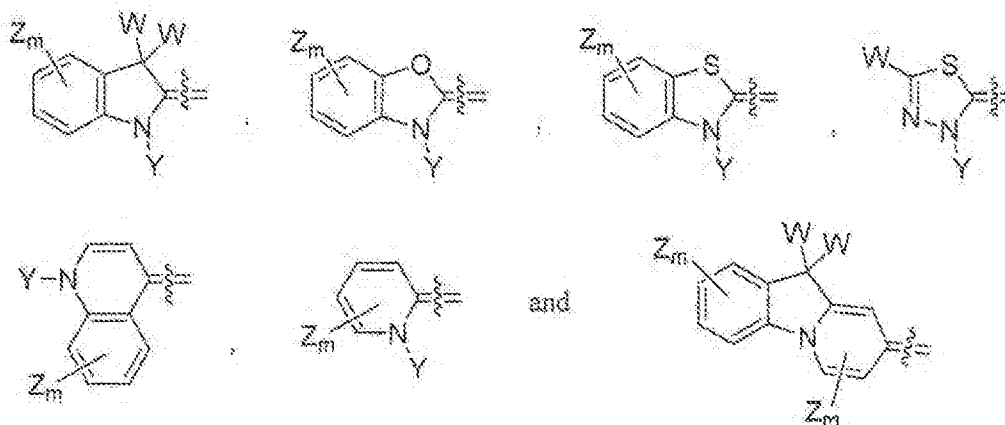
II

wherein:

n is an integer from 0 to 2;

each X is independently hydrogen or an electron withdrawing group selected from the group consisting of carbonyl, cyano, halogen, nitro, sulfonyl, and trifluoromethyl;

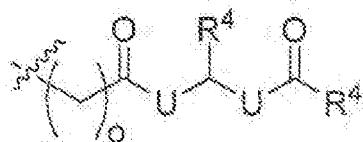
D is selected from the group consisting of:



W is optionally substituted C_{1-6} alkyl;

m is an integer from 0 to 4;

Y and Z are each independently selected from the group consisting of hydrogen, C_{1-6} alkyl, a conjugatable side chain, and

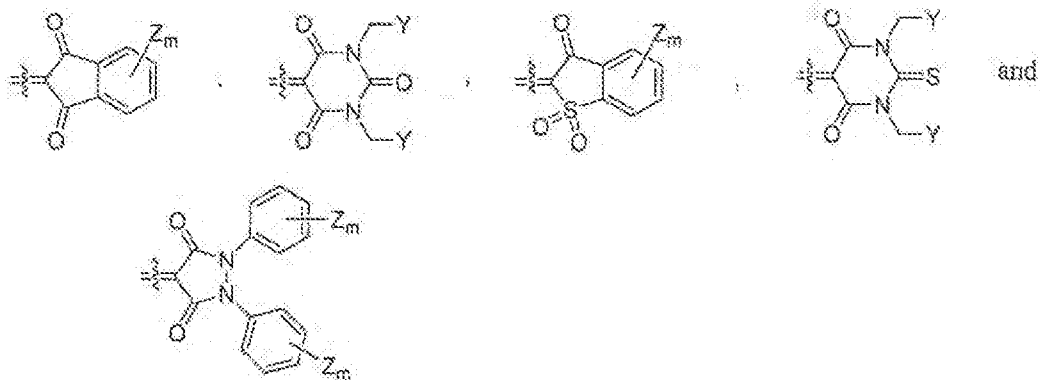


o is an integer from 0 to 5;

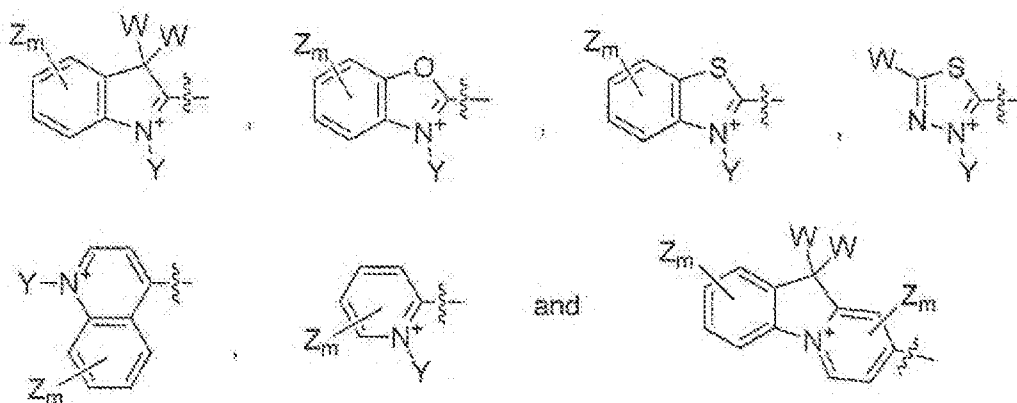
U is O or S;

R^4 is hydrogen or optionally substituted C_{1-6} alkyl;

A is selected from the group consisting of:

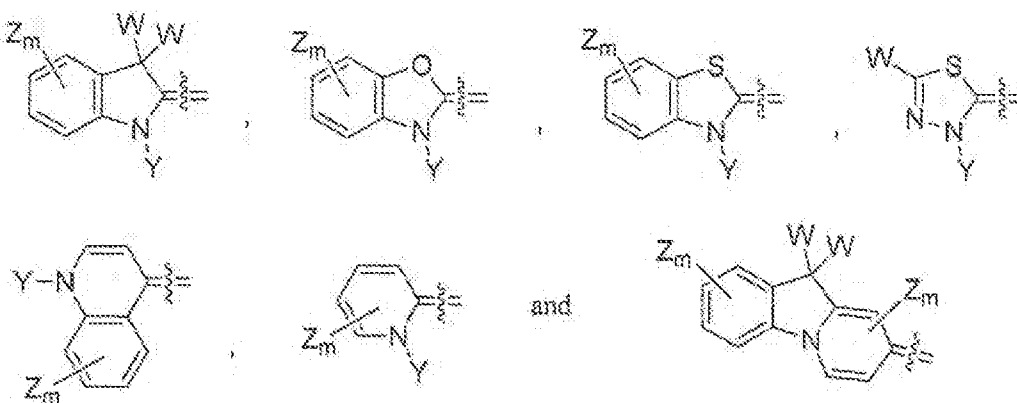


E is selected from the group consisting of:



and

G is selected from the group consisting of:



[0010] Another aspect of the invention relates to a biosensor comprising a compound of Formula I or Formula II and a binding member having a specific affinity for the target molecule, wherein the biosensor exhibits a detectable change in a fluorescence property during or after binding to the target molecule.

[0011] A further aspect of the invention relates to a method for determining the presence or amount of one or more target molecules in a sample, the method comprising: (a) providing a biosensor of the invention; (b) contacting the biosensor with a sample suspected of containing one or more target molecules to bind the one or more target molecules, if present, to the binding member; (c) irradiating the sample suspected of containing one or more target molecules with electromagnetic radiation to induce the compound of Formula I or Formula II to fluoresce; and (d) detecting a fluorescence property of the compound of Formula I or Formula II to determine the presence or amount of one or more target molecules in the sample.

[0012] An additional aspect of the invention relates to a method of detecting an activity or a location of one or more target molecule within a cell, the method comprising: (a) providing a biosensor of the invention; (b) contacting the biosensor with a cell suspected of containing one or more target molecules to bind the one or more target molecules, if present, to the binding member; (c) irradiating the cell suspected of containing one or more target molecules with electromagnetic radiation to induce the compound of Formula I or Formula II to fluoresce; and (d) detecting one or more of: (i) a fluorescence property of the compound of Formula I or Formula II; (ii) a change in a fluorescence property of the compound of Formula I or Formula II; (iii) a location of a fluorescence of the compound of Formula I or Formula II; and (iv) combinations thereof, to determine the activity or location of one or more target molecules in the cell.

[0013] Another aspect of the invention relates to a method of detecting an interaction between an endogenous target molecule and a cellular entity, the method comprising: (a) providing a cell comprising an endogenous target molecule; (b) providing a biosensor of the invention; (c) observing a background fluorescence signal from the biosensor; (d) contacting the biosensor with the cell; and (e) detecting a change in fluorescence from the biosensor to indicate an interaction between the target molecule and the cellular entity.

[0014] A further aspect of the invention relates to a kit comprising a biosensor for detecting, monitoring or observing a target molecule, wherein the biosensor comprises a compound of Formula I or Formula II and a binding member having a specific affinity for the target molecule.

[0015] Certain embodiments of the presently disclosed subject matter having been stated hereinabove, which are addressed in whole or in part by the presently disclosed subject matter, other embodiments will become evident as the description proceeds when taken in connection with the accompanying Examples and Figures as best described herein below.

BRIEF DESCRIPTION OF THE DRAWINGS

[0016] Having thus described the presently disclosed subject matter in general terms, reference will now be made to the accompanying Drawings, which are not necessarily drawn to scale.

[0017] Fig. 1 shows brightness values at the emission peak maximum when exciting the dye at its excitation maximum. Values are grouped by donor heterocycle and are presented for the set of 4 solvents (dioxane, DMSO, BuOH, MeOH) in order of increasing hydrogen bonding capacity.

[0018] Figs. 2A-2B show the change in excitation maximum (A) and emission maximum (B) from smallest value for set of four solvents.

[0019] Figs. 3A-3B show the change in the λ max of excitation (A) and emission (B) from non-polar (1,4-dioxane) to polar (methanol) solvent.

[0020] Fig. 4 shows the photostability of library dyes. Photostability values were calculated as: (photobleaching rate of fluorescein) / (photobleaching rate of dye).

[0021] Figs. 5A-5B show the *in vitro* binding assay of dye labeled CBD to Cdc42. (A) Fold change increase of emission intensity upon binding of labeled CBD to Cdc42. (B) Maximum fluorescence intensity at saturated binding of Cdc42.

Mero221 was used in previously published successful biosensor applications.

[0022] Fig. 6 shows the complete emission spectra for binding of merocyanine-labeled CBD-MBP fragment at low (0.01 μ M) and high (3.0 μ M) concentrations of Cdc42.

[0023] Fig. 7 shows the dissociation constant (K_d) values calculated for microCBD binding with Cdc42 in 50 mM NaH_2PO_4 , 150 mM NaCl buffer at pH 7.5.

[0024] Figs. 8A-8C show Cdc42 biosensors based on the new dyes show the localization of Cdc42 activation in moving MEF cells. Cells were injected with CBD biosensor labeled with **mero61** (A), **mero62** (B), or **mero87** (C). Biosensors show increased Cdc42 activation in areas where the cell edge is protruding, as indicated by white arrows. Some activity is also seen at perinuclear compartments and in the nucleus of some cells, consistent with previous observations (Nalbant *et al.*, *Science* 305:1615-1619 (2004)).

[0025] Fig. 9 shows the fluorescence emission intensity of the four conjugatable merocyanines over time at different pH. Each dye was prepared as a 1.0 μM solution in aqueous sodium phosphate/citric acid buffer.

[0026] Fig. 10 shows merocyanine dyes that consist of electron donor and acceptor components that are linked by an extended polyene system. The **R** groups indicate the various positions where acetoxymethyl ester derivatives were integrated to screen for compounds with cell membrane permeability and minimal membrane staining.

[0027] Fig. 11 shows fluorescence intensity of each dye in the initial screening set. MEF cells were incubated for 30 min at 37 °C in serum-free media containing 10 μM of a given test dye and 1 μM Hoechst 33342. Pluronic F-127 was added to the initial DMSO stock solutions of dyes to give a final concentration in the incubation solutions of < 0.01%. Fluorescence intensity (FI) values were recorded at the peak of the emission curve for each dye when excited at the excitation maximum wavelength. Values were normalized based on relative intensity of Hoechst (exc. 351 nm, em. 440 nm).

[0028] Fig. 12 shows the plot of fluorescence intensity for dye 4 (**I(AM)-BA(AM)**) vs. dye 5 (**I(2AM)-BA**) at various incubation concentrations. Cells were incubated in medium containing the test dye at 37 °C for 30 minutes. The dye solution was aspirated and the cells were washed with medium, trypsinized, and suspended in fresh medium. The fluorescence intensity (FI) of the cell suspension was obtained for each dye. Values were normalized for cell number based on the relative intensity of the nuclear stain Hoechst 33342, which was included with the dye in each test solution.

[0029] Fig. 13 shows the distribution of dyes in MEF cells. Cells were scaled individually to show subcellular distribution, so images do not indicate relative uptake of each dye. Dye 2 showed staining of vesicles, perinuclear membrane and nuclear exclusion. Some diffuse staining of membrane away from the nucleus was also seen. Dye 5 showed uniform subcellular distribution and no organelle specific interactions. Variations in intensity were a function of cell volume, as indicated by the volume marker YFP that was stably expressed in the same cell.

[0030] Fig. 14 shows the emission spectra for **mero76** and **mero166**. Samples diluted to 1.0 μ M in respective solvent (DMSO, MeOH, or water/0.1% DMSO). Excitation wavelength was at sample excitation maximum minus 30 nm.

[0031] Fig. 15 shows mouse embryonic fibroblasts treated with **mero166** or **G9a-mero 166** conjugate.

[0032] Figs. 16A-16B show the design and synthesis of the small molecule based biosensor CaMero. (A) Overview of the small molecule based biosensor approach. Activation of the protein of interest (POI) allows binding of the sensor, which results in measurable changes in fluorescence properties of the probe. (B) Crystal structure of trifluoperazine bound to calmodulin indicates solvent exposed *N*-methyl piperazine not involved in binding interactions (PDB: 1LIN). (C) Synthetic route to trifluoperazine analogue (2) and negative control compound (3) each with reactive tether. Conjugation to the environment-sensing, membrane permeant merocyanine dye **mero76** gives the calmodulin small molecule biosensor CaMero and calmodulin small molecule biosensor negative control compound CaMero-NC.

[0033] Fig. 17 shows inactive versus active conformations of calmodulin. Upon binding calcium, calmodulin undergoes a shift from a more globular conformation to an open, dumbbell shaped conformation (PDB: 1LIN). Only once activated will calmodulin bind with the small molecule trifluoperazine (TFP).

[0034] Figs. 18A-18B show that fluorescence changes of CaMero sensor require the presence of calcium and calmodulin. (A) The CaMero sensor shows a change in fluorescence intensity (FI) with increasing calcium concentration only in the presence of calmodulin. No change in FI is observed with the CaMero-NC compound. (B) Changes in CaMero probe FI are reversible. (C) Changes in CaMero FI with increasing concentration of calmodulin are observed only in the presence of calcium.

[0035] Figs. 19A-19C show that CaMero sensor shows fluorescence response on stimulation of intracellular calcium release. (A) Astroglial cells incubated with the dye alone or CaMero-NC probe both show uniform intracellular distribution. Cells incubated with CaMero show perinuclear localization similar to that of the endoplasmic reticulum marker ER-Tracker. (B) Ratio images of astroglial cells pre- and post-stimulation show increase in fluorescence intensity (FI) in CaMero treated cells only. (C) Quantification of FI changes shows modest change at the whole cell level due to large changes in cellular morphology. Larger changes in FI are observed within the perinuclear region ($n = 3$).

[0036] Fig. 20 shows primary murine neurons incubated with CaMero show increase in fluorescence intensity following stimulation. Cells were stimulated with KCl (75 mM). Plot is of average fluorescence intensity in the cell body region vs. time for CaMero and CaMero-NC ($n = 9$).

[0037] Figs. 21A-21C show that CaMero sensor shows calmodulin activation at the cell edge in normally migrating cells. (A) Ratio images of migrating MEF cells show localized intensity band at the leading edge only in CaMero treated cells. Far right panel shows zoom of edge section (scale bar for whole cell images = 5 μM , zoom = 10 μM). (B) Quantification of fluorescence intensity ratio relative to distance from the leading edge. (C) Plot of fluorescence intensity ratio and calculated cell edge velocity.

[0038] Fig. 22 shows a line scan analysis of cell edge intensity using carboxyfluorescein diacetate volume indicator. Edge intensity gradient apparent with CaMero with and without co-treatment with clozapine. No intensity gradient is seen with the CaMero-NC probe.

[0039] Figs. 23A-23C show whole animal treatment with CaMero. (A) Excised intestinal tissue shows intensity of CaMero probe relative to control. (B) Co-localization of GFP-LifeAct actin marker observed with CaMero but not with CaMero-NC probe in smooth muscle cells. (C) Plot of fluorescence intensity over time for CaMero (dark line) and CaMero-NC (light line). The CaMero probe shows periodic changes in intensity consistent with the frequency of tissue contraction.

DETAILED DESCRIPTION

[0040] The presently disclosed subject matter now will be described more fully hereinafter with reference to the accompanying Figures, in which some, but not all embodiments of the presently disclosed subject matter are shown. Many modifications and other embodiments of the presently disclosed subject matter set forth herein will come to mind to one skilled in the art to which the presently disclosed subject matter pertains having the benefit of the teachings presented in the foregoing descriptions and the associated Figures. Therefore, it is to be understood that the presently disclosed subject matter is not to be limited to the specific embodiments disclosed and that modifications and other embodiments are intended to be included within the scope of the appended claims. Although specific terms are employed herein, they are used in a generic and descriptive sense only and not for purposes of limitation.

[0041] Unless the context indicates otherwise, it is specifically intended that the various features of the invention described herein can be used in any combination. Moreover, the present invention also contemplates that in some embodiments of the invention, any feature or combination of features set forth herein can be excluded or omitted. To illustrate, if the specification states that a complex comprises components A, B and C, it is specifically intended that any of A, B or C, or a combination thereof, can be omitted and disclaimed singularly or in any combination.

[0042] Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. The terminology used in the description of the invention herein is for the purpose of describing particular embodiments only and is not intended to be limiting of the invention.

[0043] All publications, patent applications, patents, patent publications and other references cited herein are incorporated by reference in their entireties for the teachings relevant to the sentence and/or paragraph in which the reference is presented.

[0044] The terms "a," "an," and "the" refer to "one or more" when used in this application, including the claims. Thus, for example, reference to "a sample" includes

a plurality of samples, unless the context clearly is to the contrary (e.g., a plurality of samples), and so forth.

[0045] Also as used herein, “and/or” refers to and encompasses any and all possible combinations of one or more of the associated listed items, as well as the lack of combinations when interpreted in the alternative (“or”).

[0046] Throughout this specification and the claims, the words “comprise,” “comprises,” and “comprising” are used in a non-exclusive sense, except where the context requires otherwise.

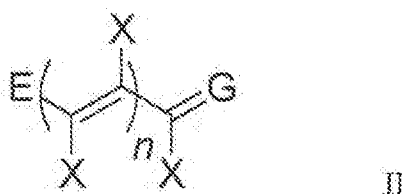
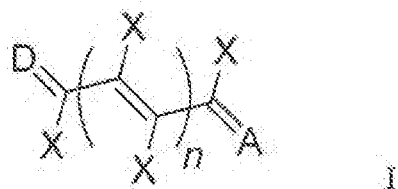
[0047] The transitional phrase “consisting essentially of” means that the scope of a claim is to be interpreted to encompass the specified materials or steps recited in the claim, “and those that do not materially affect the basic and novel characteristic(s)” of the claimed invention. *See, In re Herz*, 537 F.2d 549, 551-52, 190 USPQ 461, 463 (CCPA 1976) (emphasis in the original); *see also* MPEP § 2111.03.

[0048] As used herein, the term “about,” when referring to a value is meant to encompass variations of, in some embodiments $\pm 20\%$, in some embodiments $\pm 10\%$, in some embodiments $\pm 5\%$, in some embodiments $\pm 1\%$, in some embodiments $\pm 0.5\%$, and in some embodiments $\pm 0.1\%$ from the specified amount, as such variations are appropriate to perform the disclosed methods or employ the disclosed compositions.

I. Compounds

[0049] Cyanine and merocyanine dyes can be used as fluorescent labels for biomolecules, such as proteins or antibodies, for histochemical studies or immunobased assays. Such dyes can be attractive for use as fluorophores due, in part, to their wide range of available wavelengths, large molar extinction coefficients, and high quantum yields. Further, because of the sensitivity of the fluorescence intensity and wavelength of merocyanine dyes to solvent polarity and other characteristics of the dye environment, the presently disclosed merocyanine dyes can be used as sensors, e.g., a biosensor, of protein activation, protein-protein binding, and protein conformational changes.

[0050] One aspect of the invention relates to a compound of Formula I or Formula II:

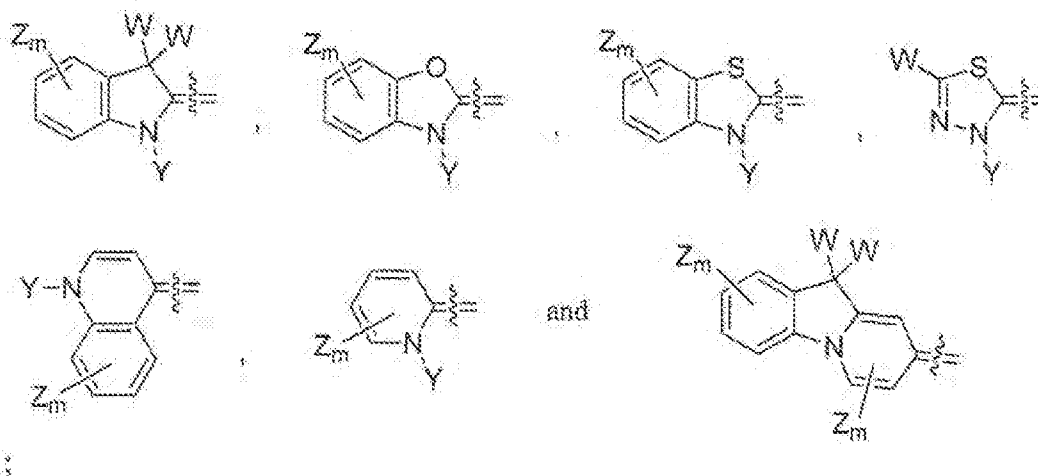


wherein:

n is an integer from 0 to 2;

each X is independently hydrogen or an electron withdrawing group selected from the group consisting of carbonyl, cyano, halogen, nitro, sulfonyl, and trifluoromethyl;

D is selected from the group consisting of:

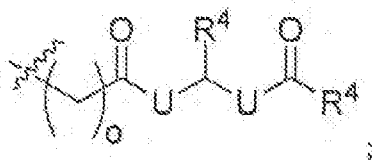


W is optionally substituted C₁₋₆ alkyl;

m is an integer from 0 to 4;

Y and Z are each independently selected from the group consisting of hydrogen,

C₁₋₆ alkyl, a conjugatable side chain, and

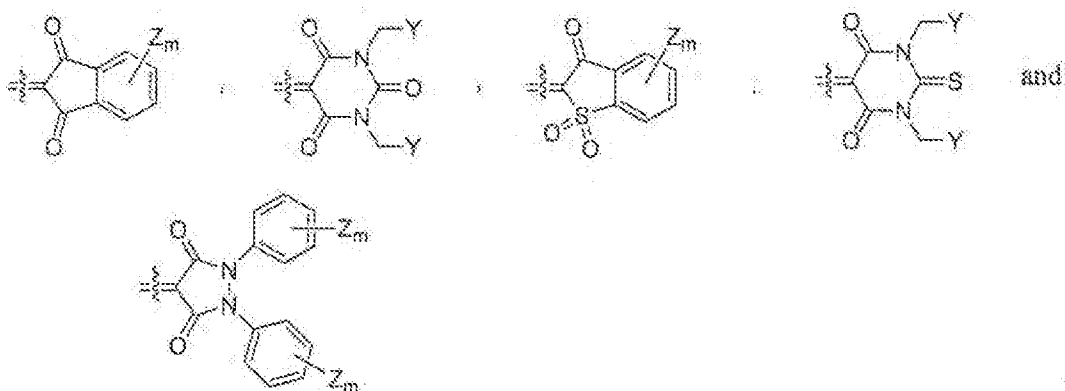


o is an integer from 0 to 5;

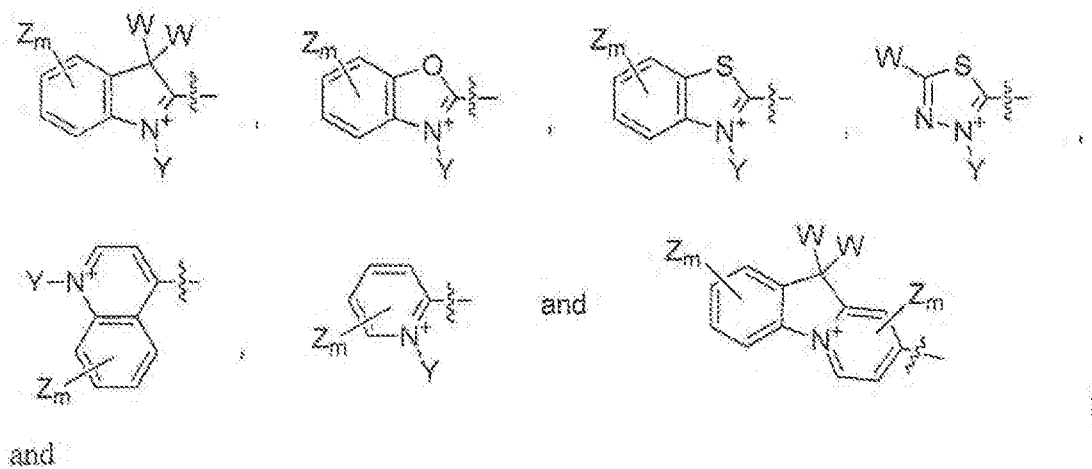
U is O or S;

R^4 is hydrogen or optionally substituted C_{1-6} alkyl;

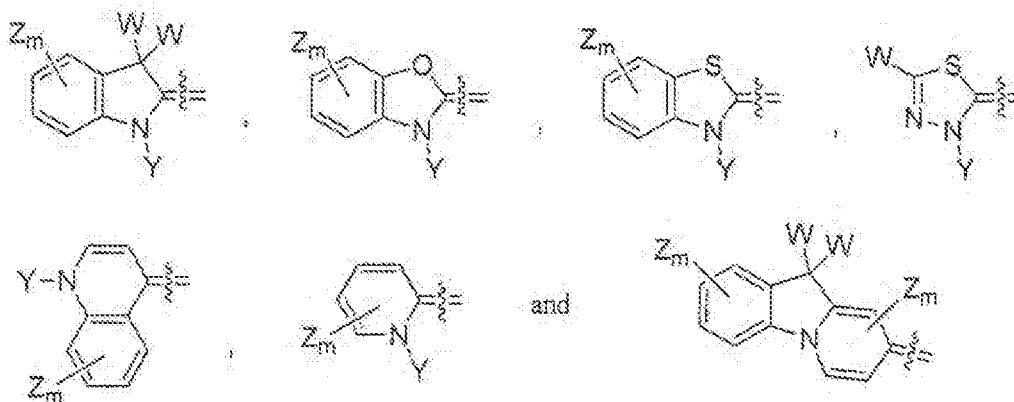
A is selected from the group consisting of:



E is selected from the group consisting of:



G is selected from the group consisting of:



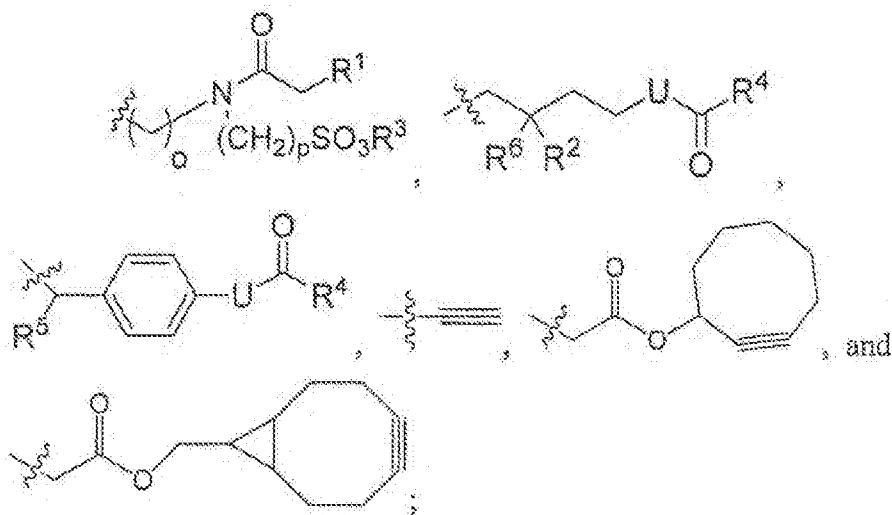
[0051] In specific embodiments, the invention encompasses compounds having any one of the described D moieties coupled with any one of the A moieties. The invention further encompasses compounds having any one of the E moieties coupled with any one of the G moieties.

[0052] In some embodiments, the compound does not contain an electron withdrawing group on the polymethine chain connecting A and D or E and G, *e.g.*, each X is a hydrogen. In other embodiments, the compound contains a single electron withdrawing group, *e.g.*, one X is an electron withdrawing group. In certain embodiments, the electron withdrawing group is attached to the polymethine chain carbon that is connected to D or E. In certain embodiments, the electron withdrawing group is attached to the polymethine chain carbon that is connected to A or G. In other embodiments, the electron withdrawing group is attached to a carbon in the middle of the polymethine chain, *e.g.*, a carbon that is not connected to A, D, E, or G. In further embodiments, the compound contains two or more electron withdrawing groups. In these embodiments, one or more of the electron withdrawing groups may be attached to a polymethine chain carbon that is connected to A, D, E, or G. Placing electron-withdrawing groups at selected positions on the polymethine chain reduces photobleaching of cyanine and merocyanine dyes.

[0053] The conjugatable side chain of Y and/or Z comprises a linker and a reactive group and is one that is capable of conjugating the compound to a binding member having an affinity for a target molecule as described below. In some embodiments, the compound comprises one conjugatable side chain. In other

embodiments, the compound comprises more than one conjugatable side chain. In some embodiments, the compound does not comprise a conjugatable side chain.

[0054] In some embodiments, the conjugatable side chain is selected from the group consisting of:



wherein:

o is an integer from 1 to 5;

p is an integer from 1 to 5;

U is O or S;

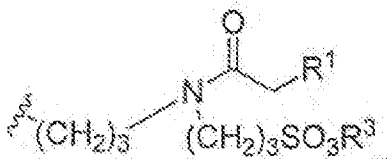
R^1 is a leaving group;

R^3 is hydrogen, Li^+ , Na^+ or K^+ ;

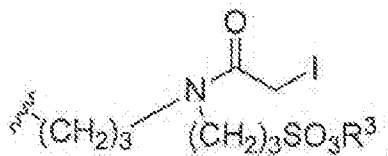
R^2 , R^4 , and R^6 are each independently C_{1-6} alkyl; and

R^5 is C_{1-6} alkyl or CF_3 .

[0055] In certain embodiments, R^1 is a halogen, *e.g.*, iodide. In some embodiments, the conjugatable side chain is

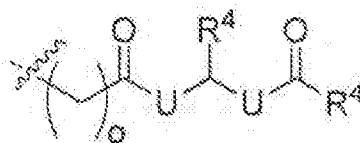


[0056] In other embodiments, the conjugatable side chain is



[0057] In some embodiments, at least one of Y and Z is a conjugatable side chain and the compound contains only one conjugatable side chain. In some embodiments, the conjugatable side chain is on the D or E moiety, *e.g.*, the Y in the D or E moiety. In other embodiments, the conjugatable side chain is on the A or G moiety, *e.g.*, the Y in the A or G moiety. In some embodiments, the Y in the D or E moiety is hydrogen or C₁₋₆ alkyl and one Z in the compound, *e.g.*, one Z in the D or E moiety or one Z in the A or G moiety, is a conjugatable side chain.

[0058] In some embodiments of the invention, the compounds comprise one or more substituents that are permeability enhancers, *i.e.*, substituents that increase the ability of the compound to cross cell membranes. In Formula I and Formula II, Y and/or Z may be one or more permeability enhancers. In some embodiments, the permeability enhancer is



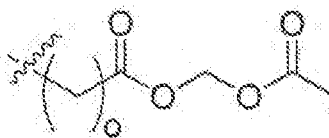
wherein:

o is an integer from 0 to 5;

U is O or S;

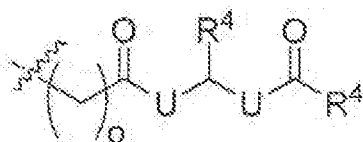
R⁴ is hydrogen or optionally substituted C₁₋₆ alkyl.

[0059] In certain embodiments, the permeability enhancer is an acetoxymethyl ester, *e.g.*,

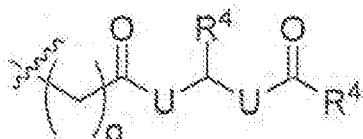


[0060] In some embodiments, the compound comprises one or more permeability enhancers, *e.g.*, 1, 2, 3, 4, or 5 permeability enhancers, *e.g.*, 1, 2, or 3 permeability enhancers. In certain embodiments, the one or more permeability enhancers are attached to the D or E moiety. In certain embodiments, the one or more permeability enhancers are attached to the A or G moiety. In certain embodiments, the compound comprises at least two permeability enhancers and the permeability enhancers are attached to both the D or E moiety and the A or G moiety. In some embodiments, the compound does not comprise a permeability enhancer.

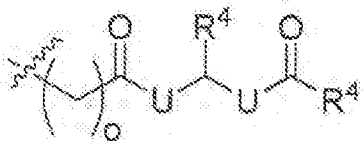
[0061] In some embodiments, the compound comprises a conjugatable side chain and one or more permeability enhancers. In certain embodiments, the Y in the D or E moiety is a conjugatable side chain and at least one Z is



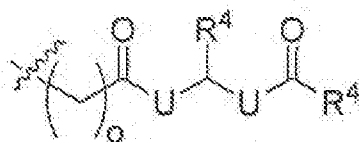
[0062] In some embodiments, the Y in the D or E moiety is hydrogen and at least one Z is



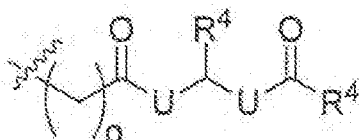
[0063] In some embodiments, the Y in the D or E moiety is hydrogen, one Z is a conjugatable side chain, and at least one Z is



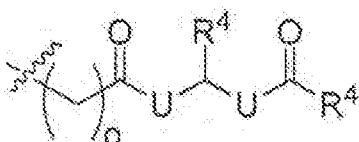
[0064] In some embodiments, the Y in the D or E moiety and at least one Z are



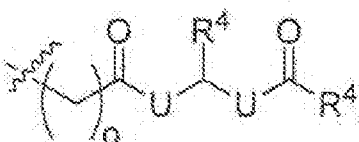
[0065] In some embodiments, one Z is a conjugatable side chain and the Y in the D or E moiety and at least one Z are



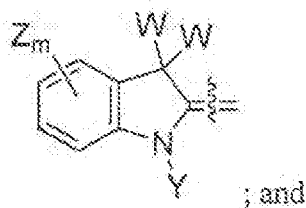
[0066] In some embodiments, the compound comprises at least two Z moieties that are



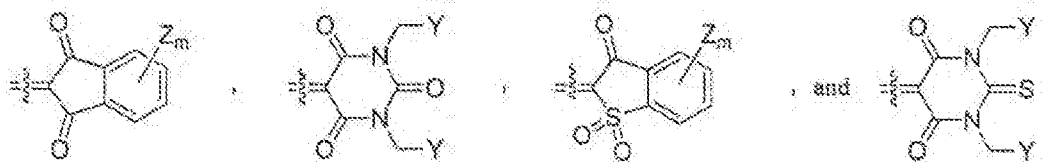
[0067] In some embodiments, the compound comprises at least three Z moieties that are



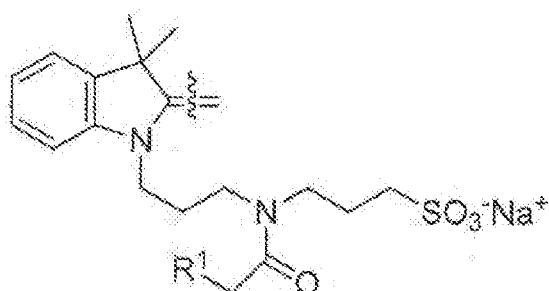
[0068] In certain embodiments of the compounds of the invention, D is



A is selected from the group consisting of



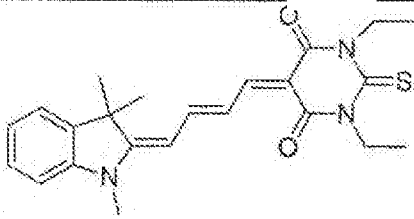
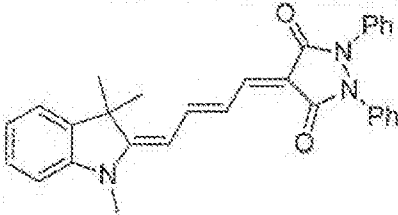
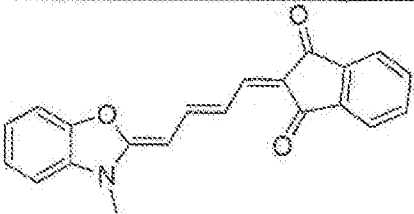
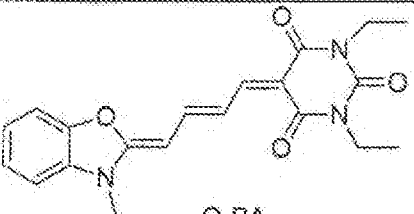
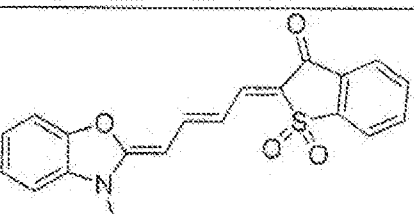
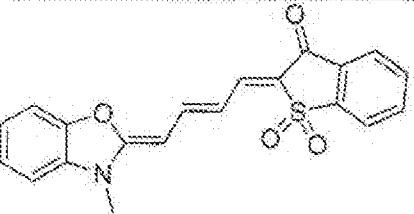
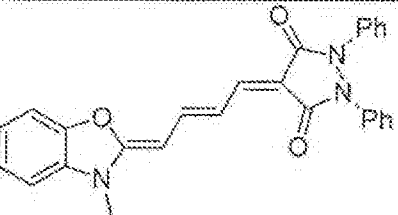
[0069] In certain embodiments, D is

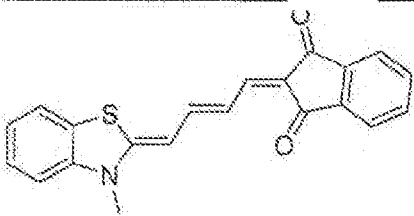
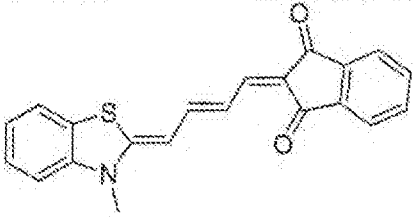
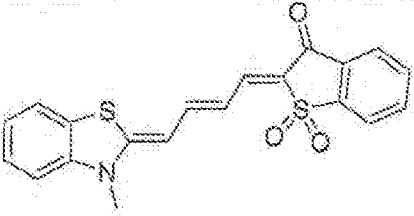
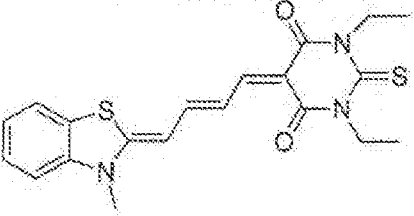
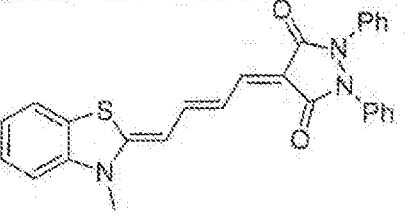
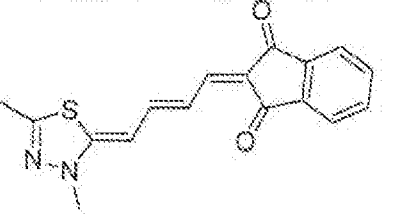
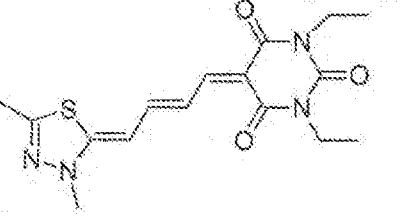


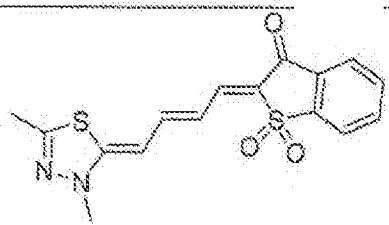
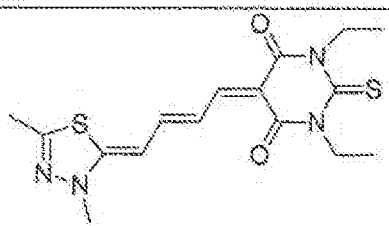
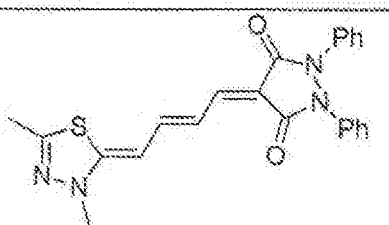
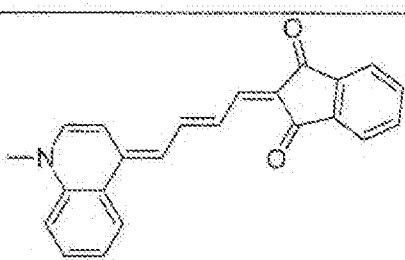
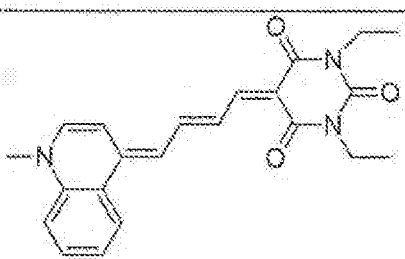
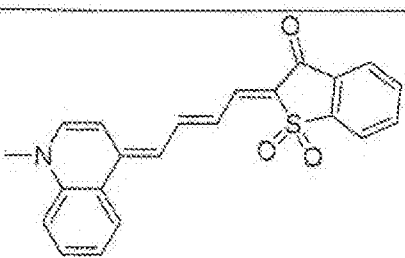
e.g., wherein R^1 is halogen, *e.g.*, I.

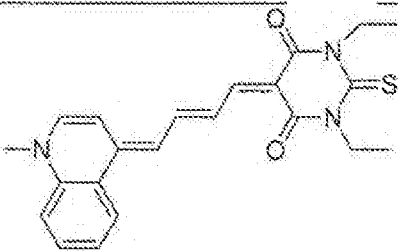
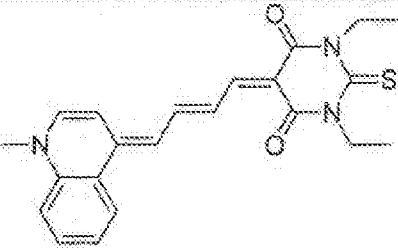
[0070] In some embodiments, the compound is selected from the following list or any subcombination thereof. Each of these compounds may further comprise a conjugatable sidechain and/or one or more permeability enhancers.

Compound	Structure
I-Pht	
I-BA	
I-SO	

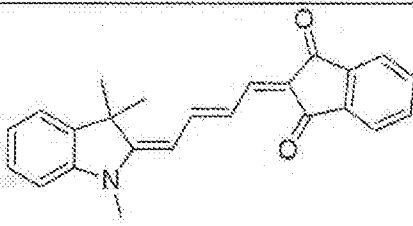
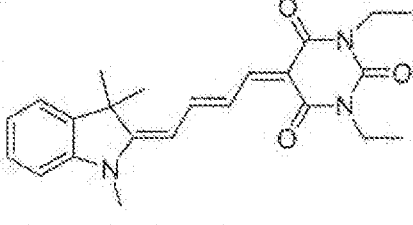
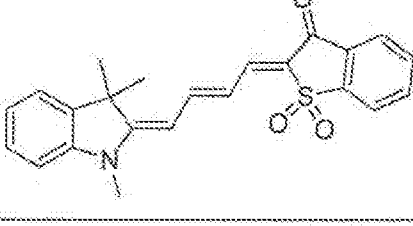
I-TBA	
I-Pz	
O-Pht	
O-BA	 O-BA
O-SO	
O-TBA	
O-Pz	

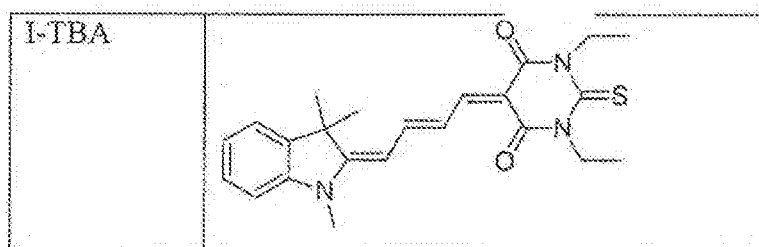
S-Pht	
S-BA	
S-SO	
S-TBA	
S-Pz	
TD-Pht	
TD-BA	

TD-SO	
TD-TBA	
TD-Pz	
Q-Pht	
Q-BA	
Q-SO	

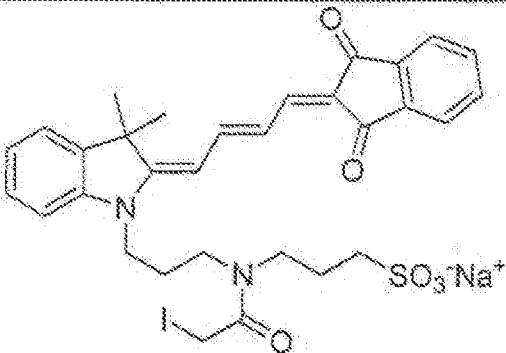
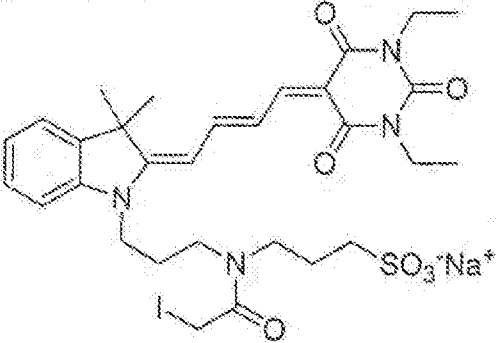
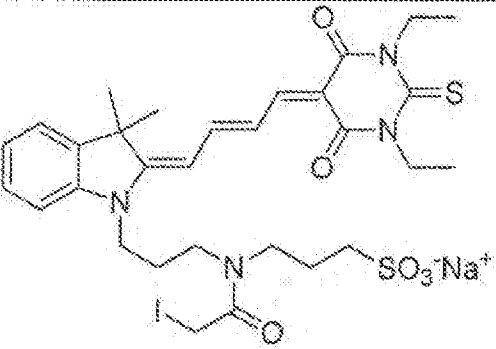
Q-TBA	
Q-Pz	

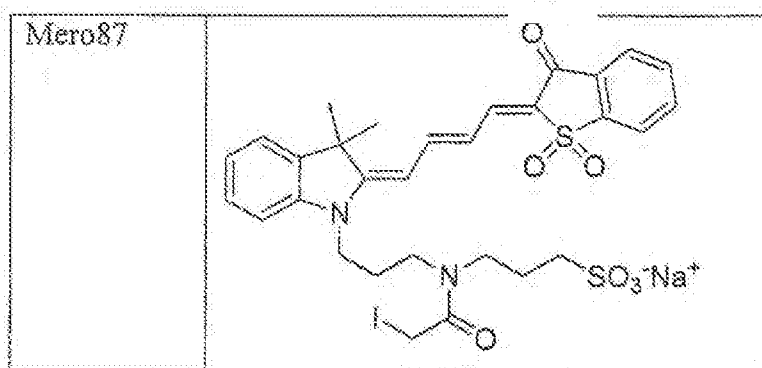
[0071] In some embodiments, the compound is selected from the following list or any subcombination thereof. Each of these compounds may further comprise a conjugatable sidechain and/or one or more permeability enhancers.

Compound	Structure
I-Pht	
I-BA	
I-SO	

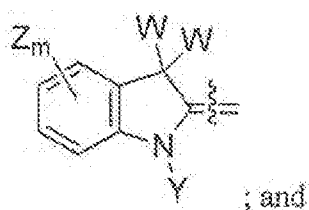


[0072] In particular embodiments, the compound is selected from the following list or any subcombination thereof.

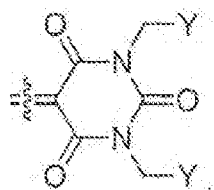
Compound	Structure
Mero60	
Mero61	
Mero62	



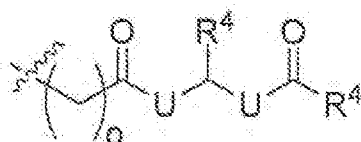
[0073] In certain embodiments of the compounds of the invention, D is



A is:

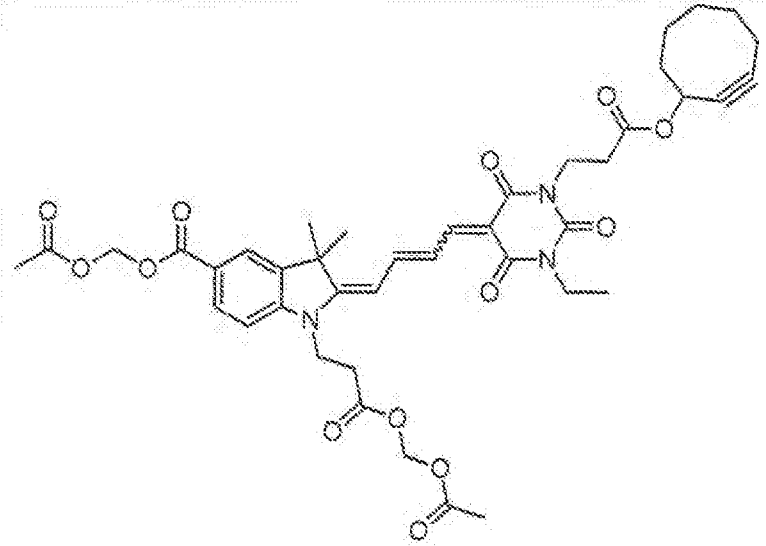
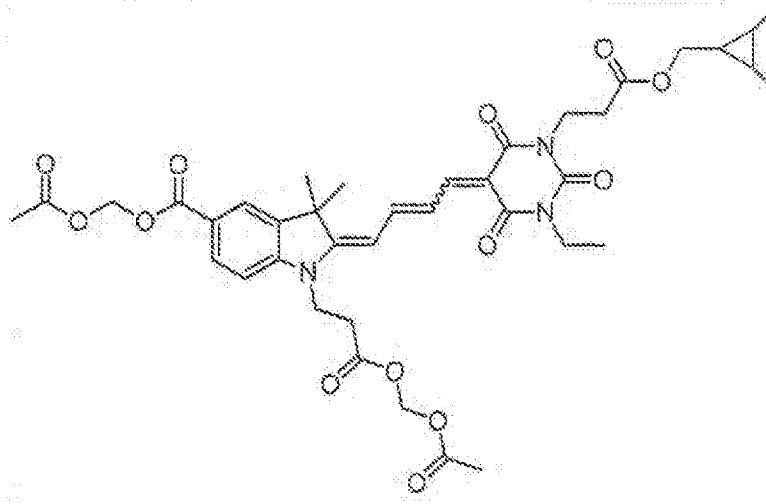


[0074] In some embodiments of these compounds, at least one Y is:



[0075] In particular embodiments, the compound is selected from the following list or any subcombination thereof. Each of these compounds may further comprise a conjugatable sidechain.

[illegible]

Merol168	 Chemical structure of Merol168, a complex molecule featuring a central indole-like core. The core is substituted with a long chain containing a double bond, a carbonyl group, and a terminal group consisting of a cyclooctene ring attached via an ester linkage. The structure also includes a side chain with a carbonyl group and a terminal group consisting of a cyclooctene ring attached via an ester linkage.
Merol169	 Chemical structure of Merol169, a complex molecule featuring a central indole-like core. The core is substituted with a long chain containing a double bond, a carbonyl group, and a terminal group consisting of a cyclooctene ring attached via an ester linkage. The structure also includes a side chain with a carbonyl group and a terminal group consisting of a cyclooctene ring attached via an ester linkage.

[0076] The presently disclosed dyes can be many times brighter than other dyes used previously to study antibody-antigen and other protein-protein interactions *in vitro*. See, for example, Renard *et al.*, *J. Mol. Biol.* 326:167-175 (2003); Sloan and Hellinga, *Protein Eng.*, 11:819-823 (1998); Iwatani *et al.*, *Biochemistry* 38:10318-10323 (1999). The quantum yield of the presently disclosed dyes can be high, for example, in hydrophobic environments. For example, for many of the presently disclosed dyes in a hydrophobic environment or solvent, the quantum yield, in some embodiments, is greater than about 0.4, in some embodiments, greater than about 0.5, in some embodiments, greater than about 0.6, in some embodiments, greater than about 0.7, in some embodiments, greater than about 0.8, and, in some embodiments, greater than about 0.9. Further, the combination of a high extinction coefficient and a

high quantum yield, as exhibited by some embodiments of the presently disclosed dyes, can provide a strong direct signal from biosensors comprising such dyes. Such high quantum yields and extinction coefficients can allow detections from very small amounts of biosensor, minimally perturbing the biological activity of the endogenous proteins being studied, and enabling high resolution kinetic studies by obtaining many images before photo-bleaching occurs.

[0077] In some embodiments, the “donor” ring structures (D and E) and/or the “acceptor” ring structures (A and G) of a compound of Formula I or Formula II can include one or more substituent groups. Such ring systems comprising one or more substituent groups can be referred to as “derivatives” of the parent compound, *i.e.*, a derivative of a compound of Formula I or Formula II. As used herein, a “derivative” refers to a chemical compound that is derived from or obtained from a parent compound and contains essential elements of the parent compound (*e.g.*, ability to fluoresce), but typically has one or more different functional groups. Such functional groups can be added to a parent compound, for example, to improve the molecule’s solubility, absorption, biological half life, fluorescent properties, and the like, or to decrease the toxicity of the molecule, eliminate or attenuate any undesirable side effect of the molecule, and the like.

[0078] In some embodiments, the substitution of one or more functional groups on the donor and/or acceptor ring structure can improve the solubility of the dye in particular solvents, such as water. For example, as described in U.S. Patent Application Publication No. 2006/0029946 to Hahn, the presently disclosed dyes can be designed to have enhanced water solubility by attaching substituent groups on the ring structure that sterically block aggregation without unduly increasing hydrophobicity. Under some circumstances, this approach can be more desirable than enhancing water solubility using highly charged groups that can affect protein interaction.

[0079] Planar merocyanine dyes known in the art are thought to form non-fluorescent hydrogen conjugates in water, thereby reducing the exposure of their hydrophobic surfaces to water. Such aggregation can be decreased by incorporating bulky, non-planar substituents with tetragonal geometry in the aromatic rings to make stacking unfavorable. This approach can provide dyes with good water solubility

while retaining substantial hydrophobic character. For example, in the presently disclosed I-SO dyes, the out-of-plane substituent groups can be two alkyl, *e.g.*, methyl, groups bound to a carbon atom of an aromatic ring, or two oxygen atoms bound to a sulfur atom of a heterocyclic ring structure. To this end, the presently disclosed dyes may have one or more groups arranged in, for example, a tetragonal geometry from rings of the dye. In many cases, the substituent groups include groups with one or more carbons providing steric hindrance to ring stacking, and/or polar to weak ionic character to enhance water solubility.

[0080] Accordingly, the substituents on one or more ring structures of the compound of Formula I or Formula II can be any functional group that inhibits stacking of the aromatic rings and/or enhances water solubility. Such substituent groups can provide steric hinderance to ring stacking, thereby inhibiting aggregation, without reducing water solubility of the dye. In some embodiments, the substituent group can have an aliphatic nature, *e.g.*, a straight-chain or branched alkyl group. In some embodiments, the substituent group can have a polar or weakly charged character. Generally, the presently disclosed dyes can be derivatized with side chains to prevent aggregation, improve solubility, affect interactions with proteins, enable covalent attachment to proteins, and improve cell membrane permeability.

[0081] In some embodiments, the presently disclosed compound of Formula I or Formula II are environmentally sensitive. As used herein, the term “environmentally sensitive” in relation to a compound of Formula I or Formula II refers to a dye in which a fluorescence signal from the dye changes when the dye is exposed to a change in environment, for example, a hydrophobicity, hydrogen bonding, polarity, or conformational change. Thus, a fluorescence signal from a compound of Formula I or Formula II detectably changes upon exposure to a change in solvent, change in hydrogen bonding, change in the hydrophobicity of the environment, changed polarity or polarization, or change affecting the conformation of the dye. In one embodiment, the fluorescence signal from the compound of Formula I or Formula II increases when the dye is exposed to an environment that is more hydrophobic. In another embodiment, the fluorescence signal provided by the compound of Formula I or Formula II increases when the dye is exposed to an environment where there is increased hydrogen binding between the dye and a component of the environment.

Such an increase in hydrophobicity or an increase in hydrogen bonding can occur when a presently disclosed biosensor comprising a compound of Formula I or Formula II binds to a target protein or subcellular component. In other embodiments, the fluorescence signal provided by the compound of Formula I or Formula II decreases when the dye is exposed to an environment that is more hydrophilic. In further embodiments, the fluorescence signal provided by the compound of Formula I or Formula II decreases when the dye is exposed to an environment that has less hydrogen bonding. Such an increase in hydrophilicity or a decrease in hydrogen binding can occur when a biosensor comprising a compound of Formula I or Formula II is exposed to an aqueous environment or when such a biosensor becomes unbound from a target protein or subcellular component.

[0082] In some embodiments, the presently disclosed dyes can have an extended zwitterion structure having a polarized ground state, which allows the presently disclosed dyes to respond to changes in solvent polarity. Such dyes can exhibit a strong solvent-dependent excitation wavelength shift, while retaining a bright fluorescence. Other solvent-sensitive dyes currently used *in vitro* are poorly suited for use in living cells. They are not sufficiently bright to quantify spectral changes from small amounts of labeled protein, and their short excitation and emission wavelengths damage cells and overlap cellular autofluorescence. To be bright and fluoresce at longer wavelengths, dyes must have extended conjugation, which can reduce water solubility and lead to self-aggregation. The dyes designed for use in living cells were specifically selected to be insensitive to their environment, because they are used to quantify protein distribution (*i.e.*, Fluorescein, Rhodamine, Alexa, Cy3/5). In such dyes, solubility problems could be overcome by incorporating charged groups around the dyes' edges. This approach might not be desirable in dyes that are intended to be used for detecting proteins, because those dyes must interact with hydrophobic protein regions to be useful for reporting protein conformational changes. Dyes must interact with the protein surface so that protein conformational changes affect their interactions with water, hydrogen bonding, or hydrophobic interactions. For domain- or antibody-based sensors, dyes must be able to move from water to hydrophobic pockets during protein binding events. Prior to the development of the presently disclosed dyes, use of solvatochromic dyes to monitor protein activity

in vivo was restricted to proteins that could be labeled with hydrophobic dyes in organic cosolvents. Hahn *et al.*, *J. Biol. Chem.* 265:20335-20345 (1990).

[0083] Such environmentally-sensitive dyes can be used as biosensors. Use of biosensors with the presently disclosed compound of Formula I or Formula II can provide advantages over currently available affinity probe schemes, such as those that involve imaging antibodies labeled with non-environmentally sensitive dyes. See, *e.g.*, Nizak *et al.*, *Science* 300:984-987 (2003). For example, the presently disclosed methods allow a change in fluorescence intensity or emission wavelength to be measured in near real time, as compared to the visualization of a target retrospectively by radioactive or ELISA formats. In many cases, the fleeting presence of target cannot be observed by methods with delayed detection.

[0084] The presently disclosed dyes can be compatible with aqueous chemistries or aqueous/organic solvent combinations used to prepare biosensors. The dyes can have reactive groups and/or protected groups to facilitate processes for binding dyes to binding members, *e.g.*, binding members having one or more binding domains. Biosensors comprising a compound of Formula I or Formula II, *e.g.*, dyes linked to a binding member, can be compatible with and move freely in intracellular and/or extracellular environments of living cells. Dyes on the biosensors can exist in or near binding regions between a sensor and target to provide a detectable signal without significantly interfering with binding.

[0085] The presently disclosed compounds of Formula I or Formula II generally have an emission wavelength in the visible spectral region and can, in some embodiments, exhibit a ratiometric response to one or more target molecules. Embodiments that exhibit a shift in excitation or emission wavelength upon ligand binding enable a biosensor comprising a presently disclosed compound of Formula I or Formula II to be self-referencing. In such embodiments, the fluorophore exhibits an increase in a first excitation or emission wavelength in the presence of a target molecule and a decrease in a second wavelength. A ratio between the first wavelength and the second wavelength can be calculated to determine the amount of target molecule in the sample under test. Such self referencing can correct for variations in excitation source intensity and other sources of noise and instability in the biosensor without requiring a reference dye. Thus, a single fluorophore can be

used to observe a ratiometric response in the biosensor. As used herein, the term “ratiometric response” means that the intensities of the first wavelength and the second wavelength vary such that the ratio of the two emission wavelengths can be used to indicate interaction with the target molecule in a sample.

II. Biosensors Comprising Compounds of Formula I or Formula II

[0086] In some embodiments, the presently disclosed subject matter provides a biosensor comprising a compound of Formula I or Formula II and a binding member having a specific affinity for a target molecule that can be used to detect and quantify the presence and/or activities of one or more target molecules. The biosensor can bind to one or more target molecules and the detection of such binding events can be used in homogeneous assays, *i.e.*, assays in which the fluorescent properties of the dye can undergo a change as a result of the binding event, and for live-cell imaging. The presently disclosed biosensors can be used to detect and/or quantify diverse protein activities, including changing subcellular locations, conformational changes, activation states, posttranslational modifications, and/or small ligand binding of proteins *in vivo*.

[0087] As used herein, the terms “biosensor” and “biosensor compound” generally refer to a compound that undergoes a detectable change in specific response to a target molecule. The presently disclosed biosensors combine the molecular recognition properties of biological macromolecules, such as a binding member of a specific binding pair, or small molecules, with a fluorophore capable of producing a detectable signal, *e.g.*, a detectable change in one or more fluorescent properties, upon binding with a target molecule. The term “producing a detectable signal” refers to the ability to recognize a change in a property of a reporter group, *e.g.*, a fluorophore, in a manner that enables the detection of ligand-protein binding.

[0088] The present invention provides methods which can use specific binding partners for a particular target molecule of interest. A specific binding partner or member, as used herein, is a member of a specific binding pair. A “specific binding pair” refers to two different molecules where one of the molecules through chemical or physical means specifically binds the second molecule. In this sense, a target

molecule is a reciprocal member of a specific binding pair. Further, specific binding pairs can include members that are analogs of the original specific binding partners, for example, an analog of a target molecule can have a similar structure to the target molecule. By "similar" it is intended that, for example, an analog has an amino acid sequence that has at least about 60% or 65% sequence identity, about 70% or 75% sequence identity, about 80% or 85% sequence identity, about 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or greater amino acid sequence identity compared to a target molecule amino acid sequence using alignment programs and standard parameters well known in the art. An analog of a target molecule also can have the same function as the target molecule.

[0089] A biosensor can translate a binding event into a directly measurable fluorescent property. Such changes in one or more fluorescence properties include, but are not limited to, an increase or decrease in fluorescence intensity, a shift in excitation or emission maxima, shape of the excitation or emission band profiles, a change in fluorescence lifetime, a change in anisotropy, a change in polarization, and combinations thereof. Further, the producing of a detectable signal can be reversible or non-reversible. The signal-producing event includes continuous, programmed, and episodic means, including one-time or reusable applications. The reversible signal-producing event can be instantaneous or can be time-dependent, so long as a correlation with the presence or concentration of target molecule is established.

[0090] As used herein, the term "fluorophore" includes a moiety of a larger molecule or conjugate that can be induced to emit fluorescence when irradiated, *i.e.*, excited, by electromagnetic radiation of an appropriate wavelength. More particularly, a fluorophore can be a functional group of a molecule or conjugate that absorbs light of a certain wavelength and emits light at different wavelength. The intensity and the wavelength of the light emitted, as well as other fluorescence properties including, but not limited to, fluorescence lifetime, anisotropy, polarization, and combinations thereof, depend on the identity of the fluorophore and its chemical environment. A fluorophore can include a fluorescent molecule, such as the presently disclosed compounds of Formula I or Formula II.

[0091] In some embodiments, the presently disclosed compounds of Formula I or Formula II include a reactive group, *e.g.*, substituents Y and Z, which can be

conjugated with another molecule, such as a binding member, *e.g.*, one member of a specific binding pair, which has an affinity for a target molecule of interest, *e.g.*, the other member of a binding pair. Such conjugates can be used as a biosensor compound to detect the presence, location, conformation, activation state, modification, binding, and the like, of a target molecule.

[0092] As used herein, the term “conjugate,” and variations thereof, refers to a molecule comprising two or more subunits bound together, optionally through a linking group, to form a single molecular structure. The binding of the two or more subunits can be through a direct chemical bond between the subunits or through a linking group. Such binding in a conjugate typically is irreversible. As used herein, the term “affinity” refers to the strength of the attraction between one binding member of a specific binding pair to another binding member of a specific binding pair at a particular binding site. The term “specificity” and derivations thereof, refer to the likelihood that a binding member will bind to another member of a specific binding pair. Such binding between one binding member, *e.g.*, a binding protein, to another binding member of a specific binding pair, *e.g.*, a ligand or analyte, can be reversible.

[0093] In some embodiments, substituents Y and/or Z can be a conjugatable sidechain and the presently disclosed dyes can be linked to a binding member of a specific binding pair, such as a binding protein, or a portion thereof, such as a binding domain, or a small molecule, to prepare a biosensor. In some embodiments, the binding member includes one or more conjugatable sidechain-reactive groups, *i.e.*, functional groups that can attach to, *e.g.*, form a covalent bond with, a conjugatable sidechain, and the dye includes a conjugatable sidechain. Alternatively, in some embodiments, the dye includes the conjugatable sidechain-reactive group and the binding member includes the conjugatable sidechain. In another embodiment, both the binding member and dye include conjugatable sidechain-reactive groups. In such embodiments, the linkage is through a bivalent linker group. Generally, a linkage can be formed by contacting the binding member and dye in solution. The conjugatable sidechain can be a flexible chain-like structure, such as an aliphatic chain or a polymer, *e.g.*, a primary alkyl, polyethylene glycol, or alkylamido chain. In some embodiments, the conjugatable sidechain is a peptide or nucleic acid having hydrophilic character.

[0094] In some embodiments, conjugatable sidechains can be provided on compounds of Formula I or Formula II. In such embodiments, the presently disclosed compounds of Formula I or Formula II can be derivatized to make reactive forms which could be site-specifically attached to proteins. As used herein, the term “derivatized,” and variations thereof, is meant to include any chemical modification, addition, deletion, or substitution to a parent compound. Further, a derivative can include any reaction product of the derivative, for example, the reaction product of the derivative with an amino acid residue. Accordingly, in some embodiments, the presently disclosed dyes can include an aromatic ring structure having a conjugatable sidechain that can be conjugated, e.g., covalently attached, to an amino acid, for example, an amino acid residue of a protein. A non-limiting example of a derivative is an ester or amide of a parent compound having a carboxylic acid functional group.

[0095] For example, derivatives of the presently disclosed compounds of Formula I or Formula II can be made with succinimidyl ester conjugatable sidechain for attachment to lysine or with an iodoacetyl linker for selective reaction with cysteine. Use of these groups for site-specific protein labeling is known in the art. *See, e.g.,* Dent and Aslarn, *Bioconjugation*, 364-482 (1998). For example, the reactive compounds of Formula I or Formula II can be synthesized from starting materials having an amino group, which can be used for attachment of a side chain comprising a reactive linker group at the end of the synthesis. Accordingly, in some embodiments, cysteine-selective iodoacetyl groups can be attached to the dye ring structure for covalent binding to a binding member or binding domain thereof.

[0096] In some embodiments, the presently disclosed dyes can include a thiol-reactive group that can be conjugated to the thiol moiety of a cysteine amino acid residue in a natural or unnatural protein. As used herein, the term “thiol-reactive group” refers to a substituent group that can react with a thiol moiety to form a carbon-sulfur bond. Examples of suitable thiol-reactive groups that can be introduced into the presently disclosed dyes include a halo-acetyl group, *i.e.*, a halo-acetamido group. In some embodiments, the halo-acetyl group includes an iodoacetyl group. One of ordinary skill in the art upon review of the presently disclosed subject matter would recognize that other thiol-reactive groups known in the art, such as maleimide

and dithiopyridyl groups, are suitable for use with the presently disclosed subject matter.

[0097] In some embodiments, the residue to which the dye is attached is a cysteine. In some embodiments, one or more such residues can include conjugatable sidechain-reactive groups other than cysteine. Conjugatable sidechain-reactive groups can include any chemical functional groups on the dye, the binding member, or both, that can react with selected conjugatable sidechains to form one or more chemical bonds. For example, natural amino acids, modified or derivatized amino acids, and/or other residues in a binding member or binding domain can provide conjugatable sidechain-reactive groups, such as amines, sulfhydryls, carboxylic acids, alcohols, aldehydes, and thiols, that can covalently bond to linker molecules.

[0098] Conjugatable sidechains can be any type suitable to react with conjugatable sidechain-reactive groups to form a linkage between selected binding members to compounds of Formula I or Formula II. Representative conjugatable sidechains suitable for use with the presently disclosed subject matter include, but are not limited to, sidechains reactive with primary amines (*e.g.*, hydroxysuccinimide, carboxyl, NHS ester, sulfo-NHS ester, isothiocyanate, imidoester, epoxide), side chains reactive with sulfhydryl groups (*e.g.*, maleimides, haloacetyls, dithiopyridyls), side chains reactive with aldehydes and other carbonyl groups (*e.g.*, hydrazines, hydrazides), sidechains reactive with carboxyl groups (*e.g.*, ethyldiethylamino propylcarbodiimide (EDC)) and sidechains reactive with bioorthogonal reactive groups such as terminal alkyne, cycloalkyne, alkene, tetrazine, 1,3-dipole (*e.g.*, azide, nitron, nitrile oxide). Conjugatable sidechains can include a flexible aliphatic or polymer chain of suitable length and hydrophilicity to bridge between linked molecules. Bivalent linker groups having the same or different linker chemistries at each end also can be used. Conjugatable sidechains can include one or more protective group to protect the linker group during storage, handling or other chemistries. The protective group can be removed under defined conditions to allow completion of the conjugation reaction.

[0099] Further, based on sequence data and the crystal structure of binding domains of the binding members, one or more compounds of Formula I or Formula II can be attached to provide the presently disclosed biosensors. For example, based on

the sequence data and the crystal structure of a free Fab fragment, several cyanine and merocyanine dyes can be attached to a series of one or more different residue positions to create a small but intelligently designed library of biosensor candidates. The resultant library can be screened, as can be appreciated by those in the art, to identify candidates with desired properties, *e.g.*, of signal strength and binding affinity.

[0100] Intelligent design of linkage reactions can be accomplished with the selection of a target reactive group in a binding member or binding domain, or engineering of a desired reactive group into the binding member or binding domain. For example, a binding member or binding domain can naturally have a cysteine residue (reactive, for example, with a haloacetyl linker) at an appropriate location, or can be mutated to include such a cysteine residue. Alternatively, a binding member or a binding domain can have a small number of cysteine residues. Binding members or binding domains can be modified or mutated to contain a single or a small number of cysteine residues by procedures known in the art.

A. Representative Binding Members

[0101] Binding members suitable for use with the presently disclosed subject matter can include any molecules that bind to a target molecule with suitable specificity. Binding members typically include binding regions of affinity molecules known in the art including, but not limited to, antibodies, antibody fragments, leucine zippers, histones, enhancers, complementary determining regions (CDRs), to single chain variable fragments (scFv's), receptors, ligands, aptamers, lectins and one of several proteins in a protein complex or a protein pair. Binding members can comprise either member of a binding pair, *e.g.*, a pair of proteins in a protein-protein interaction, with the binding member being identified as the member introduced into an assay system to probe for a target molecule. Binding members or binding domains thereof can be binding regions of, for example, full sized versions of the affinity molecule, fragments of the affinity molecule, or the smallest portion of the affinity molecule providing binding useful in the detection of a target molecule. In some embodiments, the binding member can have specific affinity to endogenous (*e.g.*, constitutive or inducible, but not recombinant) peptides of a cell.

[0102] In some embodiments, the binding member is a small molecule that has affinity for a target molecule, *e.g.*, a protein. In some embodiments, the small molecule is one that cannot be produced by genetic means, *e.g.*, is not a peptide or protein. In some embodiments, the small molecule is a compound having a molecular weight of less than about 1000 Da, *e.g.*, less than about 800 Da or about 500 Da. In some embodiments, the small molecule binds a specific protein, a protein in a specific conformation, or a protein having a specific post-translational modification. In certain embodiments, the small molecule can be a compound that is a known or later identified therapeutic, prophylactic or diagnostic agent that binds to a target in or on a cell. In certain embodiments, the small molecule is one that can enter a living cell, *e.g.*, by diffusion through the membrane, by entry through channels, or by active or passive transport. In some embodiments, the small molecule is conjugated to a compound of Formula I or Formula II. In other embodiments, the small molecule is conjugated to any dye that is known to be or later identified to be useful for visualization in living cells or animals, *e.g.*, to visualize interactions between proteins and small molecules. In some embodiments, the dye may be one that is permeable to cell membranes to enhance its usefulness in living cells or animals. In some embodiments, the dye may be one that can be visualized by microscopy, *e.g.*, one that emits a fluorescent or other signal.

[0103] In some embodiments, the binding member is a binding protein. As used herein, the term "binding protein" refers to a protein, that when conjugated with a fluorophore, interacts with a specific target molecule in a manner capable of producing a detectable fluorescence signal differentiable from when a target molecule is present or absent, or when a target molecule is present in varying concentrations over time.

[0104] The binding member can comprise a binding domain. Binding domains suitable for use with the presently disclosed biosensors can comprise polypeptide, peptide, or nucleic acid sequences. For example, binding domains can be single stranded DNA (ssDNA), double stranded DNA (dsDNA), RNA, nucleic acids with modified bases, and the like. In one embodiment, the binding domain can be an oligonucleotide probe and the target can be a complimentary target nucleic acid. In another embodiment, the binding domain can be a dsDNA strand specific to a target

enhancer protein target. The presently disclosed compounds of Formula I or Formula II can be linked to nucleic acids by any technique known in the art, such as by reaction of linker groups on the dye to reactive groups available on modified bases on the nucleic acid.

[0105] Affinity specificity of peptide binding domains can be provided by a short sequence of amino acids (*e.g.*, a sequence of amino acids comprising 3 to 20 residues, *i.e.*, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 residues), or the specificity can rely on contributions of amino acid side chains brought in proximity by the primary, secondary, tertiary, and/or quaternary structural conformations of one or more affinity proteins.

[0106] Binding domains comprising peptides can have natural amino acid side chains, modified side chains, or the like, that provide reactive groups specifically reactive with conjugatable sidechains on compounds of Formula I or Formula II. In some embodiments, the compound of Formula I or Formula II can have reactive groups specifically reactive with linker groups present on the binding domain to link the dye to the domain. The position of a dye on a domain can be determined by the location of a reactive group or linker moiety on the domain. In some embodiments, the binding domain has one or more cysteine residues reactive with groups on the dyes, for example, iodoacetyl groups on the dyes.

[0107] The presently disclosed biosensors can incorporate binding domains of naturally occurring proteins having specific binding activity. The binding domains can include, for example, full length affinity proteins, members of protein-protein interaction pairs (or portions thereof), Fv antibody fragments, aptamers, Vh antibody fragments, and the like. In one embodiment, the biosensors comprise binding domains that are members of the immunoglobulin family of proteins, or derivatives thereof. For example, the binding domain can be a complete immunoglobulin, fragment, single chain variable fragment (scFv), a heavy or light chain variable region, a CDR peptide sequence, and/or the like.

[0108] In some embodiments, an antibody or antibody fragment can be used in a binding domain to which compounds of Formula I or Formula II can be attached to form a biosensor. An antibody suitable for use with the presently disclosed subject matter can be in any of a variety of forms, including a whole immunoglobulin, an

antibody fragment such as Fv, Fab, and similar fragments, a single chain antibody which includes the variable domain complementarity determining regions (CDR), and the like forms, all of which fall under the broad term "antibody," as used herein. The use of any specificity of an antibody, polyclonal or monoclonal, is intended to fall under the scope of the presently disclosed subject matter. In some embodiments, in the context of methods described herein, an antibody, or fragment thereof is used that is immunospecific for a selected target.

[0109] As used herein, the term "antibody fragment" refers to a portion of a full-length antibody, generally the antigen binding or variable region. Examples of antibody fragments include Fab, Fab', F(ab')₂ and Fv fragments. Papain digestion of antibodies produces two identical antigen binding fragments, called Fab fragments, each with a single antigen binding site, and a residual Fc fragment. Fab fragments thus have an intact light chain and a portion of one heavy chain. Pepsin treatment yields an F(ab')₂ fragment that has two antigen binding fragments that are capable of cross-linking antigen, and a residual fragment that is termed a pFc' fragment. Fab' fragments are obtained after reduction of a pepsin digested antibody, and consist of an intact light chain and a portion of the heavy chain. Two Fab' fragments are obtained per antibody molecule. Fab' fragments differ from Fab fragments by the addition of a few residues at the carboxyl terminus of the heavy chain CH1 domain including one or more cysteines from the antibody hinge region.

[0110] Fv is a small antibody fragment that contains a complete antigen recognition and binding site. This region consists of a dimer of one heavy and one light chain variable domain in a tight, non-covalent association (V_H-V_L dimer). It is in this configuration that the three CDRs of each variable domain interact to define an antigen binding site on the surface of the V_H-V_L dimer. Collectively, the six CDRs confer antigen binding specificity to the antibody. Even a single variable domain (or half of an Fv comprising only three CDRs specific for an antigen), however, has the ability to recognize and bind antigen, although at a lower affinity than the entire binding site. Antibody fragments usefully incorporated into the presently disclosed biosensors can include, but are not limited to, single CDRs, V_H regions, V_L regions, Fv fragments, F(ab) and F(ab')₂ fragments.

[0111] Additional fragments can include diabodies, linear antibodies, single-chain antibody molecules, and multispecific antibodies formed from antibody fragments. Antibody fragments suitable for use with the presently disclosed binding domains can include, e.g., natural, synthetic, or recombinant versions. Single chain antibodies are genetically engineered molecules containing the variable region of a light chain and a variable region of the heavy chain, linked by a suitable polypeptide linker as a genetically fused single chain molecule. Such single chain antibodies also are referred to as "single-chain Fv" or "scFv" antibody fragments. Generally, the Fv polypeptide further comprises a polypeptide linker between the V_H and V_L domains that enables the scFv to form the desired structure for antigen binding. For a description of scFv, see Pluckthun, in *The Pharmacology of Monoclonal Antibodies*, vol. 113, Rosenberg and Moore eds. Springer-Verlag, N.Y., pp. 269-315 (1994).

[0112] In some embodiments, the binding domain can be a single chain variable fragment (scFv) of an antibody with an attachment site for a dye within a CDR3 region of the scFv, a modified fragment of a naturally occurring protein, or another entity that binds to a specific state of the targeted protein or polypeptide. Any available scFv can be used provided that it binds to a selected target with sufficient affinity to permit detection of a complex formed between the scFv and the selected target. A compound of Formula I or Formula II can be attached to the selected scFv at any convenient site. The fluorescence of the attached dye, however, can be influenced by its attachment position. Single chain variable fragment (scFv) binding domains can be useful in modular biosensors in which binding domain and/or target modules connected with a linker can be replaced with alternate versions to provide new desired specificities to the sensor. In some embodiments, scFvs can be combined with compounds of Formula I or Formula II for methods of probing living cells.

[0113] Antibody fragments suitable for use with the presently disclosed subject matter do not have to be full-length antibodies. Such antibody fragments, however, can have similar or improved immunological or other properties relative to a full-length antibody. For example, such antibody fragments can be smaller and more stable than full-length antibodies. Such antibody fragments can include about 4 amino acids, 5 amino acids, 6 amino acids, 7 amino acids, 9 amino acids, about 12

amino acids, about 15 amino acids, about 17 amino acids, about 18 amino acids, about 20 amino acids, about 25 amino acids, about 30 amino acids, or more.

[0114] In general, an antibody fragment can have any upper size limit so long as it has similar or improved properties relative to an antibody that binds with specificity to a target molecule. For example, smaller antibody fragments can have less than about 200 amino acids, less than about 175 amino acids, less than about 150 amino acids, or less than about 120 amino acids if the antibody fragment is related to a light chain antibody subunit. Moreover, larger antibody fragments can have less than about 425 amino acids, less than about 400 amino acids, less than about 375 amino acids, less than about 350 amino acids, less than about 325 amino acids or less than about 300 amino acids if the antibody fragment is related to a heavy chain antibody subunit.

[0115] Antibodies and antibody fragments directed against selected targets can be prepared by techniques commonly known in the art. In some embodiments, antibody fragments can be prepared from full-length antibodies. Methods for the preparation of polyclonal antibodies are available to those skilled in the art. See, for example, Green *et al.*, Production of Polyclonal Antisera, in: *Immunochemical Protocols* (Manson, ed.), pages 1-5 (Humana Press); Coligan *et al.*, Production of Polyclonal Antisera in Rabbits, Rats Mice and Hamsters, in: *Current Protocols in Immunology*, section 2.4.1 (1992), which are incorporated herein by reference in their entirety. Such polyclonal antibodies can be cleaved, e.g., by chemical or enzymatic treatment to prepare antibody fragments useful in the presently disclosed methods.

[0116] Monoclonal antibodies, and fragments thereof, also can be used in the presently disclosed methods. The term "monoclonal antibody" as used herein refers to an antibody obtained from a population of substantially homogeneous antibodies. In other words, the individual antibodies comprising the population are identical except for occasional naturally occurring mutations in some antibodies that can be present in minor amounts. Monoclonal antibodies are highly specific, being directed against a single antigenic site. Further, in contrast to polyclonal antibody preparations that typically include different antibodies directed against different determinants (epitopes), each monoclonal antibody is directed against a single determinant on the antigen. In addition to their specificity, the monoclonal antibodies are advantageous in that they are synthesized by the hybridoma culture, uncontaminated by other

immunoglobulins. The modifier "monoclonal" indicates the character of the antibody as being obtained from a substantially homogeneous population of antibodies, and is not to be construed as requiring production of the antibody by any particular method. The monoclonal antibodies herein specifically include "chimeric" antibodies in which a portion of the heavy and/or light chain is identical or homologous to corresponding sequences in antibodies derived from a particular species or belonging to a particular antibody class or subclass, while the remainder of the chain(s) are identical or homologous to corresponding sequences in antibodies derived from another species or belonging to another antibody class or subclass. Fragments of such antibodies also can be used, so long as they exhibit the desired biological activity. See U.S. Pat. No. 4,816,567; Morrison *et al.*, *Proc. Natl. Acad. Sci.* 81:6851-55 (1984).

[0117] The preparation of monoclonal antibodies is conventional. See, for example, Kohler and Milstein, *Nature*, 256:495 (1975); Coligan *et al.*, sections 2.5.1-2.6.7; and Harlow *et al.*, in: *Antibodies: A Laboratory Manual*, page 726 (Cold Spring Harbor Pub. (1988)), which are incorporated herein by reference in their entirety. Monoclonal antibodies can be isolated and purified from hybridoma cultures by a variety of established techniques. Such isolation techniques include affinity chromatography with Protein-A Sepharose, size-exclusion chromatography, and ion-exchange chromatography. See, *e.g.*, Coligan *et al.*, sections 2.7.1-2.7.12 and sections 2.9.1-2.9.3; Barnes *et al.*, Purification of Immunoglobulin G (IgG), in: *Methods in Molecular Biology*, Vol. 10, pages 79-104 (Humana Press (1992)).

[0118] Methods of *in vitro* and *in vivo* manipulation of antibodies are available to those skilled in the art. For example, the monoclonal antibodies to be used in accordance with the presently disclosed subject matter can be made by the hybridoma method or can be made by recombinant methods, *e.g.*, as described in U.S. Pat. No. 4,816,567, which is incorporated herein by reference in its entirety. Monoclonal antibodies suitable for use with the presently disclosed subject matter also can be isolated from phage antibody libraries using the techniques described in Clackson *et al.*, *Nature* 352:624-628 (1991), as well as in Marks *et al.*, *J. Mol. Biol.* 222:581-597 (1991).

[0119] Methods of making antibody fragments are also known in the art (see for example, Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor

Laboratory, New York, (1988), incorporated herein by reference). Antibody fragments suitable for use with the presently disclosed subject matter can be prepared by proteolytic hydrolysis of the antibody or by expression of nucleic acids encoding the antibody fragment in a suitable host. Antibody fragments can be obtained by pepsin or papain digestion of whole antibodies conventional methods. For example, antibody fragments can be produced by enzymatic cleavage of antibodies with pepsin to provide a 5S fragment described as $F(ab')_2$. This fragment can be further cleaved using a thiol reducing agent, and optionally using a blocking group for the sulfhydryl groups resulting from cleavage of disulfide linkages, to produce 3.5S Fab' monovalent fragments. Alternatively, enzymatic cleavage using pepsin produces two monovalent Fab' fragments and an Fc fragment directly. These methods are described, for example, in U.S. Pat. Nos. 4,036,945 and 4,331,647, and references contained therein, which patents are incorporated herein by reference in their entireties.

[0120] Other methods of cleaving antibodies, such as separation of heavy chains to form monovalent light-heavy chain fragments, further cleavage of fragments, or other enzymatic, chemical, or genetic techniques also can be used, so long as the fragments bind to the antigen that is recognized by the intact antibody. For example, Fv fragments comprise an association of V_H and V_L chains. This association can be noncovalent or the variable chains can be linked by an intermolecular disulfide bond or cross-linked by chemicals such as glutaraldehyde. Preferably, the Fv fragments comprise V_H and V_L chains connected by a peptide linker. These single-chain antigen binding proteins (sFv) are prepared by constructing a structural gene comprising DNA sequences encoding the V_H and V_L domains connected by an oligonucleotide. The structural gene is inserted into an expression vector, which is subsequently introduced into a host cell such as *E. coli*. The recombinant host cells synthesize a single polypeptide chain with a linker peptide bridging the two V domains. Methods for producing sFvs are described, for example, by Whitlow *et al.*, *Methods: a Companion to Methods in Enzymology*, Vol. 2, page 97 (1991); Bird *et al.*, *Science* 242:423-426 (1988); Ladner *et al.*, U.S. Pat. No. 4,946,778; and Pack *et al.*, *Bio/Technology* 11:1271-77 (1993).

[0121] Another form of an antibody fragment is a peptide coding for a single complementarity-determining region (CDR). CDR peptides ("minimal recognition

units") are often involved in antigen recognition and binding. CDR peptides can be obtained by cloning or constructing genes encoding the CDR of an antibody of interest. Such genes are prepared, for example, by using the polymerase chain reaction to synthesize the variable region from RNA of antibody-producing cells. See, for example, Larrick *et al.*, *Methods: a Companion to Methods in Enzymology*, Vol. 2, p. 106 (1991).

[0122] The amino acid sequence of monoclonal antibodies, polyclonal antibodies, fragments thereof, and such, can be determined by amino acid sequencing methods known in the art. Amino acid sequences of antibodies and fragments of interest can be evaluated for binding sequences, such as, *e.g.*, CDR sequences, useful in the presently disclosed biosensors. Desired sequences can be directly synthesized or translated into nucleic acid sequences for manipulation by genetic engineering techniques known in the art. Amino acid and nucleic acid sequences thus obtained can be screened for binding characteristics desirable in the presently disclosed biosensors. Optionally, sequences thus obtained can be logically modified and/or randomly mutated to generate additional binding domain candidates that can be screened to identify sequences most useful in particular biosensor systems of interest.

[0123] Human and humanized forms of non-human (*e.g.*, murine) antibodies also are suitable for use with the presently disclosed subject matter. Such humanized antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, scFv, Fab, Fab', F(ab')₂ or other antigen-binding subsequences of antibodies) that contain minimal sequence derived from non-human immunoglobulin. For the most part, humanized antibodies are human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a nonhuman species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity.

[0124] In some embodiments, Fv framework residues of the human immunoglobulin can be replaced by corresponding non-human residues. Furthermore, humanized antibodies can comprise residues that are found neither in the recipient antibody nor in the imported CDR or framework sequences. These modifications can be made to further refine and optimize antibody performance. In

general, humanized antibodies will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody optionally can comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin. For further details, see Jones *et al.*, *Nature* 321:522-525 (1986); Reichmann *et al.*, *Nature* 332:323-329 (1988); Presta, *Curr. Op. Struct. Biol.* 2:593-596 (1992); Holmes *et al.*, *J. Immunol.* 158:2192-2201 (1997) and Vaswani *et al.*, *Annals Allergy, Asthma & Immunol.* 81:105-115 (1998).

[0125] In the immune system, specific antibodies are selected and amplified from a large library (affinity maturation). The combinatorial techniques employed in immune cells can be mimicked by mutagenesis and generation of combinatorial libraries of binding entities. Variant binding entities, antibody fragments and antibodies therefore also can be generated through display-type technologies. Such display-type technologies include, for example, phage display, retroviral display, ribosomal display, yeast display and other techniques. Techniques available in the art can be used for generating libraries of binding entities, for screening those libraries and the selected binding entities can be subjected to additional maturation, such as affinity maturation. Wright and Harris, *supra*, Hanes and Pluchau, *PNAS USA* 94:4937-4942 (1997) (ribosomal display), Parmley and Smith, *Gene* 73:305-318 (1988) (phage display), Scott, *TIBS* 17:241-245 (1992), Cwirla *et al.*, *PNAS USA* 87:6378-6382 (1990), Russel *et al.*, *Nucl. Acids Research* 21:1081-1085 (1993), Hoganboom *et al.*, *Immunol. Reviews* 130:43-68 (1992), Chiswell and McCafferty, *TIBTECH* 10:80-84 (1992), and U.S. Pat. No. 5,733,743.

[0126] Methods for mutating antibodies, CDRs or binding domains can be used to optimize their affinity, selectivity, binding strength and/or other desirable properties. A mutant binding domain refers to an amino acid sequence variant of a selected binding domain (*e.g.*, a CDR). In general, one or more of the amino acid residues in the mutant binding domain is different from what is present in the reference binding domain. Such mutant antibodies necessarily have less than 100% sequence identity or similarity with the reference amino acid sequence. In general, mutant binding

domains have at least 75% amino acid sequence identity or similarity with the amino acid sequence of the reference binding domain. Preferably, mutant binding domains have at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% amino acid sequence identity or similarity with the amino acid sequence of the reference binding domain.

[0127] For example, affinity maturation using phage display can be utilized as one method for generating mutant binding domains. Affinity maturation using phage display refers to a process described in Lowman *et al.*, *Biochemistry* 30(45):10832-10838 (1991), see also Hawkins *et al.*, *J. Mol Biol.* 254: 889-896 (1992). While not strictly limited to the following description, this process can be described briefly as involving mutation of several binding domains or antibody hypervariable regions at a number of different sites with the goal of generating all possible amino acid substitutions at each site. The binding domain mutants thus generated are displayed in a monovalent fashion from filamentous phage particles as fusion proteins. Fusions are generally made to the gene III product of M13. The phage expressing the various mutants can be cycled through several rounds of selection for the trait of interest, *e.g.*, binding affinity or selectivity. The mutants of interest are isolated and sequenced. Such methods are described in more detail in U.S. Pat. Nos. 5,750,373, 6,290,957 and Cunningham *et al.*, *EMBO J.* 13(11):2508-2515 (1994).

[0128] Therefore, in one embodiment, the presently disclosed subject matter provides methods of manipulating binding entity or antibody polypeptides or the nucleic acids encoding them to generate binding entities, antibodies and antibody fragments with improved binding properties that recognize selected targets. Such methods of mutating portions of an existing binding entity or antibody involve fusing a nucleic acid encoding a polypeptide that encodes a binding domain to a nucleic acid encoding a phage coat protein to generate a recombinant nucleic acid encoding a fusion protein, mutating the recombinant nucleic acid encoding the fusion protein to generate a mutant nucleic acid encoding a mutant fusion protein, expressing the mutant fusion protein on the surface of a phage, and selecting phage that bind to a target.

[0129] Accordingly, the presently disclosed subject matter provides antibodies, antibody fragments, and binding entity polypeptides that can recognize and bind to

selected target molecules. Thus, methods for manipulating those antibodies, antibody fragments, and binding entity polypeptides can be used to optimize their binding properties or other desirable properties (e.g., stability, size, ease of use).

[0130] In one embodiment, the biosensor molecule includes an HIV-1 neutralizing antibody Fab fragment (X5), which binds to HIV envelope protein gp120 after forming a complex with the host cell receptor CD4. The preparation of such biosensors is described in U.S. Patent Application Publication No. 2006/0029946 to Hahn, which is incorporated herein by reference in its entirety.

[0131] In one embodiment, a fragment of WASP, which binds only to the activated state of Cdc42, can be used to detect Cdc42 activation, as also described in U.S. Patent Application Publication No. 2006/0029946 to Hahn. Such biosensors can be used to detect and/or monitor the level of Cdc42 activity *in vitro*, in cell lysates, and in living cells. In another embodiment, a selective ligand for the DNA methyltransferase G9a can be used to detect the enzyme. In a further embodiment, the compound trifluoperazine, which binds selectively to the active conformation of calmodulin, can be used to detect calmodulin activity.

[0132] As provided hereinabove, in some embodiments, the presently disclosed biosensors can be linked to a target molecule. In embodiments wherein the presently disclosed biosensors comprise a target molecule of interest, the dye can be linked to the target molecule at a position that results in a change in a signal from the dye upon changes to the target conformation, ligand binding to the target, protein-protein interactions with the target, phosphorylation of the target, or posttranslational modification of the target. The presently disclosed compounds of Formula I or Formula II can be linked to the target molecule in biosensors using the techniques of dye linkage known in the art and described in U.S. Patent Application Publication No. 2006/0029946 to Hahn. Suitable positions for linkage to provide, for example, improved signal intensity or minimal perturbation of normal biological activity can be identified using screening techniques also described in U.S. Patent Application Publication No. 2006/0029946 to Hahn.

[0133] In some embodiments, the presently disclosed dyes can be attached to a protein of interest, wherein an activity, location, and/or conformational change in the protein can result in a change in the fluorescence of the dye. In some embodiments,

the protein of interest can be subject to a changed phosphorylation state and/or protein-ligand interaction, where the ligand can be a small molecule or a second protein. Compounds of Formula I or Formula II can be covalently attached to the protein of interest to provide a signal associated with the phosphorylation state and/or protein-protein interaction. This type of detection is in contrast to previously described detections, *see, e.g.,* Hahn *et al.*, *J. Am. Chem. Soc.* 125:4132-4145 (2003), wherein the biosensors detect a conformational change in the target protein induced by the action of a third element. In the presently disclosed methods, detection can be extended to conformational changes induced by phosphorylation and does not require induction of a conformational change in the protein of interest. One advantage of the presently disclosed methods is that proteins within multiprotein complexes can be monitored in situations where other types of biosensors, for example, those requiring a domain to find the target protein, would be blocked.

[0134] Further, the presently disclosed biosensors can allow the detection from a single dye, *e.g.*, without the use of two fluorophores, as in FRET, to provide direct excitation and a brighter signal. In some embodiments, however, two or more of the presently disclosed compounds of Formula I or Formula II having different emissions wavelengths can be used to enable imaging of multiple protein activities in the same cell simultaneously. Thus, in some embodiments, it can be desirable to provide more than one dye molecule on the binding domain of the biosensor to facilitate monitoring of a binding event.

[0135] The presently disclosed subject matter provides methods to examine protein activity, structure, or protein-protein interactions. The presently disclosed methods and biosensors can report protein localization, protein activation and/or report specific aspects of changing protein structure or interaction. The biosensors can be used in living cells and/or for homogeneous assays. For example, the activity of a protein in a complex mixture, such as cell lysate, can be determined by adding and detecting the biosensors without additional steps, such as wash steps, and the like.

[0136] In some embodiments, the presently disclosed subject matter provides a biosensor comprising one or more binding domains having a specific affinity for one or more target molecules in a specific state of phosphorylation. For example, the binding domain of the biosensor can have a specific affinity for a binding site

produced upon a protein when that protein binds to a guanosine triphosphate, thereby activating the signal transduction protein. Optionally, the biosensor can have a significantly lower affinity for the inactive form of the protein (e.g., the protein bound to guanosine diphosphate instead of guanosine triphosphate). Thus, for example, a dye of the biosensor, e.g., a compound of Formula I or Formula II, can occupy a position on the binding domain near the binding site such that, upon binding to its target, the binding environment alters the fluorescent signal from the dye. Further, the position can provide a detectable signal change in the dye without significantly inhibiting the biological activity of the targeted protein. Detection of the activated target in this embodiment does not require conformational changes in the target or the presence of third molecules interacting with the target.

[0137] The presently disclosed biosensors also can detect protein-protein interactions. Such interactions can be high affinity interactions, such as, for example, interactions between antigens and antibodies, or they can be lower affinity interactions, such as, for example, interactions between enzymes and substrates, signal cascade members, or members of protein complexes. In a protein-protein interaction between two binding members, one binding member can be considered the "target" and the other binding member the biosensor "probe." While the distinction between target and probe can be arbitrary, because the two interact with each other these terms are used to facilitate discussion of protein interaction. Generally, as used herein, the target can be the molecule to be detected and the probe can be a binding member introduced to interrogate a sample for the presence of the target in a state or form of interest. In a representative protein-protein biosensor, one member of the binding pair can have a binding domain complimentary to a binding site of the second binding pair member. One or both of the protein-protein interaction pair members can include an attached compound of Formula I or Formula II. The protein-protein interaction pair members can be full-length, naturally occurring proteins, synthetic analogs of naturally occurring proteins, recombinant analogs of naturally occurring proteins, or fragments thereof.

[0138] The presently disclosed biosensors capable of having protein-protein interactions can have compounds of Formula I or Formula II at one or more positions on one or both binding members such that the dyes are positioned between the

binding members during binding. In such embodiments, binding can cause a detectable change in the signal from the dye without significantly inhibiting the binding interaction between the binding members. Suitable dye positions can be identified by screening alternative positions for improved signal and/or binding function in the presently disclosed biosensor systems.

[0139] In some embodiments, the biosensor can have specific affinity for a target molecule in a specific conformation, bound to a specific ligand, or with a specific posttranslational modification other than phosphorylation.

[0140] In some embodiments, the presently disclosed subject matter provides a modular biosensor. Modular biosensors are sensors of unified design that allow certain components to be changed, e.g., to change the target specificity and/or signal character. Modular biosensors typically provide for convenient alteration of binding domain specificity. For example, an affinity molecule can be expressed from a recombinant expression construct, such as an expression vector. The genetic construct can include unique endonuclease sites bracketing the region encoding the binding domain so that alternate binding domains can be readily inserted into the construct for expression as part of a biosensor. The modular binding domains can be selected, e.g., from a library of binding domains. Optionally, the modules can be encoded scFv domains. The modular systems can include affinity molecules with alternate binding domain sites, alternate dye linkage sites, alternate dye linkage reactive molecules, alternate linked targets, alternate linkers, and/or the like. Alternately, the binding domain can be of broad specificity, and the target domain (i.e., different kinase consensus sequence peptides) can be switched in the modular design.

III. Detecting Presence and/or Activity of Target Molecules in Living Cells

[0141] Cell behavior can be regulated through transient activation of protein activities at specific subcellular locations. The ability to study translocation of proteins has been enhanced by advances in the microscopy of fluorescent protein analogues within living cells. In many cases, however, localized protein activities are controlled not by translocating proteins to the site of action, but by localized

activation of a small portion of the protein pool. See Hahn and Touchkine, *Curr. Opin. Cell Biol.* 14:167-172 (2002); Wouters *et al.*, *Trends Cell Biol.* 11:203-211 (2001). Such behaviors typically are not apparent when studying protein translocations or when using *in vitro* biochemical approaches. Further, the outcome of signaling protein activation can depend on subtle variations in activation kinetics that are not discernible in the population averages generated by biochemical techniques. For precise quantification of rapid activation kinetics and of the level of protein activation, it also is necessary to measure protein activity in living cells. See Wouters *et al.*, *Trends Cell Biol.* 11:203-211 (2001); Williams *et al.*, *Nature* 318:558-561 (1985); Berridge, *J. Biol. Chem.* 265:9583-9586 (1990). Accordingly, a need exists for methods for monitoring protein movements in living cells.

[0142] The presently disclosed dyes exhibit many properties that make them suitable for detecting target molecules and their interactions in living cells. The dyes are, for example, bright, with long wavelengths outside of cellular autofluorescence background frequencies and that are less damaging to cells. Addition or deletion of parts of the aromatic system, or substituent groups thereon, can shift excitation and/or emission wavelengths of the dyes so that more than one event can be monitored in a cell at the same time. The dyes have increased cell membrane permeability with minimal staining of intracellular membranes.

[0143] The presently disclosed dyes can be detected in cells by observing changes in intensity, a change in the shape or maxima of the excitation or emission peak, and/or dye lifetime, to permit ratio imaging and other techniques that can eliminate effects of uneven illumination, cell thickness, and the like.

[0144] In some embodiments, the presently disclosed subject matter provides a method of detecting an activity or a location of one or more target molecule within a cell, the method comprising: (a) providing the biosensor of the present invention; (b) contacting the biosensor with a cell suspected of containing one or more target molecules to bind the one or more target molecules, if present, to the binding member; (c) irradiating the cell suspected of containing one or more target molecules with electromagnetic radiation to induce the compound of Formula I or Formula II to fluoresce; and (d) detecting one or more of: (i) a fluorescence property of the compound of Formula I or Formula II; (ii) a change in a fluorescence property of the

compound of Formula I or Formula II; (iii) a location of a fluorescence of the compound of Formula I or Formula II; and (iv) combinations thereof, to determine the an activity or location of one or more target molecules in the cell providing a biosensor comprising a dye of Formula (I) or Formula (II); (b) contacting the biosensor with a cell suspected of containing one or more target molecules to bind the one or more target molecules, if present, to the binding member; (c) irradiating the cell suspected of containing one or more target molecules with electromagnetic radiation to induce the dye of Formula (I) or Formula (II) to fluoresce; and (d) detecting one or more of: (i) a fluorescence property of the dye of Formula (I) or Formula (II); (ii) a change in a fluorescence property of the dye of Formula (I) or Formula (II); (iii) a location of a fluorescence of the dye of Formula (I) or Formula (II); and (iv) combinations thereof, to determine the an activity or location of one or more target molecules in the cell.

[0145] The presently disclosed methods and biosensors can be useful for detecting target molecules in or on the surface of various types of biological cells, including, mammalian, bacterial, fungal, yeast, insect, and plant cells. In some embodiments, target molecules can be detected in freshly isolated cells from mammals (*e.g.*, humans), insects, fungal, or bacterial cells. For example, blood cells, such as B cells, T cells, monocytes, and neutrophils, and the like, can be probed with the presently disclosed biosensors. In other embodiments, stably maintained cell lines such as CHO, HEK-293, L-cells, 3T3 cells, COS, or THP-1 cells can be investigated using the presently disclosed methods.

[0146] Useful information can be obtained from any type of cell using the presently disclosed biosensors and methods. For example, mammalian cells, such as human cells or animal cells, that naturally or recombinantly express human proteins can be evaluated to identify potential human therapeutics, observed for interactions between biomolecules, and/or studied for the effects of ligands, drugs, and other molecules on mammalian and human systems. In another example, bacterial or fungal cells can be used to screen for potential antibiotic or anti-fungal agents.

[0147] In some embodiments, well characterized cell lines known to provide predictive models of human cell functions can be used to obtain results correlated with human systems in pharmaceutical and medical research. Exemplary cell lines

useful in such research include, for example, COS cells, CHO cells, HEK-293 cells, RBL-1, Jurkat, U937, and YB-1 cells.

[0148] The cells to be monitored can be provided in either immobilized form or as a suspension culture. Immobilized cells, such as, *e.g.*, cell lawns, tissue slices, or libraries, can be monitored, *e.g.*, by microscopy, scanners, or with imaging systems. The immobilized cells can be monitored live or fixed for detection of target molecules in killed cells.

[0149] In many embodiments, cells subject to target molecule detection with the presently disclosed biosensors are in suspension. Suspended cells can be cells from suspension cell cultures or cells liberated from tissues or lawns. Suspended cells are particularly well suited to handling and monitoring in fluidic systems, such as cell sorters, cell counters, and microfluidic systems. Cell suspensions can be provided at cell densities appropriate to the handling system and detection method that is being employed. Determination of optimal cell densities is routine for one of ordinary skill in the art. In the case of flow-through embodiments, cell densities of monitored suspensions generally range from about 1 cell/nL to about 30 cells/nL in, *e.g.*, a reaction vessel or detection channel. In the case of test tube or multiwell plate based reactions, cell densities typically range from about 1,000 cells/mm² to about 100,000 cells/mm². Of course, these ranges can vary depending upon, *e.g.*, the cell types used, the type of biosensor employed, the type of interaction to be studied, the relative adherence of the cells to the vessel surfaces, as well as each other factors.

[0150] In sum, the presently disclosed subject matter provides biosensors that are capable of studying proteins in living cells. The dyes can be directly attached to proteins, enabling conformational changes to be followed *in vivo* for proteins incorporated in large molecular machines. For *in vitro* applications, the dyes exhibit an increased photostability and provide a substantially brighter signal than current solvent-sensitive fluorophores. Thus, they can enhance sensitivity for studying high-affinity binding interactions, small amounts of protein in high-throughput assays, or protein changes that produce only small effects on other dyes.

A. *Administration of Biosensors Comprising Compounds of Formula I or Formula II*

[0151] The presently disclosed biosensors can be used *in vitro* and/or *in vivo* to detect target molecules of interest. In some embodiments, the biosensors can be added to test samples in a homogenous assay and do not require addition of multiple reagents and/or wash steps before detection of the target molecule.

[0152] The presently disclosed biosensors can be contacted with target molecules *in vitro* by addition to a test sample containing the target molecules. Test samples for *in vitro* assays include molecular libraries, cell lysates, analyte eluates from chromatographic columns, and the like. The *in vitro* assay can take place in a chamber, such as, e.g., a well of a multiwell plate, a test tube, an Eppendorf tube, a spectrophotometer cell, conduit of an analytical system, channels of a microfluidic system, and the like. In an exemplary *in vitro* assay, an enzyme protein of interest is coated to the bottom of 96-well plates also containing solutions representing a library of possible enzyme substrates. A presently disclosed biosensor having a specific affinity for enzyme-substrate complex is added to each well. A multiwell scanning fluorometer can be used to observe each well for fluorescence. Wells containing enzyme substrate can be identified as those in which fluorescent emissions at the wavelength of the biosensor dye. That is, in this example, the binding domain of the biosensor only binds to enzyme acting on substrate; the binding placing the dye into a binding pocket environment that significantly changes the emissions intensity of the dye.

[0153] In embodiments wherein the presently disclosed biosensors are administered to living cells, binding can take place with targets on the cell surface, or the biosensor is transferred into the cell to make contact with an intracellular target molecule. In some embodiments, the biosensor can penetrate a cell suspected of containing a selected target passively by mere exposure of the cell to a medium containing the biosensor. In other embodiments, the biosensor is actively transferred into the cell by mechanisms known in the art, such as, e.g., poration, injection, transduction along with transfer peptides, and the like.

[0154] In some embodiments, a translocation functionality can be incorporated on the biosensor to facilitate the translocation or internalization of that biosensor from

the outside to inside the cell. As used herein, the term "translocation functionality" refers to a chemical compound, group or moiety that increases the cell's ability to internalize another compound or material, for example, a biosensor. Examples of such translocation functionalities include peptide recognition/transport sequences, liposomal compositions, or the like. Alternative translocation methods and compositions also can be used in accordance with the presently disclosed subject matter to induce uptake of the second component, including, *e.g.*, electroporation, cell permeating compositions containing, *e.g.*, PEG, porins, saponins, streptolysin, or the like.

[0155] Techniques useful for promoting uptake of biosensors include optoporation, for example, as described in Schneckenburger *et al.*, *J. Biomed. Opt.* 7:410-6 (2002); or Soughayer *et al.*, *Anal. Chem.* 72:1342-7 (2000). A variety of transduction peptides also are useful for promoting uptake of biosensors, including those described in Zelphati *et al.*, *J. Biol. Chem.* 276:35103-10 (2001); Yang *et al.*, *FEBS Lett.* 532:36-44 (2002); and Torchilin *et al.*, *Proc. Natl. Acad. Sci. USA* 100:1972-7 (2003).

[0156] Additional techniques such as electroporation also can be used. Examples of electroporation procedures are provided in Glogauer and McCulloch, *Exp. Cell Res.* 200:227-34 (1992); Teruel and Meyer, *Science* 295:1910-2 (2002); and Teruel *et al.*, *J. Neurosci. Methods* 93:37-48 (1999).

[0157] Another procedure for introducing molecules such as biosensors into cells is the osmotic shock procedure. Examples of osmotic shock procedures include those described in Okada and Rechsteiner, *Cell* 29:33-41 (1982); and Park *et al.*, *J. Cell Physiol.* 135:443-50 (1988).

[0158] One of skill in the art also can use bead/syringe loading to introduce the presently disclosed biosensors into cells. Bead/syringe loading procedures are described in McNeil *et al.*, *J. Cell Biol.* 98:1556-1564 (1984); and McNeil and Warder, *J. Cell Sci.* 88:669-678 (1987).

[0159] Nucleic acids encoding binding domains as disclosed herein can be introduced into cells in expression plasmids, *e.g.*, by transduction or other forms of transformation. Once inside the living cells, the binding domain can be translated from the nucleic acid to a functional peptide. Compounds of Formula I or Formula II

can enter the cell, *e.g.*, by injection or diffusion, to become linked to the expressed binding domain to generate a biosensor *in situ*. The presently disclosed dyes also can be introduced into cells by other methods known in the art, including, but not limited to viral-based delivery systems, including viral-based nanoparticles, and lipid delivery vehicles, including liposomes, and the like.

B. Detection of Target-Biosensor Binding Reactions

[0160] A wide variety of binding reactions can be detected and monitored using the presently disclosed biosensors, for example, protein-protein interactions, receptor-ligand interactions, nucleic acid interactions, protein-nucleic acid interactions, and the like. Detection of a target molecule can provide identification of the target in a specified state, quantification of the target, and/or localization of the target. Multiple measurements can allow determination of kinetics. The ability to monitor multiple targets can permit monitoring of the balance between different signaling activities. In the intracellular environment, many of these reaction types are involved in the multiplicity of steps of signal transduction within cells. For example, activation of a particular cellular event is often triggered by the interaction between a cell surface receptor and its ligand. The signal from the receptor is often transmitted along via the binding of enzymes to other proteins, for example, kinases, which then pass the signal on through the cell until the ultimate cell system response is achieved. In many cases, the signal or ultimate response can be detected using the presently disclosed biosensors. For example, signal transduction often involves phosphorylation of system molecules that can be detected directly with the phosphate involved in the binding site, or indirectly through conformational changes induced by the phosphorylation.

[0161] In one embodiment, the presently disclosed subject matter provides methods for identifying the activation status of endogenous proteins in living cells. The presently disclosed biosensors can permit identification, quantification, and resolution of the spatial, temporal and compartmental regulation of receptor phosphorylation and activation during various processes, for example, endocytosis. In another embodiment, the presently disclosed biosensors and methods can permit observation of epidermal growth factor receptor (EGFR) effects on the development

and progression of breast cancer. In a further embodiment, complex formation between HIV gp120 and CD4 cell receptors can be monitored.

[0162] Accordingly, in some embodiments, the presently disclosed subject matter provides a method of detecting an interaction between an endogenous target molecule and a cellular entity, the method comprising: (a) providing a cell comprising an endogenous target molecule; (b) providing the biosensor of the invention; (c) observing a background fluorescence signal from the biosensor; (d) contacting the biosensor with the cell; and (e) detecting a change in fluorescence from the biosensor to indicate an interaction between the target molecule and the cellular entity.

[0163] In some embodiments, the cellular entity is selected from the group consisting of a cellular nucleic acid, a protein, a peptide, an enzyme, a receptor, a cytokine, a cytoskeleton, and a signal transduction protein. In some embodiments, the binding member of the probe binds the target molecule at a phosphorylation site. In some embodiments, the binding member has a specific affinity for a conformation, a ligand interaction, or posttranslational modification of the target molecule. In some embodiments, the detecting a change in fluorescence comprises quantifying a protein amount, locating a protein, detecting a conformational change in the target molecule, detecting activation of the target molecule, or detecting phosphorylation of the target molecule.

[0164] In some embodiments, the compound of Formula I or Formula II is linked to the binding member at a site that does not substantially interfere with binding between the probe and the target biomolecule. Further, in some embodiments, the site is selected by an examination of a crystal structure for the binding member or the target molecule. In some embodiments, the method further comprises introducing the probe into the cell by using a technique selected from the group consisting of electroporation, transduction, microporation, and microinjection.

C. Methods for Detecting the Presence or Amount of a Target Molecule

[0165] In accordance with the presently disclosed subject matter, binding interactions can occur between a biosensor and one or more target molecules or components of the cell. A 'target molecule of interest' is a molecule that is known by one of skill in the art and is selected for interaction with a presently disclosed

biosensor. A target molecule often comprises an endogenous unlabeled and/or untagged component of a test solution or cell. Endogenous components can be, *e.g.*, expressed by the cell naturally, or present as a result of introduction of an appropriate genetic construct within the cell. For example, nucleic acid or protein target molecules can be expressed in the cell, either naturally (*e.g.*, constitutively) or by induction of an appropriate genetic construct introduced into the cell line.

[0166] In some embodiments, the presently disclosed subject matter provides a method for determining the presence or amount of one or more target molecules in a sample, the method comprising: (a) providing the biosensor of the invention; (b) contacting the biosensor with a sample suspected of containing one or more target molecules to bind the one or more target molecules, if present, to the binding member; (c) irradiating the sample suspected of containing one or more target molecules with electromagnetic radiation to induce the compound of Formula I or Formula II to fluoresce; and (d) detecting a fluorescence property of the compound of Formula I or Formula II to determine the presence or amount of one or more target molecules in the sample.

[0167] In some embodiments, the method further comprises: (a) measuring a fluorescence intensity at a first emission wavelength before contacting the biosensor with a sample suspected of containing one or more target molecules; (b) measuring a fluorescence intensity at a second emission wavelength after contacting the biosensor with a sample suspected of containing one or more target molecules; and (c) determining a ratio of the second emission wavelength to the first emission wavelength to determine the presence or amount of one or more target molecules in the sample.

[0168] In some embodiments, the method further comprises continuously: (a) contacting the biosensor with the sample suspected of containing one or more target molecules; (b) irradiating the sample with electromagnetic radiation; and (c) detecting the fluorescence property of the compound of Formula I or Formula II.

[0169] As used herein, the term "sample" includes any liquid or fluid sample, including a sample derived from a biological source, such as a physiological fluid, including whole blood or whole blood components, such as red blood cells, white blood cells, platelets, serum and plasma; ascites; urine; saliva; sweat; milk; synovial

fluid; peritoneal fluid; amniotic fluid; perocerebrospinal fluid; lymph fluid; lung embolism; cerebrospinal fluid; pericardial fluid; cervicovaginal samples; tissue extracts; cell extracts; and other constituents of the body that are suspected of containing the analyte of interest. In addition to physiological fluids, other liquid samples, such as water, food products and the like, for the performance of environmental or food production assays are suitable for use with the presently disclosed subject matter. A solid material suspected of containing the analyte also can be used as the test sample. In some instances it might be beneficial to modify a solid test sample to form a liquid medium or to release the analyte.

[0170] In some embodiments, the sample can be pre-treated prior to use, such as preparing plasma from blood, diluting viscous fluids, or the like. Such methods of treatment can involve filtration, distillation, concentration, inactivation of interfering compounds, and the addition of reagents.

[0171] The sample can be any sample obtained from a subject. The term "subject" refers to an organism, tissue, or cell from which a sample can be obtained. A subject can include a human subject for medical purposes, such as diagnosis and/or treatment of a condition or disease, or an animal subject for medical, veterinary purposes, or developmental purposes. A subject also can include sample material from tissue culture, cell culture, organ replication, stem cell production and the like. Suitable animal subjects include mammals and avians. The term "avian" as used herein includes, but is not limited to, chickens, ducks, geese, quail, turkeys, and pheasants. The term "mammal" as used herein includes, but is not limited to, primates, *e.g.*, humans, monkeys, apes, and the like; bovines, *e.g.*, cattle, oxen, and the like; ovines, *e.g.*, sheep and the like; caprines, *e.g.*, goats and the like; porcines, *e.g.*, pigs, hogs, and the like; equines, *e.g.*, horses, donkeys, zebras, and the like; felines, including wild and domestic cats; canines, including dogs; lagomorphs, including rabbits, hares, and the like; and rodents, including mice, rats, and the like. Preferably, the subject is a mammal or a mammalian cell. More preferably, the subject is a human or a human cell. Human subjects include, but are not limited to, fetal, neonatal, infant, juvenile, and adult subjects. Further, a "subject" can include a patient afflicted with or suspected of being afflicted with a condition or disease. Thus, the terms "subject" and "patient" are used interchangeably herein. A subject

also can refer to cells or collections of cells in laboratory or bioprocessing culture in tests for viability, differentiation, marker production, expression, and the like.

IV. Biosensor Kits Comprising Compounds of Formula I or Formula II

[0172] The presently disclosed subject matter further provides a packaged composition such as a kit or other container for detecting, monitoring or otherwise observing a target molecule. The kit or container can hold a biosensor comprising a compound of Formula I or Formula II and instructions for using the biosensor for detecting, monitoring or otherwise observing a target molecule. The biosensor includes at least one binding member and a compound of Formula I or Formula II. In one embodiment, the kit comprises a container containing a biosensor comprising scFv binding domain and a dye. Alternatively, the kit or container holds a compound of Formula I or Formula II and instructions for using the dye. In some embodiments, kits containing dyes can contain instructions for attaching a dye to a binding domain selected by one of skill in the art.

[0173] The presently disclosed kits also can comprise containers with solutions or tools useful for manipulating or using the presently disclosed dyes or biosensors. Such tools include buffers, reaction tubes, reagents for coupling compounds of Formula I or Formula II to selected binding domains and the like. In one embodiment, the kit can contain a solution of solvents and/or buffers to facilitate coupling of a compound of Formula I or Formula II to a selected binding domain and/or a solution of mercaptoethanol for quenching the dye-binding domain conjugation reaction. The kit also can include a container of buffer at roughly neutral pH (e.g., sodium phosphate buffer, pH 7.5).

V. Chemical Definitions

[0174] While the following terms are believed to be well understood by one of ordinary skill in the art, the following definitions are set forth to facilitate explanation of the presently disclosed subject matter. Unless otherwise defined, all technical and

scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this presently described subject matter belongs.

[0175] Throughout the specification and claims, a given chemical formula or name shall encompass all optical and stereoisomers, as well as racemic mixtures where such isomers and mixtures exist.

[0176] When the term “independently selected” is used, the substituents being referred to (e.g., R groups, such as groups R_1 , R_2 , and the like, or groups X and Y), can be identical or different. For example, both R_1 and R_2 can be substituted alkyls, or R_1 can be hydrogen and R_2 can be a substituted alkyl, and the like.

[0177] A named “R” or “X” group will generally have the structure that is recognized in the art as corresponding to a group having that name, unless specified otherwise herein. For the purposes of illustration, certain representative “R” and “X” groups as set forth above are defined below. These definitions are intended to supplement and illustrate, not preclude, the definitions that would be apparent to one of ordinary skill in the art upon review of the present disclosure.

[0178] As used herein the term “electron withdrawing group” refers to a atom or a group of atoms that draws electrons away from a reaction center. Examples include, without limitation, carbonyl, cyano, halogen, nitro, sulfonyl, and trifluoromethyl groups.

[0179] As used herein the term “leaving group” refers to a charged or uncharged atom or group of atoms that departs during a substitution or displacement reaction. Examples include, without limitation, halogens, such as chloro, bromo, and iodo; alkoxide, triflate, tosylate, mesylate, fluorosulfonate and optionally substituted N-hydroxysuccinimide.

[0180] As used herein the term “conjugatable side chain” refers to side chain comprising, consisting essentially of, or consisting of a linker and a reactive group at the end of the linker that can be used to conjugate a dye compound to a binding member.

[0181] As used herein the term “alkyl” refers to C_{1-20} inclusive, linear (*i.e.*, “straight-chain”), branched, or cyclic, saturated or at least partially and in some cases fully unsaturated (*i.e.*, alkenyl and alkynyl) hydrocarbon chains, including for example, methyl, ethyl, propyl, isopropyl, butyl, isobutyl, tert-butyl, pentyl, hexyl,

octyl, ethenyl, propenyl, butenyl, pentenyl, hexenyl, octenyl, butadienyl, propynyl, butynyl, pentynyl, hexynyl, heptynyl, and allenyl groups. "Branched" refers to an alkyl group in which a lower alkyl group, such as methyl, ethyl or propyl, is attached to a linear alkyl chain. "Lower alkyl" refers to an alkyl group having 1 to about 6 carbon atoms (*i.e.*, a C₁₋₆ alkyl), *e.g.*, 1, 2, 3, 4, 5, or 6 carbon atoms. "Higher alkyl" refers to an alkyl group having about 8 to about 20 carbon atoms, *e.g.*, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 carbon atoms. In certain embodiments, "alkyl" refers, in particular, to C₁₋₆ straight-chain alkyls. In other embodiments, "alkyl" refers, in particular, to C₁₋₆ branched-chain alkyls.

[0182] Alkyl groups can optionally be substituted (a "substituted alkyl") with one or more alkyl group substituents, which can be the same or different. The term "alkyl group substituent" includes but is not limited to alkyl, substituted alkyl, halo, arylamino, acyl, hydroxyl, aryloxy, alkoxyl, alkylthio, arylthio, aralkyloxy, aralkylthio, carboxyl, alkoxycarbonyl, oxo, and cycloalkyl. There can be optionally inserted along the alkyl chain one or more oxygen, sulfur or substituted or unsubstituted nitrogen atoms, wherein the nitrogen substituent is hydrogen, lower alkyl (also referred to herein as "alkylaminoalkyl"), or aryl.

[0183] Thus, as used herein, the term "substituted alkyl" includes alkyl groups, as defined herein, in which one or more atoms or functional groups of the alkyl group are replaced with another atom or functional group, including for example, alkyl, substituted alkyl, halogen, aryl, substituted aryl, alkoxyl, hydroxyl, nitro, amino, alkylamino, dialkylamino, sulfate, and mercapto.

[0184] "Cyclic" and "cycloalkyl" refer to a non-aromatic mono- or multicyclic ring system of about 3 to about 10 carbon atoms, *e.g.*, 3, 4, 5, 6, 7, 8, 9, or 10 carbon atoms. The cycloalkyl group can be optionally partially unsaturated. The cycloalkyl group also can be optionally substituted with an alkyl group substituent as defined herein, oxo, and/or alkylene. There can be optionally inserted along the cyclic alkyl chain one or more oxygen, sulfur or substituted or unsubstituted nitrogen atoms, wherein the nitrogen substituent is hydrogen, alkyl, substituted alkyl, aryl, or substituted aryl, thus providing a heterocyclic group. Representative monocyclic cycloalkyl rings include cyclopentyl, cyclohexyl, and cycloheptyl. Multicyclic

cycloalkyl rings include adamantyl, octahydronaphthyl, decalin, camphor, camphane, and noradamantyl.

[0185] The term “cycloalkylalkyl,” as used herein, refers to a cycloalkyl group as defined hereinabove, which is attached to the parent molecular moiety through an alkyl group, also as defined above. Examples of cycloalkylalkyl groups include cyclopropylmethyl and cyclopentylethyl.

[0186] The terms “cycloheteroalkyl” or “heterocycloalkyl” refer to a non-aromatic ring system, such as a 3- to 7-member substituted or unsubstituted cycloalkyl ring system, including one or more heteroatoms, which can be the same or different, and are selected from the group consisting of N, O, and S, and optionally can include one or more double bonds. The cycloheteroalkyl ring can be optionally fused to or otherwise attached to other cycloheteroalkyl rings and/or non-aromatic hydrocarbon rings. Representative cycloheteroalkyl ring systems include, but are not limited to pyrrolidinyl, pyrrolinyl, imidazolidinyl, imidazolynyl, pyrazolidinyl, pyrazolynyl, piperidyl, piperazinyl, indolynyl, quinuclidinyl, morpholynyl, thiomorpholynyl, thiadiazinanyl, tetrahydrofuranyl, and the like.

[0187] The term “alkenyl” as used herein refers to a straight or branched hydrocarbon of a designed number of carbon atoms containing at least one carbon-carbon double bond. Examples of “alkenyl” include vinyl, allyl, 2-methyl-3-heptene, and the like.

[0188] The term “cycloalkenyl” as used herein refers to a cyclic hydrocarbon containing at least one carbon-carbon double bond. Examples of cycloalkenyl groups include cyclopropenyl, cyclobutenyl, cyclopentenyl, cyclopentadiene, cyclohexenyl, 1,3-cyclohexadiene, cycloheptenyl, cycloheptatrienyl, and cyclooctenyl.

[0189] The term “alkynyl” as used herein refers to a straight or branched hydrocarbon of a designed number of carbon atoms containing at least one carbon-carbon triple bond. Examples of “alkynyl” include propargyl, propyne, and 3-hexyne.

[0190] “Alkylene” refers to a straight or branched bivalent aliphatic hydrocarbon group having from 1 to about 20 carbon atoms, *e.g.*, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 carbon atoms. The alkylene group can be straight, branched or cyclic. The alkylene group also can be optionally unsaturated and/or substituted with one or more “alkyl group substituents.” There can be optionally

inserted along the alkylene group one or more oxygen, sulfur or substituted or unsubstituted nitrogen atoms (also referred to herein as "alkylaminoalkyl"), wherein the nitrogen substituent is alkyl as previously described. Exemplary alkylene groups include methylene ($-\text{CH}_2-$); ethylene ($-\text{CH}_2-\text{CH}_2-$); propylene ($-(\text{CH}_2)_3-$); cyclohexylene ($-\text{C}_6\text{H}_{10}-$); $-\text{CH}=\text{CH}-\text{CH}=\text{CH}-$; $-\text{CH}=\text{CH}-\text{CH}_2-$; $-(\text{CH}_2)_q-\text{N}(\text{R})-(\text{CH}_2)_r-$, wherein each of q and r is independently an integer from 0 to about 20, *e.g.*, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20, and R is hydrogen or lower alkyl; methylenedioxy ($-\text{O}-\text{CH}_2-\text{O}-$); and ethylenedioxy ($-\text{O}-(\text{CH}_2)_2-\text{O}-$). An alkylene group can have about 2 to about 3 carbon atoms and can further have 6-20 carbons.

[0191] The term "aryl" is used herein to refer to an aromatic substituent that can be a single aromatic ring, or multiple aromatic rings that are fused together, linked covalently, or linked to a common group, such as, but not limited to, a methylene or ethylene moiety. The common linking group also can be a carbonyl, as in benzophenone, or oxygen, as in diphenylether, or nitrogen, as in diphenylamine. The term "aryl" specifically encompasses heterocyclic aromatic compounds. The aromatic ring(s) can comprise phenyl, naphthyl, biphenyl, diphenylether, diphenylamine and benzophenone, among others. In particular embodiments, the term "aryl" means a cyclic aromatic comprising about 5 to about 10 carbon atoms, *e.g.*, 5, 6, 7, 8, 9, or 10 carbon atoms, and including 5- and 6-membered hydrocarbon and heterocyclic aromatic rings.

[0192] The aryl group can be optionally substituted (a "substituted aryl") with one or more aryl group substituents, which can be the same or different, wherein "aryl group substituent" includes alkyl, substituted alkyl, aryl, substituted aryl, aralkyl, hydroxyl, alkoxyl, aryloxy, aralkyloxy, carboxyl, acyl, halo, nitro, alkoxycarbonyl, aryloxy carbonyl, aralkoxy carbonyl, acyloxy, acylamino, aroylamino, carbamoyl, alkylcarbamoyl, dialkylcarbamoyl, arylthio, alkylthio, alkylene, and $-\text{NR}'\text{R}''$, wherein R' and R'' can each be independently hydrogen, alkyl, substituted alkyl, aryl, substituted aryl, and aralkyl.

[0193] Thus, as used herein, the term "substituted aryl" includes aryl groups, as defined herein, in which one or more atoms or functional groups of the aryl group are replaced with another atom or functional group, including for example, alkyl,

substituted alkyl, halogen, aryl, substituted aryl, alkoxy, hydroxyl, nitro, amino, alkylamino, dialkylamino, sulfate, and mercapto.

[0194] Specific examples of aryl groups include, but are not limited to, phenyl, furan, thiophene, pyrrole, pyran, pyridine, imidazole, benzimidazole, isothiazole, isoxazole, pyrazole, pyrazine, triazine, pyrimidine, quinoline, isoquinoline, indole, carbazole, and the like.

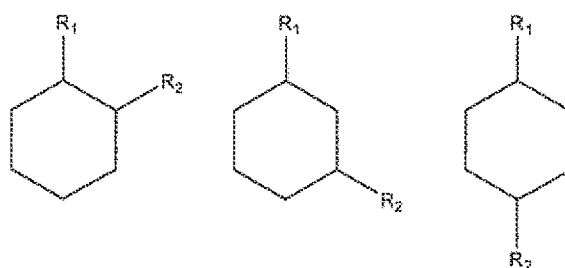
[0195] The term "heteroaryl" refers to an aromatic ring system, such as, but not limited to a 5- or 6-member ring system, including one or more heteroatoms, which can be the same or different, and are selected from the group consisting of N, O, and S. The heteroaryl ring can be fused or otherwise attached to one or more heteroaryl rings, aromatic or non-aromatic hydrocarbon rings, or heterocycloalkyl rings. Representative heteroaryl ring systems include, but are not limited to, pyridyl, pyrimidyl, pyrrolyl, pyrazolyl, azolyl, oxazolyl, isoxazolyl, oxadiazolyl, thiazolyl, isothiazolyl, imidazolyl, furanyl, thienyl, quinolinyl, isoquinolinyl, indolinyl, indolyl, benzothienyl, benzothiazolyl, benzofuranyl, benzimidazolyl, benzisoxazolyl, benzopyrazolyl, triazolyl, tetrazolyl, and the like.

[0196] A structure represented generally by the formula, wherein the ring structure can be aromatic or non-aromatic:



as used herein refers to a ring structure, for example, but not limited to a 3-carbon, a 4-carbon, a 5-carbon, a 6-carbon, and the like, aliphatic and/or aromatic cyclic compound, including a saturated ring structure, a partially saturated ring structure, and an unsaturated ring structure as defined herein, comprising a substituent R group, wherein the R group can be present or absent, and when present, one or more R groups can each be substituted on one or more available carbon atoms of the ring structure. The presence or absence of the R group and number of R groups is determined by the value of the integer n. Each R group, if more than one, is substituted on an available carbon of the ring structure rather than on another R group. For example, the structure above where n is 0 to 2 would comprise compound groups including, but not limited to:

68



and the like.

[0197] A dashed line representing a bond in a cyclic ring structure indicates that the bond can be either present or absent in the ring. That is, a dashed line representing a bond in a cyclic ring structure indicates that the ring structure is selected from the group consisting of a saturated ring structure, a partially saturated ring structure, and an unsaturated ring structure.

[0198] When a named atom of an aromatic ring or a heterocyclic aromatic ring is defined as being “absent,” the named atom is replaced by a direct bond.

[0199] As used herein, the term “acyl” refers to an organic acid group wherein the –OH of the carboxyl group has been replaced with another substituent (*i.e.*, as represented by RCO–, wherein R is an alkyl or an aryl group as defined herein). As such, the term “acyl” specifically includes arylacyl groups, such as an acetylfuran and a phenacyl group. Specific examples of acyl groups include acetyl and benzoyl.

[0200] “Alkoxyl” refers to an alkyl-O– group wherein alkyl is as previously described. The term “alkoxyl” as used herein can refer to C₁₋₂₀ inclusive, linear, branched, or cyclic, saturated or unsaturated oxo-hydrocarbon chains, including, for example, methoxyl, ethoxyl, propoxyl, isopropoxyl, butoxyl, t-butoxyl, and pentoxyl.

[0201] The term “alkoxyalkyl” as used herein refers to an alkyl-O-alkyl ether, for example, a methoxyethyl or an ethoxymethyl group.

[0202] “Aryloxyl” refers to an aryl-O– group wherein the aryl group is as previously described, including a substituted aryl. The term “aryloxyl” as used herein can refer to phenyloxyl or hexyloxyl, and alkyl, substituted alkyl, halo, or alkoxyl substituted phenyloxyl or hexyloxyl.

[0203] The term “alkyl-thio-alkyl” as used herein refers to an alkyl-S-alkyl thioether, for example, a methylthiomethyl or a methylthioethyl group.

[0204] “Aralkyl” refers to an aryl-alkyl-group wherein aryl and alkyl are as previously described, and included substituted aryl and substituted alkyl. Exemplary aralkyl groups include benzyl, phenylethyl, and naphthylmethyl.

[0205] “Aralkyloxy” refers to an aralkyl-O— group wherein the aralkyl group is as previously described. An exemplary aralkyloxy group is benzyloxy.

[0206] “Alkoxycarbonyl” refers to an alkyl-O—CO— group. Exemplary alkoxycarbonyl groups include methoxycarbonyl, ethoxycarbonyl, butyloxycarbonyl, and t-butyloxycarbonyl.

[0207] “Aryloxycarbonyl” refers to an aryl-O—CO— group. Exemplary aryloxycarbonyl groups include phenoxy- and naphthoxy-carbonyl.

“Aralkoxycarbonyl” refers to an aralkyl-O—CO— group. An exemplary aralkoxycarbonyl group is benzyloxycarbonyl.

[0208] “Carbamoyl” refers to an $\text{H}_2\text{N}-\text{CO}-$ group. “Alkylcarbamoyl” refers to a $\text{R}'\text{RN}-\text{CO}-$ group wherein one of R and R' is hydrogen and the other of R and R' is alkyl and/or substituted alkyl as previously described. “Dialkylcarbamoyl” refers to a $\text{R}'\text{RN}-\text{CO}-$ group wherein each of R and R' is independently alkyl and/or substituted alkyl as previously described.

[0209] “Acyloxy” refers to an acyl-O— group wherein acyl is as previously described.

[0210] The term “amino” refers to the $-\text{NH}_2$ group and also refers to a nitrogen containing group as is known in the art derived from ammonia by the replacement of one or more hydrogen radicals by organic radicals. For example, the terms “acylamino” and “alkylamino” refer to specific N-substituted organic radicals with acyl and alkyl substituent groups respectively.

[0211] The term “alkylamino” refers to an $-\text{NHR}$ group wherein R is an alkyl group and/or a substituted alkyl group as previously described. Exemplary alkylamino groups include methylamino, ethylamino, and the like.

[0212] “Dialkylamino” refers to an $-\text{NRR}'$ group wherein each of R and R' is independently an alkyl group and/or a substituted alkyl group as previously described. Exemplary dialkylamino groups include ethylmethylamino, dimethylamino, and diethylamino.

- [0213] "Acylamino" refers to an acyl-NH- group wherein acyl is as previously described.
- [0214] "Aroylamino" refers to an aroyl-NH- group wherein aroyl is as previously described.
- [0215] A "quaternary nitrogen atom," is a nitrogen atom bound to four atoms, for example, four carbon atoms, and having a positive charge available for binding ionically to an anion for the remaining valence. A quaternary nitrogen atom is designated herein as "N⁺."
- [0216] The term "carbonyl" refers to the $-(C=O)-$ group.
- [0217] The term "carboxyl" refers to the $-COOH$ group.
- [0218] The term "cyano" refers to the $-C\equiv N$ group.
- [0219] The terms "halo", "halide", or "halogen" as used herein refer to fluoro, chloro, bromo, and iodo groups.
- [0220] The term "hydroxyl" refers to the $-OH$ group.
- [0221] The term "hydroxyalkyl" refers to an alkyl group substituted with an $-OH$ group.
- [0222] The term "mercapto" refers to the $-SH$ group.
- [0223] The term "oxo" refers to the $=O$ group.
- [0224] The term "nitro" refers to the $-NO_2$ group.
- [0225] The term "thio" refers to a compound described previously herein wherein a carbon or oxygen atom is replaced by a sulfur atom.
- [0226] The term "sulfate" refers to the $-SO_4$ group.
- [0227] The term "sulfonyl" refers to the $-SO_2-$ group.

EXAMPLES

- [0228] The following Examples have been included to provide guidance to one of ordinary skill in the art for practicing representative embodiments of the presently disclosed subject matter. In light of the present disclosure and the general level of skill in the art, those of skill can appreciate that the following Examples are intended to be exemplary only and that numerous changes, modifications, and alterations can be employed without departing from the scope of the presently disclosed subject

matter. The following Examples are offered by way of illustration and not by way of limitation.

Example 1

Experimental Procedures

[0229] **Materials and Methods.** 1,2-Diphenyl-3,5-pyrazolidinedione was purchased from Acros Organics. 2,3,3-Trimethylindolenine was purchased from TCI America. All other reagents were purchased from Sigma-Aldrich. Reactions were run using anhydrous solvents. Microwave reactions were carried out on a CEM Discovery instrument. Automated column chromatography was performed using a Teledyne-Isco Combiflash Rf system. UV-visible spectra were obtained with a Hewlett-Packard 8453 diode array spectrophotometer. HPLC separations were performed on a Shimadzu Prominence system using a 250 x 21.2 mm, 15 micron Phenomenex C18 preparative column and elution at 8 mL/min with a gradient of 10% solvent B (H₂O/ACN 5:95, TFA 0.05%) 90% solvent A (H₂O/ACN 95:5, TFA 0.05%) for 2 min, increasing to 90% solvent B over 30 min and held for a total of 45 min. Emission and excitation spectra were obtained using a Spex Fluorolog 2 spectrofluorometer at room temperature. Mass spectra were obtained on a Hewlett-Packard 1100 mass-selective detector (MS-ESI) using direct infusion. NMR spectra were obtained on Varian Mercury-300 or Inova-400 spectrometers. Reference peaks of 7.26 ppm (CDCl₃), 3.31 ppm (MeOH-d₄) and 2.50 ppm (DMSO-d₆) were used for ¹H spectra and 77.23 ppm (CDCl₃) and 39.50 (DMSO-d₆) for ¹³C spectra. All operations with dyes were performed under dim light.

[0230] **General procedure for the quaternization of donor heterocycles.**

Representative example for compound I: 2,3,3-Trimethylindolenine (0.228 mL, 1.42 mmol) and iodomethane (0.444 mL, 7.09 mmol) were added to a small microwave vial and heated to 120 °C for 5 min. The mixture was cooled, filtered, and washed with ice-chilled ether to give 0.341 g (80%) pale pink powdery solid. Spectral data matched those previously reported.³⁰ ¹H NMR (400 MHz, DMSO) δ 7.96 - 7.87 (m, 1H), 7.85 - 7.78 (m, 1H), 7.67 - 7.57 (m, 2H), 3.97 (s, 3H), 2.77 (s, 3H), 1.53 (s, 6H). ¹³C NMR (100 MHz, DMSO) δ 195.9, 142.0, 141.5, 129.2, 128.7, 123.2, 115.1, 53.9, 34.7, 21.7, 14.2.

[0231] **General procedure for the preparation of acceptor heterocycles as the activated enol ether.** Representative example for compound **Pht**: 1,3-Indandione (0.146 g, 1.00 mmol) was added to a small microwave vial with 1,1,3,3-tetramethoxypropane (0.823 mL, 5.00 mmol), followed by addition of trifluoroacetic acid (7.7 μ L, 0.10 mmol). The vial was capped and heated to 150 °C for 15 min. The reaction was cooled and the precipitate was filtered and rinsed with ice chilled 3:1 hexanes/Et₂O. Isolated 0.158 g (74%) brown solid. ¹H NMR (400 MHz, CDCl₃) δ 7.95 – 7.85 (m, 2H), 7.76 – 7.70 (m, 2H), 7.51 (d, J = 12.2 Hz, 1H), 7.42 (d, J = 12.3 Hz, 1H), 7.27 (dd, J = 12.4, 11.6 Hz, 1H), 3.93 (s, 3H). ¹³C NMR (100 MHz, CDCl₃) δ 191.4, 190.5, 168.9, 145.0, 142.0, 140.7, 134.8, 134.7, 124.1, 123.0, 122.7, 103.9, 58.5.

[0232] **General procedure for the synthesis of merocyanine dyes not bearing reactive side chains.** Representative example for **I-Pht**: Donor **I** (38 mg, 0.13 mmol), sodium acetate (11 mg, 0.13 mmol), and acceptor **Pht** (21 mg, 0.10 mmol) were added to a small microwave vial and diluted in 1:1 MeOH/CHCl₃ (1 mL). The vial was capped and heated to 75 °C for 30 min. The reaction mixture was cooled and concentrated with 250 mg SiO₂ and eluted on a 12 g silica column with 0 – 3% MeOH in CH₂Cl₂ over 20 min. Isolated 29 mg (82%) dark blue solid. ¹H NMR (400 MHz, DMSO) δ 8.12 (t, J = 13.1 Hz, 1H), 7.82 – 7.63 (m, 5H), 7.52 (dd, J = 17.4, 10.2 Hz, 2H), 7.33 (t, J = 7.7 Hz, 1H), 7.23 (d, J = 8.0 Hz, 1H), 7.12 (t, J = 7.4 Hz, 1H), 6.14 (d, J = 13.5 Hz, 1H), 3.49 (s, 3H), 1.62 (s, 6H). ¹³C NMR (100 MHz, CDCl₃) δ 191.8, 168.1, 152.5, 146.6, 143.7, 142.2, 140.8, 140.1, 133.9, 133.7, 128.4, 123.0, 122.2, 122.1, 122.0, 120.8, 120.3, 108.5, 99.7, 47.8, 30.1, 28.5. MS-ESI m/z 356.2 ([M + H]⁺ requires 356.2).

[0233] **3-[(3-bromopropyl)ammonio]propane-1-sulfonate (3).**

Bromopropylamine hydrobromide (10.95 g, 50.0 mmol) and propane sultone (12.21 g, 100 mmol) were added together to a flame dried round bottom flask. The flask was evacuated and argon flushed and EtOH (100 mL) was added. After complete dissolution, Et₃N (7.67 mL, 55.0 mmol) was added quickly drop wise and the reaction mixture was stirred at room temp under argon for 24 h. The resulting white precipitate was filtered and recrystallized from methanol to give 4.83 g (37%) white solid. ¹H NMR (400 MHz, DMSO) δ 8.59 (s, 2H), 3.59 (t, J = 6.6 Hz, 2H), 3.14 –

2.93 (m, 4H), 2.62 (t, $J = 6.7$ Hz, 2H), 2.12 (p, $J = 6.6$ Hz, 2H), 1.93 (p, $J = 6.8$ Hz, 2H). ^{13}C NMR (100 MHz, DMSO) δ 48.7, 46.7, 45.4, 30.9, 28.8, 21.9.

[0234] 2,3,3-Trimethyl-1-(3-((3-sulfopropyl)amino)propyl)indolinium

bromide (5). Compound **3** (1.44 g, 5.54 mmol) was added to a flame dried round bottom flask and the flask was evacuated and flushed with argon. 2,3,3-Trimethylindolenine (4.44 mL, 27.7 mmol) was added and the flask was added to a pre-heated oil bath at 150 °C and the reaction stirred for 30 min. The flask was cooled to room temperature, and then further cooled in an ice bath. Excess liquid was decanted and further removed by syringe. The residue was rinsed with Et₂O (3 X 10 mL), then dried under vacuum. The product was isolated as 1.86 g pink granular solid and used in the following step without further purification.

[0235] General procedure for the synthesis of conjugatable merocyanine

dyes. Representative example for dye **mero60**: Compound **5** (0.628 g, 1.50 mmol), chloroacetic anhydride (0.855 g, 5.00 mmol), and NaOAc (0.246 g, 3.00 mmol) were dissolved in DMF (4 mL) under argon. After 10 min, acceptor **Pht** (0.214 g, 1.00 mmol) was added. The flask was covered in foil and allowed to stir at room temperature under argon for 24 h. The DMF was removed with addition of toluene (3 X 40 mL) in the presence of CELITE™ (3 g) and the cake was eluted on a 24 g silica column with a 0 – 10% MeOH in CH₂Cl₂ gradient over 10 min, followed by 15 min at 10% MeOH. The main product containing fractions were combined, concentrated, and dried under vacuum to give a crude yield of 0.291 g dark blue crystalline solid.

[0236] The chloroacetamide dye intermediate (0.060 g, 0.100 mmol) was diluted in 1:1 MeOH/CHCl₃ (2 mL) and sodium iodide (0.150 g, 1.00 mmol) was added. The flask was fitted with a condenser topped with a drying tube and the reaction was heated at reflux for 16 h. The solvent was then removed and the crude reaction mixture was concentrated, diluted in methanol, and submitted to HPLC column purification to give 0.030 g (21% from **Pht**) dark blue crystalline solid. ^1H NMR (400 MHz, DMSO) δ 8.12 (t, $J = 13.1$ Hz, 1H), 7.79 – 7.67 (m, 5H), 7.57 (dd, $J = 13.1, 9.5$ Hz, 1H), 7.50 (d, $J = 7.5$ Hz, 1H), 7.38 – 7.28 (m, 1H), 7.23 (d, $J = 8.0$ Hz, 1H), 7.13 (t, $J = 7.2$ Hz, 1H), 6.26 (d, $J = 13.4$ Hz, 1H), 4.05 – 3.82 (m, 4H), 3.52 – 3.36 (m, 4H), 2.48 – 2.37 (m, 2H), 2.00 – 1.72 (m, 4H), 1.64 (s, 6H). MS-ESI m/z 687.3 ($[\text{M} - \text{H}]^+$ requires 687.1).

[0237] **General procedure for the synthesis of β -mercaptoethanol derivatives of conjugatable merocyanines.** Representative example for dye mero60: Mero60 (6.6 mg, 9.6 μ mol) was added to a 5 mL conical vial with DMSO (10 μ L) and 50 mM aqueous Na_2HPO_4 , pH 7.5 (490 μ L). The mixture was stirred and β -mercaptoethanol (3.4 μ L, 48 μ mol) was added. The vial was capped, covered in foil, and stirred at room temp for 24 h. The reaction mixture was subjected neat to HPLC column purification. The main product fractions were combined, concentrated, and lyophilized to give 2.1 mg (33%) dark blue powdery solid.

[0238] **Quantum yield determinations.** Quantum yield (QY) values were determined using the following equation: $\text{QY}_{\text{samp}} = \text{QY}_{\text{std}} (\text{grad}_{\text{samp}} / \text{grad}_{\text{std}}) (\eta_{\text{samp}}^2 / \eta_{\text{std}}^2)$, where "grad" is equal to the slope of a plot relating the integrated emission to the absorbance (x axis) at a given concentration, using a minimum of four different concentrations per sample, and η is the refractive index of the solvent used for the fluorescence readings. Merocyanine 540³¹ and sulforhodamine 101³² were used as standards.

[0239] **Photostability measurements.** Each sample dye and fluorescein standard solution was prepared at 1 mM in degassed *n*-BuOH and placed in a capped 1 cm X 1 cm cuvette. Samples were irradiated with a 90 W halogen tungsten lamp with fan cooling to maintain a temperature of $25^\circ\text{C} \pm 1^\circ\text{C}$ during the study. Absorbance measurements were taken prior to irradiation and at designated intervals during irradiation.

[0240] **CBD-MBP dye labeling and Cdc42 binding assay.** The Cdc42 binding domain (CBD) from Wiskott Aldrich Syndrome Protein (WASP) was covalently derivatized with dye to generate a biosensor, using protein sequences and methods described previously (Nalbant *et al.*, *Science* 305:1615-1619 (2004)). Freshly prepared CBD fused to Maltose-binding protein for enhanced solubility (CBD-MBP, 100 μ L of a 132.5 μ M stock solution in 50 mM Na_2HPO_4 , pH 7.5 buffer) was mixed with 5 molar equivalents conjugatable dye in DMSO (3 μ L of 20–25 mM solution) and stirred at room temperature for 2 h. The reaction was quenched by adding 1 μ L of β -mercaptoethanol. Excess dye was removed by passing the reaction mixture through a 1 x 20 cm G-15 column pre-equilibrated with 50 mM Na_2HPO_4 , pH 7.5 buffer and wrapped in aluminum foil. Labeling efficiency and concentrations of

CBD-MBP dye conjugates were determined using absorption measurements at protein and dye maximum absorption wavelengths, using protein dissolved in DMSO to overwhelm solvent effects on dye and produce a sample with known extinction coefficient (see Table 5 for dye absorption maxima and extinction coefficients, CBD-MBP abs max = 280 nm, $\epsilon = 74830 \text{ M}^{-1} \text{ cm}^{-1}$). Dye-labeled CBD-MBP (100 nM in 50 mM Na_2HPO_4 , 150 mM NaCl, pH 7.5 buffer) aliquots were treated with varying concentrations of Cdc42 Q61L (constitutively active mutant) and fluorescence intensity measurements were taken after incubation for 10 min at room temperature.

[0241] **Live cell imaging of Cdc42 biosensors.** Imaging experiments utilized a CBD-Cerulean C49S construct for fluorescence ratio analysis. Labeling reactions were carried out as for CBD-MBP. After size exclusion chromatography, the dye-labeled proteins were dialyzed in 50 mM Na_2HPO_4 , pH 7.5 buffer at 4 °C overnight and then concentrated to 60 μM by centrifugation at 12000 rpm, 4 °C (Millipore, Amicon Ultra-0.5mL, MWCO 10K). NIH 3T3 mouse embryonic fibroblasts (MEFs) were microinjected with biosensor solution and imaged for 20 min with acquisition at every 1 min. Image processing was carried out as previously described (Hodgson *et al.*, *Methods Enzymol.* 406:140-156 (2006); Hodgson *et al.*, *Curr. Protoc. Cell Biol.* pp 1-26, Chapter 14, Unit 14.11 (2010)).

[0242] **Protein expression.** The Cdc42-binding domain (CBD), derived from Wiskott-Aldrich syndrome protein (WASP, residues 201-320), was mutated to include a single cysteine (F271C) for dye attachment and fused at its C-terminus with maltose-binding protein (MBP) through a GSGSGS linker. The CBD-MBP fragment was subcloned into a pET23b plasmid to generate a C-terminal 6xHis fusion protein. The CBD-Cerulean fusion protein was prepared with C49S mutation in Cerulean to eliminate off-target dye labeling and subcloned into a pET23b plasmid to generate a C-terminal 6xHis fusion protein. The constitutively active mutant, Cdc42 Q61L, was subcloned into the pQE-80L plasmid to generate an N-terminal 6xHis fusion protein. The resulting constructs were transformed into the BL21(DE3) strain from *E. coli* (Stratagene) and the bacteria was cultured at 37 °C, 245 rpm (New Brunswick Scientific, Innova 4000) in Leuria-Bertani medium in the presence of 100 $\mu\text{g/mL}$ carbenicillin until OD600 reached 0.8. Protein expression was conducted at 32 °C for 2 hours after addition of 40 mM isopropyl β -D-1-thiogalactopyranoside (IPTG). The

bacteria pellet was collected by centrifugation (Beckman, model J-68) at 4000 rpm for 20 min, 4 °C, and stored at -80 °C prior to purification.

[0243] Protein purification. For lysis, the bacterial pellets were resuspended in 30 mL of the lysis buffer (50 mM NaH_2PO_4 , pH 7.0, 300 mM NaCl) and disrupted by sonication (Branson Digital Sonifier). The supernatant was collected by centrifugation (Sorvall, model T21) at 6000 rpm for 15 min, 4 °C, followed by incubation with Talon Metal Affinity Resin (Clontech) on a nutator mixer (BD Clay Adams, model 1105) for 1 hour at 4 °C. The protein-bound resin was collected after centrifugation (Eppendorf, model 5810) at 1600 rpm for 2 min, 4 °C, washed with the lysis buffer, and loaded onto a polypropylene column (BioRad). The immobilized proteins were washed with 20 mL of the washing buffer (50 mM NaH_2PO_4 , pH 7.5) and eluted with 1–2 mL elution buffer (300 mM imidazole, pH 7.0). To fully reduce the F271C residue of CBD-MBP, 1 μL of β -mercaptoethanol (BioRad) was added to the eluate and the resulting protein solution was concentrated to 200 μL by centrifuge (VivaSpin 20 MWCO 10K; Eppendorf, model 5810). The excess imidazole and β -mercaptoethanol were removed by buffer exchange on a dextran desalting column (Thermo Scientific Pierce, 5K MWCO) using the protein labeling buffer (50 mM NaH_2PO_4 , pH 7.5). The protein aliquots were stored at -80 °C.

[0244] Cell culture and microinjection. NIH 3T3 mouse embryonic fibroblasts (MEF) were maintained in 10% CO_2 at 37 °C in Dulbecco's modified Eagle's medium (DMEM, Cellgro) with 10% fetal bovine serum (HyClone, Thermo Scientific) and 2 mM GlutaMax (Gibco, Life Technologies). The cells were plated on coverslips coated with fibronectin (Sigma-Aldrich) overnight. Micropipettes were made fresh with a flaming/brown micropipette puller (Sutter Instrument Co., Model P-97) and thin-walled glass capillary tubes (World Precision Instruments, 1.2 O.D./0.9 I.D., TW120F-4). Cells were microinjected on an inverted microscope (Zeiss Axiovert S100 TV) with a 20X objective (LD Achrostatigmat, 0.30 N.A., Ph1), followed by recovery at 37 °C in the cell incubator in Ham's F-12K medium without phenol red (Gibco) with 5% fetal bovine serum, 15 mM HEPES, and 2mM GlutaMax for at least 30 min prior to imaging.

[0245] Image acquisition. Live cell imaging was carried out on an Olympus IX81 microscope with UPLFLN 40X oil objective (NA 1.3) and mercury lamp

excitation (103W HBO bulb). Filters used for Cerulean were 430/24 excitation and 470/24 emission. Excitation was through a ND1.5 (3% transmission) neutral density filter, with 200 ms exposure. For **mero61**, filters were 545/50 excitation and 630/45 emission. Excitation was through a ND2.0 (1%T) filter with 200 ms exposure. Imaging of **mero62** and **mero87** was carried out using 590/30 excitation, 630/40 emission, ND2.0 (1%T), and 200 ms exposure. Images were acquired with a Coolsnap ES2 camera (Photometrics) with a Sony 6.45 x 6.45 μM /pixel chip using 2 x 2 binning. All image acquisition, processing, and analysis was carried out with Metamorph software.

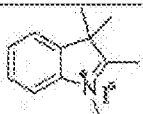
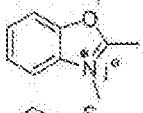
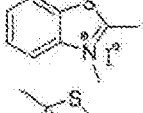
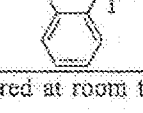
Example 2

Synthesis and Evaluation of Merocyanine Library

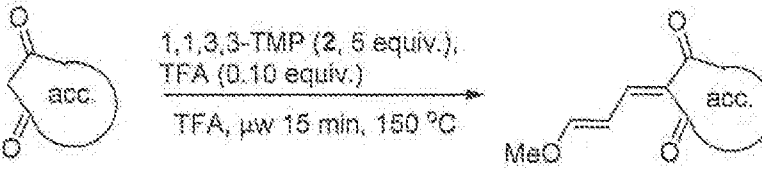
[0246] **Synthesis of the merocyanine library.** Synthesis of the dyes utilized commercially available donor and acceptor components, with the exception of diethylbarbituric acid, which was synthesized using a previously described method (Prajapati *et al.*, *J. Chem. Soc. Perkin Trans. 1*:607-610 (1988)). Donor and acceptor groups will be henceforth referred to by their letter code names, derived from the parent heterocycle, in order to clarify discussion (Ernst *et al.*, *Cytometry* 10:3-10 (1989)) (**Table 1**). Donor groups are ordered according to their Brooker basicity, with indolenine (**I**) being the weakest and quinoline (**Q**) the strongest (Brooker *et al.*, *J. Am. Chem. Soc.* 73:1087-1093 (1951); Brooker, *Rev. Mod. Phys.* 14:275-293 (1942)). Donor components of the dyes were activated for conjugation through quaternization with iodomethane (**Table 1**). Each acceptor ring was prepared as its corresponding methyl enol ether through acid catalyzed reaction with malonaldehyde bis(dimethyl) acetal (**Table 2**).

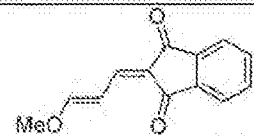
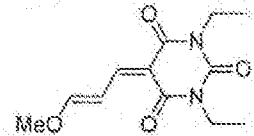
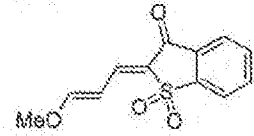
Table 1. Synthesis of quaternized donor heterocycles.

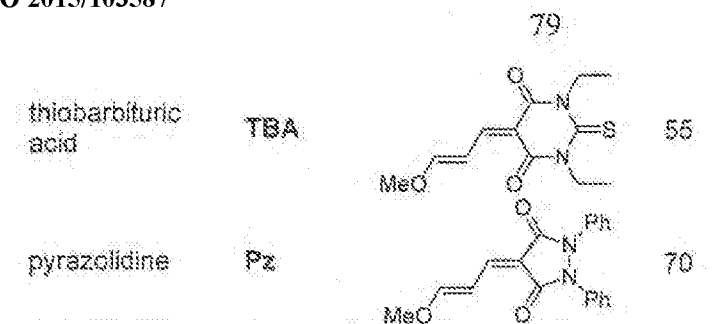


heterocycle	donor code	Brooker basicity ^a	product	temp (°C)	time (min)	yield (%)
indolenine	I	260		120	5	80
benzoxazole	O	345		100	20	51
benzothiazole	S	375		120	20	93
thiadiazole	TD	-		90	20	98
quinoline	Q	730		r.t. ^b	15	95

^a From reference 15.^b Reaction stirred at room temperature for time indicated without microwave heating.**Table 2.** Synthesis of activated acceptor enol ethers.

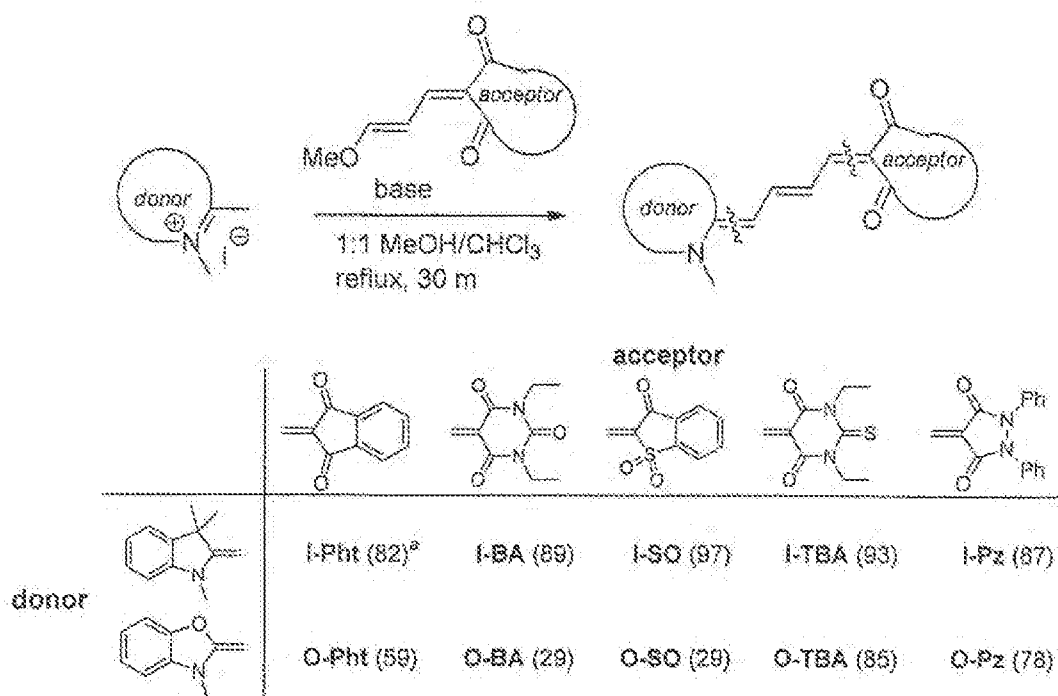


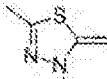
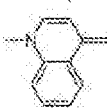
heterocycle	acceptor code	product	% yield
phthal.	Pht		74
barbituric acid	BA		61
benzothiophene	SO		73



[0247] The dyes for initial screening were synthesized through reaction of an acceptor enol ether with a donor in the presence of 1.25 equivalents base (Table 3). All dyes were made with three intervening double bonds between the terminal heterocycles as this provided the optimum compromise between solvent-sensitive fluorescence and photostability (Mostovnikov *et al.*, *Zh. Prikl. Spektrosk.* 20:42-47 (1974); Ishchenko, *Usp. Khim.* 60:1708-1743 (1991)). Following reflux, the dyes were purified by column chromatography or in some cases by precipitation through concentration of the reaction mixture, followed by addition of diethyl ether and washing with MeOH. Donor **Q** did not react successfully with acceptors when using NaOAc and instead required the use of the stronger base *N*-methylpiperidine.

Table 3. Preparation of the merocyanine library



	S-Pht (72)	S-BA (70)	S-SO (78)	S-TBA (93)	S-Pz (70)
	TD-Pht (82)	TD-BA (61)	TD-SO (77)	TD-TBA (70)	TD-Pz (53)
	Q-Pht (56)	Q-BA (45)	Q-SO (20)	Q-TBA (89)	Q-Pz (39)

^a compound name (% yield)

[0248] **Brightness and solvatochromism of the dyes.** The synthesized library dyes were screened for relative brightness based on the fluorescence intensity value at the emission peak maximum (EPM) when the dye is excited at its excitation maximum. Given that the emission peak shapes for the dyes were generally similar, the EPM value provided a straightforward measurement that is useful for the comparison of brightness across a broad range of dye structures. The solvents 1,4-dioxane, dimethylsulfoxide (DMSO), butanol (BuOH), and methanol (MeOH) were used to evaluate the solvent-dependence of dye brightness (Table 4). These solvents were selected to provide a range of polarities as well as examples of both weak and strong hydrogen bonding (Hansen, *Hansen Solubility Parameters: A User's Handbook, Second Edition* (Hansen, C. M., Ed.) pp 1-26, Chapter 1, CRC Press, Boca Raton (2007); Hansen *et al.*, *Dan. Kemi* 48:81-84 (1967)).

Table 4: Polarity characteristics of solvents selected for brightness screening.

solvent	solubility parameter		
	dispersion	polar	H-bonding
MeOH	14.7	12.3	22.3
BuOH	16.0	5.7	15.8
DMSO	18.4	16.4	10.2
1,4-dioxane	17.5	1.8	9.0

[0249] Dyes with weaker donor groups such as indolenine (I), benzoxazole (O), and benzothiazole (S) tended to have the largest EPM values (Fig. 1). As a whole, dyes containing the quinoline (Q) donor group were significantly dimmer than the other dyes. The presence of BA, TBA, or SO acceptor groups led to the highest

brightness levels, regardless of the donor component of the dye. Conversely, the Pz acceptor group was associated with diminished relative brightness, regardless of the donor group with which it was paired. The reduced brightness for Pz-containing dyes may be attributable to an alternative de-excitation pathway generated through the unrestricted rotation of phenyl substituents around the C-N bonds (Sharafy *et al.*, *J. Am. Chem. Soc.*, 93:4119-4125 (1971)). No clear trend was observed when the solvents were ordered based on dipolar characteristics but dye brightness in all cases was maximized in solvents with intermediate hydrogen bonding capacity. The fluorescence intensity of several of the brightest dyes ($\text{EPM} > 8 \times 10^5$) was strongly solvent-dependent. As solvent-dependent changes in brightness can be harnessed to report protein activity (Gulyani *et al.*, *Nat. Chem. Biol.* 7:437-444 (2011)), this was an important consideration when selecting dyes for synthesis of reactive derivatives (see below).

[0250] Excitation and emission maxima of the dyes in each of the four solvents were also recorded. Most of the solvatochromic shifts were modest (<20 nm), especially among the weaker donor groups (Fig. 2). The largest shifts were seen among the Q donor group dyes, which uniformly exhibited negative solvatochromism (Fig. 3). A hypsochromic shift when going from low to high polarity solvent (e.g. from dioxane to MeOH, Fig. 3) is typically observed when the ground state structure of the dye is more dipolar than that of the excited state. This result would be consistent with an increased propensity of the quinoline heterocycle to establish aromaticity in the ground state relative to the other donor groups.

[0251] **Photostability measurements.** A previously reported method (Toutchkine *et al.*, *J. Am. Chem. Soc.* 125:4132-4145 (2003)) was used to determine the photobleaching rates of each dye relative to fluorescein (Song *et al.*, *Biophys. J.* 68:2588-2600 (1995)) (Fig. 4). The most photostable dyes contained the weakest donor (I) and values fell sharply with increasing donor strength. Dyes containing the Pz acceptor group tended to show greater photostability relative to other acceptor groups in the library. Photobleaching in merocyanines is thought to be mediated by singlet oxygen attacking the polyene chain due to its strongly electrophilic character (Toutchkine *et al.*, *J. Am. Chem. Soc.* 125:4132-4145 (2003)). In dyes with weaker donors, such as I, electron density in the polyene is reduced and therefore the extent

of reaction with excited oxygen species and subsequent disproportionation reactions would also be decreased.

[0252] Selection and synthesis of water soluble, conjugatable merocyanines.

Several dyes showed extraordinary and useful combinations of brightness, photostability, and solvent-dependent fluorescence. We therefore developed these further by incorporating water solubilizing and reactive side chains to generate derivatives that could be used for protein labeling. Dyes containing the **I** donor group were selected based on their photostability and because the geminal dimethyl group of **I** is known to assist in limiting dye aggregation that can result in reduced solubility and fluorescence (Toutchkine *et al.*, *J. Am. Chem. Soc.* 125:4132-4145 (2003)). Despite the considerable photostability of **I-Pz**, this dye was not selected for further development due to its low brightness. Although the dyes containing donor **Q** showed the greatest solvatochromism, their poor photostability and relatively low brightness led us to eliminate them from further consideration.

[0253] Synthesis of the water soluble, conjugatable versions of the selected dyes began with reaction of 1-bromopropylamine with propane sultone to give 3-((3-bromopropyl)amino)propane-1-sulfonic acid (**3**, **Scheme 1**). The donor heterocycle was then quaternized with compound **3** and subsequently coupled with the acceptor enol ether in the presence of chloroacetic anhydride and sodium acetate using a one-pot procedure (Toutchkine *et al.*, *Bioconjugate Chem.* 18:1344-1348 (2007)). Refluxing in MeOH for 3 h in the presence of an excess of NaI gave merocyanines containing iodoacetamide (IAA) for protein labeling. Consistent with the nomenclature for merocyanines used in our lab (Gulyani *et al.*, *Nat. Chem. Biol.* 7:437-444 (2011)), these conjugatable dyes were named **mero60**, **mero61**, **mero62**, and **mero87** (**Table 5**). Physical constants were obtained for both the reactive derivatives and for beta-mercaptoethanol adducts that eliminate effects of iodine on quantum yield, and better mimic dye attached to protein.

Scheme 1. Synthesis of water soluble, conjugatable merocyanines.

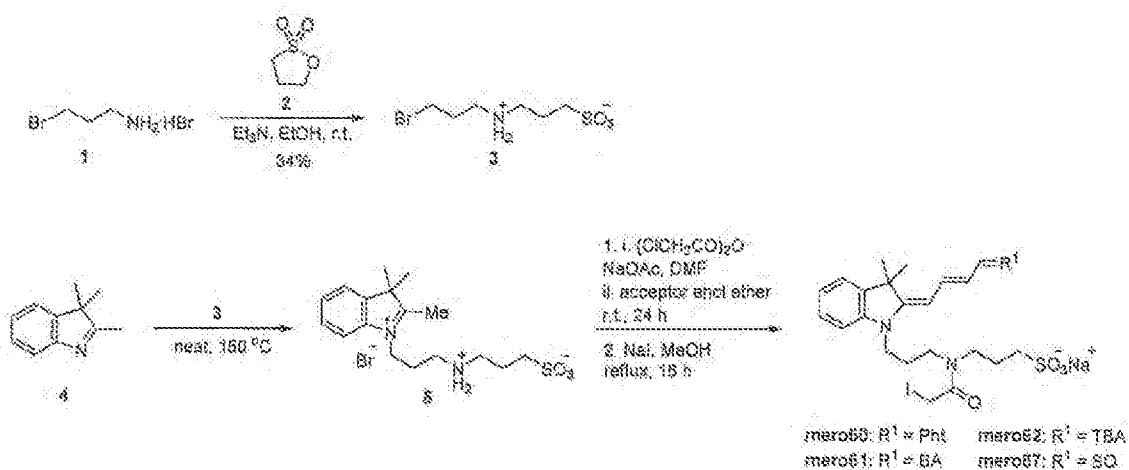


Table 5. Photophysical properties of functionalized merocyanines.

dye name	% yield (from enol ether)	solvent	QY ^a	ϵ^b	QY x ϵ	λ_{max} (nm) absorption	λ_{max} (nm) emission
mero60	21	DMSO	0.37	126000	46620	593	620
		MeOH	0.02	133000	2660	591	614
		BuOH	0.16	132000	21120	593	616
		H ₂ O	<0.01	106000	11	595	608
mero61	26	DMSO	0.37	100000	37035	570	595
		MeOH	0.07	167000	11690	565	588
		BuOH	0.17	146000	24820	570	592
		H ₂ O	0.03	131000	3930	565	586
mero62	21	DMSO	0.20	135000	27000	597	615
		MeOH	0.09	140000	12600	590	608
		BuOH	0.18	139000	25020	594	614
		H ₂ O	0.10	36000	3600	585	608
mero87	33	DMSO	0.34	140000	47600	592	620
		MeOH	0.06	107000	6420	588	620
		BuOH	0.21	111000	23310	591	622
		H ₂ O	0.02	81000	1620	592	616
mero60-BME	-	DMSO	0.66	124000	81870	595	618
mero61-BME	-	DMSO	0.55	116000	63800	570	594
mero62-BME	-	DMSO	0.41	142000	58220	596	614
mero87-BME	-	DMSO	0.27	105000	28350	590	620
I-SO ^c	-	BuOH	0.54	134000	72000	587	618
		OcOH	0.98	125000	123000	587	617
		MeOH	0.08	143000	11440	586	615

^aQuantum yield of fluorescence, error $\pm$ 10%. ^bMolar extinction coefficient, error $\pm$ 10%. ^cValues for I-SO as reported in ref.9 are included for comparison (see Discussion and Conclusions).

Compound characterization data:Quaternized donor heterocycles:

[0254] O: Yield = 51%. ^1H NMR (400 MHz, DMSO) δ 8.17 – 8.08 (m, 2H), 7.83 – 7.74 (m, 2H), 4.08 (s, 3H), 3.07 (s, 3H); ^{13}C NMR (100 MHz, DMSO) δ 169.0, 147.2, 130.3, 128.6, 127.6, 114.4, 112.8, 32.6, 13.4.

[0255] S: Yield = 93%. ^1H NMR (400 MHz, DMSO) δ 8.48 – 8.42 (m, 1H), 8.29 (d, J = 8.4 Hz, 1H), 7.89 (ddd, J = 8.5, 7.3, 1.3 Hz, 1H), 7.80 (ddd, J = 8.2, 7.3, 1.0 Hz, 1H), 4.21 (s, 3H), 3.18 (s, 3H); ^{13}C NMR (100 MHz, DMSO) δ 177.2, 141.5, 129.2, 128.6, 128.0, 124.5, 116.7, 36.3, 17.2.

[0256] TD: Yield = 98%. ^1H NMR (400 MHz, DMSO) δ 4.17 (s, 3H), 3.02 (s, 3H), 2.81 (s, 3H); ^{13}C NMR (100 MHz, DMSO) δ 174.4, 166.7, 41.4, 39.5, 15.9, 14.9.

[0257] Q: Yield = 95%. ^1H NMR (400 MHz, DMSO) δ 9.37 (d, J = 6.1 Hz, 1H), 8.58 – 8.43 (m, 2H), 8.27 (ddd, J = 8.7, 7.0, 1.4 Hz, 1H), 8.12 – 8.00 (m, 2H), 4.58 (s, 3H), 3.01 (s, 3H); ^{13}C NMR (100 MHz, DMSO) δ 158.1, 148.9, 137.6, 134.9, 129.6, 128.4, 126.7, 122.4, 119.5, 45.0, 39.5, 19.6.

Acceptor enol ethers:

[0258] BA: Yield = 61%. ^1H NMR (400 MHz, CDCl_3) δ 8.17 – 7.97 (m, 1H), 7.58 – 7.38 (m, 2H), 3.98 (m, 4H), 3.92 (s, 3H), 1.22 (m, 6H); ^{13}C NMR (100 MHz, CDCl_3) δ 171.4, 162.4, 162.2, 157.5, 151.0, 110.9, 105.8, 58.7, 37.2, 36.6, 13.6, 13.5.

[0259] SO: Yield = 73%. Spectra matched those previously reported (Touichkine *et al.*, *J. Am. Chem. Soc.* 125:4132 (2003)). ^1H NMR (400 MHz, CDCl_3) δ 8.04 (d, J = 7.6 Hz, 1H), 7.98 (d, J = 8.0 Hz, 1H), 7.86 (t, J = 7.6 Hz, 1H), 7.78 (t, J = 7.6 Hz, 1H), 7.67 (d, J = 12.8 Hz, 1H), 7.45 (d, J = 12.0 Hz, 1H), 6.45 (t, J = 12.4 Hz, 1H), 3.96 (s, 1H).

[0260] TBA: Yield = 55%. ^1H NMR (400 MHz, CDCl_3) δ 8.11 (d, J = 11.8 Hz, 1H), 7.60 – 7.45 (m, 2H), 4.51 (q, J = 7.0 Hz, 4H), 3.96 (s, 3H), 1.28 (dt, J = 12.8, 6.5 Hz, 6H); ^{13}C NMR (100 MHz, CDCl_3) δ 179.3, 172.6, 161.0, 160.4, 159.2, 111.5, 106.7, 77.2, 59.1, 43.8, 43.2, 12.7, 12.6.

[0261] Pz: Yield = 70%. ^1H NMR (400 MHz, CDCl_3) δ 7.85 – 7.67 (m, 1H), 7.50 (d, J = 12.3 Hz, 1H), 7.44 – 7.27 (m, 9H), 7.20 – 7.08 (m, 2H), 3.91 (s, 3H); ^{13}C NMR

(100 MHz, CDCl₃) δ 170.4, 164.4 (2C), 152.1, 137.2 (2C), 129.1, 129.0 (2C), 126.3, 126.2, 122.4, 122.3, 111.0, 104.1, 58.7.

Library dyes:

[0262] **I-BA:** Yield = 89%. ¹H NMR (300 MHz, DMSO) δ 8.12 (dt, J = 13.1, 6.5 Hz, 2H), 7.71 (t, J = 13.0 Hz, 1H), 7.50 (d, J = 6.8 Hz, 1H), 7.40 – 7.21 (m, 2H), 7.21 – 7.06 (m, 1H), 6.13 (d, J = 13.5 Hz, 1H), 3.84 (q, J = 6.8 Hz, 3H), 3.51 (s, 3H), 1.62 (s, 6H), 1.10 (td, J = 6.8, 1.6 Hz, 6H). MS-ESI m/z 394.1 ([M + H]⁺ requires 394.2).

[0263] **I-SO:** Yield = 97%. ¹H NMR (400 MHz, CDCl₃) δ 8.00 (d, J = 7.4 Hz, 1H), 7.94 (d, J = 7.4 Hz, 1H), 7.83 – 7.67 (m, 4H), 7.34 – 7.26 (m, 2H), 7.11 (t, J = 7.5 Hz, 1H), 6.92 (d, J = 7.9 Hz, 1H), 6.86 (t, J = 13.2 Hz, 1H), 5.89 (d, J = 12.8 Hz, 1H), 3.46 – 3.31 (s, 3H), 1.62 (s, 6H). MS-ESI m/z 392.0 ([M + H]⁺ requires 392.1).

[0264] **I-TBA:** Yield = 93%. ¹H NMR (400 MHz, CDCl₃) δ 8.17 – 8.04 (m, 1H), 7.95 – 7.81 (m, 2H), 7.36 – 7.28 (m, 2H), 7.16 (dt, J = 7.5, 0.7 Hz, 1H), 6.97 (d, J = 7.9 Hz, 1H), 6.05 – 5.90 (m, 1H), 4.58 (q, J = 6.9 Hz, 4H), 3.38 (s, 3H), 1.62 (s, 6H), 1.38 – 1.26 (m, 6H). MS-ESI m/z 410.1 ([M + H]⁺ requires 410.2).

[0265] **I-Pz:** Yield = 87%. ¹H NMR (400 MHz, CDCl₃) δ 7.76 (dd, J = 25.0, 13.0 Hz, 2H), 7.57 (t, J = 13.1 Hz, 1H), 7.48 – 7.40 (m, 4H), 7.33 – 7.26 (m, 6H), 7.14 – 7.07 (m, 3H), 6.91 (d, J = 7.8 Hz, 1H), 5.87 (d, J = 13.0 Hz, 1H), 3.36 (s, 3H), 1.65 (s, 6H). MS-ESI m/z 462.1 ([M + H]⁺ requires 462.2).

[0266] **O-Pht:** Yield = 59%. ¹H NMR (400 MHz, DMSO) δ 7.91 (t, J = 12.5 Hz, 1H), 7.70 – 7.56 (m, 6H), 7.55 – 7.32 (m, 4H), 6.10 (d, J = 13.5 Hz, 1H), 3.72 (s, 3H). MS-ESI m/z 330.1 ([M + H]⁺ requires 330.1).

[0267] **O-BA:** Yield = 29%. ¹H NMR (300 MHz, CDCl₃) δ 8.08 (d, J = 12.6 Hz, 1H), 7.83 (td, J = 25.3, 13.1 Hz, 2H), 7.40 (d, J = 8.0 Hz, 1H), 7.35 – 7.24 (m, 2H), 7.13 – 7.06 (m, 1H), 5.45 (d, J = 11.9 Hz, 1H), 4.02 (dq, J = 7.0, 1.3 Hz, 4H), 3.53 (s, 3H), 1.25 (dt, J = 7.0, 3.0 Hz, 6H). MS-ESI m/z 368.1 ([M + H]⁺ requires 368.2).

[0268] **O-SO:** Yield = 29%. ¹H NMR (400 MHz, DMSO) δ 8.01 (t, J = 13.0 Hz, 1H), 7.92 (d, J = 6.5 Hz, 1H), 7.85 – 7.76 (m, 3H), 7.72 (d, J = 8.2 Hz, 1H), 7.69 – 7.57 (m, 2H), 7.48 (t, J = 7.2 Hz, 1H), 7.41 (t, J = 7.2 Hz, 1H), 6.55 (t, J = 13.1 Hz, 1H), 6.26 (d, J = 13.5 Hz, 1H), 3.75 (s, 3H). MS-ESI m/z 366.1 ([M + H]⁺ requires 366.1).

[0269] O-TBA: Yield = 85%. ^1H NMR (400 MHz, DMSO) δ 8.04 – 7.96 (m, 1H), 7.83 – 7.72 (m, 4H), 7.53 (dt, J = 7.7, 1.3 Hz, 1H), 7.48 (dt, J = 7.6, 1.2 Hz, 1H), 6.42 (d, J = 13.8 Hz, 1H), 4.43 (q, J = 6.9 Hz, 4H), 3.83 (s, 3H), 1.17 (t, J = 6.9 Hz, 6H). MS-ESI m/z 384.1 ($[\text{M} + \text{H}]^+$ requires 384.1).

[0270] O-Pz: Yield = 78%. ^1H NMR (400 MHz, CDCl_3) δ 7.81 (t, J = 12.8 Hz, 1H), 7.66 (d, J = 13.2 Hz, 1H), 7.51 (t, J = 13.2 Hz, 1H), 7.46 – 7.36 (m, 5H), 7.31 – 7.21 (m, 6H), 7.12 – 7.05 (m, 3H), 5.42 (d, J = 12.4 Hz, 1H), 3.51 (s, 3H). MS-ESI m/z 437.0 ($[\text{M} + \text{H}]^+$ requires 436.2).

[0271] S-Pht: Yield = 72%. ^1H NMR (400 MHz, DMSO, 90 °C) δ 7.87 (d, J = 7.9 Hz, 1H), 7.74 – 7.45 (m, 9H), 7.33 (t, J = 7.1 Hz, 1H), 6.43 (d, J = 12.4 Hz, 1H), 3.76 (s, 3H). MS-ESI m/z 346.0 ($[\text{M} + \text{H}]^+$ requires 346.1).

[0272] S-BA: Yield = 70%. ^1H NMR (300 MHz, DMSO) δ 7.95 – 7.81 (m, 2H), 7.72 – 7.60 (m, 3H), 7.51 (t, J = 7.7 Hz, 1H), 7.34 (t, J = 7.5 Hz, 1H), 6.50 (d, J = 11.5 Hz, 1H), 3.83 (q, J = 6.9 Hz, 4H), 3.80 (s, 3H), 1.10 (t, J = 6.8 Hz, 6H). MS-ESI m/z 384.0 ($[\text{M} + \text{H}]^+$ requires 384.1).

[0273] S-SO: Yield = 78%. ^1H NMR (400 MHz, DMSO) δ 7.93 (d, J = 7.3 Hz, 1H), 7.91 – 7.72 (m, 5H), 7.63 (t, J = 10.7 Hz, 2H), 7.58 – 7.50 (m, 1H), 7.44 – 7.30 (m, 1H), 6.98 – 6.68 (m, 1H), 6.55 (d, J = 13.0 Hz, 1H), 3.82 (s, 3H). MS-ESI m/z 382.0 ($[\text{M} + \text{H}]^+$ requires 382.1).

[0274] S-TBA: Yield = 93%. ^1H NMR (400 MHz, DMSO) δ 8.10 (d, J = 8.0 Hz, 1H), 7.94 – 7.83 (m, 4H), 7.63 (t, J = 7.6 Hz, 1H), 7.51 (t, J = 7.6 Hz, 1H), 6.82 (d, J = 13.5 Hz, 1H), 4.43 (q, J = 6.8 Hz, 4H), 3.94 (s, 3H), 1.17 (t, J = 7.0 Hz, 6H). MS-ESI m/z 400.1 ($[\text{M} + \text{H}]^+$ requires 400.1).

[0275] S-Pz: Yield = 70%. ^1H NMR (400 MHz, CDCl_3) δ 7.66 (d, J = 12.9 Hz, 1H), 7.55 (t, J = 11.3 Hz, 2H), 7.46 – 7.37 (m, 6H), 7.31 – 7.22 (m, 5H), 7.18 (d, J = 8.0 Hz, 1H), 7.13 – 7.04 (m, 2H), 6.02 (d, J = 12.3 Hz, 1H), 3.61 (s, 3H). MS-ESI m/z 452.0 ($[\text{M} + \text{H}]^+$ requires 452.1).

[0276] TD-Pht: Yield = 82%. ^1H NMR (400 MHz, DMSO) δ 7.68 – 7.49 (m, 5H), 7.38 (d, J = 5.4 Hz, 2H), 6.32 (d, J = 12.8 Hz, 1H), 3.84 (s, 3H), 2.57 (s, 3H). MS-ESI m/z 311.0 ($[\text{M} + \text{H}]^+$ requires 311.1).

[0277] **TD-BA:** Yield = 61%. ^1H NMR (400 MHz, DMSO) δ 7.85 – 7.70 (m, 1H), 7.67 – 7.47 (m, 2H), 6.44 – 6.19 (m, 1H), 3.86 (s, 3H), 3.82 (q, J = 6.9 Hz, 4H), 2.58 (s, 3H), 1.07 (t, J = 6.9 Hz, 6H). MS-ESI m/z 349.1 ($[\text{M} + \text{H}]^+$ requires 349.1).

[0278] **TD-SO:** Yield = 77%. ^1H NMR (400 MHz, DMSO) δ 7.88 (d, J = 6.2 Hz, 1H), 7.84 – 7.73 (m, 3H), 7.68 (t, J = 12.6 Hz, 1H), 7.56 (d, J = 13.9 Hz, 1H), 6.45 (t, J = 14.0 Hz, 2H), 3.88 (s, 3H), 2.60 (s, 3H). MS-ESI m/z 347.0 ($[\text{M} + \text{H}]^+$ requires 347.0).

[0279] **TD-TBA:** Yield = 70%. ^1H NMR (400 MHz, DMSO) δ 7.76 – 7.55 (m, 3H), 6.68 – 6.48 (m, 1H), 4.42 (q, J = 6.9 Hz, 4H), 3.96 (s, 3H), 2.64 (s, 3H), 1.16 (t, J = 6.9 Hz, 6H). MS-ESI m/z 365.0 ($[\text{M} + \text{H}]^+$ requires 365.1).

[0280] **TD-Pz:** Yield = 53%. ^1H NMR (400 MHz, CDCl_3) δ 7.58 (d, J = 13.0 Hz, 1H), 7.50 – 7.40 (m, 4H), 7.31 – 7.24 (m, 4H), 7.15 (d, J = 12.7 Hz, 2H), 7.11 – 7.05 (m, 2H), 5.75 (d, J = 12.7 Hz, 1H), 3.69 (s, 3H), 2.49 (s, 3H). MS-ESI m/z 417.1 ($[\text{M} + \text{H}]^+$ requires 417.1).

[0281] **Q-Pht:** Yield = 56%. ^1H NMR (400 MHz, DMSO) δ 8.57 (d, J = 8.5 Hz, 1H), 8.27 (d, J = 7.3 Hz, 1H), 8.14 (t, J = 12.0 Hz, 1H), 7.91 (m, 2H), 7.70 – 7.47 (m, 7H), 7.35 (d, J = 13.8 Hz, 1H), 7.06 (d, J = 13.6 Hz, 1H), 4.05 (s, 3H). MS-ESI m/z 340.1 ($[\text{M} + \text{H}]^+$ requires 340.1).

[0282] **Q-BA:** Yield = 45%. ^1H NMR (400 MHz, CDCl_3) δ 8.09 (d, J = 8.6 Hz, 1H), 8.04 (d, J = 13.2 Hz, 1H), 7.93 (t, J = 12.8 Hz, 1H), 7.79 (t, J = 12.7 Hz, 1H), 7.69 (ddd, J = 8.5, 7.1, 1.4 Hz, 1H), 7.46 (ddd, J = 8.3, 7.1, 1.1 Hz, 1H), 7.42 – 7.38 (m, 1H), 7.28 (d, J = 7.4 Hz, 1H), 6.95 (d, J = 7.5 Hz, 1H), 6.75 (d, J = 12.8 Hz, 1H), 4.02 (q, J = 7.0 Hz, 4H), 3.82 (s, 3H), 1.24 (dt, J = 7.2, 6.8 Hz, 6H). MS-ESI m/z 378.1 ($[\text{M} + \text{H}]^+$ requires 378.2).

[0283] **Q-SO:** Yield = 20%. ^1H NMR (400 MHz, DMSO) δ 8.65 (d, J = 8.7 Hz, 1H), 8.46 (d, J = 7.0 Hz, 1H), 8.25 – 8.15 (m, 1H), 8.04 – 7.94 (m, 2H), 7.86 – 7.82 (m, 1H), 7.79 – 7.67 (m, 5H), 7.54 – 7.34 (m, 1H), 7.19 (d, J = 14.0 Hz, 1H), 6.77 – 6.51 (m, 1H), 4.14 (s, 3H). MS-ESI m/z 376.0 ($[\text{M} + \text{H}]^+$ requires 376.1).

[0284] **Q-TBA:** Yield = 69%. ^1H NMR (400 MHz, DMSO) δ 8.72 – 8.65 (m, 2H), 8.19 – 8.10 (m, 2H), 8.08 – 8.00 (m, 2H), 7.86 – 7.76 (m, 2H), 7.63 (d, J = 14.4 Hz, 1H), 7.26 (d, J = 14.4 Hz, 1H), 4.45 (q, J = 7.0 Hz, 4H), 4.24 (s, 3H), 1.17 (t, J = 6.9 Hz, 6H). MS-ESI m/z 394.1 ($[\text{M} + \text{H}]^+$ requires 394.2).

[0285] **Q-Pz**: Yield = 39%. ^1H NMR (400 MHz, DMSO) δ 8.62 (d, J = 8.7 Hz, 1H), 8.46 (d, J = 7.0 Hz, 1H), 8.32 – 8.11 (dd, J = 14.0, 11.2 Hz, 1H), 7.99 (q, J = 8.9 Hz, 2H), 7.79 (d, J = 7.2 Hz, 1H), 7.71 (t, J = 7.3 Hz, 1H), 7.49 – 7.21 (m, 10H), 7.18 – 6.96 (m, 3H), 4.14 (s, 3H). MS-ESI m/z 446.2 ($[\text{M} + \text{H}]^+$ requires 446.2).

Intermediates for conjugatable dye synthesis:

[0286] **INS**: ^1H NMR (400 MHz, D_2O) δ 7.86 – 7.75 (m, 2H), 7.73 – 7.65 (m, 2H), 4.80 (s, 3H), 4.67 – 4.57 (m, 2H), 3.40 – 3.32 (m, 2H), 3.32 – 3.25 (m, 2H), 3.05 (t, J = 7.3 Hz, 2H), 2.91 – 2.83 (m, 1H), 2.50 – 2.35 (m, 2H), 2.25 – 2.11 (m, 2H), 1.61 (s, 6H); ^{13}C NMR (100 MHz, D_2O) δ 197.7, 141.6, 140.5, 130.0, 129.2, 123.5, 117.7, 114.8, 114.6, 54.6, 47.7, 46.4, 44.6, 44.2, 23.9, 21.6, 21.1.

Example 3

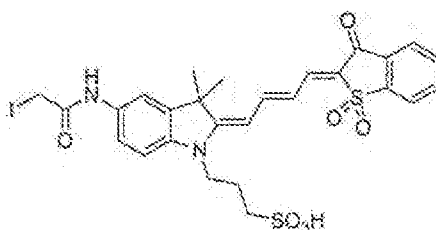
Biosensor Applications

[0287] Conjugates of the reactive dyes with the Cdc42 binding domain (CBD) of Wiskott Aldrich Syndrome Protein (WASP) were prepared and evaluated for their ability to report the activated, GTP-bound conformation of Cdc42. As shown in previous studies using dye-labeled CBD as a biosensor for Cdc42 activity *in vitro* and *in vivo* (Nalbant *et al.*, *Science* 305:1615-1619 (2004)), CBD binds selectively to the activated, GTP-bound conformation of Cdc42. Upon binding, a solvent-sensitive dye on CBD can respond to a change in environment as it is brought near Cdc42.

[0288] All of the newly prepared dyes exhibited an increase in fluorescence intensity on binding to activated Cdc42 *in vitro*, with **mero60** showing a remarkable 14.7 fold increase (Fig. 5a). The **mero61** and **mero62** dyes exhibited the greatest absolute brightness at saturating levels of Cdc42 (Fig. 5b) but their brightness changes upon binding Cdc42 were smaller relative to **mero60** and **mero87** dyes due to their much higher initial brightness in aqueous solution (Table 6). Biosensors based on these new dyes showed significant improvements in both brightness and extent of change upon Cdc42 binding relative to the **mero221**-based biosensor that has been used in other successful live cell studies (Fig. 6) (Machacek *et al.*, *Nature* 461:99-103 (2009); Nalbant *et al.*, *Science* 305:1615-1619 (2004)). Dissociation constant (K_d) values calculated for the binding of the meroCBD constructs to Cdc42

ranged from 144 ± 62 nM (**mero61**-CBD) to 186 ± 40 nM (**mero87**-CBD) (Fig. 7). These values are similar to that of the previously reported K_d for **mero221**-CBD (150 ± 50 nM) (Nalbant *et al.*, *Science* 305:1615-1619 (2004)) and are slightly higher than that reported for the unlabeled CBD fragment (77 ± 9 nM) (Rudolph *et al.*, *J. Biol. Chem.* 273:18067-18076 (1998)). This relatively modest difference indicates that the presence of the dye does not substantially perturb the overall binding affinity.

Table 6: Structure and photophysical properties of **mero221**.



Solvent	λ_{max}	M_{max}	ϵ	QY	$\epsilon \times QY$
Water	599	630	143000	0.004	600
MeOH	601	634	138000	0.01	1400
BuOH	607	639	150000	0.06	9000

[0289] To test the dyes in live cell imaging applications, Cdc42 biosensors were prepared with each of the four new conjugatable dyes as well as the previously reported **mero221**. For these biosensors, CBD was fused to Cerulean fluorescent protein for ratio image analysis and mutation F271C was included for conjugation of the dye (Hodgson *et al.*, *Methods Enzymol.* 406:140-156 (2006); Hodgson *et al.*, *Curr. Protoc. Cell Biol.* pp 1-26, Chapter 14, Unit 14.11 (2010)). The biosensors were microinjected into mouse embryonic fibroblasts (MEFs) and used to examine the Cdc42 activity in constitutive cell protrusions and retractions (Fig. 8). The biosensors based on **mero61**, **mero62**, and **mero87** each reported endogenous Cdc42 activation during protrusion at the cell's leading edge, consistent with previously reports of Cdc42 behavior (Fig. 8). Although the **mero60** biosensor had shown the greatest fluorescence change upon activation *in vitro*, its signal bleached too rapidly to provide Cdc42 data *in vivo* without using concentrations that clearly caused cell contraction and otherwise affected cell behavior.

[0290] For imaging, it is useful to define dye brightness as the product of quantum yield and extinction coefficient ($QY \times \epsilon$), as this gauges the total light output per unit illumination intensity. By this measure, all dyes except **mero62** showed maximum brightness values comparable to those of fluorophores frequently used as intracellular markers but which have little solvent-dependent fluorescence (e.g. Cy5, $QY \times \epsilon = 67500$ (Mujumdar *et al.*, *Bioconjugate Chem.* 4:105-111 (1993)), rhodamine = 61000 (Leytus *et al.*, *Biochem. J.* 209:299-307 (1983)), and eGFP = 33000 (Patterson *et al.*, *Biophys. J.* 73:2782-2790 (1997))).

[0291] It is difficult to use our solvent-dependent fluorescence data to predict how each dye will behave in a given application. The brightness of these dyes on proteins will depend on the specific protein environment, and optimization of positioning is usually required to place the fluorophore in an environment mimicking a hydrophobic solvent, where most dyes are at their brightest. The fluorescence response of merocyanines is complex, involving interplay of multiple factors including solvent viscosity, local solvent polarity, and hydrogen bonding of specific dye heteroatoms (Han *et al.*, *ChemPhysChem* 4:1084-1094 (2003); Liu *et al.*, *J. Phys. Chem. A* 108:3545-3555 (2004); Touthkine *et al.*, *J. Phys. Chem. A* 111:10849-10860 (2007)). Physical constants were determined for β -mercaptoethanol adducts of the dyes, to mimic cysteine conjugates and eliminate the quenching effects of iodine (Touthkine *et al.*, *J. Am. Chem. Soc.* 125:4132-4145 (2003)). However even these brightness values may not be as large as when the dye is in a very hydrophobic protein environment. The reactive dyes could not be dissolved in highly nonpolar solvents, but the parent compound I-SO has been reported to reach a brightness value ($QY \times \epsilon$) of 123,000 in octanol, which is among the largest values reported for organic dyes (Lavis *et al.*, *ACS Chem. Biol.* 3:142-155 (2008)).

[0292] Although the solvent dependence studies of the screening library can be useful in guiding the selection of fluorophores likely to show a strong response to changes in protein environment, only direct testing can ultimately determine which dye will be best for any given biosensor application. The dyes showed different trends in their response to solvent, implying that each may show unique behavior on proteins. **Mero62** achieved its maximum brightness in butanol, unlike all the other reactive dyes, which achieved maximum brightness in DMSO. Fluorescence in water

was not measured in our initial screening assay because side chains for water solubilization had not been added to the dyes. For the reactive dye derivatives, $QY \times \epsilon$ values in water were very low, consistent with previous studies indicating that biosensor response is due to increased shielding from water when biosensors interact with their targets (Han *et al.*, *ChemPhysChem* 4:1084-1094 (2003); Liu *et al.*, *J. Phys. Chem. A* 108:3545-3555 (2004); Touthkine *et al.*, *J. Phys. Chem. A* 111:10849-10860 (2007)). For the nonreactive versions of compounds that were later converted to reactive forms, solvent-dependent differences in brightness ranged from three to twenty fold. For the reactive versions of these same dyes, with values in water included, brightness varied from seven to twenty-nine fold. The brightness of **mero60** was so low in water that the extent of increase could not be precisely quantified. However, **mero60** was among the brightest dyes in organic solvents, indicating a potential for very large changes when it is used in favorable protein environments. None of the dyes showed a significant pH-dependent change in fluorescence intensity when tested from pH 5.0 to 8.0 for 6 hours (Fig. 9).

[0293] There were several parent fluorophores that showed not only changes in fluorescence intensity, but also changes in excitation and emission maxima. Such fluorophores are very valuable for ratio imaging applications, as they can be used without adding a second fluorophore to the biosensor; however these dyes were not pursued because of their poor photostability. Additional structural modifications aimed at reducing the susceptibility of these dyes to photodegradative attack by singlet oxygen may make them useful for protein applications in the future (Griebel, *Ber. Bunsenges. Phys. Chem.* 84:84-91 (1980)).

[0294] The potential of the new dyes in biosensor applications was demonstrated by making new versions of dye-based Cdc42 biosensors. When substituted for the dye **mero221**, which was utilized in a previously published biosensor that had proven successful *in vivo*, each of the new dyes produced a several fold improvement in brightness or change in fluorescence intensity upon protein binding. Dyes **mero61**, **mero62**, and **mero87** proved to be superior to the original **mero221**, but **mero60** was difficult to use because of its more rapid photobleaching. This dye underwent the largest fluorescence change upon protein activation, so still may prove valuable in

screening assays and other applications not requiring repeated measurements over time.

[0295] In conclusion, the systematic variation of donor and acceptor groups on merocyanine fluorophores, coupled with the synthesis of reactive and water soluble derivatives, has led to a set of new dyes with proven utility for biosensor applications requiring solvent sensitive fluorescence. The combined properties of these dyes should enable experimentalists to develop biosensors that can effectively report endogenous protein activity at low intracellular biosensor concentrations, thereby reducing perturbation of normal cell physiology. Biosensors based on these dyes may be used to enhance the sensitivity of biophysical measurements across a range of potential applications in which a dye is directly conjugated to a protein or peptide to report conformational changes or ligand interactions.

Example 4

Experimental Procedures

[0296] **Cell Culture.** NIH 3T3 mouse embryonic fibroblasts (MEF) were maintained in 10% CO₂ at 37°C in Dulbecco's modified Eagle's medium (DMEM, Cellgro) with 10% fetal bovine serum (HyClone, Thermo Scientific) and 2 mM GlutaMax (Gibco, Life Technologies). The cells were plated on coverslips coated with fibronectin (Sigma-Aldrich) overnight.

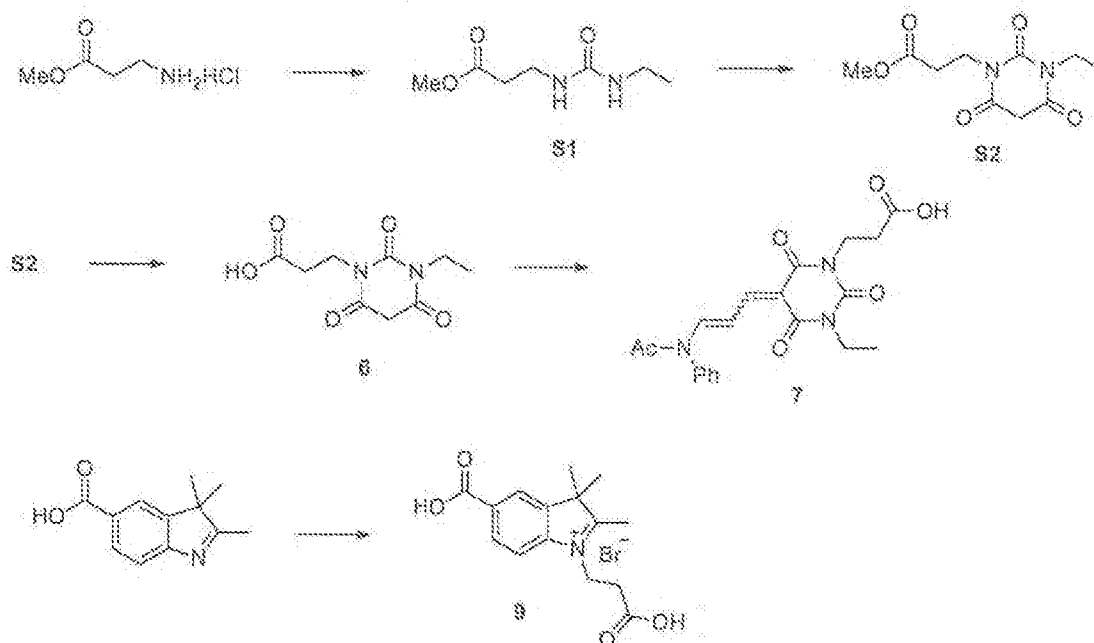
[0297] **Image Acquisition.** Live cell imaging was carried out on an Olympus IX81 microscope with a UPLFLN 40X oil objective (NA 1.3) and mercury lamp excitation (103W HBO bulb). Filters were 545/50 for excitation and 630/45 for emission. Excitation was through a ND 2.0 (1%T) filter. Images were acquired with a Coolsnap ES2 camera (Photometrics) with a Sony 6.45 x 6.45 μ M/pixel chip using 2 x 2 binning. All image acquisition, processing, and analysis was carried out with Metamorph software. Images were shade corrected, background subtracted, and linearly contrast stretched to reveal subcellular features as previously described.^[1]

[0298] **Chemistry.** β -Alanine methyl ester hydrochloride (Sigma-Aldrich) and malonaldehyde dianilide hydrochloride (TCI America) were purchased and used as received. The following compounds were prepared according to literature procedures: 1,3-diethylbarbituric acid, 1,2,3,3-tetramethylindolium iodide, 1,3-

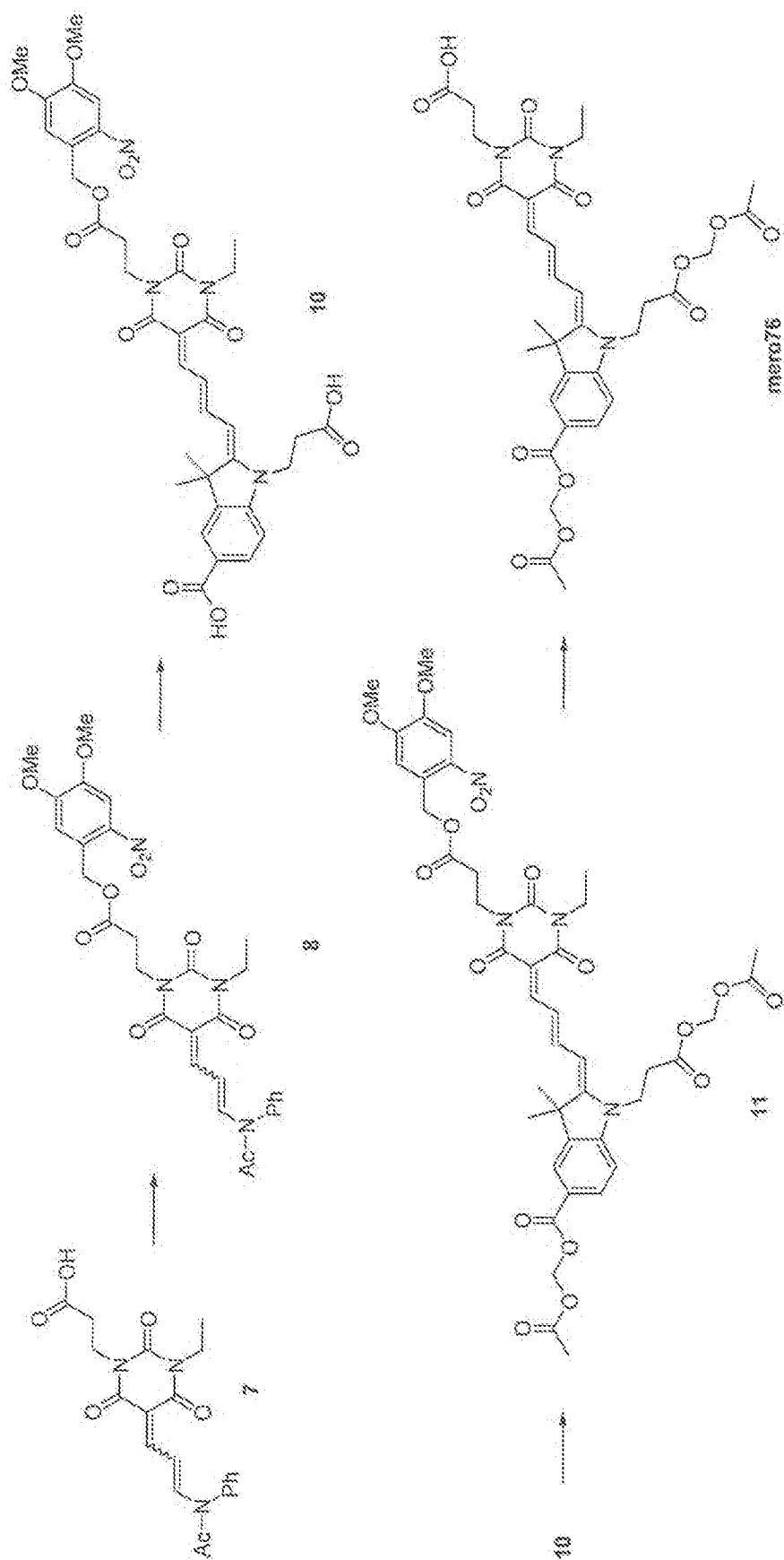
diethyl-5-(3-methoxyallylidene)pyrimidine-2,4,6(1*H*,3*H*,5*H*)-trione, 2,3,3-trimethyl-5-carboxyindolenine and 5-carboxy-1,2,3,3-tetramethyl-3*H*-indolium iodide, 1-(2-carboxyethyl)-2,3,3-trimethyl-3*H*-indolium bromide, 4,5-dimethoxy-2-nitrobenzyl alcohol, and 2-azidoethylamine. Reaction progress was monitored via thin-layer chromatography (TLC) on EM Science pre-coated glass-backed plates (silica gel 60 Å F₂₅₄, 0.25 mm thickness). Flash chromatography was carried out with silica gel 60 Å (230 - 400 mesh) and automated chromatography was performed on an Isco Combiflash Companion. Unless otherwise stated, organic extracts were dried over commercially available magnesium sulfate and the solvents were removed by rotary evaporation. ¹H and ¹³C NMR spectra were recorded on either a Varian Inova 400 MHz in deuterated chloroform (CDCl₃) unless otherwise noted and referenced to the residual solvent peak (for CDCl₃ = ¹H δ 7.26 ppm, ¹³C δ 77.23 ppm; for DMSO-d₆ = ¹H δ 2.50 ppm, ¹³C δ 39.51 ppm). Mass spectra were obtained on an Hewlett-Packard 1100 high-performance liquid chromatograph equipped with a 1100 mass-selective detector (MS-ESI). UV-visible spectra were obtained with a Hewlett-Packard 8453 diode array spectrophotometer. Emission and excitation spectra were obtained using a Spex Fluorolog 2 spectrofluorometer at 23°C. All operations with dyes were performed under dim light. Photoreactions were carried out using a Rayonet RMR-600 equipped with 350 nm bulbs. High performance liquid chromatography was performed on a Shimadzu Prominence system with a Vydac 218TP152022 C₁₈ column (15-20 μm particle size, 300-Å, 250 X 22 mm) in the water/methanol or acetonitrile/TFA mixture: Solvent A (water 95 parts, organic solvent 5 parts, TFA 0.05%), Solvent B (water 5 parts, organic solvent 95 parts, TFA 0.05 %). Standard gradient for separation of final products is 10% Solvent B, 30 min 90% solvent B, total run 45 min.

Example 5

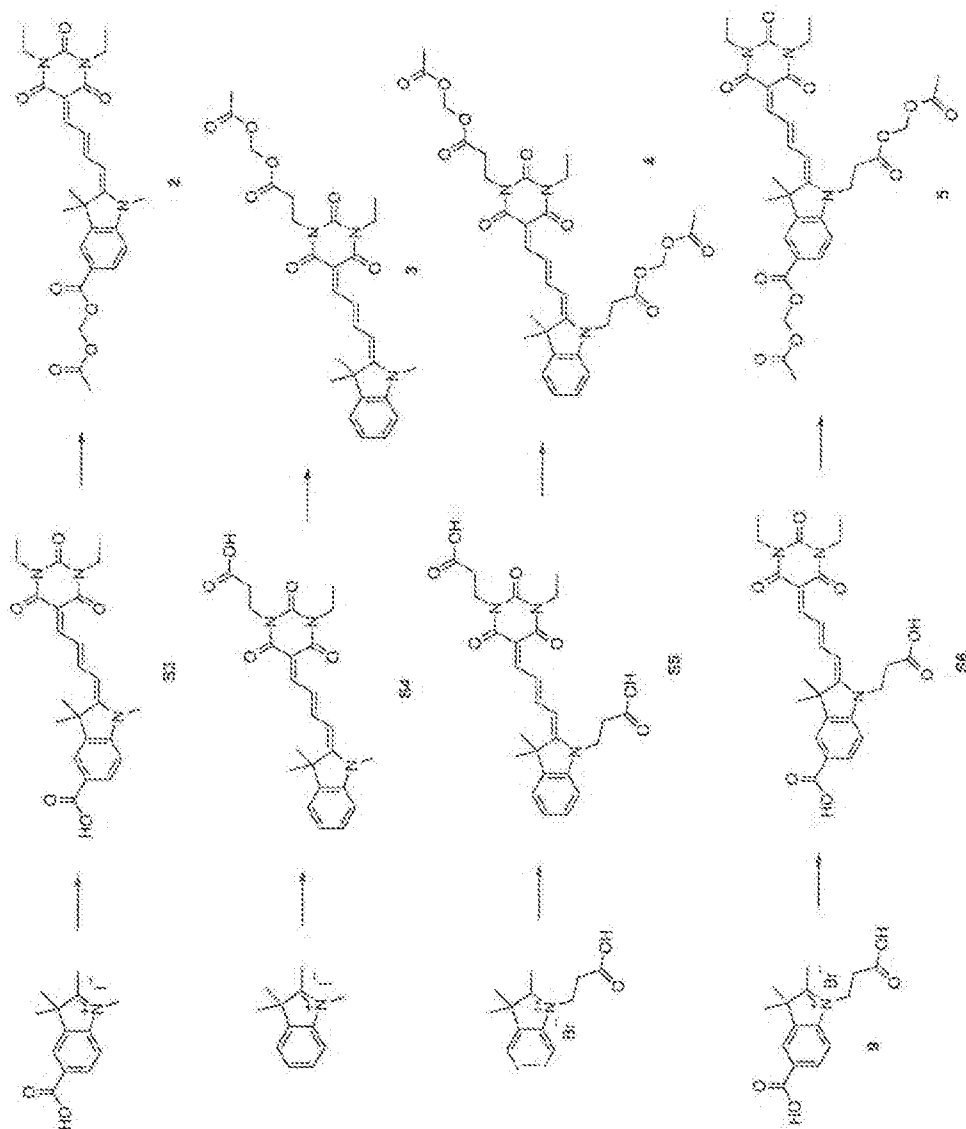
Synthesis of Compounds



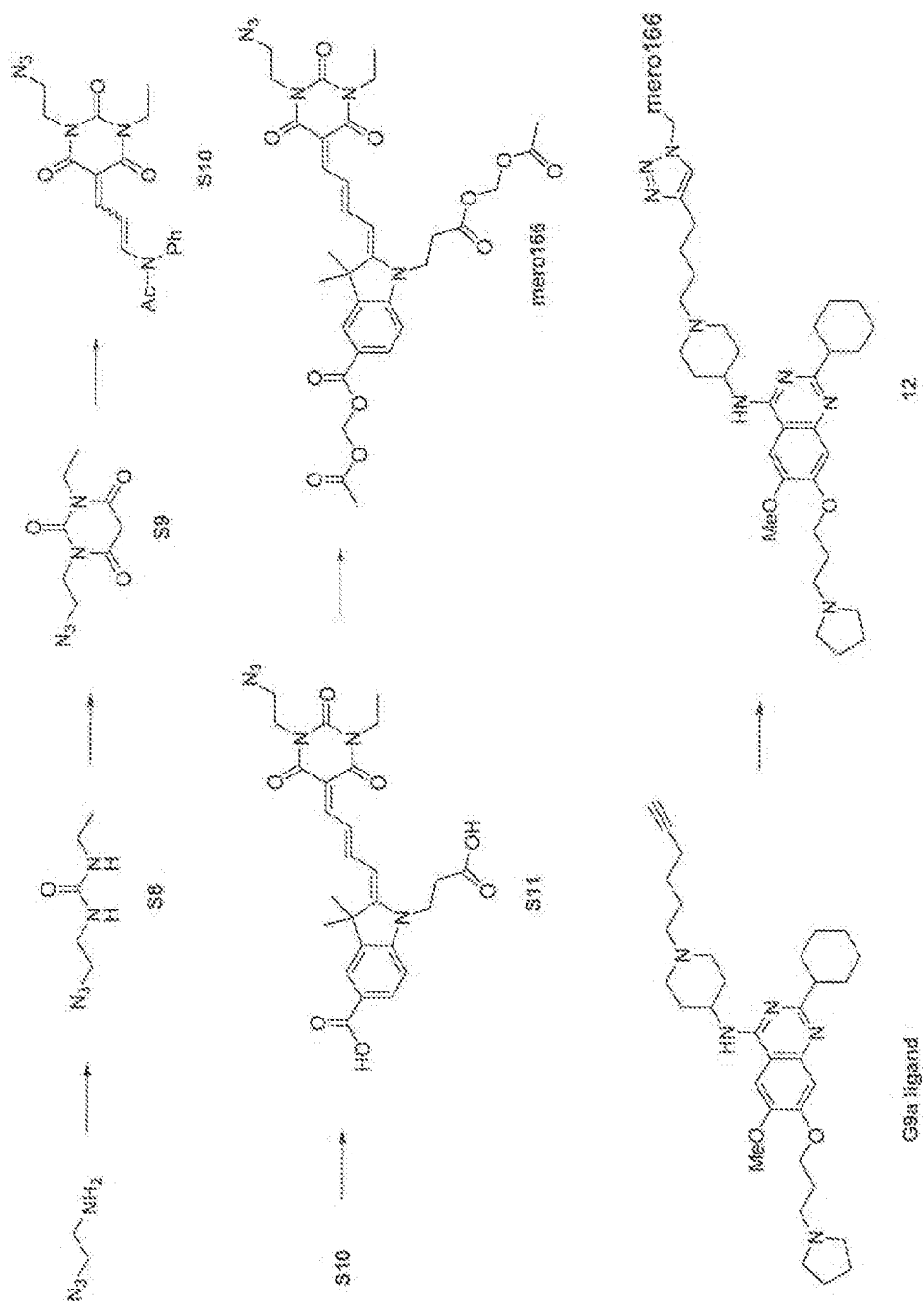
Scheme 2. Synthesis of key intermediates.



Scheme 3. Synthesis of mero76.



Scheme 4. Synthesis of structure permeability screening compound set.



Scheme 3. Synthesis of mero166 and G9a ligand --- mero166 conjugate (12).

[0299] **Methyl 3-(3-ethylureido)propanoate (S1).** β -Alanine methyl ester hydrochloride (34.9 g, 250 mmol) was added to a flame dried flask. The flask was sealed, evacuated, and flushed with argon and CH_2Cl_2 (350 mL) was added via cannula. Triethylamine (38.3 mL, 275 mmol) was added and the reaction was stirred at room temp for 45 min. The flask was cooled in an ice bath and ethyl isocyanate (20.8 mL, 262.5 mmol) was added quickly dropwise. The reaction was allowed to gradually equilibrate to room temp. After stirring for 16 h, the reaction was quenched with half-saturated aq. NH_4Cl (400 mL). The organic layer was separated and the aq. layer was further extracted with CH_2Cl_2 (200 mL X 2). The organic layers were combined, washed with water and brine, dried with MgSO_4 , filtered, and concentrated. The resulting solid was further dried under high vacuum to give 25.2 g (58%) white solid suitable for use without further purification. ^1H NMR (400 MHz, CDCl_3) δ 5.24 (bs, 1H), 4.82 (bs, 1H), 3.66 (s, 3H), 3.44 (q, $J = 6.0$ Hz, 2H), 3.20 – 3.12 (m, 2H), 2.52 (t, $J = 6.0$ Hz, 2H), 1.10 (t, $J = 7.2$ Hz, 3H). ^{13}C NMR (100 MHz, CDCl_3) δ 173.6, 158.4, 51.9, 36.0, 35.4, 34.9, 15.6. MS-ESI m/z 197.0 ($[\text{M} + \text{Na}]^+$ requires 197.1).

[0300] **Methyl 3-(3-ethyl-2,4,6-trioxotetrahydropyrimidin-1(2H)-yl)propanoate (S2).** Compound S1 (24.9 g, 143 mmol) and malonic acid (15.6 g, 150 mmol) were suspended in glacial AcOH (200 mL). The mixture was heated to 60 °C and stirred at that temp for 10 min. Ac_2O (53.9 mL, 572 mmol) was then added in bulk and the reaction was heated further to 100 °C and stirred for 1.5 h. The reaction mixture was concentrated and set at room temp open to air. The resulting crystals were re-crystallized from 3:1 hexanes/EtOAc (50 mL). The re-crystallized solid was filtered and washed with ice cold Et_2O (100 mL X 2) to give 23.9 g (69%) white solid suitable for use in the next step. ^1H NMR (400 MHz, CDCl_3) δ 4.18 (t, $J = 7.2$, 2H), 3.92 (q, $J = 7.2$ Hz, 2H), 3.67 (s, 3H), 3.65 (s, 2H), 2.62 (t, $J = 7.2$ Hz, 2H), 1.20 (t, $J = 7.2$ Hz, 3H). ^{13}C NMR (100 MHz, CDCl_3) δ 171.5, 164.7, 164.4, 151.2, 52.1, 39.9, 37.9, 37.6, 32.5, 13.4. MS-ESI m/z 240.9 ($[\text{M} - \text{H}]^-$ requires 241.1).

[0301] **3-(3-Ethyl-2,4,6-trioxotetrahydropyrimidin-1(2H)-yl)propanoic acid (6).** Compound S2 (8.30 g, 34.3 mmol) was diluted in THF (70 mL) and DI (35 mL) was added. $\text{LiOH} \cdot \text{H}_2\text{O}$ (5.75 g, 4.00 eq) was added and the reaction was stirred at

room temp for 16 h. The flask was set in an ice bath and the reaction was quenched with 6 N aq. HCl (27 mL) to a pH of ~4. The reaction mixture was concentrated and the resulting precipitate was filtered, washed with DI (2x) and ice chilled Et₂O (2x). The product was dried under high vacuum to give 7.30 g (93%) white solid that required no further purification. ¹H NMR (400 MHz, DMSO-*d*₆) δ 12.36 (bs, 1H), 3.93, (t, *J* = 7.8 Hz, 2H), 3.75 (q, *J* = 7.2 Hz, 2H), 3.71 (s, 2H), 2.45 (t, *J* = 7.8 Hz, 2H), 1.08 (t, *J* = 7.2 Hz, 3H). ¹³C NMR (100 MHz, DMSO-*d*₆) δ 172.2, 165.6, 165.4, 151.4, 40.1, 36.7, 36.1, 32.0, 12.9. MS-ESI *m/z* 226.8 ([*M* - H]⁻ requires 227.1).

[0302] 3-(3-ethyl-2,4,6-trioxo-5-((*E*)-3-(*N*-

phenylacetamido)allylidene)tetrahydropyrimidin-1(2H)-yl)propanoic acid (7).

Compound 6 (0.500 g, 2.19 mmol), *N*-(3-(phenylamino)-2-propenylidene)aniline hydrochloride (0.624 g, 2.41 mmol), and sodium acetate (0.198 g, 2.41 mmol) were added together to an oven-dried flask. The flask was sealed, evacuated, and argon flushed and acetic anhydride (2.2 mL) was added. The flask was added to a pre-heated oil bath at 135 °C and stirred for 1 h. The reaction mixture was cooled to room temp and diluted with dichloromethane. The organic layer was washed with saturated sodium bicarbonate, saturated ammonium chloride, and brine, and dried with sodium sulfate. The organic layer was filtered, concentrated, and loaded neat onto a 120 g silica column. The column was eluted with 0 – 75% EtOAc in hexanes over 25 min. Main product fractions were combined and concentrated to give 0.387 g (44%) yellow solid as an approximate 1:1 mixture of *E* and *Z* isomers. Isomer A: ¹H NMR (400 MHz, DMSO-*d*₆) δ 12.28 (bs, 1H), 8.67 (d, *J* = 13.6 Hz, 1H), 8.23 (d, *J* = 12.4 Hz, 1H), 7.65 – 7.53 (m, 3H), 7.44 – 7.36 (m, 2H), 6.70 (t, *J* = 13.0 Hz, 1H), 3.86 (t, *J* = 7.8 Hz, 2H), 3.80 (q, *J* = 7.2 Hz, 2H), 2.37 (t, *J* = 7.8 Hz, 2H), 2.09 (s, 3H), 1.08 (t, *J* = 7.0 Hz, 3H). ¹³C NMR (100 MHz, DMSO-*d*₆) δ 172.3, 169.9, 161.5, 161.0, 157.2, 150.3, 137.8, 130.3, 129.5, 128.2, 110.5, 109.0, 37.0, 36.2, 35.5, 32.2, 23.1, 13.0. Isomer B: ¹H NMR (400 MHz, DMSO-*d*₆) δ 12.28 (bs, 1H), 8.67 (d, *J* = 13.6 Hz, 1H), 8.22 (d, *J* = 12.4 Hz, 1H), 7.65 – 7.53 (m, 3H), 7.44 – 7.36 (m, 2H), 6.71 (t, *J* = 13.0 Hz, 1H), 3.99 (t, *J* = 7.8 Hz, 2H), 3.68 (q, *J* = 6.8 Hz, 2H), 2.46 (t, *J* = 7.8 Hz, 2H), 2.09 (s, 3H), 0.99 (t, *J* = 7.0 Hz, 3H). ¹³C NMR (100 MHz, DMSO-*d*₆) δ 172.2, 169.9, 161.4, 161.1, 157.2, 150.3, 137.8, 130.3, 129.6, 128.2, 110.5, 109.0, 37.0, 36.3, 35.5, 32.1, 23.1, 13.1. MS-ESI *m/z* 400.0 ([*M* + H]⁺ requires 400.2).

[0303] **5-Carboxy-1-(2-carboxyethyl)-2,3,3-trimethyl-3*H*-indolium bromide (9).** 2,3,3-Trimethyl-5-carboxyindolenine (1.00 g, 4.92 mmol) and bromopropionic acid (0.828 g, 5.41 mmol) were added to a flame-dried 2-necked flask fitted with a condenser and rubber septum. The flask was sealed, evacuated, and argon flushed, and *o*-dichlorobenzene (10 mL) was added. The suspension was heated to 150 °C for 16 h under argon. The reaction mixture was cooled gradually to 0 °C and filtered. The precipitate was washed with chloroform (3 x 20 mL) and dried under high vacuum to give 1.167 g pale red solid that was carried on to the next step without further purification. MS-ESI *m/z* 276.2 ($[M - Br]^+$ requires 276.1).

[0304] **(*E*)-2-((*E*)-4-(1,3-diethyl-2,4,6-trioxotetrahydropyrimidin-5(2*H*)-ylidene)but-2-en-1-ylidene)-1,3,3-trimethylindoline-5-carboxylic acid (S3).** 1,3-Diethyl-5-(3-methoxyallylidene)pyrimidine-2,4,6(1*H*,3*H*,5*H*)-trione (0.126 g, 0.500 mmol) was added to a flame dried flask fitted with a condensor. A 1:1 mixture of MeOH and CHCl₃ (5 mL) was added and the flask was flushed with argon. 5-Carboxy-1,2,3,3-tetramethyl-3*H*-indolium iodide (0.207 g, 0.600 mmol) was added, followed quickly by NaOAc (0.049 g, 0.600 mmol). The reaction was allowed to stir for 16 h at reflux. Solvent was removed and the residue was redissolved in CHCl₃ and loaded onto silica. The silica cake was eluted on a 12 g silica column using 0 – 5% MeOH in CH₂Cl₂. The product was isolated as 0.075 g (34%) purple powdery solid. ¹H NMR (400 MHz, DMSO-*d*₆) δ 12.75 (bs, 1H), 8.21 (d, *J* = 12.8 Hz, 1H), 8.15 (t, *J* = 13.2 Hz, 1H), 7.98 (s, 1H), 7.93 (d, *J* = 8.4 Hz, 1H), 7.78 (t, *J* = 13.2 Hz, 1H), 7.27 (d, *J* = 8.4 Hz, 1H), 6.14 (d, *J* = 13.2 Hz, 1H), 3.90 – 3.80 (m, 4H), 3.47 (s, 3H), 1.64 (s, 6H), 1.14 – 1.07 (m, 6H). MS-ESI *m/z* 438.2 ($[M + H]^+$ requires 438.2).

[0305] **(*E*)-acetoxymethyl-2-((*E*)-4-(1,3-diethyl-2,4,6-trioxotetrahydropyrimidin-5(2*H*)-ylidene)but-2-en-1-ylidene)-1,3,3-trimethylindoline-5-carboxylate (2).** Compound S3 (50.0 mg, 0.114 mmol) was diluted in 2.5 mL anhydrous acetonitrile and the flask was flushed with argon. Diisopropylethylamine (30.0 μL, 0.171 mmol) was added, followed by addition of bromomethylacetate (13.0 μL, 0.137 mmol) and the reaction was stirred at room temperature for 16 h. The reaction was concentrated and the residue was redissolved in a minimum amount of dichloromethane. The crude mixture was eluted on a 12 g silica column with 0 – 50% hexanes in EtOAc. The product was isolated as 52 mg

(89%) purple sticky solid. ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 8.23 (d, $J = 12.8$ Hz, 1H), 8.15 (t, $J = 13.2$ Hz, 1H), 8.00 (d, $J = 1.6$ Hz, 1H), 7.95 (dd, $J = 8.0, 1.6$ Hz, 1H), 7.78 (t, $J = 12.8$ Hz, 1H), 7.28 (d, $J = 8.4$ Hz, 1H), 6.14 (d, $J = 13.2$ Hz, 1H), 5.92 (s, 2H), 3.90 – 3.80 (m, 4H), 3.46 (s, 3H), 2.11 (s, 3H), 1.65 (s, 6H), 1.14 – 1.07 (m, 6H). MS-ESI m/z 510.2 ($[\text{M} + \text{H}]^+$ requires 510.2).

[0306] 3-((E)-3-ethyl-2,4,6-trioxo-5-((2E,4E)-4-(1,3,3-trimethylindolin-2-ylidene)but-2-en-1-ylidene)tetrahydropyrimidin-1(2H)-yl)propanoic acid (S4). 1,2,3,3-Tetramethylindolium iodide (0.151 g, 0.500 mmol) and compound 7 (0.300 g, 0.750 mmol) were diluted in 1:1 $\text{CHCl}_3/\text{MeOH}$ (5 mL) and NaOAc (0.051 g, 0.625 mmol) was added in bulk. The reaction was stirred at room temp for 16 h and concentrated. The concentrate was prepared as a silica cake and eluted on a 24 g silica column with 0 – 8% MeOH in CH_2Cl_2 over 25 min. Pure product containing fractions were combined and concentrated to give 0.123 g (56%) dark blue solid. ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 12.27 (bs, 1H), 8.17 – 8.03 (m, 2H), 7.72 (dt, $J = 13.2, 4.0$ Hz, 1H), 7.51 (d, $J = 7.2$ Hz, 1H), 7.35 (dt, $J = 7.2, 1.2$ Hz, 1H), 7.30 – 7.25 (m, 1H), 7.16 (dt, $J = 7.2, 1.2$ Hz, 1H), 6.16 (d, $J = 13.6$ Hz, 1H), 4.03 (t, $J = 7.6$ Hz, 2H), 3.84 (q, $J = 6.8$ Hz, 2H), 3.51 (s, 3H), 2.47 (t, $J = 7.6$ Hz, 2H), 1.62 (s, 6H), 1.10 (t, $J = 6.8$ Hz, 3H). MS-ESI m/z 438.1 ($[\text{M} + \text{H}]^+$ requires 438.2).

[0307] Acetoxymethyl 3-((E)-3-ethyl-2,4,6-trioxo-5-((2E,4E)-4-(1,3,3-trimethylindolin-2-ylidene)but-2-en-1-ylidene)tetrahydropyrimidin-1(2H)-yl)propanoate (3). Compound S4 (33.8 mg, 0.077 mmol) was diluted in 2 mL anhydrous ACN and the flask was flushed with argon. Diisopropylethylamine (20.2 μL , 0.116 mmol) was added, followed by addition of bromomethylacetate (11.4 μL , 0.116 mmol) and the reaction was stirred at room temperature for 16 h. The reaction was concentrated and the residue was redissolved in a minimum amount of CH_2Cl_2 . The crude mixture was eluted on a 12 g silica column with 0 – 50% hexanes in EtOAc. The product was isolated as 33.8 mg (93%) dark violet solid. ^1H NMR (400 MHz, $\text{DMSO}-d_6$, 50 $^\circ\text{C}$) δ 8.17 – 8.03 (m, 2H), 7.80 – 7.65 (m, 1H), 7.50 (d, $J = 7.2$ Hz, 1H), 7.35 (dt, $J = 7.2, 1.2$ Hz, 1H), 7.26 (d, $J = 8.0$ Hz, 1H), 7.16 (dt, $J = 7.2, 1.2$ Hz, 1H), 6.15 (d, $J = 14.0$ Hz, 1H), 5.66 (s, 2H), 4.09 (t, $J = 7.4$ Hz, 2H), 3.85 (q, $J = 7.2$ Hz, 2H), 3.52 (s, 3H), 2.63 (t, $J = 7.2$ Hz, 2H), 2.07 (s, 3H), 1.63 (s, 6H), 1.11 (t, $J = 7.0$ Hz, 3H). MS-ESI m/z 510.2 ($[\text{M} + \text{H}]^+$ requires 510.2).

[0308] 3-((E)-5-((2E,4E)-4-(1-(2-carboxyethyl)-3,3-dimethylindolin-2-ylidene)but-2-en-1-ylidene)-3-ethyl-2,4,6-trioxotetrahydropyrimidin-1(2H)-yl)propanoic acid (S5). 1-(2-Carboxyethyl)-2,3,3-trimethyl-3H-indolium bromide (47 mg, 0.150 mmol) compound 7 (90 mg, 0.225 mmol) were diluted in 1:1 $\text{CHCl}_3/\text{MeOH}$ (1.5 mL) and NaOAc (18 mg, 0.225 mmol) was added in bulk. The reaction was stirred at room temp for 16 h and concentrated. The residue was diluted in CH_2Cl_2 , prepared as a silica cake, and eluted on a 24 g silica column with 0 – 8% MeOH in CH_2Cl_2 over 25 min. Pure product containing fractions were combined and concentrated to give 27 mg (36%) dark blue solid. ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 12.35 (bs, 1H), 8.19 – 8.08 (m, 2H), 7.74 (dt, $J = 13.1, 4.3$ Hz, 1H), 7.50 (dd, $J = 7.5, 1.2$ Hz, 1H), 7.33 (dt, $J = 7.4, 1.2$ Hz, 1H), 7.26 (d, $J = 8.0$ Hz, 1H), 7.14 (dt, $J = 7.2, 1.2$ Hz, 1H), 6.23 (d, $J = 13.3$ Hz, 1H), 4.23 (t, $J = 7.2$ Hz, 2H), 4.03 (t, $J = 7.8$ Hz, 2H), 3.84 (q, $J = 7.2$ Hz, 2H), 2.66 (t, $J = 7.2$ Hz, 2H), 2.47 (t, $J = 7.6$ Hz, 2H), 1.62 (s, 6H), 1.10 (dt, $J = 6.8, 2.0$ Hz, 3H). MS-ESI m/z 496.2 ($[\text{M} + \text{H}]^+$ requires 496.2).

[0309] Acetoxymethyl 3-((E)-5-((2E,4E)-4-(1-(3-(acetoxymethoxy)-3-oxopropyl)-3,3-dimethylindolin-2-ylidene)but-2-en-1-ylidene)-3-ethyl-2,4,6-trioxotetrahydropyrimidin-1(2H)-yl)propanoate (4). Compound S5 (14 mg, 0.028 mmol) was diluted in 1 mL anhydrous ACN and the flask was flushed with argon. Diisopropylethylamine (20 μL , 0.113 mmol) was added, followed after 5 min by addition of bromomethylacetate (11 μL , 0.113 mmol) and the reaction was stirred at room temperature for 16 h. The reaction was concentrated and the residue was redissolved in a minimum amount of CH_2Cl_2 . The crude mixture was eluted on a 12 g silica column with 0 – 3% MeOH in CH_2Cl_2 . The product was isolated as 16 mg (89%) dark violet solid. ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 8.20 – 8.10 (m, 2H), 7.75 (dt, $J = 13.0, 9.0$ Hz, 1H), 7.50 (d, $J = 7.6$ Hz, 1H), 7.33 (t, $J = 7.6$ Hz, 1H), 7.25 (d, $J = 8.0$ Hz, 1H), 7.14 (t, $J = 7.4$ Hz, 1H), 6.23 (dd, $J = 13.2, 5.2$ Hz, 1H), 5.66 (s, 2H), 5.63 (s, 2H), 4.27 (t, $J = 6.8$ Hz, 2H), 4.07 (t, $J = 6.8$ Hz, 2H), 3.84 (q, $J = 7.2$ Hz, 2H), 2.82 (t, $J = 7.2$ Hz, 2H), 2.62 (t, $J = 7.2$ Hz, 2H), 2.08 (s, 3H), 2.02 (s, 3H), 1.62 (s, 6H), 1.10 (t, $J = 7.2$ Hz, 3H). MS-ESI m/z 662.2 ($[\text{M} + \text{Na}]^+$ requires 662.2).

[0310] (E)-1-(2-carboxyethyl)-2-((E)-4-(1,3-diethyl-2,4,6-trioxotetrahydropyrimidin-5(2H)-ylidene)but-2-en-1-ylidene)-3,3-dimethylindoline-5-carboxylic acid (S6). Compound 9 (0.178 g, 0.500 mmol) and

NaOAc (0.051 g, 0.625 mmol) were diluted in 1:1 MeOH/CHCl₃ (3.0 mL) and 1,3-diethyl-5-(3-methoxyallylidene)pyrimidine-2,4,6(1*H*,3*H*,5*H*)-trione (0.173 g, 0.650 mmol) was added. The flask was fitted with a condenser and flushed with argon and the reaction was heated to reflux for 16 h. The reaction was concentrated and the residue was diluted in CH₂Cl₂, prepared as a silica cake, and eluted on a 24 g silica column with 0 – 8% MeOH in CH₂Cl₂ over 25 min. Pure product containing fractions were combined and concentrated to give 53 mg (21%) dark blue solid. ¹H NMR (400 MHz, DMSO-*d*₆) δ 12.61 (bs, 2H), 8.22 (d, *J* = 13.2 Hz, 1H), 8.13 (t, *J* = 13.4 Hz, 1H), 7.97 (s, 1H), 7.91 (d, *J* = 8.4 Hz, 1H), 7.79 (t, *J* = 13.0 Hz, 1H), 7.26 (d, *J* = 8.4 Hz, 1H), 4.20 (t, *J* = 6.8 Hz, 2H), 3.90 – 3.80 (m, 2H), 2.65 (t, *J* = 6.8 Hz, 2H), 1.63 (s, 6H), 1.11 (t, *J* = 6.8 Hz, 6H). ¹³C NMR (100 MHz, DMSO-*d*₆) δ 172.1, 168.9, 167.0, 162.0, 161.3, 156.4, 155.7, 150.6, 146.1, 140.2, 130.4, 125.0, 123.0, 122.6, 109.3, 101.5, 47.3, 35.9, 35.3, 31.2, 27.4, 13.3. MS-ESI *m/z* 493.9 ([*M* - H]⁻ requires 494.2).

[0311] (E)-Acetoxymethyl 1-(3-(acetoxymethoxy)-3-oxopropyl)-2-((E)-4-(1,3-diethyl-2,4,6-trioxotetrahydropyrimidin-5(2*H*)-ylidene)but-2-en-1-ylidene)-3,3-dimethylindoline-5-carboxylate (5). Compound S6 (36 mg, 0.073 mmol) was diluted in 2.0 mL anhydrous ACN and the flask was flushed with argon.

Diisopropylethylamine (51 μL, 0.291 mmol) was added, followed after 5 min by addition of bromomethylacetate (28 μL, 0.291 mmol) and the reaction was stirred at room temperature for 3 h. The reaction was concentrated and the residue was redissolved in a minimum amount of CH₂Cl₂. The crude mixture was eluted on a 12 g silica column with 0 – 3% MeOH in CH₂Cl₂. The product was isolated as 31 mg (67%) dark violet solid. ¹H NMR (400 MHz, DMSO-*d*₆) δ 8.25 (d, *J* = 12.8 Hz, 1H), 8.14 (t, *J* = 13.2 Hz, 1H), 7.99 (d, *J* = 1.6 Hz, 1H), 7.93 (dd, *J* = 8.4, 1.6 Hz, 1H), 7.81 (t, *J* = 13.2 Hz, 1H), 7.28 (d, *J* = 8.4 Hz, 1H), 6.22 (d, *J* = 13.2 Hz, 1H), 5.92 (s, 2H), 5.62 (s, 2H), 4.23 (t, *J* = 6.8 Hz, 2H), 3.90 – 3.80 (m, 4H), 2.81 (t, *J* = 6.8 Hz, 2H), 2.10 (s, 3H), 2.01 (s, 3H), 1.64 (s, 6H), 1.15 – 1.08 (m, 6H). ¹³C NMR (100 MHz, DMSO-*d*₆) δ 169.6, 169.4, 169.1, 167.8, 164.2, 162.0, 161.3, 156.6, 155.2, 150.5, 147.2, 140.4, 130.9, 123.1, 122.2, 109.4, 105.9, 101.7, 79.6, 79.2, 47.0, 38.4, 35.9, 35.3, 30.8, 30.7, 27.3, 20.5, 20.4, 13.3. MS-ESI *m/z* 640.1 ([*M* + H]⁺ requires 640.3).

[0312] 4,5-dimethoxy-2-nitrobenzyl 3-((E)-3-ethyl-2,4,6-trioxo-5-((E)-3-(N-phenylacetamido)allylidene) tetrahydropyrimidin-1(2H)-yl)propanoate (8).

Compound 7 (0.399 g, 1.00 mmol), 4,5-dimethoxy-2-nitrobenzyl alcohol (0.213 g, 1.00 mmol), and DMAP (0.012 g, 0.100 mmol) were added to a flame-dried flask. The flask was sealed, evacuated, and argon flushed. Dichloromethane (12.5 mL) was added, followed by *N,N'*-dicyclohexylcarbodiimide. The reaction was stirred at room temperature under argon for 16 h. The reaction mixture was then filtered through a short pad of CELITE™. The filtrate was concentrated and filtered through a 0.45 µM nylon filter. This filtrate was loaded onto a 40 g silica column and the column was eluted with 0 – 50% EtOAc in hexanes. The product was isolated as 0.538 g (82%) dark orange foamy solid. Mixture of *E/Z* isomers (1:1 ratio). Isomer A: ¹H NMR (400 MHz, DMSO-*d*₆) δ 8.50 (d, *J* = 13.6 Hz, 1H), 8.16 (d, *J* = 12.8 Hz, 1H), 7.71 (s, 1H), 7.64 – 7.49 (m, 3H), 7.25 – 7.19 (m, 2H), 7.01 (s, 1H), 6.80 (t, *J* = 12.6 Hz, 1H), 5.45 (s, 2H), 4.17 (t, *J* = 7.4 Hz, 2H), 3.98 – 3.91 (m, 2H), 3.97 (s, 3H), 3.93 (s, 3H), 2.67 (t, *J* = 7.4 Hz, 2H), 2.01 (s, 3H), 1.18 (t, *J* = 7.0 Hz, 3H). Isomer B: ¹H NMR (400 MHz, DMSO-*d*₆) δ 8.50 (d, *J* = 13.6 Hz, 1H), 8.13 (d, *J* = 12.8 Hz, 1H), 7.70 (s, 1H), 7.64 – 7.49 (m, 3H), 7.25 – 7.19 (m, 2H), 7.05 (s, 1H), 6.83 (t, *J* = 12.6 Hz, 1H), 5.50 (s, 2H), 4.28 (t, *J* = 7.4 Hz, 2H), 3.98 – 3.91 (m, 2H), 3.99 (s, 3H), 3.95 (s, 3H), 2.76 (t, *J* = 7.4 Hz, 2H), 2.01 (s, 3H), 1.10 (t, *J* = 7.0 Hz, 3H). MS-ESI *m/z* 595.1 ([*M* + *H*]⁺ requires 595.2).

[0313] (E)-1-(2-carboxyethyl)-2-((2E,4E)-4-(1-(3-((4,5-dimethoxy-2-nitrobenzyl)oxy)-3-oxopropyl)-3-ethyl-2,4,6-trioxotetrahydropyrimidin-5(2H)-ylidene)but-2-en-1-ylidene)-3,3-dimethylindoline-5-carboxylic acid (10).

Compound 8 (1.784 g, 3.00 mmol), compound 9 (1.603 g, 4.50 mmol), and NaOAc (0.369 g, 4.50 mmol) were added to a flame dried flask and diluted in 1:1 MeOH/CHCl₃ (30 mL). The reaction was heated to reflux and stirred under argon for 16 h. The reaction mixture was concentrated with 10 g silica and the resulting cake was eluted on an 80 g silica column with 0 – 5% MeOH in CH₂Cl₂. The main product was isolated as 0.385 g (17%) dark amber foamy solid. ¹H NMR (400 MHz, DMSO-*d*₆) δ 12.60 (bs, 2H), 8.22 – 8.07 (m, 2H), 7.97 (d, *J* = 1.6 Hz, 1H), 7.91 (dd, *J* = 8.4 Hz, 1.6 Hz, 1H), 7.81 – 7.69 (m, 1H), 7.68 (s, 1H), 7.27 (d, *J* = 8.4 Hz, 1H), 7.19 (d, *J* = 3.6 Hz, 1H), 6.23 (dt, *J* = 10.0, 2.4 Hz, 1H), 5.37 (s, 2H), 4.21 (t, *J* = 6.6 Hz, 2H),

4.11 (dt, $J = 7.0, 2.6$ Hz, 2H), 3.90 (s, 3H), 3.87 (s, 3H), 3.85 – 3.77 (m, 2H), 2.72 – 2.61 (m, 4H), 1.63 (s, 6H), 1.08 (t, $J = 7.0$ Hz, 3H). MS-ESI m/z 735.1 ($[M + H]^+$ requires 735.3).

[0314] (E)-acetoxymethyl 1-(3-(acetoxymethoxy)-3-oxopropyl)-2-((2E,4E)-4-(1-(3-((4,5-dimethoxy-2-nitrobenzyl)oxy)-3-oxopropyl)-3-ethyl-2,4,6-trioxotetrahydropyrimidin-5(2H)-ylidene)but-2-en-1-ylidene)-3,3-dimethylindoline-5-carboxylate (11). Compound 10 (0.270 g, 0.367 mmol) was diluted in ACN (10 mL) and diisopropylethylamine (257 μ L, 1.47 mmol) was added, followed by addition of bromomethyl acetate (144 μ L, 1.47 mmol). The reaction was allowed to stir under argon for 16 h. Solvent was removed and the residue was redissolved in EtOAc. The organic layer was washed with water and brine, dried with $MgSO_4$, filtered, and concentrated. The residue was prepared as a silica cake and eluted on a 24 g silica column with 0 – 75% EtOAc in hexanes over 25 min. Main product isolated as 0.163 g (51%) dark blue solid. 1H NMR (400 MHz, $CDCl_3$) δ 8.09 (dd, $J = 12.3, 4.2$ Hz, 1H), 8.05 (dd, $J = 8.4, 1.5$ Hz, 1H), 7.92 (d, $J = 1.5$ Hz, 1H), 7.88 – 7.81 (m, 1H), 7.80 – 7.74 (m, 1H), 7.71 (d, $J = 4.3$ Hz, 1H), 7.08 (d, $J = 8.6$ Hz, 1H), 6.96 (dd, $J = 8.5, 3.7$ Hz, 1H), 5.99 (s, 2H), 5.87 (dd, $J = 12.2, 3.2$ Hz, 1H), 5.72 (d, $J = 2.8$ Hz, 2H), 5.54 (d, $J = 3.6$ Hz, 2H), 4.33 (dt, $J = 7.4, 3.2$ Hz, 2H), 4.13 (q, $J = 7.0$ Hz, 2H), 4.00 (d, $J = 4.4$ Hz, 3H), 4.00 – 3.95 (m, 2H), 3.96 (d, $J = 1.9$ Hz, 3H), 2.84 – 2.76 (m, 4H), 2.14 (s, 3H), 2.09 (d, $J = 4.7$ Hz, 3H), 1.65 (d, $J = 1.5$ Hz, 6H), 1.22 (t, $J = 7.0$ Hz, 3H). MS-ESI m/z 879.0 ($[M + H]^+$ requires 879.3).

[0315] Mero76. Compound 11 (132 mg, 0.150 mmol) was diluted in degassed ACN (75 mL). The flask was sealed, evacuated, and argon flushed. The reaction was irradiated at 365 nm for 24 h. The reaction mixture was concentrated and the residue was loaded in a minimum amount of CH_2Cl_2 onto a 12 g silica column and eluted with 0 – 5% MeOH in CH_2Cl_2 . Isolated 88 mg (86%) desired product as a dark blue solid. 1H NMR (400 MHz, $DMSO-d_6$) δ 12.31 (bs, 1H), 8.24 (dd, $J = 12.9, 3.3$ Hz, 1H), 8.15 (t, $J = 13.1$ Hz, 1H), 7.99 (d, $J = 1.5$ Hz, 1H), 7.94 (dd, $J = 8.4, 1.7$ Hz, 1H), 7.81 (dt, $J = 13.1, 3.6$ Hz, 1H), 7.29 (d, $J = 8.5$ Hz, 1H), 6.23 (d, $J = 12.9$ Hz, 1H), 5.92 (s, 2H), 5.62 (s, 2H), 4.24 (t, $J = 6.6$ Hz, 2H), 4.03 (t, $J = 7.4$ Hz, 2H), 3.88 – 3.80 (m, 2H), 2.81 (t, $J = 6.9$ Hz, 2H), 2.49 – 2.45 (m, 2H), 2.10 (s, 3H), 2.01 (s, 3H),

1.64 (s, 6H), 1.11 (t, $J = 7.0$ Hz, 3H). HRMS m/z 684.2408 ($[M + H]^+$ requires 684.2405).

[0316] **1-(2-Azidoethyl)-3-ethylurea (S8).** 2-Azidoethylamine (2.15 g, 25.0 mmol) was diluted in THF (25 mL) in an oven dried flask under argon and the solution was cooled in a cold water bath. Ethylisocyanate (1.98 mL, 25.0 mmol) was added dropwise and the reaction was stirred for 16 h. The reaction was concentrated and all residual solvents were removed under high vacuum to give 3.02 g (77%) off-white solid suitable for use without further purification. ^1H NMR (400 MHz, CDCl_3) δ 5.53 (bs, 1H), 5.26 (bs, 1H), 3.42 – 3.31 (m, 4H), 3.22 – 3.14 (m, 2H), 1.11 (t, $J = 7.2$ Hz, 3H). ^{13}C NMR (100 MHz, CDCl_3) δ 158.8, 52.0, 39.9, 35.3, 15.6. MS-ESI m/z 158.0 ($[M + H]^+$ requires 158.1).

[0317] **1-(2-Azidoethyl)-3-ethylpyrimidine-2,4,6(1H,3H,5H)-trione (S9).** Compound S8 (1.57 g, 10.0 mmol) and malonic acid (1.09 g, 10.5 mmol) were added to a flame dried flask and the flask was sealed, evacuated, and argon flushed. AcOH (14 mL) was added and the flask was heated to 60 °C for 10 min. Acetic anhydride (3.8 mL, 40.0 mmol) was added and the reaction was further heated to 95 °C and stirred for 1 h. The reaction was then concentrated and the residue was loaded in a minimum amount of CH_2Cl_2 onto a silica column and eluted with 0 – 2% MeOH in CH_2Cl_2 . Main product containing fractions were combined and concentrated to give a pale amber oil that became a waxy solid on cooling. Isolated 1.86 g (83%). ^1H NMR (400 MHz, CDCl_3) δ 4.13 (t, $J = 6.0$ Hz, 2H), 3.95 (q, $J = 7.2$ Hz, 2H), 3.70 (s, 2H), 3.52 (t, $J = 6.0$ Hz, 2H), 1.22 (t, $J = 7.2$ Hz, 3H). ^{13}C NMR (100 MHz, CDCl_3) δ 164.8, 164.2, 151.3, 48.8, 40.5, 39.8, 37.7, 13.4. MS-ESI m/z 223.9 ($[M - H]^-$ requires 224.1).

[0318] ***N*-((1E,3E)-3-(1-(2-Azidoethyl)-3-ethyl-2,4,6-trioxotetrahydropyrimidin-5(2H)-ylidene)prop-1-en-1-yl)-*N*-phenylacetamide (S10).** Compound S9 (1.126 g, 5.00 mmol), malonaldehyde dianilide hydrochloride (1.423 g, 5.50 mmol), and sodium acetate (0.451 g, 5.50 mmol) were diluted in acetic anhydride (6.0 mL). The reaction mixture was added to a pre-heated oil bath at 100 °C and stirred for 1 h. The reaction was cooled and added in portions to saturated aq. NaHCO_3 (50 mL). The aq. layer was extracted with EtOAc (3 X 50 mL). The organic layers were combined, washed with brine, dried with MgSO_4 , filtered, and

concentrated. The residue was diluted in CH_2Cl_2 and concentrated with CELITE™. The CELITE™ cake was eluted on an 80 g silica column with 0 – 50% EtOAc in hexanes over 25 min. The product was isolated as 0.479 g (24%) yellow solid as an approximate 1:1 mixture of *E* and *Z* isomers. Isomer A: ^1H NMR (400 MHz, CDCl_3) δ 8.52 (d, J = 13.6 Hz, 1H), 8.19 (d, J = 12.8 Hz, 1H), 7.64 – 7.54 (m, 3H), 7.26 – 7.22 (m, 2H), 6.86 (t, J = 13.2 Hz, 1H), 4.17 (t, J = 6.2 Hz, 2H), 3.87 (q, J = 7.2 Hz, 2H), 3.52 (t, J = 6.0 Hz, 2H), 2.01 (s, 3H), 1.13 (t, J = 7.0 Hz, 3H). ^{13}C NMR (100 MHz, CDCl_3) δ 170.0, 162.1, 161.4, 158.8, 151.1, 150.6, 138.0, 131.0 (2C), 130.4, 128.1 (2C), 111.6, 110.5, 49.0, 40.5, 36.6, 23.6, 13.5. Isomer B: ^1H NMR (400 MHz, CDCl_3) δ 8.52 (d, J = 13.6 Hz, 1H), 8.19 (d, J = 12.4 Hz, 1H), 7.64 – 7.54 (m, 3H), 7.26 – 7.22 (m, 2H), 6.83 (t, J = 13.0 Hz, 1H), 4.04 (t, J = 6.4 Hz, 2H), 3.98 (q, J = 7.2 Hz, 2H), 3.43 (t, J = 6.2 Hz, 2H), 2.01 (s, 3H), 1.21 (t, J = 7.0 Hz, 3H). ^{13}C NMR (100 MHz, CDCl_3) δ 170.0, 162.5, 161.7, 158.8, 151.1, 150.5, 138.0, 131.0 (2C), 130.4, 128.1 (2C), 111.6, 110.6, 48.9, 39.7, 37.4, 23.7, 13.5. MS-ESI m/z 418.4 ($[\text{M} + \text{Na}]^+$ requires 419.1).

[0319] (E)-2-((2E,4E)-4-(1-(2-Azidoethyl)-3-ethyl-2,4,6-trioxotetrahydropyrimidin-5(2H)-ylidene)but-2-en-1-ylidene)-1-(2-carboxyethyl)-3,3-dimethylindoline-5-carboxylic acid (S11). Compound S10 (0.458 g, 1.155 mmol) and compound 9 (0.617 g, 1.733 mmol) were diluted in 1:1 MeOH/ CHCl_3 (7.5 mL). The reaction flask was topped with a condenser with argon inlet and heated for 16 h. The reaction was concentrated and re-diluted in EtOAc and 1 M aq. NaOH (25 mL) was added. The aq. layer was washed with EtOAc (3 X 25 mL), cooled in an ice bath, and acidified with 2 M HCl. The aq. layer was then extracted with CH_2Cl_2 . The organic layers were combined, dried with MgSO_4 , filtered, and concentrated. The residue was prepared as a CELITE™ cake and eluted on a 40 g silica column with 0 – 5% MeOH in CH_2Cl_2 . Main product isolated as 0.134 g (22%) dark blue solid. ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 12.58 (bs, 2H), 8.25 (d, J = 13.2 Hz, 1H), 8.18 (t, J = 13.2 Hz, 1H), 7.98 (s, 1H), 7.92 (d, J = 8.8 Hz, 1H), 7.85 – 7.74 (m, 1H), 7.28 (d, J = 8.4 Hz, 1H), 6.31 – 6.21 (m, 1H), 4.30 – 4.15 (m, 2H), 4.10 – 4.00 (m, 2H), 3.92 – 3.78 (m, 2H), 3.49 (t, J = 5.8 Hz, 2H), 2.66 (t, J = 6.8 Hz, 2H), 1.64 (s, 6H), 1.11 (t, J = 6.8 Hz, 3H). MS-ESI m/z 536.9 ($[\text{M} + \text{H}]^+$ requires 537.2).

[0320] **Merol66.** Compound **S11** (56 mg, 0.094 mmol) was diluted in ACN (3.0 mL) and bromomethyl acetate (37 μ L, 0.38 mmol) was added, followed by addition of *N,N*-diisopropylethylamine (66 μ L, 0.38 mmol). The flask was capped and fitted with an argon inlet line and the reaction was stirred at room temp. for 16 h. The reaction mixture was then concentrated and re-diluted in CH₂Cl₂ and prepared as a CELITE™ cake. The sample was eluted on a 12 g silica column with 0 – 50% EtOAc in hexanes over 25 min. Isolated 37 mg (52%) dark blue solid. ¹H NMR (400 MHz, CDCl₃) δ 8.11 (dd, *J* = 12.2, 2.2 Hz, 1H), 8.03 (dd, *J* = 8.4, 1.6 Hz, 1H), 7.89 (d, *J* = 1.6 Hz, 1H), 7.88 – 7.72 (m, 2H), 6.95 (d, *J* = 8.4 Hz, 1H), 5.95 (s, 2H), 5.88 (d, *J* = 12.4 Hz, 1H), 5.69 (s, 2H), 4.18 (q, *J* = 6.4 Hz, 2H), 4.11 (t, *J* = 7.0 Hz, 2H), 4.03 – 3.94 (m, 2H), 3.52 (t, *J* = 6.2 Hz, 2H), 2.77 (t, *J* = 7.2 Hz, 2H), 2.12 (s, 3H), 2.07 (s, 3H), 2.00 (s, 3H), 1.63 (s, 6H), 1.22 (dt, *J* = 7.0, 1.6 Hz, 3H). HRMS *m/z* 681.2539 ([M + H]⁺ requires 681.2520).

[0321] **G9a ligand – merol66 conjugate (12).** Merol66 (5.1 mg, 7.5 μ mol) and G9a ligand (6.2 mg, 11 μ mol) were diluted in 500 μ L CH₂Cl₂. A 15 mM stock solution of tetrakis(acetonitrile)copper(I) hexafluorophosphate was prepared in CH₂Cl₂ and 100 μ L was added to the reaction mixture dropwise. After 24 h, the reaction was concentrated and submitted to preparative LC. The product was isolated after lyophilization as 6.8 mg (74%) dark blue solid. MS-ESI *m/z* 1228.2 ([M + H]⁺ requires 1228.6).

Example 6

Preparation and Characterization of Membrane Permeable Dyes

[0322] We and others have optimized merocyanine dyes for live cell imaging applications (Toutchkine *et al.*, *J. Am. Chem. Soc.* 125:4132-4145 (2003); MacNevin *et al.*, *Bioconjug. Chem.* 24:215-223 (2013); Toutchkine *et al.*, *Org. Lett.* 9:2775-2777 (2007); Toutchkine *et al.*, *Bioconjug. Chem.* 18:1344-1348 (2007); Kulinich *et al.*, *Russian Chem. Rev.* 78:141-164 (2009)). Merocyanines are characterized by the presence of electron donor and acceptor components that are linked by conjugation, usually a system of double bonds. These dyes are especially well suited for use in living cells or animals because they are very bright, they can be tuned to emit at wavelengths longer than those of cellular autofluorescence, and they can exhibit

substantial fluorescence changes in response to solvent environment (Kulinich *et al.*, *Russian Chem. Rev.* 78:141-164 (2009); Loving *et al.*, *Trends Biotechnol.* 28:73-83 (2010); Touthikine *et al.*, *J. Phys. Chem. A* 111:10849-10860 (2007); Buncel *et al.*, *Accounts Chem. Res.* 23:226-231 (1990); Bondarev *et al.*, *J. Luminescence* 124:178-186 (2007); Kulinich *et al.*, *Russian J. Gen. Chem.* 76:1441-1457 (2006)).

Merocyanine dyes have been used in several different biosensor designs that were able to report the activity of endogenous proteins (Nalbant *et al.*, *Science* 305:1615-1619 (2004); Garrett *et al.*, *Biochemistry* 47:986-996 (2008); Hahn *et al.*, *J. Biol. Chem.* 265:20335-20345 (1990); Hahn *et al.*, *Biophys. J.* 57:A248-A248 (1990); Gulyani *et al.*, *Nat. Chem. Biol.* 7:437-444 (2011)).

[0323] Like almost all bright, long wavelength fluorophores, water soluble merocyanine dyes are poorly membrane permeant, so require chemical modification to allow them to enter cells through passive translocation. If water solubility is reduced to enhance membrane passage, the less polar dyes tend to stain intracellular membranes such as the Golgi apparatus, mitochondria, and endoplasmic reticulum. One method for achieving cell permeability has been the use of acetoxymethyl (AM) ester groups that lead to charged carboxylate side chains upon hydrolysis by intracellular esterases (Tsien, *Nature* 290:527-528 (1981)). The esters mask charge, permitting membrane translocation until they are cleaved to produce carboxylic acids. To investigate the effect of ester number and orientation on membrane permeability and staining of intracellular membranes, several versions of a representative parent merocyanine dye were synthesized with AM esters at different positions (Fig. 10). Based on our earlier work optimizing merocyanines for live cell biosensors, the indolenine-barbituric acid (I-BA) scaffold was chosen (an indolenine ring (I) linked to a barbituric acid (BA) by four carbons). This dye showed a favorable combination of brightness, photostability, fluorescence wavelengths, and solvent sensitivity (Touthikine *et al.*, *J. Am. Chem. Soc.* 125:4132-4145 (2003); MacNevin *et al.*, *Bioconjug. Chem.* 24:215-223 (2013)). We examined the membrane permeability and subcellular distribution of the underivatized parent merocyanine 1 (Fig. 10), two versions of the dye with single AM esters (on the donor I component (2) or on the acceptor BA component (3)), and two versions of the dye containing two AM esters (on the donor and the acceptor (4), and with both esters on the donor (5)).

[0324] Mouse embryonic fibroblast cells were incubated in medium containing the test compound at 37 °C for 30 minutes, followed by aspiration of the dye solution, and washing with medium. The cells were then trypsinized, suspended in medium, and the fluorescence intensity of the cell suspension was obtained for each dye. Values were normalized for cell number based on the relative intensity of the nuclear stain Hoechst 33342, which was included with the dye in each test solution. As expected, the unmodified merocyanine dye **1** showed negligible import into cells (Fig. 11). The merocyanines containing a single AM ester both showed similar improvements in fluorescence uptake relative to **1**. However, the addition of a two AM esters provided an even greater enhancement of import. Compound **5**, which contains two AM esters on the donor component, was significantly brighter than compound **4**, which has an AM ester on both the indolenine and barbituric acid rings. Subsequent studies showed that compound **5** also had more efficient import relative to **4** at lower concentrations (Fig. 12).

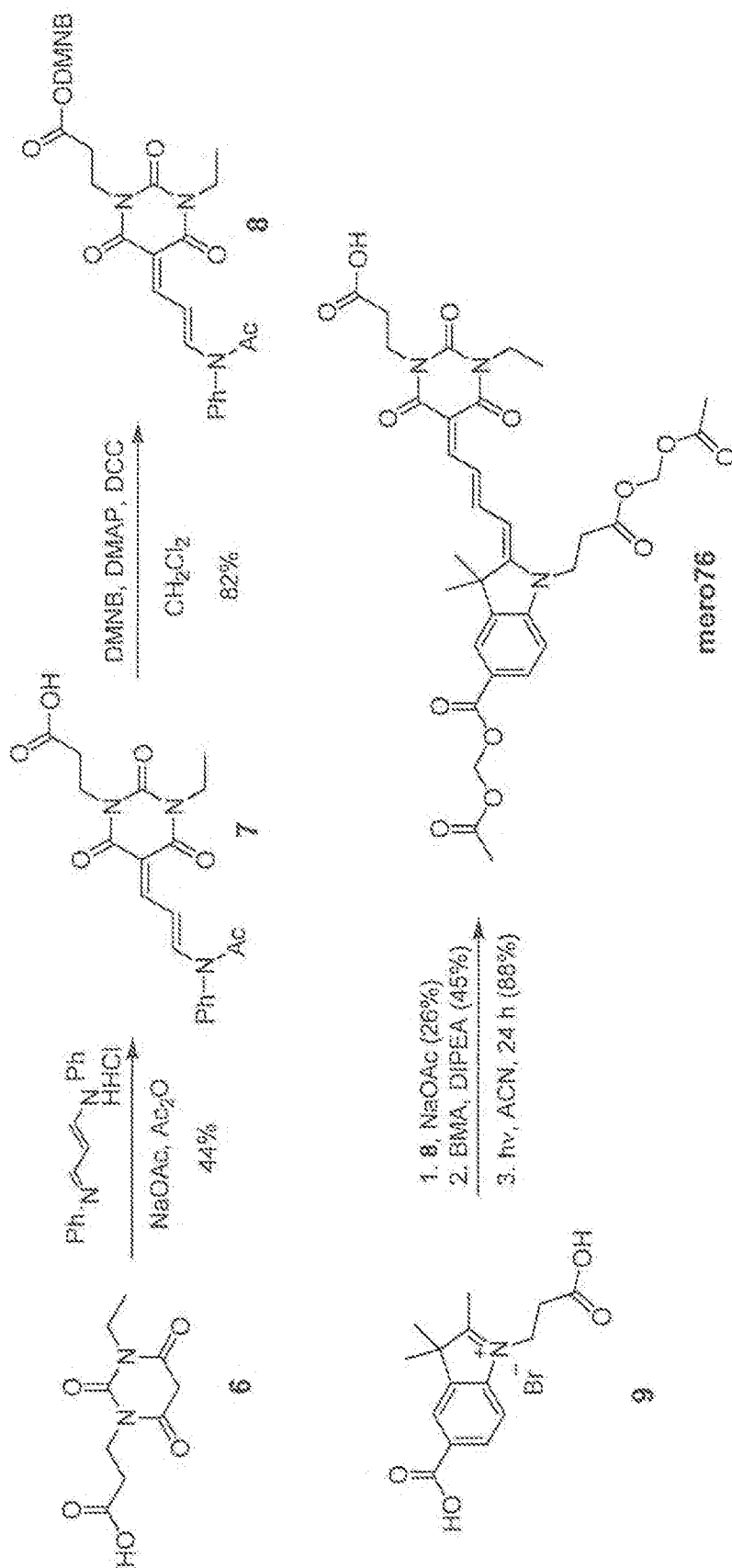
[0325] Importantly, dye **5** showed uniform distribution within the cells, unlike the other dyes which stained intracellular membranes and produced fluorescent vesicles (Fig. 13). Subcellular distribution of the dyes was compared to the distribution of molecules that are known to show uniform distribution (stably expressed yellow fluorescent protein). The dyes and these control molecules were visualized in the same cell. The combination of intrinsic fluorescent properties and efficient uptake for dye **5** led to exceptionally bright images relative to other dyes. Cells tolerated the presence of this dye and appeared healthy for the duration of the 90 minute experiments. Having both AM esters on the donor component of **5** was advantageous as it provided ready synthetic accessibility to dyes with different acceptor components; varying the acceptor has proven useful when tailoring fluorophores to different applications (MacNevin *et al.*, *Bioconjug. Chem.* 24:215-223 (2013)). Based on these results, merocyanine **5** was chosen as the lead structure to carry forward for further development into a conjugatable version.

[0326] Synthesis of mero**76**, a conjugatable version of dye **5** bearing a carboxylic acid side chain, began from the barbituric acid derivative **6**, obtained in 3 steps from β -alanine methyl ester hydrochloride (Scheme 6). Compound **6** was first converted to the activated diene **7** and the free carboxylic acid was then coupled with the

dimethoxy nitrobenzyl (DMNB) photolabile protecting group to give compound **8** (Bochet, *J. Chem. Soc.-Perkin Transactions 1* 125-142 (2002); Blanc *et al.*, *J. Org. Chem.* 67:5567-5577 (2002)). The di-carboxy indolenine salt **9** was reacted with **8** under mildly basic conditions to give a merocyanine product that after installation of the two acetoxymethyl esters and photodeprotection of DMNB provided **mero76**. The azide containing merocyanine **mero166**, for use in click chemistry, was prepared thorough an analogous route involving reaction of compound **9** with an azide-containing barbituric acid, but without the need to utilize protecting group chemistry. The emission spectra of **mero76** and **mero166** illustrate the significant solvent-sensitivity of these dyes (Fig. 14). Large increases in fluorescence intensity are observed when moving from water to less polar solvents. Both dyes maintained good brightness and solvent sensitivity characteristics relative to the non-derivatized parent structure (Table 7) (MacNevin *et al.*, *Bioconjug. Chem.* 24:215-223 (2013)).

[0327] We sought to demonstrate the potential of these dyes for use in live cell imaging applications. A selective ligand for the DNA-methyltransferase G9a (Vedadi *et al.*, *Nat. Chem. Biol.* 2011, 7:648-648 (2011)) was functionalized with an alkynyl side chain and conjugated to **mero166** via click chemistry to give compound **10** (Scheme 5). MEF cells were incubated with either the dye alone or compound **10** for 20 minutes, washed, and imaged. Cells treated with **mero166** showed uniform fluorescence intensity throughout the cell, with no apparent organelle or lipid specific interactions (Fig. 15). In contrast, the ligand-dye conjugate **10** showed distinct nuclear localization, which is consistent with the localization expected to result from binding of the probe to G9a, a well-characterized histone methyltransferase (Shinkai *et al.*, *Genes Dev.* 25:781-788 (2011)).

[0328] In conclusion, we have developed new merocyanine dyes that are able to freely cross cellular membranes and distribute uniformly without appreciable staining of cellular membranes. We have further prepared these dyes with reactive side chains and demonstrated that they can be conjugated to a small molecule, passively diffuse into cells, and there be used to observe the distribution of the small molecule as it binds to its ligand. We believe the dyes described here have the potential for use in a variety of applications that aim to monitor protein localization and dynamics in live cells and animals.



Scheme 6. Synthesis of the conjugatable dye **mero76** with carboxylic acid side chain, via photolabile protecting group strategy.

Example 7Experimental Procedures

[0329] **In vitro assays:** Stock calcium solutions were prepared using 100 mM NaCl, 50 mM Tris-HCl, pH 7.5 and buffered with the appropriate amount of EGTA to maintain the target free calcium concentration at an ionic strength of 0.05 using the online calculator located at: maxchelator.stanford.edu/CaEGTA-NIST.htm.

Calmodulin was purchased from Sigma-Aldrich (cat. P1431). Fluorescence intensity experiments were run in triplicate in 386-well plates. For calcium titration and reversibility experiments, CaM and probe (CaMero or CaMero-NC) concentrations were both 1.0 μ M and the stock EGTA solution was prepared at 200 μ M.

[0330] **Cell culture and treatment with imaging probes:** Human 1321N1 astrocytoma cells and NIH 3T3 mouse embryonic fibroblasts (MEF) stably expressing YPet fluorophore were maintained in 10% CO₂ at 37 °C in Dulbecco's modified Eagle's medium (DMEM, Cellgro) with 10% fetal bovine serum (HyClone, Thermo Scientific) and 2 mM GlutaMax (Gibco, Life Technologies). Cells were plated 3-4 hours before imaging on coverslips coated with fibronectin 20 μ g/mL (Sigma-Aldrich) overnight. Cells were incubated with the test probe (5.0 μ M) for 20 min in 10% CO₂ at 37 °C in imaging medium [HAM'S F-12 (K) Medium (SKU #RR080058L1, Gibco, Life Technologies) with 5% fetal bovine serum (HyClone, Thermo Scientific) and 2 mM GlutaMax (Gibco, Life Technologies)]. The cells were then gently washed with imaging medium (2 X 2 mL) and imaged in the same medium.

[0331] **Live cell imaging:** Live cell imaging was carried out on an Olympus IX81 microscope with UPLFLN 40X oil objective (NA 1.3) and mercury lamp excitation (103W HBO bulb). Filters used for CaMero, CaMero-NC, and the dye alone were 545/50 excitation and 620/60 emission. Excitation was through an ND 2.0 (1% transmission) neutral density filter, typically using 50 ms exposure. For YPet, filters were 500/20 for excitation and 535/30 for emission. Excitation was through a ND 2.0 (1% transmission) filter with 500 ms exposure. Excitation was through a ND 1.3 (5%T) neutral density filter with 500 ms exposure. Images were acquired with a Coolsnap ES camera (Photometrics) with a Sony 6.45 x 6.45 μ m/pixel chip using 2 x 2 binning. All image acquisition, processing, and analysis was carried out with

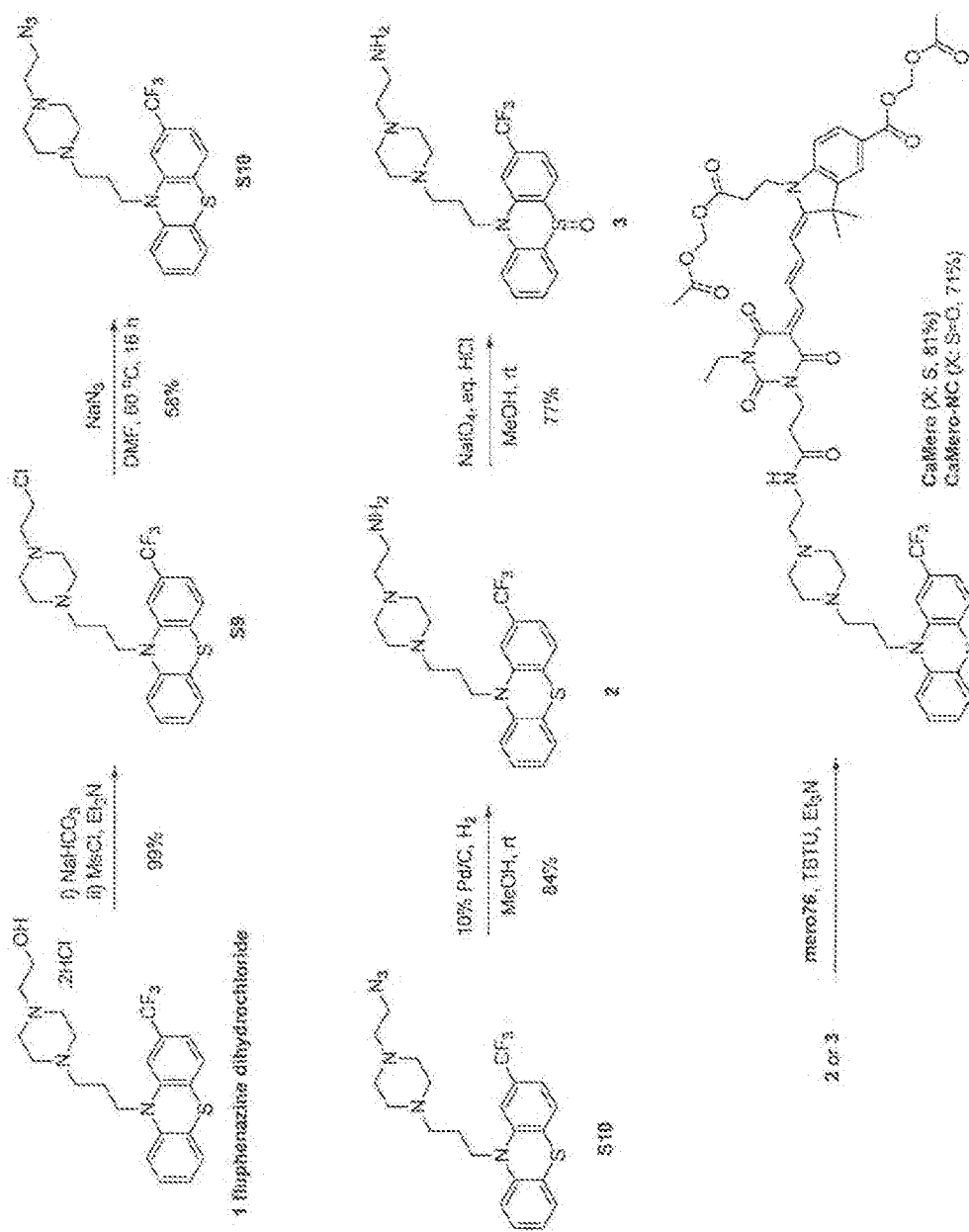
Metamorph software. Ratio images were generated from unprocessed images following shade correction, background subtraction, masking, registration for image alignment, and photobleaching correction. The display threshold of ratio images was restricted to the pixels falling between 5 - 95% of the overall frequency histogram.

[0332] **Whole animal experiments:** Treatment solutions of the probes were prepared as follows: 8.0 μ L stock solution of probe (CaMero or CaMero-NC at 20 mM in DMSO), 8.0 μ L 20% w/v Pluronic F-127 in DMSO, 15 μ L stock solution of Hoechst 33342 (10mM in DMSO), and 120 μ L phosphate buffered saline were combined and mixed in the order listed to give a final concentration for the probe of approx. 1.0 μ M and a total working volume of 150 μ L. Treatment solutions were administered to GFP-LifeAct mice via tail-vein injection GFP-LifeAct mice. Mice were allowed to recover for a period of 1 - 2 hours, sacrificed by cervical dislocation, and short sections (~1 cm) of the intestine were removed at a distance of approximately 10 - 15 cm from the stomach, washed once in PBS, and placed into Fish Ringer's solution in glass bottom dishes (Matek) for imaging. Tissue was imaged using a LaVision TRIM scope attached to a Nikon Eclipse Ti inverted stand with an Olympus 20x, 0.95 NA water immersion objective. Two photon excitation of CaMero/CaMero-NC was achieved with an optical parametric oscillator laser tuned to 1110 nm and emission was detected using a 630/60 emission filter for CaMero/CaMero-NC and 549/15 filter for the SHG signal of collagen. Hoechst was excited with a Ti-sapphire laser tuned to 800 nm and emission was collected using a 435/40 filter. For localization studies, tissue samples were imaged after beginning from the crypt of Lieberkühn and extending through to the peripheral smooth muscle layer using incremental steps in the Z plane of 4 μ M (30 frames in total). For imaging of muscle contraction, CaMero and CaMero-NC samples were acquired within approximately 10 minutes of each other using identical acquisition parameters (laser power and PMT gain). CaMero and GFP-LifeAct images are displayed using different contrast settings for each axial position; however, the same contrast settings have been applied to images of the CaMero and CaMero-NC probes.

Example 8Synthesis of Compounds

[0333] 10-(3-(4-(2-Chloroethyl)piperazin-1-yl)propyl)-2-(trifluoromethyl)-10H-phenothiazine (S9). Fluphenazine dihydrochloride (1, 1.00 g, 1.96 mmol) was diluted in saturated aqueous sodium bicarbonate (50 mL) and stirred for 10 min. The aqueous layer was extracted with CH₂Cl₂ (3 x 50 mL). Organic layers were combined, washed with saturated aq. NaCl, dried with MgSO₄, and filtered. The filtrate was concentrated to give 0.85 g pale amber oil (quant.) that was carried forward directly to the next step.

[0334] The crude product (0.845 g, 1.93 mmol) was diluted in CH₂Cl₂ (25 mL) in an oven dried 100 mL round-bottom flask. The flask was sealed with a rubber septum, flushed with argon, and chilled in an ice bath. Triethylamine (0.417 mL, 2.99 mmol) was added, followed by dropwise addition of methanesulfonyl chloride (0.233 mL, 2.99 mmol). The mixture was stirred for 1 h at 0 °C, allowed to equilibrate to RT and stirred for 1.5 h. Water was added (25 mL) and the organic layer was separated, dried with MgSO₄, filtered, and the filtrate was concentrated. Isolated 0.893 g (quant.) pale amber oil that appeared pure by NMR analysis. The product was carried forward without further purification. ¹H NMR (400 MHz, CDCl₃) δ 7.24 – 7.08 (m, 4H), 7.03 (s, 1H), 6.97 – 6.88 (m, 2H), 3.95 (t, *J* = 6.8 Hz, 2H), 3.56 (t, *J* = 7.1 Hz, 2H), 2.70 (t, *J* = 7.1 Hz, 2H), 2.62 – 2.25 (m, 10H), 1.98 – 1.87 (m, 2H). ¹³C NMR (100 MHz, CDCl₃) δ 145.9, 144.5, 130.1, 129.7 (q, *J* = 32 Hz), 127.8, 127.7, 127.6, 124.3 (q, *J* = 270 Hz), 124.2, 123.2, 119.1 (q, *J* = 3.9 Hz), 116.1, 112.1 (q, *J* = 3.8 Hz), 59.9, 55.5, 53.3 (four C), 45.4, 41.0, 24.3. MS-ESI *m/z* 456.2 ([M + H]⁺ requires 456.1).



Scheme 7. Full synthetic route to CalMero and CalMero-NC.

[0335] 10-(3-(4-(2-Azidoethyl)piperazin-1-yl)propyl)-2-(trifluoromethyl)-10H-phenothiazine (S10). Compound 89 (1.52 g, 3.33 mmol) was added to a 50 mL oven dried round-bottom flask and diluted in DMF (20.0 mL). Sodium iodide (0.599 g, 3.99 mmol) was added, followed by sodium azide (0.433 g, 6.67 mmol). The flask was sealed with a septum, argon flushed, and the reaction mixture was heated to 60 °C for 16 h. Aqueous LiCl (15% w/v, 50 mL) was added and the aqueous layer was extracted with EtOAc (3 x 50 mL). The organic layers were combined, washed with water (200 mL) and brine (200 mL), dried with Na₂SO₄, filtered, and concentrated. The crude product was suspended in a minimum amount of CH₂Cl₂ and loaded onto a 24 g SiO₂ column. The column was eluted with 0 – 3% MeOH in CH₂Cl₂ over 20 min. Main product containing fractions were combined, concentrated, and dried under vacuum to give 0.890 g (58%) pale amber oil. ¹H NMR (400 MHz, CDCl₃) δ 7.22 – 7.09 (m, 4H), 7.04 (s, 1H), 6.97 – 6.88 (m, 2H), 3.96 (t, *J* = 6.8 Hz, 1H), 3.32 (t, *J* = 6.1 Hz, 1H), 2.56 (t, *J* = 6.1 Hz, 2H), 2.54 – 2.20 (m, 8H), 1.97 – 1.88 (m, 2H). ¹³C NMR (100 MHz, CDCl₃) δ 145.9, 144.5, 130.0, 129.7 (q, *J* = 32 Hz), 127.8, 127.7, 127.6, 124.3 (q, *J* = 271 Hz), 124.2, 123.2, 119.1 (q, *J* = 3.9 Hz), 116.1, 112.1 (q, *J* = 3.8 Hz), 57.2, 55.5, 53.3 (four C), 48.4, 45.5, 24.3. MS-ESI *m/z* 463.2 ([*M* + *H*]⁺ requires 463.2).

[0336] 2-(4-(3-(2-(trifluoromethyl)-10H-phenothiazin-10-yl)propyl)piperazin-1-yl)ethan-1-amine (2). Palladium on carbon (10 wt. %, 0.060 g) was added to a 50 mL flame dried two-neck round-bottom flask. One neck of the flask was sealed with a rubber septum fitted with an argon line and the other with a glass stopper. The flask was evacuated and argon flushed two times. Compound 3 (0.305 g, 0.659 mmol) was diluted in anhydrous methanol (10 mL) and added to the reaction flask. The glass stopper was replaced with a hydrogen containing balloon. The argon line was removed and the flask was flushed with hydrogen (2 x). The balloon was left open to the system and the reaction mixture was stirred at RT for 16 h. The reaction mixture was filtered through a sand-topped CELITE™ column and the column was rinsed with methanol (75 mL). The filtrate was concentrated, diluted in CH₂Cl₂, dried with Na₂SO₄, and filtered with CH₂Cl₂ washes. The filtrates were concentrated and loaded in a minimum amount of CH₂Cl₂ onto a 12 g SiO₂ column and eluted with 0 – 10%

MeOH (containing 1% NH_3) in CH_2Cl_2 over 25 min. Main peak containing fractions were combined, concentrated, and dried under high vacuum to give 0.242 g (84%) colorless oil. ^1H NMR (400 MHz, CDCl_3) δ 7.22 – 7.07 (m, 4H), 7.03 (s, 1H), 6.97 – 6.88 (m, 2H), 3.95 (t, J = 6.8 Hz, 2H), 2.76 (t, J = 6.2 Hz, 2H), 2.63 – 2.20 (m, 12H), 2.01 – 1.85 (m, 2H), 1.41 (bs, 2H). ^{13}C NMR (100 MHz, CDCl_3) δ 145.9, 144.5, 130.0, 129.7 (q, J = 32.1 Hz), 127.8, 127.7, 127.6, 124.3 (q, J = 271 Hz), 124.2, 123.2, 119.1 (q, J = 3.9 Hz), 116.1, 112.0 (q, J = 3.8 Hz), 61.2, 55.6, 53.5 (two C), 53.4 (two C), 45.5, 39.0, 24.3. MS-ESI m/z 437.2 ($[\text{M} + \text{H}]^+$ requires 437.2).

[0337] 10-(3-(4-(2-aminoethyl)piperazin-1-yl)propyl)-2-(trifluoromethyl)-10H-phenothiazine 5-oxide (3). Compound 2 (0.050 g, 0.11 mmol) was diluted in MeOH (1.0 mL) in a 10 mL round-bottom flask open to air. Sodium periodate (0.028 g, 0.132 mmol) was dissolved in a 3:1 MeOH/1.0 M aq. HCl solution (2.0 mL) with sonication and added to the reaction flask quickly dropwise. The reaction mixture was stirred for 4 h at RT. Saturated aq. NaHCO_3 (2.0 mL) was added and the mixture was concentrated. Some additional water was added and the aqueous layer was extracted with CH_2Cl_2 (10 mL $\times$ 4). Organic layers were combined, washed with saturated aq. NaCl, dried with MgSO_4 , filtered, and concentrated to give 0.040 g (77%) clear oil that foamed on drying. The product required no further purification. ^1H NMR (400 MHz, CDCl_3) δ 8.04 (d, J = 8.0 Hz, 1H), 7.95 (dd, J = 7.7, 1.5 Hz, 1H), 7.72 (s, 1H), 7.68 – 7.62 (m, 1H), 7.59 (d, J = 8.3 Hz, 1H), 7.47 (d, J = 8.0 Hz, 1H), 7.30 (t, J = 7.4 Hz, 1H), 4.50 – 4.31 (m, 2H), 2.79 (t, J = 6.2 Hz, 2H), 2.69 – 2.27 (m, 12H), 2.12 – 1.99 (m, 2H), 1.63 (bs, 2H). ^{13}C NMR (100 MHz, CDCl_3) δ 138.7, 138.1, 134.4 (q, J = 32.3 Hz), 133.1, 132.2, 131.5, 127.2, 124.6, 123.5 (q, J = 272 Hz), 122.6, 117.9 (q, J = 3.5 Hz), 116.3, 113.2 (q, J = 4.0 Hz), 61.1, 54.6, 53.4 (two C), 53.3 (two C), 45.6, 38.8, 23.9. MS-ESI m/z 453.2 ($[\text{M} + \text{H}]^+$ requires 453.2).

[0338] CaMero. Compound mero76 (10.0 mg, 0.015 mmol), compound 2 (9.6 mg, 0.022 mmol), and TBTU (7.0 mg, 0.022 mmol) were added to a 5 mL conical vial with spin vane and dissolved in DMF (0.70 mL), followed by addition of Et_3N (3.1 μL , 0.022 mmol). The vial was capped and covered in foil and the reaction was stirred at RT for 16 h. The reaction mixture was submitted to preparative HPLC. Main product containing fractions were combined, concentrated, and lyophilized. Isolated 13.0 mg (81%) dark blue powdery solid. ^1H NMR (400 MHz, DMSO) δ 8.24

(t, $J = 12.2$ Hz, 1H), 8.20 – 8.04 (m, 2H), 7.99 (s, 1H), 7.94 (d, $J = 8.4$ Hz, 1H), 7.81 (dd, $J = 24.9, 12.9$ Hz, 1H), 7.39 (d, $J = 7.9$ Hz, 1H), 7.32 – 7.26 (m, 4H), 7.22 (d, $J = 7.5$ Hz, 1H), 7.10 (dd, $J = 7.9, 2.7$ Hz, 1H), 7.03 (t, $J = 7.4$ Hz, 1H), 6.22 (dd, $J = 15.6, 13.1$ Hz, 1H), 5.93 (s, 2H), 5.62 (d, $J = 2.6$ Hz, 2H), 4.30 – 4.15 (m, 2H), 4.09 – 3.97 (m, 4H), 3.90 – 3.80 (m, 2H), 3.35 – 3.22 (m, 2H), 3.13 – 2.84 (m, 4H), 2.81 (t, $J = 6.3$ Hz, 2H), 2.37 (t, $J = 6.4$ Hz, 2H), 2.10 (s, 3H), 2.05 – 1.95 (m, 2H), 2.01 (d, $J = 2.2$ Hz, 3H), 1.63 (s, 6H), 1.10 (t, $J = 7.0$ Hz, 3H). HR-FTMS m/z 1102.4218 ($[M + H]^+$ requires 1102.4208).

[0339] CaMero-NC. Compound mero76 (10.2 mg, 0.015 mmol), compound 3 (10.2 mg, 0.023 mmol), and TBTU (7.2 mg, 0.022 mmol) were added to a 5 mL conical vial with spin vane and dissolved in DMF (0.70 mL), followed by addition of Et₃N (3.2 μ L, 0.023 mmol). The vial was capped and covered in foil and the reaction was stirred at RT for 16 h. The reaction mixture was submitted to preparative HPLC. Main product containing fractions were combined, concentrated, and lyophilized. Isolated 12.0 mg (71%) dark blue powdery solid. ¹H NMR (400 MHz, DMSO) δ 8.28 – 7.90 (m, 6H), 7.98 (s, 1H), 7.97 – 7.89 (m, 1H), 7.86 – 7.73 (m, 3H), 7.61 (d, $J = 8.1$ Hz, 1H), 7.38 (t, $J = 7.4$ Hz, 1H), 7.29 (t, $J = 9.0$ Hz, 1H), 6.29 – 6.15 (m, 1H), 5.93 (s, 2H), 5.62 (d, $J = 3.4$ Hz, 2H), 4.65 – 4.60 (m, 2H), 4.30 – 4.17 (m, 2H), 4.02 (t, $J = 6.5$ Hz, 2H), 3.90 – 3.78 (m, 2H), 3.37 – 3.22 (m, 2H), 3.05 – 2.75 (m, 6H), 2.37 (t, $J = 6.4$ Hz, 2H), 2.10 (s, 3H), 2.09 – 2.02 (m, 2H), 2.00 (d, $J = 2.8$ Hz, 3H), 1.62 (s, 6H), 1.10 (t, $J = 7.0$ Hz, 3H). HR-FTMS m/z 1118.4162 ($[M + H]^+$ requires 1118.4157).

Example 9

Preparation and Characterization of Biosensors

[0340] The tremendous importance of spatial and temporal factors to the regulation of cellular signaling events is evidenced by the multi-faceted roles that a given protein may play in diverse and sometimes opposing cellular processes. Even when the primary molecular constituents of a signaling network are known, it often remains difficult to understand the dynamics of their collective behavior without taking into account spatiotemporal aspects of the endogenous system. Biosensors are quantitative tools that can be used to provide insight into how signaling networks are

transiently constructed to produce a specific behavior in real time and with spatial resolution. Despite the power of these tools, the widespread application of biosensors has been somewhat limited. It has recently been estimated that sensors have become available for little more than 100 of the vastly greater number of potential molecular targets of interest, due in large part to the complexity associated with their development and optimization (Lemke *et al.*, *Nature Chem. Biol.* 7:480-483 (2011)).

[0341] One approach to the design of biosensors is through the use of dye-based affinity reagents (Kummer *et al.*, *Chem. Biol.* 20:847-856 (2013); Gulyani *et al.*, *Nature Chem. Biol.* 7:437-444 (2011); Garrett *et al.*, *Biochemistry-U.S.* 47:986-996 (2008); Nalbant *et al.*, *Science* 305:1615-1619 (2004)). The affinity reagent, such as a protein fragment or antibody, binds selectively to the active state of its target and this binding event is associated with a measurable change in fluorescence. Affinity reagent based sensors afford access to endogenous, unmodified protein targets. The use of a dye for signal transduction provides a bright signal due to direct excitation, rather than indirect excitation as in designs utilizing Förster resonance energy transfer between fluorescent proteins. The enhanced signal-to-noise of dye-based sensors relative to other designs also allows for lower excitation light and lower biosensor loading to be used, which minimizes perturbation to normal cellular activity. However one significant drawback associated with affinity reagent based biosensors is that they have historically required microinjection. This is a challenging technique that requires specialized equipment and many labs are not sufficiently motivated to make this part of their routine.

[0342] We report here a new type of biosensor, consisting of a small molecule recognition component coupled to an environmentally sensitive dye, which extends upon the affinity reagent based biosensor concept to provide several additional advantages (Fig. 16a). As opposed to a protein fragment or antibody, a small molecule based biosensor can be introduced into cells through simple passive diffusion. Further, one of the bottlenecks in the development of biosensors for new molecular targets is the lack of suitable affinity reagents. Small molecule based biosensors can be constructed from compounds having known activity profiles and can thus potentially provide access to a wide variety of molecular targets for which no natural binding partner is known. This type of biosensor design has the potential to

afford access to unknown aspects of signaling pathway dynamics in both genetically intractable cells as well as whole organisms.

[0343] **Rational design and synthesis of CaMero:** For the reporter component of the sensor, we chose to work with merocyanine dyes. Merocyanines are characterized by the presence of electron donor and acceptor components linked by a system of conjugation (Kulinich *et al.*, *Russ. Chem. Rev.* 78:141-164 (2009)). These dyes are especially well suited for use in live cell imaging as they are very bright, having some of the largest extinction coefficients reported for organic dyes, they can be tuned to emit at long wavelengths that avoid overlap with cellular autofluorescence, and they are sufficiently photostable to withstand the repeated excitation required in experiments aimed at monitoring protein dynamics over time (Loving *et al.*, *Trends Biotechnol.* 28:73-83 (2010)). In addition to these features, merocyanines are perhaps most notable for their "environment-sensing" capabilities. Changes in solvation environment, such as when going from a more polar, solvated state to one in which the dye is brought into proximity with a protein surface, can be associated with significant changes in fluorescence (Toutchkine *et al.*, *J. Am. Chem. Soc.* 125:4132-4145 (2003); MacNevin *et al.*, *Bioconjug. Chem.* 24:215-223 (2013)). However, as is the case with most organic dyes, the merocyanines are not inherently membrane permeant and therefore required structural optimization to enable the envisioned passive delivery of the small molecule based biosensor. Several analogues of a lead merocyanine compound were prepared with acetoxymethyl ester functional groups in various positions and these dyes were screened for cellular membrane permeability and uniformity of intracellular distribution. The best dye among this screening set was further prepared as a conjugatable version with a carboxylic acid side chain (**Scheme 6**). This new dye, **mero76**, maintained the good brightness and solvent sensitivity characteristics of its non-functionalized parent dye (**Table 7**).

Table 7. Photophysical properties of conjugatable and membrane permeant merocyanines mero76 and mero166. Values are reported as the average of three experiments.					
mero#	solvent	ex $\lambda_{\max}$ (nm)	em $\lambda_{\max}$ (nm)	ϵ	QY
76	DMSO	564	594	70000	0.12
76	MeOH	555	588	56000	0.08
76	water ^a	569	590	28000	0.02
166	DMSO	565	594	100000	0.24
166	MeOH	559	588	91000	0.11
166	water ^a	568	591	30000	0.03
167	DMSO	566	594	116000	0.21
167	MeOH	558	588	85000	0.09
167	water ^a	569	591	21000	0.06

^a Water contained 0.1% DMSO.

[0344] For the recognition component of the sensor, we sought to take advantage of the known binding selectivity of the small molecule trifluoperazine (TFP) for the active conformation of calmodulin (CaM) (Massom *et al.*, *Biochemistry-Us* 29:671-681 (1990)). The calcium-binding protein calmodulin, which is ubiquitously expressed in eukaryotic cells, is involved in numerous processes throughout the cell, including growth, proliferation, and motility (Chin *et al.*, *Trends Cell Biol.* 10:322-328 (2000)). Calmodulin undergoes a distinct conformational shift from a more globular form in its calcium-free state, to an open, "dumbbell" shape once in complex with calcium ions (Fig. 17) (Zhang *et al.*, *Nature Structural Biol.* 2:758-767 (1995)). The subsequent binding of trifluoperazine to calmodulin is contingent upon this conformational change. We therefore hypothesized that a biosensor for calmodulin activation could be developed through the construction of a TFP conjugate containing an environment-sensing merocyanine dye.

[0345] Several crystal structures of the TFP-CaM complex are available (Vandonselaar *et al.*, *Nature Structural Biol.* 1:795-801 (1994); Cook *et al.*, *Biochemistry-Us* 33:15259-15265 (1994); Vertessy *et al.*, *Biochemistry-Us* 37:15300-15310 (1998)). In each case, the structures reveal that the general orientation of the TFP molecule once bound with calmodulin places the *N*-methyl substituent of the TFP piperazine moiety towards solvent (Fig 16b). We saw this position of TFP to be

an ideal point for elaboration with a reactive side chain for the attachment of the dye. Commercially available fluphenazine hydrochloride (**1**, Fig 16c) was converted to the *N*-ethylamine substituted piperazine analogue of TFP, compound **2**, in three steps. Compound **2** was then conjugated with the optimized dye **mero76** to give the small molecule based biosensor CaMero. Previous screening studies had identified the sulfoxide derivative of TFP as a compound having significantly reduced affinity for CaM relative to TFP itself (Levin *et al.*, *J. Pharmacol. Exp. Ther.* 208:454-459 (1979)). Due to the structural similarity of the two compounds, the sulfoxide derivative was thought to be an ideal target for the development of a control probe. Treatment of compound **2** with sodium periodate in dilute acid provided the desired sulfoxide derivative compound **3** (Morrow *et al.*, *Helv. Chim. Acta* 88:962-967 (2005)). Compound **3** was then conjugated with **mero76** to give the negative control probe, CaMero-NC.

[0346] **Biosensor response characteristics *in vitro*:** We first examined the behavior of the two probe molecules *in vitro*. The CaMero probe showed a significant increase in fluorescence intensity (8-fold) with increasing calcium concentration (Fig 18a). No change in CaMero fluorescence intensity was observed with increasing calcium concentration in the absence of calmodulin. This result demonstrated that the probe was not acting as calcium indicator but required the presence of activated calmodulin to generate a fluorescence response. Minimal changes in fluorescence were observed for the negative control probe CaMero-NC with increasing calcium under the same conditions. Gradual reduction in calcium concentration below the level required to generate maximal fluorescence response resulted in a corresponding reduction in fluorescence intensity of CaMero (Fig 18b). Further, only in the presence of saturating levels of calcium ($\geq 10 \mu\text{M}$) did CaMero show an increase in fluorescence intensity with increasing concentration of calmodulin (Fig 18c). Thus the fluorescence response of CaMero was not the result of non-specific binding with calmodulin. The specificity of CaMero fluorescence response for the active, calcium-bound state of calmodulin, in conjunction with the demonstrated reversibility of response, indicated that CaMero had the potential to act as a sensor of calmodulin activation dynamics in more complex systems.

[0347] **Localization and dynamics of CaMero in live cells:** We next examined the intracellular localization of the different probes using live cell imaging. Localization studies were carried out in 1321N1 astrocytoma cells stably expressing the nucleotide receptor P_2Y_6 . Cells were briefly incubated in media containing the test compound, then the treatment media was removed and the cells were washed and imaged. Treatment with the membrane permeant dye component of the sensor resulted in uniformity of fluorescence and no apparent organelle specific interactions in the 1321N1 cells (Fig 19a). The negative control probe CaMero-NC, which contained both dye and small molecule components, proved to maintain membrane permeability and also showed even distribution throughout the cells. Treatment with the active CaMero probe however showed distinct localization in the cells, with maximal fluorescence appearing to be localized generally in the perinuclear region. Co-treatment of cells with a marker for the endoplasmic reticulum (ER) revealed clear co-localization with CaMero.

[0348] We next sought to assess the ability of CaMero to report calmodulin activation dynamics in live cells through modulation of intracellular calcium levels. The P_2Y_6 receptor can be selectively stimulated with the nucleoside uridine diphosphate (UDP) (Nicholas *et al.*, *Mol. Pharmacol.* 50:224-229 (1996)). It is well established that activation of P_2Y_6 leads to elevation of intracellular calcium levels via initiation of phospholipase C mediated cleavage of phosphatidyl-inositol-4,5-bisphosphate (PIP₂) to diacylglycerol and inositol-triphosphate (IP₃) and the subsequent release of calcium from stores located within the ER following IP₃ binding with ER localized IP₃ receptors. Cells were co-treated with either CaMero or CaMero-NC along with a second, non-environment sensing small molecule fluorophore known to distribute evenly throughout the cell that could be used as a control for fluctuation in cellular volume. Stimulation of CaMero treated P_2Y_6 astrocytoma cells with UDP generated a significant increase in fluorescence intensity, whereas no change in intensity was observed in CaMero-NC treated cells (Fig. 19b). Average fluorescence intensity when quantified across the entire cell showed little change following UDP treatment (Fig 19c). Stimulation resulted in characteristic cellular protrusions and this increase in overall area served to offset localized increases in fluorescence intensity. When analysis was constrained to ER rich

perinuclear areas, fluorescence changes were most clearly evident. This particular localization of intensity change is consistent with the signaling cascade leading to intracellular calcium release from the ER. Additional cellular stimulation experiments were carried out using primary murine neurons. Depolarization following treatment with KCl resulted in increased fluorescence intensity when analyzed within the entire cell body region (Fig. 20).

[0349] We also observed the behavior of the CaMero probes in normally migrating mouse embryonic fibroblasts (MEFs). Fluorescence intensity was most noticeably present in perinuclear regions, as previously observed in the 1321N1 cells. However we were surprised to also see a band of intensity at the cell edge in CaMero, but not CaMero-NC, treated cells (Fig. 21a). This apparent area of calmodulin activation appeared predominantly within 3 μ M from the cell edge (Fig 21b). The small molecule fluorophore carboxyfluorescein diacetate (CFDA) and yellow fluorescent protein (YFP) were both used as volume indicators in separate experiments in order to control for any potential artifacts due to differential localization associated with molecular size. Experiments using either indicator gave similar results with respect to both intensity of signal and depth of range (Fig. 22). Pre-treatment of the cells with clozapine to minimize potential off-target binding of the probes with cell surface GPCRs did not affect the localization or intensity of signal of CaMero (Fig. 22). Cell edge velocity was calculated such that protruding areas of the cell were given a positive velocity and retracting areas a negative velocity. Correlation analysis of cell edge velocity with respect to fluorescence intensity of CaMero revealed an interaction between the two variables that was especially prominent in the positive direction (Fig 21c, pos. velocity: $p < 0.001$, $n = 58$; neg. velocity: $p < 0.05$, $n = 56$). Although limited correlation between CaMero intensity at lower negative velocity values was found, at the more extreme ranges of negative velocity, such as in areas of tail retraction, a strong correlation was present. Increased intensity of the probe was clearly visible in these areas.

[0350] **Whole animal treatment with CaMero:** CaMero or CaMero-NC was injected into GFP-LifeAct mice and sections of intestine were imaged ex vivo using multi-photon microscopy. CaMero showed distinctly more intense fluorescence than CaMero-NC in the excised tissue samples taken from probe treated mice (Fig 23a).

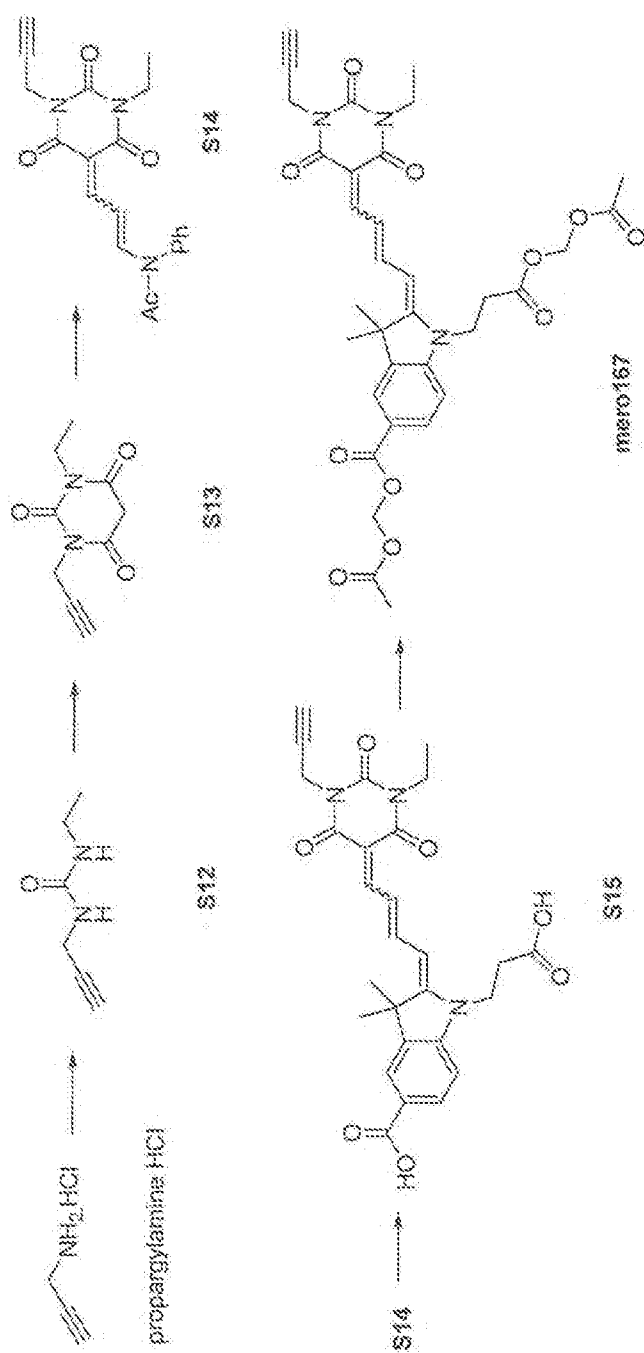
Image stacks were acquired starting at the base of the crypt of Lieberkühn ($Z = 0 \mu\text{m}$, Fig 23b) downward through the circular ($Z = -56 \mu\text{m}$) and longitudinal smooth muscle layers ($-84 \mu\text{m}$), which were visualized using GFP-LifeAct (left column). The CaMero localization pattern (right column) corresponds to the fiber direction in each muscle layer. CaMero-NC did not clearly localize to smooth muscle cells as indicated by minimal fluorescence present in those areas. The excised tissue maintains contractility ex vivo for sustained periods. An analysis of fluorescence intensity over time showed an oscillation in intensity for the CaMero probe at a frequency consistent with that of tissue contraction. No such periodic oscillation was observed in the CaMero-NC treated tissue.

[0351] Biosensors applied in living cells have provided valuable insight into the dynamics of signaling networks, enabling quantification of the kinetics and localization of protein activity. That calmodulin may play such diverse roles within the cell underscores the complexity associated with understanding the molecular mechanisms by which calcium signals are integrated into specific cellular responses. We have shown that a biosensor based on a small molecule that is specific for the active state of calmodulin can be attached to an environment-sensing dye, introduced into cells using non-invasive methods, and be applied therein to monitor target protein activation dynamics.

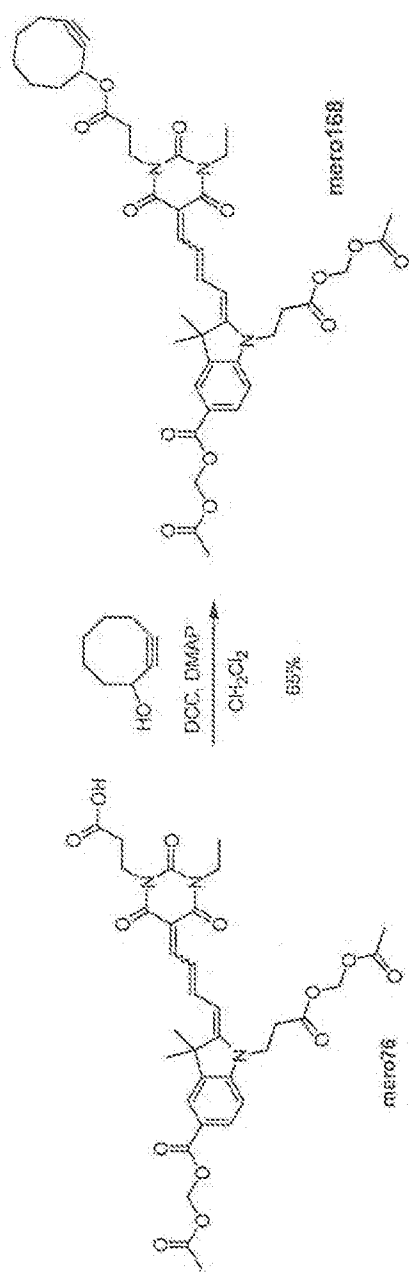
Example 10

Additional Membrane Permeable Dyes

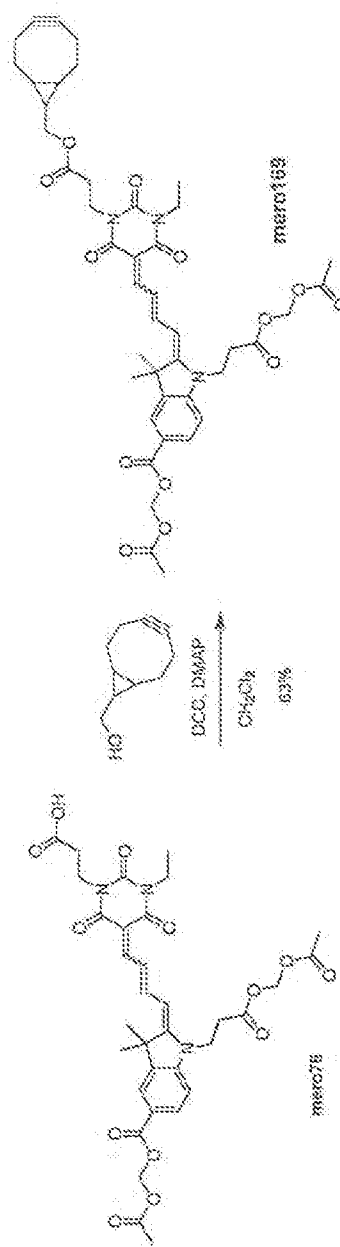
[0352] Membrane permeable dyes **mero167**, **mero168**, and **mero169** were synthesized as shown in schemes 8-10.



Scheme 8. Synthesis of macro 167



Scheme 9. Synthesis of mero163



Scheme 10. Synthesis of mero169

[0353] **1-ethyl-3-(prop-2-yn-1-yl)urea (S12).** Propargylamine hydrochloride (0.976 g, 10.92 mmol) was diluted in THF (10 mL) in an oven dried flask under argon. Triethylamine (1.6 mL, 11.47 mmol) was added dropwise and the solution was cooled in a cold water bath. Ethylisocyanate (0.86 mL, 10.92 mmol) was then added dropwise and the reaction was stirred for 16 h. The solution was combined with half saturated aq. NH_4Cl (15 mL) and diluted with CH_2Cl_2 (15 mL). The aqueous layer was extracted with CH_2Cl_2 (3 x 15 mL). Organic layers were combined, washed with brine, dried with MgSO_4 , filtered, concentrated, and dried under vacuum. Isolated 0.761 g (57%) white solid that required no further purification. ^1H NMR (400 MHz, CDCl_3) δ 5.38 (t, $J = 5.9$ Hz, 1H), 5.19 (t, $J = 5.9$ Hz, 1H), 3.97 (dd, $J = 5.7$ Hz, 2H), 3.21 (dq, $J = 7.3$ Hz, 2H), 2.19 (t, $J = 2.7$ Hz, 1H), 1.12 (t, $J = 7.2$ Hz, 3H). ^{13}C NMR (100 MHz, CDCl_3) δ 158.4, 81.2, 71.0, 35.4, 30.2, 15.6. MS-ESI m/z 127.10 ($[\text{M} + \text{H}]^+$ requires 127.09).

[0354] **1-ethyl-3-(prop-2-yn-1-yl)pyrimidine-2,4,6(1H,3H,5H)-trione (S13).** Compound S12 (0.708 g, 5.61 mmol) and malonic acid (0.61 g, 5.89 mmol) were added to a flame dried flask and the flask was sealed, evacuated, and argon flushed. Acetic acid (8.0 mL) was added and the flask was heated to 60 °C for 10 min. Acetic anhydride (2.12 mL, 22.4 mmol) was added and the reaction was further heated to 95 °C and stirred for 1 h. The reaction mixture was then concentrated and the residue was loaded in a minimum amount of CH_2Cl_2 onto a 24 g silica column and eluted with 0–2% MeOH in CH_2Cl_2 . Main product containing fractions were combined and concentrated to give 0.852 g (78%) amber oil. ^1H NMR (400 MHz, CDCl_3) δ 4.64 (d, $J = 2.7$, 2H), 3.96 (q, $J = 7.0$, 2H), 3.70 (s, 2H), 2.22 (t, $J = 2.6$ Hz, 1H), 1.22 (t, $J = 7.0$, 3H). ^{13}C NMR (100 MHz, CDCl_3) δ 164.2, 163.9, 150.6, 77.6, 71.8, 39.8, 37.7, 31.2, 13.4. MS-ESI m/z 195.17 ($[\text{M} + \text{H}]^+$ requires 195.08).

[0355] ***N*-((1*E*,3*E*)-3-(1-ethyl-2,4,6-trioxo-3-(prop-2-yn-1-yl)tetrahydropyrimidin-5(2*H*)-ylidene)prop-1-en-1-yl)-*N*-phenylacetamide (S14).** Compound S13 (0.84 g, 4.31 mmol), malonaldehyde dianilide hydrochloride (1.41 g, 6.46 mmol), and sodium acetate (0.40 g, 6.46 mmol) were diluted in acetic anhydride (7.0 mL). The reaction mixture was added to a pre-heated oil bath at 100 °C and stirred for 1 h. The reaction was cooled and added in portions to saturated aq. NaHCO_3 (50 mL). The aq. layer was extracted with EtOAc (3 X 50 mL). The

organic layers were combined, washed with brine, dried with NaSO₄, filtered, and concentrated. The residue was diluted in CH₂Cl₂ and concentrated with Celite. The Celite cake was eluted on an 80 g silica column with 0 – 50% EtOAc in hexanes over 30 min. The product was isolated as 0.928 g (59%) yellow solid as an approximate 1:1 mixture of *E* and *Z* isomers. Isomer A: ¹H NMR (400 MHz, CDCl₃) δ 8.53 (d, *J* = 13.6 Hz, 1H) 8.20 (d, *J* = 12.4 Hz, 1H), 7.65 – 7.53 (m, 3H), 7.26 – 7.20 (m, 2H), 6.86 (dd, *J* = 12.6 Hz, 1H), 4.58 (d, *J* = 2.8 Hz, 2H), 3.99 (q, *J* = 7.2 Hz, 2H), 2.18 (t, *J* = 2.4 Hz, 1H), 2.00 (s, 3H), 1.21 (t, *J* = 7.2 Hz, 3H), Isomer B: ¹H NMR (400 MHz, CDCl₃) δ 8.52 (d, *J* = 14.0 Hz, 1H), 8.20 (d, *J* = 12.4 Hz, 1H), 7.65 – 7.53 (m, 3H), 7.26 – 7.20 (m, 2H), 6.85 (dd, *J* = 12.4 Hz, 1H), 4.69 (d, *J* = 2.8 Hz, 2H), 3.87 (q, *J* = 7.2 Hz, 2H), 2.14 (t, *J* = 2.6 Hz, 1H), 2.01 (s, 3H), 1.13 (t, *J* = 7.0 Hz, 3H). MS-ESI *m/z* 388.26 ([M + Na]⁺ requires 388.13).

[0356] (*E*)-1-(2-carboxyethyl)-2-((2*E*,4*E*)-4-(1-ethyl-2,4,6-trioxo-3-(prop-2-yn-1-yl)tetrahydropyrimidin-5(2*H*)-ylidene)but-2-en-1-ylidene)-3,3-dimethylindoline-5-carboxylic acid (S15). Compound S14 (0.868 g, 2.38 mmol), sodium acetate (0.293 g, 3.57 mmol), and compound 9 (1.272 g, 3.57 mmol) were diluted in 1:1 MeOH/CHCl₃ (15 mL). The reaction flask was topped with a condenser with argon inlet and heated to reflux for 16 h. The reaction was concentrated and re-diluted in EtOAc and minimal amount of MeOH. The residue was prepared as a Celite cake and eluted on a 40 g silica column with 0 – 5% MeOH in CH₂Cl₂ over 36 min. Main product isolated as 0.352 g (29%) dark blue solid. δ ¹H NMR (400 MHz, DMSO-*d*₆) δ 12.59 (bs, 2H), 8.29 – 8.14 (m, 2H), 7.99 (d, *J* = 1.6 Hz, 1H), 7.92 (dd, *J* = 8.4, 1.6 Hz, 1H), 7.83 – 7.73 (m, 1H), 7.29 (d, *J* = 8.4 Hz, 1H), 6.27 (d, *J* = 13.2 Hz, 1H), 4.55 (s, 2H), 4.23 (t, *J* = 6.8 Hz, 2H), 3.90 – 3.80 (m, 2H), 3.11 – 3.08 (m, 1H), 2.66 (t, *J* = 7.1 Hz, 2H), 1.64 (s, 6H), 1.12 – 1.10 (m, 3H). MS-ESI *m/z* 528.38 ([M + Na]⁺ requires 528.17).

[0357] Acetoxymethyl (*E*)-1-(3-(acetoxymethoxy)-3-oxopropyl)-2-((2*E*,4*E*)-4-(1-ethyl-2,4,6-trioxo-3-(prop-2-yn-1-yl)tetrahydropyrimidin-5(2*H*)-ylidene)but-2-en-1-ylidene)-3,3-dimethylindoline-5-carboxylate (mero167). Compound S15 (0.323 g, 0.638 mmol) was diluted in ACN (35 mL) and bromomethyl acetate (0.27 mL, 2.55 mmol) was added, followed by addition of *N,N*-diisopropylethylamine (0.49 mL, 2.55 mmol). The flask was capped and fitted with an argon inlet line and the

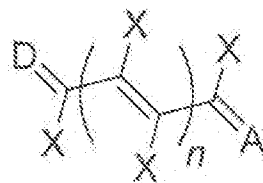
reaction was stirred at room temperature for 16 h. The reaction mixture was then concentrated, re-diluted in CH_2Cl_2 , and prepared as a Celite cake. The sample was eluted on a 24 g silica column with 0 – 50% EtOAc in hexanes over 25 min. Isolated 0.292 g (71%) dark blue solid. ^1H NMR (500 MHz, $\text{DMSO}-d_6$) δ 8.28 (d, $J = 13.0$ Hz, 0.5H), 8.27 (d, $J = 13.0$ Hz, 0.5H), 8.21 (t, $J = 13.0$ Hz, 1H), 8.01 (d, $J = 1.0$ Hz, 1H), 7.94 (dd, $J = 8.5, 1.5$ Hz, 1H), 7.82 (t, $J = 13.3$ Hz, 0.5H), 7.80 (t, $J = 13.0$ Hz, 0.5H), 7.32 (d, $J = 8.5$ Hz, 1H), 6.28 (d, $J = 13.0$ Hz, 1H), 5.92 (s, 2H), 5.62 (s, 2H), 4.55 (t, $J = 2.5$ Hz, 2H), 4.25 (t, $J = 6.8$ Hz, 2H), 3.90 – 3.80 (m, 2H), 3.14 – 3.10 (m, 1H), 2.82 (t, $J = 7.0$ Hz, 2H), 2.10 (s, 3H), 2.01 (s, 3H), 1.64 (s, 6H), 1.12 (t, $J = 7.0$ Hz, 1.5H), 1.11 (t, $J = 7.0$ Hz, 1.5H). MS-ESI m/z 672.40 ($[\text{M} + \text{Na}]^+$ requires 672.22).

[0358] Although the foregoing subject matter has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be understood by those skilled in the art that certain changes and modifications can be practiced within the scope of the appended claims.

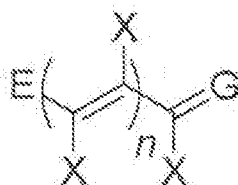
[0359] All publications, patent applications, patents, and other references are herein incorporated by reference to the same extent as if each individual publication, patent application, patent, and other reference was specifically and individually indicated to be incorporated by reference. It will be understood that, although a number of patent applications, patents, and other references are referred to herein, such reference does not constitute an admission that any of these documents forms part of the common general knowledge in the art.

THAT WHICH IS CLAIMED:

I. A compound of Formula I or Formula II:



I



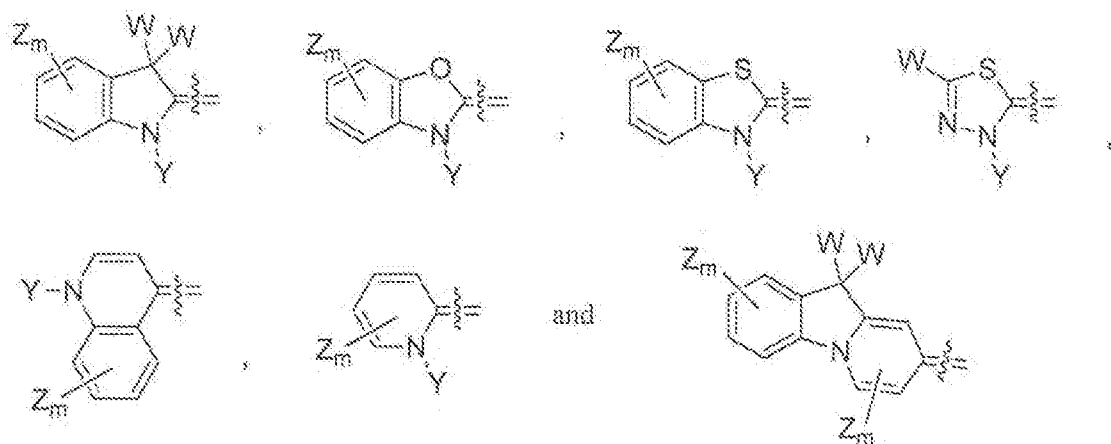
II

wherein:

n is an integer from 0 to 2;

each X is independently hydrogen or an electron withdrawing group selected from the group consisting of carbonyl, cyano, halogen, nitro, sulfonyl, and trifluoromethyl;

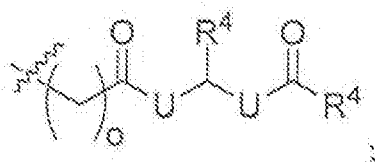
D is selected from the group consisting of:



W is optionally substituted C_{1-6} alkyl;

m is an integer from 0 to 4;

Y and Z are each independently selected from the group consisting of hydrogen, C₁₋₆ alkyl, a conjugatable side chain, and

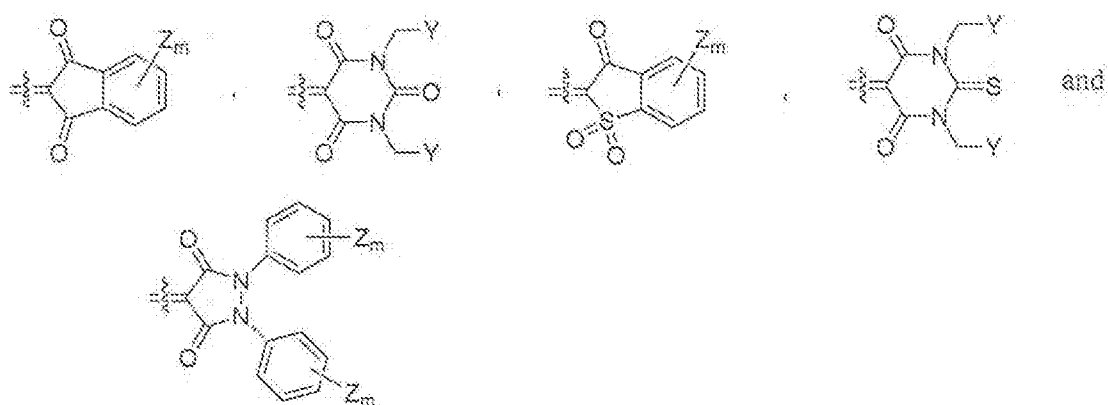


o is an integer from 0 to 5;

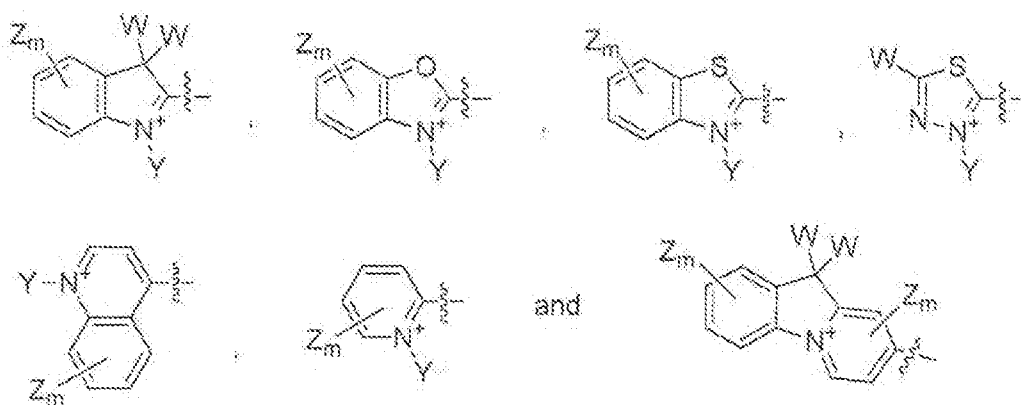
U is O or S;

R⁴ is hydrogen or optionally substituted C₁₋₆ alkyl;

A is selected from the group consisting of:

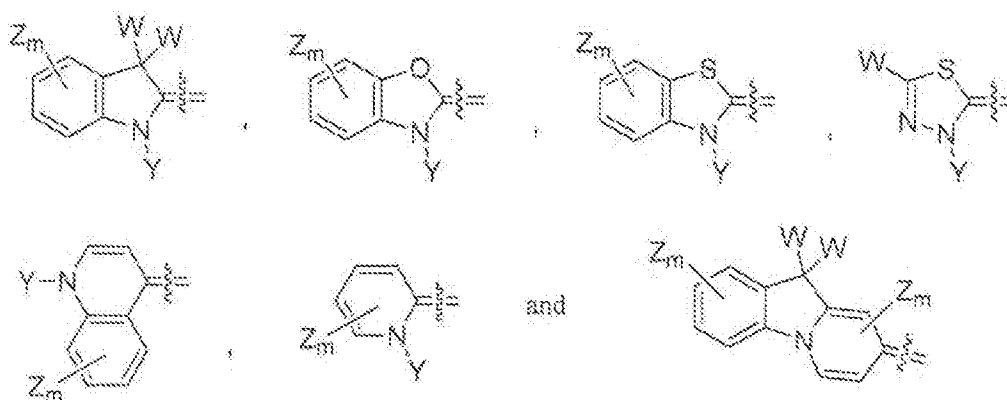


E is selected from the group consisting of:

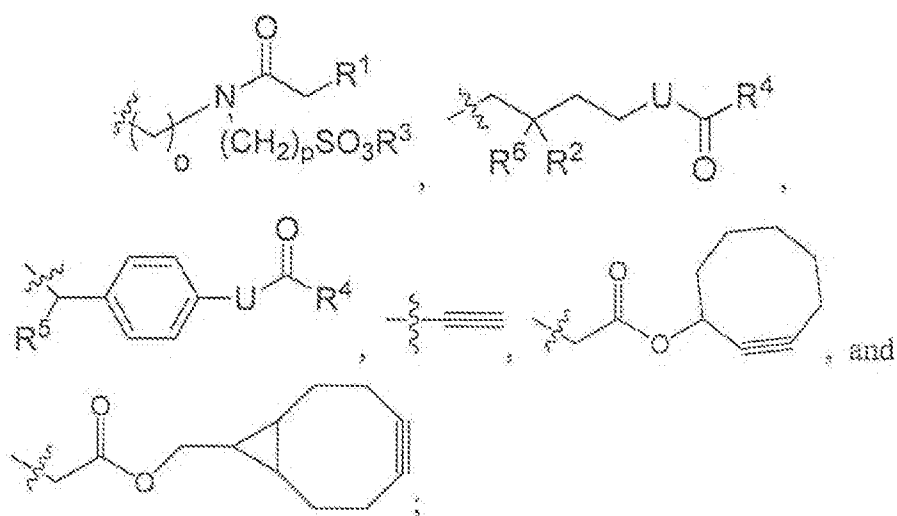


and

G is selected from the group consisting of:



2. The compound of claim 1, wherein the conjugatable side chain is selected from the group consisting of:



wherein:

o is an integer from 1 to 5;

p is an integer from 1 to 5;

U is O or S;

R¹ is a leaving group;

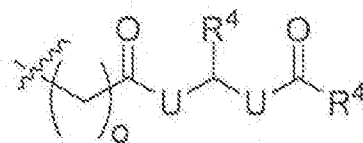
R³ is hydrogen, Li⁺, Na⁺ or K⁺;

R², R⁴, and R⁶ are each independently C₁₋₆ alkyl; and

R⁵ is C₁₋₆ alkyl or CF₃.

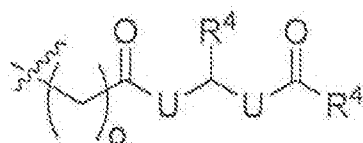
3. The compound of claim 1 or 2, wherein at least one of Y and Z is a conjugatable side chain and the compound contains only one conjugatable side chain.

4. The compound of any one of claims 1-3, wherein the Y in the D or E moiety is a conjugatable side chain and at least one Z is

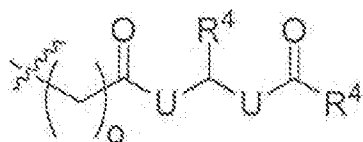


5. The compound of any one of claims 1-3, wherein the Y in the D or E moiety is hydrogen and one Z is a conjugatable side chain.

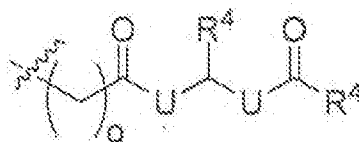
6. The compound of any one of claims 1-3, wherein the Y in the D or E moiety is hydrogen and at least one Z is



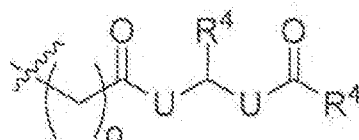
7. The compound of any one of claims 1-3, wherein the Y in the D or E moiety is hydrogen, one Z is a conjugatable side chain, and at least one Z is



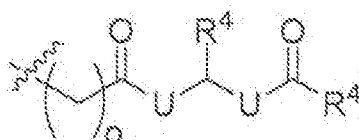
8. The compound of any one of claims 1-3, wherein the Y in the D or E moiety and at least one Z are



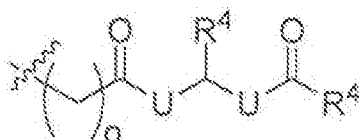
9. The compound of any one of claims 1-3, wherein one Z is a conjugatable side chain and the Y in the D or E moiety and at least one Z are



10. The compound of any one of claims 1-9, wherein the compound comprises at least two Z moieties that are

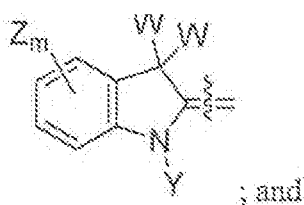


11. The compound of any one of claims 1-10, wherein the compound comprises at least three Z moieties that are

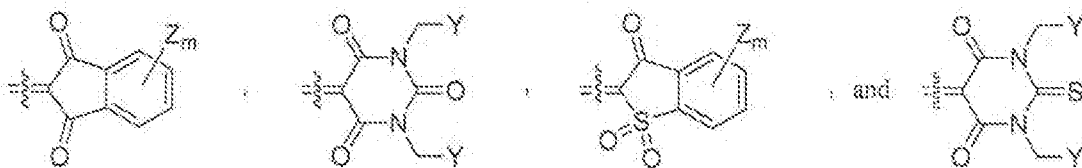


12. The compound of any one of claims 1-11, wherein only one X is an electron withdrawing group.

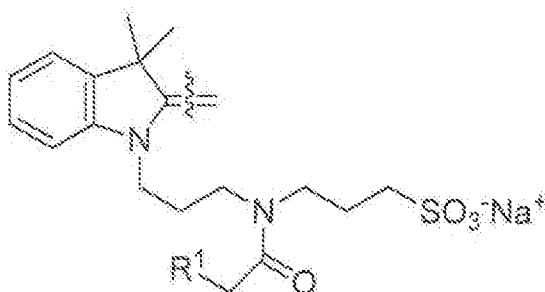
13. The method of claim 1 or 2, wherein:
D is



A is selected from the group consisting of

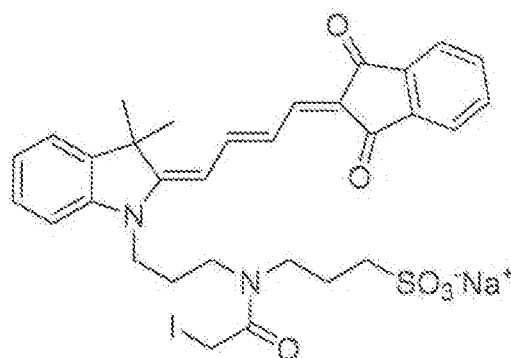


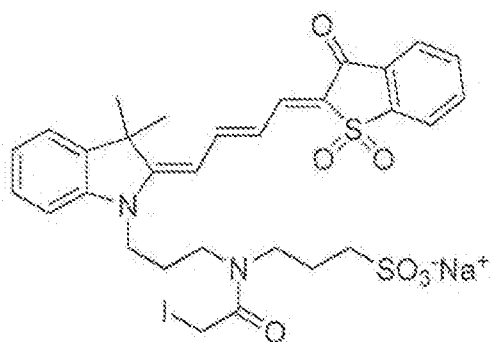
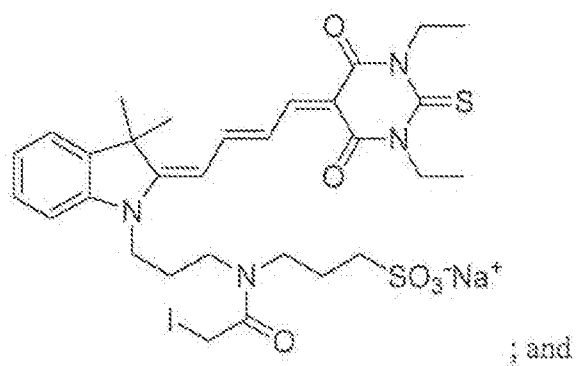
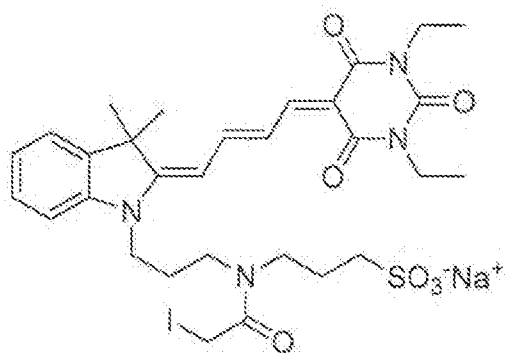
14. The method of claim 13, wherein D is



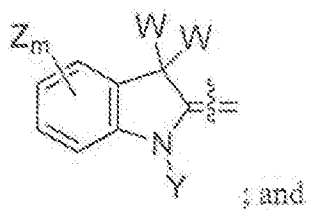
15. The method of claim 14, wherein R¹ is I.

16. The method of claim 15, wherein the compound is selected from:

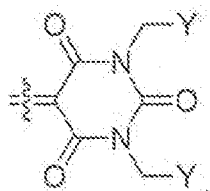




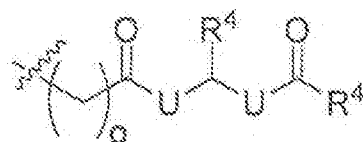
17. The method of claim 1, wherein:
Dis:



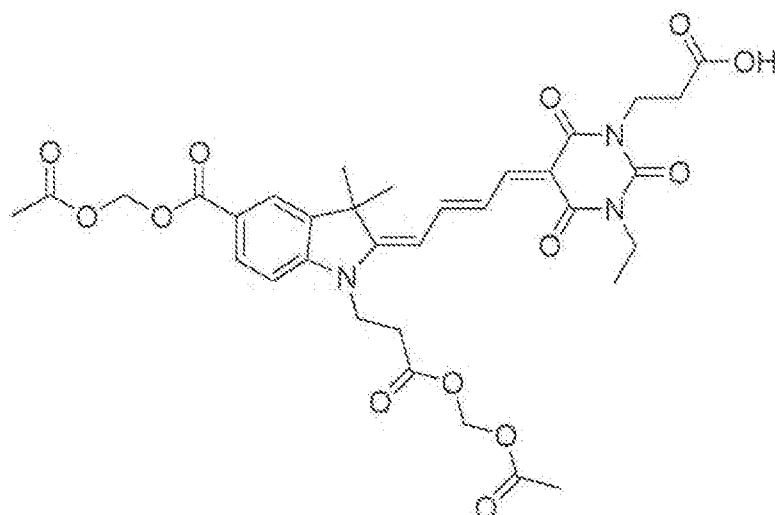
Abstract



18. The method of claim 17, wherein at least one Y is:

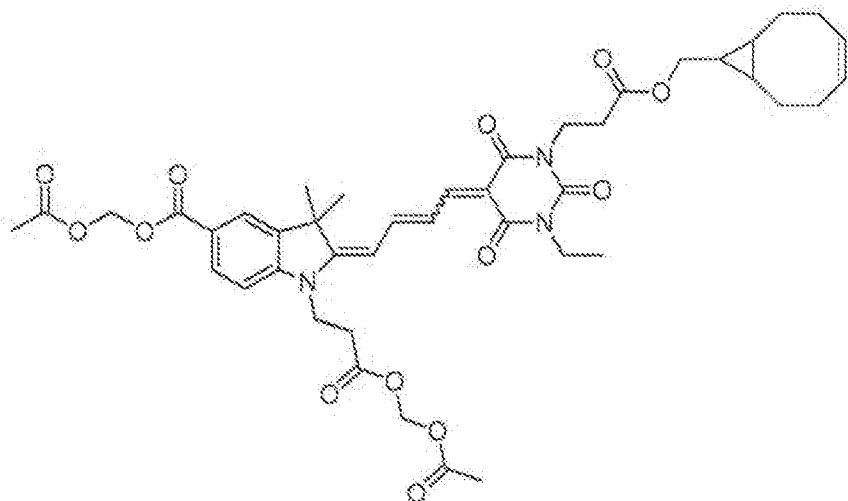


19. The method of claim 18, wherein the compound is selected from:



CC(=O)OCCOC(=O)c1ccc2c(c1)c3c(c2)C(C)(C)C(=C/C=C/C=C/C=C/C4C(=O)N(CC)C(=O)N(CC)C4=O)N2CCOC(=O)C[illegible]CC(=O)OCCOC(=O)c1ccc2c(c1)c(c3c2C(C)(C)C=C/C=C/C4=CC(=O)N(CC)C(=O)N4CCOC(=O)C5=CC=CC=C5)N2CCOC(=O)C(=O)C

, and



20. A biosensor comprising the compound of any one of claims 1-19 conjugated to a binding member having an affinity for a target molecule.
21. The biosensor of claim 20, wherein the compound is conjugated to the binding member through a conjugatable side chain on the compound.
22. The biosensor of claim 20 or 21, wherein the binding member is selected from the group consisting of a nucleic acid, a polypeptide, and a binding protein.
23. The biosensor of any one of claims 20-22, wherein the binding member has a binding domain comprising at least one of an aptamer, a complementary determining region (CDR), a V_H region, a V_L region, a Fv fragment, an F(ab) fragment, an F(ab')₂ fragment, an antibody, an antibody fragment, a leucine zipper, a histone, an enhancer, a single chain variable fragment (scFv), a ligand, a receptor, one protein in a protein complex, a lectin, and combinations thereof.
24. The biosensor of any one of claims 20-23, wherein the target molecule is selected from the group consisting of a nucleic acid and a protein.

25. The biosensor of any one of claims 20-24, wherein the binding member binds to the target molecule when the target molecule is in a specific conformation, is bound to a specific ligand, or has been posttranslationally modified.
26. The biosensor of any one of claims 20-25, wherein the target molecule has been posttranslationally modified by phosphorylation.
27. The biosensor of any one of claims 20-26, wherein the compound of Formula I or Formula II is conjugated with the binding member at a position on the binding member selected from the group consisting of a cysteine, a lysine, an arginine, a natural amino acid side chain, a derivatized amino acid side chain, and an unnatural amino acid side chain.
28. The biosensor of any one of claims 20-27, wherein two or more compounds of Formula I or Formula II are conjugated with the binding member.
29. A biosensor device comprising the biosensor of any one of claims 20-28.
30. A reagent for determining the presence or amount of one or more target molecules in a sample, the reagent comprising the biosensor of any one of claims 20-28.
31. A kit for determining the presence or amount of one or more target molecules in a sample, the kit comprising the biosensor device of claim 29 or the reagent of claim 30.
32. A method for determining the presence or amount of one or more target molecules in a sample, the method comprising:
- (a) providing the biosensor any one of claims 20-28;
 - (b) contacting the biosensor with a sample suspected of containing one or more target molecules to bind the one or more target molecules, if present, to the binding member;

- (c) irradiating the sample suspected of containing one or more target molecules with electromagnetic radiation to induce the compound of Formula I or Formula II to fluoresce; and
 - (d) detecting a fluorescence property of the compound of Formula I or Formula II to determine the presence or amount of one or more target molecules in the sample.
33. The method of claim 32, further comprising:
- (a) measuring a fluorescence intensity at a first emission wavelength before contacting the biosensor with a sample suspected of containing one or more target molecules;
 - (b) measuring a fluorescence intensity at a second emission wavelength after contacting the biosensor with a sample suspected of containing one or more target molecules; and
 - (c) determining a ratio of the second emission wavelength to the first emission wavelength to determine the presence or amount of one or more target molecules in the sample.
34. The method of claim 33, further comprising continuously:
- (a) contacting the biosensor with the sample suspected of containing one or more target molecules;
 - (b) irradiating the sample with electromagnetic radiation; and
 - (c) detecting the fluorescence property of the compound of Formula I or Formula II.
35. A method of detecting an activity or a location of one or more target molecule within a cell, the method comprising:
- (a) providing the biosensor of any one of claims 20-28;
 - (b) contacting the biosensor with a cell suspected of containing one or more target molecules to bind the one or more target molecules, if present, to the binding member;

- (c) irradiating the cell suspected of containing one or more target molecules with electromagnetic radiation to induce the compound of Formula I or Formula II to fluoresce; and
- (d) detecting one or more of:
 - (i) a fluorescence property of the compound of Formula I or Formula II;
 - (ii) a change in a fluorescence property of the compound of Formula I or Formula II;
 - (iii) a location of a fluorescence of the compound of Formula I or Formula II; and
 - (iv) combinations thereof,

to determine the an activity or location of one or more target molecules in the cell.

36. The method of claim 35, wherein the target molecule is selected from the group consisting of a protein, a receptor, a ligand, and an enzyme.

37. The method of claim 35, wherein the cell is selected from the group consisting of a living cell, a cell lysate, a cell library, or a cell culture.

38. The method of claim 35, wherein the activity or location of the one or more target molecules is selected from the group consisting of a phosphorylation state, a subcellular location, an interaction with one or more subcellular structures, and an interaction with one or more cellular proteins.

39. A method of detecting an interaction between an endogenous target molecule and a cellular entity, the method comprising:

- (a) providing a cell comprising an endogenous target molecule;
- (b) providing the biosensor of any one of claims 20-28;
- (c) observing a background fluorescence signal from the biosensor;
- (d) contacting the biosensor with the cell; and
- (e) detecting a change in fluorescence from the biosensor to indicate an interaction between the target molecule and the cellular entity.

40. The method of claim 39, wherein the cellular entity is selected from the group consisting of a cellular nucleic acid, a protein, a peptide, an enzyme, a receptor, a cytokine, a cytoskeleton, and a signal transduction protein.
41. The method of claim 39, wherein the binding member of the biosensor binds the target molecule at a phosphorylation site.
42. The method of claim 39, wherein the binding member has a specific affinity for a conformation, a ligand interaction, or posttranslational modification of the target molecule.
43. The method of claim 39, wherein the compound of Formula I or Formula II is linked to the binding member at a site that does not substantially interfere with binding between the probe and the target biomolecule.
44. The method of claim 43, wherein the site is selected by an examination of a crystal structure for the binding member or the target molecule.
45. The method of claim 39, further comprising introducing the biosensor into the cell by using a technique selected from the group consisting of electroporation, transduction, microporation, and microinjection.
46. The method of claim 39, wherein the detecting a change in fluorescence comprises quantifying a protein amount, locating a protein, detecting a conformational change in the target molecule, detecting activation of the target molecule, or detecting phosphorylation of the target molecule.
47. A kit comprising the biosensor of any one of claims 20-28 for detecting, monitoring or observing a target molecule, wherein the biosensor comprises a compound of Formula I or Formula II and a binding member having a specific affinity for the target molecule.

48. The kit of claim 47 further comprising instructions for using the biosensor to detect, monitor, or observe a target molecule.
49. A biosensor comprising a compound having an affinity for a target molecule conjugated to a dye.
50. The biosensor of claim 49, wherein the compound has a molecular weight less than about 1000 Da.
51. The biosensor of claim 49 or 50, wherein the dye is a cell membrane permeable dye.
52. The biosensor of any one of claims 49-51, wherein the dye can be visualized by microscopy.

FIG. 1

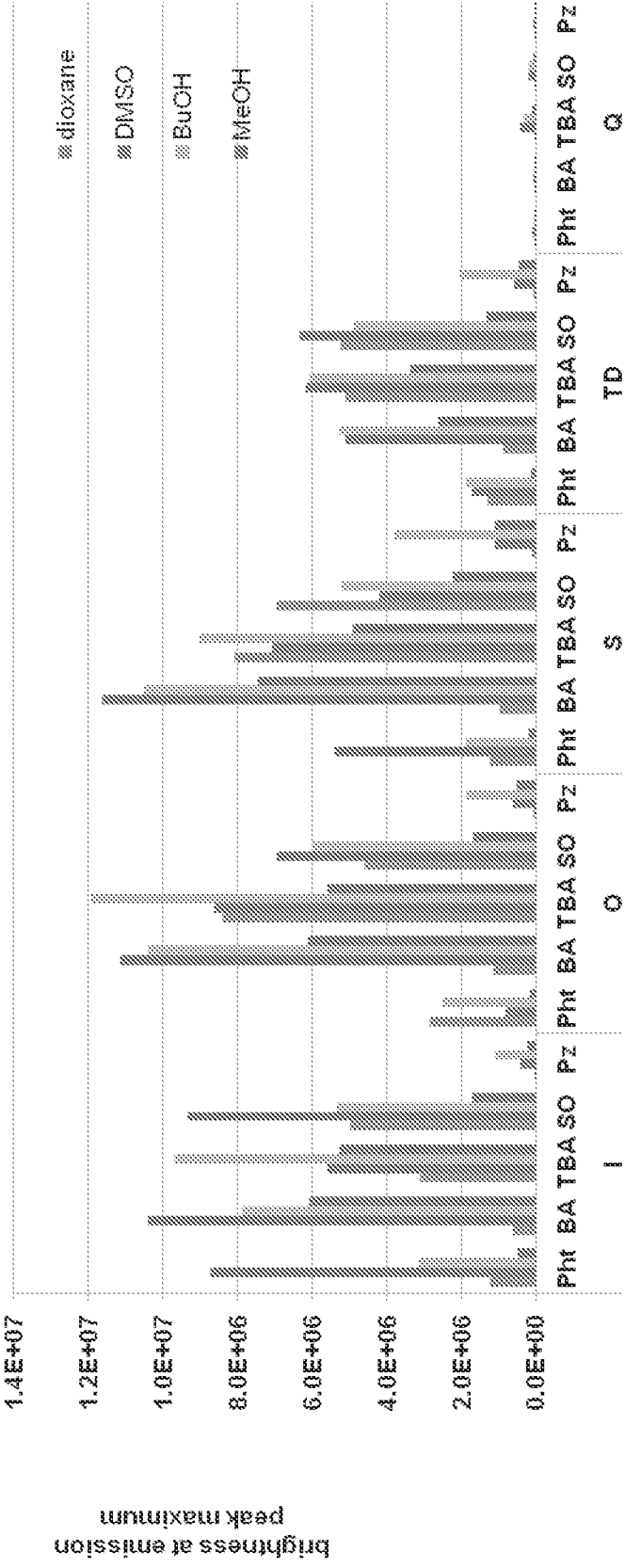


FIG. 2A

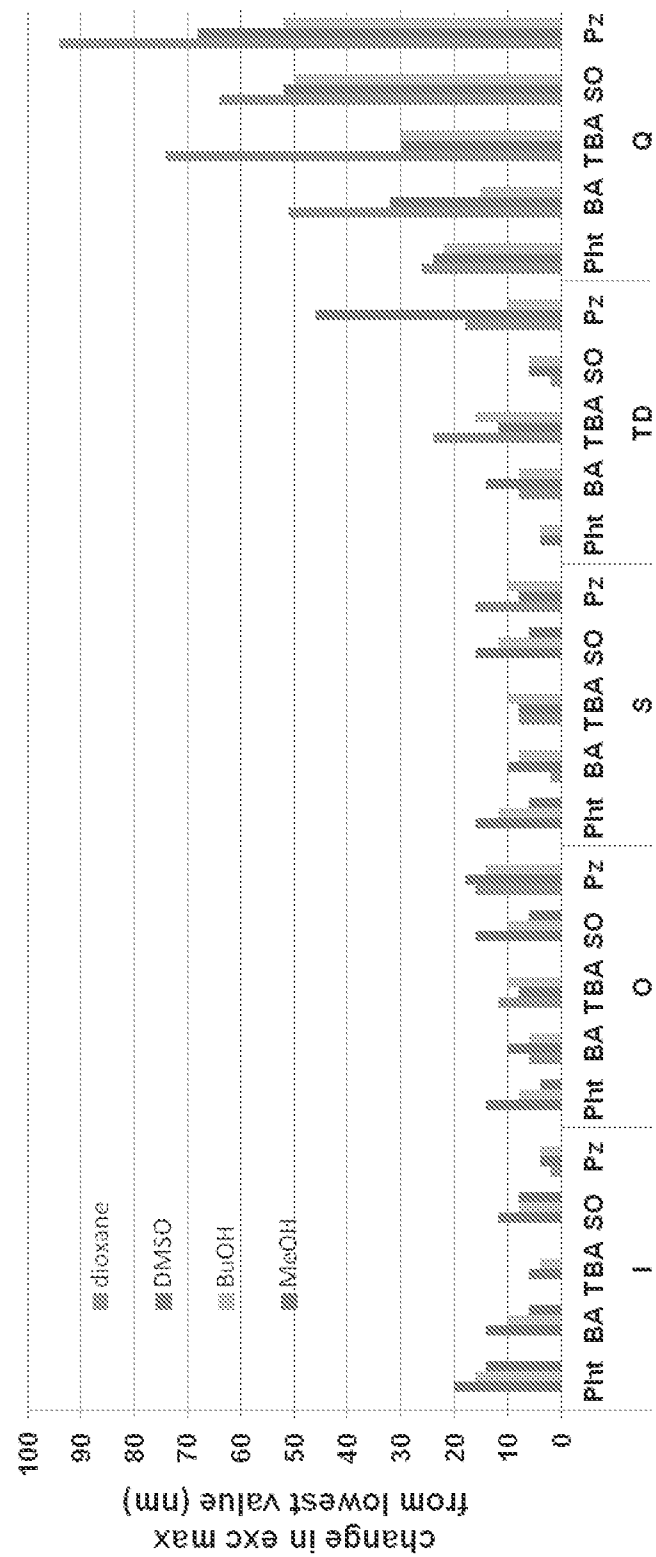


FIG. 2B

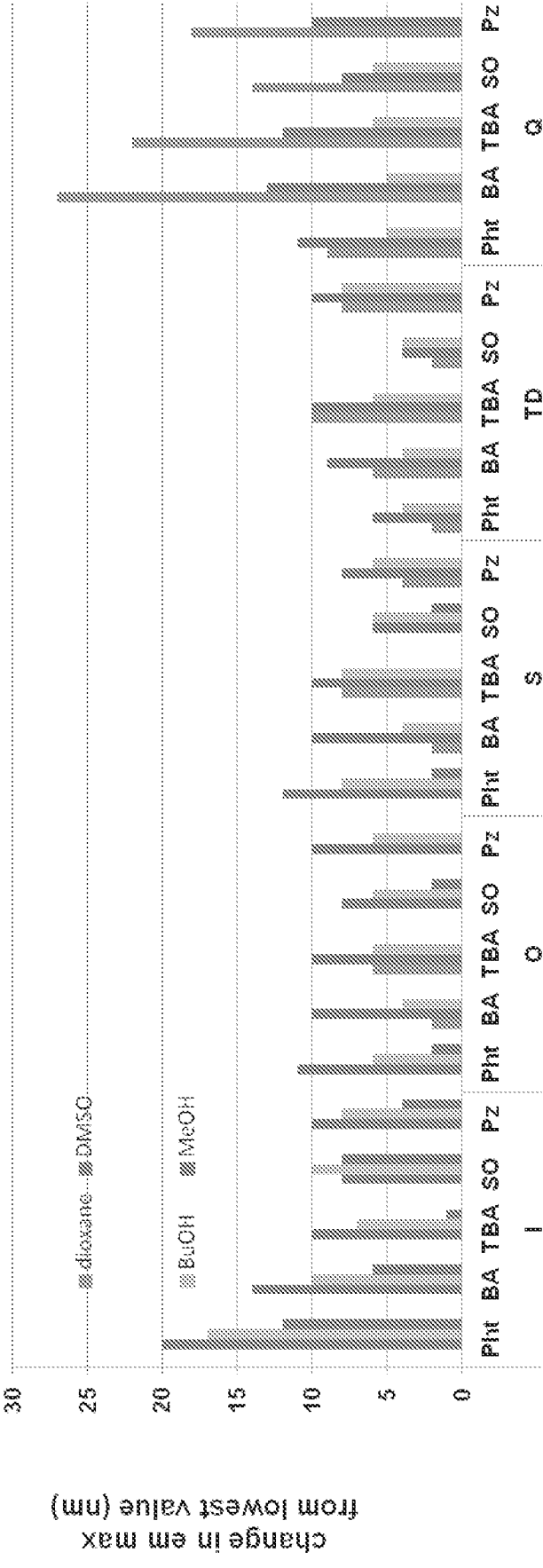


FIG. 3A

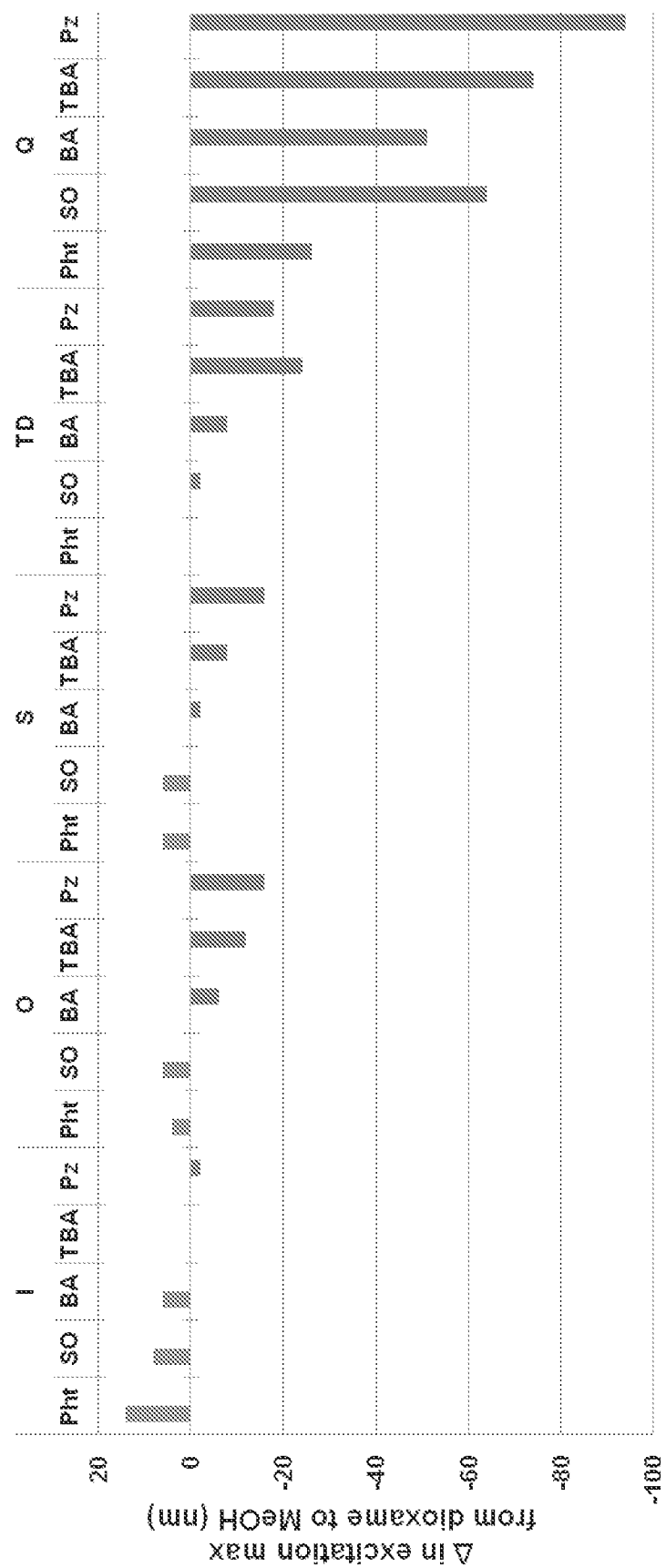


FIG. 3B

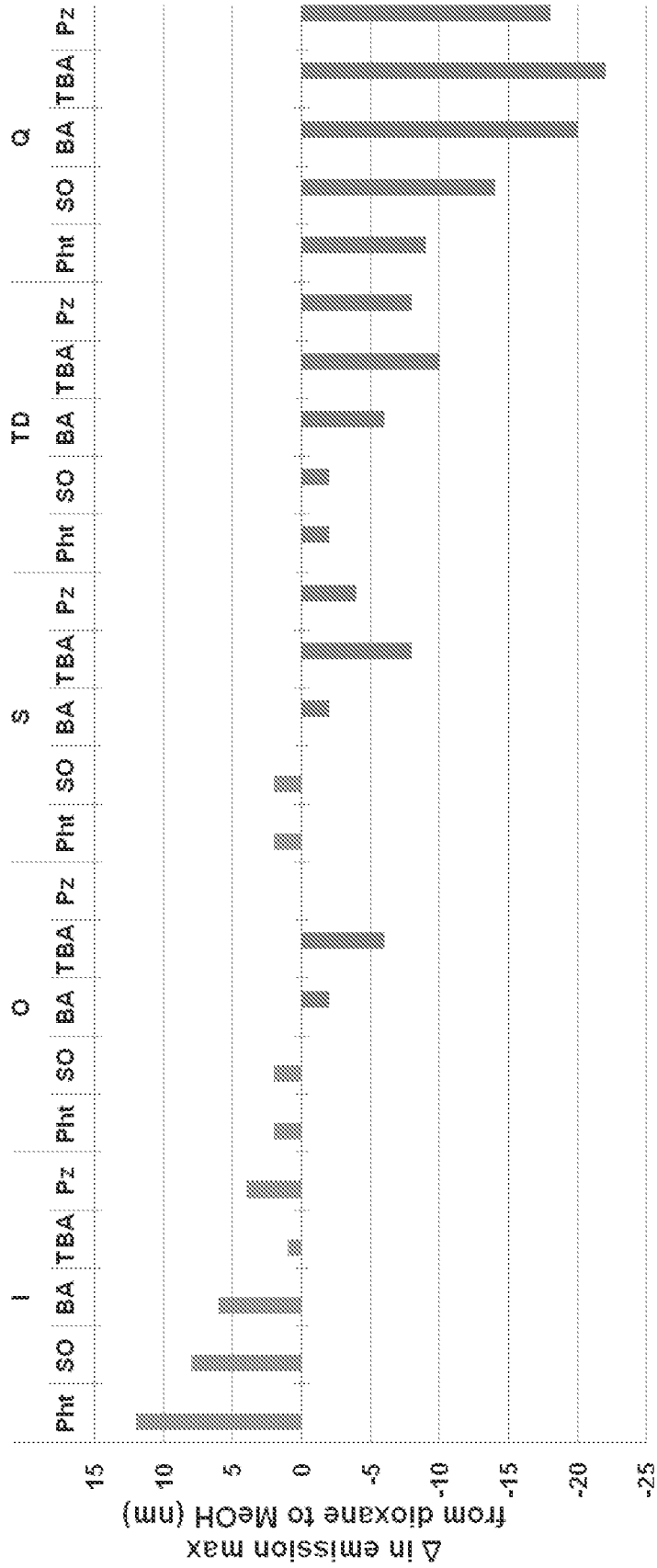


FIG. 4

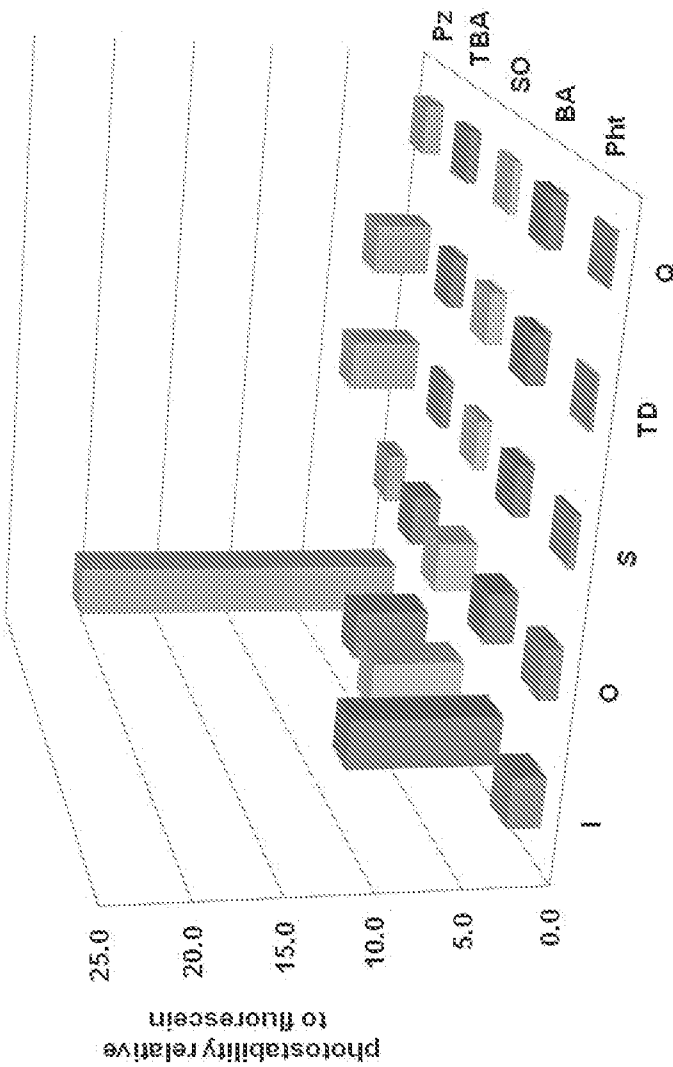


FIG. 5A

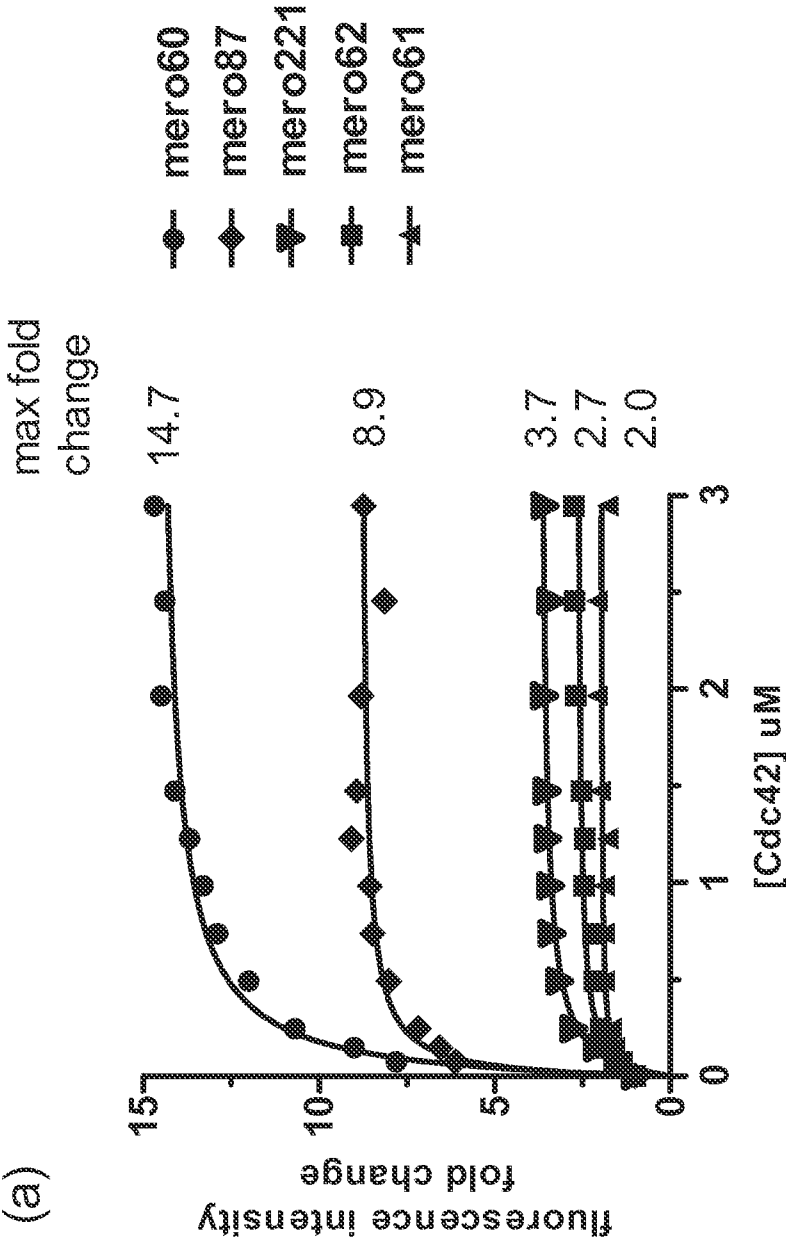


FIG. 5B

(b)

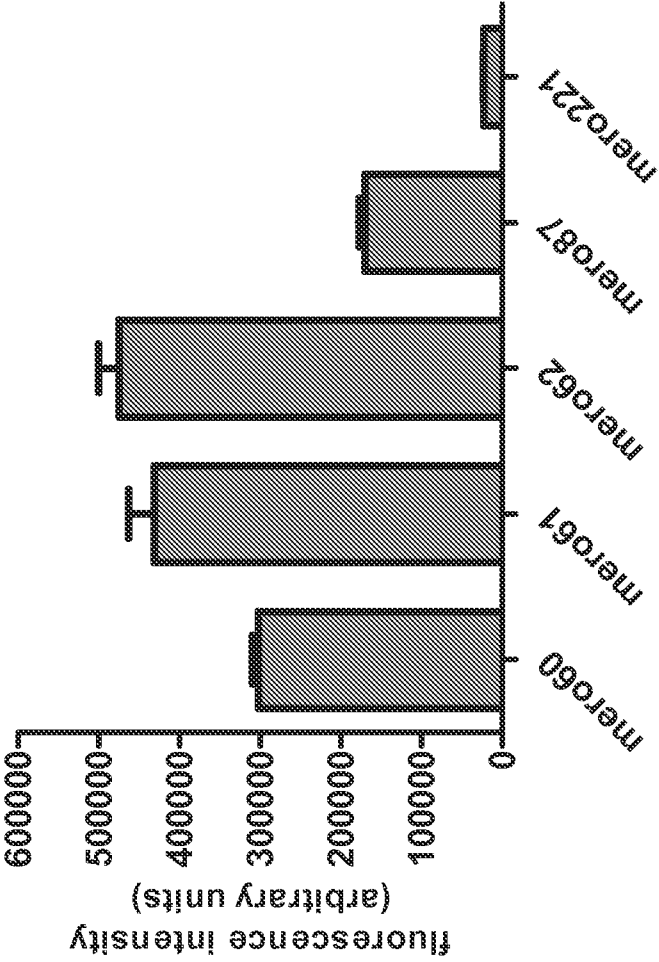
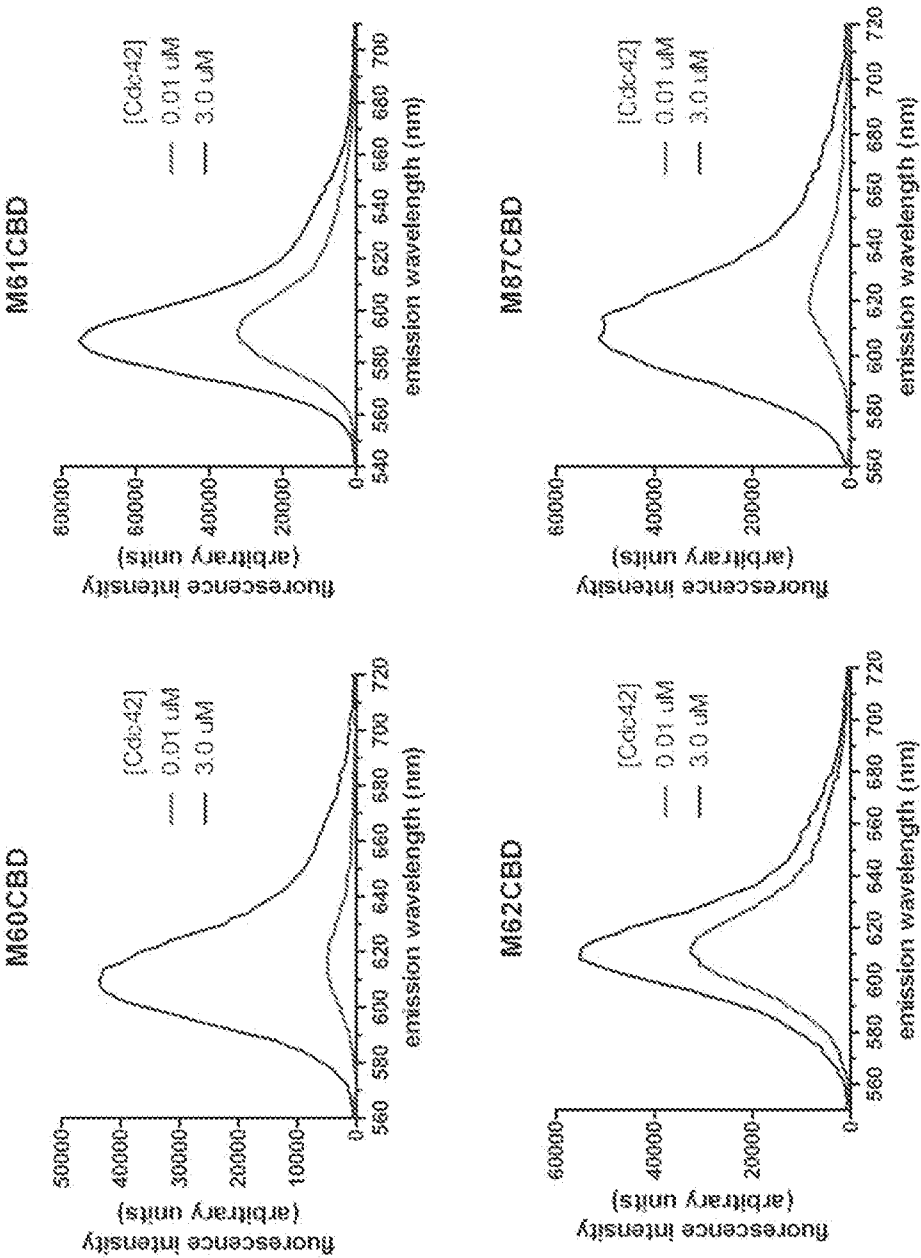


FIG. 6



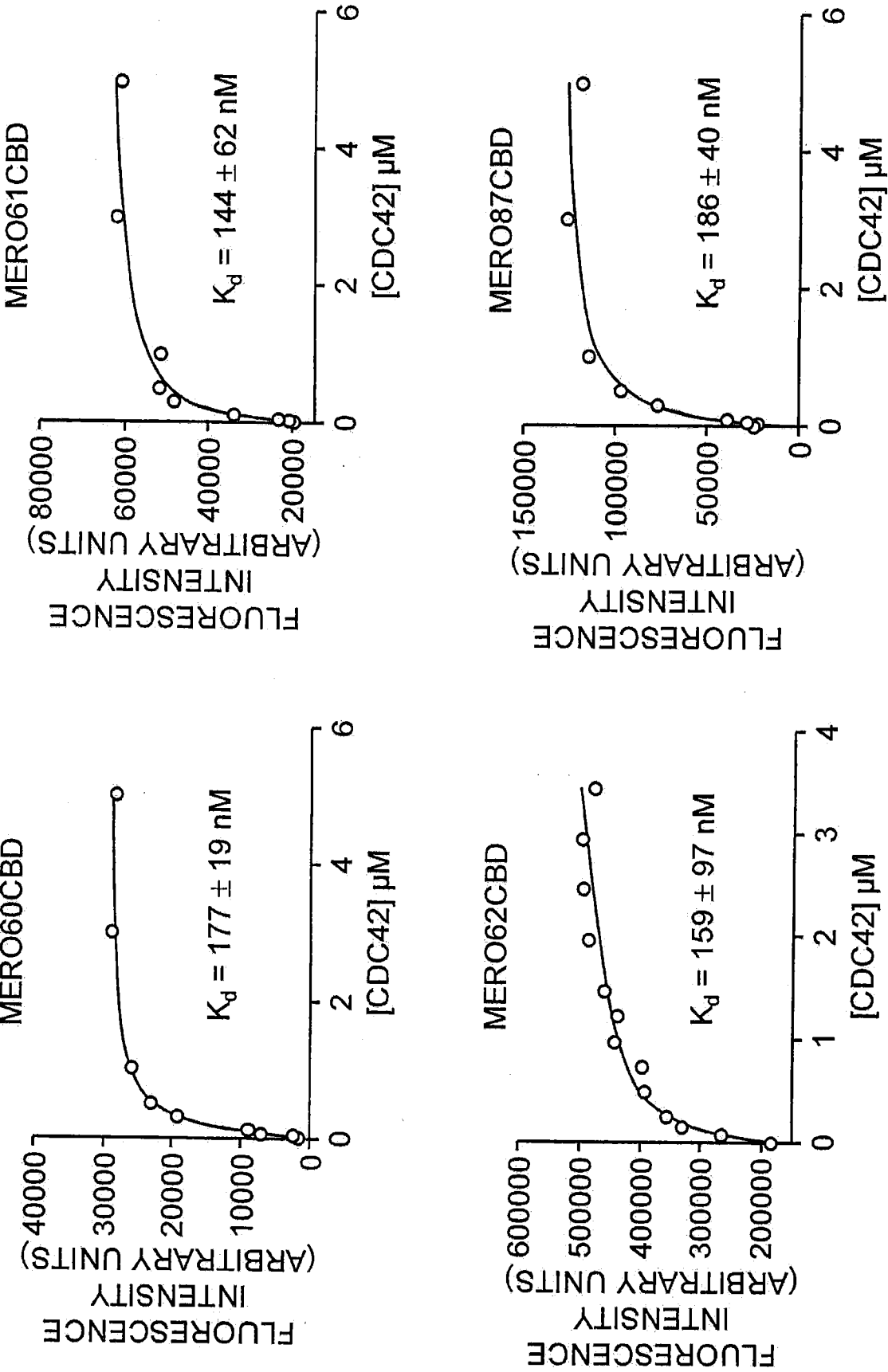


FIG. 7

FIGS. 8A-8C

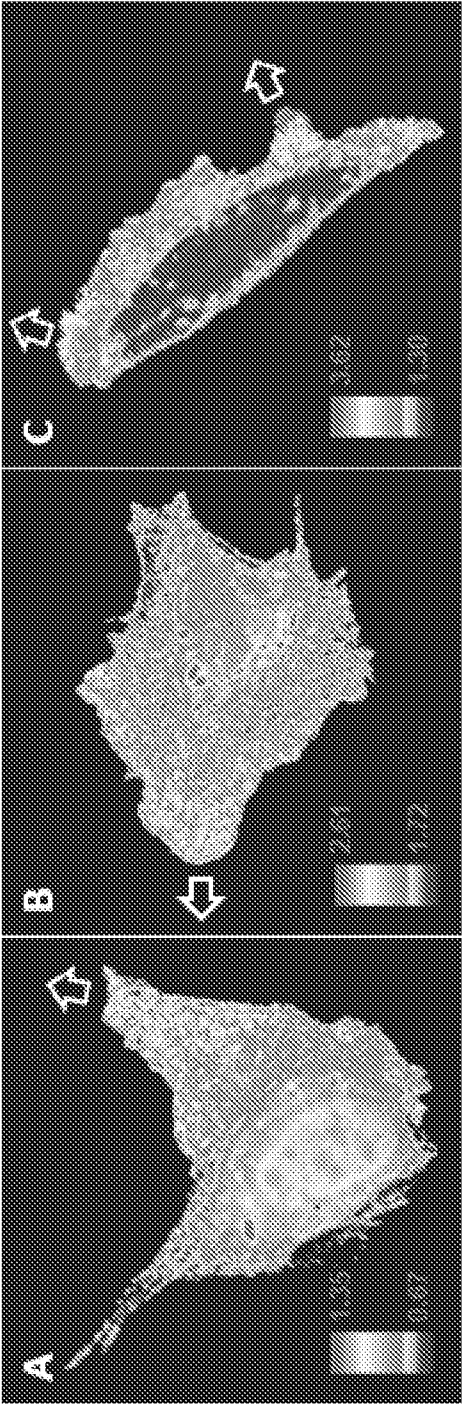


FIG. 9

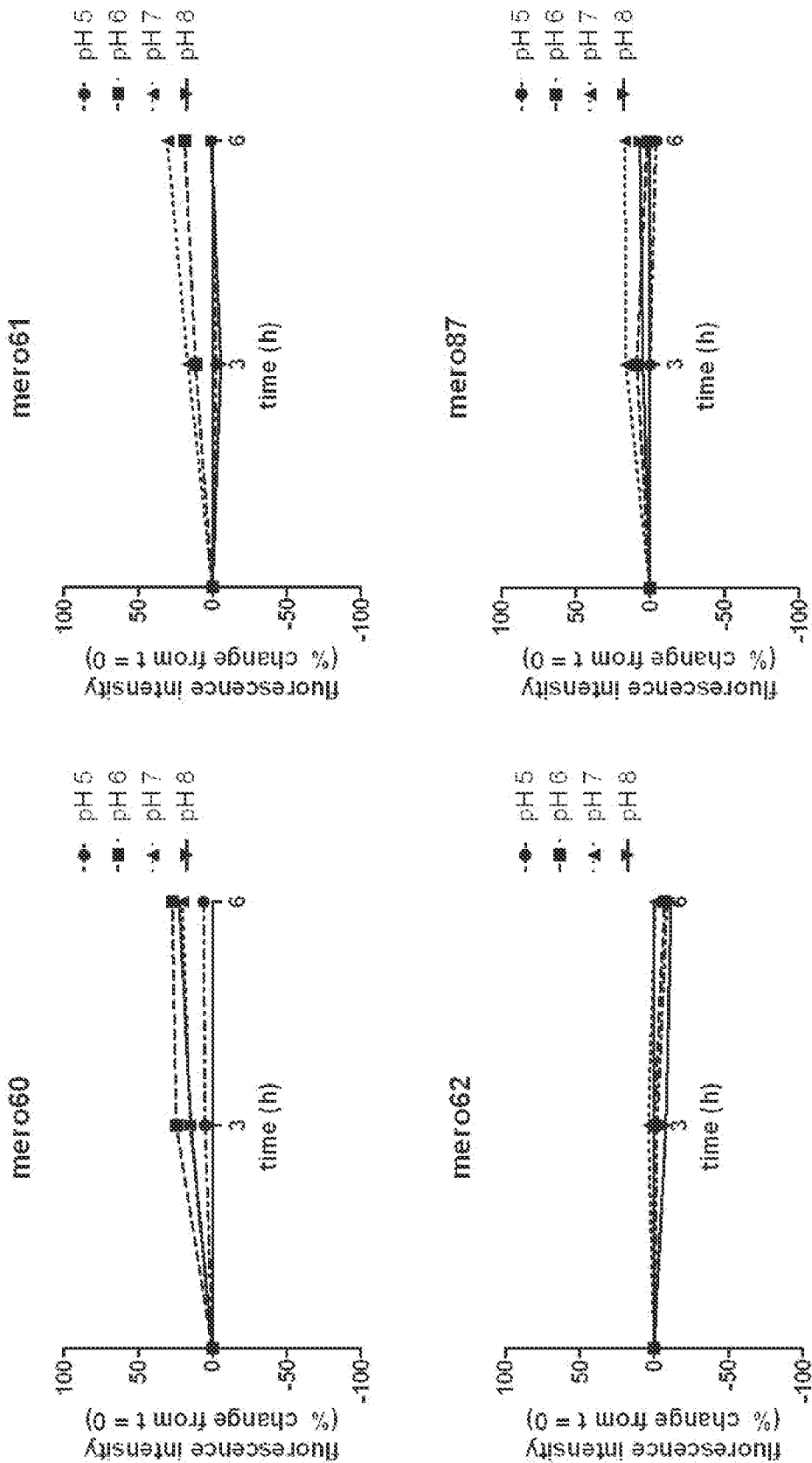
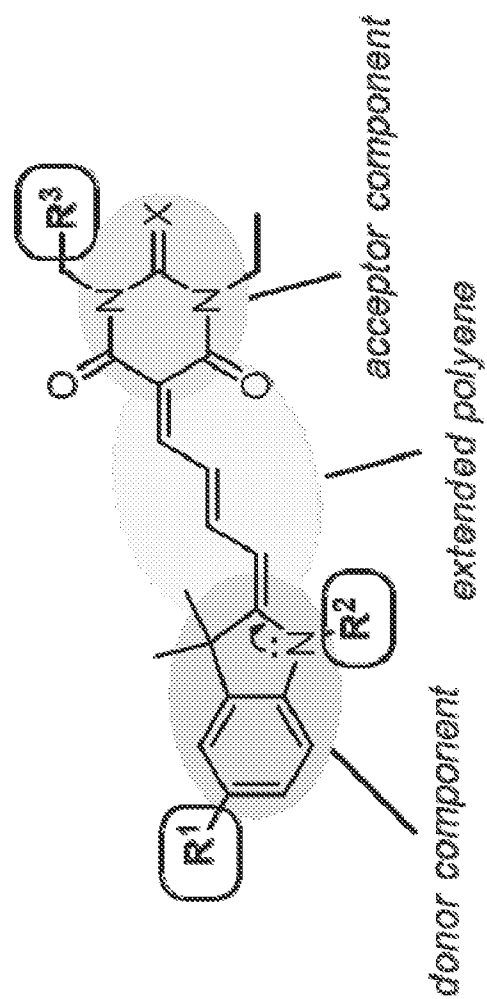


FIG. 10



- 1: $R^1 = H$, $R^2 = CH_3$, $R^3 = CH_3$
 2: $R^1 = AM^*$, $R^2 = CH_3$, $R^3 = CH_3$
 3: $R^1 = H$, $R^2 = CH_3$, $R^3 = (CH_2)AM$
 4: $R^1 = H$, $R^2 = (CH_2)_2AM$, $R^3 = (CH_2)AM$
 5: $R^1 = AM$, $R^2 = (CH_2)_2AM$, $R^3 = CH_3$
 *AM = $CO_2CH_2OCOCH_3$

FIG. 11

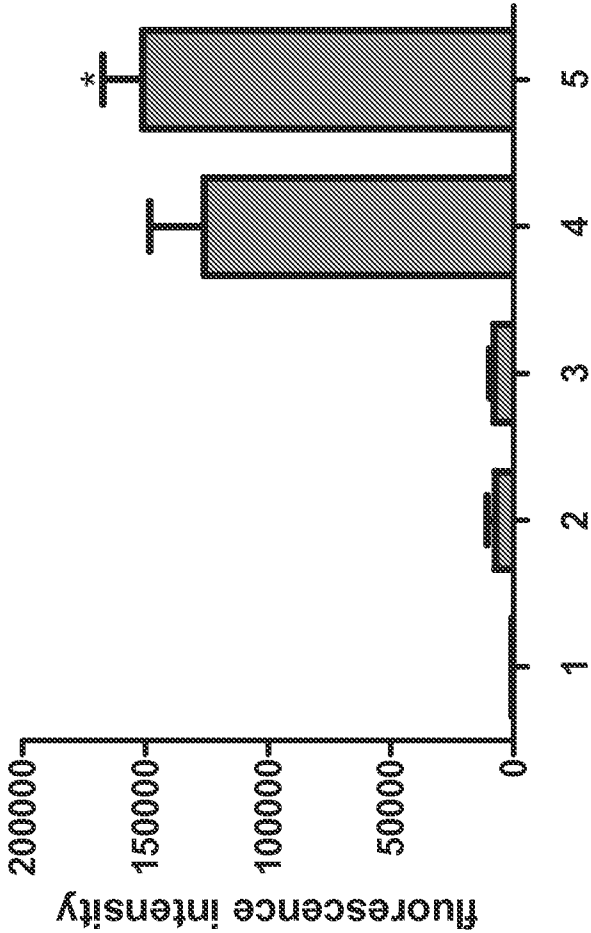


FIG. 12

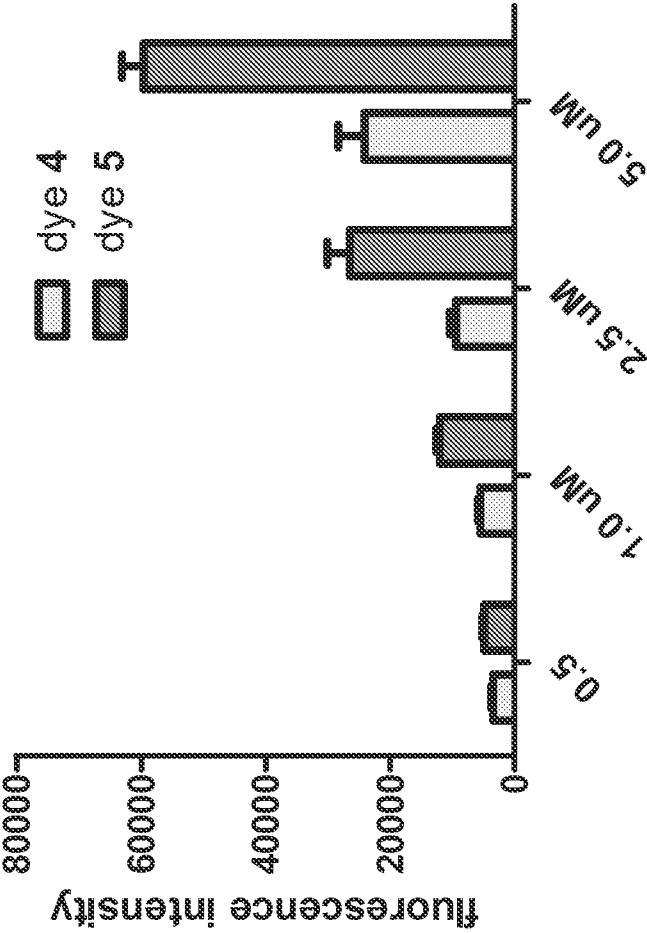


FIG. 13

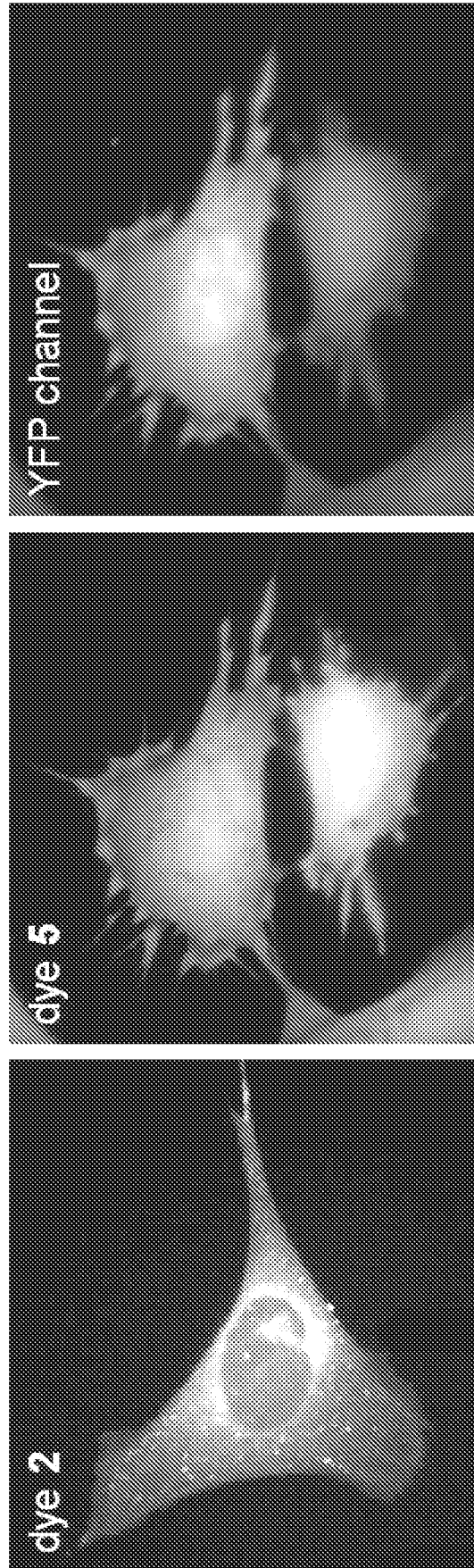


FIG. 14

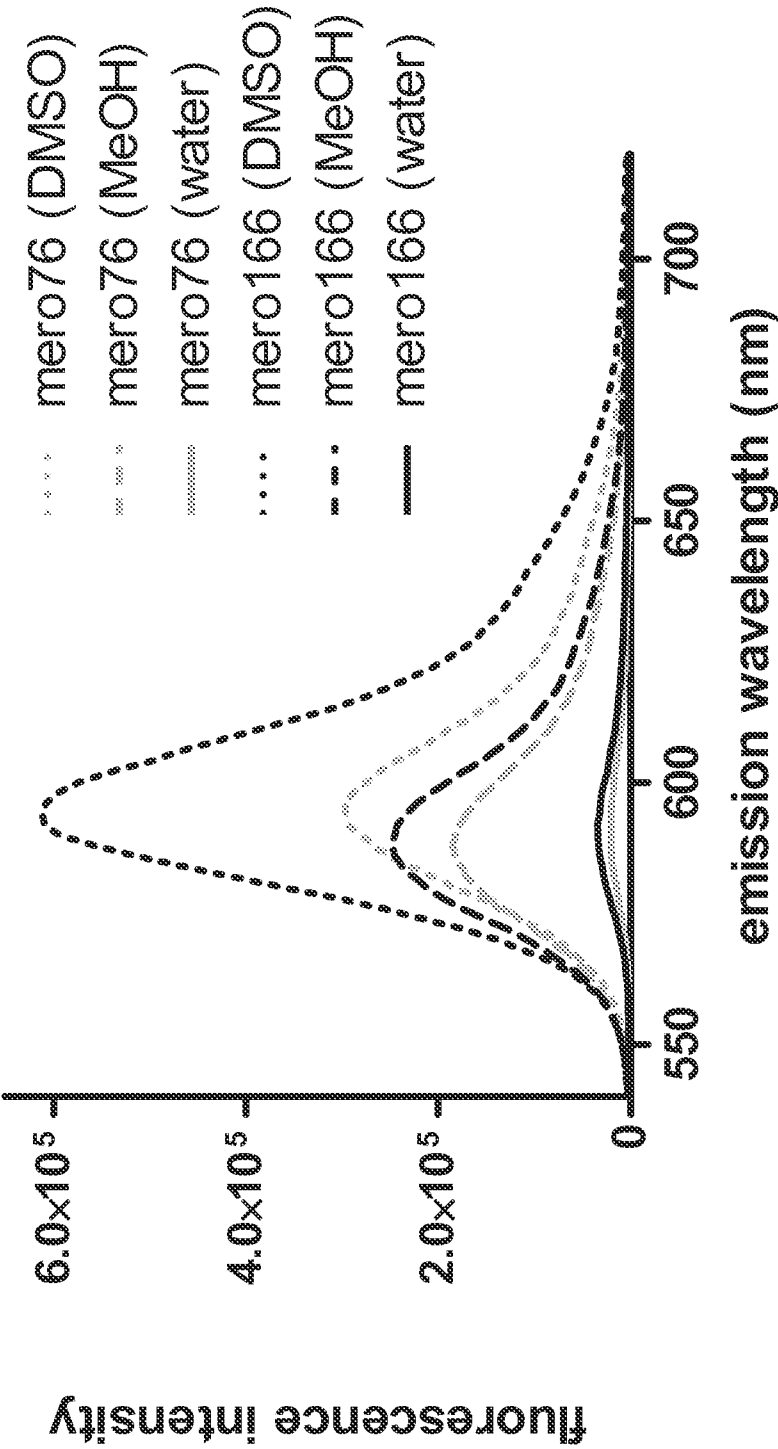
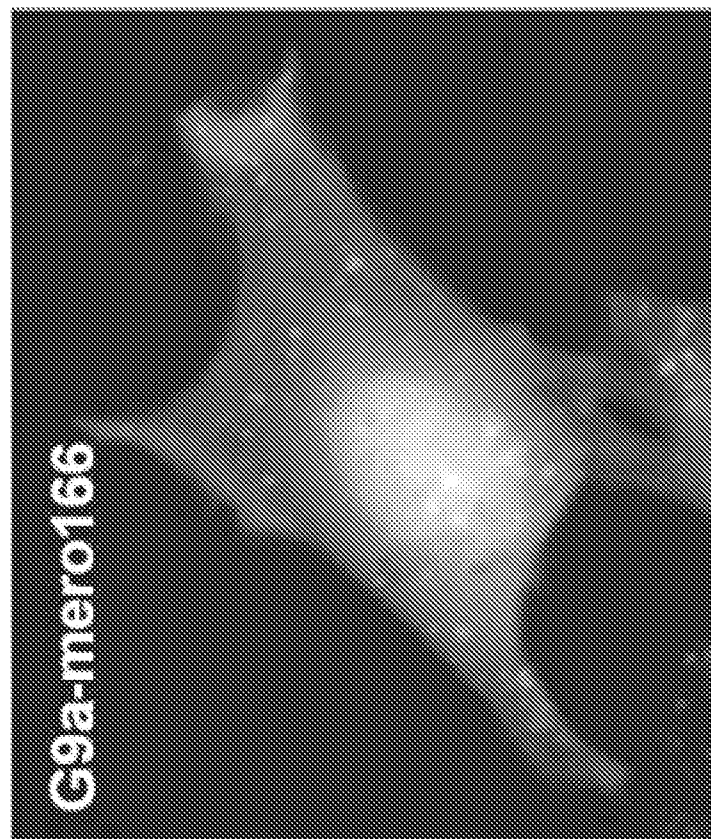
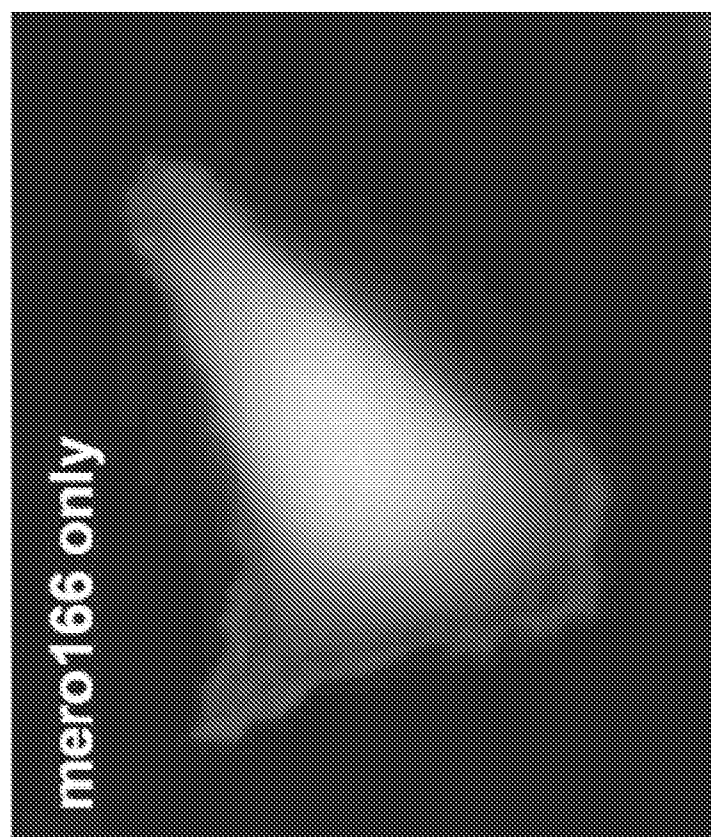


FIG. 15



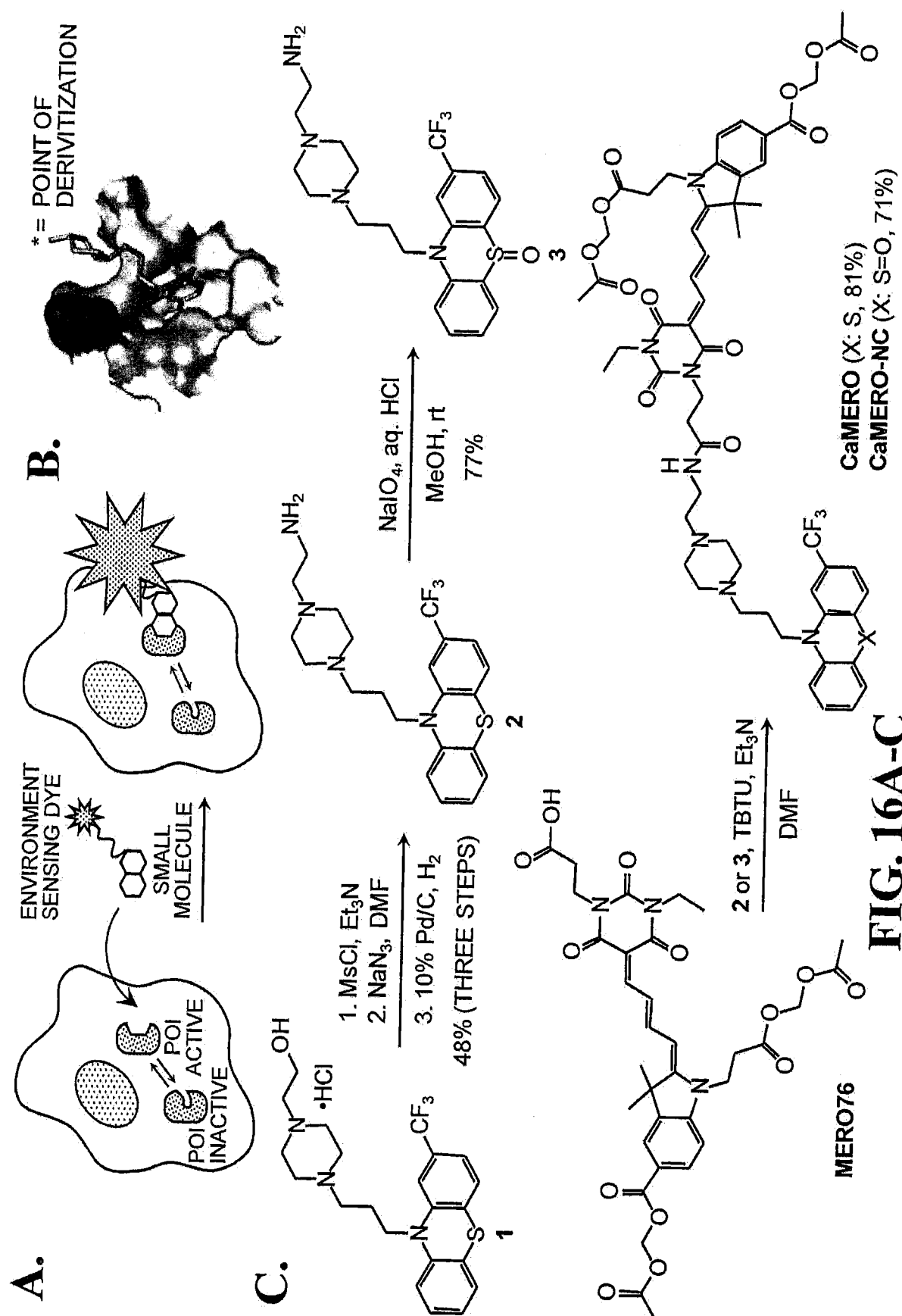
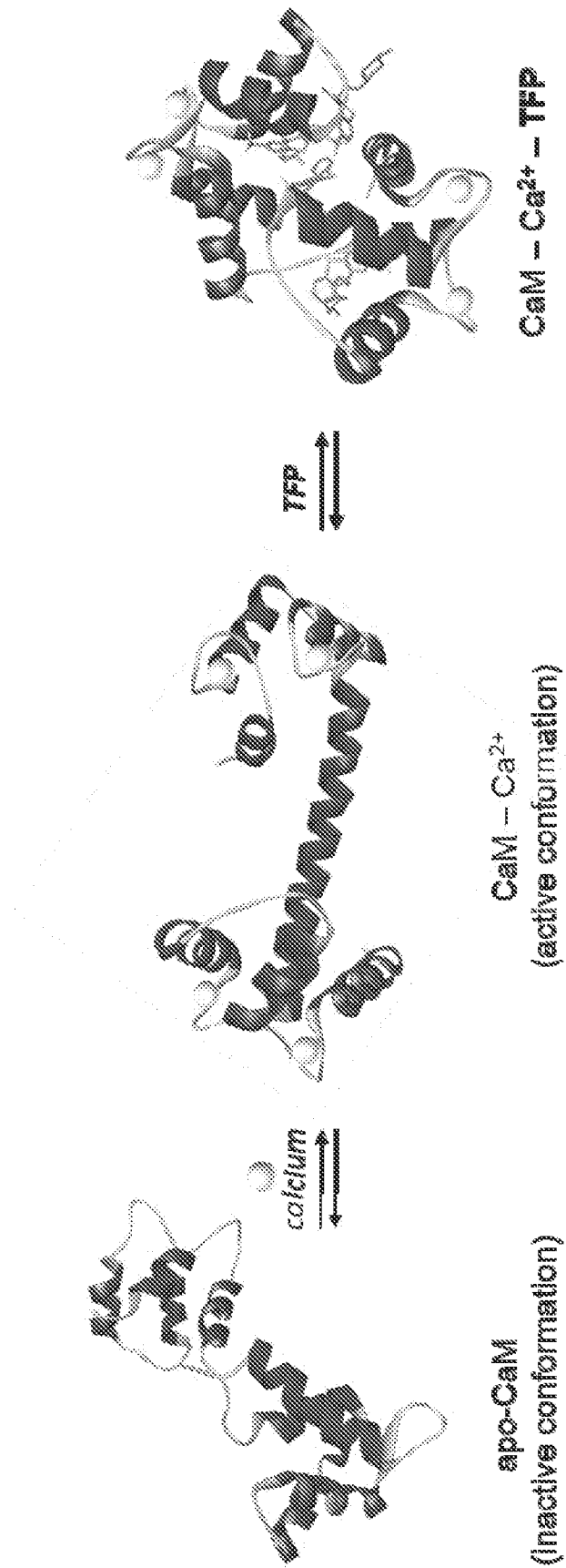
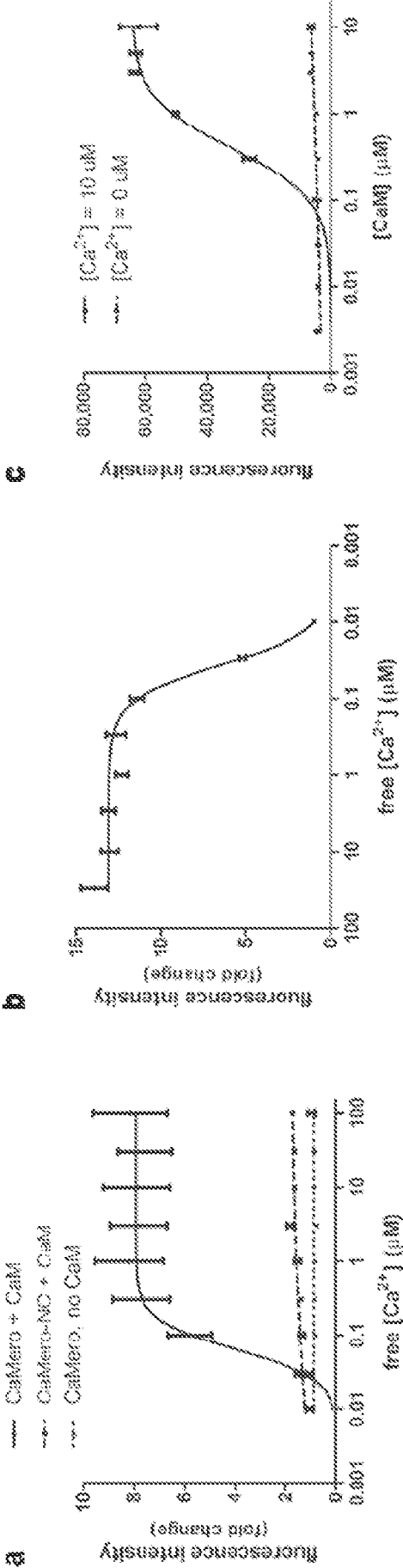
**FIG. 16A-C**

FIG. 17



FIGS. 18A-18C



FIGS. 19A-19C

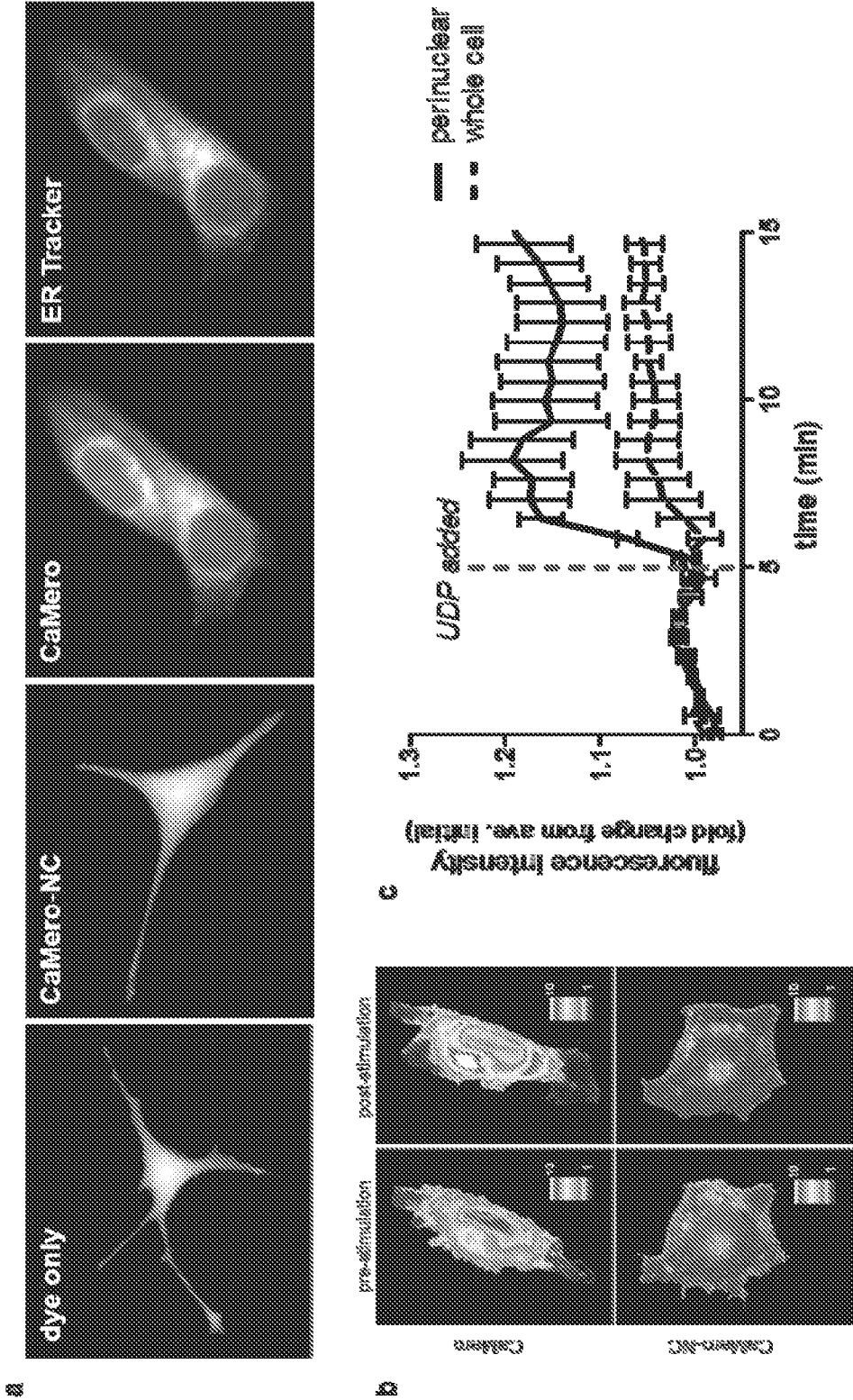
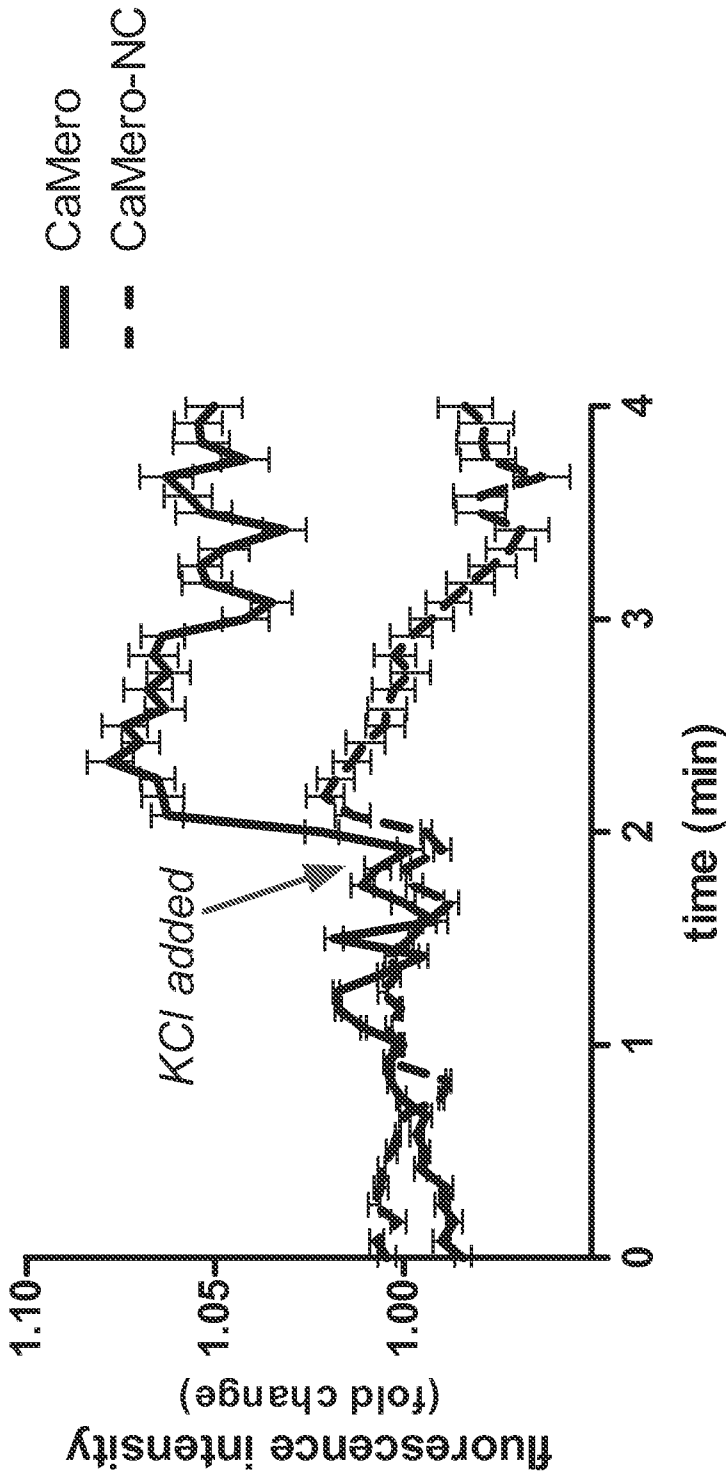


FIG. 20



FIGS. 21A-21C

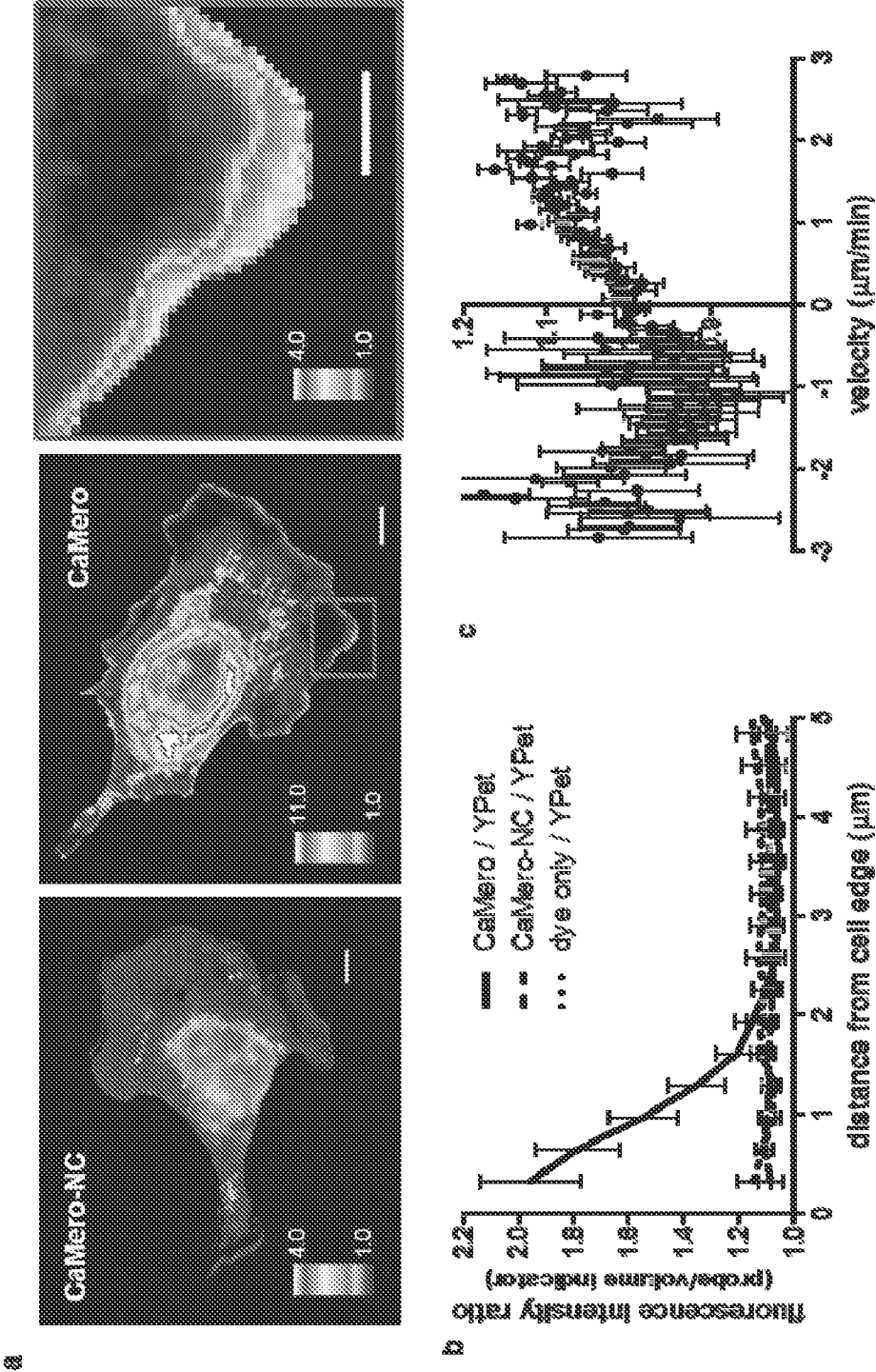
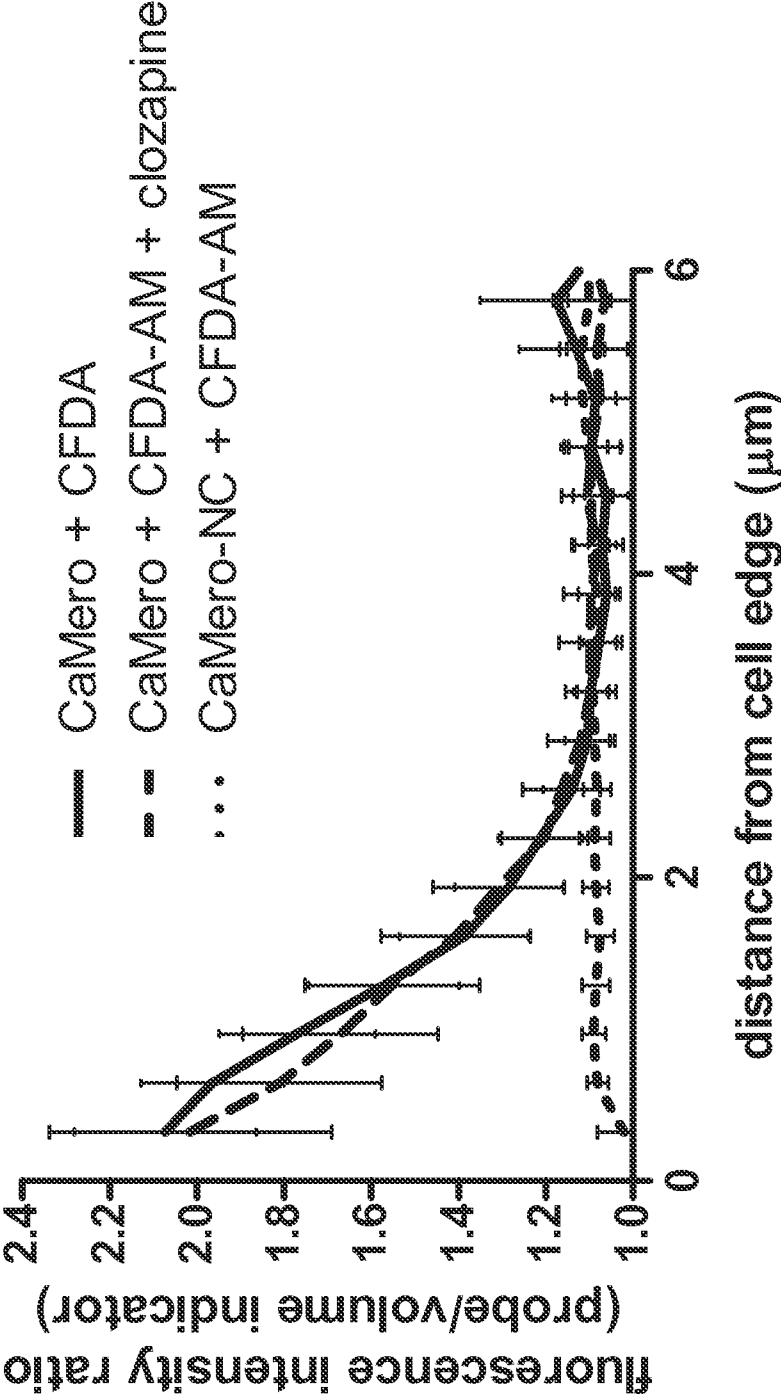


FIG. 22



FIGS. 23A-23C

