

- 1 -

CHALCONE STRUCTURE FLUORESCENCE DYE FOR EMBRYONIC STEM CELL PROBE

RELATED APPLICATION

This application claims the benefit of U.S. Provisional Application No. 61/377,570, filed on August 27, 2010. The entire teachings of the above application is incorporated herein by reference.

BACKGROUND OF THE INVENTION

Fluorescence detection is an indispensable technique for the analysis of a variety of biological phenomena. During the past few decades, a number of small molecule fluorescent chemosensors have been developed for use in biological analyses, which typically are elaborately designed to selectively detect a target substance or phenomenon. (See (a) Geddes, C. D.; Lakowicz, J. R., *Topics in Fluorescence Spectroscopy*, Vol. 9; Springer: New York, 2005. (b) Geddes, C. D.; Lakowicz, J. R. *Topics in Fluorescence Spectroscopy*, Vol. 10; Springer: New York, 2005, the teachings of both are incorporated herein by reference).

Embryonic stem cells (ESCs) are derived from the inner cell mass (ICM) of blastocyst-stage embryos and have the ability to self-renew indefinitely. ESCs are pluripotent because they have the capacity to give rise to differentiated progeny representative of all three embryonic germ layers, as well as the extraembryonic tissues that support development. In mammals, the property of pluripotentiality is restricted to the oocyte, the zygote, early embryonic cells, primordial germ cells, and the stem cells of tumors derived from pluripotential cells (embryonal carcinomas).

ESCs can be characterized by high level expression of Oct-3/4 (also termed Oct-3 or Oct-4), a member of the POU transcription factors, and Nanog. A critical amount of Oct-3/4 and Nanog expression is required to sustain stem-cell pluripotency. When ES cells are induced to differentiate, Oct-3/4 and Nanog are down-regulated, which has proven to be essential for a proper and divergent developmental program. Each stage of ESC differentiation can be monitored by

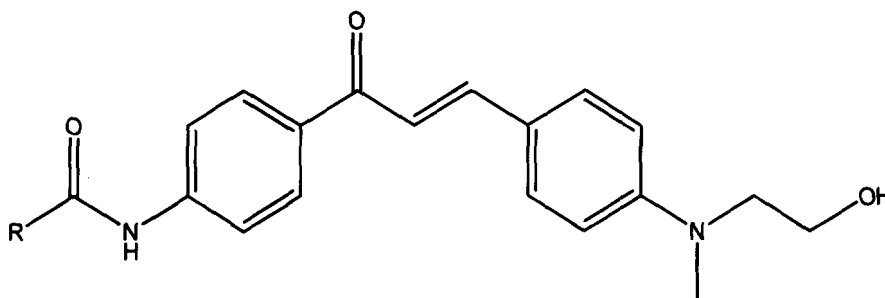
immunocytochemistry methods. The methods for probing mouse embryonic stem cell are generally diagnosed by cell surface antibody such as Nanog, Oct4 or GFP fluorescence protein. However, these protocols are limited mainly because of late stage of stem cell differentiation and selectivity.

- 5 Therefore, there is a need for a means by which the differentiation process can be monitored, which can reduce or eliminate one or more of the above-mentioned problems.

SUMMARY OF THE INVENTION

- 10 The invention is based on the synthesis of a new fluorescence dye of chalcone structure which shows fluorescence emission sensitivity depending on the polarity and pH of the solvent. The chalcone derivatives possess a central ketone group having electron accepting properties which can make these compounds suitable for use as fluorescence emission dyes.

- 15 One embodiment of the invention is a compound represented by the following structural formula:



or a salt thereof, wherein:

- 20 R is (C₁-C₁₆)alkyl, (C₂-C₁₆)alkenyl, (C₂-C₁₆)alkynyl, carbocyclyl or heterocyclyl, each of which is optionally substituted with one or more groups independently selected from the following: halo, NO₂, -NHCO(C₁-C₆)alkyl, -OCO(C₁-C₆)alkyl, -COO(C₁-C₆)alkyl, (C₁-C₁₀)alkyl, halo(C₁-C₁₀)alkyl, (C₁-C₁₀)alkoxy, -S(C₁-C₆)alkyl, (C₃-C₉)cycloalkyl, aryl, and phenoxy, wherein each phenoxy or aryl are further optionally substituted
- 25 with one or more groups selected from halo, NO₂, (C₁-C₁₀)alkyl, halo(C₁-C₁₀)alkyl, and (C₁-C₁₀)alkoxy.

The chalcone derivatives described herein possess fluorescence emission profiles useful for staining specific embryonic stem cell. The chalcone scaffold is compatible with a range of chemical functional groups, and can be bioconjugated to proteins as well as other macromolecules of interest. As described herein, chalcone derivatives of the invention selectively stain mouse embryonic stem population and MEF feeder cells. The chalcone derivatives can be used as new probes for ES and are useful for understanding early iPS generation mechanisms. Furthermore, the chalcone derivatives may be used to elucidate new biomarkers in ESC.

CDg4 (compound identified as CLA-7 in FIG. 1) stained the early stage of embryonic stem cell differentiation, which allowed for selective sorting and isolation from mixture of Feeder cells even up to day 5. Therefore, sorting of embryonic stem cell from feeder cell after inducer treatment can be accomplished using the chalcone derivatives of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a chart showing the chalcone compounds of the invention. The chart includes mass (calculated), mass (found), absorption (Abs), emission (Em), extinction coefficient (Ext Coeff) and quantum yield (5 μ M) data for the compounds.

FIG. 2 are plots demonstrating the selective staining of 4 dpi (day of post induction) iPS by CDg4 using FACS. The iPS cells were maintained in mESC culture medium which was changed every day until designated time point. 4dpi iPS was incubated with 2 μ M of CDg4 for 1hr at 37°C. The cells were harvested by trypsin treatment, washed with PBS and resuspended in PBS. The cells gated into SSClow CDg4 bright (R2) and SSChigh CDg4 dim (R1) regions were collected using a FACS machine (BD FACSAria™).

FIG. 3 is a representative graph depicting the fluorescence measurement of the chalcone derivatives of this invention.

FIG. 4 shows a flow cytometry dot plot image of CDg4 stained mESC and MEF.

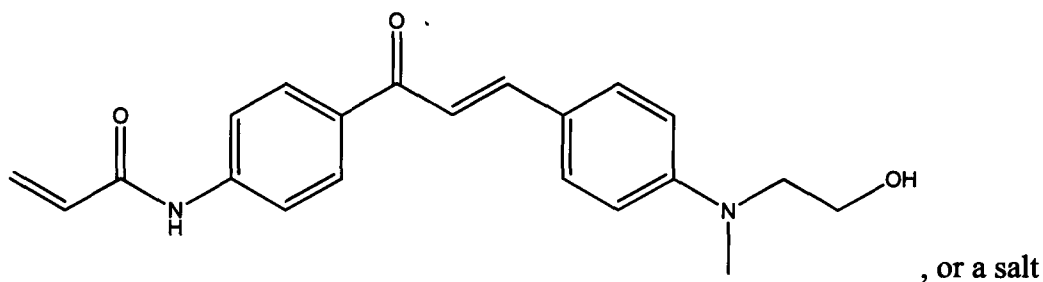
DETAILED DESCRIPTION OF THE INVENTION

A description of example embodiments of the invention follows. The teachings of all patents, published applications and references cited herein are incorporated by reference in their entirety.

5 Chalcones are 1,3-diphenyl-2-propene-1-one compounds, in which two aromatic rings are linked by a three carbon α,β -unsaturated carbonyl system. Chalcones are abundant in edible plants and are considered to be precursors of flavonoids and isoflavonoids. Compounds incorporating the chalcone backbone have been reported to possess various biological activities, such as antimicrobial, anti-inflammatory, analgesic, antiplatelet, antiulcerative, antimalarial, anticancer, antiviral, antileishmanial, antioxidant, antitubercular, antihyperglycemic, immunomodulatory activities. (See (a) Mokle S S, *et al.*, *Int. J. Chem. Sci.* 2004, 2(1), 96. (b). Hsieh H K, *et al.*, *J. Pharm. Pharmacol.*, 2000, 52, 163.(c) Viana G S, *et al.*, *J. Phytomedicine*, 2003, 10, 189. (d) Zhao L M, *et al.*, *Bioorg. Med. Chem. Lett*, 2005,15, 5027.(e) Mukarami S, *et al.*, *Biochem. Pharmacol*, 1991, 42, 1447.(f) Liu M, *et al.*, *J. Med. Chem*, 2001, 44, 4443.(g) Francesco E, *et al.*, *Phytochem*, 2007, 68, 939.(h) Onyilagna J C, *et al.*, *Can. J. Plant Pathol*, 1997,19, 133.(i) Nielsen S F, *et al.*, *Bioorg. Med. Chem. Lett*, 1995, 5, 449.(j) Miranda C L, *et al.*, *J. Agric. Food Chem*, 2000, 48, 3876.(k) Siva Kumar P M, *et al.*, *Chem. Pharm. Bull.*, 2007, 55(1), 44.(l) Satyanarayana M, *et al.*, *Bioorg. Med. Chem.*, 2004, 12, 883.(m) Barford L, *et al.*, *Int. Immunopharmacol*, 2002, 2, 545, the entire teaching of are incorporated herein by reference.)

25 The photophysical and photochemical properties of chalcones derivatives have been of considerable interest for a long time because of their donor-acceptor ability and because they are readily functionalized, but not many cases of their use as a fluorescence dye have been reported. (See (a) P.F. Wang, S.K. Wu, *J. Photochem. Photobiol. A* 86 (1995) 109.(b) C-G Niu, *et al.*, *Analytica Chimica Acta* 577 (2006) 264–270, the entire teachings of each are incorporated herein by reference.) We report herein novel fluorescent chalcone compounds synthesized via solid phase support synthetic method. One embodiment of the invention is a compound represented by the following chemical formula, designated CDgreen4 (or CDg4), $\lambda_{ex}/\lambda_{em} = 430/560$ nm):

30



thereof.

CDgreen4 selectively stains embryonic stem cells (ESC) and iPSC in a feeder cell-based culture system.

- 5 One embodiment of this invention is a method of isolating embryonic stem cells comprising the step of staining the cells with a chalcone derivative, or salt thereof, of the invention. In a particular embodiment, the chalcone derivative is CDg4 or a salt thereof.

- Each stage of embryonic stem cell (ESC) can be monitored by
10 immunocytochemistry methods. ESC can be characterized by high level expression of Oct-3/4 (also termed Oct-3 or Oct-4), a member of the POU transcription factors, and Nanog. A critical amount of Oct-3/4 and Nanog expression is required to sustain stem-cell pluripotency. When ESC are induced to differentiate, Oct-3/4 and Nanog are down-regulated, which has been proven to be essential for a proper and
15 divergent developmental program.

- After the ESCs have been stained with the chalcone compounds of the invention, the cells can be sorted. Cell surface antigens and alkaline phosphatase are biomarkers and may be used for stem cell confirmation. We have done for this confirmation experiment after sorting out by CDg4 with the biomarkers described
20 in the following paragraphs.

- The undifferentiated state of ES cells is often characterized by the expression of the cell surface antigens, SSEA-1, SSEA-3 and SSEA-4. SSEA-1 is expressed on the surface of preimplantation-stage murine embryos at the eight-cell stage and has been found on the surface of teratocarcinoma stem cells but not on
25 their differentiated derivatives. SSEA-3 and SSEA-4 are synthesized during oogenesis and are present in the membranes of oocytes, zygotes and early cleavage-stage embryos. It has been suggested that the biological role of these carbohydrate-associated molecules is to control cell surface interactions during development.

Undifferentiated primate ES cells, human ES cells and human embryonic carcinoma (EC) cells express SSEA-3 and SSEA-4, but not SSEA-1. Undifferentiated mouse ES cells, however, do express SSEA-1, but do not express SSEA-3 or SSEA-4.

Alkaline phosphatase is an enzyme in the blood, intestines, liver, and bone cells and exists as membrane-bound isoforms of glycoproteins sharing a common protein structure but differing in carbohydrate content. These enzymes are most active at alkaline pH - hence the name. Undifferentiated human EC, ES and embryonic germ (EG) cells have been shown to express a very high level of the liver/bone/kidney isozyme of alkaline phosphatase. Expression levels of alkaline phosphatase decrease following stem cell differentiation.

“Alkyl” means an optionally substituted saturated aliphatic branched or straight-chain monovalent hydrocarbon radical having from 1 to 16 carbon atoms. Alternatively, an alkyl group may have from 1 to 11 carbon atoms, 1 to 6 carbon atoms, 1 to 4 carbon atoms, 4 to 16 carbon atoms, 6 to 12 carbon atoms, 6 to 10 carbon atoms, or 10 to 16 carbon atoms. For example, “(C1-C6) alkyl” means a radical having from 1-6 carbon atoms in a linear or branched arrangement. “(C1-C6)alkyl” includes methyl, ethyl, propyl, butyl, pentyl and hexyl.

As used herein, the term “alkenyl” means a straight chain or branched non-cyclic hydrocarbon having from 2 to 16 carbon atoms and having at least one carbon-carbon double bond. Alternatively, an alkenyl group may have from 2 to 12 carbon atoms, 2 to 6 carbon atoms, 2 to 4 carbon atoms, 4 to 16 carbon atoms, 6 to 12 carbon atoms, 6 to 10 carbon atoms, or 10 to 16 carbon atoms. Alkenyl groups may be optionally substituted with one or more substituents at any substitutable hydrogen atom. For example, $-\text{CH}=\text{CH}_2$ may be substituted at a hydrogen atom with fluorine to yield $-\text{CF}=\text{CH}_2$, $-\text{CH}=\text{CFH}$, $-\text{CH}=\text{CF}_2$, $-\text{CF}=\text{CFH}$, $-\text{CH}=\text{CF}_2$, or $-\text{CF}=\text{CF}_2$.

As used herein, the term “alkynyl” means a straight chain or branched non-cyclic hydrocarbon having from 2 to 16 carbon atoms and having at least one carbon-carbon triple bond. Alternatively, an alkynyl group may have from 2 to 12 carbon atoms, 2 to 6 carbon atoms, 2 to 4 carbon atoms, 4 to 16 carbon atoms, 6 to 12 carbon atoms, 6 to 10 carbon atoms, or 10 to 16 carbon atoms. Alkynyl groups

may be optionally substituted with one or more substituents at any substitutable hydrogen atom.

“Halogen” and “halo” are interchangeably used herein and each refers to fluorine, chlorine, bromine, or iodine. As used herein,

- 5 halo-substituted-(C₁-C₁₆)alkyl means a (C₁-C₁₆)alkyl substituted with one or more halogen groups. Examples of halo-substituted-(C₁-C₄)alkyl include, but are not limited to -CF₃, -CH₂CF₃, -CH₂CF₂H, -CH₂CH₂F, -CH₂CH₂CF₃, -CCl₃, -CH₂CCl₃, -CH₂CCl₂H, -CH₂CH₂Cl and -CH₂CH₂CCl₃.

“Alkoxy” means an alkyl radical attached through an oxygen linking atom.

- 10 “(C₁-C₆)-alkoxy” includes methoxy, ethoxy, propoxy, butoxy, pentoxy and hexoxy.

“Haloalkyl” and “halocycloalkyl” include mono, poly, and perhaloalkyl groups where each halogen is independently selected from fluorine, chlorine, and bromine.

- 15 “Carbocyclyl” means a cyclic group with only ring carbon atoms.

“Carbocyclyl” includes 3-12-membered saturated or unsaturated aliphatic cyclic hydrocarbon rings or 6-12-membered aryl rings. Carbocyclyls are saturated or unsaturated aliphatic cyclic hydrocarbon rings or aromatic hydrocarbon rings having the specified number of carbon atoms. Carbocyclyls include cycloalkyl, cycloalkenyl, cycloalkynyl and aryl.

- 20 “Cycloalkyl” means a saturated aliphatic cyclic hydrocarbon ring. Thus, “C₃-C₉ cycloalkyl” means a hydrocarbon radical of a (3-9-membered) saturated aliphatic cyclic hydrocarbon ring. A C₃-C₉ cycloalkyl includes, but is not limited to cyclopropyl, cyclobutyl, cyclopentyl, cyclohexyl, cycloheptyl, cyclooctyl, and cyclononyl.

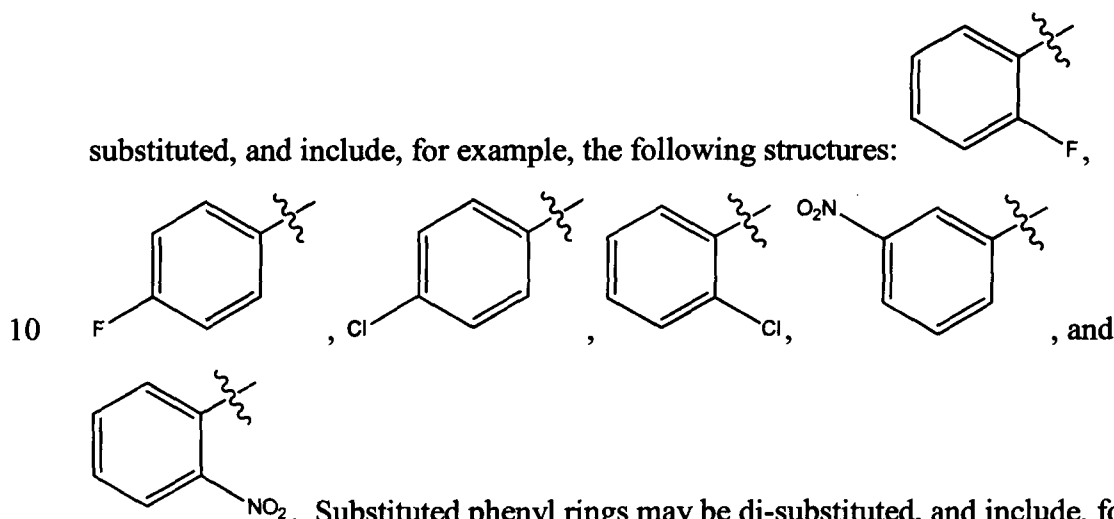
“Cycloalkene” means an aliphatic cyclic hydrocarbon ring having one or more double bonds in the ring. For example, cyclohexene is a cycloalkene.

“Cycloalkyne” means an aliphatic cyclic hydrocarbon ring having one or more triple bonds in the ring.

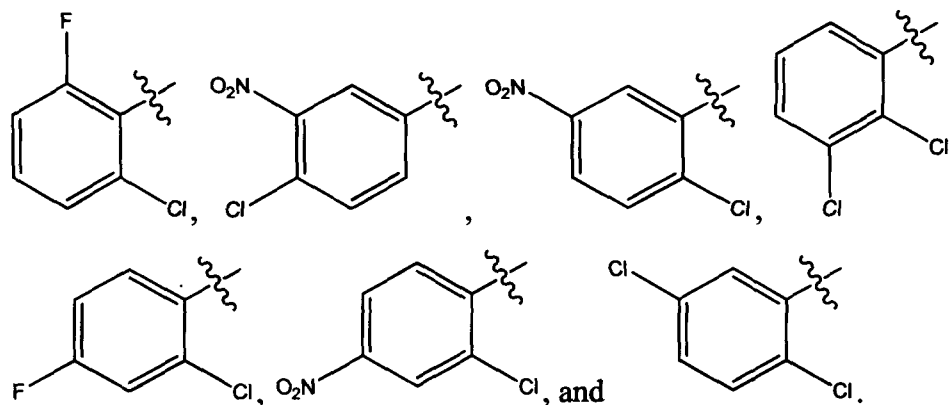
- 30 “Aryl” or “aromatic” means an aromatic monocyclic or polycyclic (*e.g.*, bicyclic or tricyclic) carbocyclic ring system. In one embodiment, “aryl” is a 6-12-membered monocyclic or bicyclic system. Aryl systems include, but are not limited

to, phenyl, naphthalenyl, fluorenyl, indenyl, azulenyl, and anthracenyl. Aryl rings are optionally substituted with one or more groups independently selected from the following: halo, NO₂, -NHCO(C₁-C₆)alkyl, -OCO(C₁-C₆)alkyl, -COO(C₁-C₆)alkyl, (C₁-C₁₀)alkyl, halo(C₁-C₁₀)alkyl, (C₁-C₁₀)alkoxy, -S(C₁-C₆)alkyl, (C₃-C₉)cycloalkyl, 5 aryl, and phenoxy. More specifically, the aryl ring is phenyl and is optionally substituted with one, two, three, four or five groups independently selected from halo, NO₂, -NHCO(C₁-C₆)alkyl, -OCO(C₁-C₆)alkyl, (C₁-C₁₀)alkyl, halo(C₁-C₄)alkyl, (C₁-C₁₀)alkoxy, and phenyl. Substituted phenyl rings may be mono

substituted, and include, for example, the following structures:



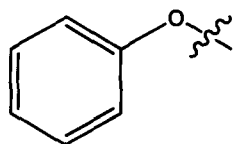
example, the following structures:



Similarly, the phenyl ring may optionally include three, four or five substituents that are the same or different.

The term "aryloxy," as used herein, means an "aryl-O-" group, wherein aryl is defined above. Examples of an aryloxy group include phenoxy or naphthoxy groups.

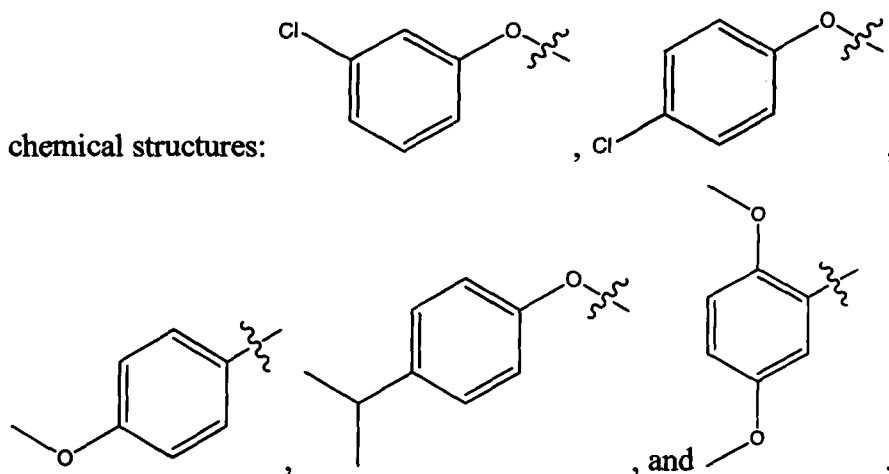
“Phenoxy” means a phenyl ring connected through an oxygen atom and is represented by the following structural formula:



, wherein  indicates connectivity to the remainder of the

5 molecule. Phenoxy moieties may be optionally substituted with one, two, three, four or five groups independently selected from halo, NO₂, -NHCO(C₁-C₆)alkyl, -OCO(C₁-C₆)alkyl, (C₁-C₁₀)alkyl, halo(C₁-C₄)alkyl, (C₁-C₁₀)alkoxy, and phenyl.

Examples of substituted phenoxy include, for example the following representative



“Hetero” refers to the replacement of at least one carbon atom member in a ring system with at least one heteroatom selected from N, S, and O. “Hetero” also
 15 refers to the replacement of at least one carbon atom member in an acyclic system. A hetero ring system or a hetero acyclic system may have 1, 2, 3 or 4 carbon atom members replaced by a heteroatom.

“Heterocyclyl” means a cyclic 4-12-membered saturated or unsaturated aliphatic or aromatic ring containing 1, 2, 3, 4 or 5 heteroatoms independently
 20 selected from N, O or S. When one heteroatom is S, it can be optionally mono- or di-oxygenated (*i.e.*, -S(O)- or -S(O)₂-). The heterocyclyl can be monocyclic, bicyclic, or polycyclic.

“Saturated heterocyclyl” means an aliphatic heterocyclyl group without any degree of unsaturation (*i.e.*, no double bond or triple bond).

Examples of monocyclic saturated heterocyclyls include, but are not limited to, azetidine, pyrrolidine, piperidine, piperazine, azepane, hexahydropyrimidine,
5 tetrahydrofuran, tetrahydropyran, morpholine, thiomorpholine, thiomorpholine 1,1-dioxide, tetrahydro-2H-1,2-thiazine, tetrahydro-2H-1,2-thiazine 1,1-dioxide, isothiazolidine, isothiazolidine 1,1-dioxide.

A bicyclic heterocyclyl has two rings which have two adjacent ring atoms in common. The first ring is a monocyclic heterocyclyl and the second ring is a
10 monocyclic carbocycle (such as a cycloalkyl or phenyl) or a monocyclic heterocyclyl. For example, the second ring is a (C₃-C₆)cycloalkyl, such as cyclopropyl, cyclobutyl, cyclopentyl and cyclohexyl. Alternatively, the second ring is phenyl. Examples of fused bicyclic heterocyclyls include, but are not limited to, octahydrocyclopenta[c]pyrrolyl, indoline, isoindoline,
15 2,3-dihydro-1H-benzo[d]imidazole, 2,3-dihydrobenzo[d]oxazole, 2,3-dihydrobenzo[d]thiazole, octahydrobenzo[d]oxazole, octahydro-1H-benzo[d]imidazole, octahydrobenzo[d]thiazole, octahydrocyclopenta[c]pyrrole, 3-azabicyclo[3.1.0]hexane, and 3-azabicyclo[3.2.0]heptane.

20 Polycyclic heterocyclyls have more than two rings, one of which is a heterocyclyl (*e.g.*, three rings resulting in a tricyclic ring system) and adjacent rings having at least one ring atom in common.

“Heteroaryl” or “heteroaromatic ring” means a 5-12-membered monovalent heteroaromatic monocyclic or bicyclic ring radical. A heteroaryl contains 1, 2, 3 or 4
25 heteroatoms independently selected from N, O, and S. Heteroaryls include, but are not limited to furan, oxazole, thiophene, 1,2,3-triazole, 1,2,4-triazine, 1,2,4-triazole, 1,2,5-thiadiazole 1,1-dioxide, 1,2,5-thiadiazole 1-oxide, 1,2,5-thiadiazole, 1,3,4-oxadiazole, 1,3,4-thiadiazole, 1,3,5-triazine, imidazole, isothiazole, isoxazole, pyrazole, pyridazine, pyridine, pyridine-*N*-oxide, pyrazine, pyrimidine, pyrrole,
30 tetrazole, and thiazole. Bicyclic heteroaryl rings include, but are not limited to indolizine, indole, isoindole, indazole, benzimidazole, benzthiazole, benzothiophene,

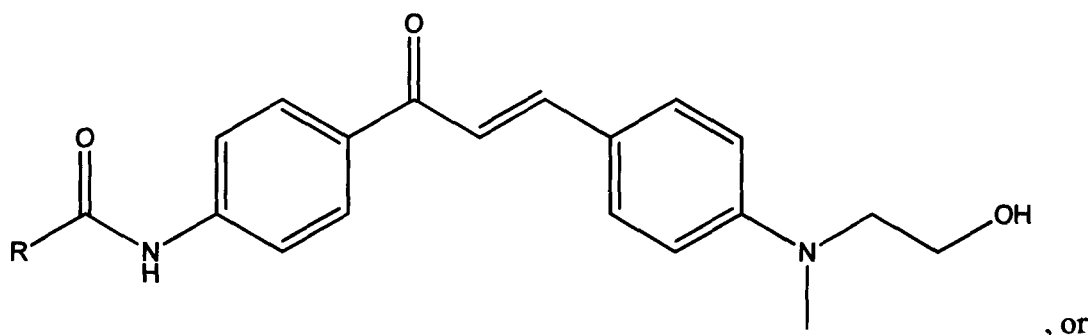
purine, quinoline, isoquinoline, cinnoline, phthalazine, quinazoline, quinoxaline, 1,8-naphthyridine, and pteridine.

EXEMPLIFICATION

5 Design of Chalcone Derivatives

Most fluorescent chemosensors employ an increase or decrease in their emission intensity as a sensing signal in response to the surrounding medium or through specific molecular recognition events. Due to their simplicity and high sensitivity, fluorescent sensors have been widely utilized as tools for chemical, biological, and medical applications. Sensors are prepared via two different strategies. For "Analyte Directed Sensors," the fluorescence dye molecules are combined with designed receptors for a specific analyte. Alternatively, "Diversity Directed Sensors" are generated via a combinatorial dye library and can be developed for a diverse set of dyes.

15 The invention is based on the synthesis of a new fluorescence dye of chalcone structure:

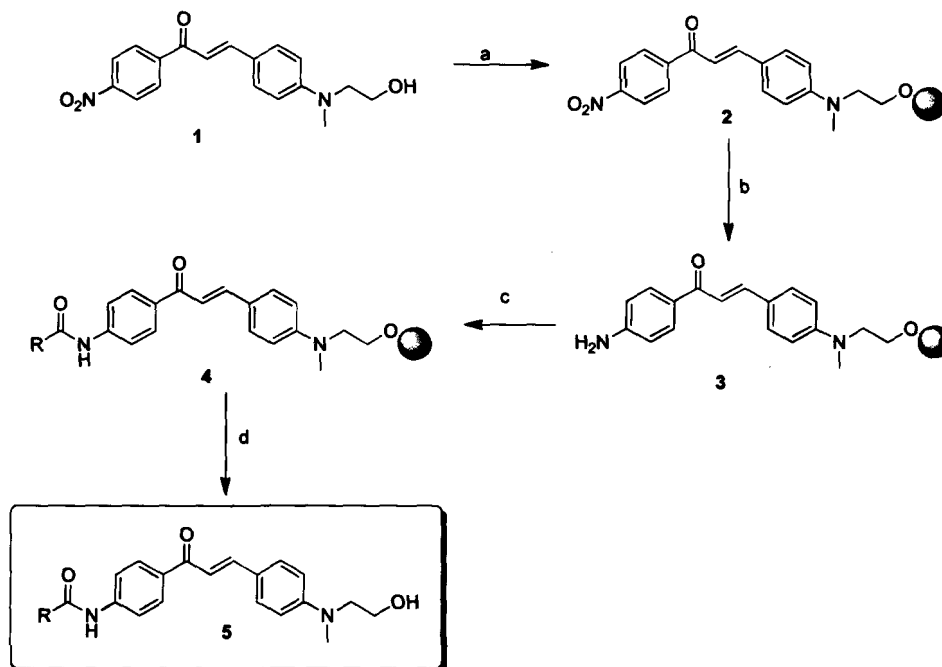


a salt thereof, where R is as defined above, which shows fluorescence emission sensitivity depending upon the polarity of solvent and pH. The central ketone group has electron accepting properties which can make it suitable for use as a fluorescence emission dye.

General Methods. All reactions were performed in oven-dried glassware under a positive pressure of nitrogen. Unless otherwise noted, starting materials and solvents were purchased from Aldrich and Acros Organics and used without further purification. Analytical TLC was carried out on Merck 60 F254 silica gel plate (0.25 mm layer thickness) and visualization was done with UV light. Column chromatography was performed on Merck 60 silica gel (230-400 mesh). NMR

spectra were recorded on a Bruker Avance 300 NMR spectrometer. Chemical shifts are reported as δ in units of parts per million (ppm) and coupling constants are reported as a J value in Hertz (Hz). Mass of all the compounds was determined by LC-MS of Agilent Technologies with an electrospray ionization source. All

5 fluorescence assays were performed with a Gemini XS fluorescence plate reader.



(a) ClTr-Cl resin, Py, DCM, 5hrs, RT, (b) 2M SnCl₂·2H₂O in DMF, RT, 16hrs, (c) RCOCl, Py, DCM, 30min, (d) 2% TFA in DCM

10

Chemical synthesis

Synthesis of compound 1: To the solution of 4'-nitroacetophenone (100mg, 0.6 mmol) and *N,N'*-2-hydroxyethyl, methyl nitrobenzaldehyde (100 mg, 0.55 mmol) in ethanol (EtOH), pyrrolidine (0.1 mL) was added and heated with

15 commercial Microwave Reactor (150MW) for 3 minutes. The resultant solution was cooled and kept at the room temperature for 2 hours, after which a dark red solid precipitated. The solid was filtered and washed with ethyl acetate and hexanes (1:1) solution and dried as a dark red solid (100 mg, 60%). Obtained solid was used for subsequent reaction without any purification.

20

Synthesis of compound 2: 2-Chlorotrityl resin (200 mg, 1mmol/g) was pre-swelled in DCM. To the resin solution, compound 1 (100 mg) and 3 equivalents of pyridine were added and the reaction mixture was stirred overnight. The reaction

mixture was filtered and washed with DMF, MeOH, DCM after 1 hour MeOH capping procedure for extra resin chloride deactivation. Washed resin was dried with high vacuum dessicator and used without further purification.

Synthesis of compound 5: Resin loaded with compound 2 was treated with SnCl₂ hydrate (20 equivalents) in DMF solution. The reaction mixture was shaken overnight and washed with DMF, MeOH, DCM. After high vacuum drying, the resin (60 mg) was aliquoted into each 5 mL syringe. The reaction syringes were washed with DCM and pyridine (3 equivalents) in DCM (1mL) was loaded in syringe and acid chlorides (RCOCl) in 2 mL loading was followed to yield compound 3, which was used without further purification. The reaction syringes were shaken for 30 minutes, and then were washed out with DCM, MeOH, DCM.

Cleavage protocol: 1% TFA in DCM solution was loaded for cleavage. For 30 minutes, incubated with cleavage solution, aliquot was squeezed out and collected in 20 mL vial. 2.5% ammonia water in ACN was added for TFA neutralization and then filtered through a silica end-filled tip. Dark brown solid was obtained after removing solvent and purity was analyzed by LC/MS.

Characteristics of representative compounds

CDg4 was synthesized according to the general synthetic procedure. ¹H-NMR (CDCl₃) δ 10.61 (s, 1H, NH), 8.11 (d, J = 8.1, 2H), 7.85 (d, J = 8.1, 2H), 7.66 (d, J = 8.7, 2H), 7.63 (bs, 2H), 6.74 (d, J = 8.4, 2H), 6.52(m, 1H), 6.30 (d, J = 16.5, 1H), 5.80 (d, J = 9.9, 1H), 5.74 (s, 1H), 4.76 (bs, 1H), 3.58 (bs, 2H), 3.48 (bd, 2H), 3.01 (s, 3H) ; ¹³C-NMR (CDCl₃) δ 187.50, 163.94, 151.45, 144.95, 143.28, 133.70, 131.11, 129.80, 127.97, 122.10, 119.09, 116.17, 111.85, 104.90, 58.51, 54.27, 48.97. ESI-MS m/z (M+H) calc'd: 351.1, found 351.1, FAB-MS MS m/z (M+H) calc'd: 351.1, found 351.1.

Fluorescence property measurement

10 μM solutions in DMSO were prepared and the absorption and emission of library compounds was measured. FIG. 3 is a representative absorbance and fluorescence spectrum. The spectral information is as follows:

Spectral Information	
ABS (nm)	429
Excitation (nm)	430nm
C(abs)	50 μ m
Em(nm)	563
Max RFU	25709
C(em)	10 μ m
Quantum Yield	0.202

Measurements for the other chalcone compounds are presented in FIG. 1.

Identification of Chalcone Derivatives

To evaluate the application of chalcone derivatives as ESC-selective probes,
5 320 chalcone compounds were screened in mouse ESC (mESC) and mouse embryonic fibroblast (MEF) feeder cells prepared on 384-well cell culture microplates. From the image-based primary screening, 8 compounds that stained mESC consistently with stronger intensity than MEF were selected. For the secondary screening, mESC and MEF were incubated separately with each of the hit
10 compounds and subsequently analyzed using flow cytometry. The flow cytometry data of side scatter (SSC) and fluorescence intensity in dot plot overlay of mESC and MEF indicated CDg4 as the most selective for mESCs among the 8 hit compounds.

Enrichment of mESC using FACS

15 To confirm the secondary screening data, a mixture of mESC and MEF was incubated with CD4. The cell mixture was sorted using Fluorescence-Activated Cell Sorting (FACS). Fluorescence microscopy demonstrated that live mESC cultured on MEF feeder are much more brightly stained by CDg4 than MEF. The data demonstrated that CD4 can be used for mESC enrichment from a mixed-cell
20 population. FIG. 4 shows a flow cytometry dot plot image of CDg4 stained mESC and MEF.

Co-localization of CDg4 and CellMask- plasma membrane. mESC were incubated with 2 μ M CDg4, 5ug/mL CELLMASK® Deep Red Plasma membrane stain and 4 μ g/ml Hoechst for 30 minutes at 37° C and washed with PBS two times. The images
25 were taken using a Nikon ECLIPSE Ti fluorescence microscope equipped with a 60x objective lens.

Use of CDg4 to isolate early iPS cells to study DNA profile with DNA microarray.

CDg4 was applied to induced pluripotent stem cell (iPSC) which was generated from MEF of transgenic mice that express GFP under the control of Oct4 (also known as Pou5f1) promoter. CDg4 selectively stained the iPSC colony. FACS was used to isolate early iPS cell population stained by CDg4. (FIG. 2). Isolation of the early iPS cell population will facilitate gene profiling using DNA microarray of this cell population. While gene expression profiling for mixture populations or using cell lines are reported, with lack of direct marker or probes for early iPS, it was not possible to study with purified early iPS.

DNA microarray analysis for their gene expression pattern will be performed to understand the mechanism and events for the difference of early and late stage of cell reprogramming. CDg4-stained cells will be sorted out using FACS and their RNA will be isolated. Biotinylated cRNA using Illumina RNA Amplification Kit (Ambion Inc.) will be synthesized and hybridized to BeadChips. Arrays will be scanned and analyzed using iScan Scanner System and GenePlex software. Our CDg4 will provide a unique chance to isolate and enrich the specific populations.

General Experimental Methods:

Cell Culture. mESC was cultured in a culture dish coated with 0.1% gelatin using a high-glucose Dulbecco's Modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum, 2 mM L-glutamine, 100 U/ml penicillin, 100 µg/ml streptomycin, 0.1 mM non-essential amino acids, 0.1% β-mercaptoethanol and 100 U/ml leukemia inhibitory factor (LIF, Chemicon). MEF was maintained in the same media as used for mESC without LIF and treated with mitomycin C (10 µg/ml) before used as feeder cell.

Image-based cell screening. MEFs were plated on 0.1% gelatin-coated 384-well cell culture microplates (Greiner bio-one, Germany). On the next day, mESCs were plated both on MEF feeder and feeder-free wells for overnight culture and then incubated with 1µM of chalcone fluorescent compounds. After 0.5 hours, 24 hours and 48 hours, FITC fluorescence and bright field images were taken using ImageXpressMICRO imaging system (Molecular Devices).

Flow cytometry and FACS. MEFs and mESCs were cultured in 60 mm culture dish for 1 hour in the presence of 2 μ M of hit compounds, including CDg4. The cells were harvested by trypsin treatment, washed with PBS and resuspended in PBS. The fluorescence intensity of the cells was measured on Flow Cytometry (BDTM LSR II) with FITC (excitation at 530 nm, emission at 560 nm). The dot plot images for each cell type were overlaid using FlowJo7 (Three Star Inc.) For FACS MEF and mESC were cultured in a same dish and the sample was prepared as described above for flow cytometry. The cells gated into SSClow CDgreen4 bright and SSChigh CDgreen4 dim regions were collected using a FACS machine (BD FACS AriaTM).

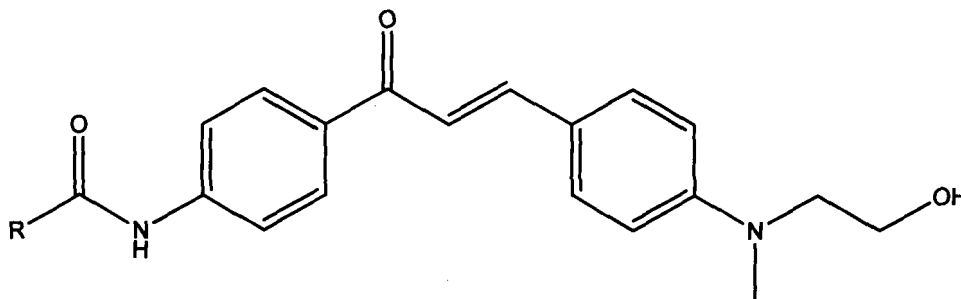
iPSC generation. The retroviral packaging cell, PlatE cell, was maintained in DMEM containing 10% FBS, 100 U/ml penicillin, 100 μ g/ml streptomycin, 10 μ g/ml blasticidin S and 1 μ g/ml puromycin. For the generation of retroviruses, 40 μ g of each retroviral vectors expressing each mouse Oct4, Sox2, Klf4, and c-Myc was transfected together using Lipofectamine 2000TM (Invitrogen) into PlatE cells plated at a density of 8×10^6 cells/10cm cell culture dish and cultured overnight in DMEM containing only 10% FBS. Medium was changed 5 hours after transfection to normal maintenance medium and incubated further in a humidified atmosphere of 5% CO₂ and 95% air at 37 °C. Culture supernatant was harvested 48 hours after transfection, filtered through a 0.45- μ m syringe filter, concentrated by using Amicon Ultra-15 Centrifugal filter units (Millipore) and stored at -80 °C until use. For reprogramming, MEFs prepared from E13.5 B6;CBA-Tg(Pou5f1-EGFP)2Mnn/J mouse (Jackson Laboratory) embryo were plated at a density of 3×10^4 cells/well in a 24-well plate and infected with retroviruses mixture 8 hours later. Polybrene (Sigma) was added at a concentration of 10 μ g/ml to increase virus adsorption. Infected MEFs were harvested 48 hours later by trypsinization and plated onto mytomycin C treated MEF feeder cells. The cells were maintained in mESC culture medium which was changed every day until designated time points.

While this invention has been particularly shown and described with references to example embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the scope of the invention encompassed by the appended claims.

CLAIMS

What is claimed is:

1. A compound represented by the following structural formula:



- 5 or a salt thereof, wherein:

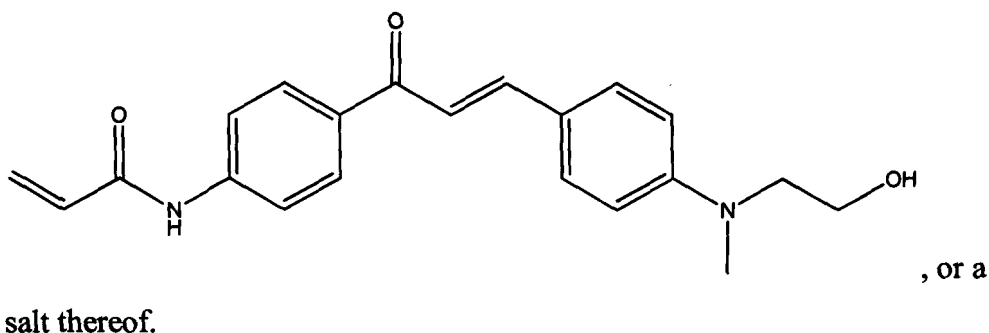
R is (C₁-C₁₆)alkyl, (C₂-C₁₆)alkenyl, (C₂-C₁₆)alkynyl, carbocyclyl or heterocyclyl, each of which is optionally substituted with one or more groups independently selected from the following: halo, NO₂, -NHCO(C₁-C₆)alkyl, -OCO(C₁-C₆)alkyl, -COO(C₁-C₆)alkyl, (C₁-C₁₀)alkyl, halo(C₁-C₁₀)alkyl, (C₁-C₁₀)alkoxy, -S(C₁-C₆)alkyl, (C₃-C₉)cycloalkyl, aryl, and phenoxy, wherein each phenoxy or aryl are further optionally substituted with one or more groups selected from halo, NO₂, (C₁-C₁₀)alkyl, halo(C₁-C₁₀)alkyl, and (C₁-C₁₀)alkoxy.

- 15 2. The compound of Claim 1, or a salt thereof, wherein R is (C₁-C₁₆)alkyl.
- 20 3. The compound of Claim 2, or a salt thereof, wherein R is (C₁-C₁₁)alkyl, optionally substituted with one, two or three groups independently selected from chloro, bromo, methyl, ethyl, methoxy, -COOCH₃, -OCOCH₃, cyclopentyl, phenyl, and phenoxy, wherein the phenyl or phenoxy is optionally and independently substituted with one, two, three, four, or five groups independently selected from fluoro, chloro, and methoxy.
- 25 4. The compound of Claim 1, or a salt thereof, where the compound is represented by the following structural formulas:

5. The compound of Claim 1, or a salt thereof, wherein R is (C₂-C₁₆)alkenyl.

5 6. The compound of Claim 5, or a salt thereof, wherein R is -CH=CH₂, -CH₂CH₂C=CH₂, (CH₂)₈C=CH₂, optionally substituted with one or more substituents selected from methyl and phenyl, wherein the phenyl is optionally substituted with chloro.

10 7. The compound of Claim 6, or a salt thereof, wherein the compound is represented by the following structural formula:



15 8. The compound of Claim 1, or a salt thereof, wherein R is carbocyclyl.

9. The compound of Claim 8, or a salt thereof, wherein R is an optionally substituted aryl.

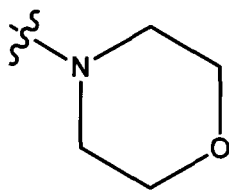
20 10. The compound of Claim 9, or a salt thereof, wherein R is an optionally substituted phenyl.

25 11. The compound of Claim 10, or a salt thereof, wherein R is phenyl, optionally substituted with one, two, three, four or five groups independently selected from halo, NO₂, -NHCO(C₁-C₆)alkyl, -OCO(C₁-C₆)alkyl, (C₁-C₁₀)alkyl, halo(C₁-C₄)alkyl, (C₁-C₁₀)alkoxy, and phenyl.

12. The compound of Claim 11, or a salt thereof, wherein R is phenyl, optionally substituted with one, two, three, four or five groups independently selected

from fluoro, chloro, bromo, iodo, -NO₂, -NHCOCH₃, -OCOCH₃, methyl, ethyl, propyl, isopropyl, butyl, pentyl, hexyl, heptyl, nonyl, decyl, -CF₃, -CH₂Cl, methoxy, ethoxy, butoxy, hexoxy, heptoxy, and phenyl.

- 5 13. The compound of Claim 9, or a salt thereof, wherein R is naphthyl, optionally substituted with one, two, three, four or five groups independently selected from halo, NO₂, -NHCO(C₁-C₆)alkyl, -OCO(C₁-C₆)alkyl, (C₁-C₁₀)alkyl, halo(C₁-C₄)alkyl, (C₁-C₁₀)alkoxy, and phenyl.
- 10 14. The compound of Claim 13, or a salt thereof, wherein R is naphthyl, optionally substituted with ethoxy.
- 15 15. The compound of Claim 8, or a salt thereof, wherein R is (C₃-C₉)cycloalkyl, optionally substituted with one or two groups independently selected from halo, NO₂, -NHCO(C₁-C₆)alkyl, -OCO(C₁-C₆)alkyl, (C₁-C₁₀)alkyl, halo(C₁-C₄)alkyl, (C₁-C₁₀)alkoxy, and phenyl.
- 20 16. The compound of Claim 15, or a salt thereof, wherein R is cyclohexyl, cyclopentyl, cyclobutyl, or cyclopropyl.
- 25 17. The compound of Claim 8, or a salt thereof, wherein R is (C₃-C₉)cycloalkenyl, optionally substituted with one or two groups independently selected from halo, NO₂, -NHCO(C₁-C₆)alkyl, -OCO(C₁-C₆)alkyl, (C₁-C₁₀)alkyl, halo(C₁-C₄)alkyl, (C₁-C₁₀)alkoxy, and phenyl.
- 30 18. The compound of Claim 17, or a salt thereof, wherein R is cyclohexenyl.
19. The compound of Claim 1, or a salt thereof, wherein R is heterocyclyl.
20. The compound of Claim 19, or a salt thereof, wherein R is



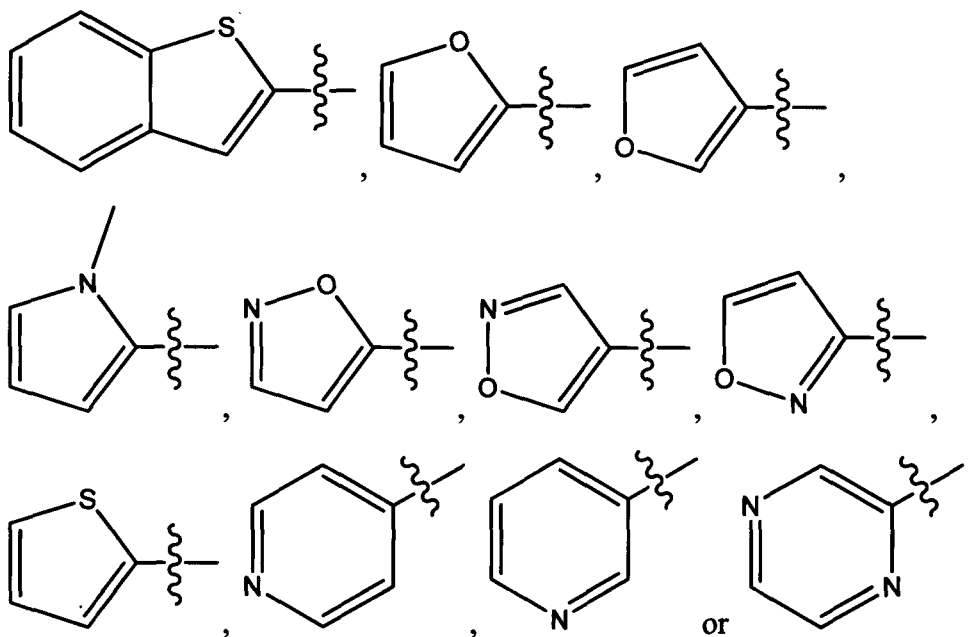
21. The compound of Claim 19, or a salt thereof, wherein R is optionally substituted heteroaryl.

5

22. The compound of Claim 21, or a salt thereof, wherein R is optionally substituted heteroaryl selected from benzo[b]thiophene, furanyl, pyrrole, isoxazole, thiophene, pyridine, and pyrimidine.

10

23. The compound of Claim 22, or a salt thereof, wherein R is selected from



wherein each R is optionally substituted with one, two, or three groups independently selected from fluoro, chloro, bromo, -NO₂, methyl, methoxy, -SCH₃, and phenyl, wherein the phenyl is optionally substituted with one, two or three groups independently selected from fluoro, chloro, and NO₂.

15

20

24. A method of isolating embryonic stem cells comprising the step of staining the cells with a compound of Claim 1.