



(51) International Patent Classification:

A61K 31/4164 (2006.01) C07D 233/66 (2006.01)

A61K 31/4166 (2006.01) C07D 233/70 (2006.01)

A61K 31/4178 (2006.01)

(21) International Application Number:

PCT/US2019/021175

(22) International Filing Date:

07 March 2019 (07.03.2019)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/639,952 07 March 2018 (07.03.2018) US

(71) Applicant: **THE PENN STATE RESEARCH FOUNDATION** [US/US]; 304 Old Main, University Park, Pennsylvania 16802 (US).

(72) Inventors: **ZHANG, Xin**; 225 Ghaner Drive, State College, Pennsylvania 16803 (US). **LIU, Yu**; 225 Ghaner Drive, State College, Pennsylvania 16803 (US). **WOLSTEN-HOLME, Charles**; 70 Rowland Drive, Fairfield, Connecticut 06824 (US).

(74) Agent: **STEWART, Duane A., III** et al.; Buchanan Ingersoll & Rooney PC, P. O. Box 1404, Alexandria, Virginia 22313-1404 (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, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, 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).

Declarations under Rule 4.17:

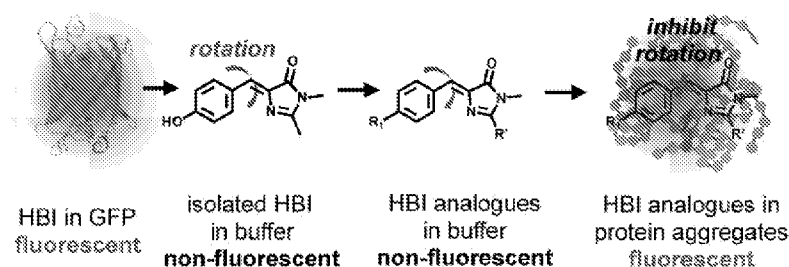
- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))
- as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))

Published:

- with international search report (Art. 21(3))

(54) Title: DYES FOR ANALYSIS OF SOLUBLE PROTEIN AGGREGATES OR MISFOLDED PROTEIN OLIGOMERS

Figure 1



(57) Abstract: Dye and compositions to monitor the multistep protein aggregation process in both test tubes and live cells are provided. These dyes can detect misfolded protein oligomers and distinguish insoluble protein aggregates from misfolded oligomers. Applications of these dyes include measuring kinetics of protein aggregation, monitoring aggregation of specific proteins in intact live cells, monitoring aggregation of cellular proteome in intact live cells, and tracking the time course of protein aggregation in cells under stress conditions.



DYES FOR ANALYSIS OF SOLUBLE PROTEIN AGGREGATES OR MISFOLDED PROTEIN OLIGOMERS

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of United States Provisional Patent Application No. 62/639,952, filed on March 7, 2018, which is incorporated by reference in its entirety.

FIELD OF THE INVENTION

[0002] The present invention relates to dyes and compositions for studying protein aggregation processes.

BACKGROUND OF THE INVENTION

[0003] Mechanisms of protein aggregation

[0004] Environmental stresses and pathogenic mutations of proteins lead to aberrant misfolding and aggregation, causing neurodegenerative diseases, including Alzheimer's disease, Parkinson disease, familial amyloidosis, and amyotrophic lateral sclerosis. Protein aggregation is a multistep process that has been associated with a growing number of human diseases, including neurodegenerative disorders, metabolic disorders, some cancers.¹⁻⁶ Misfolding yields misfolded monomers, which subsequently associate with one another to form misfolded oligomers. Misfolded oligomers evolve into insoluble aggregates in forms of amyloid- β fibrils, amorphous aggregates, or stress granules. Studying the multistep process of protein aggregation, in particular the intermediate misfolded oligomers, is increasingly being recognized as an important field in the biomedical and biochemical communities.

[0005] Misfolded protein oligomers in diseases

[0006] Protein homeostasis (proteostasis) dynamically adapts to diverse environmental factors and cellular events.¹ To achieve an appropriate level of proteostasis, the endogenous proteome has evolved to maintain a specific balance between the folded, misfolded and aggregate states of its protein components. However, exogenous stress conditions (including environmental perturbations, chemical toxins, and pathogen invasion) impair the integrity of proteostasis by shifting the free energy landscape of protein folding and/or inducing chemical or conformational changes in folded proteins.^{4,7} Failure to maintain proteostasis during stress leads to global

misfolding and aggregation of the endogenous proteome, resulting in a series of aberrant conformations that include misfolded proteins in the form of soluble oligomers, disordered or amorphous aggregates, and fibrils containing ordered hydrogen-bonded β -sheet structures. Formation of these structures often leads to the loss of essential functions and the formation of toxic aggregates. Both of these phenomena have been increasingly associated with a growing number of diseases, such as cancer, neurodegeneration, metabolic disorders, cardiovascular disease and inflammation.⁸⁻¹⁰

[0007] Decades of studies have delineated the structure, interaction, and activity of proteins in either their natively folded structures or in insoluble aggregates such as amyloid fibrils.

However, a variety of intermediate species exist between these two extreme states in the protein folding landscape. Herein, these conformations are collectively termed as misfolded oligomers, including soluble oligomers and pre-amyloid oligomers whose formation is driven by misfolded proteins. Accumulating evidence suggests that the misfolded oligomers may play key roles in both cell physiology and pathology.¹¹⁻¹⁵ Firstly, they may exert toxicity in diseases. For instance, soluble oligomers, but not the insoluble deposits, can confer synaptic dysfunction in neurodegenerative disorders². Secondly, beneficial functions have been demonstrated for oligomeric prion or prion-like proteins in processes including development, neuroprotection and metabolism³. Finally, they may be implicated in evolution. It has been shown that oligomers formed by prion proteins induce phenotypic changes in evolution⁴. With the existing knowledge, the biomedical community is in need of establishing methods to study misfolded protein oligomers in living cells.

[0008] Detection of protein aggregation in live cells

[0009] The need to reliably detect protein aggregation in live cells has promoted the emergence of a group of fluorescent-based methods.

[0010] First, chemical dyes (such as the PROTEOSTAT assay kit) can detect intracellular insoluble aggregates. However, this assay requires cell fixation and membrane permeabilization. Therefore, this method is not suited for live cells.

[0011] Second, fusion of Fluorescent Proteins (FP) or labeling of fluorescent probes to Protein-of-Interest (POI) has been used to visualize POI's aggregation by observing fluorescent granules in live cells.¹⁶⁻¹⁹ The limitation is that FP-fused POIs exhibit fluorescence before AND after aggregation (non-fluorogenic), and this non-fluorogenic nature makes these methods not suited

to visualize soluble oligomers because these oligomers do not have granular structures, nor visualize protein aggregation in certain subcellular compartments (such as mitochondria and stress granules) because of their granular morphology.

[0012] Third, diffusion constants of FP-fused POI can be quantified to differentiate insoluble aggregates from folded proteins.^{20,21} However, such assays may not easily distinguish misfolded oligomers from folded proteins because both exhibit similar diffusion constants.

[0013] Finally, fluorescence resonance energy transfer (FRET) of FP-fused POI has been used to distinguish misfolded oligomers from folded proteins.²¹⁻²³ However, this method is laborious and can encounter complications because these two conformations may exhibit similar FRET signals. Thus, despite efforts in past few decades, no simple and direct method is available to directly visualize the multistep process of how a POI aggregates in live cells.

[0014] Fluorescent protein chromophores

[0015] Fluorescent proteins (FPs) have been widely used as genetic tags to provide spatial and temporal information of a protein-of-interest (POI) in live organisms.^{25,26} Since its discovery, GFP has been used for various biological applications.^{38,39} Variations of the GFP chromophore, 4-hydroxybenzylidene-imidazolinone (HBI), have expanded FPs with diverse photophysical properties, including spectral range, quantum yield, photostability, and photoswitchability.²⁴ These chromophores, however, become mostly non-fluorescent when synthesized outside their protein cavity, largely due to rapid non-radiative decay via twisted-intramolecular charge transfer (TICT).^{28,29}

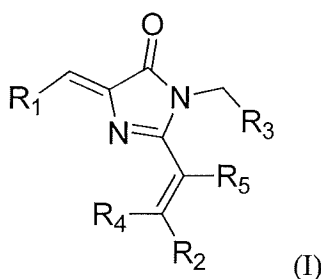
[0016] Although such behavior undermines the potential of FP analogues as a valuable class of fluorophores with broad applications, both chemical and biological restriction of TICT restores fluorescence of synthetic FP chromophores. Inspired by this property, fluorophores have been derived from HBI to be fluorescent with supreme photophysical properties. In addition, FP analogues locked in supra-molecular hosts,²⁹ metal-organic framework,³⁰ aggregated solid,³¹ and host proteins³² have been reported to fluoresce strongly. Applications of HBI analogues in biological imaging are represented by visualization of RNA aptamers and DNA quadruplex.³³⁻³⁷ Beyond these applications, these FP mimics are rarely reported in the detection of other biological processes in live cells.

SUMMARY OF THE INVENTION

[0017] The present invention designs, synthesizes and applies analogues of FP chromophores as fluorescent probes to visualize the multistep process of protein aggregation in live cells.

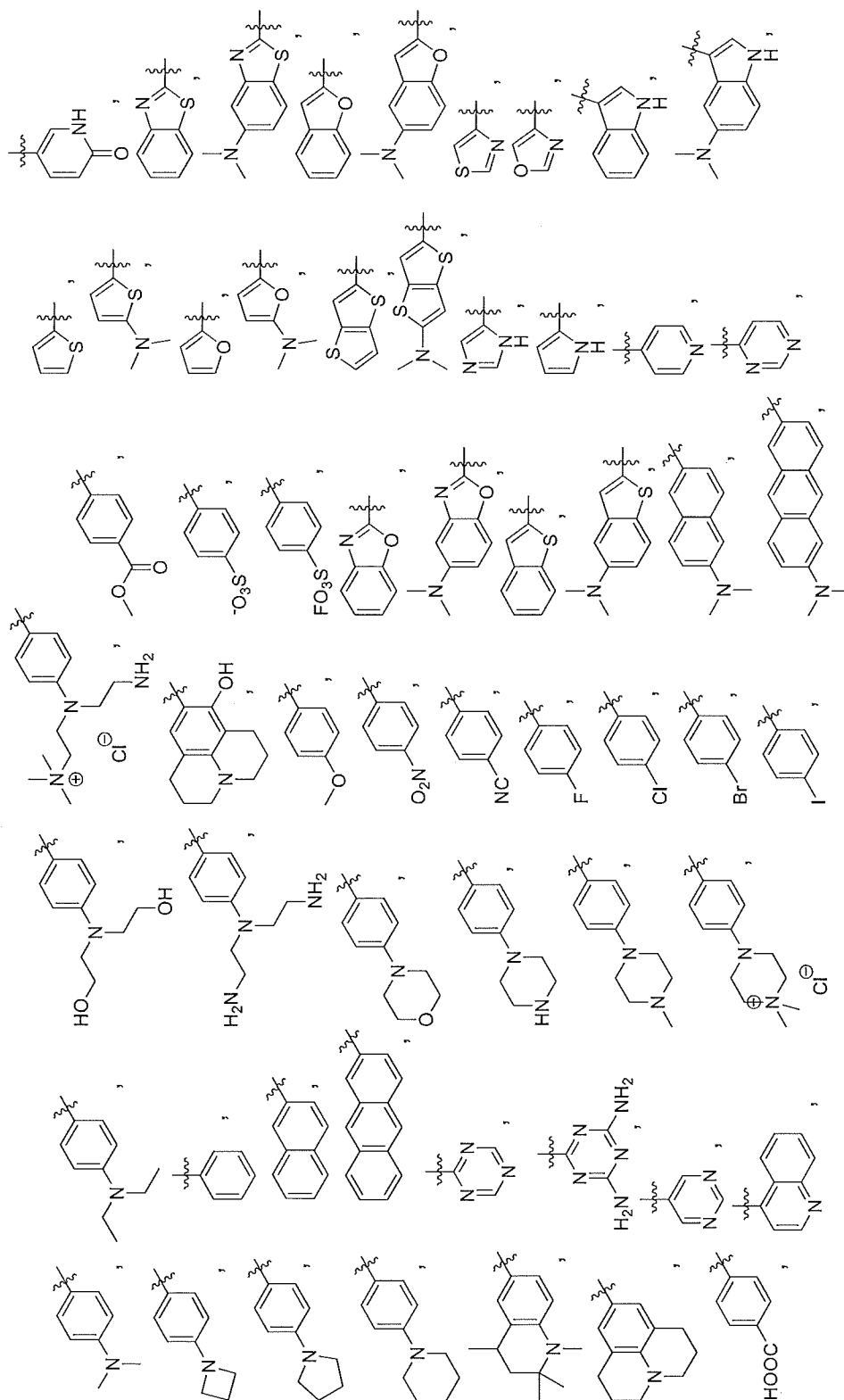
[0018] FP chromophores are used to visualize protein misfolding and aggregation, using turn-on fluorescence, both in test tube and in live cells. Thus, the inventors sensitize FP chromophores, whose TICT can be inhibited in the rigid environment within protein aggregates to turn on fluorescence (Figure 1).

[0019] One aspect of the present invention is directed to a compound of Formula I:



wherein:

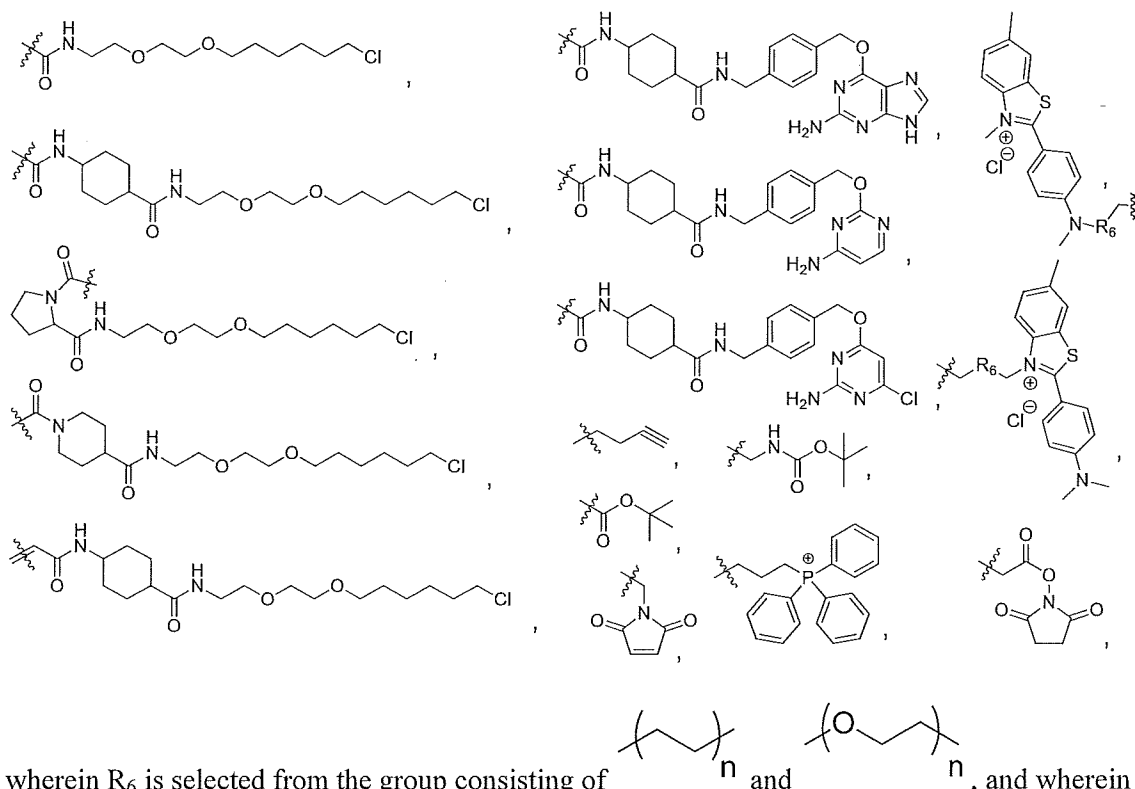
R₁, R₂, and R₅ are independently selected from the moieties of Group 1, and R₄ is -H;
 R₁, R₂, and R₄ are independently selected from the moieties of Group 1, and R₅ is -H;
 R₁ and R₂ are independently selected from the moieties of Group 1, R₄ is -H, and R₅ is -H;
 R₁ and R₂ are independently selected from the moieties of Group 1, R₄ is -H and R₅ is -CH₃;
 R₁ and R₂ are independently selected from the moieties of Group 1, R₄ is -CH₃ and R₅ is -H; or
 R₁, R₂, R₄, and R₅ are independently selected from the moieties of Group 1; and
 wherein the moieties of Group 1 are selected from the group consisting of



wherein substituents may be at para-, meta-, or ortho- positions on the aromatic ring, even if shown in only a single position above;

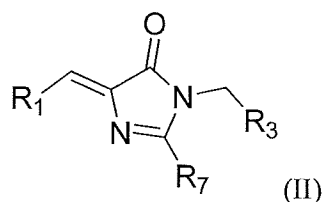
wherein R_3 is a directing moiety that binds and bioconjugates at least one biological target,

wherein R_3 is selected from the group consisting of:



wherein R_6 is selected from the group consisting of $(CH_2CH_2)_n$ and $(OCH_2CH_2)_n$, and wherein n is 0, 1, 2, 3, 4, or 5.

[0020] Another aspect of the present invention is directed to a compound of Formula II:

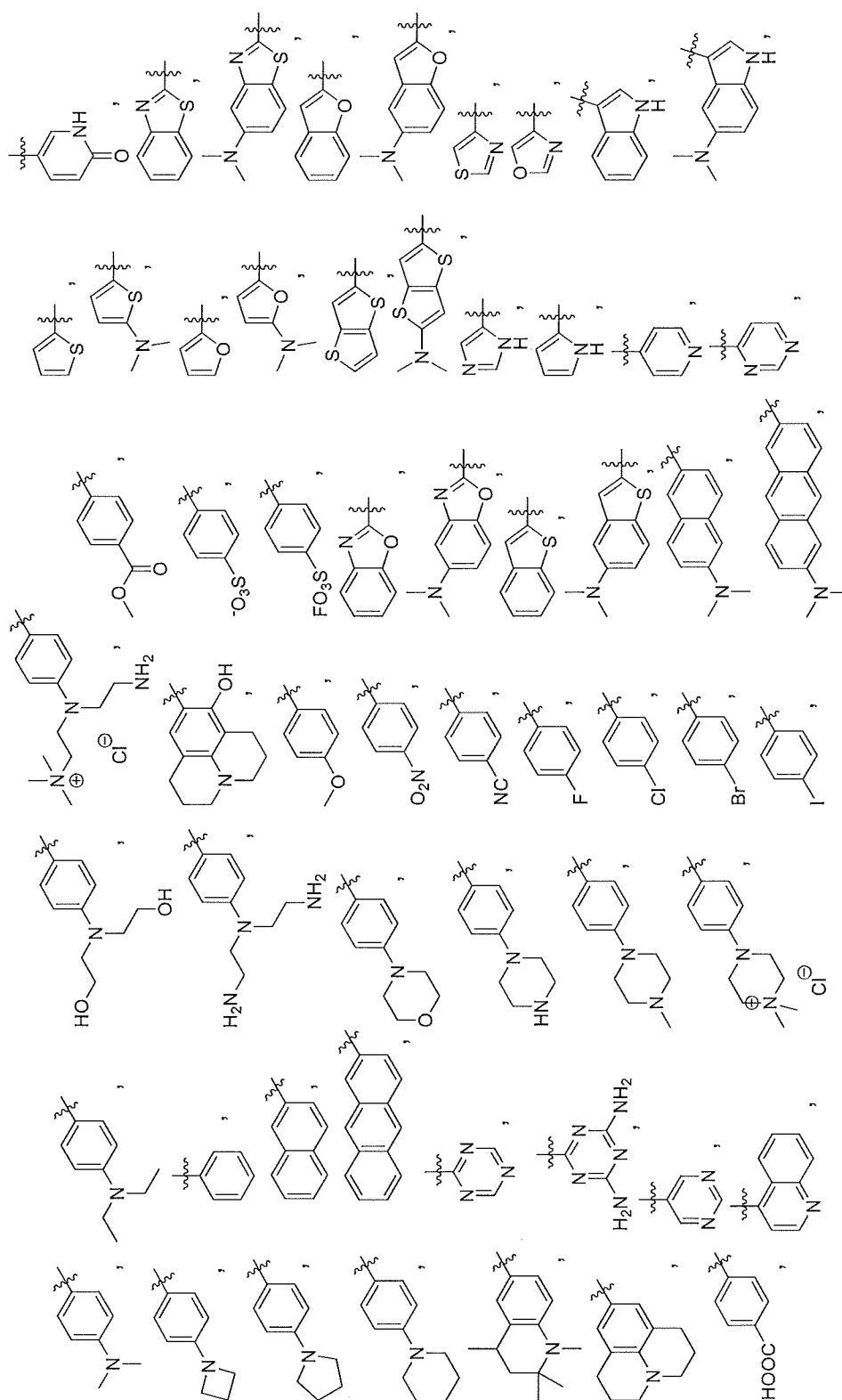


wherein

R_1 and R_7 are independently selected from a moiety of Group 1; or

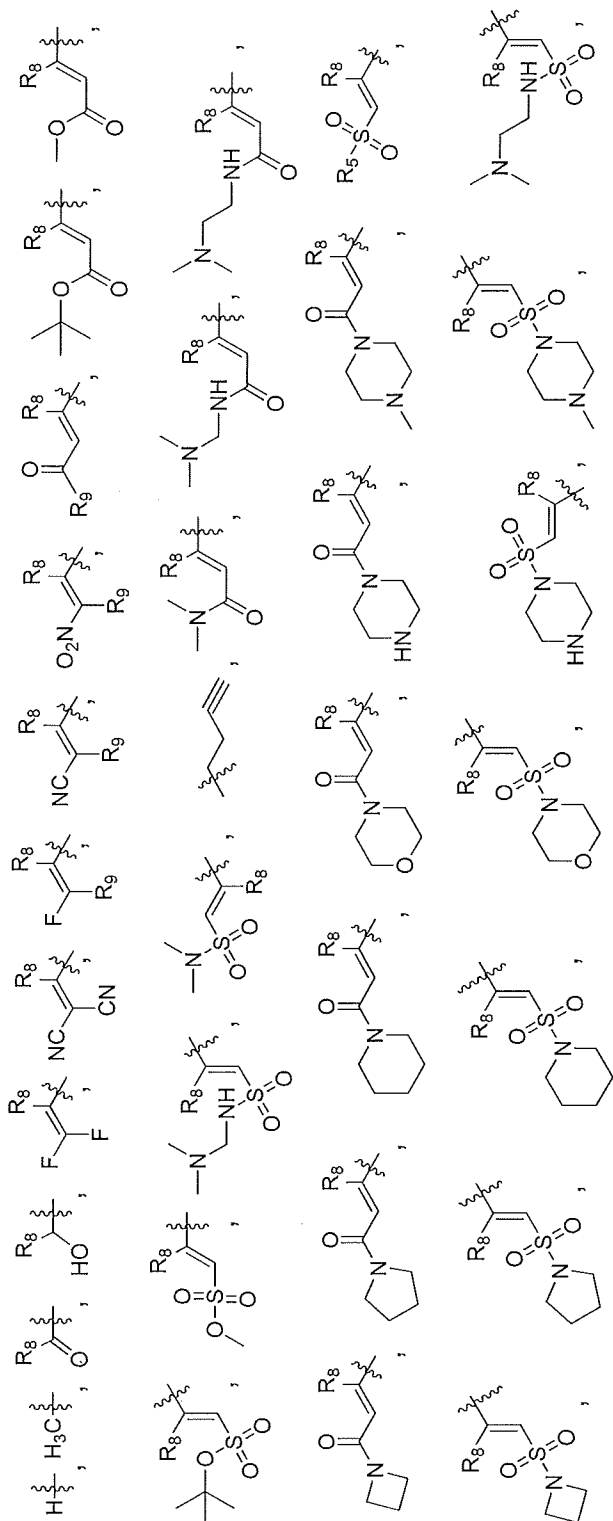
R_1 is a moiety of Group 1, R_7 is a moiety of Group 2, and R_8 and R_9 are independently selected from $-H$, $-CH_3$, and a moiety of Group 1; wherein

a moiety of Group 1 is selected from the group consisting of:

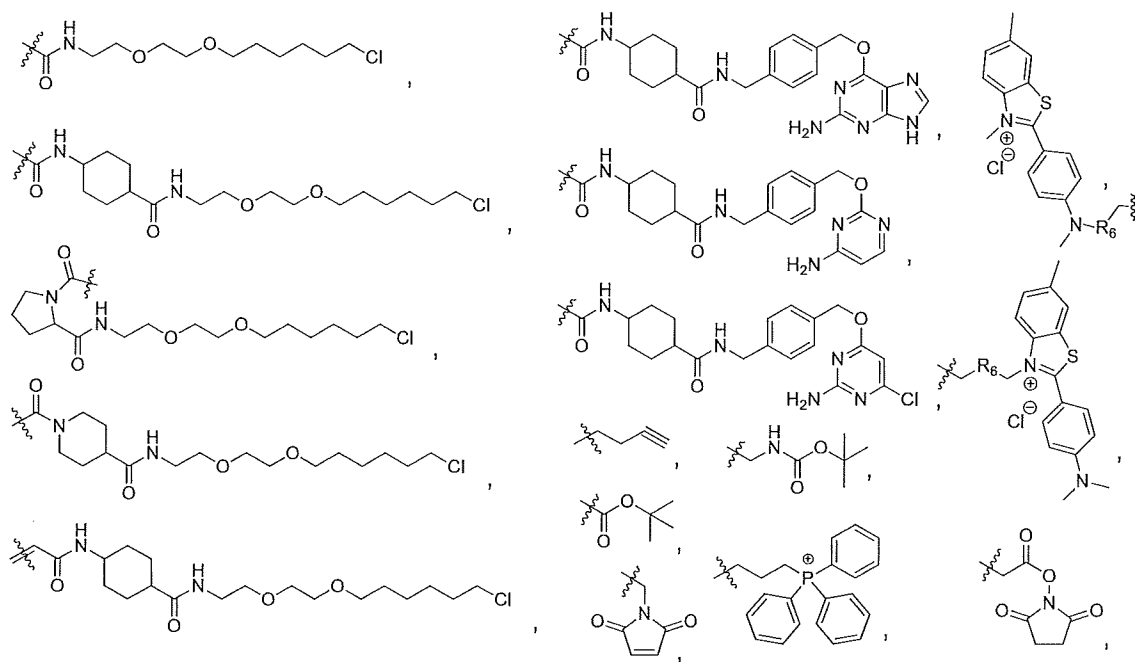


and wherein

a moiety of Group 2 is selected from the group consisting of:



wherein substituents may be at para-, meta-, or ortho- positions on the aromatic ring of Group 1 and Group 2, even if shown in only a single position above; and wherein R₃ is a directing moiety to bind and bioconjugate to biological targets, selected from the group consisting of:



wherein R₆ is selected from the group consisting of $\text{-(CH}_2\text{)}_n$ and $\text{-(OCH}_2\text{)}_n$, and wherein n is 0, 1, 2, 3, 4, or 5.

[0021] Another aspect of the present invention is directed to one of the compounds in Table 1.

[0022] Another aspect of the present invention is directed to a method for detecting at least one of aggregated protein, misfolded protein, and amyloid fiber in a protein of interest. The method includes performing a first measurement of fluorescence intensity of a protein of interest; adding to the protein of interest a fluorescent protein chromophore; and performing a second measurement of fluorescence intensity of the protein of interest. The increased fluorescence is indicative of at least one of aggregation, misfolding, and amyloid fiber in the protein of interest.

[0023] In one embodiment, the fluorescent protein chromophore is a compound selected from the previously listed compounds.

[0024] In another embodiment, the method is conducted in a cell in vivo.

[0025] In another embodiment, the method is conducted in vitro.

[0026] In another embodiment, the method further includes purifying the protein of interest prior to the first measurement.

[0027] In another embodiment, the fluorescent protein chromophore includes a thioflavin-T guiding group or a tert-butyloxycarbonyl guiding group.

[0028] Another aspect of the present invention is directed to a method for detecting at least one of aggregated protein, misfolded protein, and amyloid fiber in a protein of interest. The method includes performing a first measurement of fluorescence intensity of a standard protein; adding to a protein of interest a fluorescent protein chromophore; performing a measurement of fluorescence intensity of the protein of interest; and comparing the fluorescence intensity of the standard protein with the fluorescence intensity of the protein of interest. The increased fluorescence of the protein of interest is indicative of at least one of aggregation, misfolding, and amyloid fiber in the protein of interest.

[0029] Another aspect of the present invention is directed to a kit for detecting at least one of aggregated protein, misfolded protein, and amyloid fiber. The kit includes one or more fluorescent protein chromophores previously listed of a known concentration in a stock solution, one or more standard protein samples that form aggregated protein, misfolded protein, and amyloid fiber and optionally, instructions for use in detecting at least one of one of the aggregated protein, misfolded protein, and amyloid fiber.

[0030] In one embodiment, the stock solution is dimethyl sulfoxide or ethanol.

[0031] Another aspect of the present invention is directed to the use of the previously listed compounds to detect insoluble aggregates.

[0032] In one embodiment, the moiety of Group 2 can be an alkyl group wherein the alkyl group is saturated or unsaturated, linear or branched, substituted or unsubstituted.

[0033] Another aspect of the present invention is directed to a method for assaying aggregation of proteins in live cells is provided (AggGlow method). The kit includes in packaged combinations: (a) one or more of the previously listed compounds, and (b) instructions for using the compound for assaying aggregation of proteins in live cells.

[0034] Another aspect of the present invention is directed to a multi-dye composition comprising at least two dyes that are excited at different wavelengths is provided (AggGlow method).

[0035] Another aspect of the present invention is directed to a kit for assaying aggregation of a Halo-Tag fusion protein in live cells is provided (AggTag method). The kit includes in packaged

combinations: (a) one or more of the previously listed compounds, and (b) instructions for using the compound for assaying aggregation of proteins in live cells.

[0036] Another aspect of the present invention is directed to a kit for assaying aggregation of a SNAP-Tag fusion protein in live cells is provided (AggTag method). The kit comprises in packaged combinations: (a) one or more of the previously listed compounds, and (b) instructions for using the compound for assaying aggregation of proteins in live cells.

[0037] Another aspect of the present invention is directed to a multi-dye composition including at least two dyes that are (a) excited at different wavelengths and (b) conjugated to SNAP-Tag or Halo-Tag fusion proteins (AggTag method). The dyes are selected from the previously listed compounds.

[0038] Other aspects and advantages of the invention will be apparent from the following description, drawings and the appended claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[0039] The above and other objects, aspects, features, advantages and possible applications of the present innovation will be more apparent from the following more particular description thereof, presented in conjunction with the following drawings. Like reference numbers used in the drawings may identify like components.

[0040] Figure 1 shows an example of fluorescent protein chromophores modulated to serve as fluorogenic probes to detect protein aggregates.

[0041] Figure 2 shows an analogue of the Kaede chromophore is fluorescent in viscous solvent and crystal. (a) Chemical structures of the FP chromophore mimics in this work. (b) Relative brightness of fluorophores in glycerol. Calculation of brightness is shown in Table 2. (c) Crystal of **3** exhibits red fluorescence. (d) The planar structure of the fluorophore moiety of **3** in the crystal (the t-butyl group is omitted for simplicity).

[0042] Figure 3 shows *In vitro* detection of protein aggregates. (a) Protein aggregates provide a crowded environment to restrict rotational motion of **3** and turn on its fluorescence. (b) Aggregation kinetics of α -synuclein as measured by **3** and ThT. (c) Heat-induced aggregation of SOD1(V31A) increases fluorescence intensity of **3**. (d) Kinetics of fluorescence increase is faster than that of turbidity increase during aggregation of SOD1(V31A) at 54.5 °C.

[0043] Figure 4 shows The FP chromophore analogue **4** enables a fluorogenic method (AggTag) to detect aggregates of specific proteins in live cells. (a) Diagram of the fluorogenic AggTag

method that detects POI aggregation in live cells. (b) Structure of **4**. (c) Aggregation of the Htt 97Q-Halo-**4** conjugate induces fluorescence increase. (d) Fluorescent image of cell lysate from cells expressing Q0-Halo or Q97-Halo labeled by **4** after fractionation. Quantification of fluorescent intensity is shown in Figure 18c. T: total lysate; S: supernatant; I: insoluble fraction.

[0044] Figure 5 shows The AggTag method detects stress-induced protein aggregates that are invisible using non-fluorogenic methods. (a) Fluorogenic detection of NaAsO₂-induced SOD1(V31A) aggregation in HEK293T cells. (b) NaAsO₂ treatment leads to more SOD1(V31A) aggregation as confirmed by fractionation.

[0045] Figure 6 shows fluorescence response of FP chromophore analogues in H₂O with increasing concentrations of glycerol. FP analogues (20 μM) were prepared in glycerol:H₂O mixture with increasing glycerol concentrations. In addition, fluorescence intensity in 1,4-dioxane was measured to evaluate their response to polarity. Fluorescence measurement was carried out using a Tecan infinite M1000Pro fluorescence microplate reader at excitation and emission wavelengths as indicated. All readings were normalized against the fluorescence intensity in 100% glycerol as 1. Error bars: standard error (n=3).

[0046] Figure 7 shows normalized excitation spectra of FP chromophore analogues in glycerol. FP analogues (20 μM) were prepared in glycerol. Spectra were collected with emission wavelength of 420 nm for **1**, 525 nm for **2**, 620 nm for **3**, **3a** and **3b**. All measurements were carried out using a Tecan infinite M1000Pro fluorescence microplate reader.

[0047] Figure 8 shows normalized emission spectra of FP chromophore analogues in glycerol. FP analogues (20 μM) were prepared in glycerol. Spectra were collected with excitation wavelength of 370 nm for **1**, 455 nm for **2**, 530 nm for **3** and **3b**, 485 nm for **3a**. All measurements were carried out using a Tecan infinite M1000Pro fluorescence microplate reader.

[0048] Figure 9 shows absorbance spectra of FP chromophore analogues in glycerol. FP analogues (10 μM) were prepared in glycerol. Spectra were collected with 10 mm quartz cuvette. All measurements were carried out using Agilent 300 UV-Vis spectrophotometer.

[0049] Figure 10 shows the crystal structure of **3**. (a) The crystal of **3** is fluorescent. The crystals were looped out and transferred to a 35 mm glass bottom dish. Images were taken using an inverted Biorad ZOE fluorescent cell imager. This data suggests that **3** in the crystal structure should adopt conformation that is fluorescent. Reason of the dark image in the bright field is

because light of an inverted microscope could not pass through the thick crystal. **(b)** Crystal structure and packing diagram of **3**, with thermal ellipsoids drawn at 50% probability level.

[0050] Figure 11 shows computational analyses of **3** validate its planar S1 excited state structure as the fluorescent state and identify the charge separation at S1 excited state. **(a)** The optimized geometries of **3** in ground state and excited state. **(b-c)** A dihedral angle analysis confirmed planar structures of **3** in both ground state (GS) and excited state (ES). **(d)** Charge density difference isosurfaces (isovalue = 0.0004) at the minimum energy conical intersection between ground state and S1 excited state. Positive isosurfaces are blue and indicate electron withdraw. Negative isosurfaces are cyan and indicate electron donation. This data suggests that group A is the primary electron donor that contributes to charge separation at S1 excited state.

[0051] Figure 12 shows fluorogenic detection of mature α -synuclein fibers using **3** or ThT. Aggregation of α -synuclein was carried out using method described in the Supporting Information. Subsequently, half of reaction mixture was loaded on a 96-well plate to measure fluorescence intensity of **3** **(a)** or ThT **(b)** for the reaction mixture (T). The other half was centrifuged at 21,000 g for 30 minutes at 4 °C. The supernatant fraction (S) and pellet fraction (P, after resuspension with a volume that was equal to that of the S fraction) were loaded on a 96-well plate to measure fluorescence intensity. 10 μ M **3** or ThT was also incubated in buffer as a control. Parameters for fluorescence measurement: **3** (E_x = 530 nm/ E_m = 600 nm) and ThT (E_x = 440 nm/ E_m = 480 nm). Error bars: standard error (n=3).

[0052] Figure 13 shows formation of soluble oligomers and mature α -synuclein fibers. **(a)** Photo-induced cross-linking experiment observed soluble oligomers of α -synuclein at an 8-h time point. **(b)** Electron micrographs of α -synuclein fibrils formed at 72 h. Images were taken at different resolutions. Scale bars: 200 nm, 500 nm, 1 μ m, and 5 μ m.

[0053] Figure 14 shows fluorogenic detection of SOD1(V31A) aggregates using **3**. **(a)** SOD1(V31A) aggregation (42 μ M) was induced at 59 °C for 30 min in buffer (50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 83 mM EDTA), in the presence of **3** (21 μ M). Reaction mixture was illuminated by UV trans-illuminator before and after centrifugation at 21,000 g for 10 min at 4 °C. **(b)** Fluorescence intensity of SOD1(V31A) aggregates (42 μ M) in the presence of varying concentrations of **3**. 100 μ M of **3** in buffer was used as a control. Parameters for fluorescence measurement: **3** (E_x = 530 nm/ E_m = 600 nm). Error bars: standard error (n=3).

[0054] Figure 15 shows spectroscopic characterization of **4**. (a) Fluorescence response of **4** in H₂O with increasing concentrations of glycerol. 20 μ M of **4** was prepared in glycerol:H₂O mixture with increasing glycerol concentrations. In addition, its fluorescence intensity in 1,4-dioxane was measured. All readings were normalized against the fluorescence intensity in 100% glycerol as 1. Error bars: standard error (n=3). (b-c) Normalized excitation and emission spectra of **4** in glycerol. 20 μ M of **4** was prepared in glycerol. (d) Absorbance spectra of FP chromophore analogues in glycerol. FP analogues (10 μ M) were prepared in glycerol. Spectra were collected with 10 mm quartz cuvette. Parameters for fluorescence measurement: **4** (E_x = 530 nm/ E_m = 630 nm).

[0055] Figure 16 shows fluorescence response of SBD, CCVJ, and **4** to BSA and SDS. 20 μ M of fluorophores were incubated in H₂O, glycerol, dioxane, BSA (2 mg/mL), or SDS (0.2 %). Intensities were normalized against the value in either 100% glycerol as 1 for CCVJ and **4**, or dioxane as 1 for SBD. Parameters for fluorescence measurement: SBD (E_x = 450 nm/ E_m = 545 nm), CCVJ (E_x = 450 nm/ E_m = 500 nm), **4** (E_x = 530 nm/ E_m = 630 nm). Error bars: standard error (n=3).

[0056] Figure 17 shows confocal fluorescent images of *Htt* Q0-Halo, *Htt* Q97-Halo and *Htt* Q97-mCherry in HEK293T cells. (a) Schematic diagram for classic non-fluorogenic imaging approaches to visualize protein aggregates in live cells, via identification of fluorescent punctate structures due to the slower diffusion rate of insoluble aggregates than soluble proteins. (b) The *Htt* Q0-Halo, *Htt* Q97-Halo and *Htt* Q97-mCherry proteins were transiently transfected and expressed HEK293T cells for 24 h. 1 μ M TMR Halo-Tag ligand was added during protein expression to covalently label the *Htt* Q0-Halo and *Htt* Q97-Halo proteins. In cells expressing *Htt* Q97 proteins, obvious punctate structures (white arrows) were found in the background of diffuse fluorescence. Whereas, only diffuse fluorescence was observed in cells expressing the *Htt* Q0 protein (upper panel). Red: HaloTag-TMR or mCherry. Blue: nucleus stained by Hoechst 33342. Scale bar: 10 μ m.

[0057] Figure 18 shows fluorogenic detection of *Htt* Q97-Halo aggregation in HEK293T cells. (a) The dual-probe live cell imaging experiment. Halo-tag fusion proteins were labeled in the presence of both the non-fluorogenic coumarin ligand (1 μ M) and **4** (1 μ M). The newly synthesized *Htt*-Halo proteins were equally labeled with either the coumarin ligand or **4**. The coumarin ligand is fluorescent both before and after aggregation of *Htt*-Halo. By contrast, signal

of **4** exclusively arises from *Htt*-Halo aggregates. (b) Dual probes labeling confirms expression of proteins. The *Htt* Q0-Halo and *Htt* Q97-Halo proteins were transiently transfected and expressed in HEK293T cells for 24 h, in the presence of 1 μ M coumarin Halo-Tag ligand and 1 μ M of **4**. In cells expressing *Htt* Q97-Halo proteins, punctate structures in the coumarin channel coincided with the turn-on fluorescence of **4**. Whereas, only diffuse coumarin fluorescence was observed in cells expressing the *Htt* Q0-Halo protein and no turn-on fluorescence of **4** was found. (c) The *Htt* Q97 protein, without a Halo-Tag, was transiently transfected and expressed in HEK293T cells for 24 h, in the presence of 1 μ M of **4**. No turn-on fluorescence of **4** was found. Red: fluorescence from the SOD1(V31A)-Halo-**4** conjugate. Blue: coumarin Halo-Tag ligand in (a) or nucleus stained by Hoechst 33342 in (b). Scale bar: 10 μ m. (d) Fluorescent intensity of cell lysates from samples as represented in Figure 4d. Error bars: standard error (n=3).

[0058] Figure 19 shows confocal images of Halo-Tag in HEK293T cells treated with DPBS or NaAsO₂ (50 μ M, 24 h). During protein expression, medium is supplemented with 1 μ M of coumarin ligand and 1 μ M of **4** to enable a simultaneous dual-probe labeling to Halo-Tag. Note: Halo-Tag locates in both cytoplasm and nucleus. Blue: fluorescence from the Halo-Tag•coumarin conjugate. Red: fluorescence from the Halo-Tag-**4** conjugate. Scale bar: 10 μ m.

[0059] Figure 20 shows SOD1(V31A)-Halo fusion protein formed insoluble aggregates in cells treated with NaAsO₂ (extended from Figure 5b). Protein concentration of cell lysates was determined by a Bradford assay using pre-quantified BSA as a standard. Lysates of cells with or without NaAsO₂ stress were normalized to the same concentration and centrifuged at 21,000 g for 30 min at 4 °C. Insoluble fraction of cell lysate was resuspended in SDS-PAGE loading buffer and resolved in SDS-PAGE gel. The SOD1(V31A) was either visualized via fluorescence from **4** (FL on the right) or coomassie blue stain (CB on the right).

[0060] Figure 21 shows purified SOD1(V31A)-Halo fusion protein formed insoluble aggregates and turned on fluorescence of **4** under heat with time and temperature dependence. (a) Purified SOD1(V31A)-Halo and Halo-Tag proteins (42 μ M), both of which were conjugated with **4** (21 μ M), were incubated at varying temperatures (25, 37, 38.9, 41.8, 45.6, 50.7, 54.5, 57.2, and 59 °C). As control samples, **4** (21 μ M) was prepared in buffer, 1,4-dioxane, glycerol, BSA (2 mg/mL). Fluorescence intensity was recorded at $E_x = 530$ nm/ $E_m = 600$ nm. All readings were normalized to intensity of the Halo-Tag-**4** conjugate at 25 °C. Fluorescence of protein samples was also visualized via UV transilluminator. In addition, protein samples were centrifuged at

21,000 g for 30 min at 4 °C to analyze their insoluble fraction using fluorescence of **4** as signal on SDS-PAGE gel. While obvious insoluble aggregates were only found at 57.2 and 59 °C, turn-on fluorescence of the SOD1(V31A)-Halo-**4** conjugate was observed at lower temperatures. This data suggests that **4** is able to fluoresce prior to formation of insoluble aggregates. This fluorescence was not originated from aggregation of Halo-Tag domain, as evidenced by the low fluorescence of the Halo-**4** conjugate and lack of significant insoluble fraction at these temperatures. Error bars: standard error (n=3). **(b)** Purified SOD1(V31A)-Halo and Halo-Tag proteins (42 µM) were conjugated with **4** (21 µM) and incubated at 54.5 °C. Fluorescence was recorded using Agilent Carey Eclipse fluorescence spectrophotometer with temperature control. Buffer condition: 50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 83 mM EDTA. Parameters for fluorescence measurement: **4** ($E_x = 530$ nm/ $E_m = 600$ nm).

[0061] Figure 22 shows fluorescence of the SOD1(V31A)-Halo-**4** conjugate originates from protein misfolding. **(a)** Purified SOD1(V31A)-Halo (42 µM) was conjugated with **4** (21 µM) and incubated at 54.5 °C. Fluorescence was recorded using Agilent Carey Eclipse fluorescence spectrophotometer with temperature control. Turbidity was recorded using Agilent 300 UV-Vis spectrophotometer. Buffer condition: 50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 83 mM EDTA. **(b)** Kinetics of fluorescence increases from the SOD1(V31A)-Halo-**4** conjugate (red curve) was compared to the kinetics of protein misfolding (black curve) as measured by Tryptophan fluorescence change. Fluorescence was recorded using Agilent Carey Eclipse fluorescence spectrophotometer with temperature control. Experiments were carried out at 54.5 °C. Parameters for fluorescence measurement: **4** ($E_x = 530$ nm/ $E_m = 600$ nm) and Trp ($E_x = 290$ nm/ $E_m = 330$ and 345 nm). Error bars: standard error (n=3).

[0062] Figure 23 shows fluorescence of **3** in HEK-293T cells treated with 5 µM MG132 for 24 h, 37 °C, 5% CO₂ (middle panel) and in HeLa cells treated with 1 µM 17-AAG for 24 h, 37 °C, 5% CO₂ (right panel). Untreated cells remain dark as shown in right panel.

[0063] Figure 24 shows that **5** reports on insoluble aggregates formed by SOD1-A4V-Halo **(A)**. **(B-C)** Excitation and emission spectra of **4** (with misfolded oligomers of SOD1-A4V-Halo) and **5** (with insoluble aggregates of SOD1-A4V-Halo). **(D-E)** When treated with 5 µM MG132 for 8 h, 37 °C, 5% CO₂, HEK-293T cells expressing SOD1-A4V-Halo protein shows fluorescence of **4** **(D)** but not **5** **(E)**. **(F-G)** HEK-293T cells expressing Htt-110Q-Halo protein shows fluorescence

of **4 (D)** and **5 (D)**. **(H)** HEK-293T cells expressing Halo and 110Q-mCherry did not show fluorescence of **5**.

[0064] Figure 25 shows probes for SNAP-tag conjugation. **(A)** Chemical scaffold of SNAP-based probes. **(B)** Linkers with varying length or rigidity (**6-11**). **(C)** Synthesis of SNAP-conjugating probes. **(D)** Quantum yield of P1 probes when conjugated with folded SNAP-tag.

DETAILED DESCRIPTION OF THE INVENTION

[0065] Definitions

[0066] While the terms used herein are believed to be well understood by one of ordinary skill in the art, definitions are set forth herein to facilitate explanation of the subject matter disclosed herein.

[0067] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which the subject matter disclosed herein belongs. Although any methods, devices, and materials similar or equivalent to those described herein can be used in the practice or testing of the presently disclosed subject matter, representative methods, devices, and materials are described herein.

[0068] The terms “a,” “an,” and “the” refer to “one or more” when used in this application, including the claims. The use of the word “a” or “an” when used in conjunction with the term “comprising” in the claims and/or the specification may mean “one,” but it is also consistent with the meaning of “one or more,” “at least one,” and “one or more than one.”

[0069] All references to singular characteristics or limitations of the present disclosure shall include the corresponding plural characteristic(s) or limitation(s) and vice versa, unless otherwise specified or clearly implied to the contrary by the context in which the reference is made.

[0070] All combinations of method or process steps as used herein can be performed in any order, unless otherwise specified or clearly implied to the contrary by the context in which the referenced combination is made.

[0071] The methods and devices of the present disclosure, including components thereof, can comprise, consist of, or consist essentially of the essential elements and limitations of the embodiments described herein, as well as any additional or optional components or limitations described herein or otherwise useful.

[0072] Unless otherwise indicated, all numbers expressing physical dimensions, quantities of ingredients, properties such as reaction conditions, and so forth used in the specification and

claims are to be understood as being modified in all instances by the term “about”. Accordingly, unless indicated to the contrary, the numerical parameters set forth in this specification and claims are approximations that can vary depending upon the desired properties sought to be obtained by the presently disclosed subject matter.

[0073] As used herein, ranges can be expressed as from “about” one particular value, and/or to “about” another particular value. It is also understood that there are a number of values disclosed herein, and that each value is also herein disclosed as “about” that particular value in addition to the value itself. For example, if the value “10” is disclosed, then “about 10” is also disclosed. It is also understood that each unit between two particular units are also disclosed. For example, if 10 and 15 are disclosed, then 11, 12, 13, and 14 are also disclosed.

[0074] The following description is of exemplary embodiments that are presently contemplated for carrying out the present invention. This description is not to be taken in a limiting sense, but is made merely for the purpose of describing the general principles and features of the present invention. The scope of the present invention is not limited by this description.

[0075] As noted above, embodiments apply an FP chromophore to visualize protein misfolding and aggregation, using turn-on fluorescence, both in test tube and in live cells.

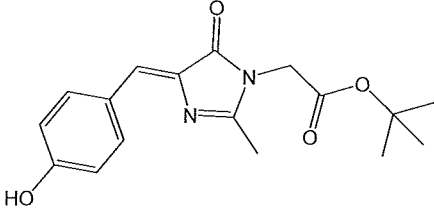
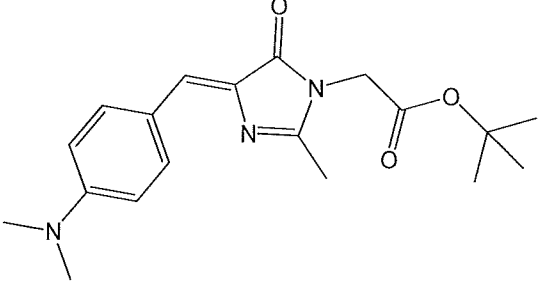
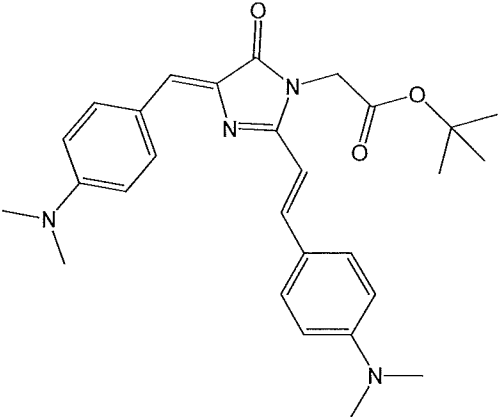
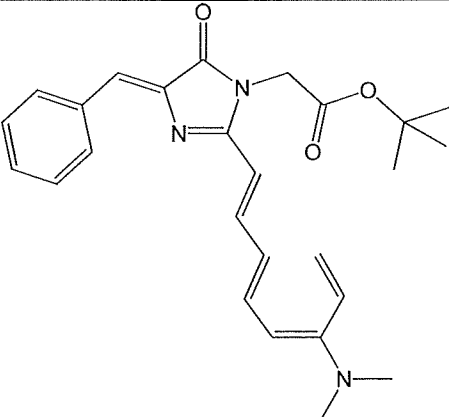
[0076] The present invention provides dyes, reagents and methods useful for detection of misfolded protein oligomers and insoluble protein aggregates in vitro and in vivo. In some embodiments, the invention provides a family of probes containing an imidazolinone core structure.

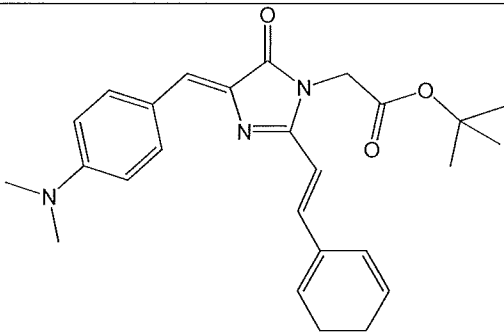
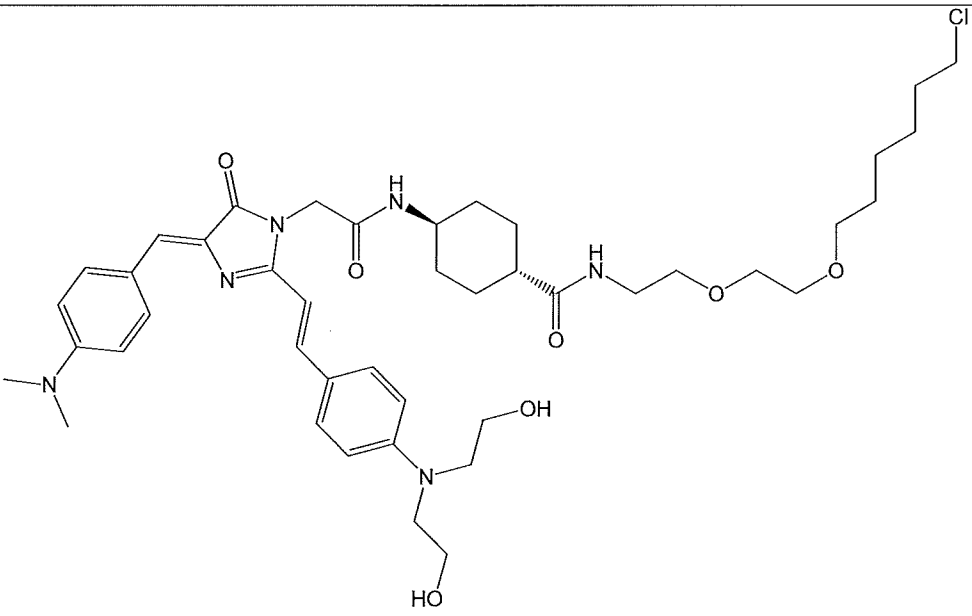
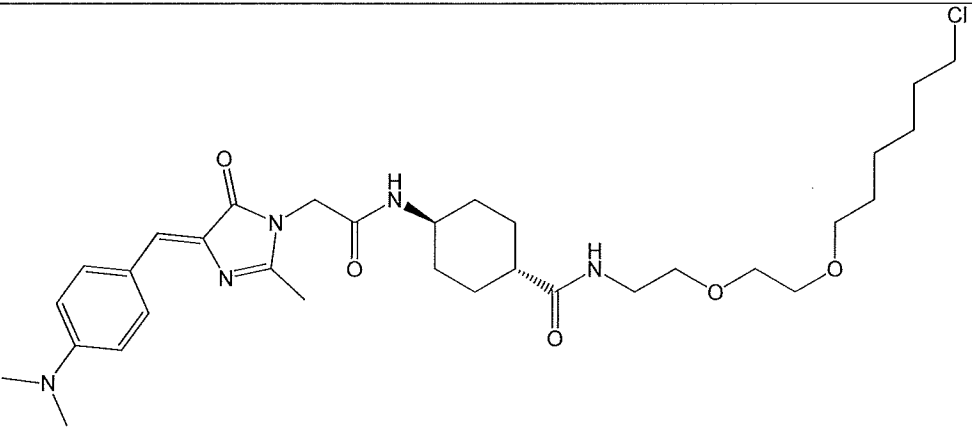
[0077] The probes of the invention are useful for generating fluorescence signals that depend upon the presence of an aggregated form of a protein, while conveying minimal levels of signals when only the native form of the protein is present.

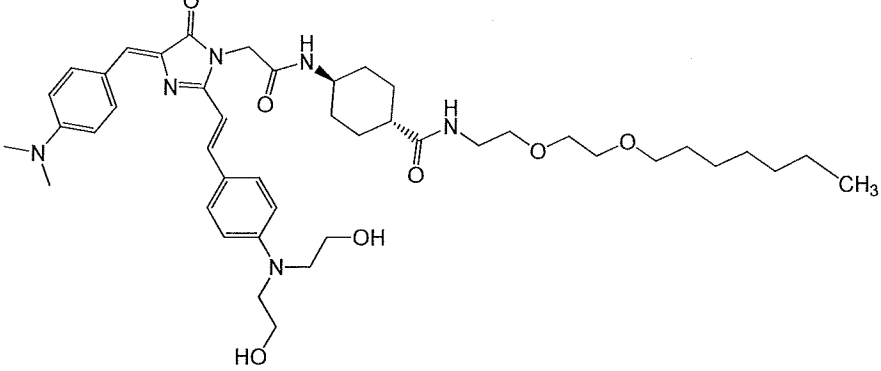
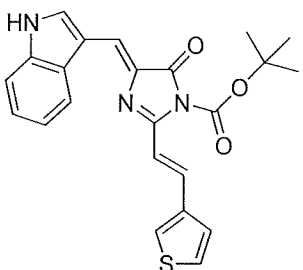
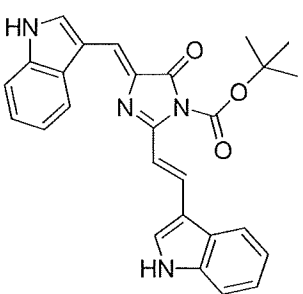
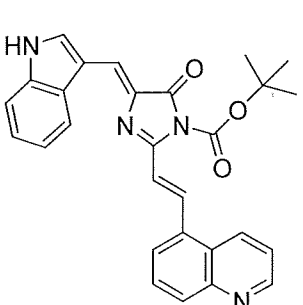
[0078] Many of these probes exhibit large variation in their excitation and emission wavelength thereby allowing for potential multicolor detection of protein aggregation and sequential imaging.

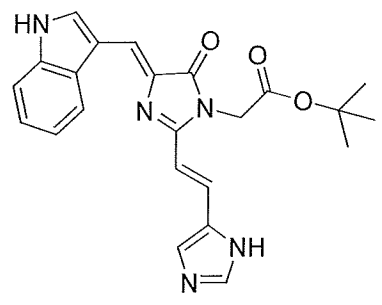
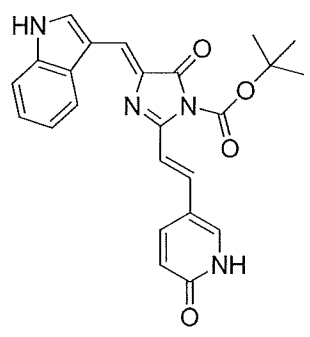
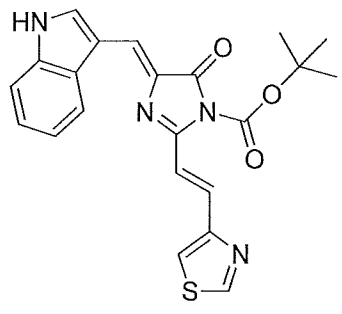
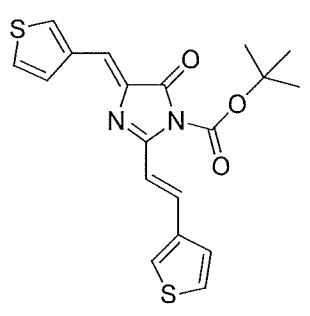
[0079] Specific embodiments of FP chromophores are reported herein. Some of these FP chromophores are included in Table 1, below. In each cell of the table, a compound and its assigned compound number are presented:

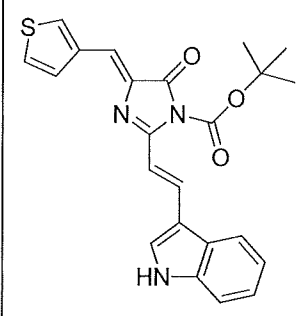
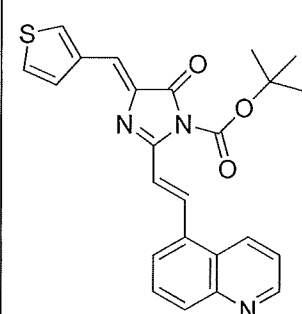
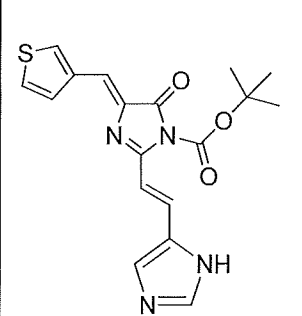
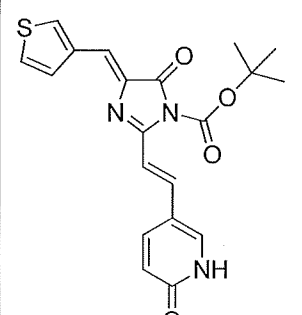
Table 1

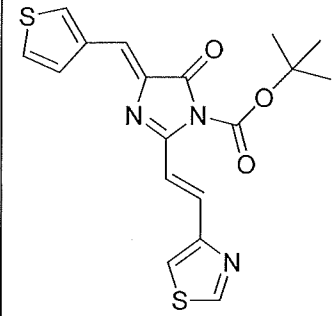
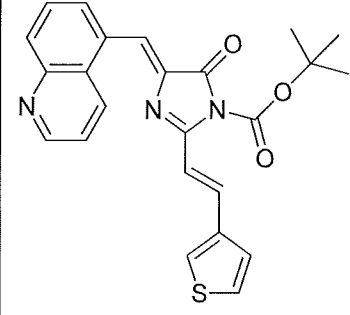
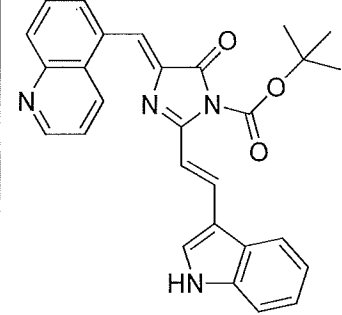
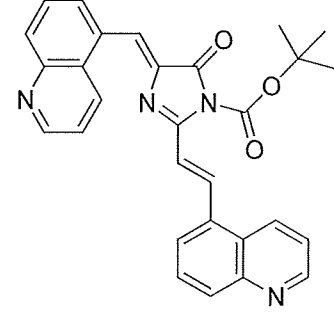
Compound Number	Compound
1	
2	
3	
3a	

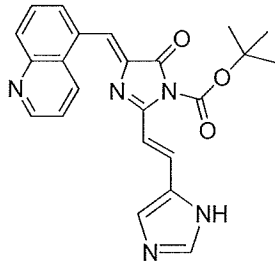
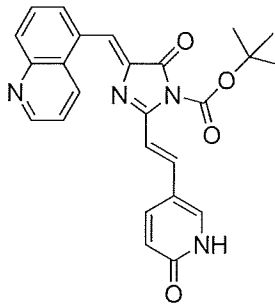
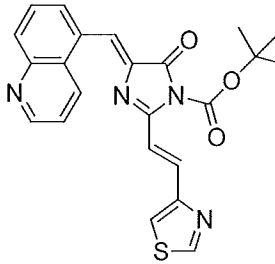
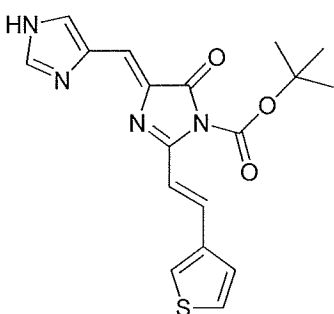
3b	
4	
5	

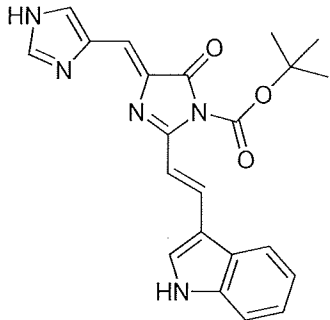
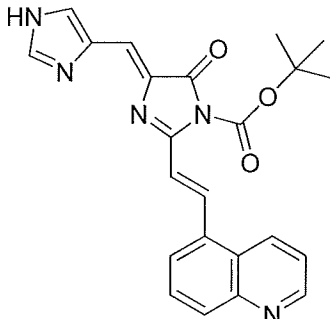
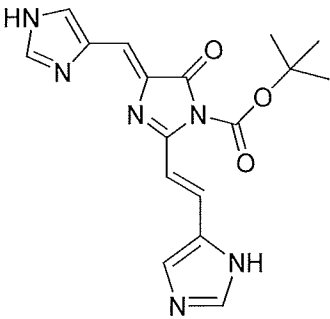
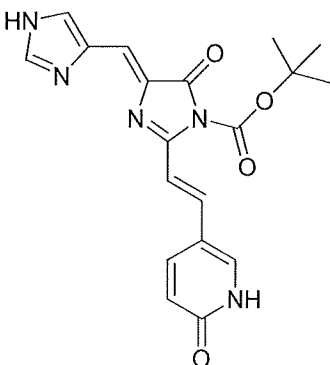
6	 <chem>CN(C)c1ccc(cc1)C#CN2C(=O)N(C2)CCN3CCCCC3C(=O)NCCOCCOCCOCCCCC</chem>
7	 <chem>CC(C)(C)OC(=O)N1C(=O)N(C=Cc2ccccc2)C1=CN</chem>
8	 <chem>CC(C)(C)OC(=O)N1C(=O)N(C=Cc2c[nH]c3ccccc23)C1=CN</chem>
9	 <chem>CC(C)(C)OC(=O)N1C(=O)N(C=Cc2c3ccncc3cc2)C1=CN</chem>

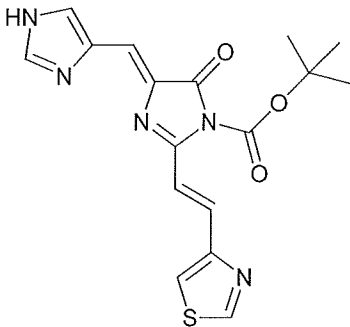
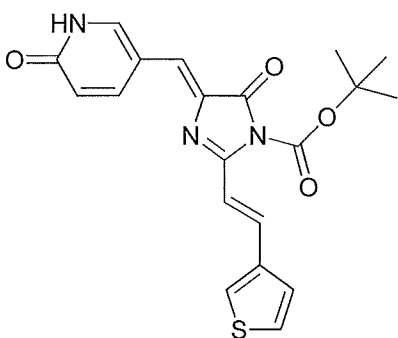
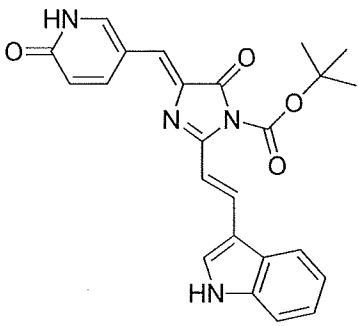
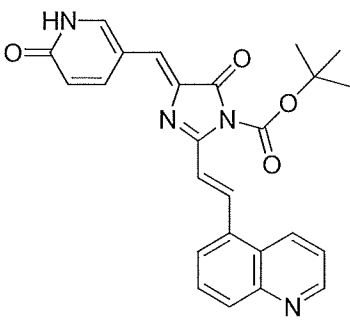
10	
11	
12	
13	

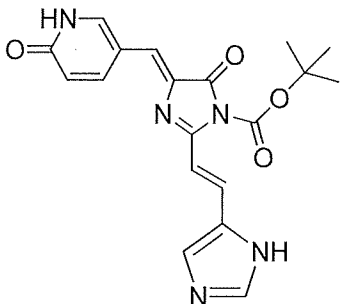
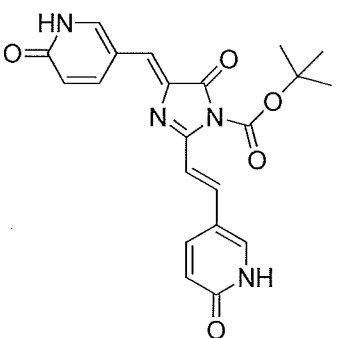
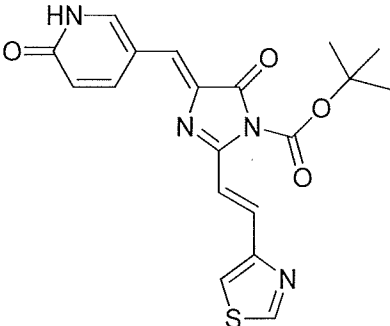
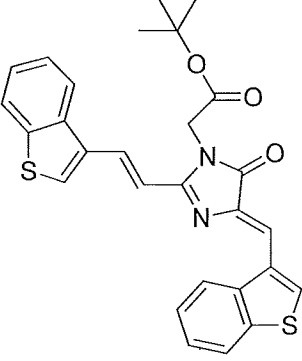
14	 <chem>CC(C)(C)OC(=O)N1C(=O)N(C=Cc2c[nH]c3ccccc23)C(=N1)/C=C/c4ccsc4</chem>
15	 <chem>CC(C)(C)OC(=O)N1C(=O)N(C=Cc2cnc3ccccc23)C(=N1)/C=C/c4ccsc4</chem>
16	 <chem>CC(C)(C)OC(=O)N1C(=O)N(C=Cc2c[nH]cn2)C(=N1)/C=C/c3ccsc3</chem>
17	 <chem>CC(C)(C)OC(=O)N1C(=O)N(C=Cc2c[nH]c(=O)c3ccccc23)C(=N1)/C=C/c4ccsc4</chem>

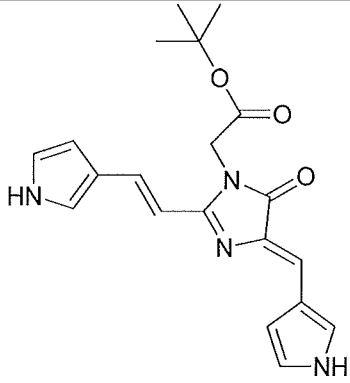
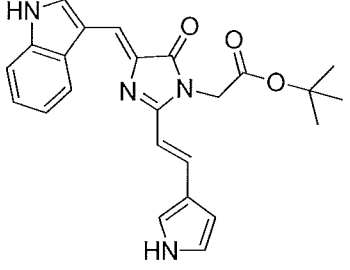
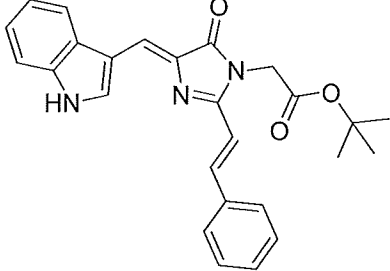
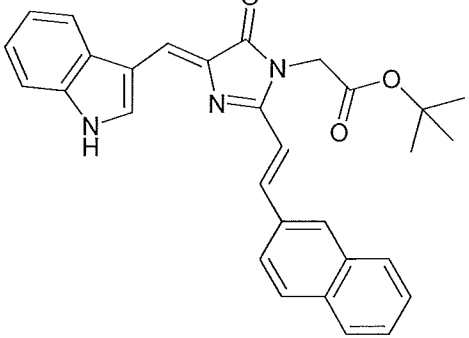
18	
19	
20	
21	

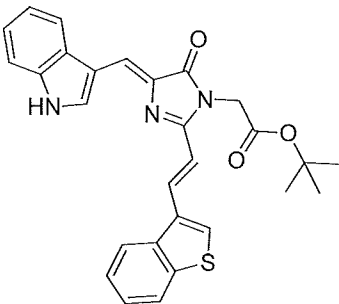
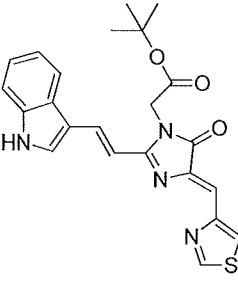
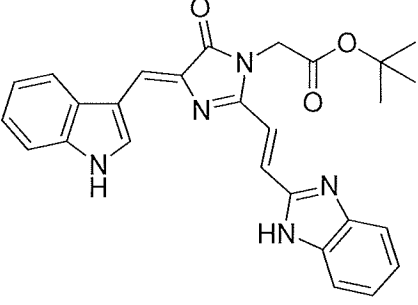
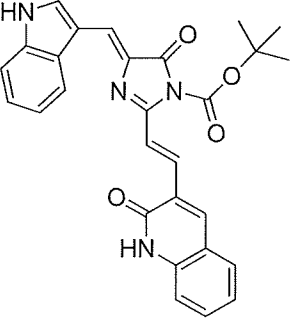
22	 <chem>CC(C)(C)OC(=O)N1C(=C(C=C1C=Cc2c[nH]cn2)C(=O)N3C=CC=C4C=CC=CC=C34)c5c[nH]cn5</chem>
23	 <chem>CC(C)(C)OC(=O)N1C(=C(C=C1C=Cc2c[nH]cn2)C(=O)N3C=CC=C4C=CC=CC=C34)c5c[nH]cn5</chem>
24	 <chem>CC(C)(C)OC(=O)N1C(=C(C=C1C=Cc2c[nH]cn2)C(=O)N3C=CC=C4C=CC=CC=C34)c5c[nH]cn5</chem>
25	 <chem>CC(C)(C)OC(=O)N1C(=C(C=C1C=Cc2c[nH]cn2)C(=O)N3C=CC=C4C=CC=CC=C34)c5c[nH]cn5</chem>

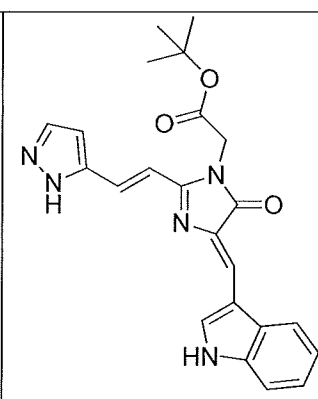
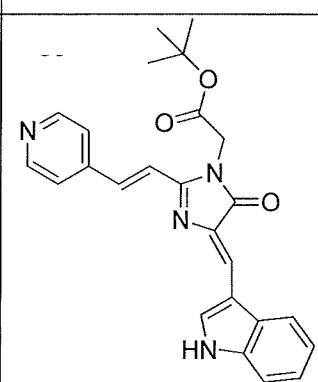
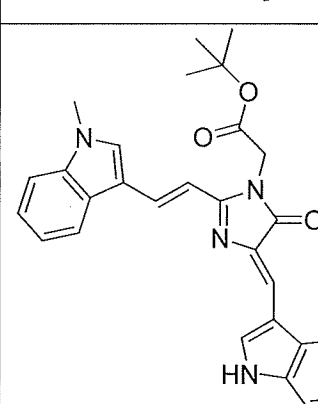
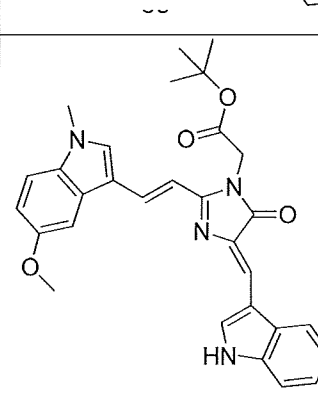
26	
27	
28	
29	

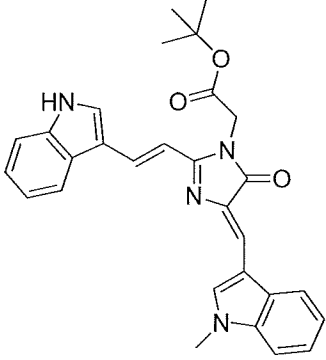
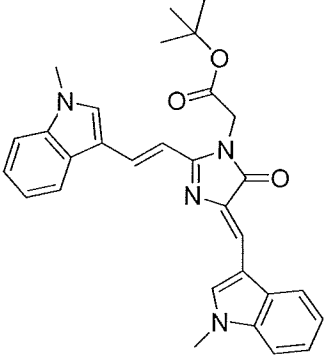
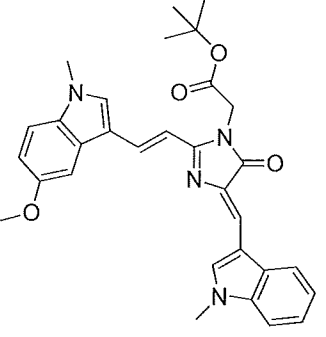
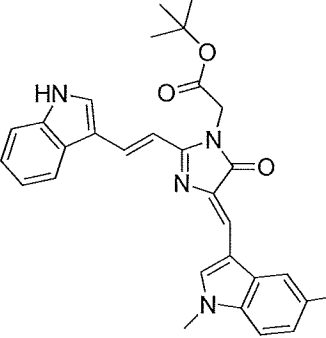
30	
31	
32	
33	

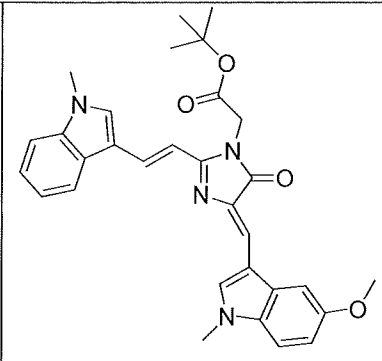
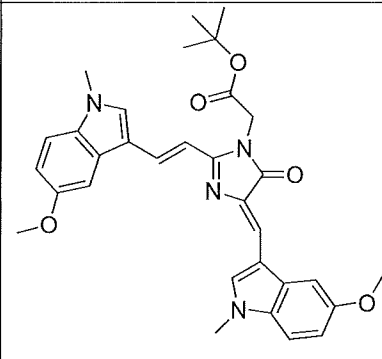
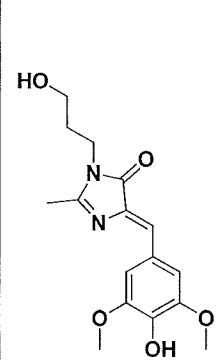
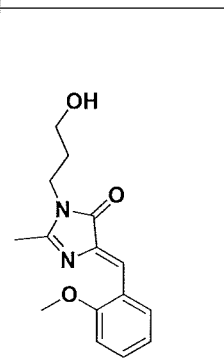
34	 <chem>CC(C)(C)OC(=O)N1C(=O)N(C=Cc2cc[nH]2)C(=N1)/C=C/c3ccc[nH]3</chem>
35	 <chem>CC(C)(C)OC(=O)N1C(=O)N(C=Cc2cc[nH]2)C(=N1)/C=C/c3ccc(=O)[nH]3</chem>
36	 <chem>CC(C)(C)OC(=O)N1C(=O)N(C=Cc2cc[nH]2)C(=N1)/C=C/c3cc[nH]s3</chem>
37	 <chem>CC(C)(C)OC(=O)CN1C(=O)N(C=Cc2c3ccccc3s2)C(=N1)/C=C/c4c5ccccc5s4</chem>

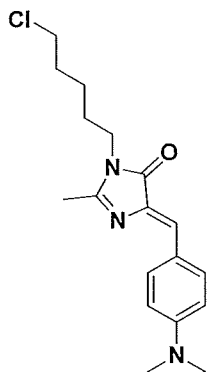
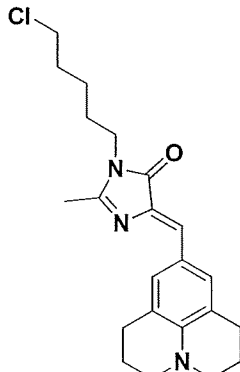
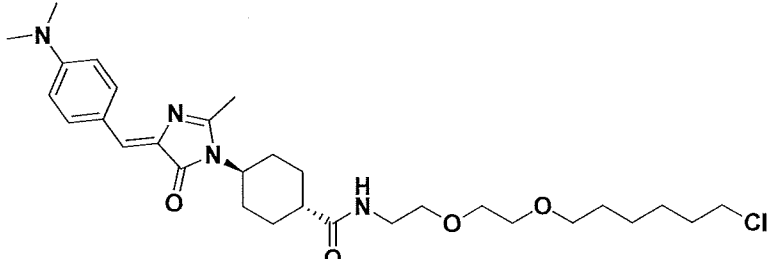
38	 <chem>CC(C)(C)OC(=O)CN1C(=O)N(C=Cc2c[nH]c[nH]2)N1C=Cc3c[nH]c[nH]3</chem>
39	 <chem>CC(C)(C)OC(=O)CN1C(=O)N(C=Cc2c[nH]c[nH]2)N1C=Cc3c[nH]c[nH]3</chem>
40	 <chem>CC(C)(C)OC(=O)CN1C(=O)N(C=Cc2c[nH]c[nH]2)N1C=Cc3c[nH]c[nH]3</chem>
41	 <chem>CC(C)(C)OC(=O)CN1C(=O)N(C=Cc2c[nH]c[nH]2)N1C=Cc3c[nH]c[nH]3</chem>

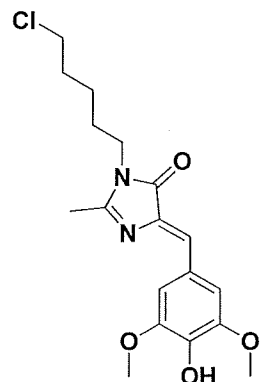
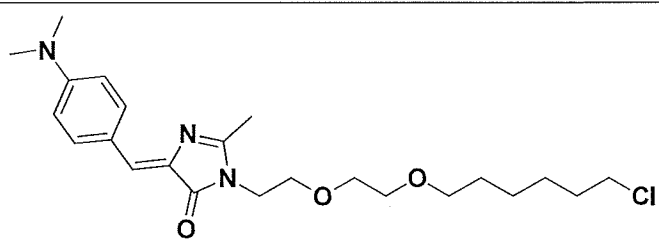
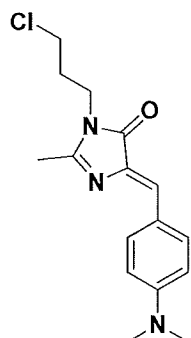
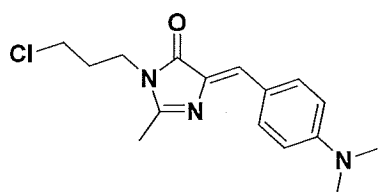
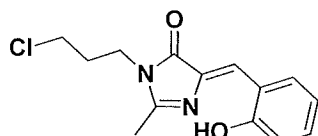
42	
43	
44	
45	

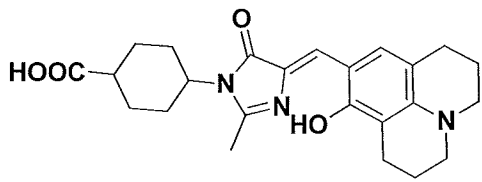
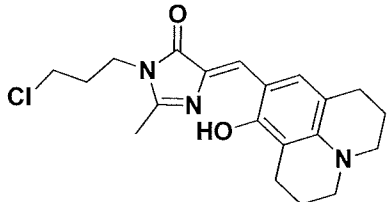
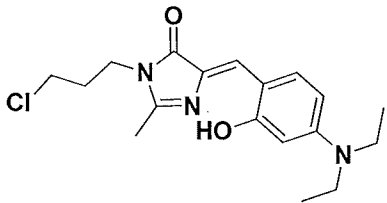
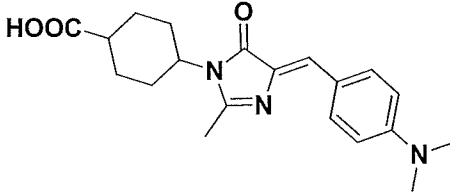
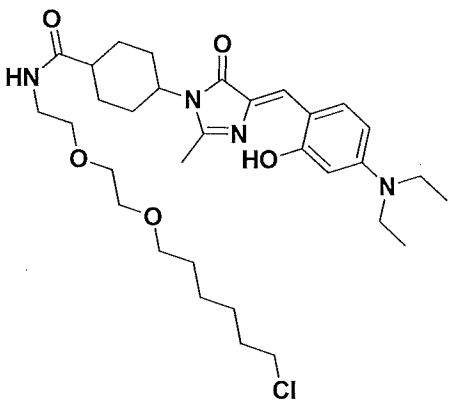
46	
47	
48	
49	

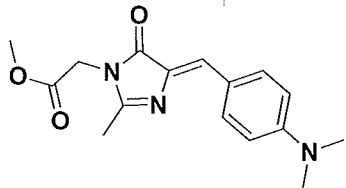
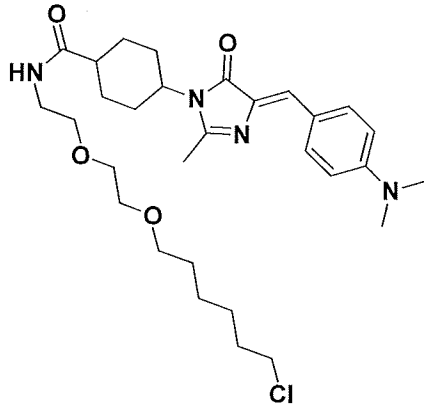
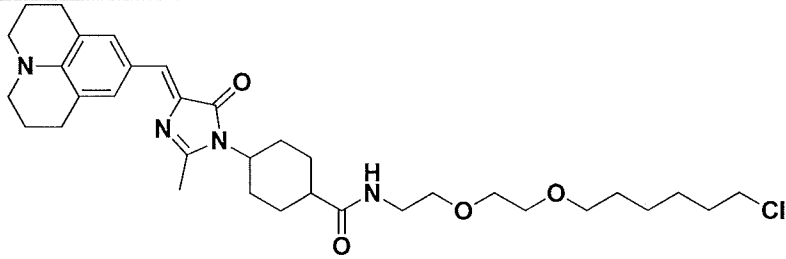
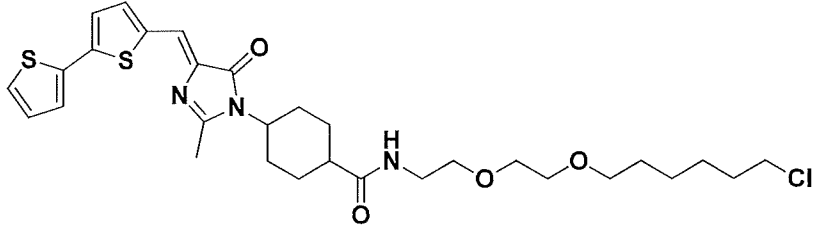
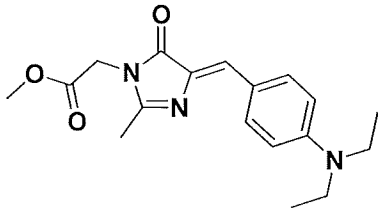
50	
51	
52	
53	

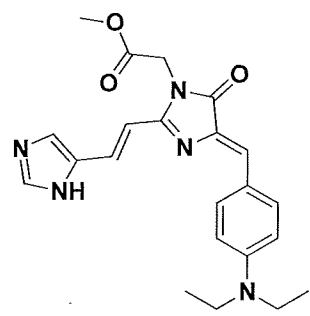
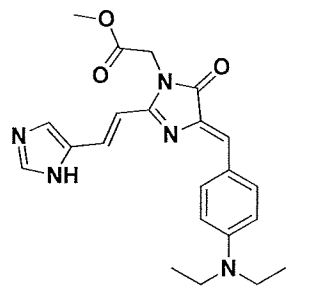
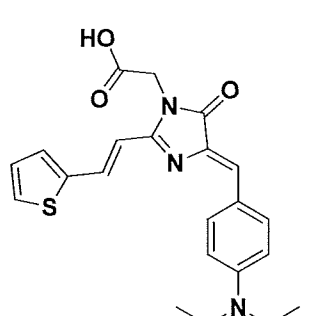
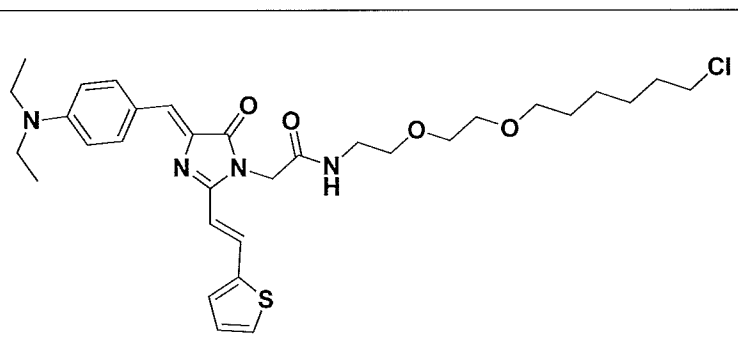
54	
55	
56	
57	

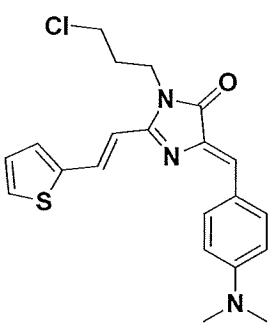
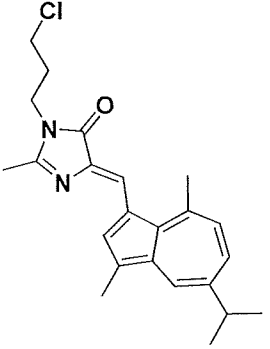
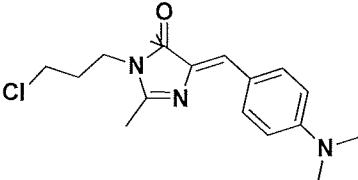
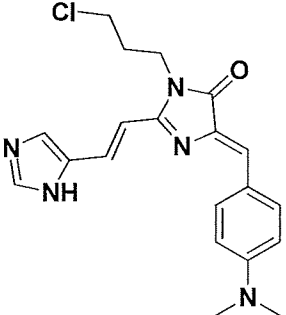
58	
59	
60	

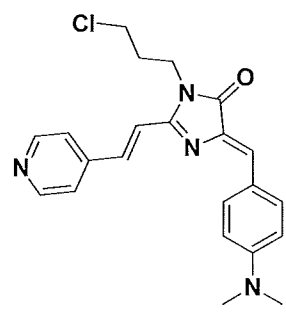
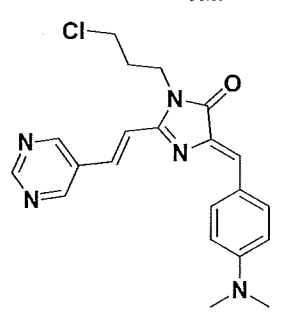
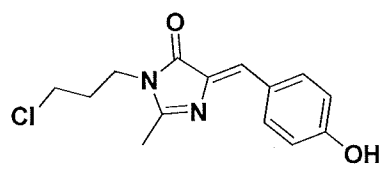
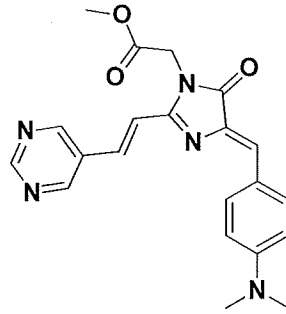
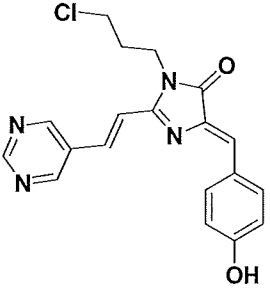
61	 <chem>ClCCCCN1C(=O)N=C(C1)/C=C/c2cc(OC)c(O)c(OC)c2</chem>
62	 <chem>CN(C)c1ccc(cc1)/C=C2C(=O)N(C)CCOCCOCCCCCl2</chem>
63	 <chem>ClCCCCN1C(=O)N=C(C1)/C=C/c2ccc(N(C)C)cc2</chem>
64	 <chem>ClCCCCN1C(=O)N=C(C1)/C=C/c2ccc(N(C)C)cc2</chem>
65	 <chem>ClCCCCN1C(=O)N=C(C1)/C=C/c2ccccc2O</chem>

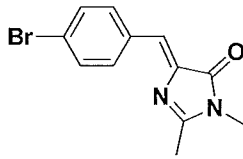
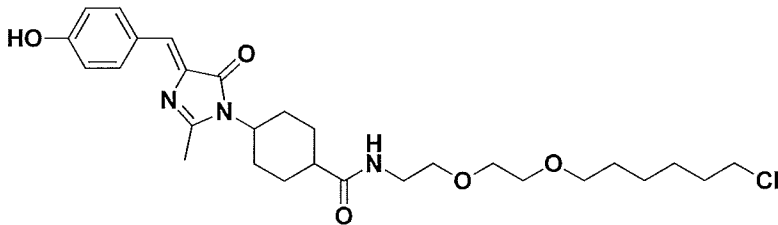
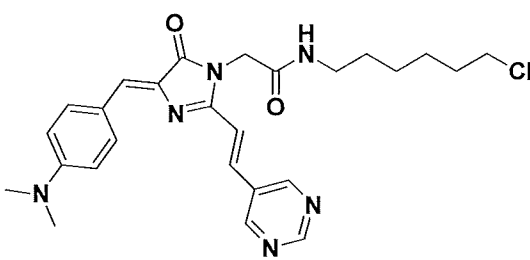
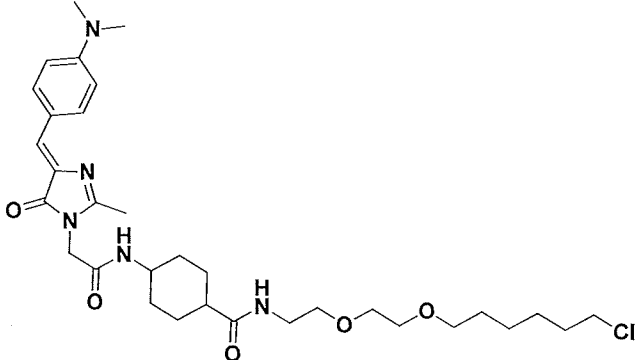
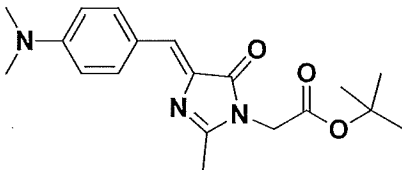
66	
67	
68	
69	
70	

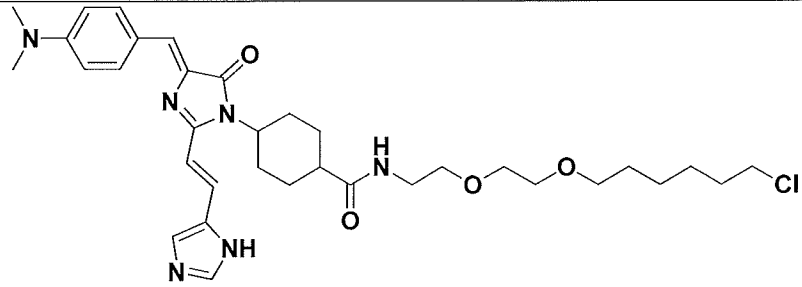
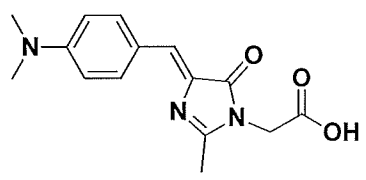
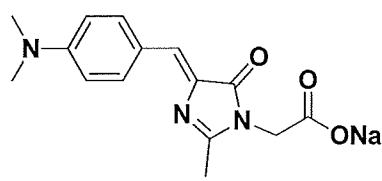
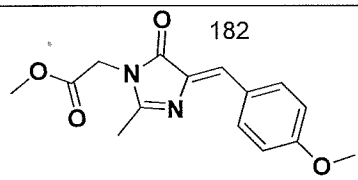
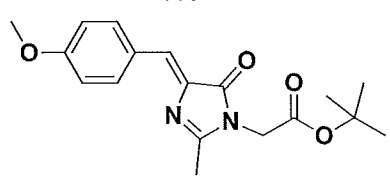
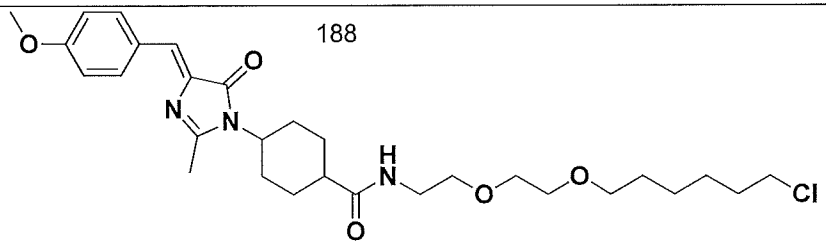
71	
72	
73	
74	
75	

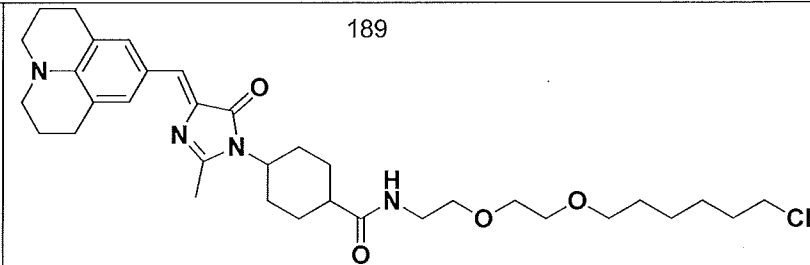
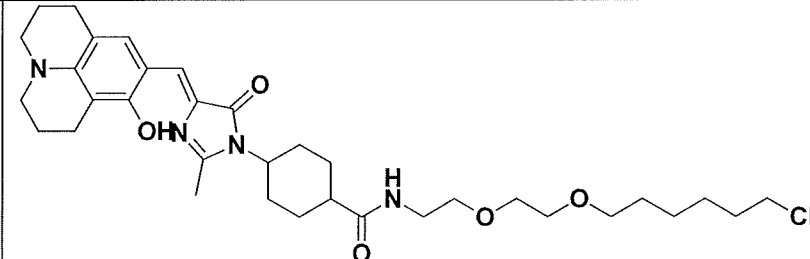
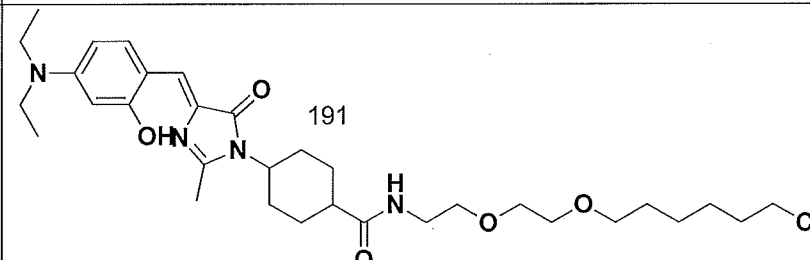
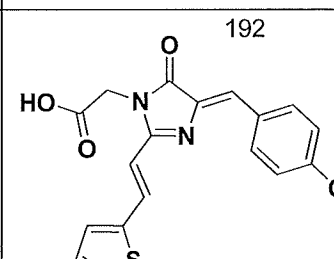
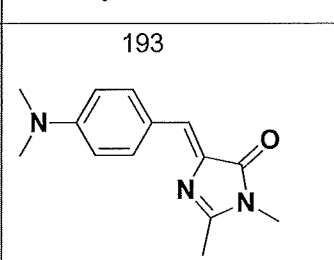
76	
77	
78	
79	

80	 <chem>CCN(C)c1ccc(cc1)/C=C/C2=CN(C(=O)N2CCCl)/C=C/c3ccsc3</chem>
81	 <chem>CC(C)c1cc(C)cc(C)/C=C/C2=CN(C(=O)N2CCCl)/C=C/c3c[nH]cn3C</chem>
82	 <chem>CCN(C)c1ccc(cc1)/C=C/C2=CN(C(=O)N2CCCl)/C=C/c3c[nH]cn3C</chem>
83	 <chem>CN1C=NC=C1/C=C/C2=CN(C(=O)N2CCCl)/C=C/c3ccc(N(C)C)cc3</chem>

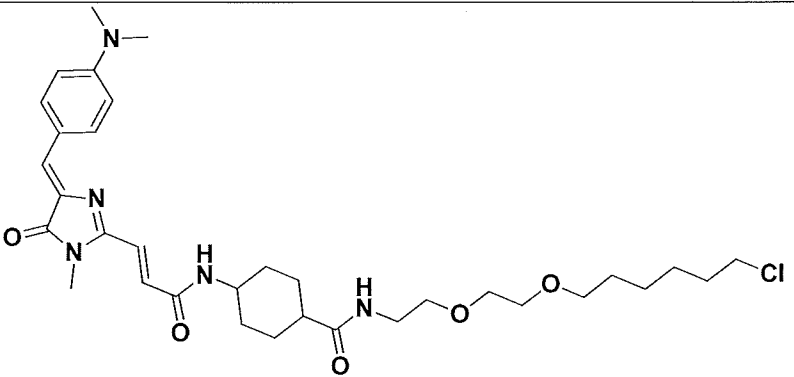
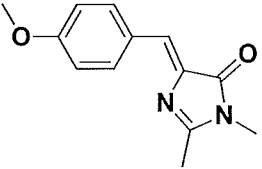
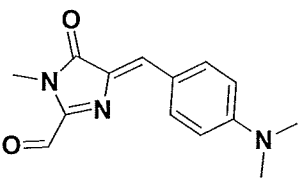
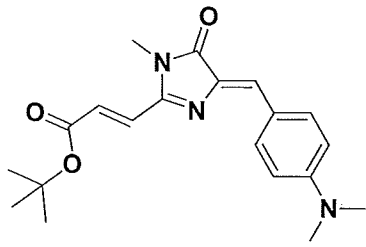
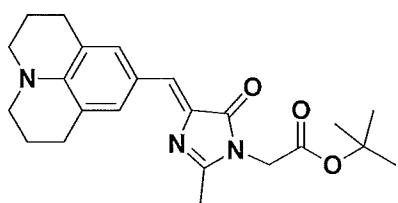
84	
85	
86	
87	
88	

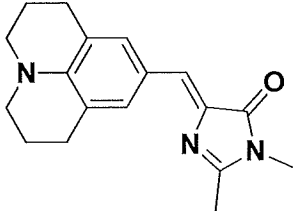
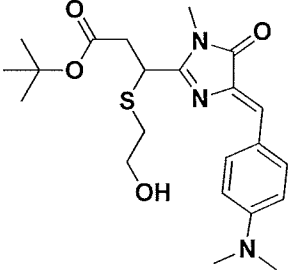
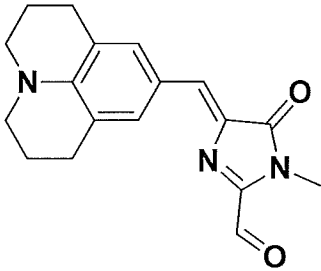
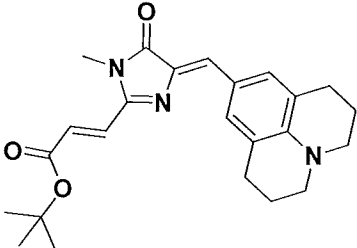
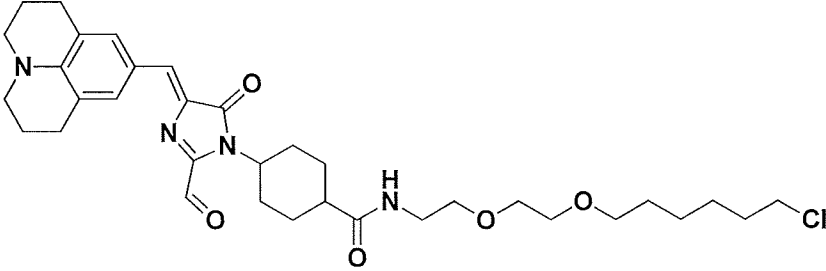
89	
90	
91	
92	
93	

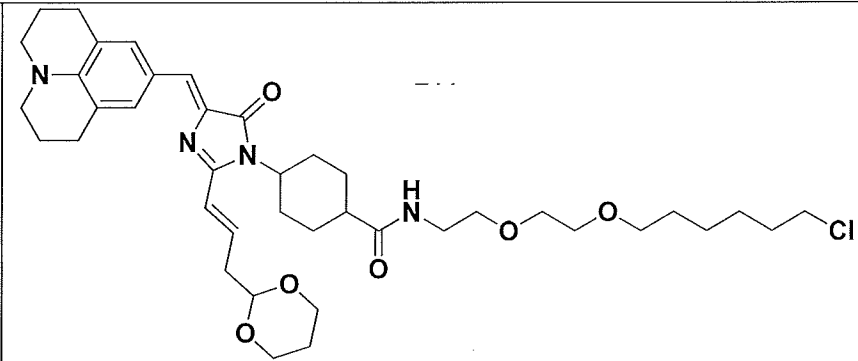
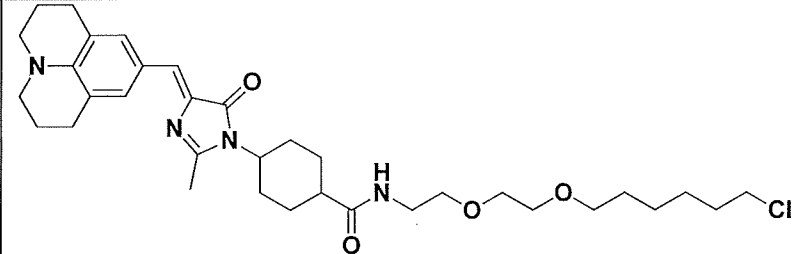
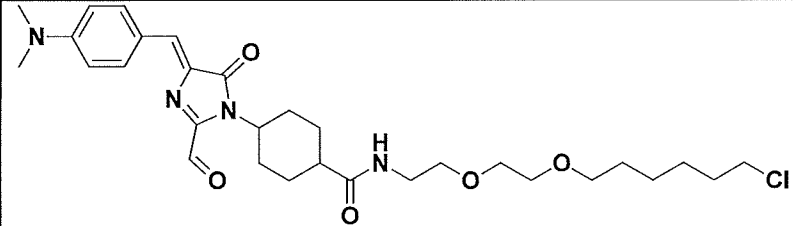
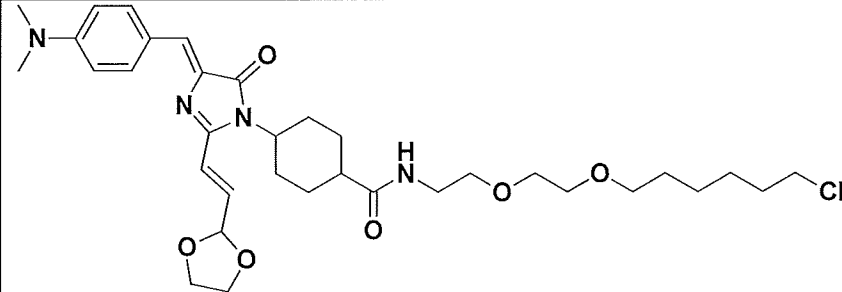
94	
95	
96	179 
97	182 
98	185 
99	188 
100	

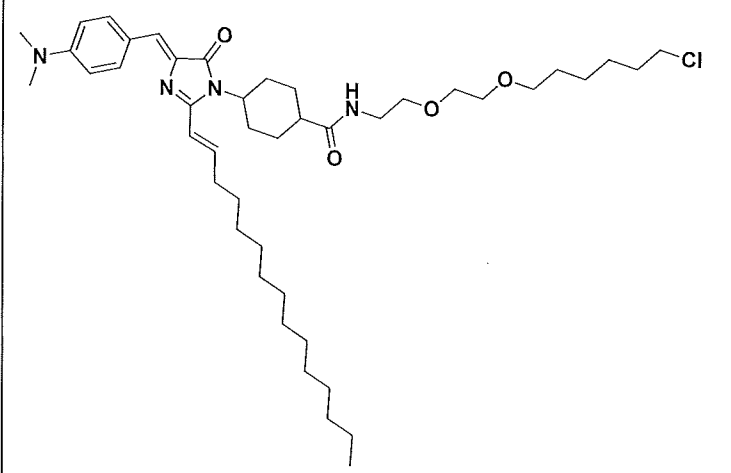
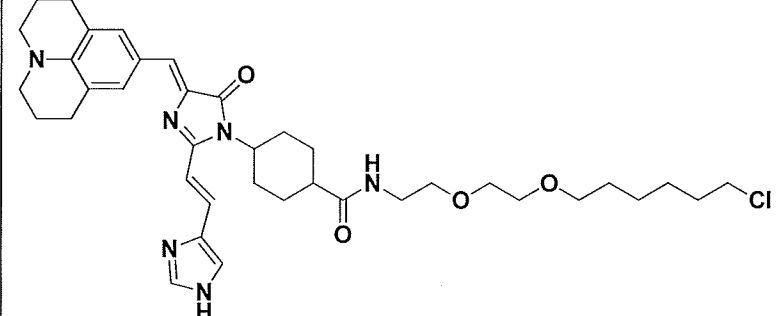
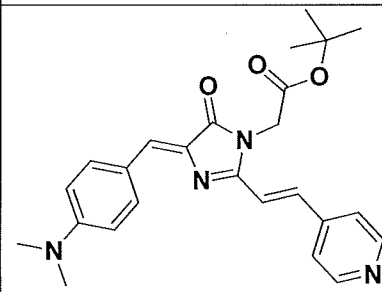
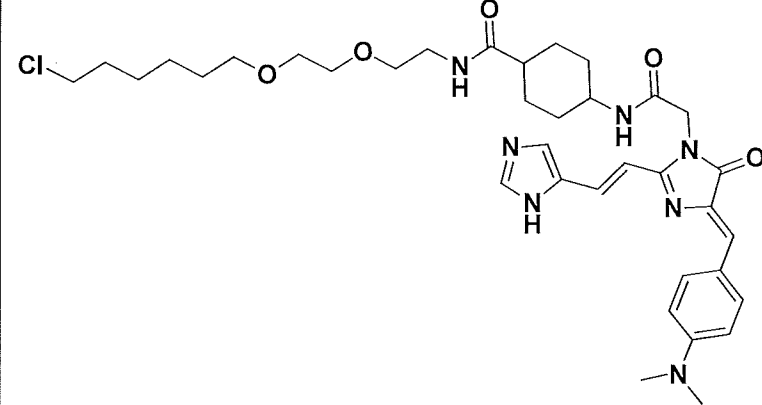
	189 
101	
102	191 
103	192 
104	193 

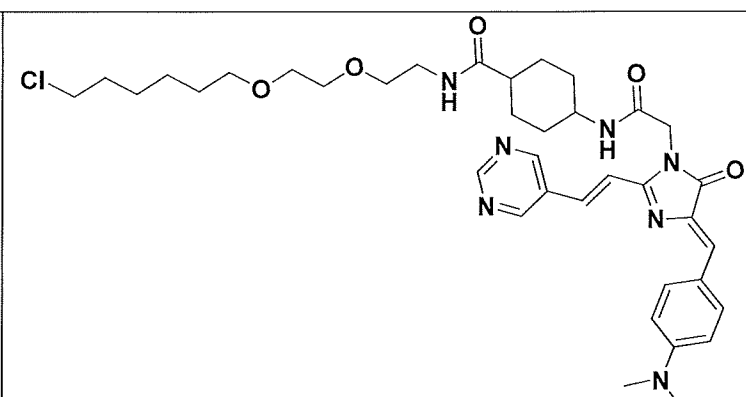
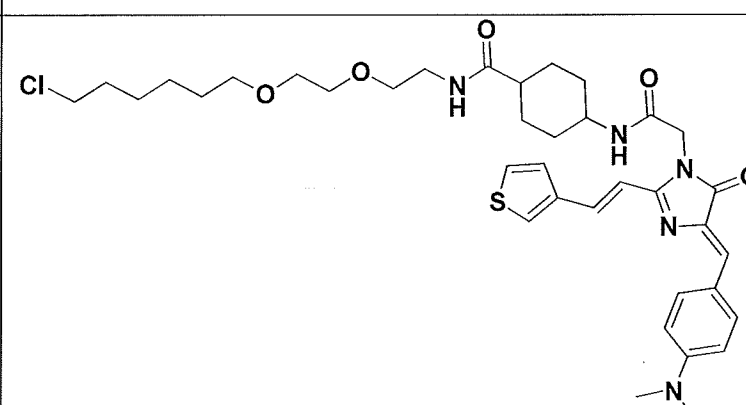
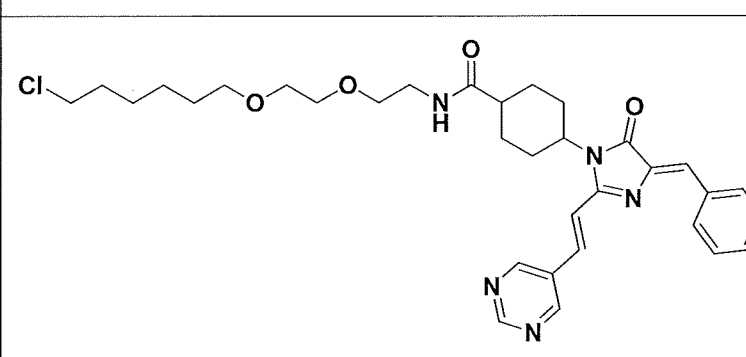
105	195 <chem>COc1ccc(C=C2N=C(NC3CCCCC3C(=O)NCCOCCOCCOCCCCl)C2=S)cc1</chem>
106	<chem>CN(C)c1ccc(C=C2N=C(C=O)N2C=O)cc1</chem>
107	<chem>CN(C)c1ccc(C=C2N=C(C(=O)OC(C)(C)C)N2C=O)cc1</chem>
108	<chem>CN(C)c1ccc(C=C2N=C(C(=O)O)N2C=O)cc1</chem>

109	
110	
111	
112	
113	

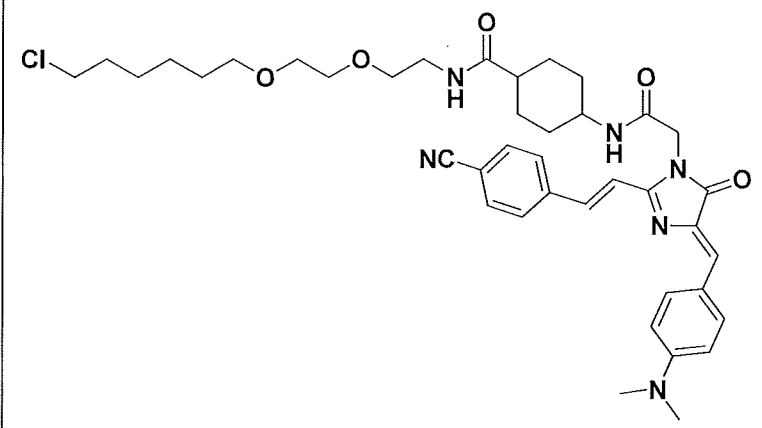
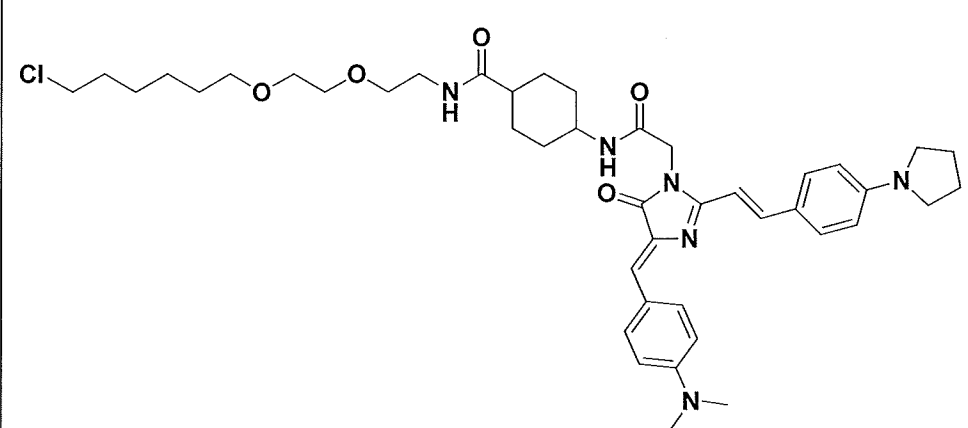
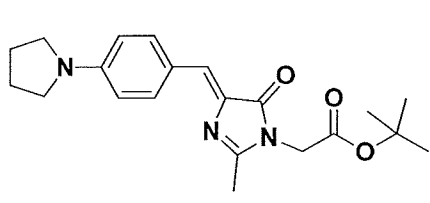
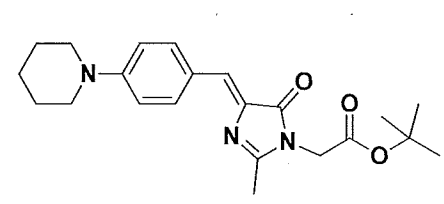
114	
115	
116	
117	
118	

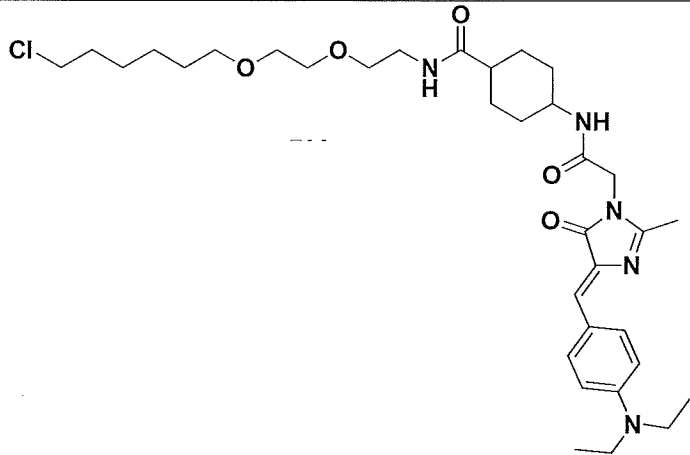
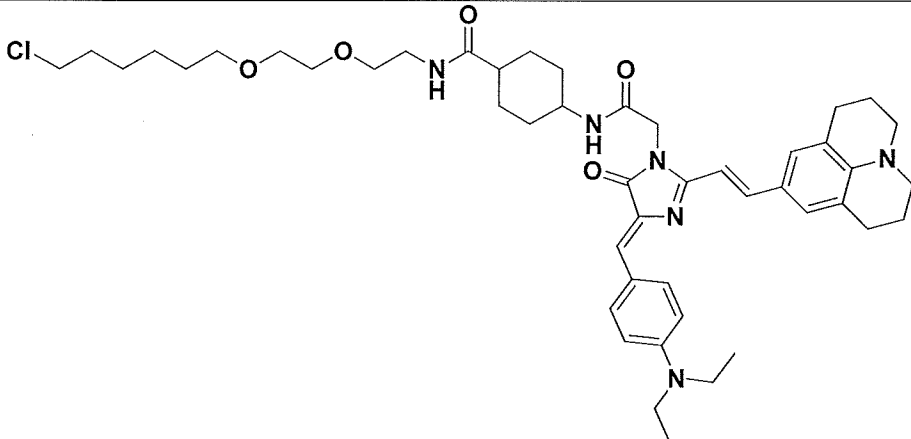
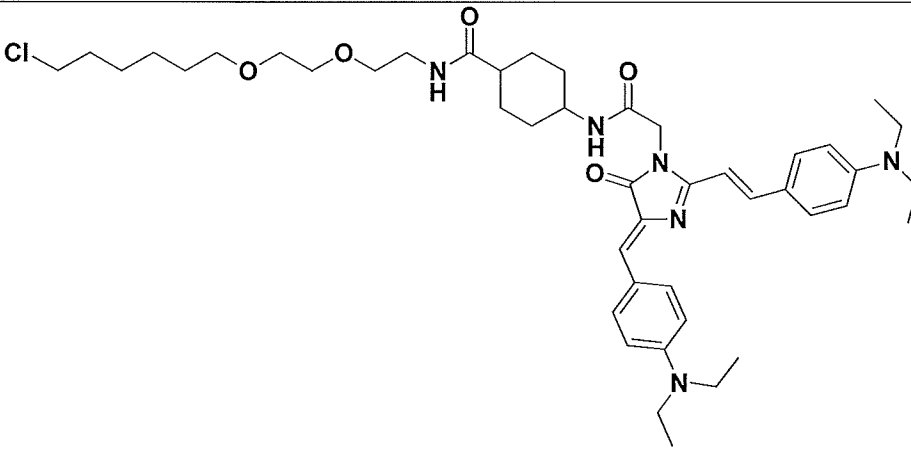
119	
120	
121	
122	

123	
124	
125	
126	

127	
128	
129	

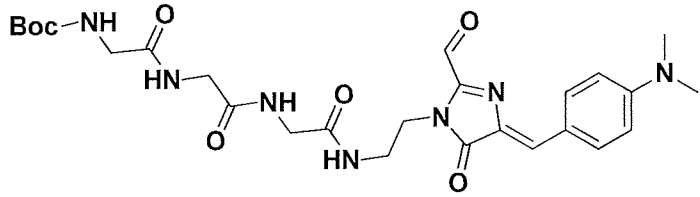
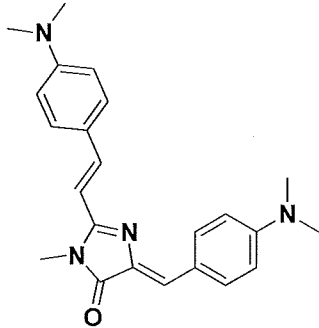
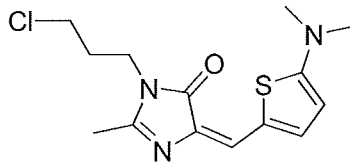
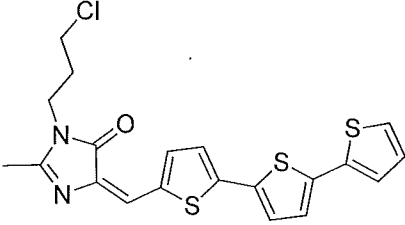
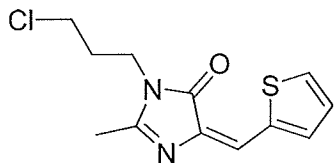
130	<chem>ClCCCCCCCCOCCOCCNC(=O)C1CCCCC1NC(=O)CN2C(=O)C(=C/C3=CC=C(C=C3)N(C)C)N2C(=O)C4=CC=C(C=C4)N(C)C</chem>
131	<chem>ClCCCCCCCCOCCOCCNC(=O)C1CCCCC1NC(=O)CN2C(=O)C(=C/C3=CC=C(C=C3)N(C)C)N2C(=O)C4=CC=C(C=C4)N(C)C</chem>
132	<chem>ClCCCCCCCCOCCOCCNC(=O)C1CCCCC1NC(=O)CN2C(=O)C(=C/C3=CC(OC)=C(OC)C=C3)N2C(=O)C4=CC=C(C=C4)N(C)C</chem>

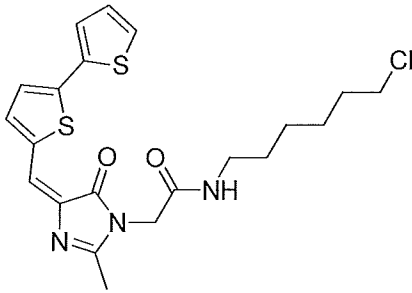
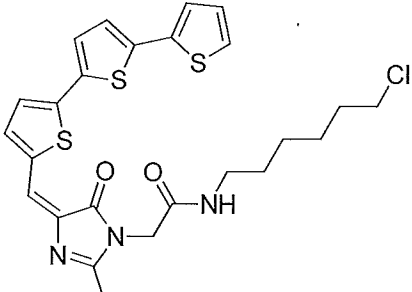
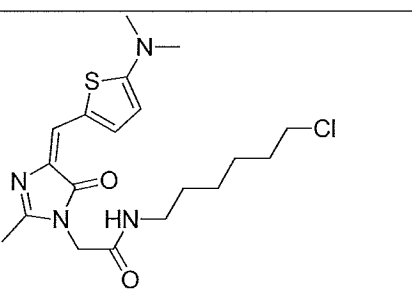
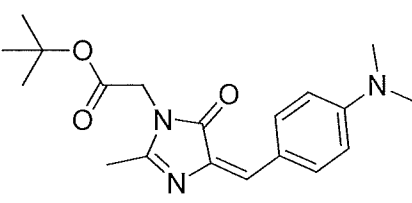
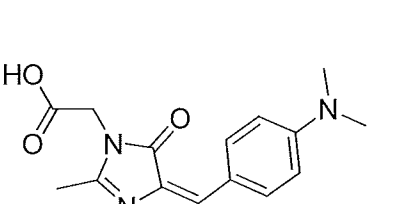
133	
134	
135	
136	

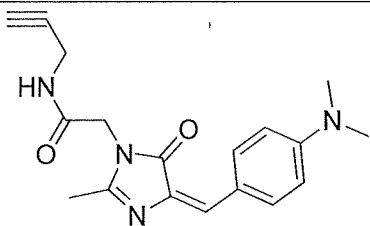
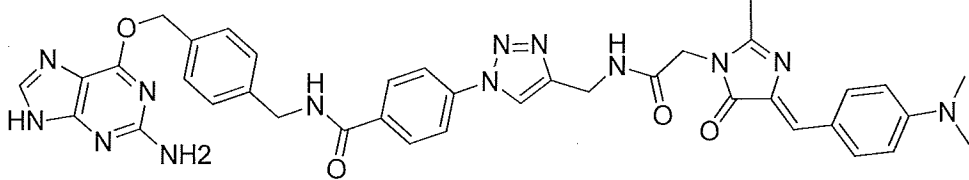
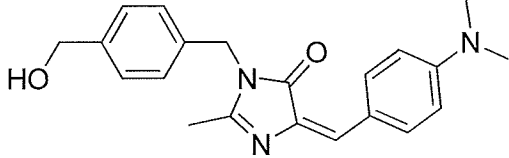
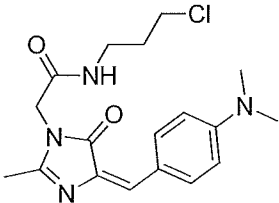
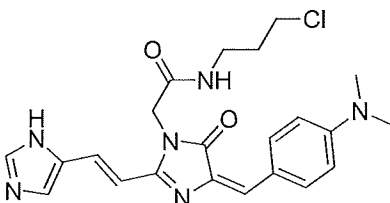
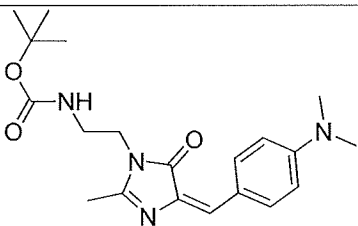
137	
138	
139	

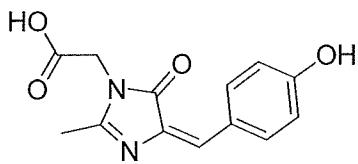
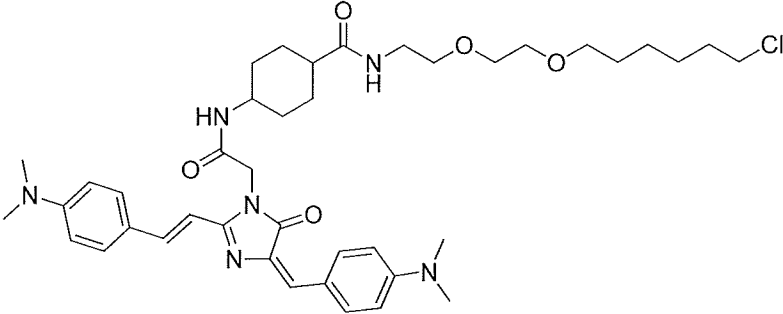
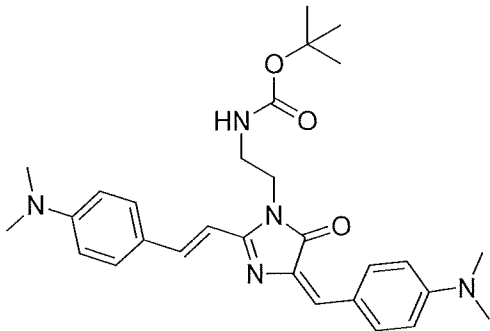
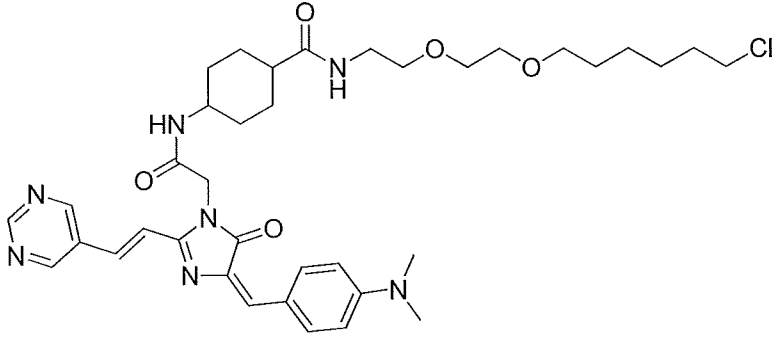
140	<chem>ClCCCCOCCOC(=O)Nc1ccccc1C(=O)NCC2=C(C)N(C(=O)N2)C=Cc3ccc(cc3)N4CCCC4</chem>
141	<chem>ClCCCCCOCCOC(=O)Nc1ccccc1C(=O)NCC2=C(C)N(C(=O)N2)C=Cc3ccc(cc3)N(C)C</chem>
142	<chem>ClCCCCCOCCOC(=O)Nc1ccccc1C(=O)NCC2=C(C)N(C(=O)N2)C=Cc3ccc(cc3)N(C)C</chem>

[illegible]

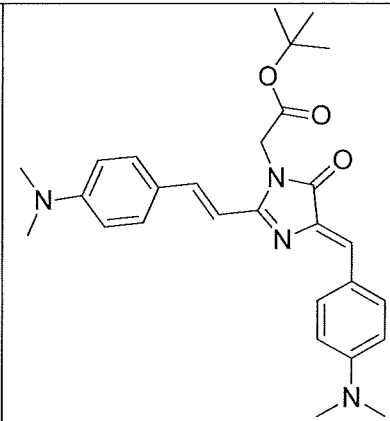
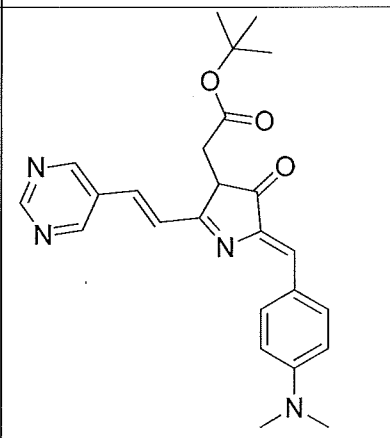
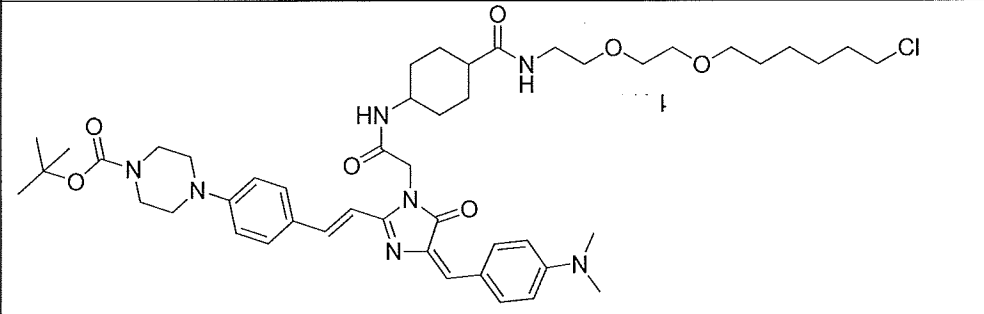
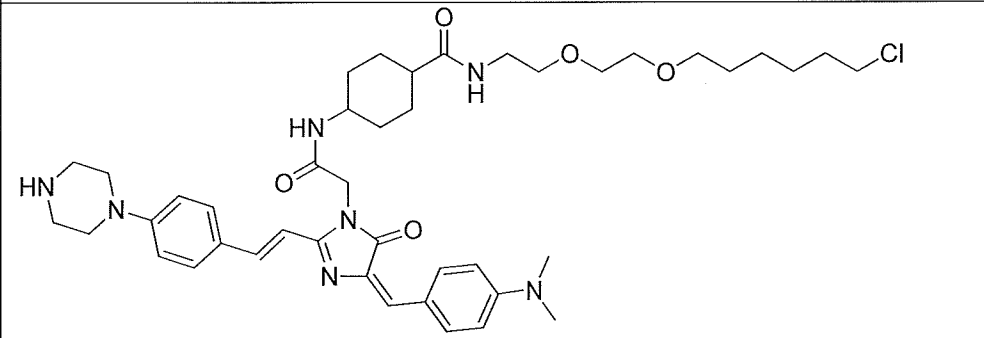
148	
149	
150	
151	
152	

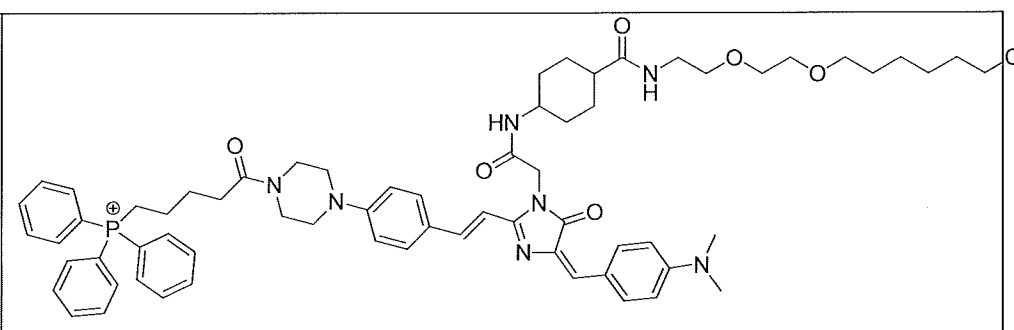
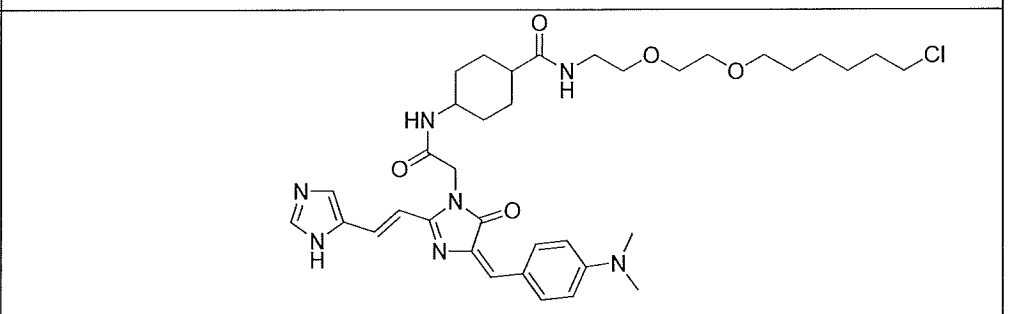
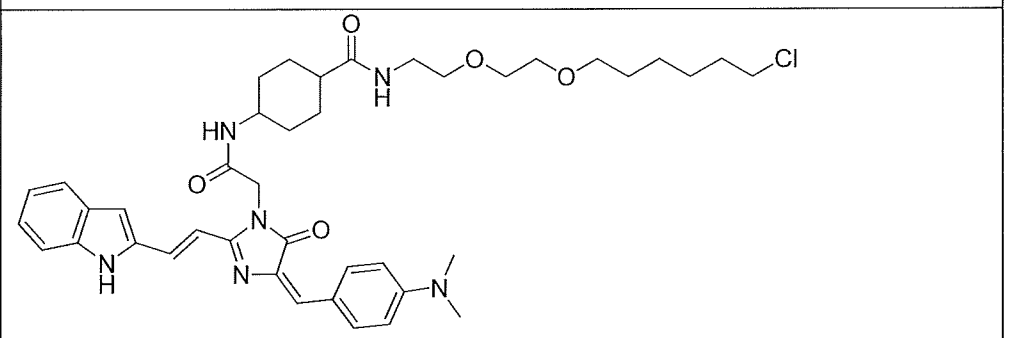
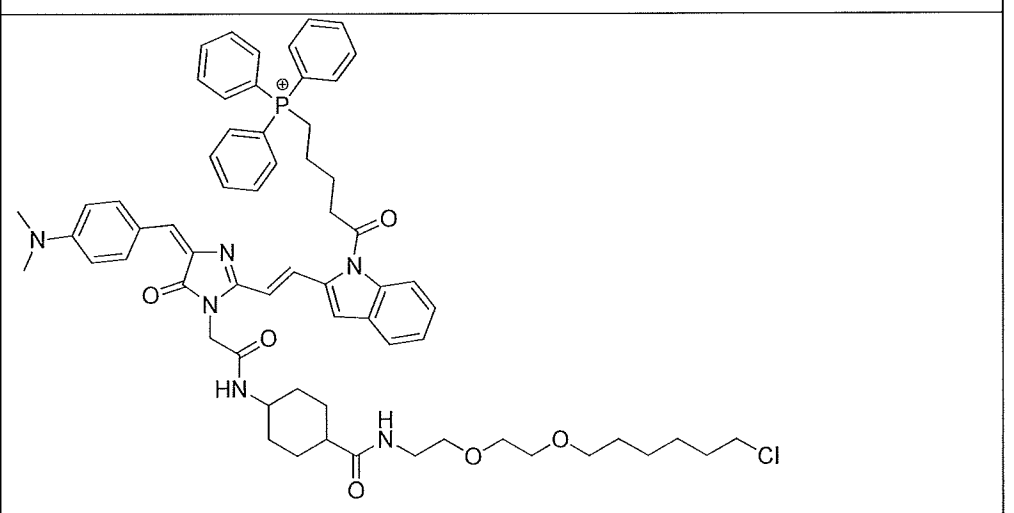
153	
154	
155	
156	
157	

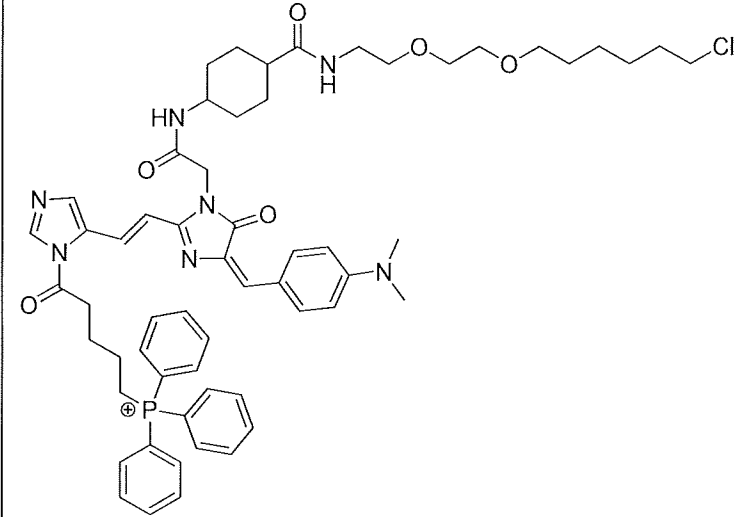
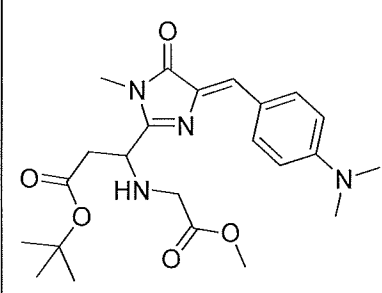
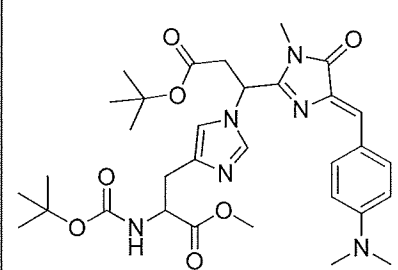
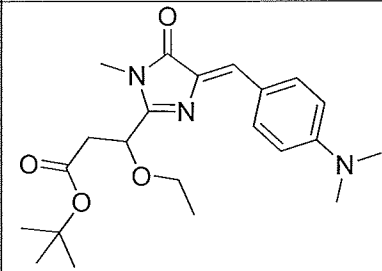
158	
159	
160	
161	
162	
163	

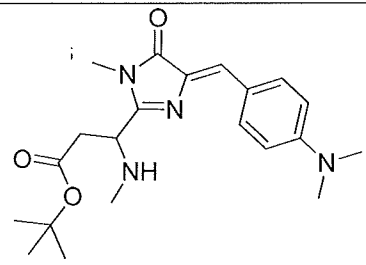
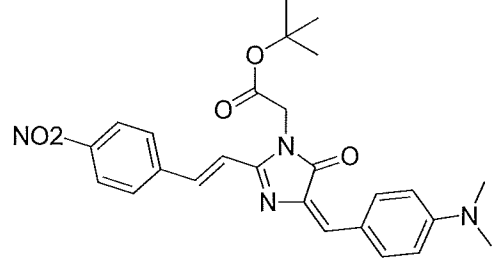
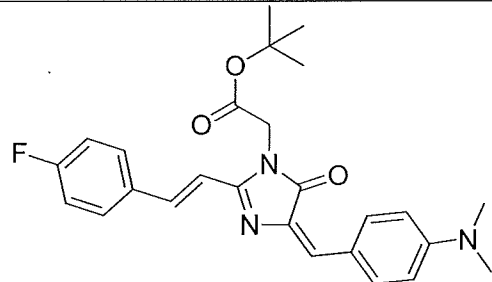
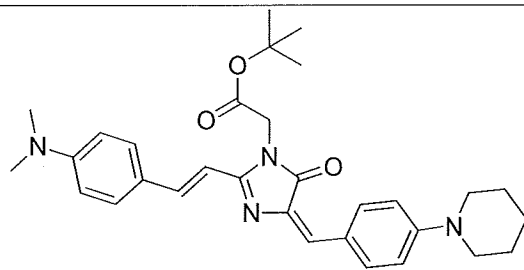
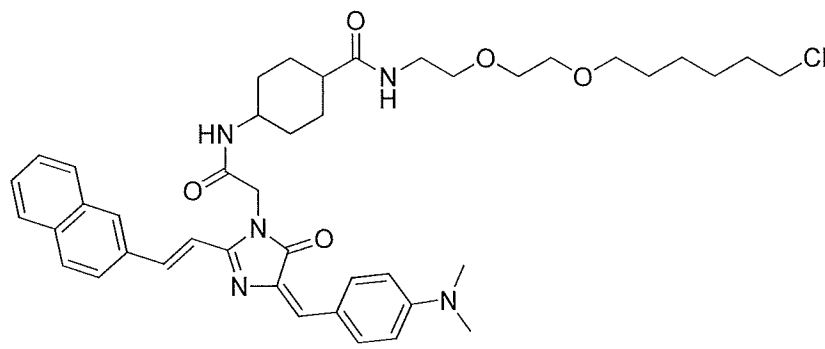
164	
165	
166	
167	

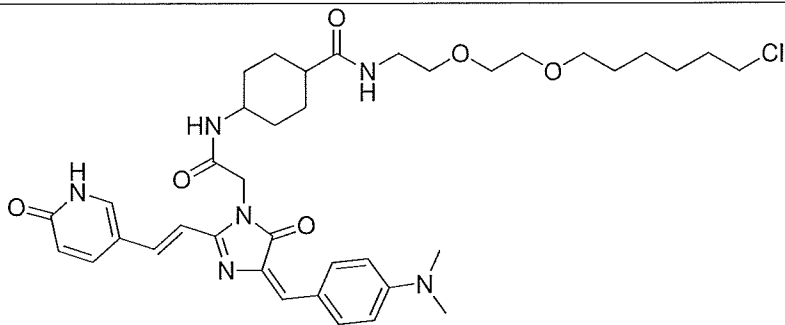
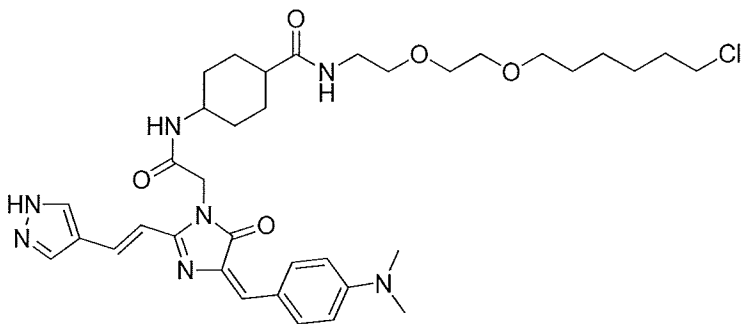
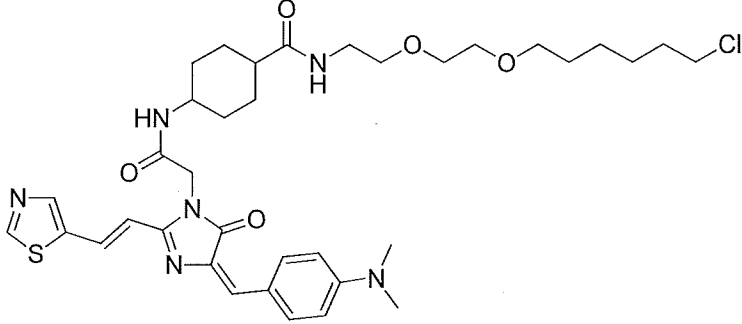
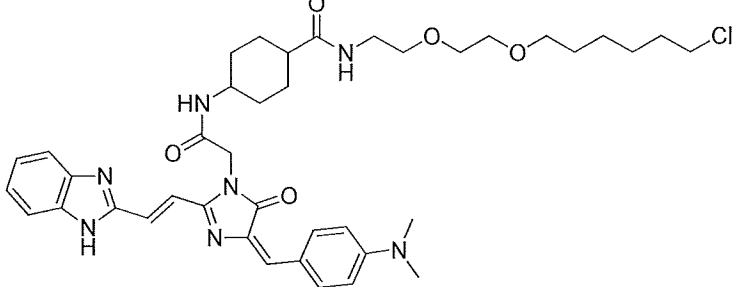
[illegible]

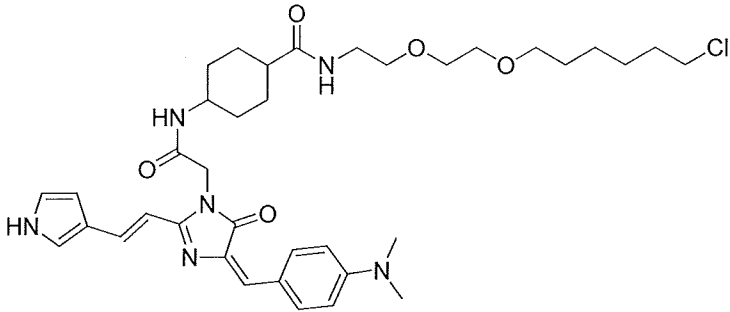
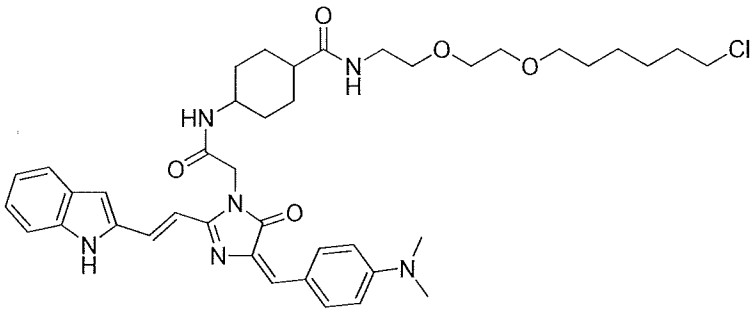
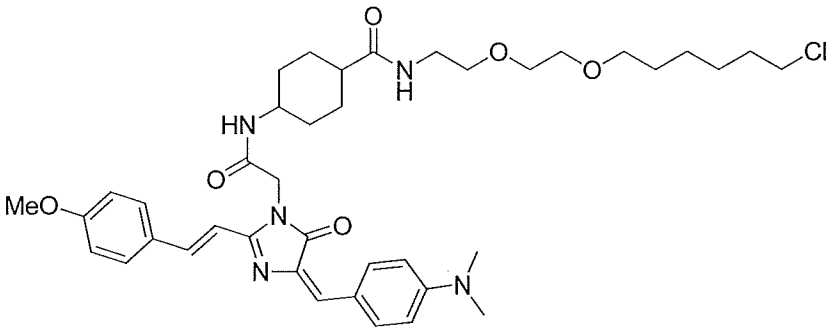
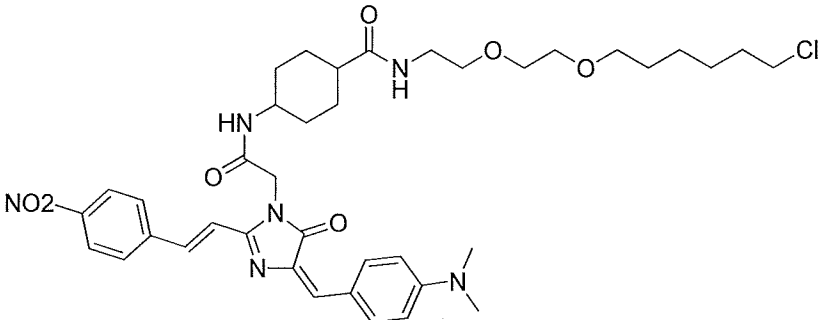
172	
173	
174	
175	

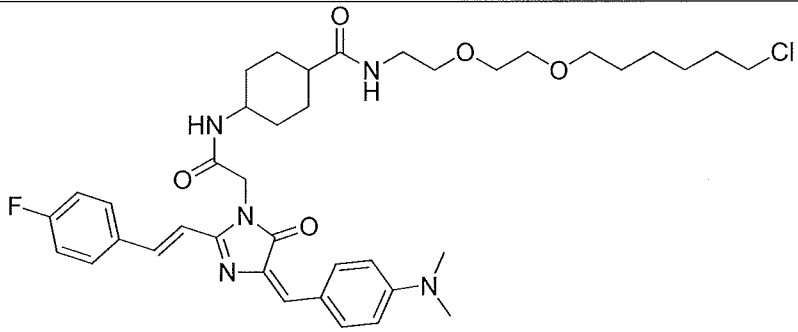
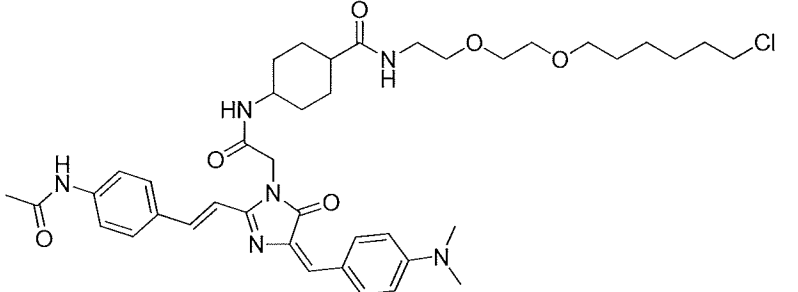
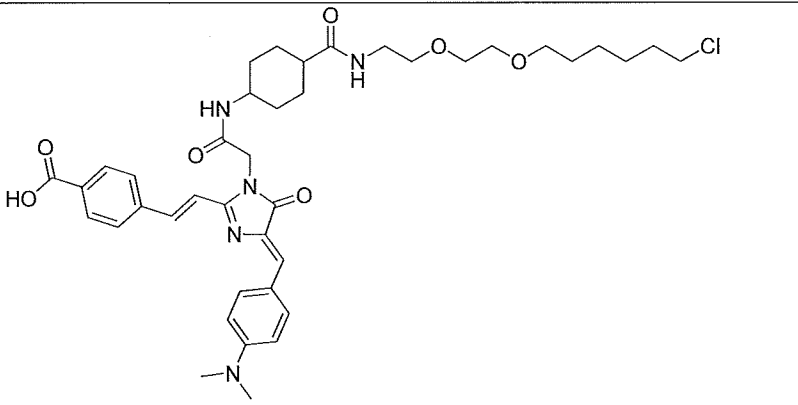
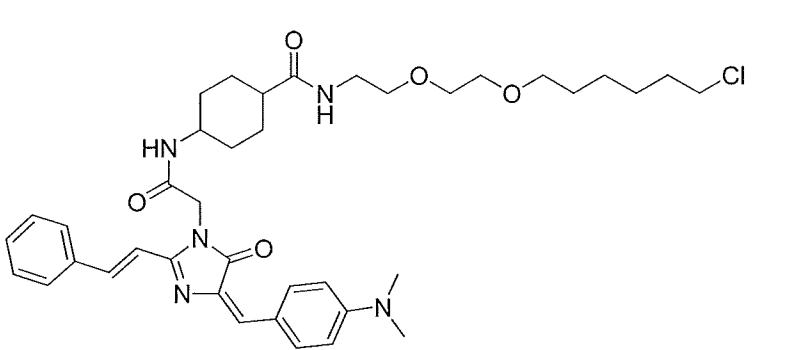
176	
177	
178	
179	

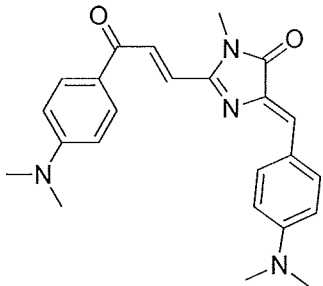
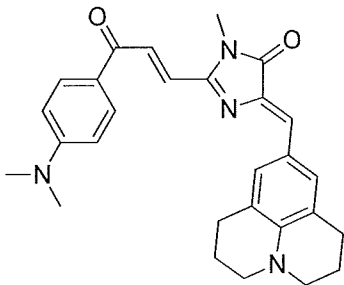
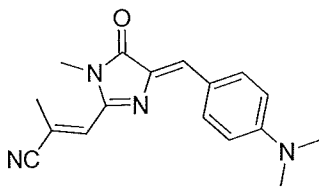
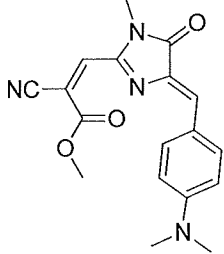
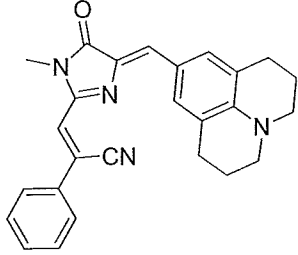
180	
181	
182	
183	

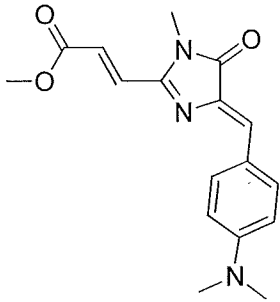
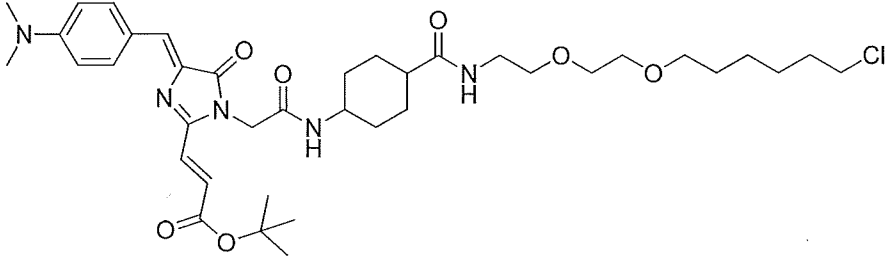
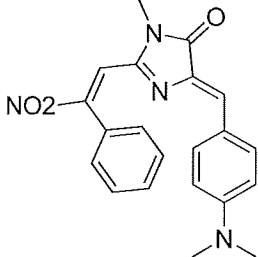
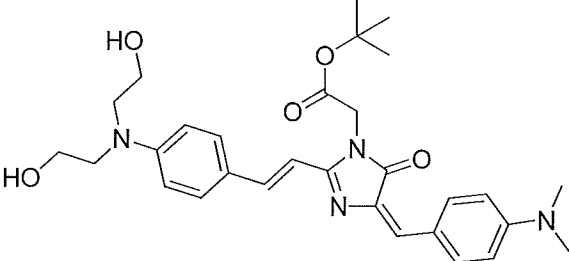
184	
185	
186	
187	
188	

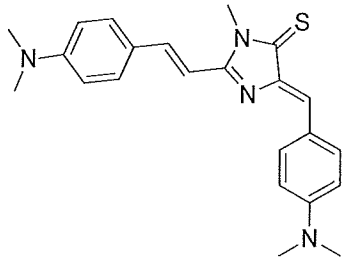
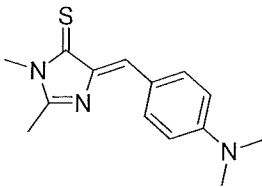
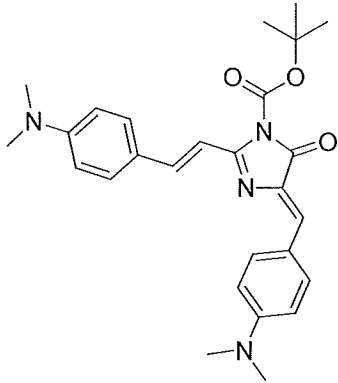
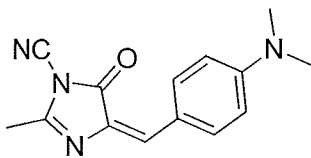
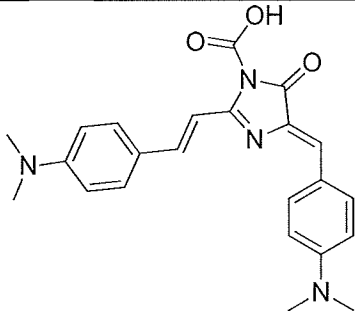
189	
190	
191	
192	

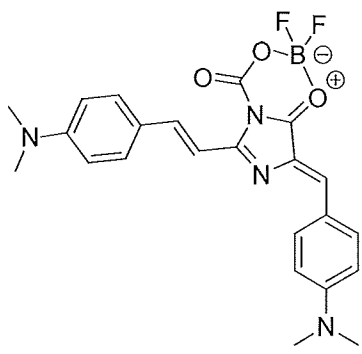
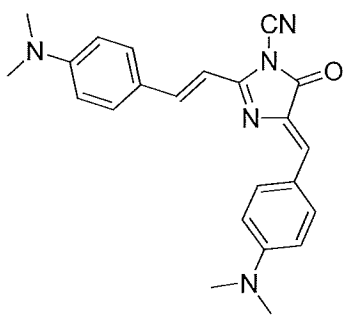
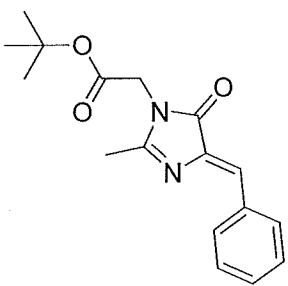
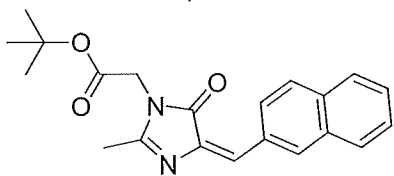
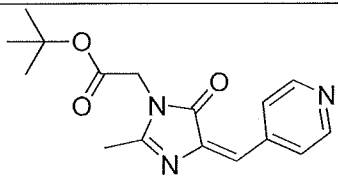
193	
194	
195	
196	

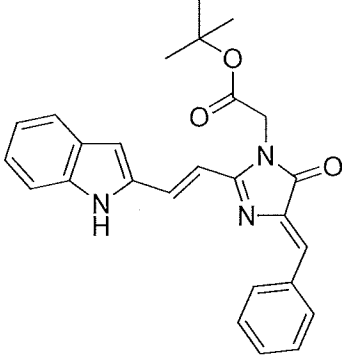
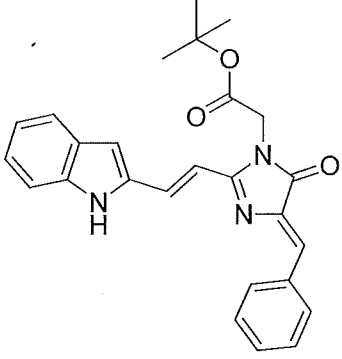
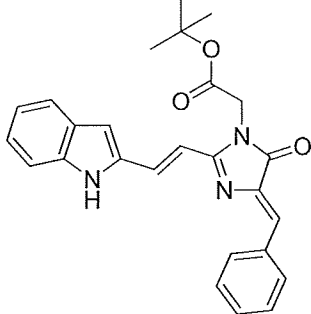
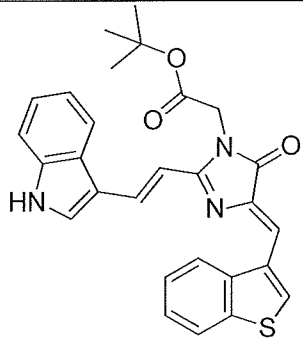
197	
198	
199	
200	

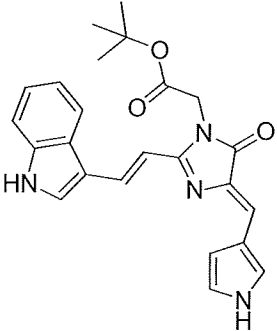
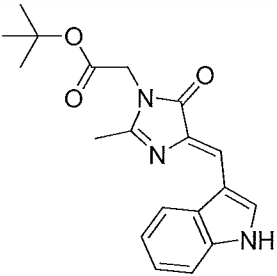
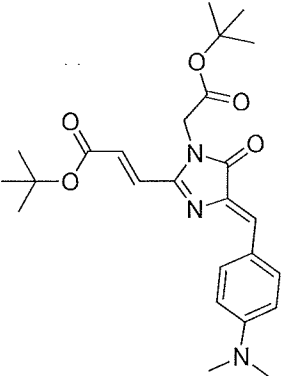
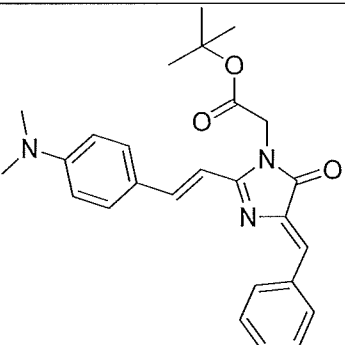
201	
202	
203	
204	
205	

206	
207	
208	
209	

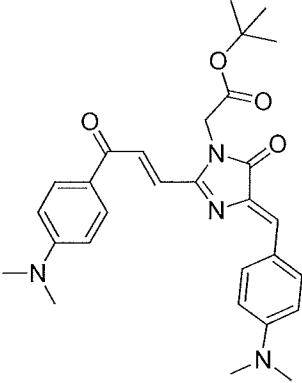
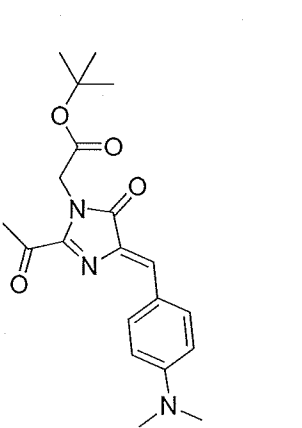
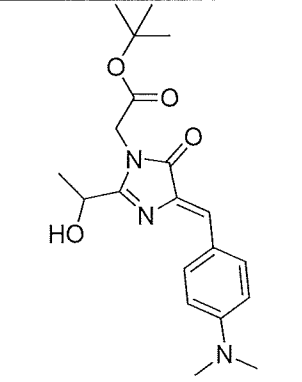
210	
211	
212	
213	
214	

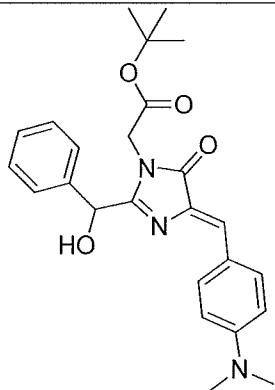
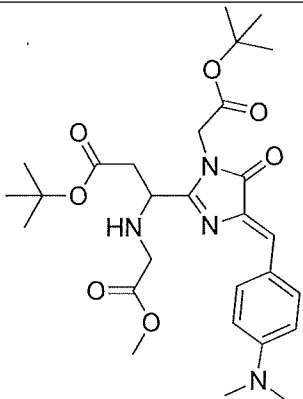
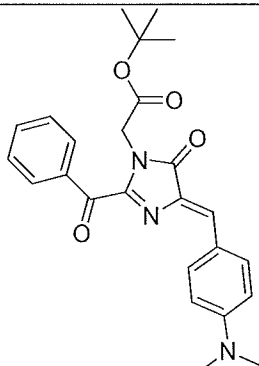
215	 <chem>CN(C)c1ccc(/C=C/c2nn(c(=O)OB(F)(F)F)c2/C=C/c3ccc(N(C)C)cc3)cc1</chem>
216	 <chem>CN(C)c1ccc(/C=C/c2nn(c(=O)N#C)c2/C=C/c3ccc(N(C)C)cc3)cc1</chem>
217	 <chem>CC1=CN(C(=O)OCC(C)(C)C)C(=O)C1/C=C/c2ccccc2</chem>
218	 <chem>CC1=CN(C(=O)OCC(C)(C)C)C(=O)C1/C=C/c2ccc3ccccc3c2</chem>
219	 <chem>CC1=CN(C(=O)OCC(C)(C)C)C(=O)C1/C=C/c2ccncc2</chem>

220	 <chem>CC(C)(C)OC(=O)CN1C(=O)N=C(C=Cc2ccccc2)N1/C=C/c3c[nH]c4ccccc34</chem>
221	 <chem>CC(C)(C)OC(=O)CN1C(=O)N=C(C=Cc2ccccc2)N1/C=C/c3c[nH]c4ccccc34</chem>
222	 <chem>CC(C)(C)OC(=O)CN1C(=O)N=C(C=Cc2ccccc2)N1/C=C/c3c[nH]c4ccccc34</chem>
223	 <chem>CC(C)(C)OC(=O)CN1C(=O)N=C(C=Cc2c3ccccc3sc2)N1/C=C/c4c[nH]c5ccccc45</chem>

224	
225	
226	
227	

[illegible]

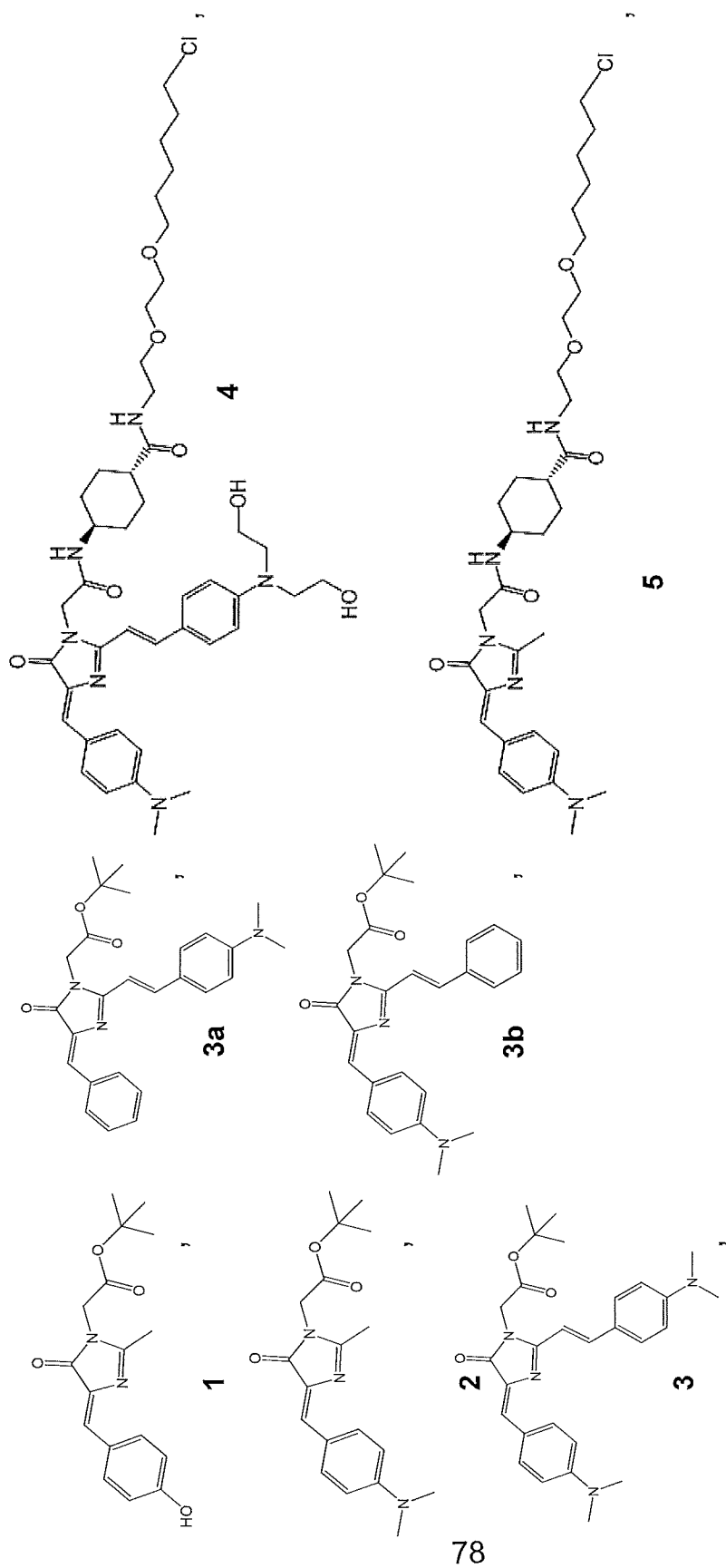
232	 <chem>CC1=CC=C(C=C1)/C=C/C2=CN(C2C(=O)OCC(C)(C)C)C(=O)/C=C/C3=CC=C(C=C3)N(C)C</chem>
233	 <chem>CC(=O)C1=CN(C1C(=O)OCC(C)(C)C)C(=O)/C=C/C2=CC=C(C=C2)N(C)C</chem>
234	 <chem>CC(O)C1=CN(C1C(=O)OCC(C)(C)C)C(=O)/C=C/C2=CC=C(C=C2)N(C)C</chem>

235	
236	
237	

[0080] In further embodiments as reported herein, protein aggregation and protein misfolding are detected by the disclosed compounds in test tubes. This detection can be conducted with or without covalent conjugation to proteins of interest. Normally, purified proteins are used in this detection. These proteins are subjected to in vitro conditions to induce protein misfolding and aggregation. Disclosed compounds can be added before, during, or after protein misfolding and

aggregation. Fluorescence intensity can be recorded by fluorescence spectrophotometers or fluorescence microplate readers.

Compounds **1-5**:



[0081] TESTING EXAMPLES

[0082] Compound **1**, which contains HBI, requires both deprotonation to phenolate and restriction of bond rotations to emit fluorescence (Figure 2a). Because the local pH inside misfolded and aggregated proteins is hardly basic enough to deprotonate phenol ($pK_a = 8-10$), the phenol in HBI may be substituted with an electron donating dimethylamino group, affording Compound **2** as reported above.

[0083] Both **1** and **2** exhibit elevated fluorescence in solvents containing increasing concentrations of glycerol (spectra shown in Figures 6a-b, 7a-b, 8a-b, and 9), which mimics the microenvironment in protein aggregates to inhibit TICT. However, the quantum yield (QY, $\Phi = 0.001$ for **1** and $\Phi = 0.027$ for **2** in glycerol, Table 2) was much lower than that of a GFP ($\Phi = 0.79$).

[0084] Further embodiments may extend π conjugation to both increase QY and restrict bond rotation (Figure 2a). Such an embodiment is shown as Compound **3**, reported above.

[0085] Compound **3** exhibited higher QY and brightness than **1** and **2** in glycerol ($\Phi = 0.22$, comparable to Φ of the Kaede protein as 0.33; brightness shown in Figure 2b; spectra shown in Figures 6c, 7c, 8c, and 9; biophysical properties shown in Table 2).

[0086] Crystals of **3** exhibited fluorescence. While not wishing to be bound by theory, this suggests its aggregation induced emission (AIE) feature (Figures 2c and 10a). The fluorescence may arise from the tight crystal packing that inhibits TICT of **3** (Figure 10b).

[0087] In the fluorescent crystal, **3** adopted a near planar structure (Figure 2d). A time-dependent density functional theory calculation based on this structure (Figures 11a-c) showed that one dimethylamino benzyl played a major role in donating electrons at the S1 excited state (position A in Figure 11d).

[0088] Compound **3a** and Compound **3b** were also prepared, as shown above.

[0089] Compound **3a** without electron-donating capacity at position A diminished quantum yield by ~50%; whereas **3b** without electron-donating capacity at position B does not affect quantum yield (Table 2). The absorptivity of both **3a** and **3b** was reduced (Table 2 and Figure 9), yielding ~5-fold decrease in its brightness (Figure 2b and Table 2). Electron-donating groups in both A and B positions appeared useful for high quantum yield and brightness.

Table 2: The photophysical parameters of fluorophores in the present invention.

Compound	λ_{ex} (nm)	λ_{em} (nm)	Φ	λ_{abs} (nm) ^c	$\epsilon(\text{M}^{-1}\text{cm}^{-1})^c$	$B(\text{M}^{-1}\text{cm}^{-1})^d$
1	370	429	0.001	376	20,592	20.6
2	450	521	0.027	452	23,718	640
3	530	629	0.221	530	39,049	8,786
3a	485	621	0.1	500	13,863	1,386
3b	530	630	0.232	533	9,004	2,089
4	530	628	0.230	534	36,241	8,734

^a Maximum absorbance and extinction coefficient measured in glycerol. ^b Brightness (B) was calculated using $B = \Phi\epsilon$.

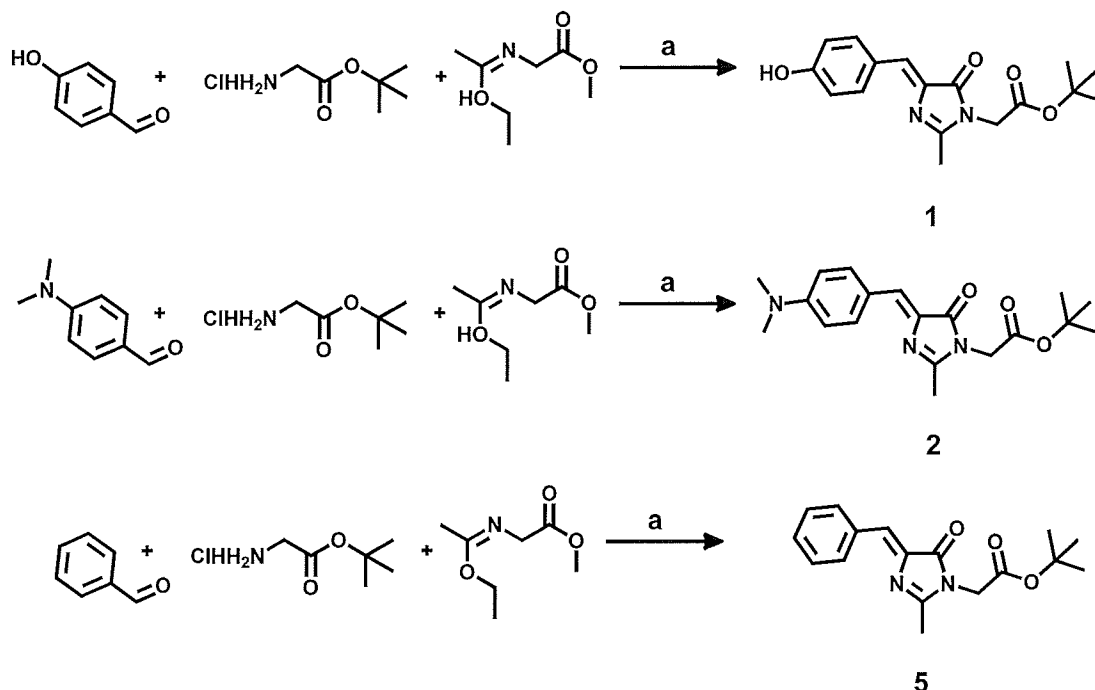
Table 3: The geometry of 3 extracted from crystal structure. The total energy at this geometry is -1531.67109379 Hartree.

Atom	X(Å)	Y(Å)	Z(Å)
C	8.589111	-1.881432	1.224387
H	9.139859	-1.768830	0.445329
H	8.998082	-2.518740	1.813209
H	8.497810	-1.039027	1.676078
H	7.677154	-4.210787	1.409144
C	1.255841	2.555928	-1.372002
C	2.079860	1.358102	-1.047783
C	0.014408	0.720154	-0.976268
C	3.430059	1.410384	-0.977822
H	3.804300	2.224322	-1.228115
C	4.388472	0.388601	-0.571549
C	5.727878	0.741428	-0.476613
H	5.985943	1.603805	-0.707828
C	6.346334	-1.447938	0.341371
C	4.994563	-1.813307	0.214286
H	4.734045	-2.680516	0.424980
C	4.042570	-0.906792	-0.220101

H	3.153915	-1.172817	-0.278479
C	6.889460	-3.694201	1.224295
H	6.345074	-3.645782	2.012474
H	6.392083	-4.110123	0.517262
C	-1.221120	-0.053142	-0.842227
H	-2.028992	0.350472	-1.064420
C	-1.230528	-1.321341	-0.411810
H	-0.405583	-1.688712	-0.190991
C	-2.393740	-2.195524	-0.248500
C	-2.222585	-3.502827	0.192167
H	-1.365663	-3.791036	0.412277
C	-3.271420	-4.390282	0.317704
H	-3.108393	-5.258111	0.608236
C	-4.572196	-3.996638	0.008796
C	-4.759688	-2.683641	-0.436285
H	-5.616644	-2.396456	-0.654614
C	-3.698918	-1.799250	-0.557339
H	-3.857202	-0.929174	-0.848739
H	-4.939809	-6.694154	-0.329524
C	-1.183415	2.949098	-1.368345
H	-1.934115	2.471932	-1.754659
H	-0.988762	3.709282	-1.938351
C	-1.562630	3.441717	0.009391
C	-3.264926	4.838528	1.141641
C	-2.270442	5.770712	1.793963
H	-1.500405	5.271937	2.076998
H	-2.677748	6.188832	2.556554
H	-2.003097	6.444013	1.164576
C	-3.736996	3.730646	2.074251
H	-4.412066	3.210115	1.635228
H	-4.099555	4.118614	2.875078

H	-2.995042	3.164813	2.301043
C	-4.455907	5.578691	0.551176
H	-4.145729	6.238137	-0.073121
H	-4.944598	6.008265	1.255781
H	-5.027917	4.953304	0.100110
N	1.227433	0.271319	-0.822233
N	-0.043023	2.084575	-1.330079
N	7.273613	-2.362966	0.822083
N	-5.636442	-4.884588	0.135566
O	1.575820	3.732509	-1.579503
O	-2.646351	4.215190	-0.068768
O	-0.970195	3.171377	1.011101
C	-6.986251	-4.455920	-0.216546
H	-7.206796	-3.659886	0.273176
H	-7.609525	-5.149602	0.002538
H	-7.028281	-4.273958	-1.158477
C	-5.403601	-6.297338	0.409664
H	-4.873618	-6.386220	1.205418
H	-6.246057	-6.740703	0.535678
C	6.689652	-0.155705	-0.046086
H	7.579008	0.113535	-0.016736

[0090] SYNTHESIS OF EXAMPLE COMPOUNDS



[0091] Scheme 1. Synthesis of GFP mimic precursors.

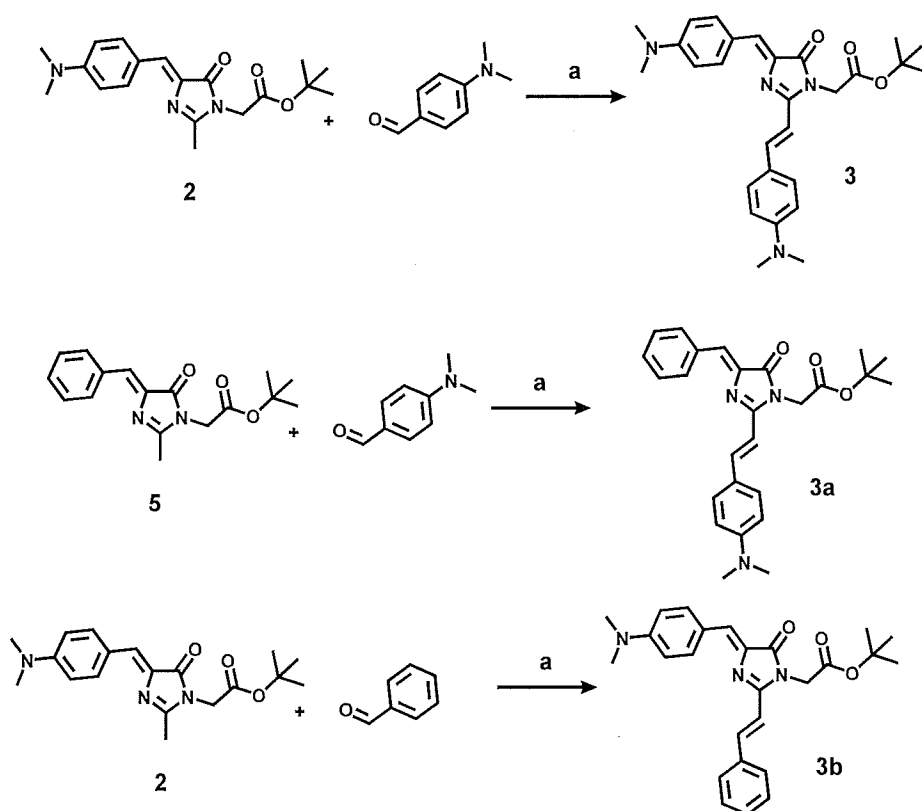
Cyclization reaction to form GFP core. Condition (a) glycine *tert*-butyl ester hydrochloride (1.1 eq) was combine with NaOH (1 eq) in EtOH and stirred for 1 hour at room temperature, aldehyde (1 eq) was added and stirred overnight, imidate (1 eq) was prepared and added in one portion.⁵ The reaction was stirred overnight, was then quenched by water and extracted with DCM. The organic fraction was collected and dried in vacuo. Compounds were further purified by flash chromatography (50% Ethyl Acetate, 50% Hexanes) to yield **1/2/5**.

[0092] 1: (Z)-*tert*-butyl 2-(4-(4-(hydroxy)benzylidene)-2-methyl-5-oxo-4,5-dihydro-1H-imidazol-1-yl)acetate. Crystalline orange solid. ¹H NMR (500 MHz, Chloroform-*d*) δ 8.11 (d, *J* = 8.9, Hz, 2H), 6.95 (s, 1H), 6.85 (d, *J* = 8.9 Hz, 2H), 4.40 (m, 3H), 2.27 (s, 3H), 1.44 (s, 9H). ¹³C NMR (126 MHz, Chloroform-*d*) δ 167.75, 161.59, 160.23, 135.95, 134.76, 126.91, 125.62, 116.27, 82.60, 42.32, 28.07, 15.53. [M+H]⁺: Calcd, 317.3570, Obsd, 317.1495

[0093] 2: (Z)-*tert*-butyl 2-(4-(4-(dimethylamino)benzylidene)-2-methyl-5-oxo-4,5-dihydro-1H-imidazol-1-yl)acetate. Crystalline orange solid. ¹H NMR (500 MHz, Chloroform-*d*) δ 8.07 (d, *J* = 8.9 Hz, 2H), 7.11 (s, 1H), 6.70 (d, *J* = 8.8 Hz, 2H), 4.29 (s, 2H), 3.05 (s, 6H), 2.31 (s, 3H), 1.48 (s, 9H). ¹³C NMR (126 MHz, Chloroform-*d*) δ 170.11, 166.90, 158.03, 151.56, 134.27,

134.22, 129.41, 122.19, 111.72, 82.83, 42.15, 40.04, 27.99, 15.43. [M+H]⁺ : Calcd, 344.1896, Obsd, 344.1983

[0094] **5**: (Z)-tert-butyl 2-(4-benzylidene-2-methyl-5-oxo-4,5-dihydro-1H-imidazol-1-yl)acetate. Crystalline yellow solid. ¹H NMR (500 MHz, Chloroform-*d*) δ 8.05 (d, *J* = 8.8, Hz, 2H), 7.30 (m, 3H), 7.01 (s, 1H), 4.19 (s, 2H), 2.23 (s, 3H), 1.39 (s, 9H). [M+H]⁺ : Calcd, 301.3580, Obsd, 301.3367



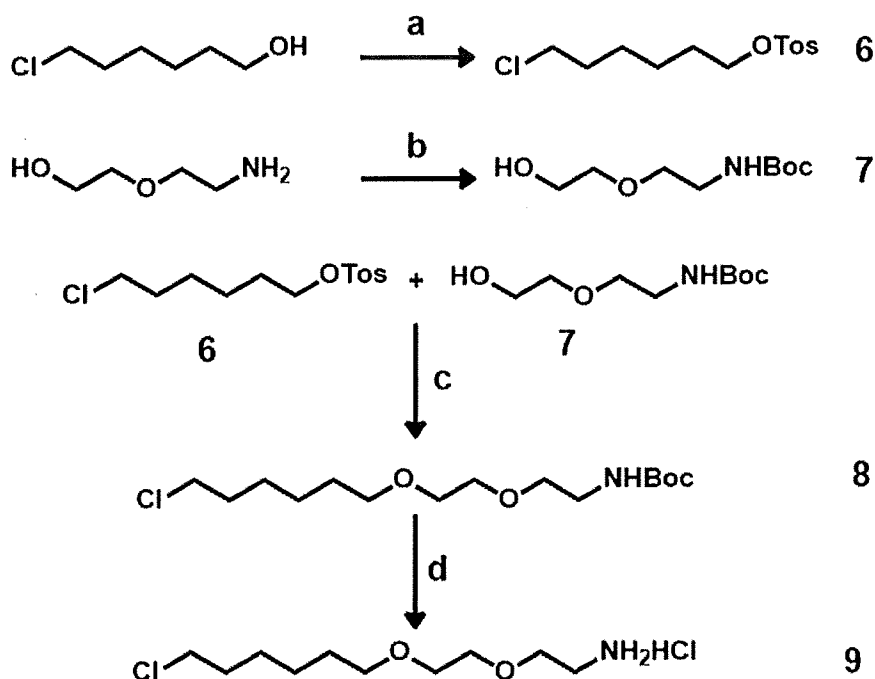
[0095] **Scheme 2. General protocol for synthesis of *in vitro* GFP mimic probe.** Carbon-carbon double bond formation via aldol condensation. Conditions, aldehyde (2eq), **2** or **5** (1 eq) were combine in dioxane under Argon. ZnCl₂ (0.01 eq) was added and the reaction was refluxed overnight. Solvent was removed and compounds were purified by flash chromatography (2:5, EA:Hexanes) to yield final GFP mimic probes **3/3a/3b**.

[0096] **3**: tert-butyl 2-((Z)-4-(4-(dimethylamino)benzylidene)-2-((E)-4-(dimethylamino)styryl)-5-oxo-4,5-dihydro-1H-imidazol-1-yl)acetate. Red Solid. ¹H NMR (500 MHz, Chloroform-*d*) δ 8.19 (d, *J* = 8.6 Hz, 2H), 7.97 (d, *J* = 15.6 Hz, 1H), 7.48 (d, *J* = 8.7,

2H), 6.75 (d, J = 8.6, 2H), 6.70 (d, J = 8.6, 2H), 6.42 (d, J = 15.6 Hz, 1H), 4.45 (s, 2H), 3.07 (d, J = 15.1 Hz, 12H), 1.49 (m, 9H). ¹³C NMR (126 MHz, Chloroform-*d*) δ 170.36, 167.11, 156.60, 151.43, 151.34, 149.45, 139.93, 135.54, 134.83, 134.30, 132.48, 129.40, 129.17, 127.63, 123.53, 123.21, 112.91, 111.95, 111.85, 111.71, 107.58, 82.72, 60.39, 42.79, 42.37, 40.17, 40.11, 40.03, 31.24, 29.71, 28.03, 27.98, 21.06, 14.21. [M+H]⁺ Calcd, 475.2664, Obsd, 475.3517

[0097] 3a: tert-butyl 2-((Z)-4-benzylidene-2-((E)-4-(dimethylamino)styryl)-5-oxo-4,5-dihydro-1H-imidazol-1-yl)acetate. Red oil. ¹H NMR (500 MHz, Chloroform-*d*) δ 8.25 (d, J = 8.5, 2H), 8.15 (d, J = 15.6 Hz, 1H), 7.56 – 7.50 (m, 2H), 7.46 (t, J = 7.6 Hz, 2H), 7.41 – 7.37 (m, 1H), 7.12 (s, 1H), 6.75 (d, J = 8.6, 2H), 6.41 (d, J = 15.5 Hz, 1H), 4.46 (d, J = 8.8 Hz, 2H), 3.07 (s, 6H), 1.50 (s, 9H). ¹³C NMR (126 MHz, Chloroform-*d*) δ 171.14, 170.53, 166.78, 159.80, 151.79, 142.13, 139.38, 134.98, 132.96, 132.37, 132.31, 132.18, 129.88, 129.67, 129.62, 128.64, 128.56, 125.54, 123.02, 111.89, 110.99, 106.56, 82.98, 60.39, 42.34, 40.13, 34.67, 34.53, 31.65, 31.59, 29.71, 29.06, 28.00, 27.96, 25.28, 22.66, 21.10, 21.05, 20.70, 14.20, 14.12, 11.43. [M+H]⁺ Calcd, 432.2242, Obsd, 432.2269

[0098] 3b: tert-butyl 2-((Z)-4-(4-(dimethylamino)benzylidene)-5-oxo-2-((E)-styryl)-4,5-dihydro-1H-imidazol-1-yl)acetate. Red solid. ¹H NMR (500 MHz, Chloroform-*d*) δ 8.20 (d, J = 8.7 Hz, 2H), 8.01 (d, J = 15.8 Hz, 1H), 7.65 – 7.57 (m, 2H), 7.42 (tt, J = 8.5, 4.1 Hz, 3H), 7.20 (s, 1H), 6.80 (d, J = 8.7, 2H), 6.68 (d, J = 15.8 Hz, 1H), 4.47 (s, 2H), 3.10 (s, 6H), 1.48 (s, 9H). ¹³C NMR (126 MHz, Chloroform-*d*) δ 170.11, 166.99, 155.57, 151.68, 139.12, 135.56, 135.04, 134.71, 129.72, 129.70, 128.94, 127.67, 122.82, 113.34, 111.84, 82.92, 67.10, 42.33, 40.10, 29.72, 27.98.



[0099] **Scheme 3. Synthesis of PEG linker in large scale.** Commercially available chloro-alcohol was converted into tosylate. Condition (a) 6-Chloro-1-hexanol (1.0 eq.), p-Toluenesulfonyl chloride (1.1 eq), 4-dimethylaminopyridine (0.1 eq) in pyridine for 1.5 hours at 0°C. The reaction mixture was extracted with diethyl ether against diluted HCl. The organic phase was collected and evaporated under reduced pressure, yielding crude colorless crystal **6**. The product was carried on to the next step without further purification. Commercially available amino-alcohol was protected with t-butyl carbonate group. Condition (b) 2-(2-Aminoethoxy)ethanol (1.0 eq.), Di-tert-butyl dicarbonate (1.0 eq.) in methanol stirred for 3 h at room temperature. The reaction mixture was extracted with PBS Buffer against dichloromethane. The organic fraction was dried in vacuo. Compounds were further purified by flash chromatography (1:1 ethyl acetate/hexane) to yield **7**, a colorless oil. **6** and **7** were combined to become protected PEG linker. Condition (c) **6** (1 eq), **7** (1.1 eq), and Potassium tert-butoxide (1 M in THF, 1.5 eq) in DMF stirred overnight at room temperature. The reaction was quenched with water and extracted with diethyl ether. The product was further purified by flash chromatography (1:2 ethyl acetate: hexane). Colorless oil **8** was obtained. Deprotection of **8**

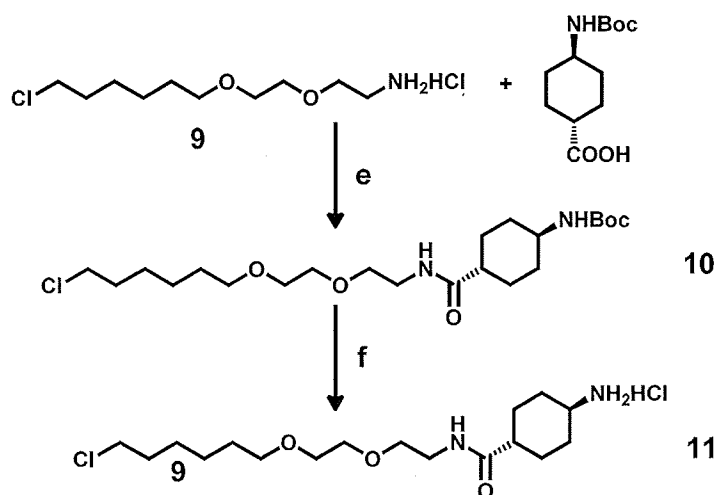
yielded the PEG amine hydrochloride. Condition (d) **8** (1 eq), HCl (4 M in Dioxane, 6 eq) stirred for 1 h at room temperature. Concentrated and then dried under high-vac to yield **9** a white solid.

[00100] **6: 6-chlorohexyl 4-methylbenzenesulfonate.** Colorless Crystal. ($R_f=0.55$, 1:9 EA:Hexane). ^1H NMR (500 MHz, CDCl_3) δ 7.73 (d, $J = 8.3$ Hz, 2H), 7.31 (d, $J = 8.4$ Hz, 2H), 3.97 (t, $J = 6.4$ Hz, 2H), 3.43 (t, $J = 6.6$ Hz, 2H), 2.39 (s, 3H), 1.73 – 1.54 (m, 4H), 1.39 – 1.24 (m, 4H). ^{13}C NMR (126 MHz, CDCl_3) δ 144.73, 132.99, 129.81, 127.73, 70.37, 44.75, 32.19, 28.54, 26.05, 24.57, 21.51. $[\text{M}+\text{Na}]^+$: Calcd, 313.0641, Obsd, 313.0639.

[00101] **7: tert-butyl (2-(2-hydroxyethoxy)ethyl)carbamate.** Colorless oil. ^1H NMR (500 MHz, CDCl_3) δ 5.29 (s, 1H), 3.73 – 3.64 (m, 2H), 3.56 – 3.45 (m, 4H), 3.31 – 3.15 (m, 3H), 1.45 – 1.34 (m, 9H). ^{13}C NMR (126 MHz, CDCl_3) δ 156.26, 79.30, 72.32, 70.31, 61.55, 40.37, 28.42. HRMS for $[\text{M}+\text{Na}]^+$: Calcd, 228.1212, Obsd, 228.1220.

[00102] **8: tert-butyl (2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)carbamate.** Colorless oil. ^1H NMR (500 MHz, CDCl_3) δ 5.02 (s, 1H), 3.58 – 3.47 (m, 8H), 3.42 (t, $J = 6.6$ Hz, 2H), 3.30 – 3.24 (m, 2H), 1.78 – 1.70 (m, 2H), 1.60 – 1.53 (m, 2H), 1.46 – 1.38 (m, 11H), 1.38 – 1.30 (m, 2H). ^{13}C NMR (126 MHz, CDCl_3) δ 156.03, 79.12, 71.28, 70.28, 70.22, 70.06, 45.02, 40.37, 32.56, 29.46, 28.44, 26.70, 25.44. $[\text{M}+\text{Na}]^+$: Calcd, 346.1761, Obsd, 346.1765.

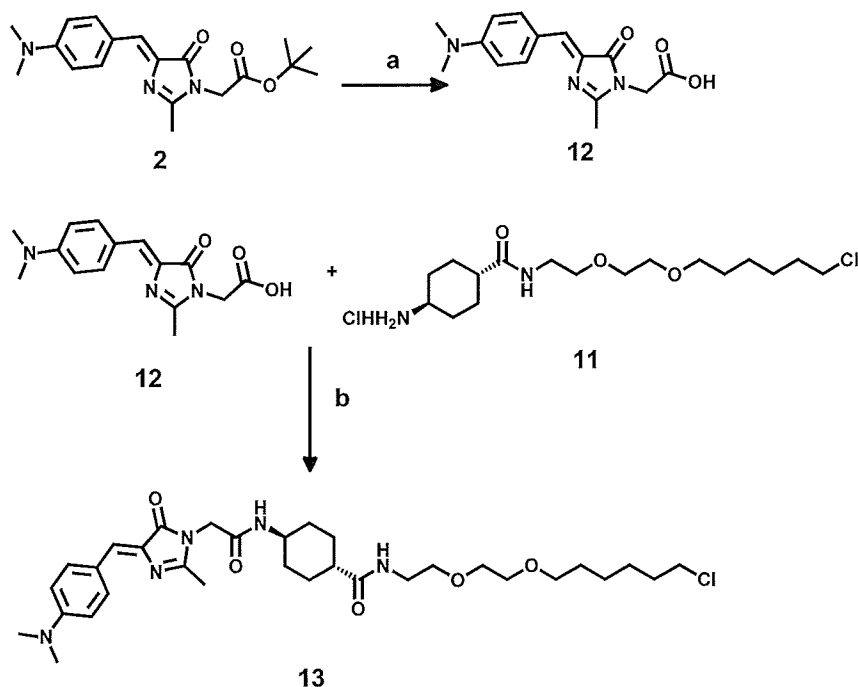
[00103] **9: 2-(2-((6-chlorohexyl)oxy)ethoxy)ethanamine hydrochloride.** White solid. ^1H NMR (500 MHz, DMSO) δ 8.26 (s, 3H), 3.65 – 3.58 (m, 4H), 3.56 – 3.52 (m, 2H), 3.51 – 3.46 (m, 2H), 3.36 (t, $J = 6.6$ Hz, 2H), 2.94 – 2.87 (m, 2H), 1.73 – 1.65 (m, 2H), 1.51 – 1.44 (m, 2H), 1.41 – 1.33 (m, 2H), 1.33 – 1.25 (m, 2H). ^{13}C NMR (126 MHz, DMSO) δ 70.23, 69.70, 69.34, 66.58, 45.39, 38.37, 32.03, 29.06, 26.13, 24.93. $[\text{M}+\text{H}]^+$: Calcd, 224.1417, Obsd, 224.1411.



[00104] Scheme 4. Synthesis of rigid linker for small molecule probe. Commercially available cyclohexane linker was coupled with homemade PEG linker. Condition (e) **9** (1 eq.), trans-4-[[[1,1-Dimethylethoxy)carbonyl]amino]cyclohexanecarboxylic acid (1 eq.), dimethylaminopyridine (0.1 eq.), N-(3-Dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride (2 eq.) in triethylamine (4 eq.) stirred overnight at room temperature. The reaction mixture was quenched by water and extracted with DCM. Product was purified by flash chromatography (100% ethyl acetate). Sticky white solid **10** was obtained. Deprotection of **10** yields final rigid linker amine hydrochloride. Condition (f) **10** (1 eq.), in HCl (4 M in Dioxane, 6 eq.) for 1h at room temperature. Reaction was concentrated and put under high-vac to yield fluffy white solid **11**.

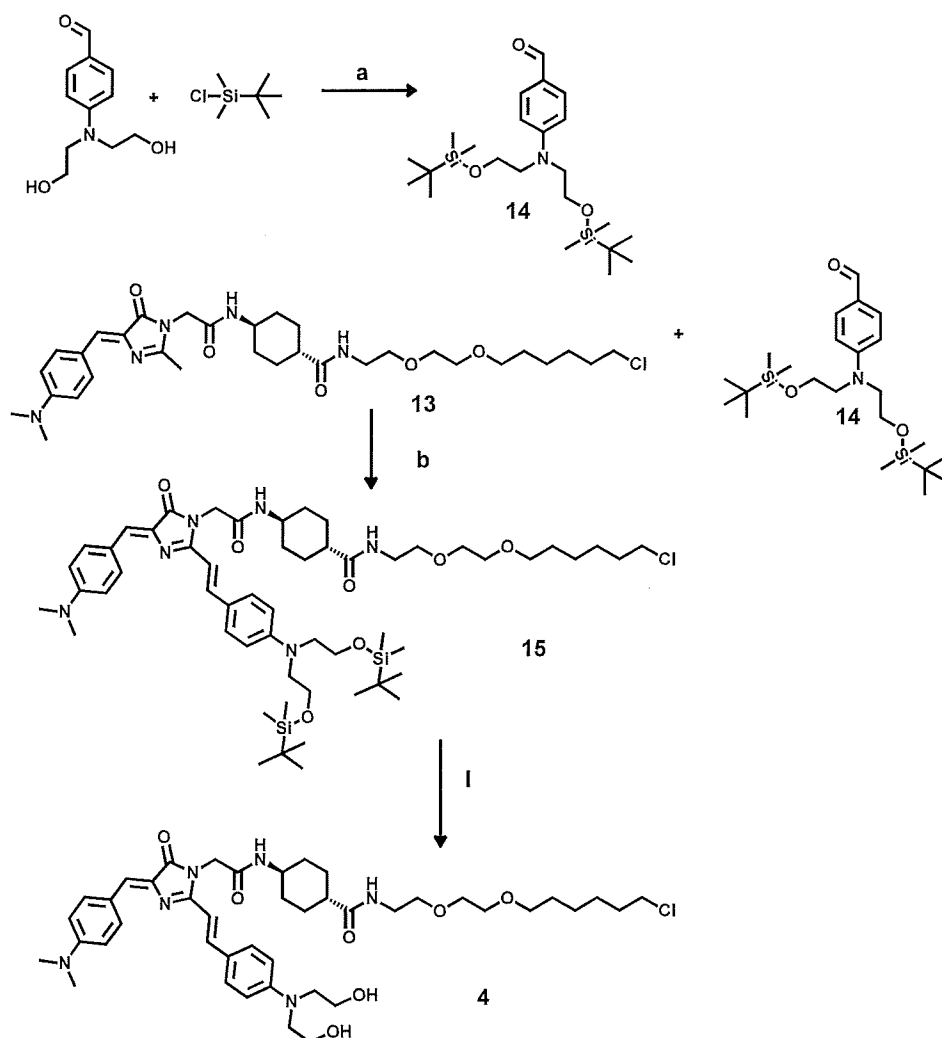
[00105] 10: tert-butyl ((1r,4r)-4-((2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)carbamoyl)cyclohexyl)carbamate. Sticky white solid. ¹H NMR (500 MHz, CDCl₃) δ 5.99 (s, 1H), 4.45 – 4.38 (m, 1H), 3.61 – 3.49 (m, 8H), 3.47 – 3.37 (m, 5H), 2.10 – 2.03 (m, 2H), 2.01 – 1.95 (m, 1H), 1.93 – 1.85 (m, 2H), 1.79 – 1.71 (m, 2H), 1.62 – 1.51 (m, 4H), 1.47 – 1.40 (m, 11H), 1.39 – 1.32 (m, 2H), 1.13 – 1.03 (m, 2H). ¹³C NMR (126 MHz, CDCl₃) δ 175.29, 155.23, 79.23, 71.32, 70.30, 70.07, 69.86, 49.13, 45.10, 44.63, 39.06, 32.80, 32.58, 29.54, 28.49, 26.75, 25.48. [M+H]⁺: Calcd, 449.2782, Obsd, 449.2783.

[00106] 11: (1r,4r)-4-amino-N-(2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)cyclohexanecarboxamide hydrochloride. Fluffy white solid. ¹H NMR (500 MHz, DMSO) δ 8.19 (s, 3H), 7.87 (s, 1H), 3.61 (t, *J* = 6.6 Hz, 2H), 3.52 – 3.43 (m, 4H), 3.40 – 3.33 (m, 4H), 3.20 – 3.13 (s, 2H), 2.96 – 2.87 (m, 1H), 2.11 – 2.03 (m, 1H), 2.01 – 1.94 (m, 2H), 1.78 – 1.65 (m, 4H), 1.53 – 1.43 (m, 2H), 1.43 – 1.25 (m, 8H). ¹³C NMR (126 MHz, DMSO) δ 174.43, 70.18, 69.59, 69.45, 69.07, 48.75, 45.39, 42.46, 38.39, 32.03, 29.43, 29.07, 27.19, 26.12, 24.94. [M+H]⁺: Calcd, 349.2258, Obsd, 349.2257.



[00107] **Scheme 5. Synthesis of GFP mimic precursors with rigid Halo linker.** Deprotection of GFP core. Condition (a) **2** deprotected according to literature protocols to yield **12** which was used without further purification. GFP core was coupled with rigid Halo linker. Condition (b) **2** (1 eq), **11** (1.0 eq), dimethylaminopyridine (0.1 eq), N-(3-Dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride (2.0 eq), triethylamine (4.0 eq), overnight, RT. The reaction mixture was quenched by water and extracted with DCM. The organic fraction was collected and dried in vacuo. Compounds were further purified by flash chromatography (20:1, DCM:MeOH) to yield **13** a yellow powder.

[00108] **13:** (1r,4r)-N-(2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)-4-(2-((Z)-4-(4-(dimethylamino)benzylidene)-2-methyl-5-oxo-4,5-dihydro-1H-imidazol-1-yl)acetamido)cyclohexanecarboxamide. Yellow powder. ¹H NMR (500 MHz, Chloroform-*d*) δ 8.09 (d, *J* = 8.8 Hz, 2H), 7.14 (s, 1H), 6.73 (d, *J* = 8.9 Hz, 2H), 6.17 – 5.87 (m, 2H), 4.23 (s, 2H), 3.60 – 3.40 (m, 10H), 3.09 (s, 6H), 2.40 (s, 3H), 2.10 – 1.32 (m, 18H), 1.22 – 1.09 (m, 2H). ¹³C NMR (126 MHz, Chloroform-*d*) δ 174.99, 170.56, 166.29, 157.54, 151.88, 134.58, 133.55, 130.56, 121.83, 111.76, 71.26, 70.24, 69.99, 69.80, 48.20, 45.07, 44.76, 44.36, 40.05, 39.01, 32.51, 32.03, 29.45, 28.25, 26.68, 25.40, 15.59. [M+H]⁺: Calcd, 618.3344, Obsd, 618.3441.



[00109] **Scheme 6. Protocol for synthesis of *in vivo* GFP mimic probe.** TBDMS protection. Conditions (a) 4-[Bis(2-hydroxyethyl)amino]benzaldehyde (1.0 eq) and imidazole (4.2 eq) were combined in DMF at room temperature, chloro(1,1-dimethylethyl)dimethylsilane (2.2 eq) was added and stirred for 1.5 hours. The reaction mixture was quenched with 1M HCl and extracted with ether. The organic fraction was collected and dried in vacuo to yield **14** a transparent brown oil, product was used without further purification. Carbon-carbon double bond formation via aldol condensation. Conditions (b) **14** (2.0 eq), **13** (1.0 eq) were combine in in dioxane under Argon. ZnCl₂ (0.01 eq) was added and the reaction was refluxed overnight. Solvent was removed and compounds were purified by flash chromatography (20:1, DCM:MeOH) to yield **15**. Product was immediately deprotected. Conditions (l) **15** (1.0 eq) was

combined with TBAF (3.0 eq) in THF and stirred overnight. Compounds were purified by flash chromatography (10:1, DCM:MeOH) to yield 4 a pink solid.

[00110] 4: (1*r*,4*r*)-4-(2-((*Z*)-2-((*E*)-4-(bis(2-hydroxyethyl)amino)styryl)-4-(4-(dimethylamino)benzylidene)-5-oxo-4,5-dihydro-1*H*-imidazol-1-yl)acetamido)-*N*-(2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)cyclohexanecarboxamide. Pink solid. ¹H NMR (500 MHz, DMSO-*d*₆) δ 8.17 (d, *J* = 8.3 Hz, 2H), 7.79 (d, *J* = 15.5 Hz, 1H), 7.55 (d, *J* = 8.7 Hz, 2H), 6.84 (s, 1H), 6.80 (dd, *J* = 9.9, 2.7 Hz, 2H), 6.75 (dd, *J* = 8.6, 5.5 Hz, 2H), 6.66 (d, *J* = 15.5 Hz, 1H), 4.82 (s, 2H), 4.38 (s, 2H), 3.59 – 3.43 (m, 12H), 3.41 – 3.35 (m, 6H), 3.20-3.13 (m, 2H) 3.04 (s, 6H), 2.12 – 1.17 (m, 18H). ¹³C NMR (126 MHz, DMSO-*d*₆) δ 175.24, 170.22, 166.39, 158.07, 151.49, 150.02, 139.33, 135.98, 134.26, 130.23, 125.20, 122.97, 122.84, 112.28, 111.89, 111.42, 108.29, 70.64, 70.04, 69.90, 69.55, 58.67, 58.60, 58.43, 53.59, 49.07, 48.09, 45.83, 43.53, 42.65, 38.85, 32.48, 32.07, 29.54, 28.61, 26.59, 25.40, 13.99. [M+H]⁺ Calcd, 809.4357, Obsd, 809.4369.

[00111] EXAMPLES OF APPLICATION

[00112] Examples provided herein are indicative of, but should not be construed as limiting of, the overall invention.

[00113] Protein aggregation is a multistep process that includes aberrant conformations in the form of soluble oligomers, disordered or amorphous aggregates, and amyloid fibrils containing ordered hydrogen-bonded β-sheet structures. Embodiments provide methods for detecting proteins as misfolded oligomers, insoluble aggregates, and amyloid fibers. Principle of detection includes providing disclosed compounds (fluorescent protein chromophores) that fluoresce when placed in an environment of enhanced rigidity indicative of protein aggregation in a local microenvironment. In a typical embodiment fluorescence occurs when the rigid microenvironment causes inhibition of non-radiative decay via twisted-intramolecular charge transfer (TICT). This principle can be applied to detection conducted in test tubes and live cells.

[00114] Detection as reported herein may be conducted using purified proteins in test tubes or in live cells bearing proteins of interest.

[00115] In one embodiment, protein aggregation and protein misfolding are detected by disclosed compounds in test tubes. This detection can be conducted with or without covalent conjugation to proteins of interest. Normally, purified proteins are used in this detection. These

proteins are subjected to in vitro conditions to induce protein misfolding and aggregation. Disclosed compounds can be added before, during, or after protein misfolding and aggregation. Fluorescence intensity can be recorded by fluorescence spectrophotometers or fluorescence microplate readers.

[00116] In one embodiment, aggregation of proteins of interest is detected in live cells either transiently or stably expressing these proteins. The inventors disclose the compounds with the following methods (varying R₃ groups) to conjugate with proteins of interest that are of at least one of the aberrant conformations in the form of soluble oligomers, disordered or amorphous aggregates, and amyloid fibrils containing ordered hydrogen-bonded β -sheet structures. These compounds can covalently conjugate with the Halo-Tag fusion domain (a product from Promega Inc) via a chloroalkane linker, the SNAP-Tag fusion domain (a product from NEB Inc.) via an O6 -benzylguanine linker, the CLIP-Tag fusion domain (a product from NEB Inc.) via an O6 -benzylcytosine linker, exposed cysteine via maleimide, or exposed lysine via N-Hydroxysuccinimide (Succinimidyl) esters. Fluorescence detection can be conducted by fluorescence spectrophotometers, fluorescence microplate readers, and epifluorescence or confocal fluorescence microscopes.

[00117] In one embodiment, aggregation of cellular proteins is detected in live cells. The inventors disclose the compounds with the following methods (varying R₃ groups) to bind to cellular proteins that are of at least one of the aberrant conformations in the form of soluble oligomers, disordered or amorphous aggregates, and amyloid fibrils containing ordered hydrogen-bonded β -sheet structures. These compounds can non-covalently bind to these conformations using a thioflavin-T guiding group or a tert-Butyloxycarbonyl guiding group. Fluorescence detection can be conducted by fluorescence spectrophotometers, fluorescence microplate readers, and epifluorescence or confocal fluorescence microscopes.

[00118] Applications as reported above can also be conducted in fixed cell samples with or without permeabilized cellular membranes.

[00119] Embodiments further provide a kit for detecting at least one of the following aberrant protein conformations: misfolded oligomers, amorphous insoluble protein aggregates, amyloid fibrils containing ordered hydrogen-bonded β -sheet structures. The kit includes the following components: one or more of disclosed fluorescent protein chromophores as reported herein of a known concentration in stock solutions of Dimethyl Sulfoxide or Ethanol, one or

more standard protein samples that form one of the abovementioned aberrant protein conformations, and optionally, instructions for use in detecting at least one of one of the abovementioned aberrant protein conformations.

[00120] Compounds detecting misfolded oligomers in test tubes (AggGlow method).

[00121] The inventors have designed and synthesized an HBI analogue, **3**, which harbors an extended π conjugation to both increase QY and restrict bond rotation (Figure 2).^{27,35,36,40,41} **3** exhibits quantum yield (Φ) value of 0.22 in glycerol, a viscous solvent that mimics misfolded oligomers, comparable to Φ of the Kaede protein as 0.33.⁴⁰ The molar extinction coefficient (ϵ) of **3** is 39,049 M⁻¹•cm⁻¹, about half of mCherry (72,000 M⁻¹•cm⁻¹). The inventors also determined photostability of **3** by measuring its absorbance spectra in glycerol and protein aggregates. When irradiated using 5 mW laser beam of 543 nm, absorbance of **3** at 530 nm only exhibits a 10% drop, suggesting that **3** is a photostable fluorophore. These photochemical properties make **3** suited for fluorescence microscopy.

[00122] Determination of Compound **3**'s detection of protein aggregation via fluorescence increase (Figure 3a), is demonstrated by analysis with the model protein α -synuclein (α -syn), whose aggregation is associated with Parkinson's disease. In the presence of **3**, aggregation of α -syn into mature fibers induced fluorogenic signal (Figure 12a). Fluorescence intensity of **3** was stoichiometric (1.0 for 140 μ M α -syn and 0.52 for 70 μ M α -syn in Figure 12a); whereas such stoichiometry was absent using Thioflavin T (ThT), a widely used fluorogenic probe to detect amyloid fibers (Figure 12b).

[00123] Compound **3** was also used to monitor kinetics of α -syn aggregation. Purified α -syn typically aggregates via a three-step process: formation of soluble oligomers, growth of amyloid fibers, and maturation of fibers (Figure 3b). Although ThT detects growth and maturation of fibers, it fails to detect soluble oligomers that are increasingly speculated to be the toxic species in PD (black curve in Figure 3b). By contrast, fluorescence of **3** started to increase at 4 h and reached a plateau at 8 h (red curve in Figure 3b). At this time point, formation of soluble oligomers was evidenced by a photo-induced crosslinking experiment (Figure 13a). Furthermore, a second phase of fluorescence increase was observed at 20 h and leveled off at 36 h (red curve in Figure 3b). These signal changes coincided with the time points of fiber growth and maturation as observed by ThT (black curve of Figure 3b). Thus, these data strongly indicate that **3** could detect multiple stages of protein aggregation.

[00124] In addition to α -synuclein, Compound **3** may detect aggregates formed by globular proteins, by using mutant superoxide dismutase 1 (SOD1), whose aggregation is commonly found in ALS disease. The recently discovered SOD1(V31A) mutant may be used to confirm this. After 20 min incubation at 59 °C, an 8 to 10-fold fluorescence intensity increase (Figures 3c and 14a) was observed using a wide concentration range of **3** (Figure 14b). The fluorescence signal was primarily found in the fraction of insoluble protein aggregates (Figure 14a), confirming the physical association between **3** and SOD1 aggregates. In addition, the kinetics of fluorescence increase from **3** was faster than that of turbidity whose signal originates from insoluble protein aggregates (Figure 3d). This suggests that **3** can detect both misfolded soluble proteins and insoluble aggregates.

[00125] **3** detects soluble oligomers via fluorescence increase. To test this point, the inventors chose α -synuclein (α -syn), whose aggregation is associated with Parkinson's disease.⁴² Purified α -syn aggregates via a three-step process: formation of soluble oligomers, growth of amyloid fibers, and maturation of fibers (Figure 3b). Although Thioflavin-T (ThT) monitors growth and maturation of fibers,⁴³ it fails to detect soluble oligomers (black curve in Figure 3b). By contrast, fluorescence of **3** started to increase at 4 h (gray curve in Figure 3b), at which soluble oligomers was detected by a chemical crosslinking assay (Figure 13). Furthermore, a second phase of fluorescence increase was observed at 20 h and leveled off at 36 h (gray curve in Figure 3b). These signal changes coincided with the time points of fiber growth and maturation as observed by ThT (black curve of Figure 3b). The inventors also demonstrated that **3** could detect misfolded oligomers formed by globular proteins, by using mutant superoxide dismutase 1 (SOD1), whose aggregation is commonly found in ALS disease.⁴⁴ Using the recently discovered V31A mutant,⁴⁵ the inventors observed 8 to 10-fold fluorescence intensity increase (ϕ from 0.02 to 0.22, Figure 3c) when purified SOD1-V31A was incubated at 59 °C. Further, the inventors found the kinetics of fluorescence increase from **3** was faster than that of turbidity whose signal originates from insoluble protein aggregates (Figures 3d and 22a) and identical to that of protein misfolding as measured by tryptophan fluorescence (Figure 22b). These data collectively suggest that **3** could detect misfolded oligomers.

[00126] **Compounds detecting misfolded oligomers in live cells (AggGlow method).**

[00127] **3** is used to monitor stress-induced proteome aggregation in HEK293T (human embryonic kidney) cells with a proteasome inhibitor MG132. Inhibition of proteasome has been

shown to form cytosolic misfolded oligomers and insoluble aggregates of cellular proteins. When 5 μ M of MG132 was used to treat HEK293T cells in the presence of 2.5 μ M **3**, obvious fluorescence pattern was developed in the cytosol of cells as a result of proteasome inhibition (Figure 23). The diffusive fluorescence structure indicates misfolded oligomers and the granular fluorescence structure indicates insoluble aggregates. Similar to this application, **3** was used to monitor stress-induced proteome aggregation in HeLa cells with a heat-shock protein 90 (Hsp90) inhibitor 17-AAG. Inhibition of Hsp90 has been shown to form cytosolic misfolded oligomers and insoluble aggregates of cellular proteins. When 1 μ M of 17-AAG was used to treat HeLa cells in the presence of 2.5 μ M **3**, obvious fluorescence pattern was developed in the cytosol of cells as a result of Hsp90 inhibition (Figure 23).

[00128] Compounds detecting misfolded oligomers of specific proteins in live cells (AggTag method).

[00129] The inventors further examined using a turn-on fluorescence to monitor aggregation of a protein-of-interest (POI) using in live cells. For this example, the inventors genetically fused Halo-Tag to the POI and synthesized **4** for bioorthogonal conjugation (Figure 4a, hereafter referred to as AggTag method). Compound **4**:

[00130] Using the AggTag method, the inventors expected that the fluorophore should remain dark when POI is folded. Aggregation of POI, however, should bury the fluorophore in protein aggregates that can restrict its rotation and trigger fluorescence (Figure 4a). **4** bears two additional hydroxyl groups to improve solubility of **3** and a warhead for bioconjugation to Halo-Tag (Figure 4b). The fluorescent properties of **4** were found to be identical to **3** (Figure 15). In addition, **4** exhibited substantially lower fluorescent background signal than typical solvatochromic fluorophores (e.g., SBD) and molecular rotor fluorophores (e.g. CCVJ), when incubated with bovine serum albumin as a representation of proteins with hydrophobic surfaces and sodium dodecyl sulfate as a mimic of lipids (Figure 16).

[00131] To test whether **4** is fluorogenic upon protein aggregation in live cells, an experiment was conducted using the Huntingtin exon 1 protein (Htt) with expansion of a polyglutamine tract within its N-terminal domain, well known for its severe protein aggregation. The inventors first observed protein aggregation using conventional non-fluorogenic methods (Figure 17a). When HEK293T cells expressed either Htt-97Q-mCherry or Htt-97Q-Halo conjugated with an always-fluorescent TMR ligand, the inventors identified obvious fluorescent punctate structures with

diffuse background fluorescence (Figure 17b). Cells expressing the non-aggregating Htt-0Q-Halo in conjugation with TMR exhibited only diffuse fluorescence (Figure 17b).

[00132] Thus, protein aggregation is primarily judged by the appearance of fluorescent puncta in non-fluorogenic methods (Figure 17a). The AggTag method, however, yielded a dark background in cells expressing Htt-0Q-Halo labeled by **4** (Figure 4c). The lack of fluorescence in these cells did not come from the lack of protein expression indicated by a dual-probe labeling experiment using **4** and an always-fluorescent coumarin ligand (Figures 18a-b). The turn-on fluorescence from **4**, however, was observed as puncta without any background diffuse fluorescence when Htt-97Q-Halo was expressed (Figures 4c and 18b).

[00133] This fluorescence was not due to non-specific binding of **4** in cells, because HEK293T cells expressing Htt-97Q without Halo-Tag showed no fluorescent signal (Figure 18c). In accordance with these observations, fluorescence from cell lysates confirmed that the fluorescence of Htt-97Q-Halo was more intense than that of Htt-0Q-Halo (Figures 4d and 18d). Furthermore, fractionation of the lysate indicated that the fluorescence of Htt-97Q-Halo primarily originated from the insoluble fraction (Figures 4d and 18d). Taken together, these experiments establish the AggTag method as a fluorogenic approach to detect protein aggregation in live cells.

[00134] Experiments further demonstrated that the AggTag method could visualize previously invisible misfolded soluble proteins in live cells. The SOD1(V31A) mutant is associated with a slow disease progression. So far, little had been known about its aggregation propensity in live cells. To this end, the inventors expressed and labeled SOD1(V31A)-Halo fusion protein simultaneously with the coumarin ligand and **4** in HEK293T cells. Using the coumarin fluorescence, the inventors found that SOD1(V31A) was primarily located in the cytoplasm and the oxidative stress inducer NaAsO₂ induced the partial translocation of SOD1(V31A) to the nucleus. Although these observations are consistent to previously reported cellular locations of other SOD1 mutants, it was unclear whether NaAsO₂ caused aggregation because the coumarin fluorescence remained diffuse before and after stress (left panel, Figure 5a). By contrast, **4** was dark before stress but exhibited both diffuse and punctate fluorescent structures in stressed cells (middle panel, Figure 5a). This turn-on fluorescence signal was not due to aggregation of the Halo-**4** conjugate, as it remained dark in stressed cells (Figure 19).

[00135] The punctate fluorescence could be rationalized by a fractionation experiment wherein more aggregates of SOD1(V31A) were found in cells treated with NaAsO₂ (Figures 5b and 20). The diffusive fluorescence of **4** was believed to arise from misfolded soluble proteins, as demonstrated by several lines of evidence from *in vitro* heat-induced aggregation experiments. First, the inventors found that fluorescence of the SOD1(V31A)-Halo-**4** conjugate increased in the absence of insoluble protein aggregates (Figure 21a). The Halo-**4** conjugate, as a control, remained non-fluorescent under all heat conditions (Figure 21). Second, the kinetics of fluorescence increase was much faster than the kinetics of turbidity (Figure 22a). Finally, the kinetics of fluorescence coincided with the kinetics of protein misfolding (Figure 22b). Collectively, these data support the belief that the AggTag method in combination with **4** is able to selectively visualize both misfolded soluble proteins and insoluble aggregates via diffuse and punctate fluorescence, respectively, in live cells.

[00136] Cellular background and control experiments of oligomer-detecting compounds.

[00137] A series of control experiments were carried out to demonstrate that the turn-on fluorescence specifically monitors misfolding of POIs. First, **4** fluorescence was not observed in cells expressing Halo-Tag in the presence of the NaAsO₂ stressor (50 μ M, 18 h, 37 °C). Second, HEK293T cells co-expressing Htt-Q97-GFP and Halo-Tag or just expressing Htt-Q97 showed no fluorescence of **4**. Third, turn-on fluorescence was not observed in cells co-expressing Halo-Tag and SOD1-V31A-GFP in the presence of the NaAsO₂ stressor (50 μ M, 18 h, 37 °C). These control experiments collectively demonstrate that the observed fluorescent signal from **4** was not due to non-specific binding between **4** and aggregates of other cellular proteins in cells.

[00138] Compounds detecting insoluble aggregates in test tubes (AggTag method).

[00139] **2** only activates fluorescence when POI forms insoluble aggregates, while it remains dark when POI is folded and forms soluble oligomers. The inventors installed Halo-tag reactive warhead to **2**, resulting in **5**. In heat-induced aggregation of SOD1-A4V-Halo that is conjugated to **5**, it was found that the kinetics of P2b fluorescence increase was similar to that of turbidity whose signal originates from insoluble aggregates (green vs black curve, Figure 24a). In contrast, when SOD1 A4V-Halo was conjugated with **5**, the kinetics of **4** fluorescence increase was faster than that of turbidity and fluorescence increase of P2b (red vs black curve, Figure 24a). **5** exhibits a quantum yield of 0.24 in insoluble aggregates and molar extinction coefficient of 23,716 M⁻¹•cm⁻¹, about 40% of EGFP (ϕ =0.6; ϵ =56,000 M⁻¹•cm⁻¹).

[00140] Compounds detecting insoluble aggregates of specific proteins in live cells (AggTag method).

[00141] It is important to distinguish insoluble aggregates from misfolded oligomers, because they do not only have distinct functions in cells, but also are managed differently by cells.^{46,47} Soluble oligomers are targeted to the Juxta-Nuclear Quality (JUNQ) control compartment, which forms under severe stress conditions and contains polyubiquitylated proteins, such as mutants of SOD1. In contrast, insoluble aggregates, such as polyglutamine-expanded Huntingtin (Htt-polyQ), are sequestered into a cytoplasmic compartment known as the Insoluble Protein Deposit (IPOD). Thus, misfolded oligomers do not necessarily display as a diffusive structure, instead they can reside in granular structures that appear to be almost identical to granules formed by insoluble aggregates.

[00142] The spectra of **4** ($E_x/E_m = 530/600$ nm; Fig 24e) and **5** ($E_x/E_m = 450/520$ nm; Figures 24b-c) make them suited for two-color imaging applications, wherein the fluorescence of **4** (red) and **5** (green) can monitor misfolded oligomers and insoluble aggregates, respectively (Figures 24b-c).

[00143] Combination of **4** and **5** enables a two-color imaging strategy to differentiate insoluble aggregates from soluble oligomers. The inventors carried out live cell imaging experiments, wherein **4** or **5** was used to visualize aggregation of POI-Halo fusion proteins in HEK293T cells. Under proteasome inhibition by a drug MG132, mutants of SOD1 has been shown to form JUNQ compartments that contain soluble oligomers.⁴⁷ If this were true, it would be expected that granules exhibit turn-on fluorescence with **4** but not **5**. To this end, the inventors labeled HEK293T cells expressing SOD1-A4V-Halo simultaneously with **4** and **5** (both at 0.5 μ M) for 24 h and treated cells with 5 μ M MG132 for 8 h. As expected, SOD1 A4V-Halo formed mostly perinuclear granules that only exhibited red fluorescence from **4** (Figure 24d) and dark fluorescence to a background level from **5** (Figure 24d). The inventors further carried out experiments using Htt-PolyQ with 110 glutamine repeats (Htt-110Q-Halo). Htt-polyQ with longer than 78Q forms IPOD inclusions that contain insoluble aggregates at both cytosolic and perinuclear localizations.⁴⁷ Consistent to this note, Htt-110Q-Halo formed inclusions that exhibit fluorescence from both **4** and **5** (Figures 24f-g).

[00144] Cellular background and control experiments of insoluble aggregates-detecting compounds.

[00145] As a control, **5** fluorescence was not observed in cells co-expressing Halo-tag and Htt-110Q-mCherry (Figure 24h), suggesting that observed fluorescent signal from **5** was not due to non-specific binding between **5** and aggregates of other cellular proteins in cells.

[00146] **Compounds enable bioconjugation to Halo-Tag and SNAP-Tag fusion proteins (AggTag method).**

[00147] Halo-Tag is an engineered dehalogenase that reacts with chloroalkane molecules to form stable covalent enzyme-ligand conjugates, and it serves as an ideal sensor platform because it exhibits fast labeling kinetics, a bioorthogonal reaction profile, and demonstrated evolvability^{48,49}. SNAP-tag is a prominent self-labelling protein tag used for live cell imaging of POIs, due to its relatively small size (19.4 kDa, two-thirds the size of GFP as 27 kDa) and fast labelling kinetics.⁵⁰⁻⁵³

[00148] Probes can be developed to detect aggregation of POI fused with SNAP-tag. Using **3** as the fluorophore (Figure 25a), **6–11** were synthesized by conjugating six different types of linkers (sarcosine, proline, cyclohexane, glycine, propane, and hexane; Figure 25b) to **3** and O⁶-BG. Synthesis is described in Figure 25c and produces the final product with a 34% yield. When conjugated with folded SNAP-Tag, **6–8** with short or rigid linkers exhibited minimal fluorescence increase ($\phi < 0.01$; Figure 25d) and **9–11** with long and flexible linkers resulted in fluorescently bright conjugates ($\phi > 0.06$; Figure 25d). These results support that **6–8** can be used for SNAP-Tag fusion proteins.

[00149] EXPERIMENTAL METHODS

[00150] **Plasmids.** *Mammalian expression:* pHTN vector (Promega, Inc) with a stop codon added to the c-terminal of Halo-Tag protein. The SOD-1 gene was amplified from the pF146 pSOD1WTAcGFP1 (a gift from Elizabeth Fisher, Addgene plasmid #26407), respectively. The V31A mutation was introduced to SOD1 via QuickChange PCR. The *Htt*-97Q gene was amplified from the pCDNA3.1-*Htt*-97Q-mCherry (Max Planck Institute of Biochemistry). These genes were sub-cloned into a pHTC HaloTag CMV-neo vector by the PIPE cloning method.

Protein expression: pET29b vectors were constructed to encode Halo-Tag-His6, SOD1(V31A)-linker-Halo-His6 (linker contains a TEV protease cleavage site), and α -synuclein.

[00151] **Protein expression and purification.** *Halo-Tag and SOD1(V31A)-Halo:* *E. coli* BL21 DE3* competent cells harboring a pBAD vector encoding σ 32-I54N were transformed with pET29b vectors containing Halo-His6 and SOD1(V31A)-TEV-Halo-His6 proteins. Expression

and purification was carried out as previously described. In brief, cells expressing recombinant proteins were thawed and lysed by sonication at 4 °C in the presence of a protease inhibitor (1 mM PMSF). Lysed cells were centrifuged for 60 min at 16,000 × g. The supernatant was collected and loaded onto a 6 mL BioRad Nuvia Ni-IMAC column and washed with 120 mL of buffer containing 50 mM Tris • HCl (pH 7.5) and 100 mM NaCl. The protein was then eluted by gradient addition of buffer containing 50 mM Tris • HCl (pH 7.5), 100 mM NaCl, and 500 mM imidazole over a volume of 48 mL. The protein fractions were identified by SDS-PAGE analysis, pooled, and concentrated. The protein was further purified using a 120 mL HiPrep™ 16/60 Sephacryl™ S-200 HR size-exclusion column. The protein containing fractions were identified by SDS-PAGE gel analysis, pooled, and concentrated. No significant impurities were identified and purity was estimated to be > 98% based on SDS-PAGE. *SOD1(V31A)*: Purified SOD1(V31A)-Halo protein was subjected to a 1 h TEV protease cleavage (0.50 μM TEV protease for every 10 μM SOD1-Halo protein) in the presence of 1 mM DTT at 25 °C. Reaction mixture was retro-purified via BioRad Nuvia Ni-IMAC resin. Flow through was collected as cleaved SOD1(V31A). No significant impurities were identified and Halo purity was estimated to be > 98% based on SDS-PAGE. *α-synuclein*: Protein expression was induced with 0.1 M IPTG, 5 hours, 37 °C. Harvest culture at 5,000 RPM, 15 min, 4 °C. Resuspend the pellet in DPBS (200 mg/L KCl, 200 mg/L KH₂PO₄, 8 g/L NaCl, 2.16 g/L Na₂HPO₄•7H₂O), transfer to a 50mL conical tube. Spin at 4,000 rpm, 30 min, 4°C. Decant the supernatant, and store the pellet in a -80 °C freezer. Next morning, resuspend defrosted pellet in osmotic shock buffer (30 mM Tris-HCl, 40% sucrose, 2 mM EDTA, pH 7.5; 100 mL for each liter of starter culture). Incubate for 10 min at room temperature. Collect pellet by centrifugation (12,000 rpm, 20 min). Quickly resuspend pellet with ice-cold water (90 mL for each liter of starter culture). Add 4 M MgCl₂ (76.5 μl for each liter of starting culture) and keep on ice for 3 min. Centrifuge at 15,000 rpm, 30 min. Collect and filter supernatant. Add 2 M Tris-HCl, pH 8 at a 1:100 (v/v) ratio. Load supernatant onto UNO Sphere Q column (5 mL, Bio-Rad), elute with a gradient from Buffer A (20 mM Tris, pH 8) to Buffer B (20 mM Tris, 500 mM NaCl, pH 8). Analyze fractions by 15% SDS-PAGE gel. Pool protein-containing fractions, and dilute by half with degassed Buffer A. Centrifuge at 15,000 rpm, 30 min to remove aggregates. Load onto an EnrichQ column (8 mL) on the NGC-Quest10 FPLC system and elute with a gradient from 0% to 100% Buffer B. Analyze fractions by 15% SDS-PAGE gel. Pool protein-containing fractions and dilute by half with degassed

Buffer A. Load onto an EnrichQ column (8 mL) and elute with a gradient from 40% to 52% Buffer B. Analyze fractions by 15% SDS-PAGE gel. Pool protein containing fractions and flash freeze small aliquots of protein. Proteins can only be thawed once before use.

[00152] Confocal microscope imaging. The HEK293T cells were seeded at 25% confluency 24 h prior to transfection in poly-D-lysine coated 35 mm glass bottom dishes (MatTek Corporation). Cells were grown in DMEM media supplemented with 10% FBS and Penicillin-Streptomycin antibiotics until they reached 50-60% confluency. Transfection was carried out using X-tremeGene™ 9 DNA transfection reagent (Roche) according to the manufacturer's instructions. Proteins were expressed for 24 h prior to analyses. To label proteins with Halo-Tag fusion, protein expression was carried out in the presence of 1 μ M **4** or 1 μ M TMR Halo-Tag ligand to form covalent conjugate with the Halo-Tag domain. For a dual-probe labeling scheme, 1 μ M **4** and 1 μ M coumarin Halo-Tag ligand to simultaneously form covalent conjugate with the Halo-Tag domain.

[00153] To wash off unbound Halo ligands, the cells were washed extensively by replacing media with fresh DMEM and incubating for 30 min at 37 °C. For confocal fluorescence imaging with either **4**-labeled Halo fusion, TMR-labeled Halo fusion, EGFP fusion, or mCherry fusion proteins, DMEM media was replaced with FluoroBrite™ DMEM media (ThermoFisher) supplemented with 10% FBS, and Hoechst 33342 (0.1 μ g/mL). For confocal fluorescence imaging with either dual-probe labeled Halo fusion with **4** and coumarin ligand, DMEM media was replaced with FluoroBrite™ DMEM media (ThermoFisher) supplemented with 10% FBS. The samples were incubated for 30 min prior to imaging. Media was replaced with fresh FluoroBrite™ DMEM media (ThermoFisher) supplemented with 10% FBS prior to imaging. Confocal images were obtained using Olympus FluoView™ FV1000 confocal microscope. The EGFP fluorescence was visualized using blue argon (488 nm) laser. Nuclear staining and coumarin fluorescence was visualized using violet laser (405 nm). Fluorescence of TMR and **4** were visualized using green HeNe laser (543 nm).

[00154] Chemical-induced proteome stress and confocal imaging of stressed cells.

HEK293T cultures were seeded at 25% confluency 24 h prior to transfection in 12-well plate for time dependent fluorescence plate reader analysis or 35 mm glass bottom culture dishes (Poly-D-lysine coated, MatTek Corporation). Cells were grown in DMEM medium supplemented with 10% FBS and penicillin-streptomycin antibiotics until they reached 50-60% confluency.

Transfection was performed using X-tremeGene™ 9 DNA transfection reagent (Roche). After 24 h of protein expression and co-translational labeling, medium was replaced with fresh DMEM to diffuse out unbound ligands. After 30 min, media was replaced by fresh DMEM medium containing DMSO vehicle or NaAsO₂ (50 μM). Cells were incubated for 18 h, 37 °C in a CO₂ incubator. Nuclear stain Hoechst 33342 (0.1 μg/mL) was added to the medium 30 min prior to confocal imaging. **Aggregation Assays.** *α-synuclein*: Aggregation solution contained 70 or 140 μM α-synuclein in 20 mM HEPES (pH 7.5) and 100 mM NaCl. 10 μM of **3** or ThT was added to the solution at the beginning of reaction. Previous studies have shown that the kinetics of α-synuclein aggregation is unaffected by the addition of ThT. A volume of 150 μL of the mixture was pipetted into a well of clear-bottom 96-well plate, which was subsequently sealed with Mylar plate sealers. The plate was loaded on a Heidolph vibrating platform shaker and shake at 1,350 rpm at 37 °C. At indicated time points, fluorescence reading was obtained using a Tecan M1000Pro fluorescence plate reader. Fluorescence intensity was recorded at E_x = 530 nm/E_m = 600 nm for **3** and E_x = 450 nm/E_m = 480 nm for ThT. Each sample was run in triplicate or quadruplicate. *SOD1(V31A)*: Aggregation was carried out with DPBS buffer containing indicated concentrations of EDTA that is used to chelate the structural metal of SOD1. Aggregation solution contained 42 μM SOD1(V31A) or SOD1(V31A)-Halo and 21 μM of **3** was added to the solution at the beginning of reaction. Reaction was carried out quiescently at indicated temperatures. Fluorescence of **3** was measured at E_x = 530 nm/E_m = 600 nm.

[00155] X-ray structure determination of 3. Single crystals of **3** (C₂₈H₃₄N₄O₃) were grown by evaporation in EtOH and ethyl acetate mixed solvents. A suitable crystal was selected and mounted on a nylon loop, with help of paratone oil on a 'SMART APEX CCD area detector' diffractometer (molybdenum target, 1600 kW). The crystal was kept at 243 K during the data collection. Using Olex2⁵⁴, the structure was solved with the XS⁵⁵ structure solution program using Direct Methods and refined with the XL refinement package using Least Squares minimization without use of any constraints/restraints. The hydrogen atoms were placed geometrically, and rode their parent atoms during the refinement. *Crystal Data for C₂₈H₃₄N₄O₃* (*M* = 474.59 g/mol): orthorhombic, space group P2₁2₁2₁ (no. 19), *a* = 9.895(4) Å, *b* = 21.470(10) Å, *c* = 25.711(12) Å, *V* = 5463(4) Å³, *Z* = 8, *T* = 243 K, μ(MoKα) = 0.076 mm⁻¹, *D*_{calc} = 1.154 g/cm³, 51829 reflections measured (3.7° ≤ 2θ ≤ 57.88°), 13915 unique (*R*_{int} = 0.1226, *R*_{sigma} =

0.1327) which were used in all calculations. The final R_1 was 0.0570 ($>2\sigma(I)$) and wR_2 was 0.1767 (all data).

[00156] Negative staining electron microscopy. Carbon-coated 400 mesh EM grids supported by Formvar film were plasma-discharged prior to use. A small amount ($\sim 5 \mu\text{L}$) of sample suspension was applied onto the grids. The grids were then blotted to remove excess of the sample and stained with 1% aqueous uranyl. Images were collected at indicated resolutions.

[00157] SUMMARY

[00158] Embodiments as reported herein demonstrate that analogues of FP chromophores can fluoresce in protein aggregates. Different from previous non-fluorogenic methods, methods and compositions, as reported herein, can visualize both misfolded soluble proteins and insoluble aggregates in intact live cells. Such fluorogenic detection can be achieved via chemical modulation of fluorophores with molecular rotor and AIE properties, providing new applications for this large family of molecules. The unique fluorogenicity of this class of probes, combined with the AggTag method, make them generally applicable to a wide range of proteins whose aggregation is associated with diseases and suited to potentiate screening platform to explore therapeutics that can ameliorate aggregation of these pathogenic proteins.

[00159] It will be apparent to those skilled in the art that numerous modifications and variations of the described examples and embodiments are possible in light of the above teachings of the disclosure. The disclosed examples and embodiments are presented for purposes of illustration only. Other alternative embodiments may include some or all of the features of the various embodiments disclosed herein. Therefore, it is the intent to cover all such modifications and alternative embodiments as may come within the true scope of this invention, which is to be given the full breadth thereof. Additionally, the disclosure of a range of values is a disclosure of every numerical value within that range, including the end points.

[00160] Therefore, while certain exemplary embodiments of apparatuses and methods of making and using the same have been discussed and illustrated herein, it is to be distinctly understood that the invention is not limited thereto but may be otherwise variously embodied and practiced within the scope of the following claims. Documents reported herein are incorporated by reference herein; however, if there is a conflict between this document and the incorporated document, this document controls.

- [00161] The following documents may be interesting or useful in reviewing and appreciating embodiments as reported herein. Inclusion of these documents is not an admission of any sort.
- [00162] Grimm, *et al.*, A general method to improve fluorophores for live-cell and single-molecule microscopy. *Nat Methods* **2015**, *12* (3), 244-50, 3 p following 250.
- [00163] Ando, *et al.*, An optical marker based on the UV-induced green-to-red photoconversion of a fluorescent protein. *Proc Natl Acad Sci USA* **2002**, *99* (20), 12651-6; Subach, *et al.*, Chromophore transformations in red fluorescent proteins. *Chem Rev* **2012**, *112* (7), 4308-27; Tomura, *et al.*, Monitoring cellular movement in vivo with photoconvertible fluorescence protein "Kaede" transgenic mice. *Proc Natl Acad Sci USA* **2008**, *105* (31), 10871-6.
- [00164] Hong, *et al.*, Aggregation-induced emission. *Chem Soc Rev* **2011**, *40* (11), 5361-88.
- [00165] Lashuel, *et al.*, The many faces of alpha-synuclein: from structure and toxicity to therapeutic target. *Nat Rev Neurosci* **2013**, *14* (1), 38-48.
- [00166] Qin, *et al.*, Highly sensitive amyloid detection enabled by thioflavin T dimers. *Mol BioSys* **2010**, *6* (10), 1791-5.
- [00167] Rosen, *et al.*, Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature* **1993**, *362* (6415), 59-62.
- [00168] Dangoumau, *et al.*, Novel SOD1 mutation p.V31A identified with a slowly progressive form of amyotrophic lateral sclerosis. *Neurobiol Aging* **2014**, *35* (1), 266 e1-4.
- [00169] Los, *et al.*, HaloTag: a novel protein labeling technology for cell imaging and protein analysis. *ACS Chem Biol* **2008**, *3* (6), 373-82; Lin, *et al.*, A Generalizable platform for interrogating target- and signal-specific consequences of electrophilic modifications in redox-dependent cell signaling. *J Am Chem Soc* **2015**, *137* (19), 6232-6244; Ballister, *et al.*, Localized light-induced protein dimerization in living cells using a photocaged dimerizer. *Nat Commun* **2014**, *5*, 5475.
- [00170] Gidalevitz, *et al.*, Progressive disruption of cellular protein folding in models of polyglutamine diseases. *Science* **2006**, *311* (5766), 1471-4; DiFiglia, *et al.*, Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain. *Science* **1997**, *277* (5334), 1990-3.
- [00171] Ramdzan, *et al.*, Tracking protein aggregation and mislocalization in cells with flow cytometry. *Nat Methods* **2012**, *9* (5), 467-476; Kitamura, *et al.*, Conformational Analysis of

Misfolded Protein Aggregation by FRET and live-cell imaging techniques. *Int J Mol Sci* **2015**, *16* (3), 6076-6092.

[00172] Mateju, *et al.*, An aberrant phase transition of stress granules triggered by misfolded protein and prevented by chaperone function. *EMBO J* **2017**, *36* (12), 1669-1687; Zhong, *et al.*, Nuclear export of misfolded SOD1 mediated by a normally buried NES-like sequence reduces proteotoxicity in the nucleus. *Elife* **2017**, *6*.

[00173] Haidekker, *et al.*, Molecular rotors - fluorescent biosensors for viscosity and flow. *Org Biomol Chem* **2007**, *5*(11), 1669-1678; (b) Qian, H.; Cousins, M. E.; Horak, E. H.; Wakefield, A.; Liptak, M. D.; Aprahamian, I., Suppression of Kasha's rule as a mechanism for fluorescent molecular rotors and aggregation-induced emission. *Nat Chem* **2017**, *9* (1), 83-87.

[00174] Kalaparthi, *et al.*, The nature of ultrabrightness of nanoporous fluorescent particles with physically encapsulated fluorescent dyes. *J Mater Chem C* **2016**, *4* (11), 2197-2210.

[00175] Wurth, *et al.*, Relative and absolute determination of fluorescence quantum yields of transparent samples. *Nat Protoc* **2013**, *8* (8), 1535-1550.

[00176] Dolomanov, *et al.*, OLEX2: a complete structure solution, refinement and analysis program. *J Appl Crystallogr* **2009**, *42*, 339-341.

[00177] Sheldrick, G. M., A short history of SHELX. *Acta Crystallogr A* **2008**, *64*, 112-122.

[00178] Baldridge, *et al.*, Efficient synthesis of new 4-arylideneimidazolin-5-ones related to the GFP chromophore by 2+3 cyclocondensation of arylideneimines with imidate ylides, *Synthesis* 2010, *14*, 2424-2436.

REFERENCES

[00179] The following references are each relied upon and incorporated herein in their entirety.

[00180] [1] Balch, W. E., Morimoto, R. I., Dillin, A., and Kelly, J. W. (2008) Adapting proteostasis for disease intervention, *Science* *319*, 916-919.

[00181] [2] Eisele, Y. S., Monteiro, C., Fearn, C., Encalada, S. E., Wiseman, R. L., Powers, E. T., and Kelly, J. W. (2015) Targeting protein aggregation for the treatment of degenerative diseases, *Nature reviews. Drug Discovery* *14*, 759-780.

- [00182] [3] Labbadia, J., and Morimoto, R. I. (2013) Huntington's disease: underlying molecular mechanisms and emerging concepts, *Trends Biochem Sci* 38, 378-385.
- [00183] [4] Kaushik, S., and Cuervo, A. M. (2015) Proteostasis and aging, *Nature Medicine* 21, 1406-1415.
- [00184] [5] Olzscha, H., Schermann, S. M., Woerner, A. C., Pinkert, S., Hecht, M. H., Tartaglia, G. G., Vendruscolo, M., Hayer-Hartl, M., Hartl, F. U., and Vabulas, R. M. (2011) Amyloid-like aggregates sequester numerous metastable proteins with essential cellular functions, *Cell* 144, 67-78.
- [00185] [6] Hipp, M. S., Park, S. H., and Hartl, F. U. (2014) Proteostasis impairment in protein-misfolding and -aggregation diseases, *Trends in Cell Biology* 24, 506-514.
- [00186] [7] Powers, E. T., Morimoto, R. I., Dillin, A., Kelly, J. W., and Balch, W. E. (2009) Biological and chemical approaches to diseases of proteostasis deficiency, *Annu Rev Biochem* 78, 959-991.
- [00187] [8] Chiti, F., and Dobson, C. M. (2006) Protein misfolding, functional amyloid, and human disease, *Annu Rev Biochem* 75, 333-366.
- [00188] [9] Aguzzi, A., and O'Connor, T. (2010) Protein aggregation diseases: pathogenicity and therapeutic perspectives, *Nature Reviews. Drug Discovery* 9, 237-248.
- [00189] [10] Jensen, M. B., and Jasper, H. (2014) Mitochondrial proteostasis in the control of aging and longevity, *Cell Metabolism* 20, 214-225.
- [00190] [11] Kaye, R., Head, E., Thompson, J. L., McIntire, T. M., Milton, S. C., Cotman, C. W., and Glabe, C. G. (2003) Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis, *Science* 300, 486-489.
- [00191] [12] Haass, C., and Selkoe, D. J. (2007) Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid beta-peptide, *Nat Rev Mol Cell Biol* 8, 101-112.
- [00192] [13] Benilova, I., Karran, E., and De Strooper, B. (2012) The toxic A β oligomer and Alzheimer's disease: an emperor in need of clothes, *Nat. Neurosci.* 15, 349.
- [00193] [14] Kim, Y. E., Hosp, F., Frottin, F., Ge, H., Mann, M., Hayer-Hartl, M., and Hartl, F. U. (2016) Soluble Oligomers of PolyQ-Expanded Huntingtin Target a Multiplicity of Key Cellular Factors, *Molecular Cell* 63, 951-964.

- [00194] [15] Ross, C. A., and Poirier, M. A. (2005) What is the role of protein aggregation in neurodegeneration?, *Nat. Rev. Mol. Cell Biol.* 6, 891.
- [00195] [16] Gupta, R., Kasturi, P., Bracher, A., Loew, C., Zheng, M., Villella, A., Garza, D., Hartl, F. U., and Raychaudhuri, S. (2011) Firefly luciferase mutants as sensors of proteome stress, *Nat Methods* 8, 879-884.
- [00196] [17] Winkler, J., Seybert, A., Konig, L., Pruggnaller, S., Haselmann, U., Sourjik, V., Weiss, M., Frangakis, A. S., Mogk, A., and Bukau, B. (2010) Quantitative and spatio-temporal features of protein aggregation in *Escherichia coli* and consequences on protein quality control and cellular ageing, *EMBO J* 29, 910-923.
- [00197] [18] Hsieh, T. Y., Nillegoda, N. B., Tyedmers, J., Bukau, B., Mogk, A., and Kramer, G. (2014) Monitoring Protein Misfolding by Site-Specific Labeling of Proteins In Vivo, *PloS one* 9.
- [00198] [19] Ignatova, Z., and Gierasch, L. M. (2004) Monitoring protein stability and aggregation in vivo by real-time fluorescent labeling, *P Natl Acad Sci USA* 101, 523-528.
- [00199] [20] Ramdhan, Y. M., Polling, S., Chia, C. P. Z., Ng, I. H. W., Ormsby, A. R., Croft, N. P., Purcell, A. W., Bogoyevitch, M. A., Ng, D. C. H., Gleeson, P. A., and Hatters, D. M. (2012) Tracking protein aggregation and mislocalization in cells with flow cytometry, *Nat Methods* 9, 467-U476.
- [00200] [21] Nath, S., Meuvlis, J., Hendrix, J., Carl, S. A., and Engelborghs, Y. (2010) Early aggregation steps in alpha-synuclein as measured by FCS and FRET: evidence for a contagious conformational change, *Biophys J* 98, 1302-1311.
- [00201] [22] Wood, R. J., Ormsby, A. R., Radwan, M., Cox, D., Sharma, A., Vopel, T., Ebbinghaus, S., Oliveberg, M., Reid, G. E., Dickson, A., and Hatters, D. M. (2018) A biosensor-based framework to measure latent proteostasis capacity, *Nat Commun* 9, 287.
- [00202] [23] Kitamura, A., Nagata, K., and Kinjo, M. (2015) Conformational Analysis of Misfolded Protein Aggregation by FRET and Live-Cell Imaging Techniques, *Int J Mol Sci* 16, 6076-6092.
- [00203] [24] Walker, C. L., Lukyanov, K. A., Yampolsky, I. V., Mishin, A. S., Bommarius, A. S., Duraj-Thatte, A. M., Azizi, B., Tolbert, L. M., and Solntsev, K. M. (2015) Fluorescence imaging using synthetic GFP chromophores, *Curr Opin Chem Biol* 27, 64-74.

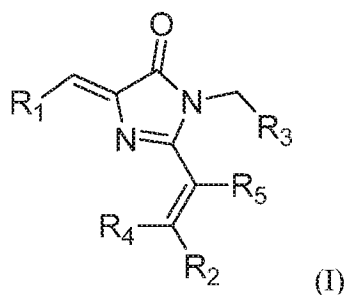
- [00204] [25] Tsien, R. Y. (1998) The green fluorescent protein, *Annual Review of Biochemistry* 67, 509-544.
- [00205] [26] Zimmer, M. (2002) Green fluorescent protein (GFP): applications, structure, and related photophysical behavior, *Chem Rev* 102, 759-781.
- [00206] [27] Subach, F. V., and Verkhusha, V. V. (2012) Chromophore transformations in red fluorescent proteins, *Chem Rev* 112, 4308-4327.
- [00207] [28] Baranov, M. S., Lukyanov, K. A., Borissova, A. O., Shamir, J., Kosenkov, D., Slipchenko, L. V., Tolbert, L. M., Yampolsky, I. V., and Solntsev, K. M. (2012) Conformationally locked chromophores as models of excited-state proton transfer in fluorescent proteins, *J Am Chem Soc* 134, 6025-6032.
- [00208] [29] Baldridge, A., Samanta, S. R., Jayaraj, N., Ramamurthy, V., and Tolbert, L. M. (2011) Steric and electronic effects in capsule-confined green fluorescent protein chromophores, *J Am Chem Soc* 133, 712-715.
- [00209] [30] Williams, D. E., Dolgoplova, E. A., Pellechia, P. J., Palukoshka, A., Wilson, T. J., Tan, R., Maier, J. M., Greytak, A. B., Smith, M. D., Krause, J. A., and Shustova, N. B. (2015) Mimic of the green fluorescent protein beta-barrel: photophysics and dynamics of confined chromophores defined by a rigid porous scaffold, *J Am Chem Soc* 137, 2223-2226.
- [00210] [31] Carayon, C., Ghodbane, A., Gibot, L., Dumur, R., Wang, J., Saffon, N., Rols, M. P., Solntsev, K. M., and Fery-Forgues, S. (2016) Conjugates of Benzoxazole and GFP Chromophore with Aggregation-Induced Enhanced Emission: Influence of the Chain Length on the Formation of Particles and on the Dye Uptake by Living Cells, *Small* 12, 6602-6612.
- [00211] [32] Baldridge, A., Feng, S., Chang, Y. T., and Tolbert, L. M. (2011) Recapture of GFP chromophore fluorescence in a protein host, *ACS Comb Sci* 13, 214-217.
- [00212] [33] Paige, J. S., Wu, K. Y., and Jaffrey, S. R. (2011) RNA mimics of green fluorescent protein, *Science* 333, 642-646.
- [00213] [34] Filonov, G. S., Moon, J. D., Svensen, N., and Jaffrey, S. R. (2014) Broccoli: rapid selection of an RNA mimic of green fluorescent protein by fluorescence-based selection and directed evolution, *J Am Chem Soc* 136, 16299-16308.
- [00214] [35] Warner, K. D., Sjekloca, L., Song, W., Filonov, G. S., Jaffrey, S. R., and Ferre-D'Amare, A. R. (2017) A homodimer interface without base pairs in an RNA mimic of red fluorescent protein, *Nat Chem Biol* 13, 1195-1201.

- [00215] [36] Feng, G., Luo, C., Yi, H., Yuan, L., Lin, B., Luo, X., Hu, X., Wang, H., Lei, C., Nie, Z., and Yao, S. (2017) DNA mimics of red fluorescent proteins (RFP) based on G-quadruplex-confined synthetic RFP chromophores, *Nucleic Acids Res* 45, 10380-10392.
- [00216] [37] Kellenberger, C. A., Wilson, S. C., Sales-Lee, J., and Hammond, M. C. (2013) RNA-based fluorescent biosensors for live cell imaging of second messengers cyclic di-GMP and cyclic AMP-GMP, *J Am Chem Soc* 135, 4906-4909.
- [00217] [38] Tsien, R. Y., and Miyawaki, A. (1998) Seeing the machinery of live cells, *Science* 280, 1954-1955.
- [00218] [39] Rodriguez, E. A., Campbell, R. E., Lin, J. Y., Lin, M. Z., Miyawaki, A., Palmer, A. E., Shu, X., Zhang, J., and Tsien, R. Y. (2017) The Growing and Glowing Toolbox of Fluorescent and Photoactive Proteins, *Trends Biochem Sci* 42, 111-129.
- [00219] [40] Ando, R., Hama, H., Yamamoto-Hino, M., Mizuno, H., and Miyawaki, A. (2002) An optical marker based on the UV-induced green-to-red photoconversion of a fluorescent protein, *Proc Natl Acad Sci U S A* 99, 12651-12656.
- [00220] [41] Tomura, M., Yoshida, N., Tanaka, J., Karasawa, S., Miwa, Y., Miyawaki, A., and Kanagawa, O. (2008) Monitoring cellular movement in vivo with photoconvertible fluorescence protein "Kaede" transgenic mice, *Proc Natl Acad Sci U S A* 105, 10871-10876.
- [00221] [42] Lashuel, H. A., Overk, C. R., Oueslati, A., and Masliah, E. (2013) The many faces of alpha-synuclein: from structure and toxicity to therapeutic target, *Nat Rev Neurosci* 14, 38-48.
- [00222] [43] Qin, L., Vastl, J., and Gao, J. (2010) Highly sensitive amyloid detection enabled by thioflavin T dimers, *Molecular BioSystems* 6, 1791-1795.
- [00223] [44] Rosen, D. R., Siddique, T., Patterson, D., Figlewicz, D. A., Sapp, P., Hentati, A., Donaldson, D., Goto, J., O'Regan, J. P., Deng, H. X., and et al. (1993) Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis, *Nature* 362, 59-62.
- [00224] [45] Dangoumau, A., Verschueren, A., Hammouche, E., Papon, M. A., Blasco, H., Cherpi-Antar, C., Pouget, J., Corcia, P., Andres, C. R., and Vourc'h, P. (2014) Novel SOD1 mutation p.V31A identified with a slowly progressive form of amyotrophic lateral sclerosis, *Neurobiol Aging* 35, 266 e261-264.

- [00225] [46] Kaganovich, D., Kopito, R., and Frydman, J. (2008) Misfolded proteins partition between two distinct quality control compartments, *Nature* 454, 1088-U1036.
- [00226] [47] Sontag, E. M., Vonk, W. I. M., and Frydman, J. (2014) Sorting out the trash: the spatial nature of eukaryotic protein quality control, *Curr Opin Cell Biol* 26, 139-146.
- [00227] [48] Long, M. J. C., Poganik, J. R., and Aye, Y. (2016) On-Demand Targeting: Investigating Biology with Proximity-Directed Chemistry, *J Am Chem Soc* 138, 3610-3622.
- [00228] [49] Lin, H. Y., Haegele, J. A., Disare, M. T., Lin, Q. S., and Aye, Y. (2015) A Generalizable Platform for Interrogating Target- and Signal-Specific Consequences of Electrophilic Modifications in Redox-Dependent Cell Signaling, *J Am Chem Soc* 137, 6232-6244.
- [00229] [50] Keppler, A., Gendreizig, S., Gronemeyer, T., Pick, H., Vogel, H., and Johnsson, K. (2003) A general method for the covalent labeling of fusion proteins with small molecules in vivo, *Nat Biotechnol* 21, 86-89.
- [00230] [51] Gautier, A., Juillerat, A., Heinis, C., Correa, I. R., Jr., Kindermann, M., Beaufils, F., and Johnsson, K. (2008) An engineered protein tag for multiprotein labeling in living cells, *Chemistry & Biology* 15, 128-136.
- [00231] [52] Xue, L., Karpenko, I. A., Hiblot, J., and Johnsson, K. (2015) Imaging and manipulating proteins in live cells through covalent labeling, *Nat Chem Biol* 11, 917-923.
- [00232] [53] Jung, K. H., Fares, M., Grainger, L. S., Wolstenholme, C. H., Hou, A., Liu, Y., and Zhang, X. (2018) A SNAP-tag fluorogenic probe mimicking the chromophore of the red fluorescent protein Kaede, *Org Biomol Chem*.
- [00233] [54] Dolomanov, O. V., Bourhis, L. J., Gildea, R. J., Howard, J. A. K., and Puschmann, H. (2009) OLEX2: a complete structure solution, refinement and analysis program, *J Appl Crystallogr* 42, 339-341.
- [00234] [55] Sheldrick, G. M. (2008) A short history of SHELX, *Acta Crystallogr A* 64, 112-122.

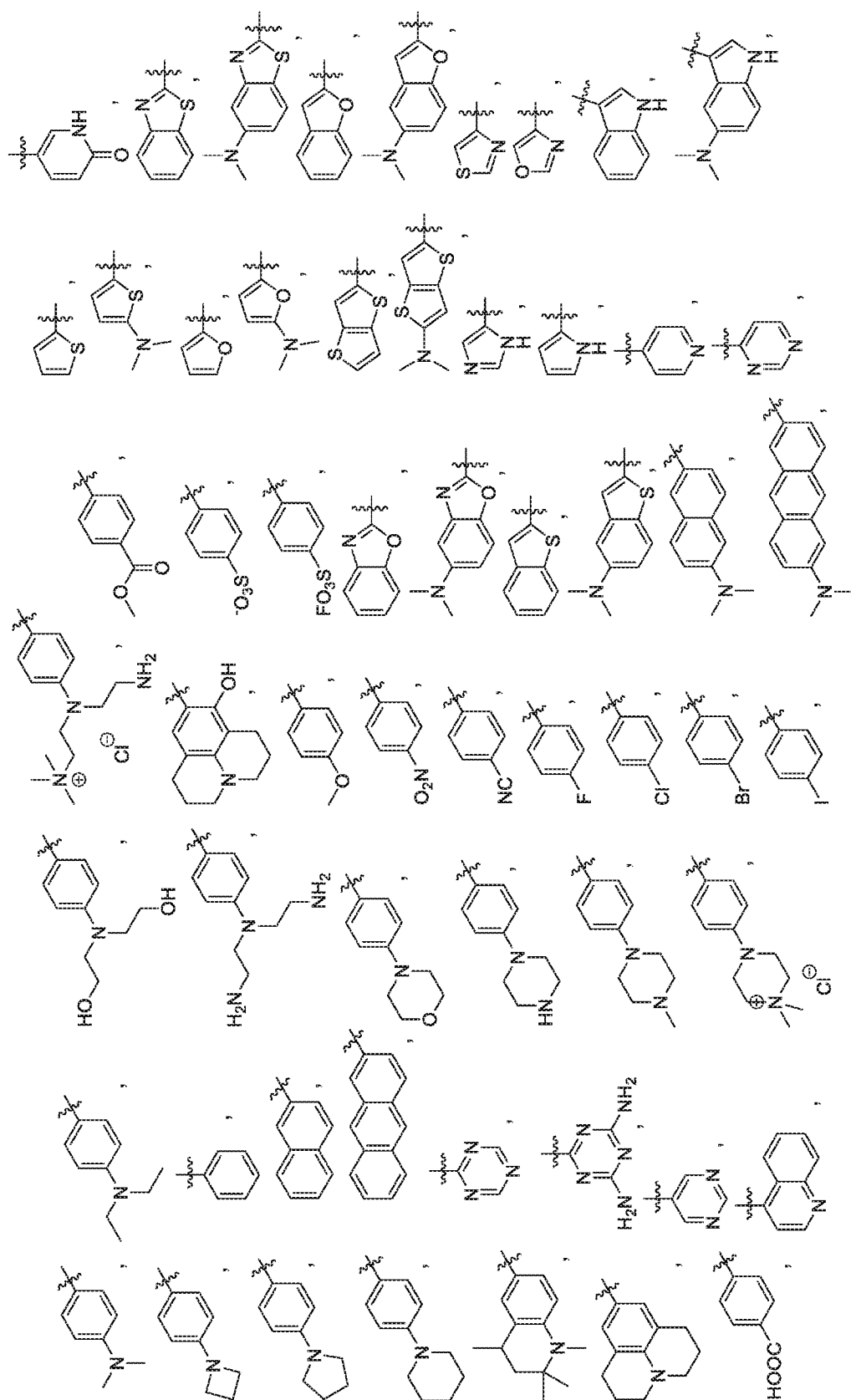
WE CLAIM:

1. A compound of Formula I:



wherein:

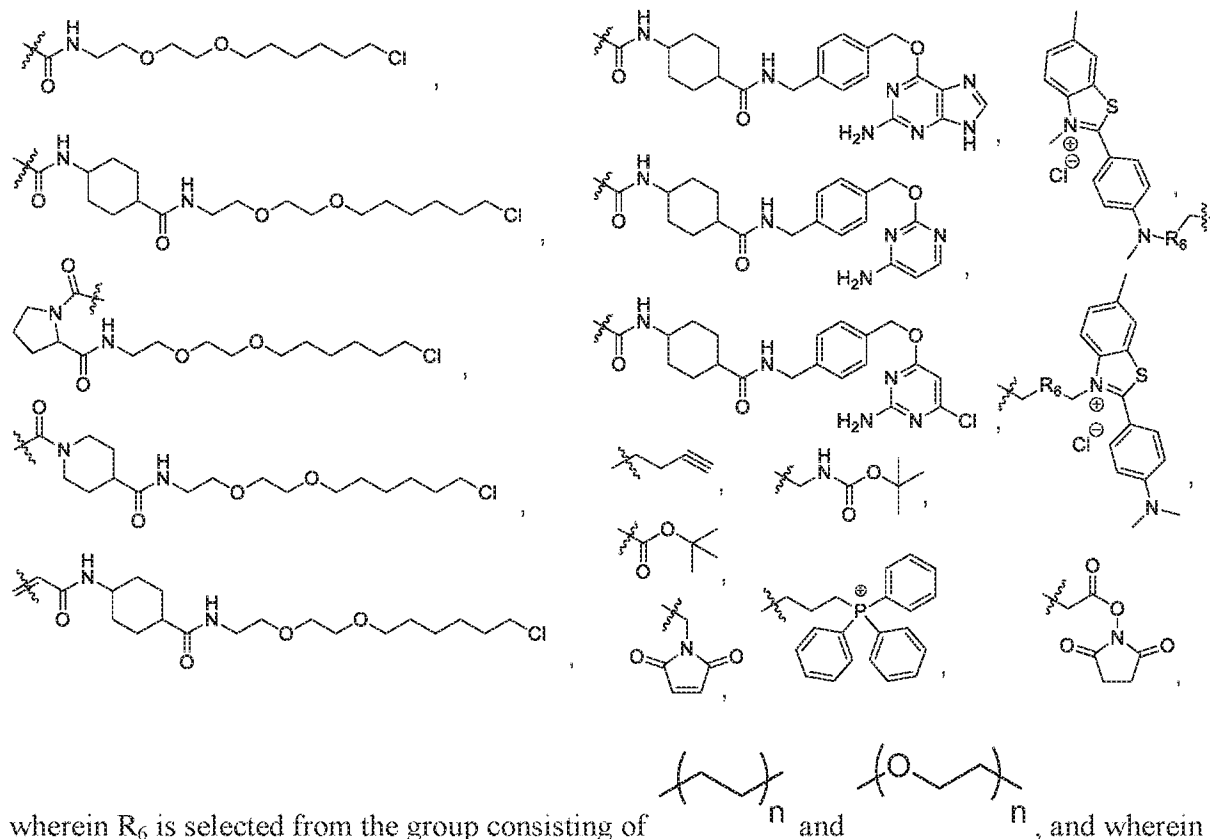
- R₁, R₂, and R₅ are independently selected from the moieties of Group 1, and R₄ is -H;
 R₁, R₂, and R₄ are independently selected from the moieties of Group 1, and R₅ is -H;
 R₁ and R₂ are independently selected from the moieties of Group 1, R₄ is -H, and R₅ is -H;
 R₁ and R₂ are independently selected from the moieties of Group 1, R₄ is -H and R₅ is -CH₃;
 R₁ and R₂ are independently selected from the moieties of Group 1, R₄ is -CH₃ and R₅ is -H; or
 R₁, R₂, R₄, and R₅ are independently selected from the moieties of Group 1; and
 wherein the moieties of Group 1 are selected from the group consisting of



wherein substituents may be at para-, meta-, or ortho- positions on the aromatic ring, even if shown in only a single position above; and

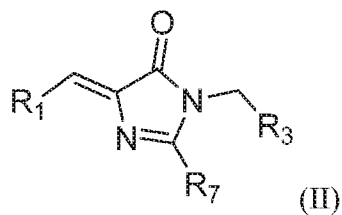
wherein R₃ is a directing moiety that binds and bioconjugates at least one biological target,

wherein R₃ is selected from the group consisting of:



wherein R_6 is selected from the group consisting of $-CH_2-$, $-CH_2CH_2-$, $-CH_2CH_2CH_2-$, $-CH_2CH_2CH_2CH_2-$, $-CH_2CH_2CH_2CH_2CH_2-$, and $-CH_2CH_2CH_2CH_2CH_2CH_2-$, and wherein n is 0, 1, 2, 3, 4, or 5.

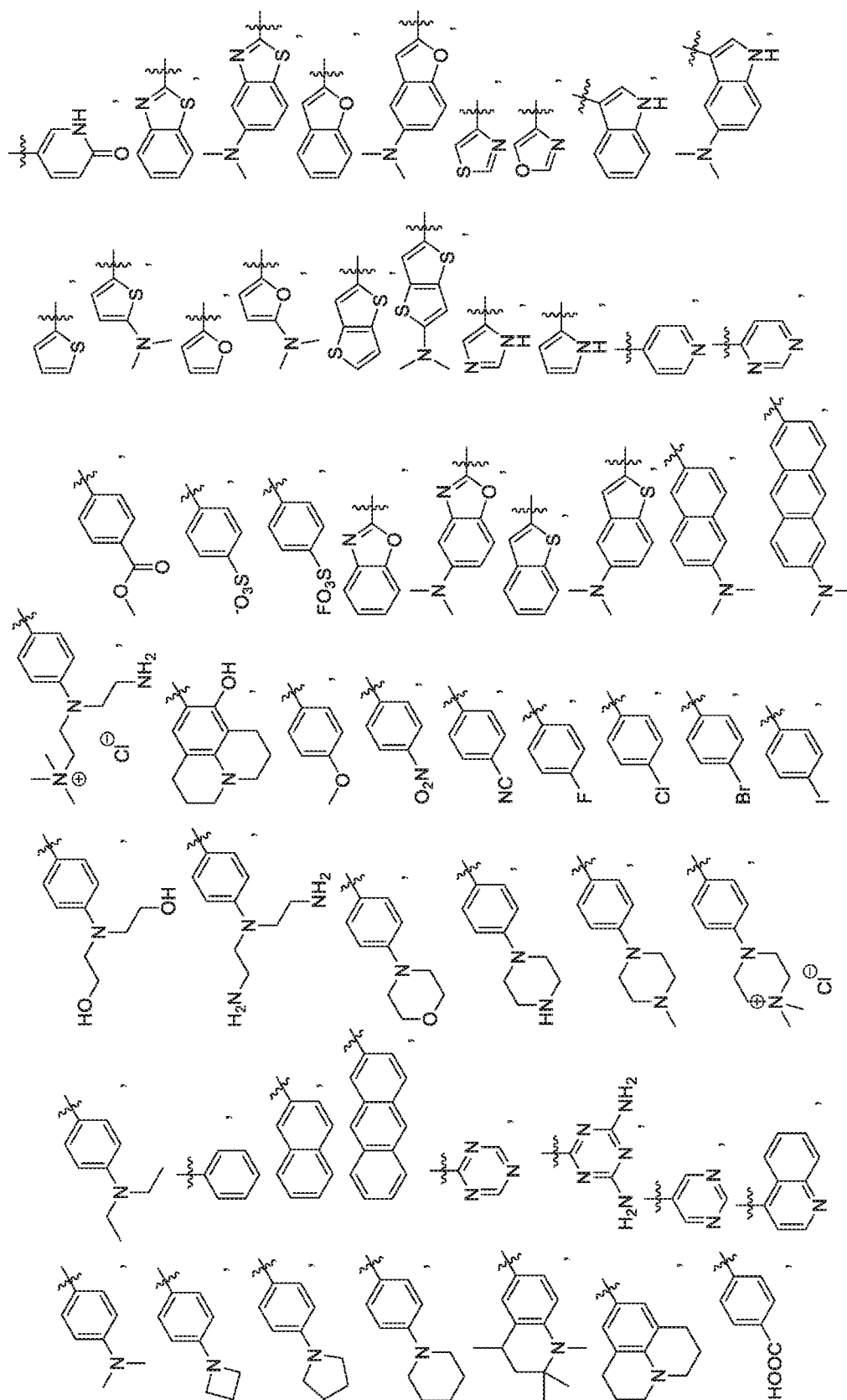
2. A compound of Formula II:



wherein

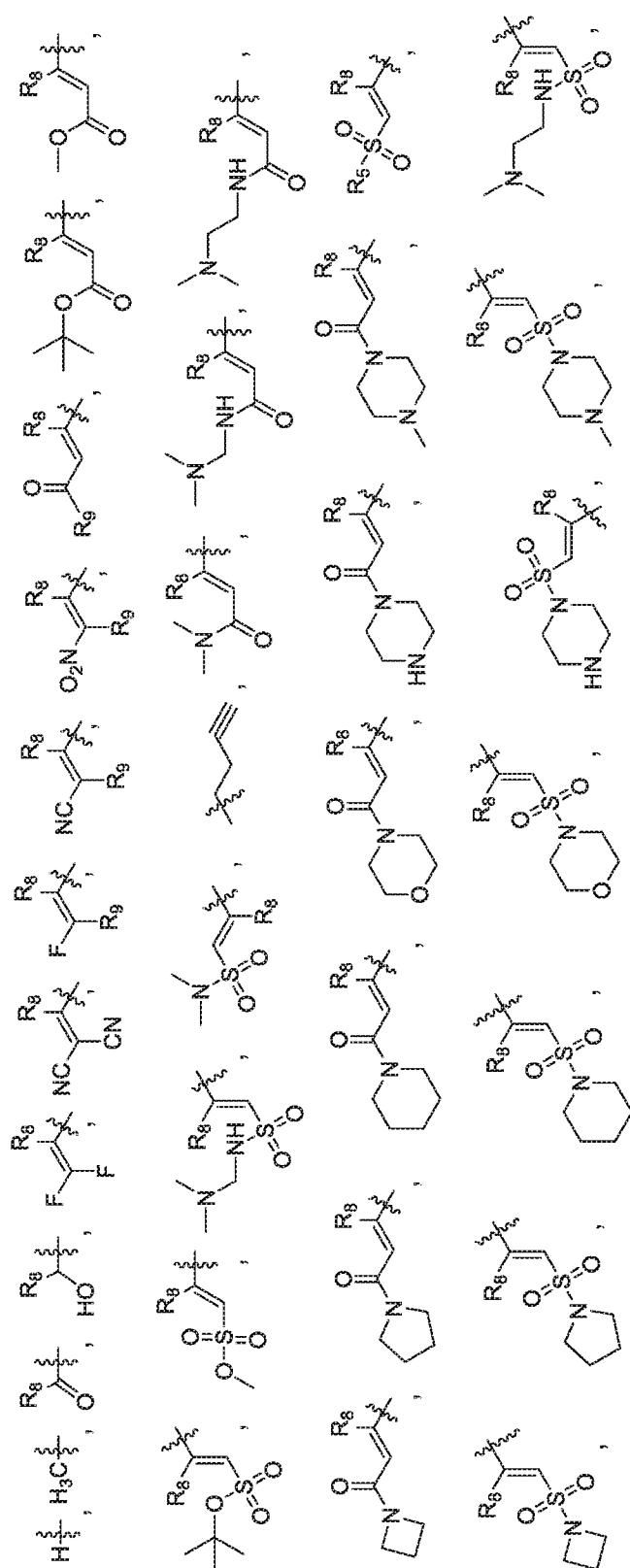
R₁ and R₇ are independently selected from a moiety of Group 1; or

R_1 is a moiety of Group 1, R_7 is a moiety of Group 2, and R_8 and R_9 are independently selected from $-H$, $-CH_3$, and a moiety of Group 1; wherein a moiety of Group 1 is selected from the group consisting of:

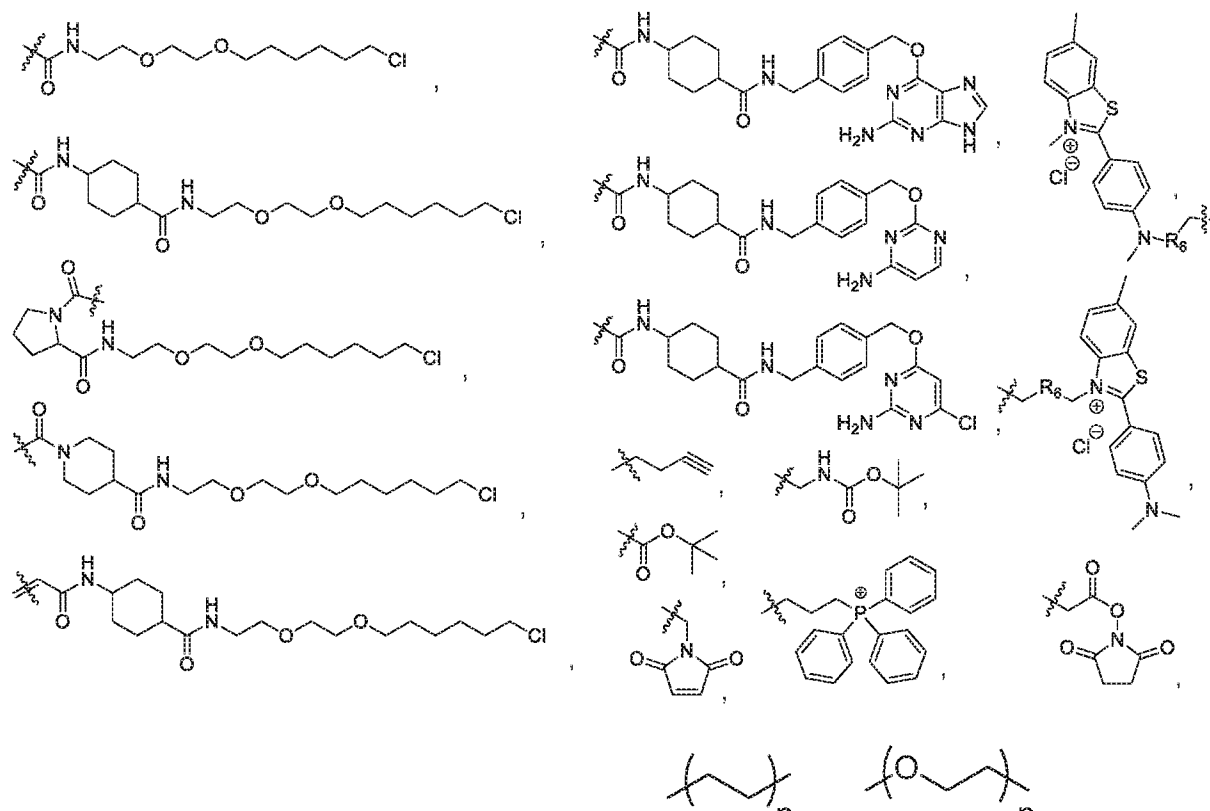


and wherein

a moiety of Group 2 is selected from the group consisting of



wherein substituents may be at para-, meta-, or ortho- positions on the aromatic ring of Group 1 and Group 2, even if shown in only a single position above; and wherein R3 is a directing moiety to bind and bioconjugate to biological targets, selected from the group consisting of:



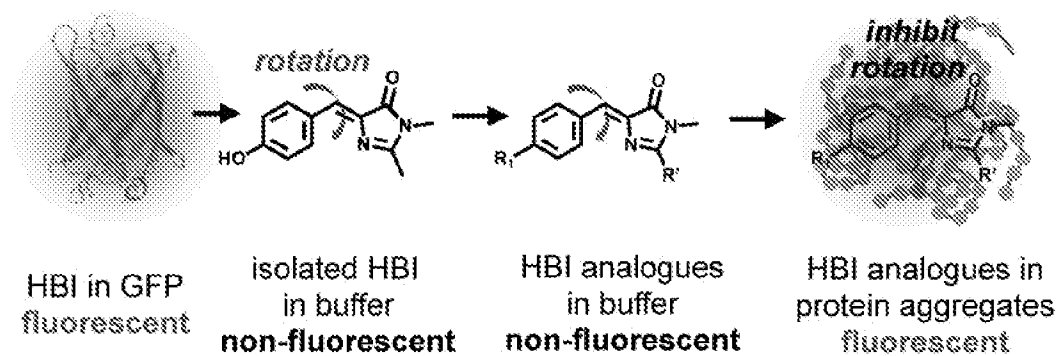
wherein, R₆ is selected from the group consisting of $\text{C}_n\text{H}_{2n+1}$ and $\text{C}_n\text{H}_{2n-1}$, and wherein n is 0, 1, 2, 3, 4, or 5.

3. A compound of Table 1.
4. A method for detecting at least one of aggregated protein, misfolded protein, and amyloid fiber in a protein of interest, comprising:
 - performing a first measurement of fluorescence intensity of a protein of interest;
 - adding to the protein of interest a fluorescent protein chromophore; and
 - performing a second measurement of fluorescence intensity of the protein of interest;wherein increased fluorescence is indicative of at least one of aggregation, misfolding, and amyloid fiber in the protein of interest.

5. The method of claim 4, wherein the fluorescent protein chromophore is a compound selected from any one of claims 1-3.
6. The method of claim 4, wherein the method is conducted in a cell in vivo.
7. The method of claim 4, wherein the method is conducted in vitro.
8. The method of claim 4, further comprising purifying the protein of interest prior to the first measurement.
9. The method of claim 4, wherein the fluorescent protein chromophore comprises a thioflavin-T guiding group or a tert-butyloxycarbonyl guiding group.
10. A method for detecting at least one of aggregated protein, misfolded protein, and amyloid fiber in a protein of interest, comprising:
 - performing a first measurement of fluorescence intensity of a standard protein;
 - adding to a protein of interest a fluorescent protein chromophore;
 - performing a measurement of fluorescence intensity of the protein of interest; and
 - comparing the fluorescence intensity of the standard protein with the fluorescence intensity of the protein of interest, wherein increased fluorescence of the protein of interest is indicative of at least one of aggregation, misfolding, and amyloid fiber in the protein of interest.
11. A kit for detecting at least one of aggregated protein, misfolded protein, and amyloid fiber, comprising:
 - one or more fluorescent protein chromophores of claims 1-3 of a known concentration in a stock solution, one or more standard protein samples that form aggregated protein, misfolded protein, and amyloid fiber and optionally, instructions for use in detecting at least one of one of the aggregated protein, misfolded protein, and amyloid fiber.
12. The kit of claim 11 wherein the stock solution is dimethyl sulfoxide or ethanol.

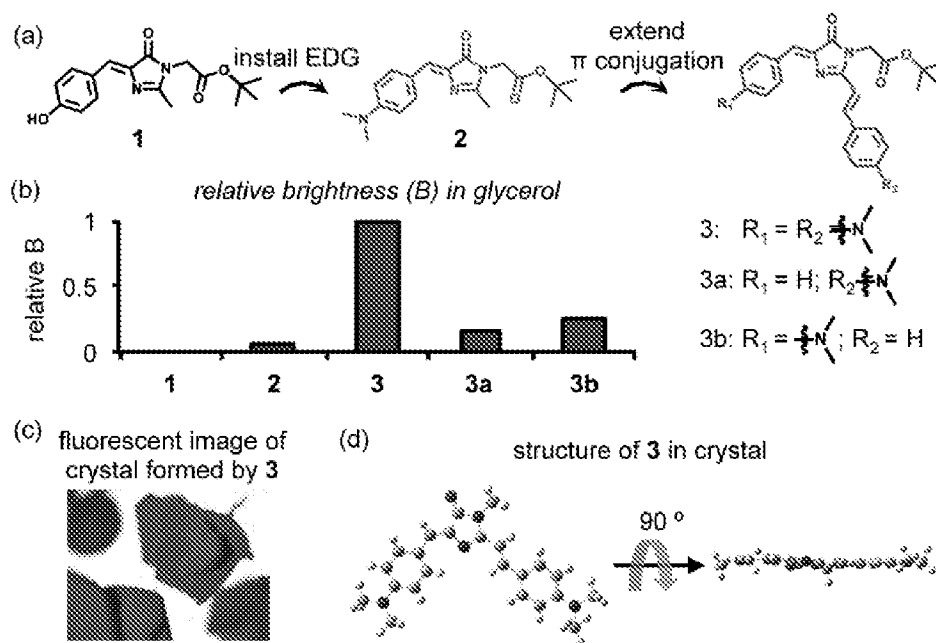
1/25

Figure 1



2/25

Figure 2



3/25

Figure 3

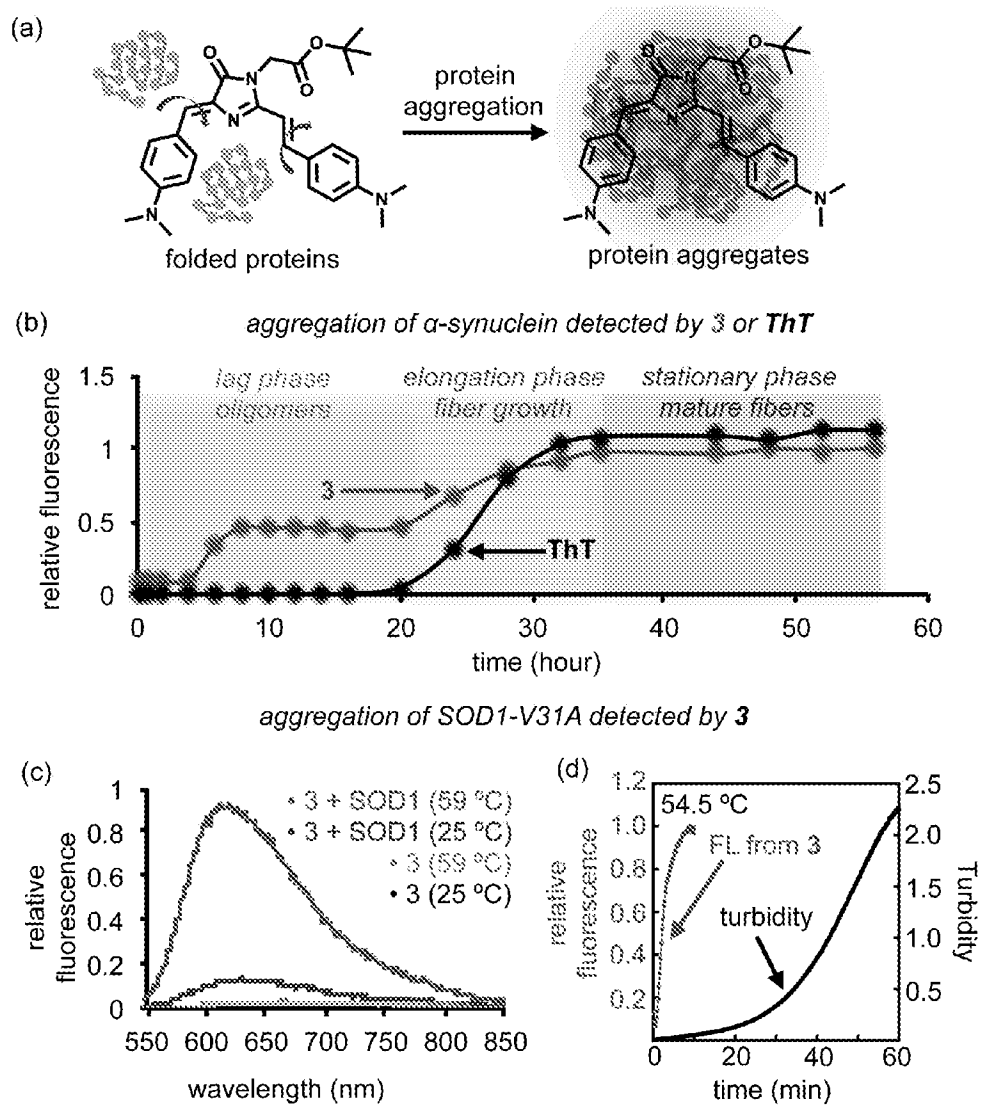


Figure 4

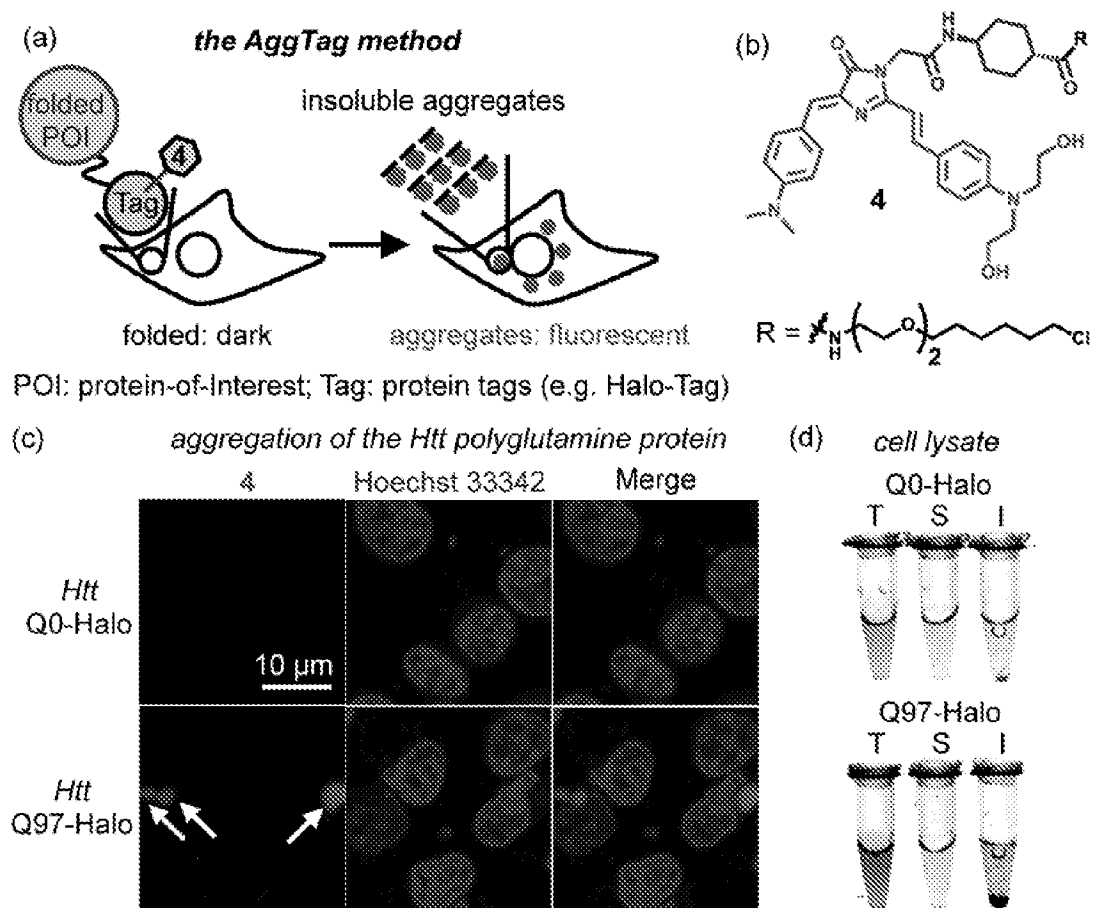
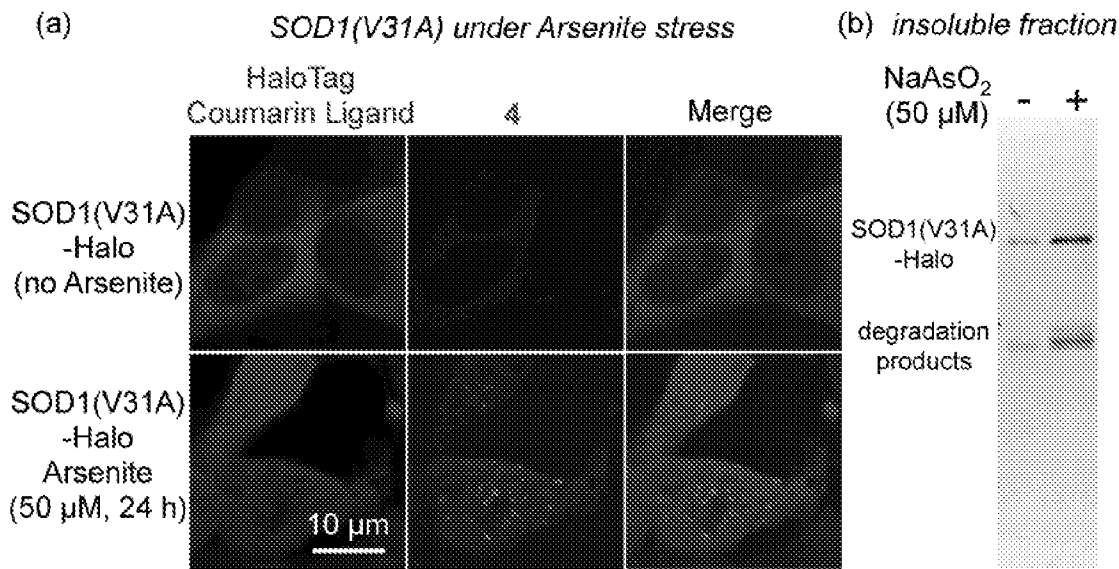
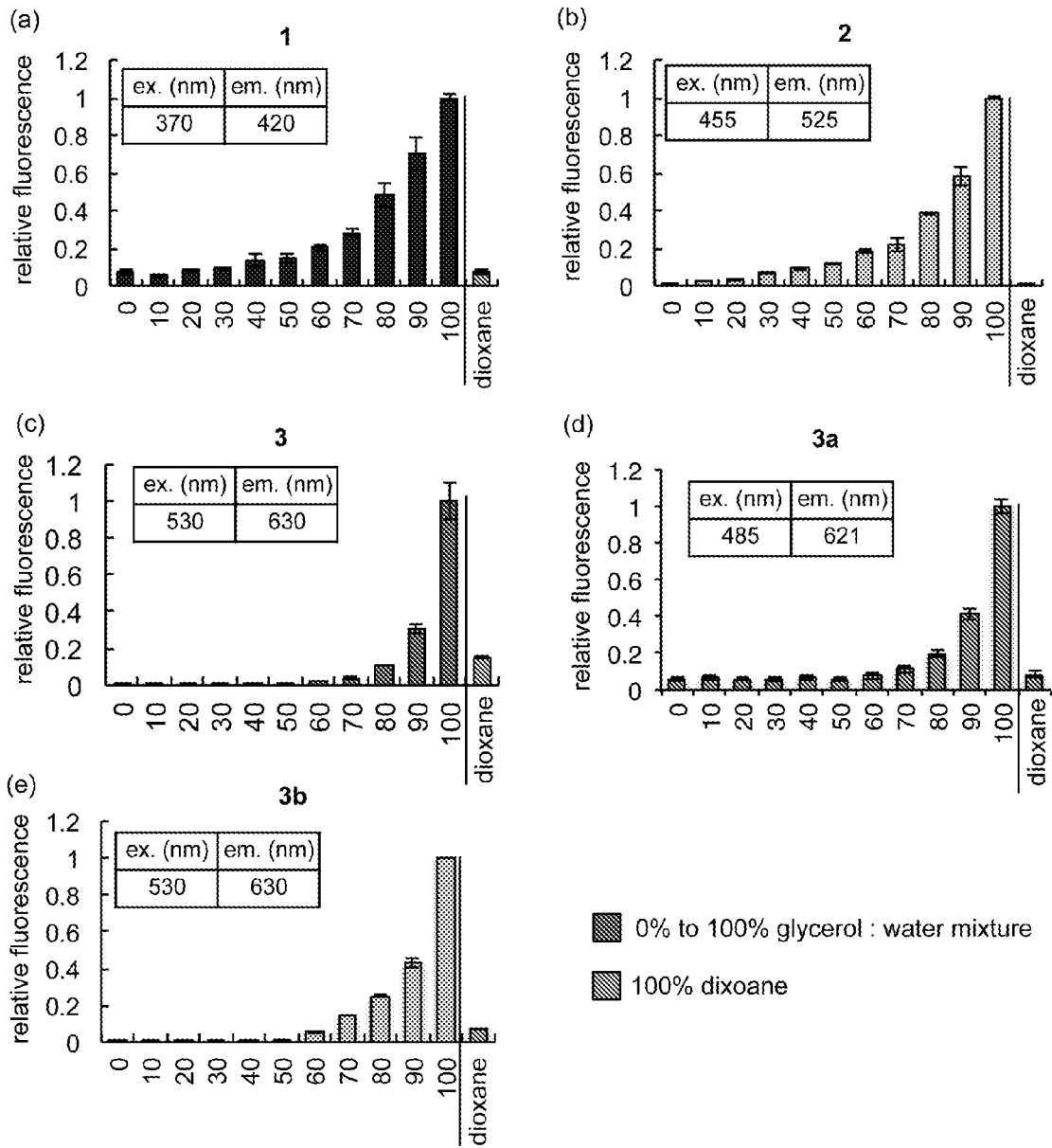


Figure 5



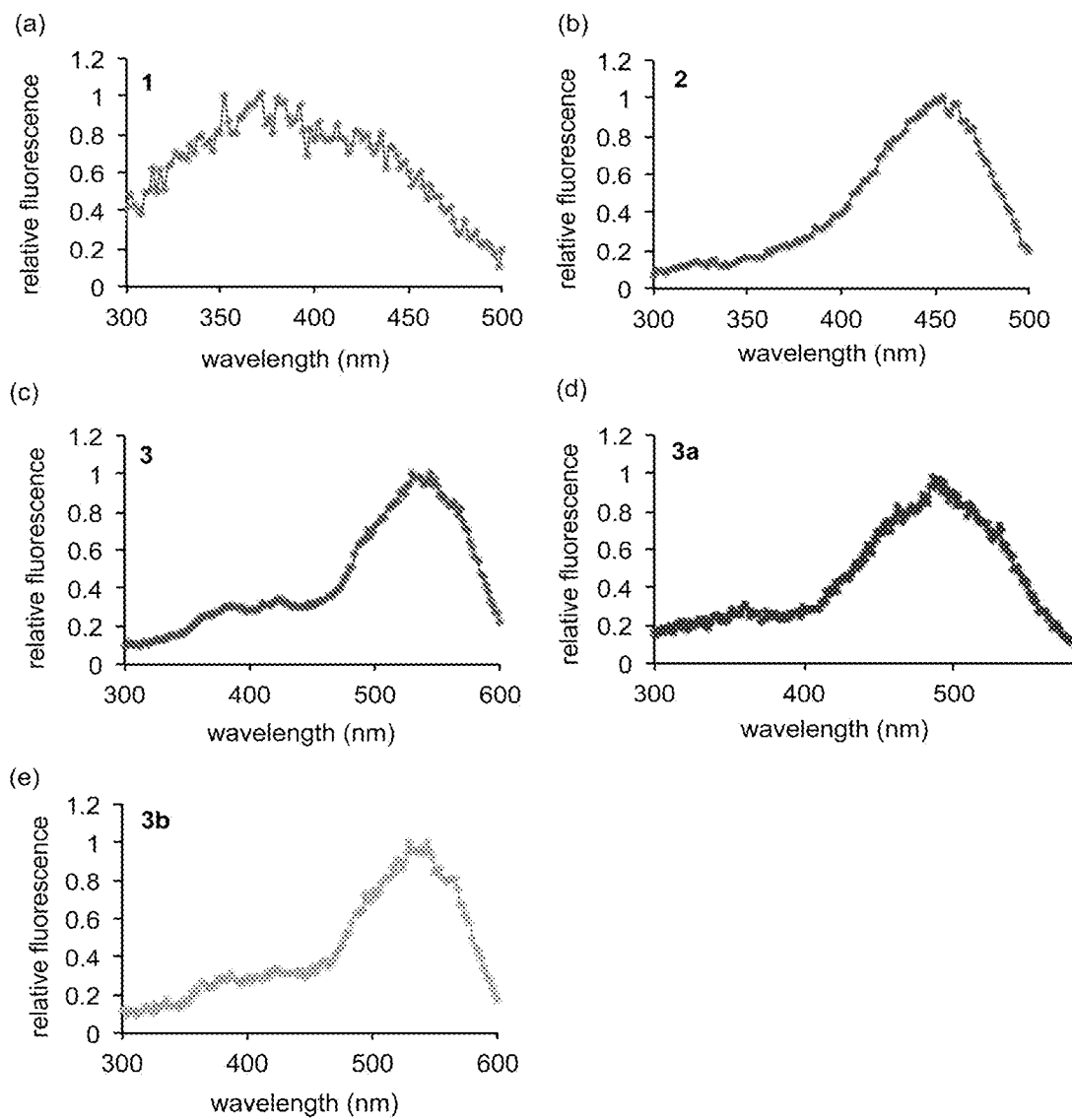
6/25

Figure 6



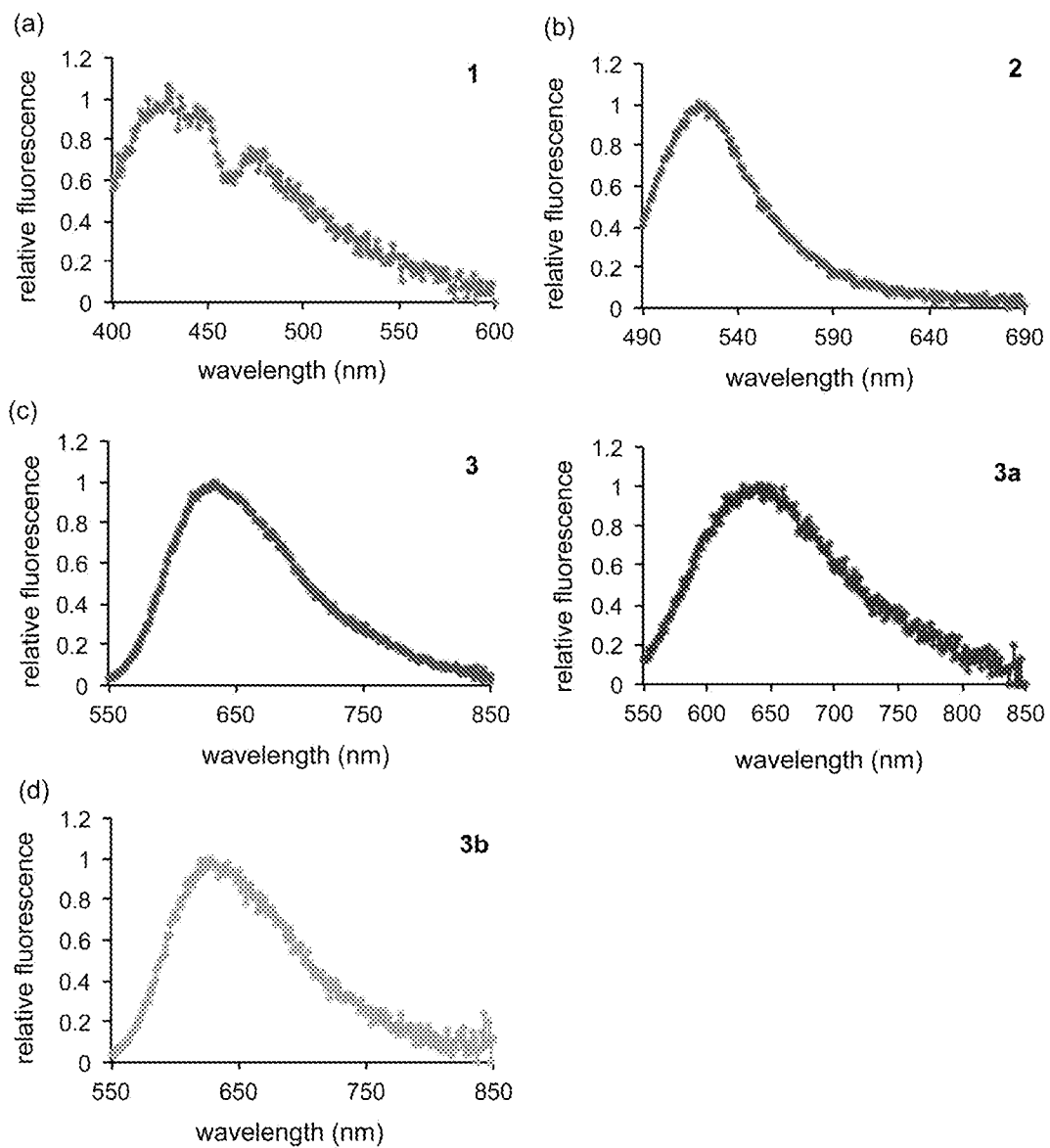
7/25

Figure 7



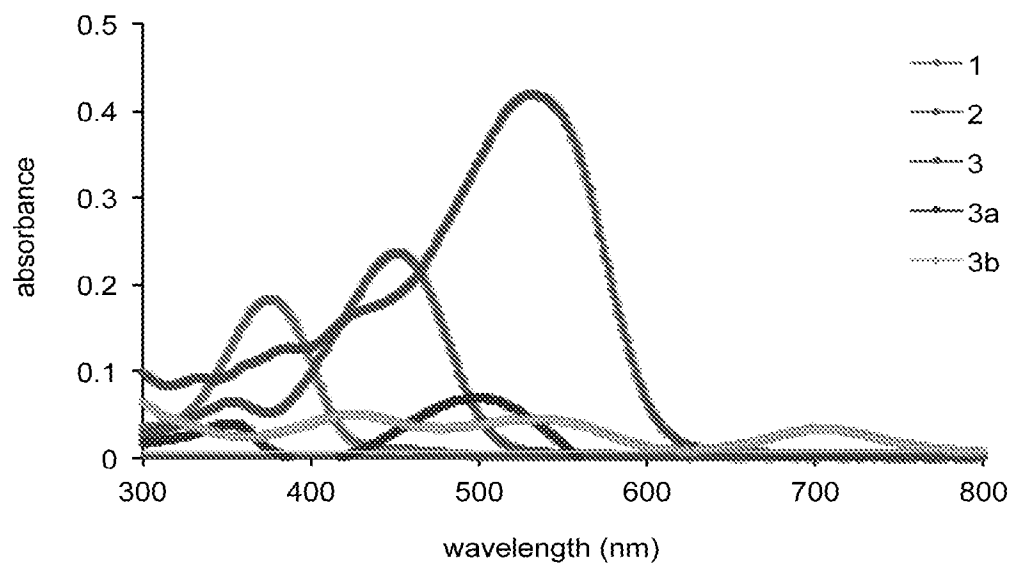
8/25

Figure 8



9/25

Figure 9



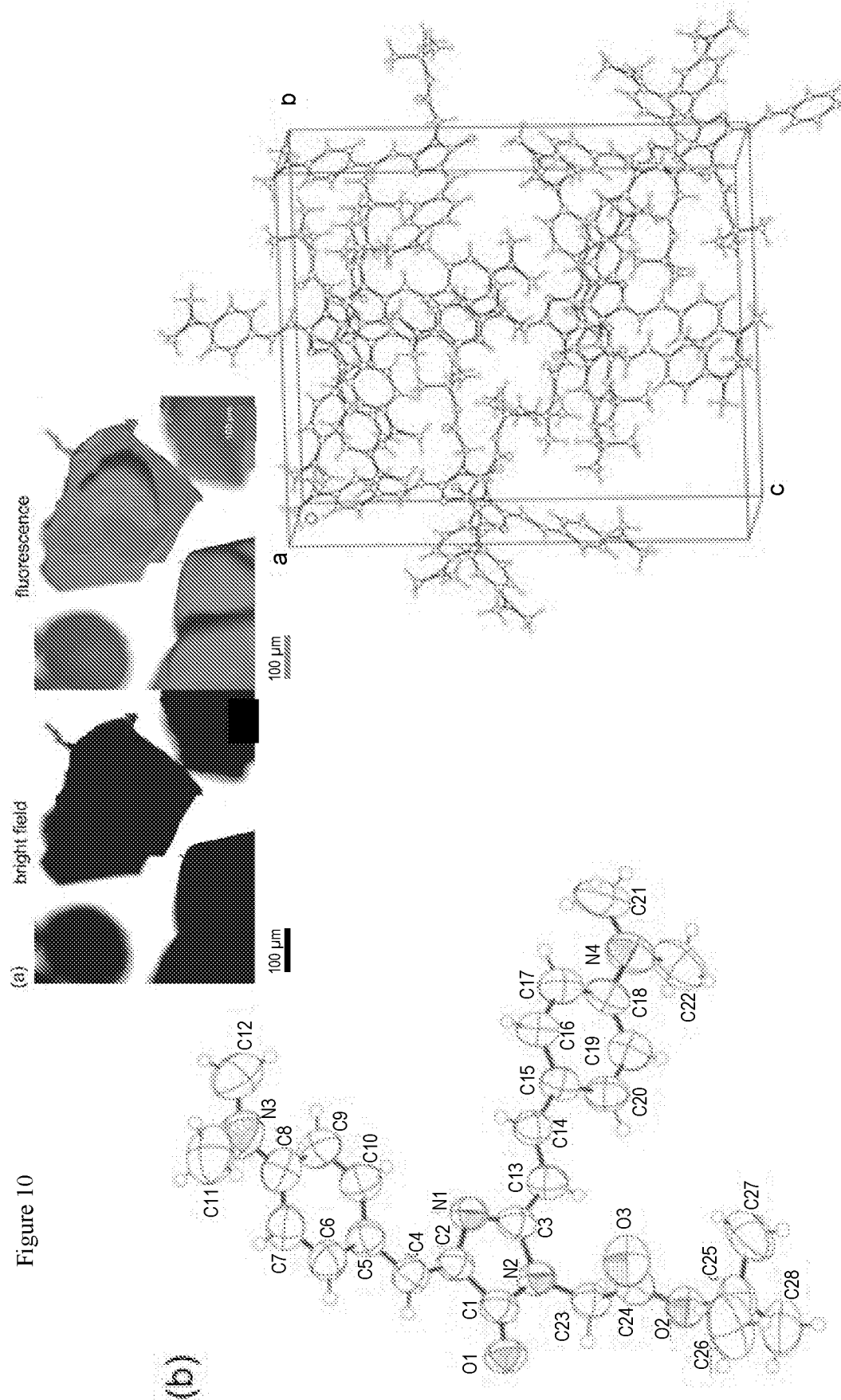


Figure 10

11/25

Figure 11

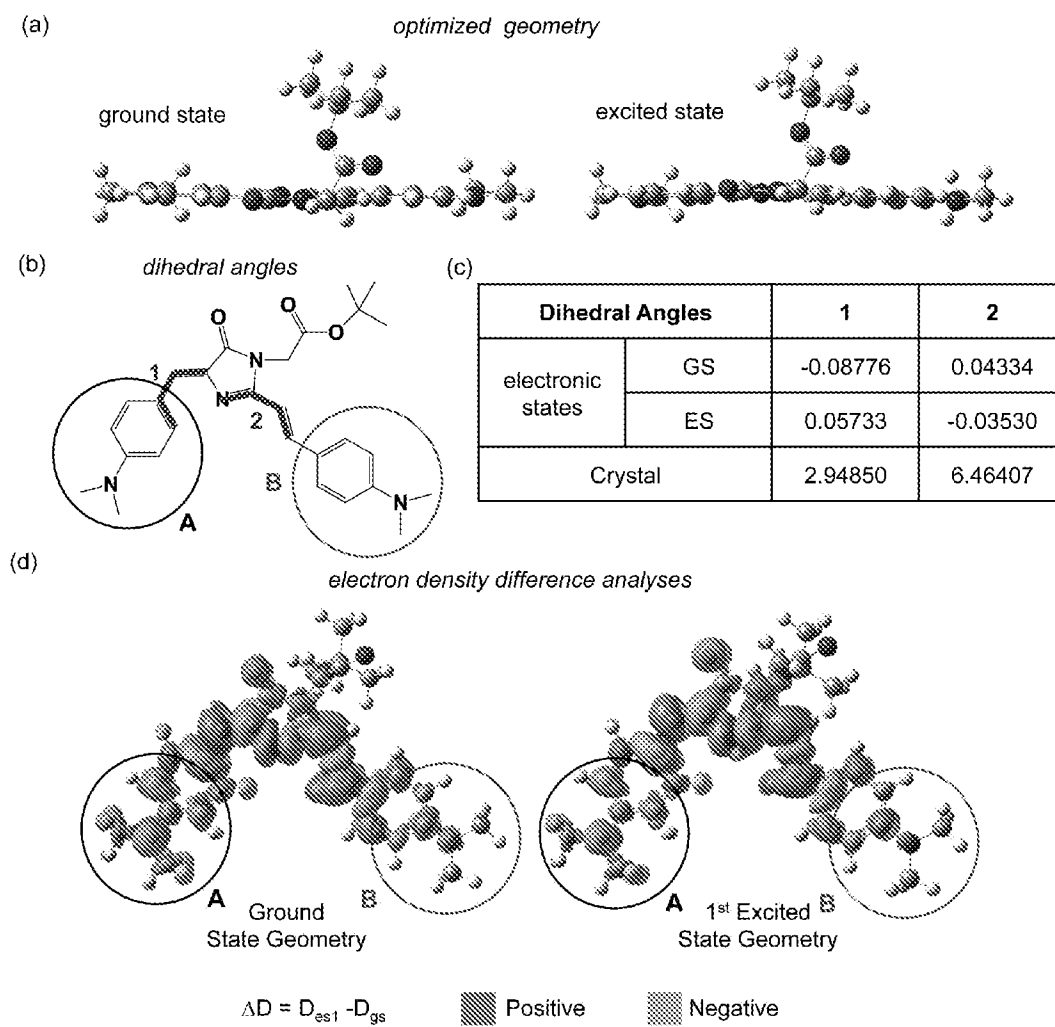


Figure 12

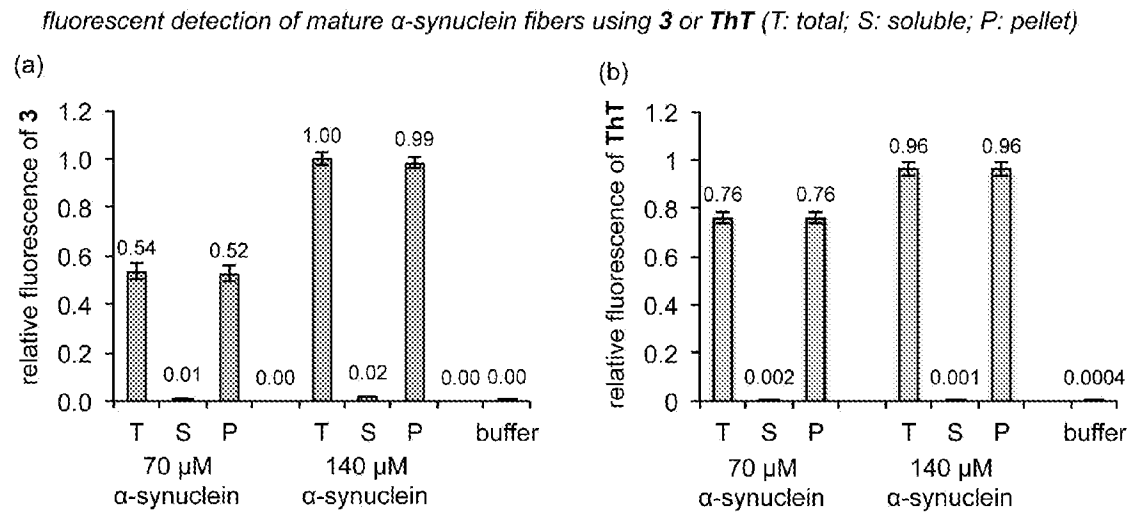
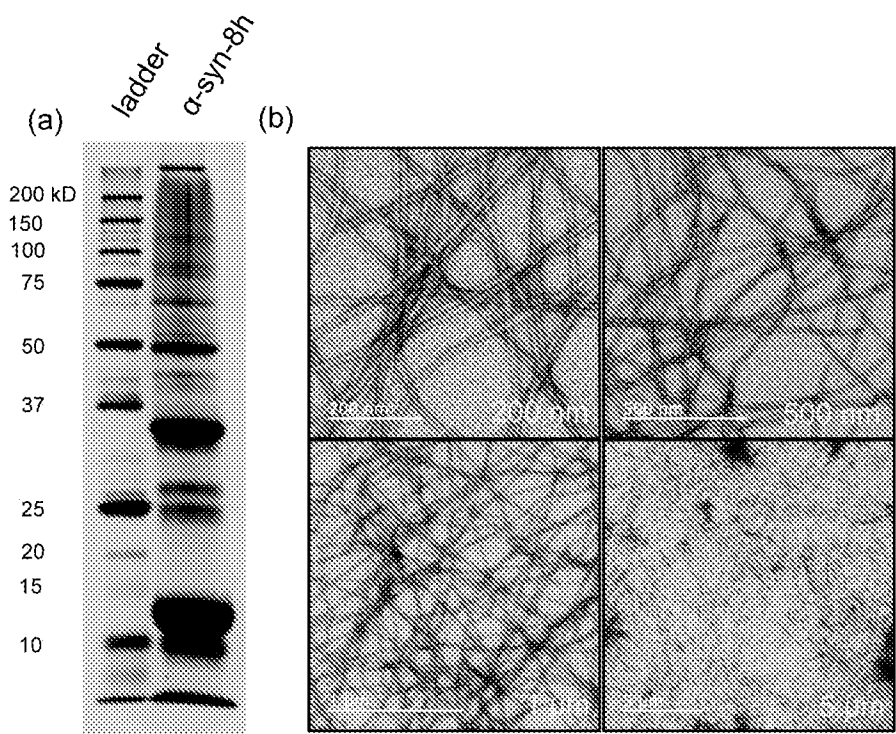


Figure 13



14/25

Figure 14

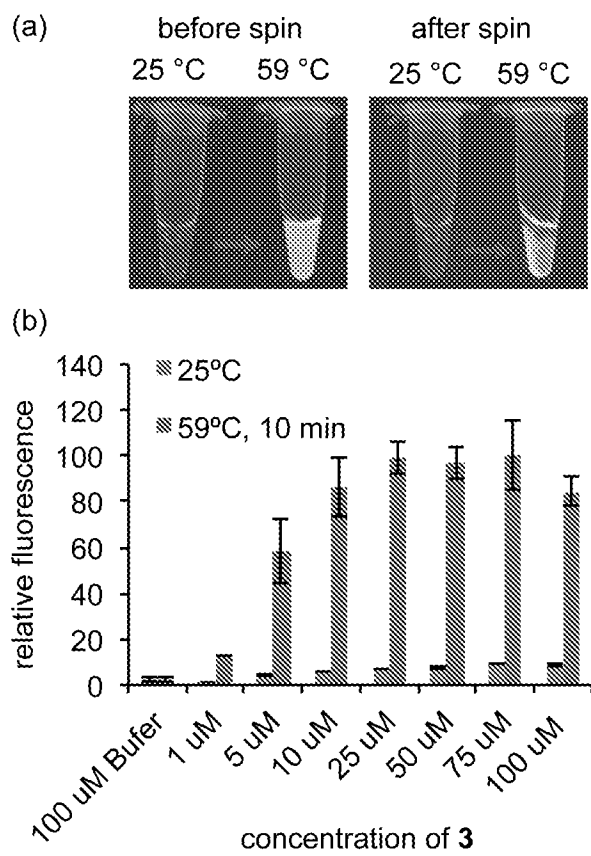


Figure 15

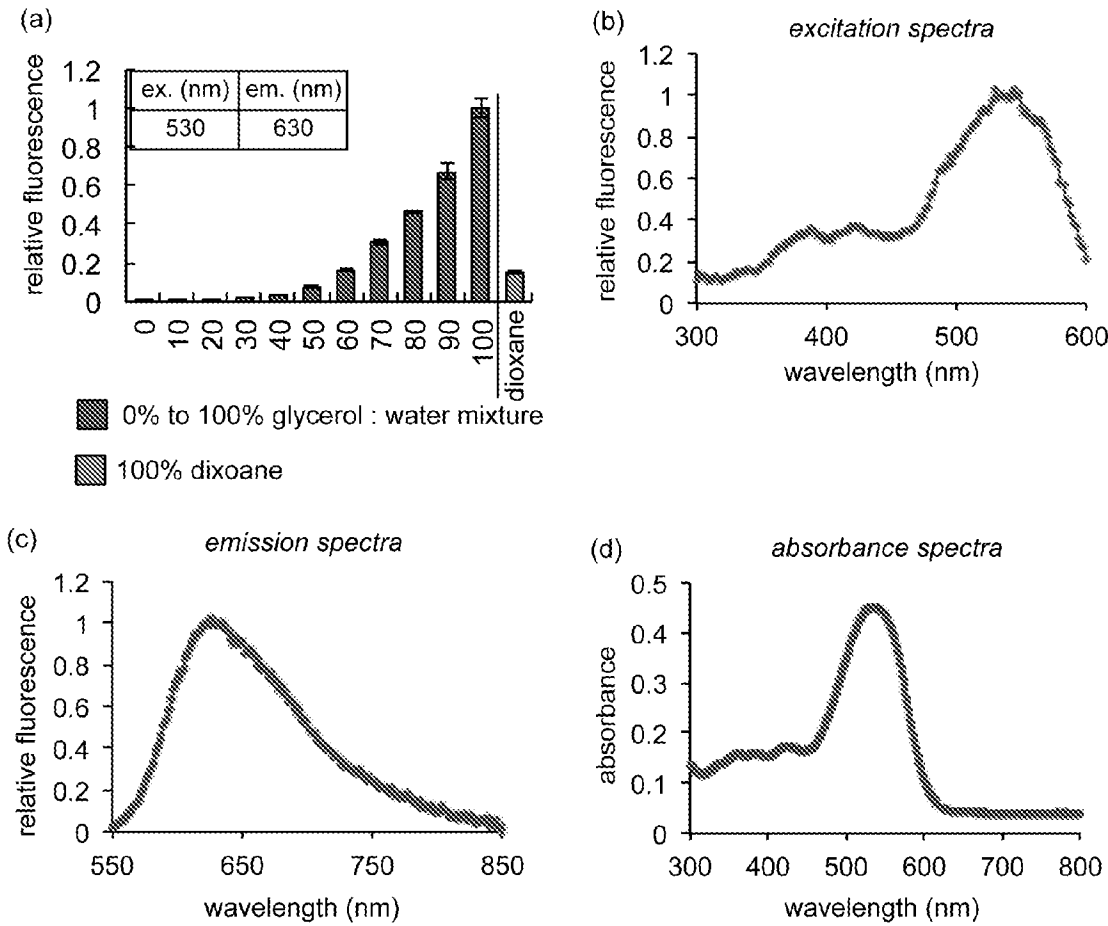
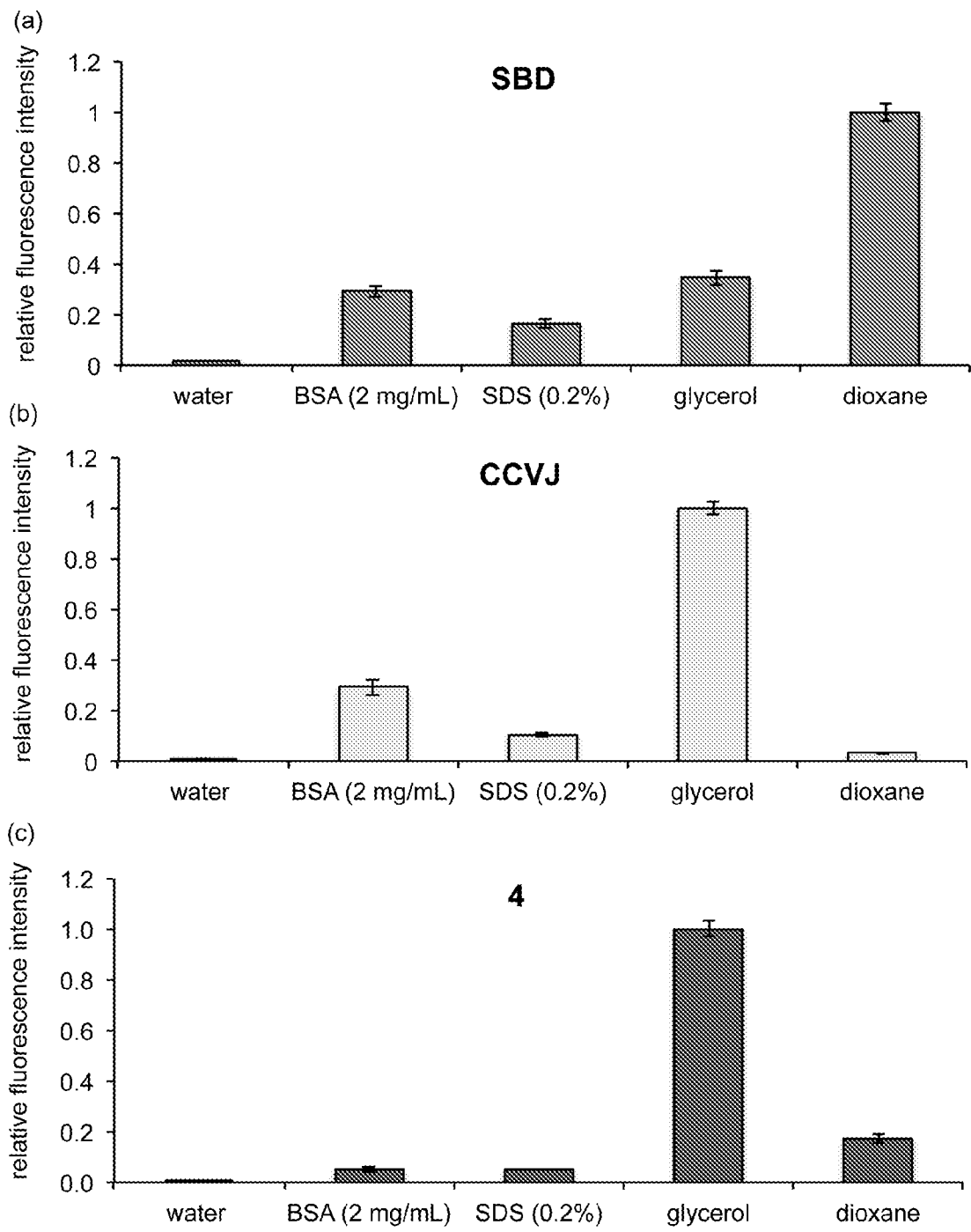


Figure 16



17/25

Figure 17

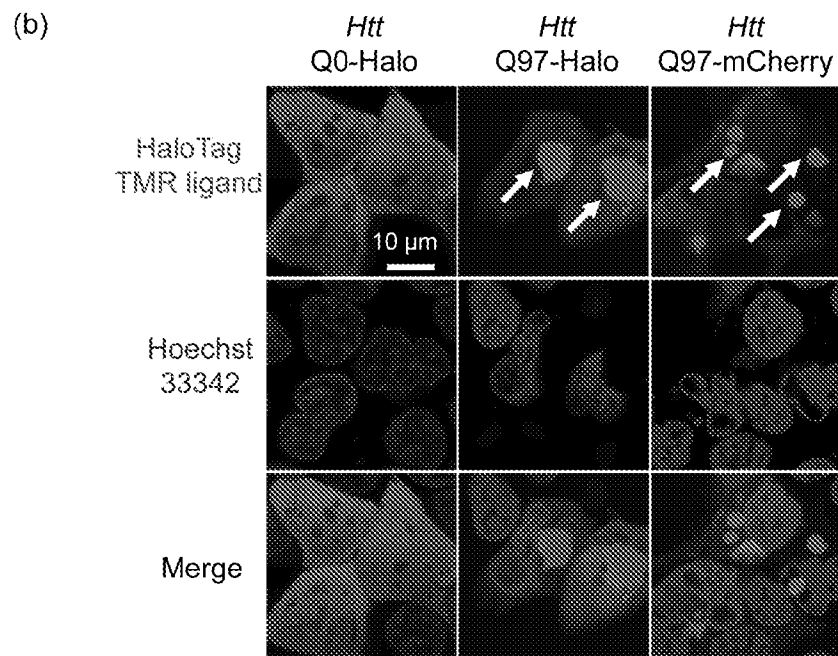
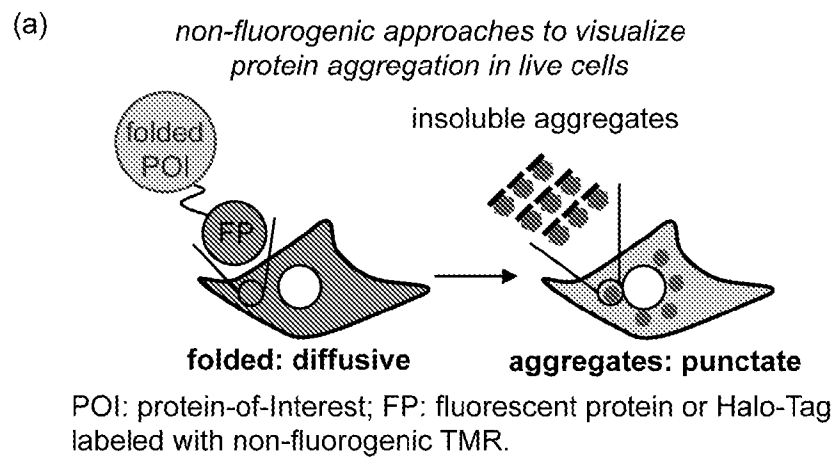


Figure 18

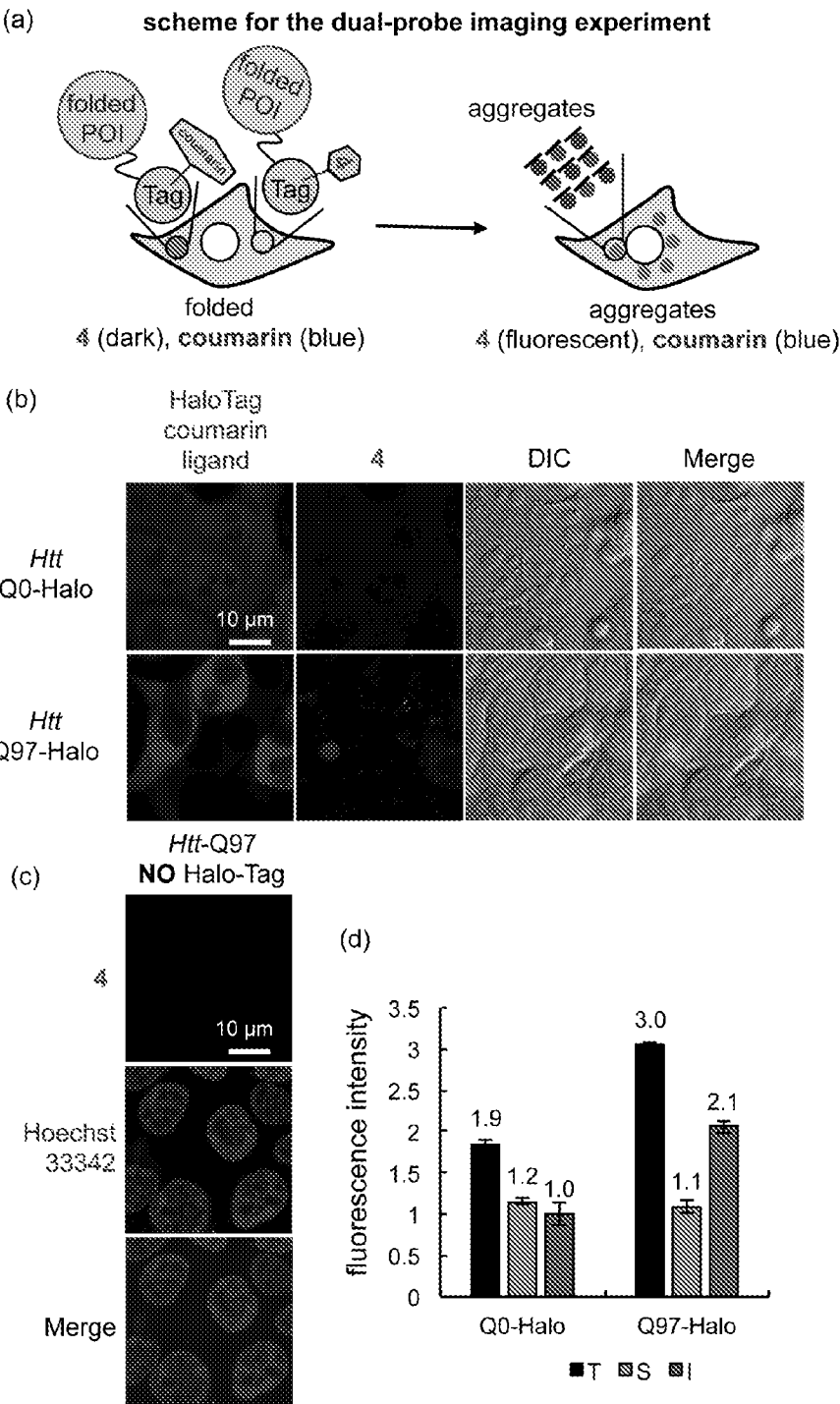
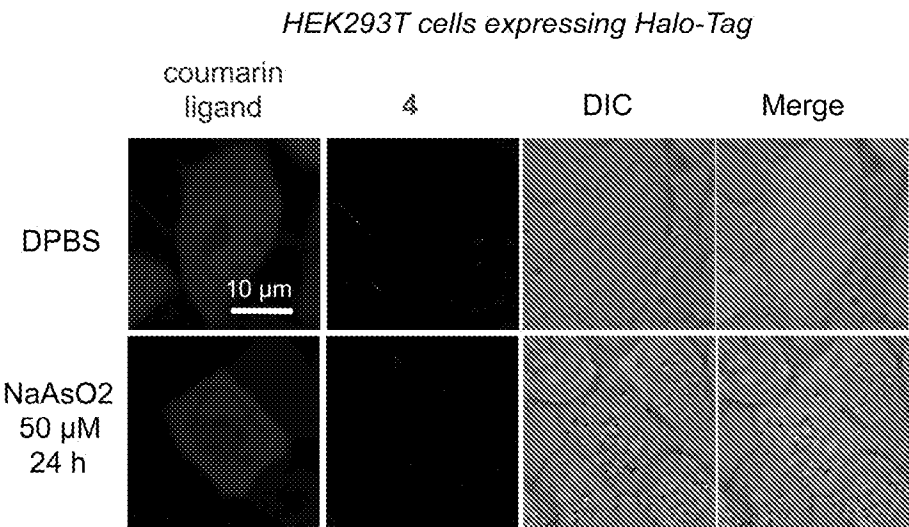


Figure 19



Note: Halo-Tag locates in both cytoplasm and nucleus

Figure 20

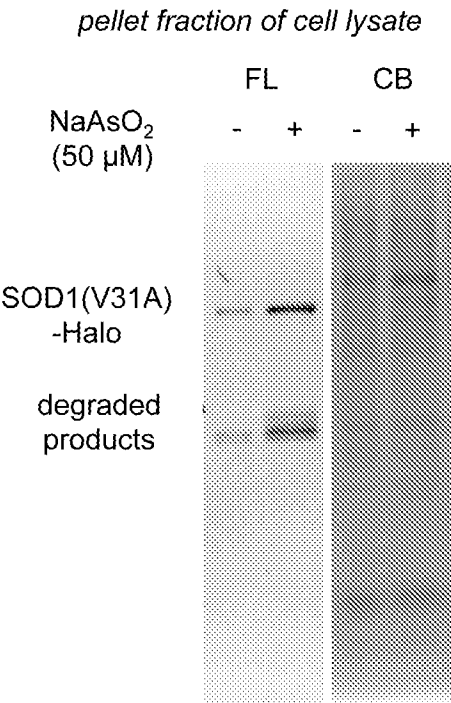


Figure 21

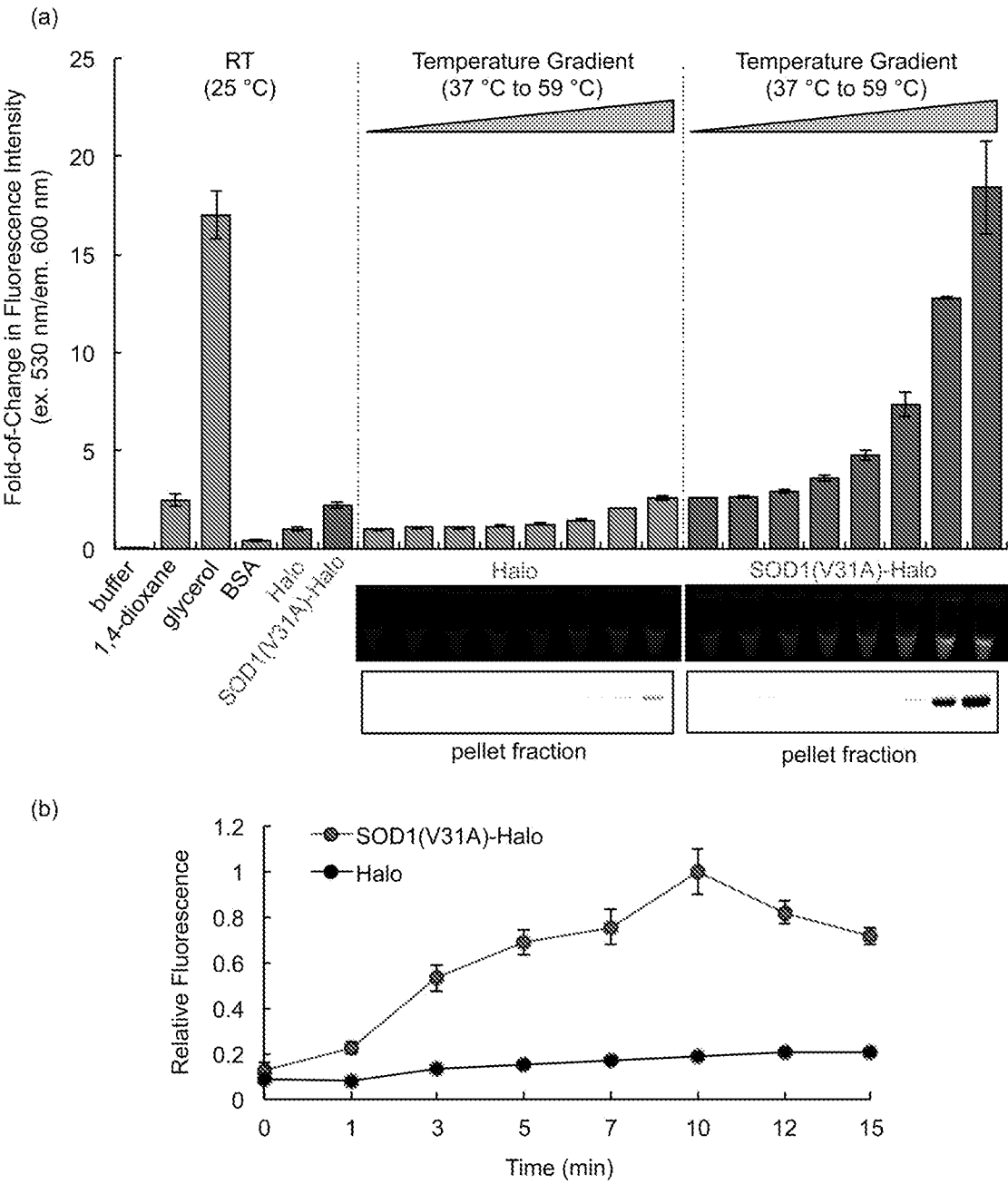
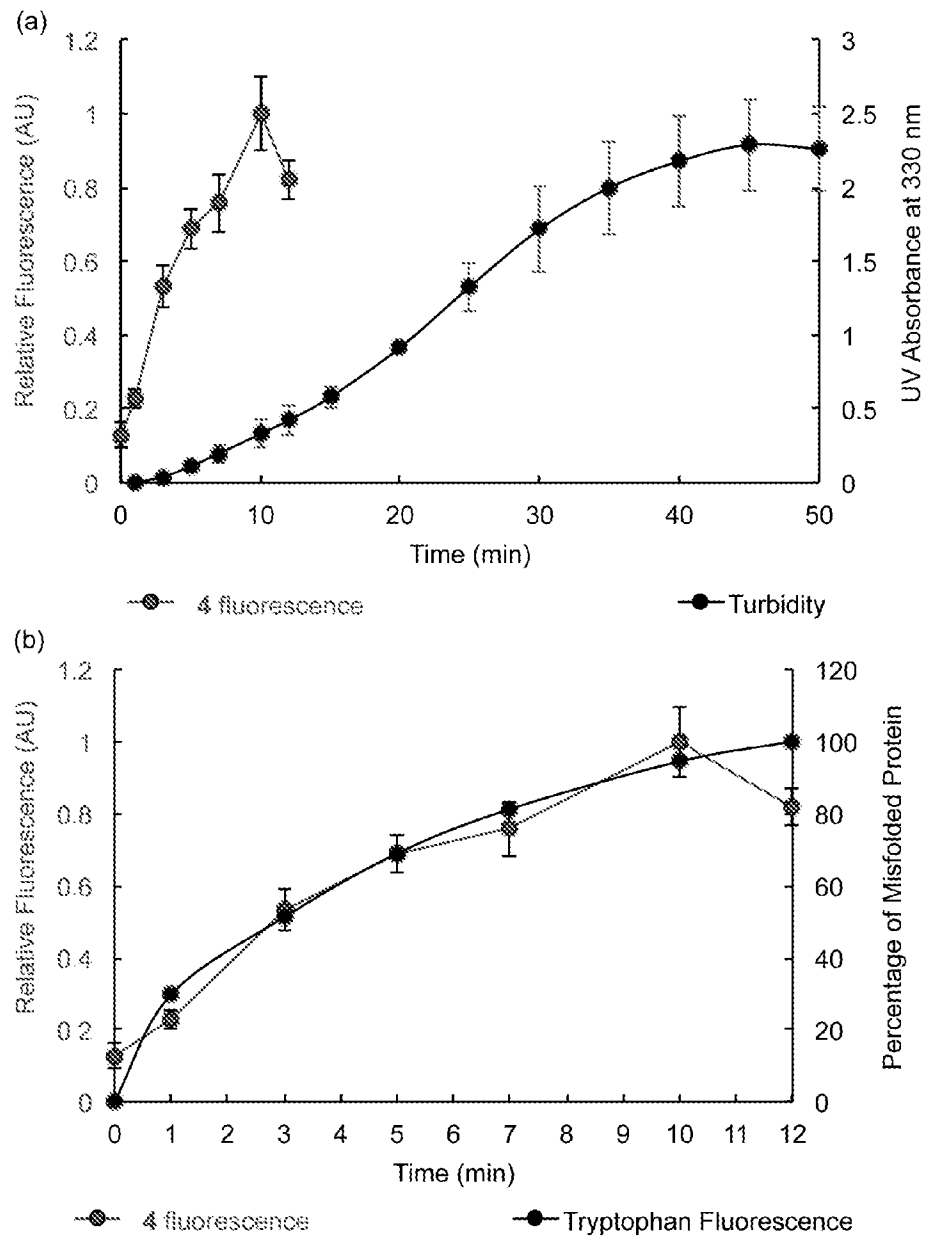


Figure 22



23/25

Figure 23

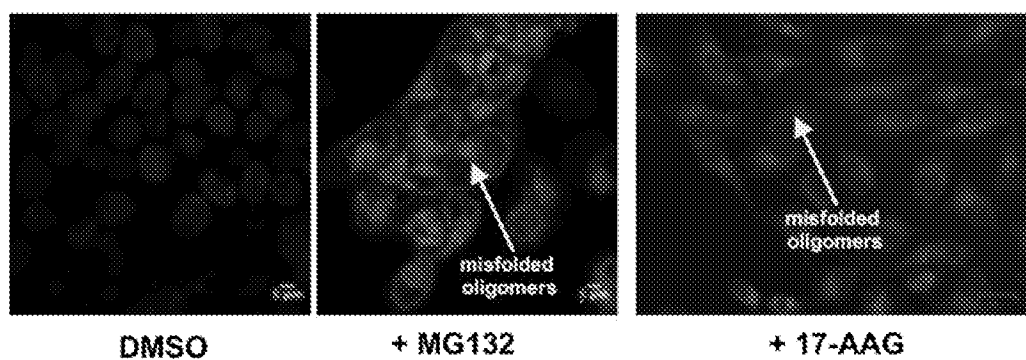


Figure 24

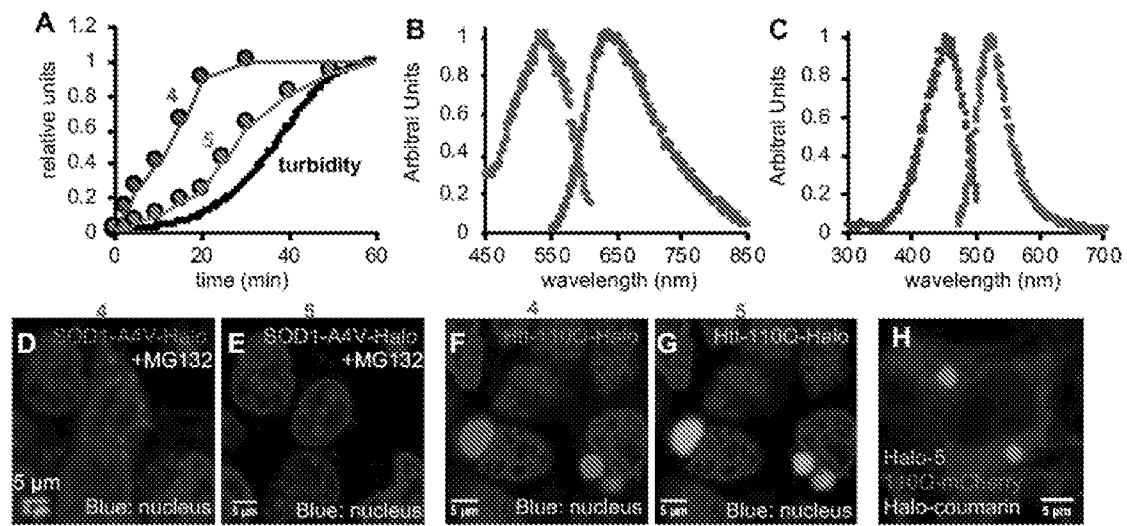
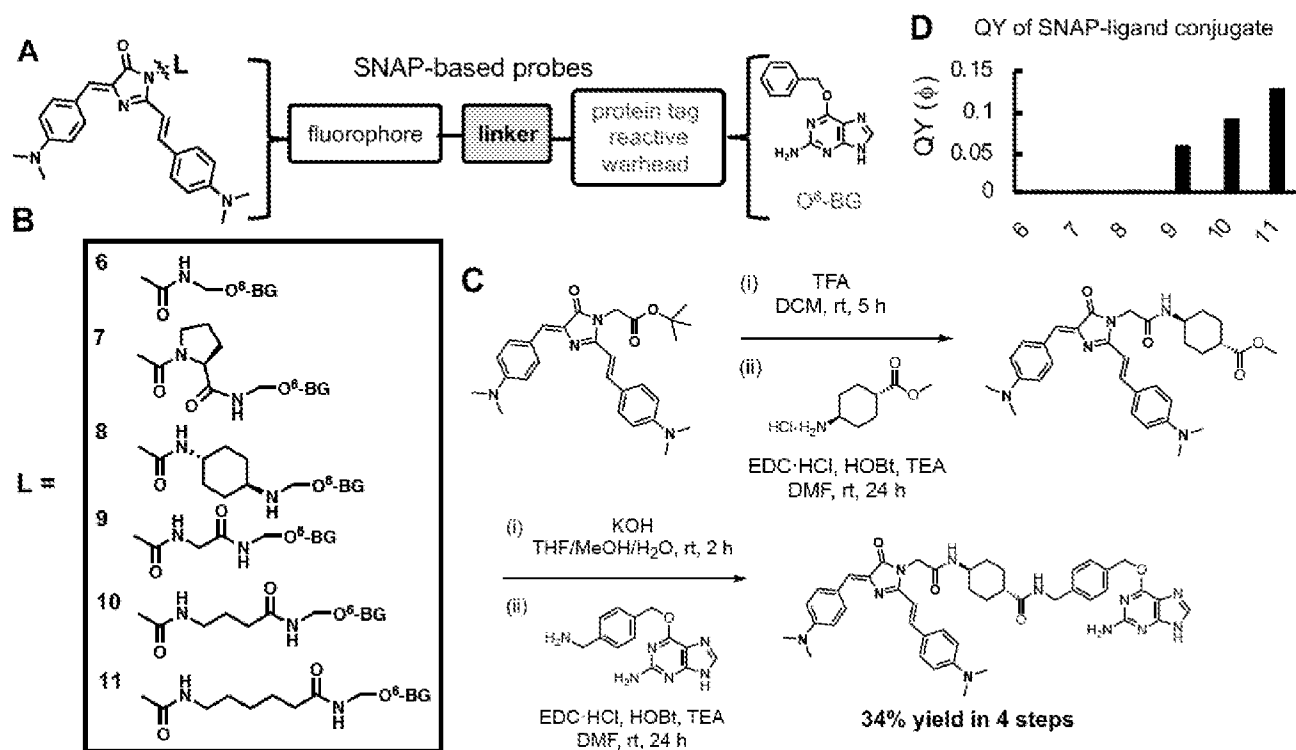


Figure 25



INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2019/021175

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - A61K 31/4164; A61K 31/4166; A61K 31/4178; C07D 233/66; C07D 233/70 (2019.01)

CPC - A61K 31/4164; A61K 31/4166; A61K 31/4178; C07D 233/66; C07D 233/70; C07D 233/96 (2019.05)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

See Search History document

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

USPC - 514/385; 514/396; 514/397 (keyword delimited)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2012/0252699 A1 (JAFFREY et al) 04 October 2012 (04.10.2012) entire document	1, 11, 12
A	WO 2008/150467 A1 (THE REGENTS OF THE UNIVERSITY OF MICHIGAN) 11 December 2008 (11.12.2008) entire document	1, 11, 12
A	US 2007/0112020 A1 (VANOTTI et al) 17 May 2007 (17.05.2007) entire document	1, 11, 12

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

04 June 2019

Date of mailing of the international search report

25 JUN 2019

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, VA 22313-1450

Facsimile No. 571-273-8300

Authorized officer

Blaine R. Copenheaver

PCT Helpdesk: 571-272-4300

PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2019/021175

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

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

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

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

3. ☒ Claims Nos.: 5
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

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

This International Searching Authority found multiple inventions in this international application, as follows:

See extra sheet(s).

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1, 11, 12

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

Continued from Box No. III Observations where unity of invention is lacking

Claims 1, 11, and 12 have been analyzed subject to the restriction that the claims read on a compound of Formula I: wherein: R1, R2, and R5 are independently selected from moieties of Group 1, wherein the moieties of Group 1 are each independently the first shown structure, and R4 is -H; wherein R3 is a directing moiety that binds and bioconjugates at least one biological target, wherein R3 is the first show structure.

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees need to be paid.

Group I+: claims 1-3, 11, and 12 are drawn to compounds of Formula I, compounds of Formula II, and kits for detecting at least one of aggregated protein, misfolded protein, and amyloid fiber.

Group II: claims 4 and 6-10 are drawn to methods for detecting at least one of a aggregated protein, misfolded protein, and amyloid fiber in a protein of interest.

The first invention of Group I+ is restricted based on the proviso R1, R2, and R5 are independently selected from the moieties of Group 1, and R4 is -H; R1, R2, and R4 are independently selected from the moieties of Group 1, and R5 is -H; R1 and R2 are independently selected from the moieties of Group 1, R4 is -H, and R5 is -H; R1 and R2 are independently selected from the moieties of Group 1, R4 is -H and R5 is -CH3; R1 and R2 are independently selected from the moieties of Group 1, R4 is -CH3 and R5 is -H; or R1, R2, R4, and R5 are independently selected from the moieties of Group 1; and is restricted to a compound of Formula I: wherein: R1, R2, and R5 are independently selected from moieties of Group 1, wherein the moieties of Group 1 are each independently the first shown structure, and R4 is -H; wherein R3 is a directing moiety that binds and bioconjugates at least one biological target, wherein R3 is the first show structure; compounds of Formula II; and kits for detecting at least one of aggregated protein, misfolded protein, and amyloid fiber. It is believed that claims 1, 11, and 12 read on this first named invention and thus these claims will be searched without fee to the extent that they read on the above embodiment.

Applicant is invited to elect additional formula(e) for each additional compound to be searched in a specific combination by paying an additional fee for each set of election. Each additional elected formula(e) requires the selection of a single definition for each compound variable. An exemplary election would be a compound of Formula I: wherein: R1, R2, and R5 are independently selected from moieties of Group 1, wherein R1 is the second shown moiety in Group 1, R2 and R5 are the first shown moieties of Group 1, and R4 is -H; wherein R3 is a directing moiety that binds and bioconjugates at least one biological target, wherein R3 is the first shown structure; compounds of Formula II; and kits for detecting at least one of aggregated protein, misfolded protein, and amyloid fiber. Additional formula(e) will be searched upon the payment of additional fees. Applicants must specify the claims that read on any additional elected inventions. Applicants must further indicate, if applicable, the claims which read on the first named invention if different than what was indicated above for this group. Failure to clearly identify how any paid additional invention fees are to be applied to the "+" group(s) will result in only the first claimed invention to be searched/examined.

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

The special technical features of Group I+, compounds of Formula I, compounds of Formula II, and kits for detecting at least one of aggregated protein, misfolded protein, and amyloid fiber, are not present in Group II; and the special technical features of Group II, methods for detecting at least one of a aggregated protein, misfolded protein, and amyloid fiber in a protein of interest, are not present in Group I+.

The Groups I+ and II formulae do not share a significant structural element requiring the selection of alternatives for the compound variables R1, R2, R3, R4, R5, R7 and accordingly these groups lack unity a priori.

Additionally, even if Groups I+ and II were considered to share the technical features of a compound having the core structure of Formula I; a compound having the core structure of Formula II; and a kit for detecting at least one of aggregated protein, misfolded protein, and amyloid fiber, comprising: one or more fluorescent protein chromophores of a known concentration in a stock solution, one or more standard protein samples that form aggregated protein, misfolded protein, and amyloid fiber and optionally, instructions for use in detecting at least one of one of the aggregated protein, misfolded protein, and amyloid fiber, these shared technical features do not represent a contribution over the prior art as disclosed by US 2012/0252699 A1 Jaffrey et al. and WO 2008/150467 A1 The Regents of the University of Michigan.

US 2012/0252699 A1 Jaffrey et al. teach a compound having the core structure of Formula I (Para. [0142]); a compound having the core structure of Formula II (Para. [0142]); and a kit (Para. [0259],...kits...) for detecting protein (Para. [0169]). The target molecule of interest can be any biomaterial or small molecule including, without limitation, proteins..., comprising: one or more fluorescent protein chromophores of a known concentration in a stock solution (Para. [0086]) Preferably, the fluorophore is soluble in an aqueous solution at a concentration of 0.1 μ M..., one or more standard protein samples (Para. [0275]). Using the extinction coefficient of DMHBI determined by standard solutions...; (Para. [0169]) and instructions for use in detecting protein (Para. [0259]). Instructions for use may also be provided.).

WO 2008/150467 A1 The Regents of the University of Michigan teaches a compound having the core structure of Formula II (Para. [0053],...fura-red...; i.e., fura-red is a resonance structure of the compound of Formula II.); and detecting amyloid fiber (Para. [0045]; Paras. [0065] through [0068]; Claim 1), one or more standard protein samples that form amyloid fiber (Para. [0041]). The relative inhibition of IgG-A β fibril interactions by test compounds is quantified using a standard ELISA protocol commonly used in the immunological arts.; Para. [0045]; Paras. [0065] through [0068]; Claim 1).

The inventions listed in Groups I+ and II therefore lack unity under Rule 13 because they do not share a same or corresponding special technical feature.