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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
02 January 2025 (02.01.2025)

WIPO | PCT

(10) International Publication Number
WO 2025/006872 A2

(51) International Patent Classification:
Not classified

(21) International Application Number:
PCT/US2024/036001

(22) International Filing Date:
28 June 2024 (28.06.2024)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

63/511,010	29 June 2023 (29.06.2023)	US
63/517,437	03 August 2023 (03.08.2023)	US

(71) Applicant: **SYNCELL (TAIWAN) INC.** [CN/CN]; C502, No. 99, Ln. 130, Sec. 1, Academia Rd, Nangang Dist., Taipei City, 115202 (TW).

(72) Inventor; and

(71) Applicant: **LIAO, Jung-Chi** [US/CN]; C502, No. 99, Ln. 130, Sec. 1, Academia Rd, Nangang Dist., Taipei City, 115202 (TW).

(72) Inventors: **CHEN, Yi-De**; C502, No. 99, Ln. 130, Sec. 1, Academia Rd, Nangang Dist., Taipei City, 115202 (TW). **CHANG, Chih-Wei**; C502, No. 99, Ln. 130, Sec. 1, Academia Rd, Nangang Dist., Taipei City, 115202 (TW). **CHEUNG, Hoi Yin**; C502, No. 99, Ln. 130, Sec. 1, Academia Rd, Nangang Dist., Taipei City, 115202 (TW). **CHUNG, Chia-Wen**; C502, No. 99, Ln. 130, Sec. 1, Academia Rd, Nangang Dist., Taipei City, 115202 (TW). **CHANG, Hsiao-Jen**; C502, No. 99, Ln. 130, Sec. 1, Academia Rd, Nangang Dist., Taipei City, 115202 (TW). **SIE, Yong-Da**; C502, No. 99, Ln. 130, Sec. 1, Academia Rd, Nangang Dist., Taipei City, 115202 (TW). **CHONG, Weng Man**; C502, No. 99, Ln. 130, Sec. 1, Academia Rd, Nangang Dist., Taipei City, 115202 (TW).

(74) Agent: **THOMAS, Justin**; Shay Glenn LLP, 155 Bovet Road, Suite 710, San Mateo, California 94402 (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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

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

(54) Title: METHOD FOR PHOTOLABELING BIOMOLECULES USING PHOTOREACTIVE PROBES

(57) Abstract: Provided is a method for photochemical labeling comprising steps of delivering a photoreactive probe to a sample, and the photoreactive probe of formula (I): T—L—W, wherein the T portion is biotin or another biotin-based moiety; the L portion includes a chemical bond or a linker; and the W portion is a photoreactive moiety, wherein the photoreactive moiety can absorb UV light and become active to crosslink with an amino acid; and selectively illuminating the sample with optical radiation to activate the photoreactive moiety of the photoreactive probe in the selected region of interest; and forming a non-specific crosslinking between the photoreactive moiety of the probe and a plurality of amino acids of different proteins of the sample in the selected region of interest.

WO 2025/006872 A2

METHOD FOR PHOTOLABELING BIOMOLECULES USING PHOTOREACTIVE PROBES

CROSS REFERENCE OF RELATED APPLICATION

[0001] This application claims priority to U.S. Provisional Patent Application No. 63/511,010, filed on June 29, 2023, titled "HIGH-SPEED ULTRA-CONTENT SPATIAL OPTOPROTEOMICS FOR SUBCELLUAR PROTEOMIC DISCOVERY," and U.S. Provisional Patent Application No. 63/517,437, filed on August 3, 2023, titled "PHOTOREACTIVE PROBES FOR TAGGING BIOMOLECULES," which are herein incorporated by reference in its entirety.

TECHNICAL FIELD

[0002] Described herein are methods for proteomics discovery. Specifically described are methods of using photoreactive probes for tagging, analyzing, and identifying biomolecules in biological samples.

BACKGROUND

[0003] Spatial proteomics allows protein mapping of a biological sample to reveal geographic organization for underlying protein-protein interactions. Both cell biologists and histologists are largely benefited by recent development in spatial proteomics, enabling, e.g., disease-associated microenvironmental protein mapping, heterogeneous protein distributions on histological samples, or protein identification of specific organelles. Targeted spatial proteomics aims to localize known proteins, whereas hypothesis-free spatial proteomics requires spatial protein identification without prior knowledge of what proteins to look for. Unlike transcriptomics where PCR is capable of amplifying signals so that hypothesis-free transcriptomics like RNAseq is possible, no PCR-equivalent technology is yet available for proteomics.

[0004] Two major techniques are feasible for hypothesis-free spatial proteomics: large-scale microscopy mapping and mass spectrometry (MS). The Protein Atlas Project mapped thousands of protein species using large-scale immunostaining, effectively creating a hypothesis-free spatial proteomic database. The limit of this approach is its applications to specific biological problems, where a multi-month exhaustive staining process has to be implemented for each project if the proteome variation due to a stimulation or a mutation is of interest.

[0005] Immunoprecipitation (IP) and MS together is a widely used biochemical approach to identify a proteome associated with a bait protein. Recent proximity labeling approaches provide better spatial precision close to a bait protein. Results from IP and proximity labeling sometimes suffer low specificity, potentially due to non-specific interactions through the pulldown process. If the regions of interest (ROIs) can only be specified by morphology characteristics without an apparent bait, proximity labeling approaches will not work. Laser-capture microdissection (LCM) enables protein isolation at specific ROIs and subsequent hypothesis-free spatial proteome identification. However, the beam size of the cutting laser is too large to achieve subcellular precision. Its non-discriminative axial cutting unavoidably includes non-specific proteins and further reduces specificity. Recent development of spatially targeted optical microproteomics (STOMP) and its derivative approaches offer another hypothesis-free spatial proteomics tool to identify the proteome at specific ROIs under a microscope. However, it lacks the fundamental scaleup requirement to reach MS need for sensitivity and specificity, making it feasible for validating known high abundant proteins but difficult to fulfill the major discovery goal, where sensitivity is required to unveil relatively low abundant proteins that are unknown for a biological problem.

SUMMARY OF THE DISCLOSURE

[0006] In view of the foregoing objectives, the invention provides a photoreactive probe, which is easy to use, hypothesis-free, and has lower noise signal in fluorescence imaging. Therefore, such photoreactive probes are better tool to process a high content of proteins for tagging, analysis, isolation, and identification in a region of interest based on user-defined microscopic image features, widely useful for cell or tissue sample experiments.

[0007] In one aspect, the disclosure provides a method for photochemical labeling, comprising: delivering a photoreactive probe to a sample, and the photoreactive probe of formula (I): T—L—W (I), wherein the T portion is a biotin or another biotin-based moiety; the L portion is a chemical bond or a linker; and the W portion is a photoreactive moiety, wherein the photoreactive moiety can absorb UV light and become active to crosslink with an amino acid; selectively illuminating the sample with optical radiation to activate the photoreactive moiety of the photoreactive probe in a selected region of interest; and forming a non-specific crosslinking between the photoreactive moiety of the probe and a plurality of amino acids of different proteins of the sample in the selected region of interest.

[0008] In one embodiment, the photoreactive probe delivered to the sample has a concentration of from 0.1 mM to 10 mM.

[0009] In one embodiment, the region of interest comprises a plurality of illumination points and the exposure time of each illumination point in the step of selectively illuminating is in a range of 10 μs - 5000 μs.

[0010] In one embodiment, the step of selectively illuminating comprises illuminating with light having a power intensity of from 1 mW to 600 mW.

[0011] In one embodiment, the sample comprises fixed cells, fixed tissues, cell extracts, or tissue extracts.

[0012] In one embodiment, the photoreactive moiety, prior to crosslink with the protein, is converted into a radical moiety upon activation with the optical radiation.

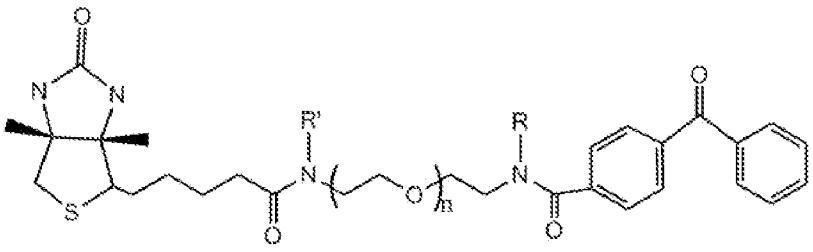
[0013] In one embodiment, the photoreactive moiety is one selecting from the group consisting of benzophenone, phenyl azide, phenyl diazirine, tetrafluorophenyl azide, hydroxyphenyl azide, and trifluoromethylphenyl diazirine.

[0014] In one embodiment, the photoreactive moiety is activatable by single-photon excitation or two-photon excitation.

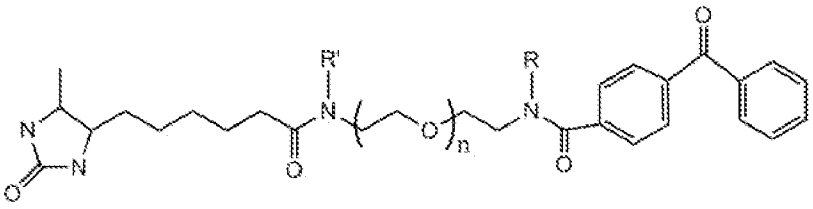
[0015] In one embodiment, the photoreactive moiety is activatable at a wavelength ranging from 200 nm to 1600 nm with a single-photon excitation.

[0016] In one embodiment, the photoreactive moiety is activatable at a wavelength ranging from 680 nm to 1600 nm with two-photon excitation.

[0017] In one embodiment, the photoreactive probe is represented by formula (I-1) or (I-2):



(I-1)



(I-2),

wherein each of R and R' independently is hydrogen, an alkyl, or a nitrogen protecting group, and n is an integer of 1-20.

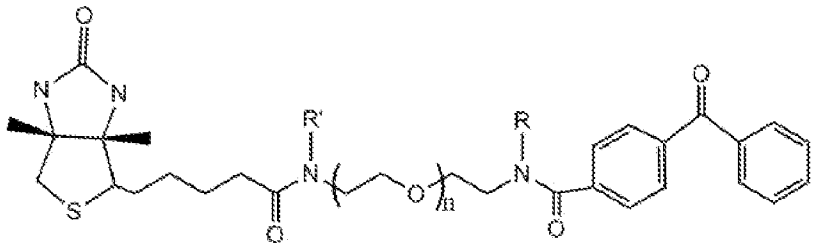
[0018] In one embodiment, the sample is affixed to a microscope slide.

[0019] In one embodiment, the non-specific crosslinking is to form a covalent bond between the active photoreactive moiety of the photoreactive probe and an alpha-carbon of an amino acid of the protein

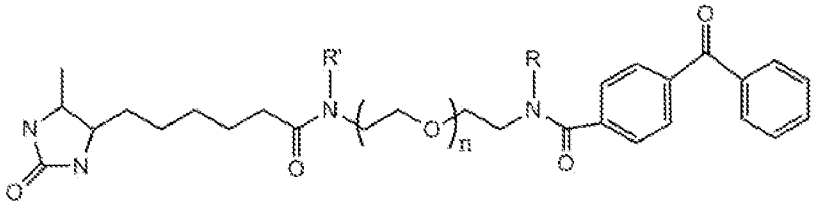
[0020] In another aspect, the disclosure provides an image-guided photolabeling method comprising: (a) delivering a photoreactive probe to a sample, wherein the photoreactive probe comprises: a detectable tag portion which is biotin or another biotin-based moiety; and a photoreactive moiety which can absorb UV light and become active to crosslink with an amino acid, wherein the detectable tag portion and the photoreactive moiety are coupled by a linker or a chemical bond; (b) imaging the sample in a first field of view and acquiring at least one image in the first field of view of the sample with an imaging light source and a camera; (c) processing the at least one image and determining a region of interest in the first of view; (d) obtaining coordinate information of the region of interest; (e) according to the coordinate information, selectively illuminating the region of interest with optical radiation to activate the photoreactive moiety of the photoreactive probe and thereby form a non-specific crosslinking between the photoreactive moiety of the photoreactive probe and a plurality of amino acids of different proteins of the sample in the region of interest; and (f) after the region of interest of the first field of view has been illuminated, moving to a second field of view of the sample, and repeating the steps (b) to (e) until all the field of views of the sample have been fully illuminated.

[0021] In one embodiment, the photoreactive moiety is one selected from the group consisting of benzophenone, phenyl azide, phenyl diazirine, tetrafluorophenyl azide, hydroxyphenyl azide, and trifluoromethylphenyl diazirine.

[0022] In one embodiment, the photoreactive probe is represented by formula (I-1) or (I-2):



(I-1)



(I-2),

wherein each of R and R' independently is hydrogen, an alkyl, or a nitrogen protecting group, and n is an integer of 1-20.

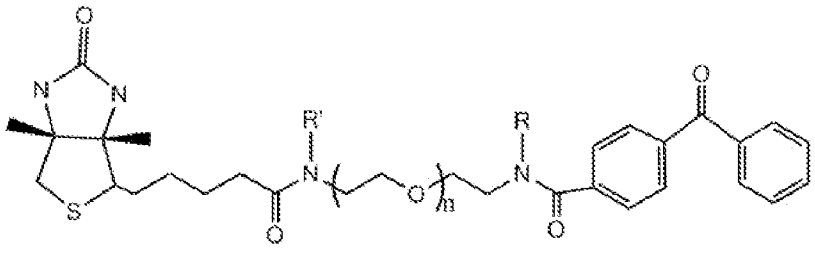
[0023] In another aspect, the disclosure provides an analytical method for a probe-labeled protein, comprising: obtaining a sample; delivering a photoreactive probe to a sample, wherein the photoreactive probe is represented by formula (I): T—L—W (I), wherein the T portion is biotin or another biotin-based moiety; the L portion is a chemical bond or a linker; and the W portion is a photoreactive moiety which can absorb UV light and become active to crosslink with an amino acid; selectively illuminating a region of interest of the sample with optical radiation to activate the photoreactive moiety of the photoreactive probe in the region of interest and form a non-specific crosslinking between the photoreactive moiety and a plurality of the amino acids of different proteins of the sample in the region of interest; and isolating the plurality of probe-labeled proteins from the sample through an affinity precipitation between the T portion of the photoreactive probe and a plurality of affinity beads.

[0024] In one embodiment, after the step of isolating, the method further comprises: subjecting the plurality of isolated proteins to mass spectrometry analysis; and identifying the plurality of isolated proteins of the sample.

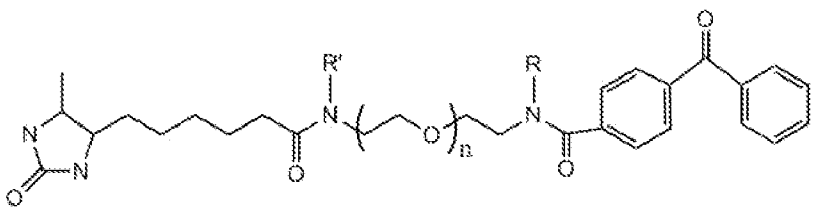
[0025] Ine one embodiment, after the step of selectively illuminating, the method further comprises removing the unbound photoreactive probes from the sample.

[0026] In one embodiment, the photoreactive moiety is one selected from the group consisting of benzophenone, phenyl azide, phenyl diazirine, tetrafluorophenyl azide, hydroxyphenyl azide, and trifluoromethylphenyl diazirine.

[0027] In one embodiment, the photoreactive probe is represented by formula (I-1) or (I-2):



(I-1)



(I-2),

wherein each of R and R' independently is hydrogen, an alkyl, or a nitrogen protecting group, and n is an integer of 1-20.

BRIEF DESCRIPTION OF THE DRAWINGS

[0028] A better understanding of the features and advantages of the methods and apparatuses described herein will be obtained by reference to the following detailed description that sets forth illustrative embodiments, and the accompanying drawings of which:

[0029] FIG. 1 is a schematic overview of the optoproteomics workflow according to some embodiments of the present disclosure. (1) Epifluorescence microscopy imaging of a field of view (FOV). (2) Real time mask generation of ROIs. (3) Illumination to selected ROIs for photochemical labeling. (4) Move to next FOV. (5) Autofocusing for next epifluorescence image. (6) Repeat steps 1-5 for each FOV to an entire specimen.

[0030] FIG. 2 shows optical setup and the controlling system of an image-guided two-photon labeling microscope system according to some embodiments of the present disclosure. A modulated femtosecond laser (780nm, 140fs pulses) was reflected off the galvanometer system (Galvo X, Galvo Y) and scanned onto the sample through the objective. The scanning patterns (masks) can be generated in real-time based on a selected mask generator for automatically photolabeling interested targets in each field of view. An all-in-one software with a user interface was developed to perform a fast high-content processing. LED refers to light-emitting diode; CAM refers to camera; ILS refers to illuminating light source; AOM refers to acousto-optic modulator; PFS refers to perfect focus system; Obj refers to objective; PRO refers to processor.

[0031] FIG. 3 shows a schematic depiction of a system useful for photoselective spatial tagging and proximity labeling of cells on a substrate according to some embodiments of the present disclosure.

[0032] FIG. 4A shows two chemical structures of one aspect of the photoreactive probe according to some embodiments of the present disclosure.

[0033] FIG. 4B shows a schematic illustration of another aspect of the photoreactive probe according to some embodiments of the present disclosure.

[0034] FIG. 4C shows a schematic illustration of illumination (IL) reaction between the photoreactive probe and alpha-carbon (Cα) of amino acid (AA) according to some embodiments of the present disclosure.

[0035] FIG. 4D schematically illustrates selectively photochemical labeling using the photoreactive probe described herein according to some embodiments of the present disclosure.

[0036] FIG. 4E shows two probes of biotin-PEG3-benzophenone (B3-BzP, top panel) and desthiobiotin-PEG3-benzophenone (DB3-BzP, bottom panel), respectively according to some embodiments of the present disclosure.

[0037] FIG. 5 shows biotin labeling efficiency significantly increased by labeling power and the concentration of B3-BzP according to some embodiments of the present disclosure.

[0038] FIG. 6A shows a thin "cross" pattern were photolabeled (λ_{ex}= 780 nm) into a PFA fixed U-2OS cells by optoproteomics and the biotin-labeled area was stained with B3-BzP and anti-biotin (neutravidin-488 fluorescent: green) according to some embodiments of the present disclosure. Scale bar: 10 μm.

[0039] FIG. 6B shows the measurement of photolabeling resolution using 40x magnification/0.95 numerical aperture (NA) objective lens by super-resolution structured illumination microscopy according to some embodiments of the present disclosure.

[0040] FIG. 6C shows confocal images of top- (xy) and side- (z) views of photolabeled subcellular compartments, the ROIs were stained with Alexa fluor 568 secondary antibody, and photolabeled signals were stained with anti-biotin (neutravidin-488 fluorescent: green) according to some embodiments of the present disclosure. NCL refers to nucleolin, NPC refers to nuclear pore complex, GM130 refers to Golgi matrix protein 130, and G3BP1 refers to GAP SH3 domain-binding protein 1. Scale bar: 10 μm.

[0041] FIG. 6D shows the top- and side views of each labeled synapse (No: 1 to 4) in spreading assay, indicating a precise photolabeling according to some embodiments of the present disclosure. The side view of photolabeling region was colocalized with CD3, a well-known marker of immune synapse on the bottom of cells. No biotin signal was found in a non-photolabeled cell (No: 5). Scale bar: 10 μm.

[0042] FIG. 6E shows photolabeled regions of immune synapse of Jurkat T cells and Raji B cells showed in green (neutravidin-488), as a precise and thin labeled layer according to some embodiments of the present disclosure. Scale bar: 10 μm.

[0043] FIG. 6F shows photolabeling on formalin-fixed and paraffin-embedded (FFPE) mouse brain tissue section according to some embodiments of the present disclosure. Cell bodies of Purkinje fibers were selected for photolabeling with biotin (green: neutravidin-488). Scale bar: 10 μm.

[0044] FIG. 7A shows overview of optoproteomics method used for proteomic profiling according to some embodiments of the present disclosure. Cells were seeded on a glass

chamber and photolabeled by optoproteomics system. Then, the photolabeled (PL) cells are lysed, enriched by streptavidin beads (STB) and digested by trypsin prior to LC-MS/MS measurement. PL refers to photolabeled, Ctrl refers to control, LP refers to labeled protein, STB refers to streptavidin bead, and LF refers to label-free.

[0045] FIG. 7B shows Dot-blot assay of streptavidin-HRP detection according to some embodiments of the present disclosure. Biotin signals are observed in photolabeled (ON) cells but not in the control cells (OFF). SA-beads refer to streptavidin beads.

[0046] FIG. 7C shows protein distribution of true positive of three biological replicates of B3-BzP (dark gray) according to some embodiments of the present disclosure. Proteins that are not annotated as nuclear proteins are shown in light grey.

[0047] FIG. 7D shows the distribution of protein copy number (< 10,000 copy number per cell: B3-BzP) according to some embodiments of the present disclosure.

[0048] FIG. 7E shows CORUM analyses of protein complexes by B3-BzP according to some embodiments of the present disclosure.

[0049] FIG. 8 shows a schematic route for synthesis of one aspect of the photoreactive probe DB3-BzP according to some embodiments of the present disclosure.

[0050] FIG. 9 shows results of selectively photochemical labeling from using the probes of DB3-BzP (right panel) and B3-BzP (left panel), respectively according to some embodiments of the present disclosure.

DETAILED DESCRIPTION

[0051] The embodiments of the invention will be apparent from the following detailed description, which proceeds with reference to the accompanying drawings, wherein the same reference characters relate to the same elements.

[0052] The ranges set forth herein may be interpreted as being inclusive of their endpoints, and open-ended ranges may be interpreted to include only commercially practical values. Similarly, lists of values may be considered as inclusive of intermediate values unless the context indicates the contrary. The recitation of ranges of values herein is merely intended to serve as a shorthand method of referring individually to each separate value falling within the range.

[0053] Methods described herein can be performed in any suitable order unless otherwise indicated herein or otherwise indicated otherwise context. The use of examples, or exemplary language (e.g., “such as”) provided with respect to certain embodiments herein is intended merely to better illuminate the invention and does not pose a limitation on the scope of the invention otherwise claimed.

[0054] The terminology used herein is for the purpose of describing particular embodiments only and is not intended to be limiting of the disclosure. As used herein, the singular forms "a", "an" and "the" are intended to include the plural forms as well, unless the context clearly indicates otherwise. It will be further understood that the terms "comprises" and/or "comprising", or "includes" and/or "including" or "has" and/or "having" when used in this specification, specify the presence of stated features, regions, integers, steps, operations, elements, and/or components, but do not preclude the presence or addition of one or more other features, regions, integers, steps, operations, elements, components, and/or groups thereof.

[0055] Microscopy-guided subcellular proteomic discovery in high sensitivity is exceedingly challenging due to the limited sensitivity of mass spectrometry and the lack of amplification tools for proteins. Described herein is an integrated technology termed optoproteomics, combining microscopy imaging, real-time deep learning-enabled image segmentation, pattern illumination-induced targeted photoaffinity labeling, affinity purification, and mass spectrometry to discover and/or capture proteins from a specific subcellular structure in cells or tissues. The process may be considered spatial proteomic discovery. In some examples, millions of targeted spots across thousands of FOVs are illuminated fully automatically through real-time image analysis and mechatronic control of one FOV at a time to achieve high-volume protein labeling of analogous subcellular ROIs.

[0056] Affinity enrichment isolates tagged proteins, effectively collecting (e.g., micro-scooping) ROI proteins at the subcellular level. The accumulated micro-scooped proteins are enough for mass spectrometric sensitivity to reveal the specific proteome in high sensitivity and specificity. This brute-force protein accumulation to overcome the current technology gap in protein amplification requires speed optimization of millions of steps to finish the entire process in hours. photoreactive probe is further used to achieve super-resolution labeling precision. Together, optoproteomics can be applicable to widely diverse cell and tissue biology problems to enable hypothesis-free subcellular proteomic discovery and generate testable hypotheses in bulk.

[0057] Described herein are methods and compositions useful for identifying, tagging, obtaining, and analyzing target biomolecules and neighboring biomolecules interaction with the target biomolecules. The methods utilize photoreactive probes particularly that can label different biomolecules, while largely maintaining naturally occurring molecular structure in the biomolecules. The probes described herein may be particularly useful for labeling subsets of biomolecules in cellular regions of cells using an image guided microscope with precision illumination control such as the system described in U.S. Patent No. 11,265,449, to enable

automatic labeling of cellular biomolecules of interest. The probes can be used for *in situ* tagging of biomolecules such as proteins inside cells or tissues and that can be followed by proximity labeling such as using Tyramide Signal Amplification (TSA). The biomolecules can be further analyzed by analytical techniques such as mass spectrometry and sequencing. These compositions may be especially useful for performing omics studies, such as genomics, proteomics, and transcriptomics, and for finding relevant biomarkers for use in disease diagnosis and treatment.

[0058] Abbreviations *and Definitions:*

[0059] The term "biotin-based moiety" refers to a biotin and variations of biotin derivatives, such as biotin with an open ring or substitutions. Typically, a biotin moiety is readily detectable with a biotin-binding entity or protein, such as avidin or streptavidin. Examples of another biotin-based moieties include diaminobiotin, biotin carbonate 5, biotin carbamate 6, and iminobiotin. In a particular example, another biotin-based moiety is desthiobiotin.

[0060] The term "linker" a structure which connects two or more substructures. A linker has at least one uninterrupted chain of atoms extending between the substructures. The atoms of a linker are connected by chemical bonds, typically covalent bonds.

[0061] The term "mass spectrometry analysis" includes linear time-of-flight (TOF), reflectron time-of-flight, single quadruple, multiple quadruple, single magnetic sector, multiple magnetic sectors, Fourier transform, ion cyclotron resonance (ICR) or ion trap.

[0062] The term "proximity molecule" or neighboring molecule refers to a molecule that is near another molecule. A proximity molecule or neighbor molecule may bound to the molecule (e.g., covalently or non-covalently) or may be close by and not bound to the molecule.

[0063] The term "photoactivated" or "light activated" refers to excitation of atoms by means of radiant energy (e.g., by a specific wavelength or wavelength range of light, UV light, etc.). In some examples, a photoactivated probe has a free radical group and can react with an alpha carbon of an amino acid.

[0064] The term "alkyl group" refers to a straight-chain (unbranched) or branched saturated group of hydrocarbons. In some embodiments, an alkyl group has 24 or fewer carbon atoms, such as having from 1 to 24 carbon atoms ("C_{1-24} alkyl"). In some embodiments, an alkyl group has 1 to 6 carbon atoms ("C_{1-6} alkyl") or 2 to 6 carbon atoms ("C_{2-6} alkyl"). Examples of C_{1-6} alkyl groups include methyl (C_1), ethyl (C_2), propyl (C_3) *(e.g.,* n-propyl, isopropyl), butyl (C_4) *(e.g.,* n-butyl, tert-butyl, sec-butyl, iso-butyl), pentyl (C_5) *(e.g.,* n-pentyl, 3-pentanyl, amyl, neopentyl, 3-methyl-2-butanyl, tertiary amyl), and

hexyl (C_6) *(e.g.,* n-hexyl). Additional examples of alkyl groups include n-heptyl, n-octyl, nonyl, decyl, dodecyl, tetradecyl, hexadecyl, eicosyl, tetracosyl, and the like. The alkyl group can be cyclic (e.g., cycloalkyl) or acyclic. By "cycloalkyl" is meant a monovalent saturated or aromatic cyclic hydrocarbon group of from 3 to 24 carbon atoms, such as from 3 to 10 carbon atoms, and is exemplified by cyclopropyl, cyclobutyl, cyclopentyl, cyclopentadienyl, cyclohexyl, cycloheptyl, bicyclo[2.2.1.]heptyl, and the like. An alkyl group can be unsubstituted or substituted. An unsubstituted alkyl group is composed only of carbon and hydrogen atoms. A substituted alkyl group has a special molecule or group bonded to a carbon atom (in place of a H atom) and can have one or more than one special molecules or groups attached. In some examples, the alkyl group of a substituted alkyl group can include haloalkyl, in which the alkyl group is substituted with one or more halo atoms (e.g., F, Cl, Br, or I). In some examples, the alkyl group can be substituted with one, two, three or, in the case of alkyl groups of two carbons or more, four or more substituents independently.

[0065] The term "nitrogen protecting group" (also referred to as an amino protecting group) refers to a substituent present on a nitrogen atom. Nitrogen can be chemically reactive and a nitrogen protecting group in a molecule can be configured to block reactivity under conditions configured for making modifications elsewhere in the molecule. Nitrogen protecting groups are well known in the art, and include, but are not limited to amide groups, carbamate groups, sulfonamide groups, and others.

[0066] The term "crosslink" or "crosslinking" refers to the process of chemically joining two or more molecules by a covalent bond, forming a three-dimensional network of connected polymers. In particular, the crosslinking is photochemical crosslinking that occurs through a photochemical reaction, which is initiated by the absorption of light. For example, the photoreactive moiety described herein can absorb UV light and generate reactive species to form a covalent bond between polymer chains.

[0067] The term "non-specific crosslinking" refers to the process where crosslinking agents form covalent bonds between biomolecules in a manner that is not selective for specific sites or interactions. This means that the crosslinking occurs randomly or broadly across various accessible sites within the biomolecules. In this disclosure, the non-specific crosslinking means when the photoreactive moiety absorbs light energy and being activated in the selected region of interest, each photoreactive moiety can form a covalent bond with an amino acid, and different photoreactive moieties can crosslink with different amino acids, either on the same protein or different proteins.

[0068] **System**

[0069] Described herein is a photoreactive probe, which can advantageously be used with a microscope system, such as the systems described herein and in U.S. Patent No. 11,265,449, to enable automatic labeling of cellular biomolecules proximal to a biomolecule of interest.

[0070] To achieve high sensitivity and high specificity of hypothesis-free subcellular proteomics, one must obtain from a biological sample enough copy numbers of protein samples for low abundant proteins. Unlike polymerase chain reaction (PCR) for nucleic acid amplification, no amplification tool is available for peptides and proteins. In this disclosure provides an automatic system and methods to achieve the same goal of protein amplification. The basic concept of the invention is to tag (e.g., biotinylate) proteins at region of interests (ROIs) of a biological sample under microscope utilizing directed photochemistry, one field of view (FOV) at a time for tens of thousands of FOVs fully automatically. The system in this disclosure performing image processing to recognize user defined ROIs by morphological features (image-guided), which providing users are able to discovery unknown and novel proteins.

[0071] Photo-induced labeling assured low background so that image-guided proteomics became feasible. However, because of the sensitivity limit of current mass spectrometry technologies, one usually needs at least femtomoles (~10^8 molecules) of the same species to be detectable. Performing image-guided photolabeling in ultrahigh content was thus necessary to be able to resolve low-abundant proteins. Therefore, the system in this disclosure performs four sequential steps repeating tens of thousands of times needed for the entire process, as shown in FIG.1.

[0072] Step 1: Epifluorescence imaging of a FOV. A biological sample (i.e. tissues, or cultivated cells) can be pretreated, such as fixed and stained on a glass slide. The fluorescence imaging allows visualization of cell morphology, cellular/subcellular compartments. For example, as shown in FIG.1, in a FOV 102 (circle), cells 101 (gray patches) were observed. According to the fluorescence signal, the researchers are able to assume target proteins existing in region of interests 103 (dark spot). In some embodiments, the photoreactive probes of the invention can be delivered to the biological sample before Step 1, or between Step 1 and Step 2.

[0073] Step 2: ROI mask generation of the FOV by traditional image processing or deep learning. The mask 104 defines all the region of interests 103 to be pattered illumination in the next step.

[0074] Step 3: Selectively illumination toward the ROIs for photochemical labeling (activation of the photoreactive probes described herein).

[0075] Step 4: Moving the stage of the microscope to the next FOV and repeating the Step 2 and Step 3, until all the FOVs were illuminated.

[0076] This repetitive process was the core to achieve protein amplification and beat the fundamental problem of protein amplification. Existing technologies or system was not optimized to perform this process repeatedly for so many times within a few hours. Without such a speed, one can only be able to identify high-abundant proteins, which are mostly known without novelty. That is, it was necessary to optimize the speeds for most of the steps and transitions during this process.

[0077] One exemplary system is shown in FIG. 2. The optical system contained a microscope 110 with a drift-free focusing setup, an sCMOS camera (CAM) 120 and LED 130 for imaging, an illuminating light source (ILS) 140 (e.g. 780-nm femtosecond laser) for two-photon illumination that triggers a photochemical reaction, an acousto-optic modulator (AOM) 150 as the illuminating light source shutter, and a pair of galvanometer scanning mirrors 160 (e.g., galvo X 162 and galvo Y 164) to direct the light illumination. The microscope 110 includes the stage 118 and the PFS 116 by a USB interface. The on/off switching of the laser was modulated by the AOM 150. The laser beam was then steered by galvo X 162 and galvo Y 164, and passed through a scan lens (SL) 113, and a tube lens (TL) 114, which are the parts of the microscope 110, to be merged in the microscope 110 through dichroic mirrors (DMs, e.g., DM1 111 and DM2 112), then focused onto the sample plane by an objective 115. The two-photon labeling can be guided from the imaging results, which can be generated by attaching a LED 130 illumination source on the microscope 110 to make epi-fluorescence imaging. The fluorescence generated within the specimen can be collected by the objective 115, and the fluorescence signals were guided with the dichroic mirrors (e.g., DM1 111 and DM2 112) and cleaned via an emission filter 117 to be projected onto a sCOMS camera 120, able of frame rates down to 19 ms.

[0078] Two-photon illumination was used to obtain better chemical labeling precision in the z direction. The use of electronically switched LED 130 and AOM 150 was to take advantage of their switching speed, at the time scale of <20 μs for LED 130 triggering speed and 25 ns for AOM 150 rise/fall time. To avoid any slowdown due to mechanical movement, multiband dichroic mirrors (e.g., DM1 111 and DM2 112) were used to allow multicolor imaging and femtosecond light illumination without movement of mechanical elements such as a turret or a shutter. A processor (PRO) 170 (e.g. field programmable gate array, FPGA) with a 1 μs conversion time was used to control AOM and galvos to assume synchronization for scanning, where 10 – 5000 μs per illumination point was implemented for a photochemical reaction. The only mechanical movements required in the process were the

fast galvo scanning and the relatively slower focal adjustment and stage movement toward the next FOV.

[0079] Photoreactive probes

[0080] According to the invention, the photoreactive probe can be represented by formula (I):

$$T—L—W \quad (I)$$

in which the T portion is a biotin or another biotin-based moiety; the L portion is a chemical bond or a linker; and the W portion is a photoreactive moiety for covalent bond formation with a biomolecule in a sample upon application of light energy.

[0081] FIG. 3 shows a schematic depiction of a system used for photoselective spatial tagging and labeling. The bottom part of FIG. 3 shows substrate 206, such as a sample slide on a microscope stage, and a monolayer of plurality of cells 208 disposed on the substrate. In some embodiments, the surface of an entire substrate, or a portion of the substrate, can be analyzed using an automated microscope system to identify a region of interest. For example, a sample can be stained or labeled to identify a region of interest. The top part of FIG. 3 shows an expanded view of cell 208a, one of a plurality of cells 208. The cell 208a has a nucleus 216 and a plurality of different types of organelles 212, such as cell membranes, mitochondria, ribosomes, and vacuoles. Microscope system 202 selectively shines narrow band of light 204 onto region of interest (ROI) 218 for analysis of the region of interest 218. The illumination can be selective, and large regions 214 of the cell and substrate are not illuminated. As explained in more detail below, narrow band of light 204 activates a photoreactive probe in only the region of interest 218.

[0082] FIG. 4A show an example of photoreactive probe that can be used in the compositions and for practicing the methods described herein. The probe of formula (I-1) and (I-2) include a benzophenone as the photoreactive moiety W and a polyethylene glycol-based (PEG-based) linker as the linker L for linkage between the photoreactive moiety W and the biotin or desthiobiotin moiety T. The linker L is not limited to the PEG-based linker, and may be, for example, peptide, amino acid, oligonucleotide, other polymeric linker, or a combination thereof. Other examples of polymeric linkers include polypropylene glycol, polyethylene, polypropylene, polyamides, and polyesters. The linker can be linear molecules in a chain of at least one or two atoms and can include more. In some embodiments, n is an integer of 1-20, and preferably is 1-6.

[0083] In some embodiments, in formula (I-1) and (I-2), R and R' each independently are hydrogen, an alkyl, or a nitrogen protecting group. As another example, the linker L may be

replaced with a chemical bond, and the biotin or desthiobiotin moiety T may be bonded to the photoreactive moiety W without the linker L therebetween, as illustrated in FIG. 4B.

[0084] FIG. 4C illustrates the reaction between the benzophenone of the photoreactive probe and alpha-carbon of an amino acid AA of the biomolecule upon illumination. The general structure of the amino acid 40 comprises alpha-carbon AA, a hydrogen, a carboxyl group 41, an amine group 42, and an R-group 43. The benzophenone group is activated by the application of light energy and converted into an excited state. The excited singlet state rapidly undergoes intersystem crossing to form a more stable triplet state. In the triplet state, benzophenone can abstract a hydrogen atom from C-H bonds (the alpha-carbon of amino acids AA) of nearby amino acids, particularly tyrosine, tryptophan, methionine, or cysteine residues. The abstraction results in the formation of a benzophenone ketyl radical and an amino acid radical. The benzophenone ketyl radical and the amino acid radical can then recombine to form a covalent bond, crossing-linking the benzophenone to the amino acid residue.

[0085] FIG. 4D shows selectively photochemical labeling using the photoreactive probe and the system described herein on a specimen to label biomolecules in a selected ROI. Prior to selectively photochemical labeling, a sample (e.g., a cell or tissue sample) containing a biomolecule of interest 310 (protein will be used herein by way of example, but other biomolecules could instead be analyzed) and a region of interest (or more) identified. The sample can be pretreated, such as fixed and stained. For example, a sample can be fixed and stained with a cell stain (e.g., hematoxylin and eosin (H&E); Masson's trichrome stain), identified with an immunofluorescent labeled antibody recognizing a biomolecule of interest 310 or by other methods.

[0086] As illustrated in FIG. 4D, the sample on substrate 309 is treated with a plurality of photoreactive probes 312 and patterned light (PLT) selectively illuminated the sample in the user-defined region of interest (grey area) so as to activate the photoreactive probes 312 to form activated photoreactive probe 312' (i.e. photoactivated probes or light activated probes). The activated photoreactive probe 312' (showed by the dotted circle) is able to form complexes with any amino acid residues by the photocrosslinking mechanism described in FIG. 4C. The activated photoreactive probes 312' can diffuse, thus not only neighbor molecules 311 near the biomolecule of interest 310, but also more distant biomolecules 331 can be labeled. The area labeled by activated photoreactive probe 312', or labeled precision, can cover a region of about 250-1000 nm. In summary, the region of interest can be determined according to the biomolecule of interest 310 *in situ*. The patterned light provides light energy to active photoreactive probe 312 so as activated photoreactive probe 312' can

non-specifically crosslink with a plurality of amino acids of different proteins within entire region of interest.

[0087] After the labeling, unreacted probes 312 (i.e. unbound probe) is washed away with wash solution, and the biomolecules labeled by the activated photoreactive probes 312' can be detected by avidin, streptavidin or NeutrAvidin, carrying a reporter group, e.g. horseradish peroxidase (HRP) or a fluorescent label.

[0088] FIG. 4E illustrates the practical chemical structure of the two photoreactive probes using in the following embodiments. One is a biotin-containing probe (biotin-benzophenone, BBzP, such as B3-BzP) (top panel), and another is a desthiobiotin-containing probe (desthiobiotin-benzophenone, DBBzP, such as DB3-BzP) (bottom panel).

[0089] Photoselective tagging and labeling as described herein can be performed in various types of samples, such as samples obtained from tissues, cells, or particles, such as from an entity (e.g., a human subject, a mouse subject, a rat subject, an insect subject, a plant, a fungi, a microorganism, a virus) or tissues samples or cell samples that are not from an organism, such as cell culture samples or artificial tissue scaffold samples (e.g., cultured laboratory cells, in vitro developed heart tissue, 3-d printed tissue, etc.). Samples for analysis using the probes, materials, and methods described herein can be living (live cells) or can be not living (e.g., fixed). A sample for tagging and labeling can include a monolayer sample, a multi-layer sample, a sample fixed to a substrate (e.g., a microscope slide), a sample not fixed to a substrate, a suspension of cells, or an extract, such as an in vitro cell extract, a reconstituted cell extract, or a synthetic extract. In some embodiments, a sample is not fixed (unfixed). In some embodiments, a sample is fixed. For example, a cell or tissue sample may be fixed with e.g., acetic acid, acetone, formaldehyde (4%), formalin (10%), methanol, glutaraldehyde, or picric acid. A fixative may be a relatively strong fixative and may crosslink molecules or may be weaker and not crosslink molecules. A cell or tissue sample for analysis may be frozen, such as using dry ice or flash frozen, prior to analysis. A cell or tissue sample may be embedded in a solid material or semisolid material such as paraffin or resin prior to analysis. In some embodiments, a cell or tissue sample for analysis may be subject to fixation followed by embedding, such as FFPE.

[0090] The concentration of the photoreactive probe treated with the sample can range from 0.1 mM to 10 mM. The wavelength of light for activation of the photoreactive probe or photo selective tagging and labeling ranges in some embodiments from about 200 nm to about 1600 nm. In some embodiments, the wavelength of light for performing photoselective tagging and labeling ranges from about 680 nm to about 1600 nm at two-photon light source; or ranges from about 300 nm to about 650 nm (e.g. 365 nm) at single-photon light source.

The wavelengths used for photoactivation of the probe are different from the wavelengths used for imaging. In some embodiments, the activation of the photoreactive probe utilizes optical radiation (light) at from around 300-450 nm, 550 nm for single-photon activation or >650 nm for multiphoton activation. The particular wavelength used in some embodiments in this disclosure is 780 nm at two-photon light source.

[0091] **Methods**

[0092] Also disclosed herein are methods of photoselectively tagging and labeling biomolecules and analytical methods. The methods may be used to tag and/or label carbohydrates, lipids, nucleic acids, proteins, either alone or in combination. The methods may include the step of treating a biological sample with a photoreactive probe having a biotin or another biotin-based moiety and a photoreactive moiety, and covalently binding the photoreactive moiety to a biomolecule in the biological sample.

[0093] In one aspect, this disclosure provides a method for photochemical labeling a biomolecule in a sample, comprising the following steps: delivering a photoreactive probe above to a sample; and selectively illuminating a selected region of interest of the sample with optical radiation to activate the photoreactive moiety of the photoreactive probe in the selected region of interest for covalent bond formation between the photoreactive moiety and a biomolecule in the sample in the selected region of interest.

[0094] As described in FIG. 2, the illumination of the ROIs is achieved by point scanning. The output of imaging processing after ROI mask generation is an XY-coordinate array (2D array) of targeted points to be illuminated covering the interested regions defined by the users. The number of illumination points per FOV can vary depending on the user's criteria. Preferably, the exposure time of each illumination point is in a range of 100-500 μs to make sure enough photochemical reaction was achieved within reasonable working hours.

[0095] The efficiency of photochemical reaction is also related to laser power intensity. Higher power intensity may shorten working time (fast photochemical reaction) but may damage sample as well. Therefore, when performing the step of selectively illumination, the power intensity is in a range of 1 mW to 600 mW. Preferably, the optimal power intensity that can be measured under a microscope is in a range of 10 mW to 300 mW.

[0096] In another aspect, this disclosure provides an image-guided photolabeling method comprising following steps: (a) delivering a photoreactive probe to a sample; (b) imaging the sample in a first field of view and acquiring at least one image in the first field of view of the sample with an imaging light source and a camera; (c) processing the at least one image and determining a region of interest in the first field of view; (d) obtaining coordinate information of the region of interest; (e) according to the coordinate information, selectively illuminating

the region of interest with optical radiation to activate the photoreactive moiety of the photoreactive probe and thereby form a non-specific crosslinking between the photoreactive moiety of the photoreactive probe and a plurality of amino acids of different proteins of the sample in the region of interest; and (f) after the region of interest of the first field of view has been illuminated, moving to a second field of view of the sample, and repeating the steps (b) to (e) until all the field of views of the sample have been fully illuminated.

[0097] In one another aspect, this disclosure provides an analytical method for a probe-labeled protein comprising the following steps: obtaining a sample; delivering a photoreactive probe above to the sample, selectively illuminating a region of interest of the sample with optical radiation to activate the photoreactive moiety of the photoreactive probe in the region of interest and form a non-specific crosslinking between the photoreactive moiety and a plurality of the amino acids of different proteins of the sample in the region of interest; and isolating the plurality of probe-labeled proteins from the sample through an affinity precipitation between the T portion of the photoreactive probe and a plurality of affinity beads .

[0098] The following examples of the invention will describe the practical application.

[0099] **Example 1** Photoactivable amino acid crosslinking enables spatial photo-induced

[0100] To achieve photo-induced labeling of proteins at a microscopic illumination point, the sample was added with molecules with three elements: a photocatalysis, a photoactivable amino acid linker to covalently bind to a protein, and a probing tag for protein pulldown. BBzP (Fig. 4E) was used as the photoreactive probe in the following Examples 1-3. Benzophenone at the illumination point was excited to become 1,2-diradical, which reacted with the C-H bond of α-carbon, forming a covalent bond of the corresponding amino acid. That is, a plurality of amino acids of different proteins within the entire patterned light area can be biotinylated.

[0101] U-2OS cells (HTB-96, ATCC, VA, USA) were cultivated at 37°C in a 5% CO_2 humidified environment in Dulbecco's Modified Eagle Medium supplemented with 10% FBS. 2×10^5 cells were seeded in glass bottom chambers (80287, ibidi) and incubated for approximately 16 h to 80-90% of confluency. Afterwards, cells were washed with phosphate buffered saline (PBS) then fixed with 2.4% paraformaldehyde solution (PFA, 15710, Electron Microscopy Sciences) or 100% ice-cold methanol. Fixed cells were incubated with PBS/0.5% Triton X-100 to permeabilize the cell membrane and blocked with 3% BSA in PBS/0.1% Triton X-100 (PBS-T) for 1 h, followed by 30 min of 0.002% streptavidin and 15 min of 40 μM biotin blocking.

[0102] Different photosensitizers, their concentrations, and illumination intensities were tested to satisfy speed and specificity needs of photochemistry. B3-BzP gave the best labeling efficiency when illuminating the entire FOV of U-2OS cells with a 780 nm femtosecond laser, illustrated by Dy488-NeutrAvidin imaging. 150 mW, 200 mW laser power with different concentrations of BBzP (100 μM, 250 μM) at the back focal plane was enough for labeling. As shown in FIG. 5, the nonspecific labeling background was low, demonstrating that BBzP was a good choice to optimize both labeling efficiency and spatial specificity.

[0103] **Example 2** Subcellular spatial biotinylation by BBzP is achieved for cell and tissue samples

[0104] To facilitate subcellular protein isolation, the labeling resolution should be at a submicron level. A line scan illumination was performed on U-2OS cells and super-resolution structured illumination microscopy (SR-SIM) was used to measure the biotinylation labeling width by staining with B3-BzP and anti-biotin (FIG. 6A). 0.39 μm labeling resolution was able to reach when using a 40x magnification/0.95 NA objective (FIG. 6B). To demonstrate its broad usage on cell samples, we photo-tagged proteins at five subcellular locales including nucleus, nucleoli, nuclear pore complexes, stress granules, and Golgi apparatuses. Image segmentation was performed based on the characteristics of each structure using traditional image processing. The segmented region was illuminated to induce targeted biotinylation. The *in situ* biotinylated regions matched well with the corresponding subcellular structures in lateral (xy) and axial (z) directions, suggesting a high spatial labeling specificity (FIG. 6C).

[0105] Moreover, deep learning was applied to segment the region within F-actin rings in a spreading assay on an anti-CD3 coated glass surface. Spreading assay was performed as follows: glass chambers were coated with poly-D-lysine, washed with PBS, and incubated with 10 mg/mL anti-CD3 (317301, BioLegend) and 10 mg/mL anti-CD28 (302933, BioLegend) antibodies in PBS overnight at 4°C. Cells were washed with PBS-T to remove excess antibodies. Jurkat T cells (Clone E6-1, TIB-152, ATCC) were dropped on glass chamber and incubated at 37°C. After 15 min of incubation, cells were fixed with 4% PFA and stained with indicated antibodies. Immune synapse induction was performed as follows: Raji B cells (CCL-86, ATCC) were used as the antigen presenting cells (APCs). To activate the APCs for immune synapse (IS) formation, 1×10^6 Raji B cells were pelleted to resuspend with 100 μL serum-free RPMI (SH30027, HyClone) containing 0.1 μg Staphylococcal Enterotoxins (ET404, Toxin Technology Inc) and incubated at 37°C for 1 h. On the last 20 min of the incubation, the Raji B cells were stained with 20 μM cell tracker Red CMTPX Dye (C34552, Invitrogen). The cells were then washed twice with PBS to remove excess dye.

6.5×10^5 B cells were resuspended with 500 µL serum-free RPMI and seeded on individual well of a poly-D-lysine (P7280, Sigma-Aldrich) coated glass chamber (80287, ibidi) for 20 min. Equal amount of Jurkat T cells (Clone E6-1, TIB-152, ATCC) were resuspended with 250 µL serum-free RPMI and then added dropwise into the B cells. After incubating at 37°C for 10 min, the T-B cell conjugates were fixed with 2% PFA solution for 10 min, followed by 100% ice-cold methanol fixation for 5 min at -20°C. To visualize IS, immunofluorescence staining of CD3 (317301, BioLegend), together with Alexa Fluor™ 647-IgG (A-21235, Invitrogen) as secondary antibody, was performed for subsequent photo-labeling experiments.

[0106] As shown in FIG. 6D, Segmented illumination at a focal level within the glass allowed photolabeling at a thin layer of the contact surface allowing specific biotinylation of proteins at the supramolecular activating complexes (SMACs) of immune synapses. Deep learning-based segmentation and targeted illumination were also applied to conjugates of Jurkat T -Raji B cells with precise biotinylation of proteins at thin immune synapses (FIG. 6E), illustrating the capability of photolabeling upon multi-color imaging.

[0107] The applicability of photolabeling on tissue samples was further examined. C57BL/B6 Mouse was perfused with 0.9% NaCl to remove the blood, then 4% PFA solution were replaced to fix tissue samples. Mouse brain was prepared by following the FFPE block procedure which included fixation, dehydration, and paraffin embedded. FFPE slices were section at 10 µm, followed by deparaffinization and rehydration. The purkinje cells were stained with anti-calbindin-D28K primary antibody (C9848, SigmaAldrich) overnight at 4°C, and subsequently stained with secondary antibody AlexaFluor-555 (A32727, Invitrogen).

[0108] As shown in FIG. 6F, Calbindin-D28K staining of a FFPE mouse brain tissue section revealed a high abundant of Purkinje cells in cerebellum with clear axon and dendrite architecture and a distinctive neuron cell body layer. Deep learning-enabled segmentation and following targeted illumination resulted in biotinylation at cell bodies, demonstrating photolabeling specificity for FFPE tissue samples.

[0109] **Example 3** Optoproteomics enables subcellular spatial proteomics in high sensitivity and specificity by BBzP

[0110] To check whether ultra-content spatial biotinylation can yield isolation of enough proteins for proteomic studies, targeted photolabeling and pulldown of nuclear proteins were first tested as a proof of principle. Photolabeling on subcellular structures was performed as follows: cells were incubated with photon labeling reagent containing 0.1-2.0 mM BBzP. Two-photon laser coupled with a microscopic system was used for photolabeling at a laser power of 100-200 mW, and the cells were subjected to an image-guided laser-exposure time at 100-1000 µs. Verification of the performance of two-photon labeling by fluorescent

microscopy was performed as follows: labeled cells were washed with PBS-T three times. Labeled cells/signals were stained with neutrAvidin-DyLight 488 (22832, Invitrogen) for 1h in 3% BSA/PBS/0.1% Triton X-100 for 1 h at RT. Cells were subsequently stained with nuclear marker (Hoechst 33258, 62249, Thermo Fisher Scientific) for 30 min at room temperature. A Zeiss LSM 880 confocal microscope was applied to verify the labeling signals.

[0111] After photolabeling, cells were scraped, harvested, and lysed to extract proteins. Protein extraction and on-bead digestion was performed as follows: labeled cells were harvested by scraping with buffer containing 10 mM Tris (pH 8.0), 1% Triton X-100, 1-fold protease inhibitor cocktail, 10 mM sodium ascorbate, 5 mM Trolox, and 1 mM sodium azide. Harvested cells were sonicated at 60% power using a Q125 sonicator (Qsonica) with 1s on/ 2s off interval for 2 min, then subjected to evaporate the scraping buffer for 2 h by SpeedVac system. 160 µL of lysis buffer containing 4% Sodium dodecyl sulfate (SDS), 1% Triton X-100, 10 mM Tris (pH 8.0), and dithiothreitol (DTT) were added to the harvested cells, and the mixture was vortex at 1 min on/2 min off interval for 5 cycles. To retrieve the cross-linked amide groups resulted from PFA fixation, the lysed cells were further heated at 99°C for 45 min followed by another vortex at 1 min on/2 min off interval for 5 cycles. Lysates were centrifuged at 16,000 g for 20 min at 20°C, and the supernatant were collected. Pierce™ 660nm Protein Assay was used to measure the protein concentrations, and 240 µg of proteins were subjected to the immunoprecipitation. Streptavidin magnetic beads were washed with dilution buffer (0.5% Triton X-100/PBS) three times, and the protein lysates were diluted 10-fold to reduce the SDS concentration to be less than 0.4%, and the diluted lysates were added to the washed beads and incubated at 2-8°C for 16 h with rotation. After which, the biotinylated-protein bonded beads were washed with the following washing buffers to reduce the non-specific binding maximally: Buffer A (2% SDS, 50 mM Tris, pH 8.0); Buffer B (0.5M NaCl, 0.1% deoxycholic acid, 0.1% SDS, 1% Triton X-100, 50 mM HEPES); Buffer C (0.5% deoxycholic acid, 0.5% Triton X-100, 10 mM Tris, 250 mM LiCl, pH 8.0). For on-bead digestion, the beads were further washed with 100-µL 50 mM TEABC three times, and the biotin-protein bonded beads were then mixed with 0.2 µg of Trypsin/ Lys-C (V5071, Promega) in a final volume of 20 µL at 37°C for 100 min for an initial digestion. After that, the supernatant was collected and subjected to the overnight digestion without adding further enzyme. Finally, the digests were acidified by adding 2 µL of 10% formic acid and were desalted by C18 Ziptip Desalted peptides were dried by Speedvac and stored at -20°C prior to LC-MS/MS analysis.

[0112] LC-MS/MS analysis was performed as follows: detection of immunoprecipitated product by data-dependent acquisition mass spectrometry. LC-MS/MS analysis was

performed using an UltiMate 3000 RSLCnano system (Thermo Fisher Scientific) coupled to an Orbitrap Fusion Lumos mass spectrometer (Thermo Fisher Scientific). The desalted peptides were resuspended in 0.1%formic acid in water and loaded onto a PepMap™ 100 C18 HPLC column (2 μm, 100 angstrom, 75 μm × 25 cm; Thermo Fisher Scientific), and peptides were eluted over 160 min gradients for nuclei-illuminated samples, over 120 min gradients for nucleoli-, SG-illuminated samples. The full MS spectra ranging from m/z 375–1500 were acquired at a resolving power of 120,000 in Orbitrap, an AGC target value of 4×10^5, and a maximum injection time of 50 ms. Fragment ion spectra were recorded in the top-speed mode at a resolving power of 30,000 in Orbitrap using a data-dependent method. Monoisotopic precursor ions were selected by the quadrupole using an isolation window of 1.2, 0.7, 0.4 Th for the ion with 2+, 3+, 4–7 charge states, respectively. An AGC target of 5×10^4, maximum injection time of 54 ms, higher-energy collisional dissociation (HCD) fragmentation with 30% collision energy, and a maximum cycle time of 3s were all applied. Dynamic exclusion was set to 60s with an exclusion window of 10 ppm. Precursor ions with the charge state of unassigned, 1+, or superior to 8+ were excluded from fragmentation selection.

[0113] Streptavidin dot-blot analysis was performed as follows: photolabeled U-2OS cells and control U-2OS cells were lysed and immunoprecipitated with streptavidin beads as described above. Dot-blot analysis was performed in a PVDF membrane. PVDF membrane was activated with 100% MEOH, and soaked in PBS-T for 2 min, then 8 μL of each sample were spot on the membrane. After dry, the membrane (blot) was rinsed with PBS-T and blocked with 5% BSA for 30 min at RT, and washed with PBS-T. Then the membrane was incubated with Streptavidin-HRP (1: 1000) in 5% BSA at 4°C overnight on a seesaw shaker. The blot was washed with PBS-T before to development with ECL substrate (Bio-Rad) using iBright FL1500 image system (Invitrogen). The blot was stripped with NaOH for 5 min at 60°C 3 times, blocked with 2.5% BSA, and re-probed with mouse anti-tubulin 1: 500 (GTX112141, GeneTex) in 2.5% BSA at 4°C overnight on a seesaw shaker, and washed with PBS-T. Then the blot was incubated with anti-mouse-HRP (1: 500) in 2.5% BSA at RT for 2 h. The blot was washed with PBS-T prior to development with ECL substrate (Bio-Rad) using iBright FL1500.

[0114] As shown in FIG. 7A, biotinylated proteins by BBzP were enriched by immunoprecipitation using streptavidin beads, tryptically digested on beads, and subject to LC-MS/MS measurement. Dot-blot analysis showed that biotinylated proteins by BBzP were enriched in the protein lysate or streptavidin beads pull-down when photolabeling was turned on (FIG. 7B), demonstrating effective photobiotinylation of the sample. A distribution of

overall protein abundances quantified in PL to control (CTL) was calculated (FIG. 7C). The sensitivity and specificity of the spatially labeled nuclear proteome were determined by calculating the percentage of the true positive proteins from a final protein list. A total of 1,150 proteins showed differentially enriched, where 1,032 were annotated as nuclear proteins, accounting for 90% of true positive rate.

[0115] Ability to identify low-abundance proteins with high sensitivity is important for discovery of novel biomarkers. Protein identification and label-free quantification was performed as follows: raw data from the same batch of two-photon illumination were processed together with Proteome Discoverer (Thermo Fisher Scientific) by Sequest HT algorithm against the UniProtKB/Swiss-Prot human protein database (version 2020.02, 20,365 entries) for feature extraction, peptide identification, and protein inference. Database search was performed as follows: tryptic peptides with up to three missed cleavages; mass tolerances of 10 ppm for peptide ions, and 0.05 Da for fragment ions; static carbamidomethylation (+57.0215 Da) on Cys residues; dynamic deamidation (+0.9840 Da) on Asp and Gln residues, oxidation (15.9949 Da) on Met residues, acetylation on protein N-termini (+42.0106 Da). The minimal peptide length was set as 6 residues. The false discovery rate (FDR) of peptide and protein were both set as 1%. For label-free quantification, the time windows for chromatographic peak alignment was set as 20 min. Peptide level data was then normalized to the total peptide intensity, and the quantification value for a given protein was derived from the sum of normalized intensities of the top three intense unique peptides belonging to that protein. Label-free quantification (LFQ) was performed to extract enriched proteome from background signals. After streptavidin beads enrichment and identification by mass spectrometry, protein IDs from both control group and experimental group were analyzed by comparing their peak intensities. The distribution of true positive and false positive proteins was separated in the axis of $\log_2$(PL/Control) and gave the PL labeled protein IDs by applying a cutoff of $\log_2$(PL/Control).

[0116] As shown in FIG. 7D, more than 10% of the nuclear proteins were identified as low-abundance proteins with less than 10,000 copy number per cell, illustrating the capability of optoproteomics to identify low abundant proteins.

[0117] The capability of protein complex discovery and carried out CORUM complex analysis was further examined by employing differentially enriched proteins as seed proteins to reveal protein complexes. Functional annotation was performed as follows: the GO (Gene Ontology) term association was obtained from Uniprot and the protein complex data was downloaded from CORUM. The Fisher's exact test with the Benjamini-Hochberg multiple testing correction was used to identify over-represented GO terms and protein complexes.

The proteins annotated by GO and curated by CORUM were used as the background set for GO and protein complex enrichment analyses. STRING analysis of stress granules protein network was performed using 124 over-represented (enriched) proteins for associations regarding functional and physical interactions as shown in a visualized network graph.

[0118] As shown in FIG. 7E, the over-represented protein complexes were within our nucleus-targeted region, such as spliceosome, histone complex, and RNA polymerase complex, suggesting optoproteomics provides not only true positive proteins but also spatial-related protein complexes.

[0119] **Example 4** Probe synthesis of desthiobiotin-benzophenone (DBBzP)

[0120] The probe of desthiobiotin-PEG3-benzophenone (DB3-BzP) was synthesized according to synthesis scheme illustrated in FIG. 8. The illustrated synthesis scheme is given as an example and not for limiting purposes.

[0121] **Example 5** Photochemical labeling by DBBzP on subcellular structures of cells.

[0122] Cells were cultivated at 37°C in a 5% CO_2 humidified environment in Dulbecco's Modified Eagle Medium supplemented with 10% FBS. 2×10^5 cells were seeded in glass bottom chambers and incubated for approximately 16 h to 80-90% of confluency. Afterwards, cells were washed with PBS and fixed with 2.4% paraformaldehyde (PFA). Fixed cells were incubated with PBS/0.5% Triton X-100 to permeabilize the cell membrane, then blocked with 3% BSA in PBS/0.1% Triton X-100 for 1 h, followed by 30 min of 0.002% streptavidin and 15 min of 40 μM biotin blocking. Cells were incubated with labeling reagent, containing 0.5-2.0 mM of DB3-BzP. A microscopic system coupled with a two-photon laser was used for spatially resolved photolabeling. The cells were subjected to a laser-exposure time at 100-1000 microseconds at a power of 50-400 mW to activate the benzophenone group to converted into an active diradical, forming a covalent bond with the alpha-carbon of amino acids in the selected regions of interest. After the labeling, cells were washed thoroughly with phosphate buffered saline (PBS) for three times to remove the residual labeling reagent.

[0123] Quality control was performed as follows: the labeling (desthiobiotinylation) efficiency and labeling reagent residues can be observed and evaluated by staining cells with 500-fold diluted NeutrAvidin-488 (488nm fluoresce dye conjugated NeutrAvidin, N488) in 0.1% PEST supplemented with 3% BSA for 1h at RT on a seesaw shaker. A signal-to-noise ratio was used as the procedure of quality control:

$$S/N = \frac{N488\ intensity\ of\ ROIs - Blank}{N488\ intensity\ of\ nonROIs - Blank}$$

where N488 intensity is the average intensity of the N488 channel from a fluorescent microscope; ROI, region of interest; Blank, average intensity of N488 channel at the areas without cells.

[0124] As shown in FIG. 9, nucleoli of U-2OS cells were labeled by B3-BzP (left) and DB3-BzP (right) and imaged by fluorescent microscopy, the signal-to-noise ratios are 5.8 and 17.8 respectively.

[0125] There were two major challenges of photochemistry: labeling efficiency and spatial labeling specificity. If one second illumination per point were implemented, it would have taken ten million seconds for ten million points, or about 4-month continuous illumination, completely impractical for an experiment. Thus, photochemical labeling efficiency should be optimized to reach the desired 100-1000 μs per illumination spot. Spatial labeling specificity was also essential because the ROIs might only occupy 1/10,000 or even smaller volume of the entire biological sample space. A small probability of nonspecific binding at the non-illuminated regions would result in large collection of unwanted proteins, shadowing the accumulated copy number of desired proteins at the ROIs. In summary, we provide photoreactive probes suitable for using in an image-guided system. The photoreactive probes have high photolabeling efficiency. Depending on the number of points of illumination, the total time to photo-label proteins of a 2 cm x 2 cm sample well using a 40x objective was 3 to 16 hours. One to ten sample wells were needed to collect enough proteins for mass spectrometry analysis. Moreover, when using BBzP and DBBzP, the nonspecific labeling background both were lower, demonstrating that both photoreactive probes have better labeling efficiency and spatial specificity.

[0126] Although the terms “first”, “second”, “third” and “fourth” may be used herein to describe various features/elements (including steps), these features/elements should not be limited by these terms, unless the context indicates otherwise. These terms may be used to distinguish one feature/element from another feature/element. Thus, a first feature/element discussed below could be termed a second feature/element, and similarly, a second feature/element discussed below could be termed a first feature/element without departing from the teachings of the present invention.

CLAIMS

What is claimed is:

1. A method for photochemical labeling, comprising:

delivering a photoreactive probe to a sample, wherein the photoreactive probe is represented by formula (I):

$$T—L—W \quad (I),$$

wherein

the T portion is biotin, or another biotin-based moiety;

the L portion is a chemical bond or a linker; and

the W portion is a photoreactive moiety, wherein the photoreactive moiety can absorb UV light and become active to crosslink with an amino acid;

selectively illuminating the sample with optical radiation to activate the photoreactive moiety of the photoreactive probe in a selected region of interest; and

forming a non-specific crosslinking between the photoreactive moiety of the probe and a plurality of amino acids of different proteins of the sample in the selected region of interest.

2. The method of claim 1, wherein the photoreactive probe delivered to the sample has a concentration of from 0.1 mM to 10 mM.

3. The method of claim 1, wherein the region of interest comprises a plurality of illumination points and the exposure time of each illumination point in the step of selectively illuminating is in a range of 10 μs - 5000 μs.

4. The method of claim 1, wherein the step of selectively illuminating comprises illuminating with light having a power intensity of from 1 mW to 600 mW.

5. The method of claim 1, wherein the sample comprises fixed cells, fixed tissues, cell extracts, or tissue extracts.

6. The method of claim 1, wherein the photoreactive moiety, prior to crosslink with the protein, is converted into a radical moiety upon activation with the optical radiation.

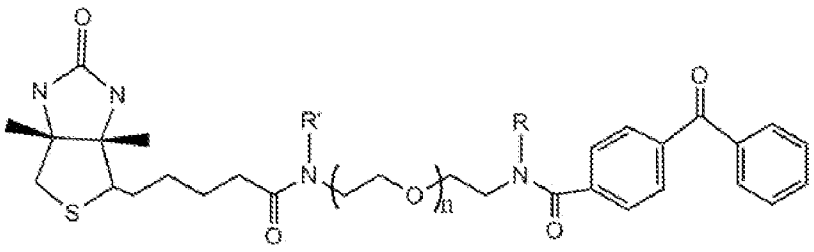
7. The method of claim 1, wherein the photoreactive moiety is one selecting from the group consisting of benzophenone, phenyl azide, phenyl diazirine, tetrafluorophenyl azide, hydroxyphenyl azide, and trifluoromethylphenyl diazirine.

8. The method of claim 1, wherein the photoreactive moiety is activated by single-photon excitation or two-photon excitation.

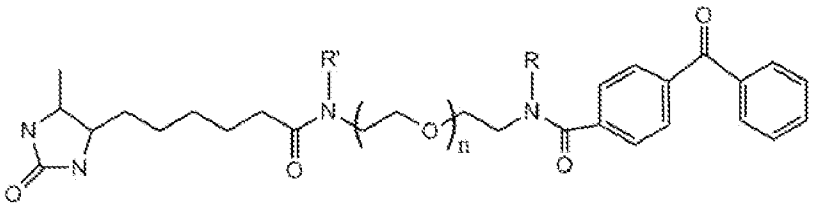
9. The method of claim 1, wherein the photoreactive moiety is activated at a wavelength ranging from 200 nm to 1600 nm with single-photon excitation.

10. The method of claim 1, wherein the photoreactive moiety is activated at a wavelength ranging from 680 nm to 1600 nm with two-photon excitation.

11. The method of claim 1, wherein the photoreactive probe is represented by formula (I-1) or (I-2):



(I-1)



(I-2),

wherein each of R and R' independently is hydrogen, an alkyl, or a nitrogen protecting group, and n is an integer of 1-20.

12. The method of claim 1, wherein the sample is affixed to a microscope slide.

13. The method of claim 1, wherein the non-specific crosslinking is to form a covalent bond between the active photoreactive moiety of the photoreactive probe and an alpha-carbon of an amino acid of the protein.

14. An image-guided photolabeling method comprising:

(a) delivering a photoreactive probe to a sample, wherein the photoreactive probe comprises:

a detectable tag portion which is biotin, or another biotin-based moiety; and

a photoreactive moiety which can absorb UV light and become active to crosslink with an amino acid,

wherein the detectable tag portion and the photoreactive moiety are coupled by a linker or a chemical bond;

(b) imaging the sample in a first field of view and acquiring at least one image in the first field of view of the sample with an imaging light source and a camera;

(c) processing the at least one image and determining a region of interest in the first field of view;

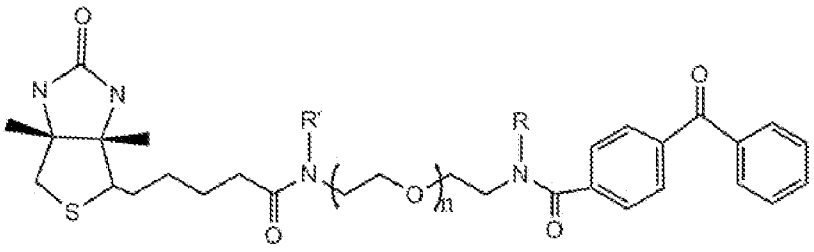
(d) obtaining coordinate information of the region of interest;

(e) according to the coordinate information, selectively illuminating the region of interest with optical radiation to activate the photoreactive moiety of the photoreactive probe and thereby form a non-specific crosslinking between the photoreactive moiety of the photoreactive probe and a plurality of amino acids of different proteins of the sample in the region of interest; and

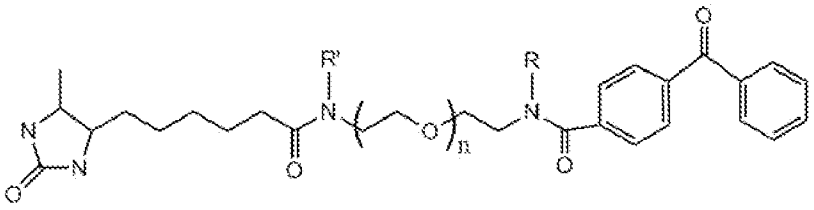
(f) after the region of interest of the first field of view has been illuminated, moving to a second field of view of the sample, and repeating the steps (b) to (e) until all the field of views of the sample have been fully illuminated.

15. The method of claim 12, wherein the photoreactive moiety is one selected from the group consisting of benzophenone, phenyl azide, phenyl diazirine, tetrafluorophenyl azide, hydroxyphenyl azide, and trifluoromethylphenyl diazirine.

16. The method of claim 12, wherein the photoreactive probe is represented by formula (I-1) or (I-2):



(I-1)



(I-2),

wherein each of R and R' independently is hydrogen, an alkyl, or a nitrogen protecting group, and n is an integer of 1-20.

17. An analytical method for a probe-labeled protein, comprising:

obtaining a sample;

delivering a photoreactive probe to the sample, wherein the photoreactive probe is represented by formula (I):

T—L—W (I),

wherein

the T portion is biotin or another biotin-based moiety;

the L portion is a chemical bond or a linker; and

the W portion is a photoreactive moiety which can absorb UV light and become active to crosslink with an amino acid;

selectively illuminating a region of interest of the sample with optical radiation to activate the photoreactive moiety of the photoreactive probe in the region of interest and form a non-specific crosslinking between the photoreactive moiety and a plurality of the amino acids of different proteins of the sample in the region of interest; and

isolating the plurality of probe-labeled proteins from the sample through an affinity precipitation between the T portion of the photoreactive probe and a plurality of affinity beads.

18. The method of claim 17, wherein after the step of isolating, the method further comprises:

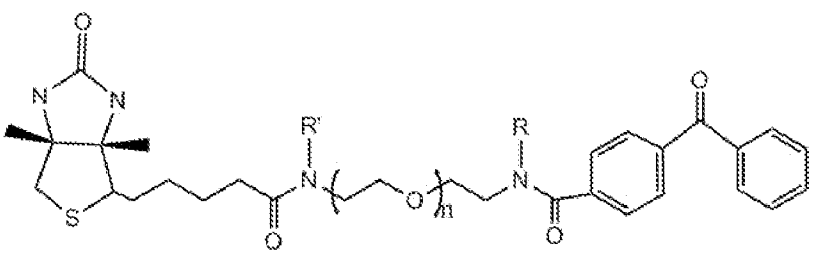
subjecting the plurality of isolated proteins to mass spectrometry analysis; and

identifying the plurality of isolated proteins of the sample.

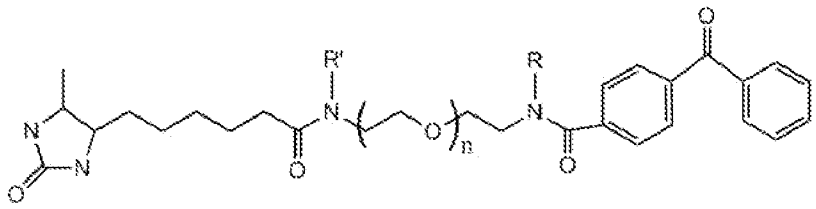
19. The method of claim 17, wherein after the step of selectively illuminating, the method further comprises removing the unbound photoreactive probes from the sample.

20. The method of claim 17, wherein the photoreactive moiety is one selected from the group consisting of benzophenone, phenyl azide, phenyl diazirine, tetrafluorophenyl azide, hydroxyphenyl azide, and trifluoromethylphenyl diazirine.

21 The method of claim 17, wherein the photoreactive probe is represented by formula (I-1) or (I-2):



(I-1)



(I-2),

wherein each of R and R' independently is hydrogen, an alkyl, or a nitrogen protecting group, and n is an integer of 1-20.

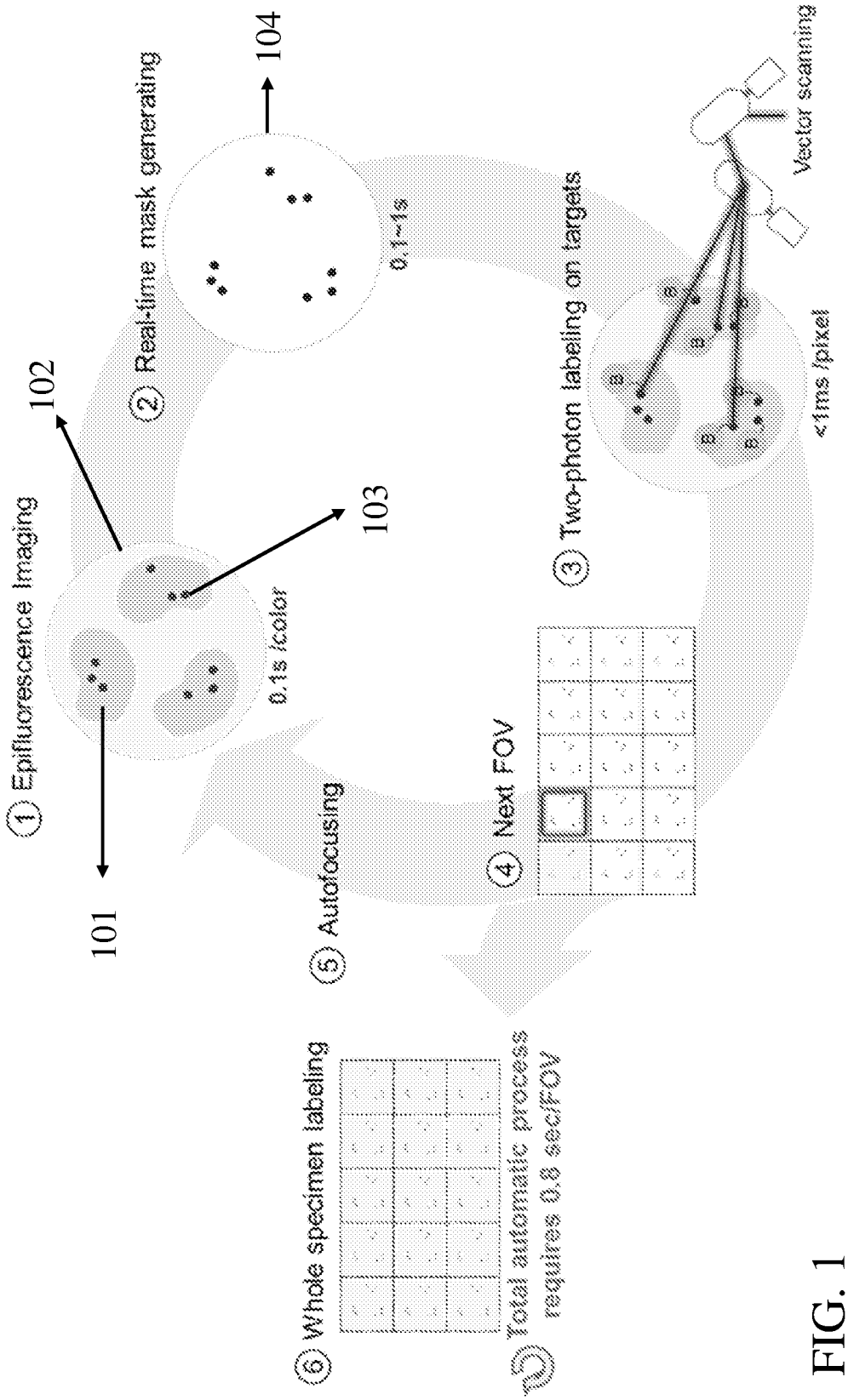

① Epifluorescence Imaging
101
102
0.1s /color
② Real-time mask generating
104
0.1~1s
③ Two-photon labeling on targets
Vector scanning
<1ms /pixel
④ Next FOV
103
⑤ Autofocusing
⑥ Whole specimen labeling
Total automatic process requires 0.8 sec/FOV

FIG. 1

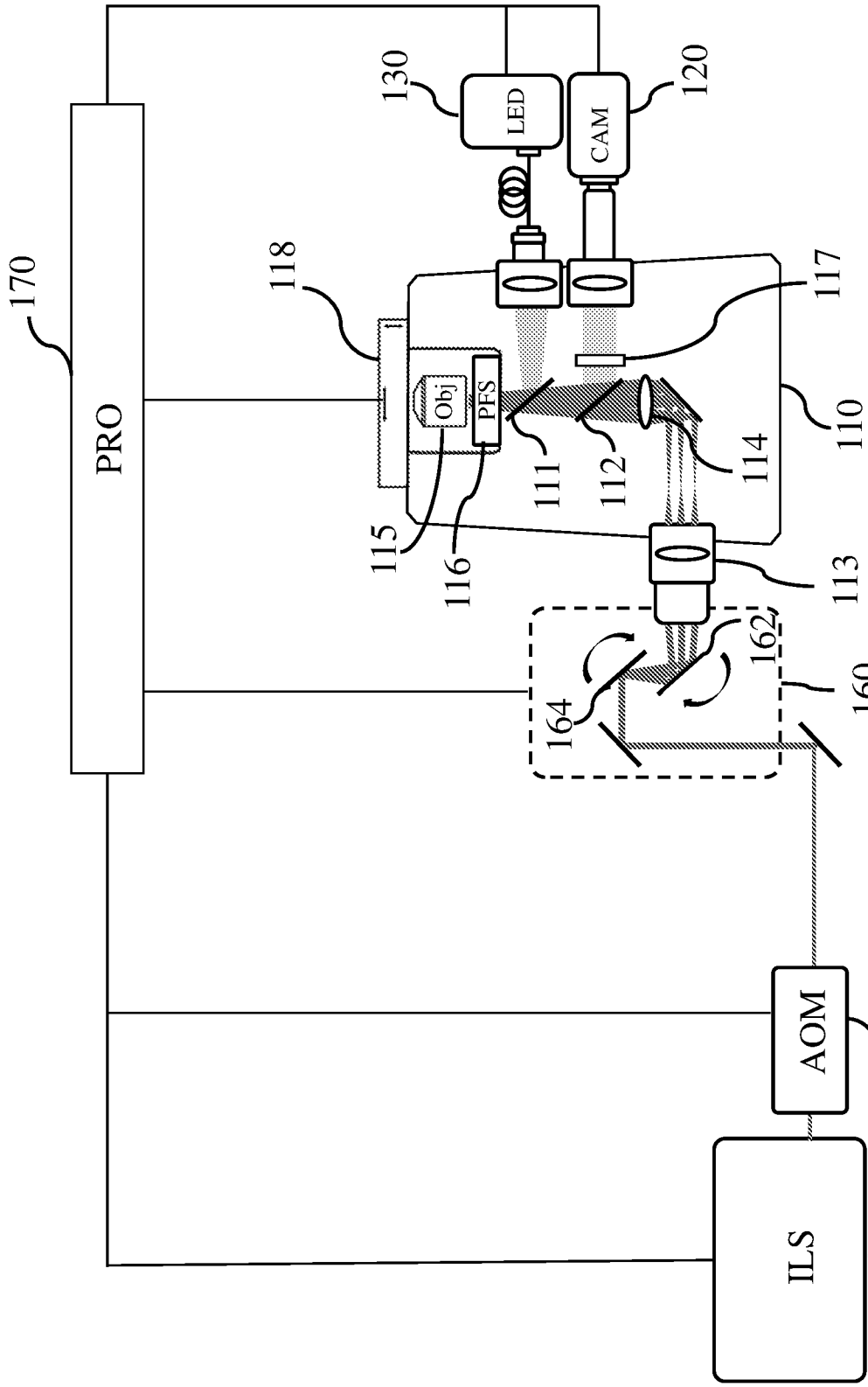

170
PRO
130
LED
120
CAM
118
117
Obj
PFS
111
112
114
110
115
116
113
162
160
164
AOM
150
ILS
140

FIG. 2

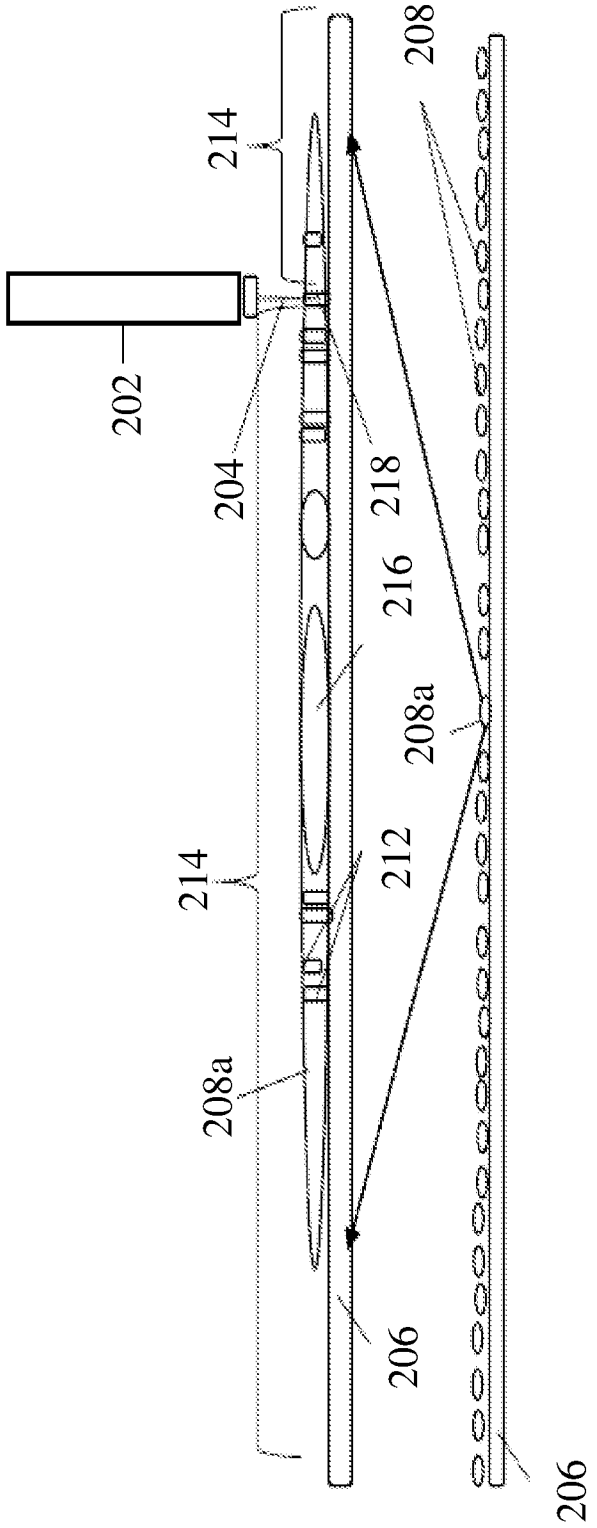
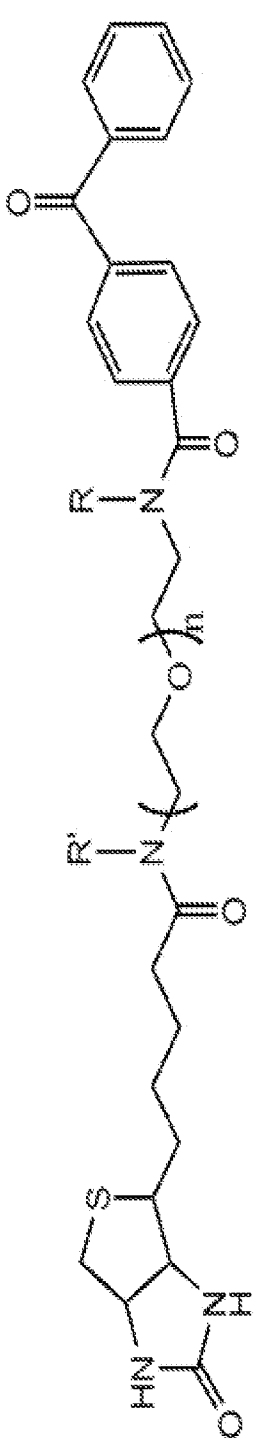


FIG. 3

(I-1)



(I-2)

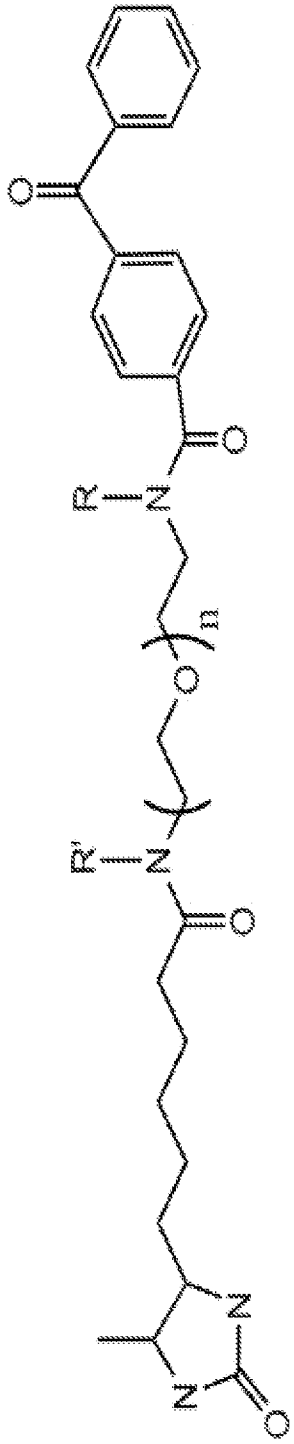


FIG. 4A

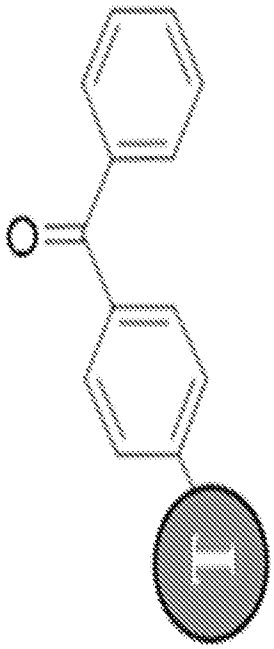


FIG. 4B

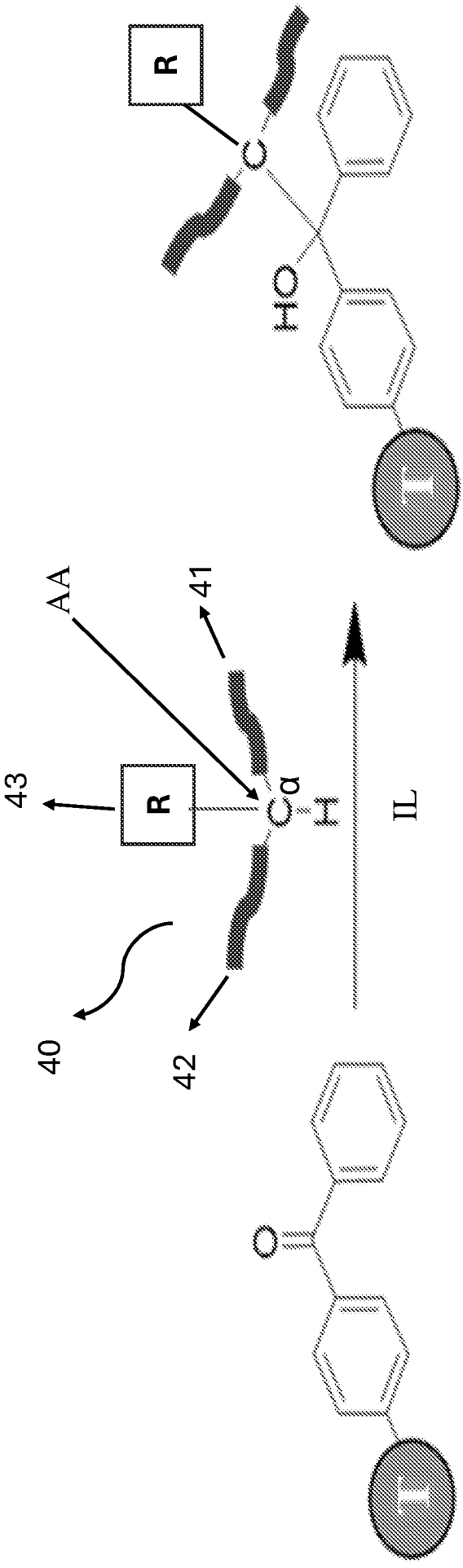


FIG. 4C

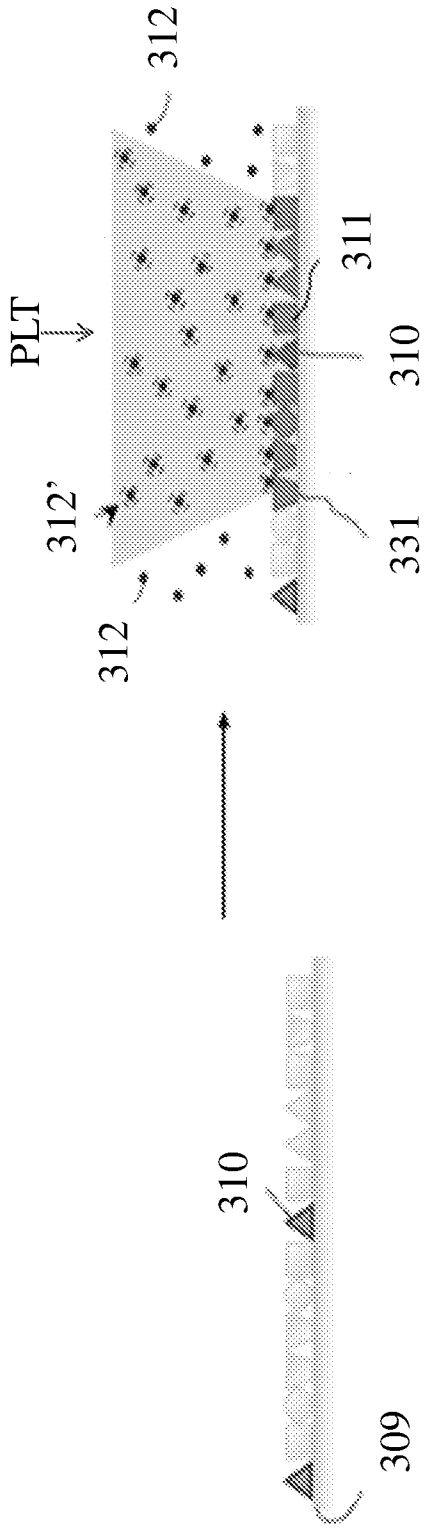


FIG. 4D

B3-BzP

DB3-BzP

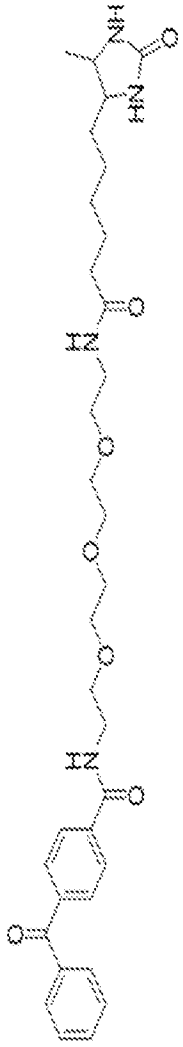


FIG. 4E

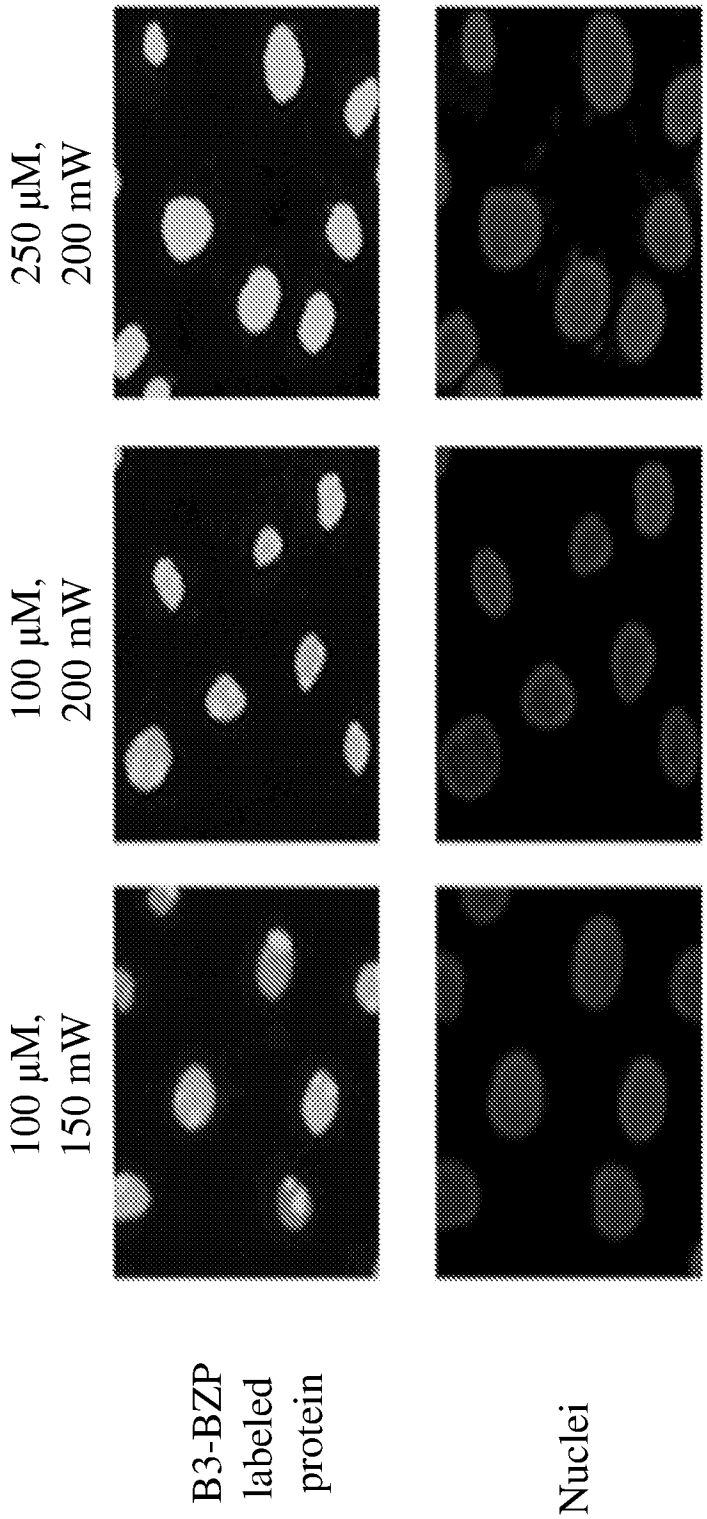

100 µM, 150 mW
100 µM, 200 mW
250 µM, 200 mW
B3-BZP labeled protein
Nuclei

FIG. 5

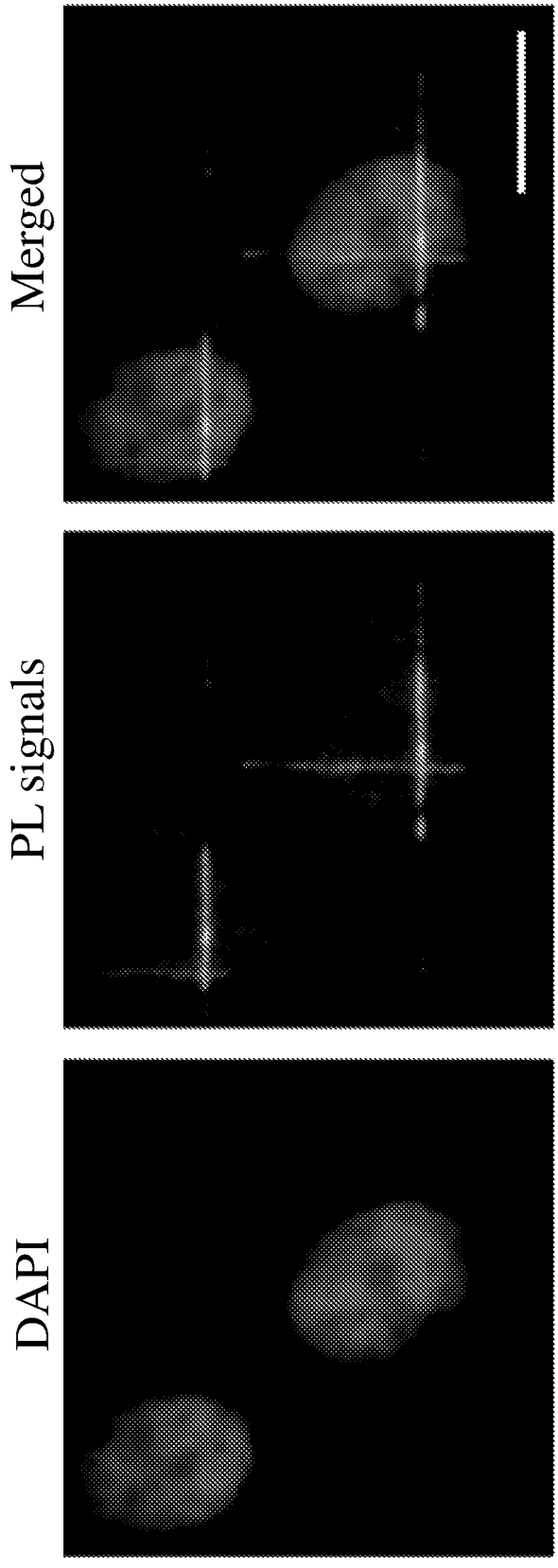


FIG. 6A

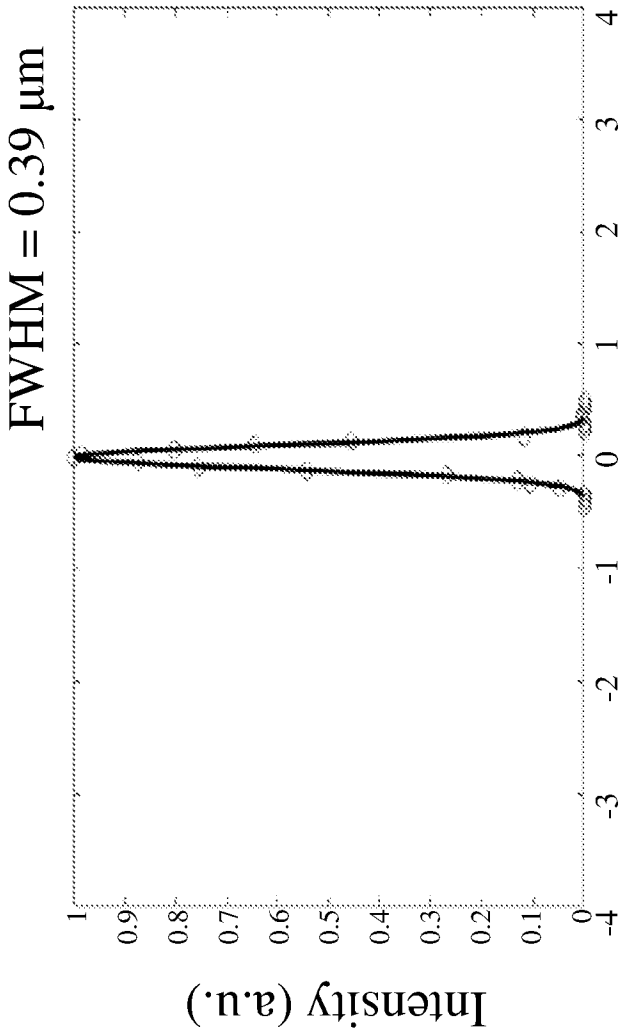

FWHM = 0.39 µm
Intensity (a.u.)
Distance (µm)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-4
-3
-2
-1
0
1
2
3
4

FIG. 6B

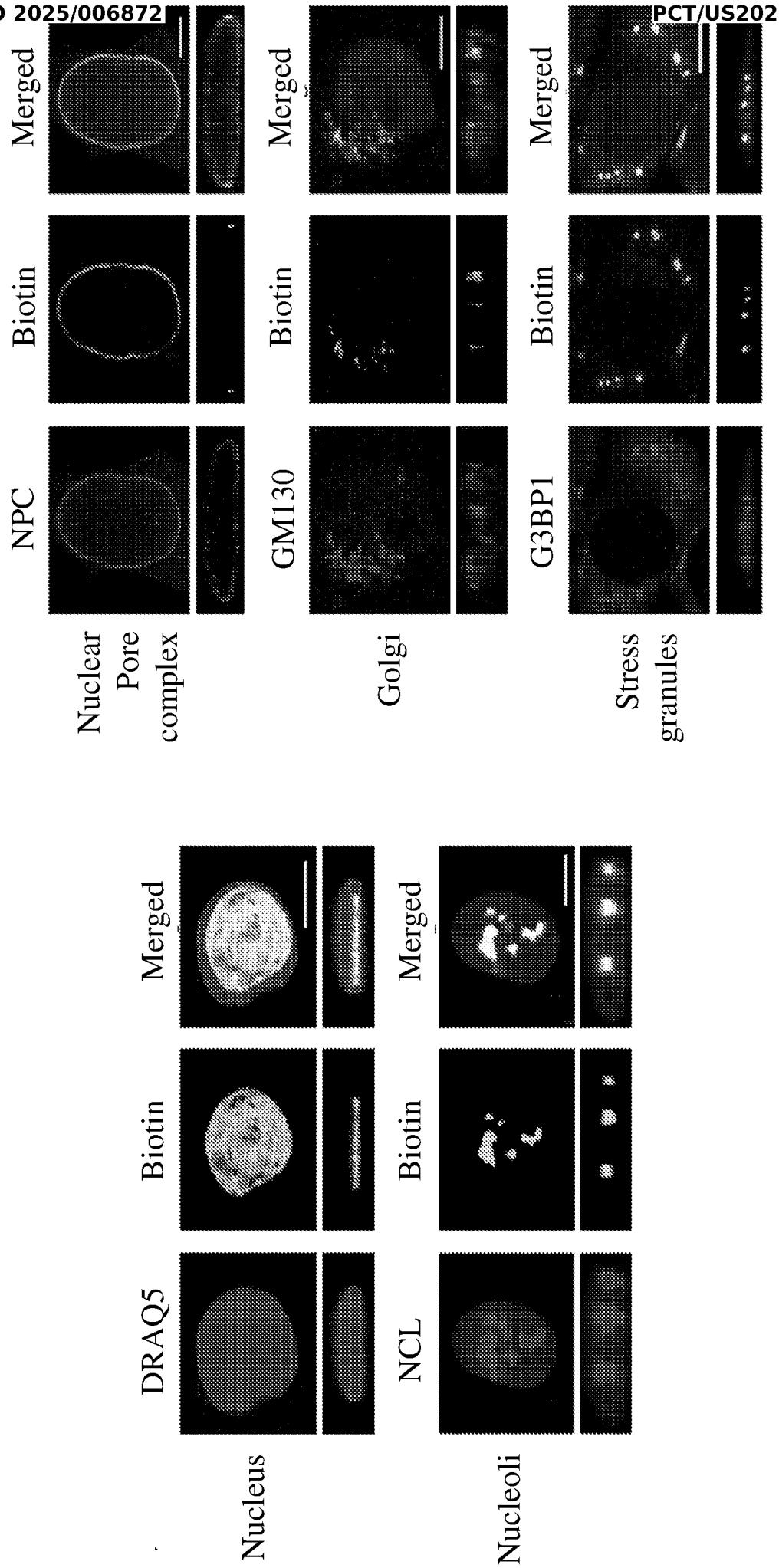

Nucleus
Nucleoli
Nuclear Pore complex
Golgi
Stress granules
DRAQ5
Biotin
Merged
NCL
Biotin
Merged
NPC
Biotin
Merged
GM130
Biotin
Merged
G3BP1
Biotin
Merged

FIG. 6C

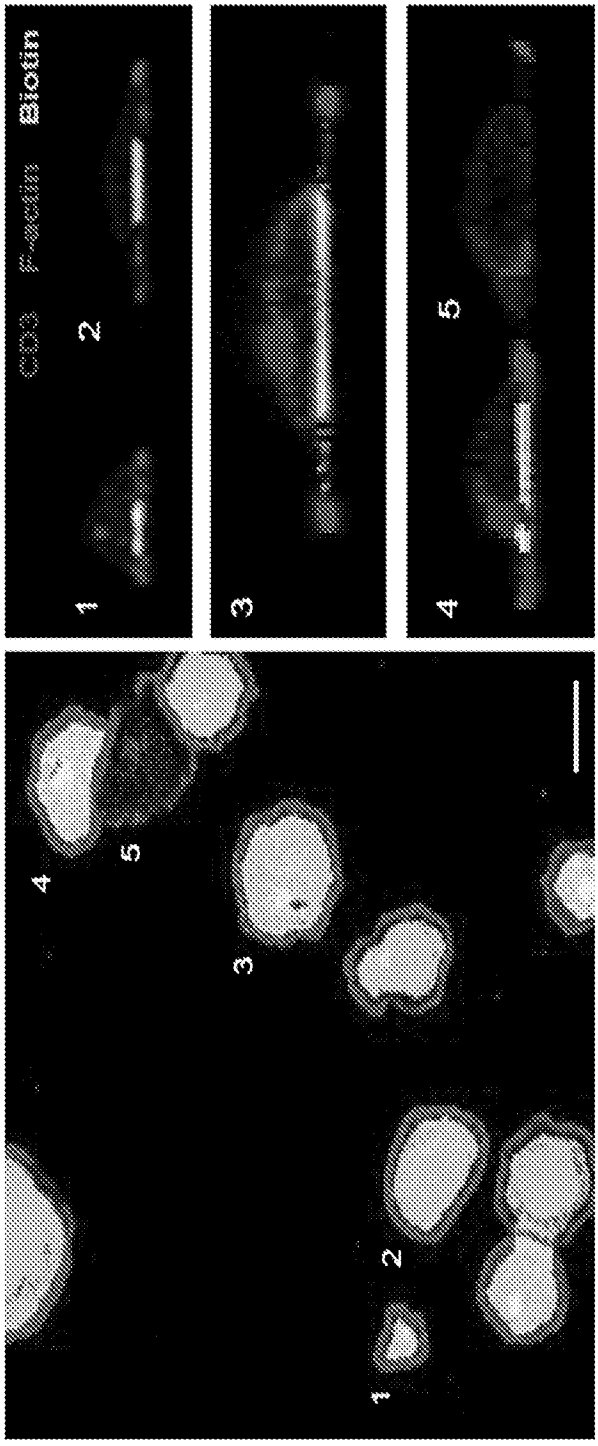


FIG. 6D

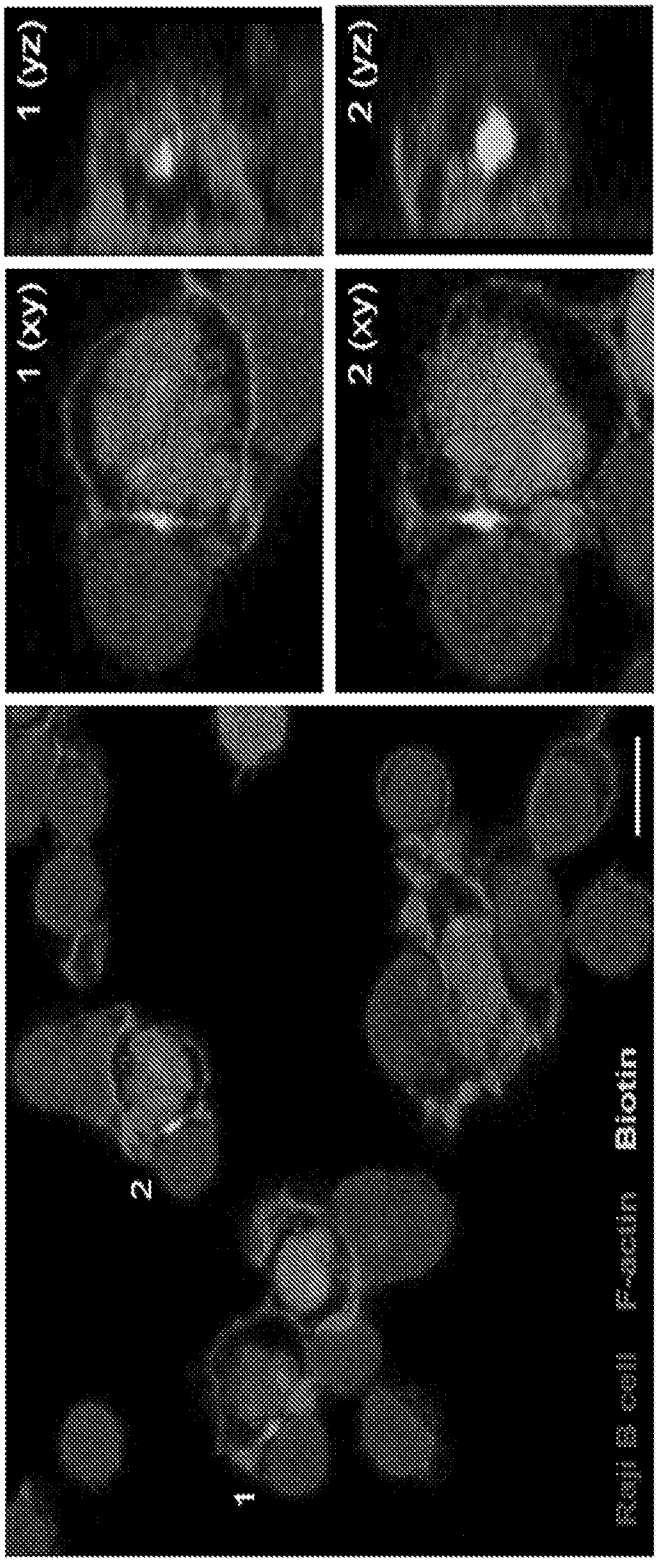


FIG. 6E

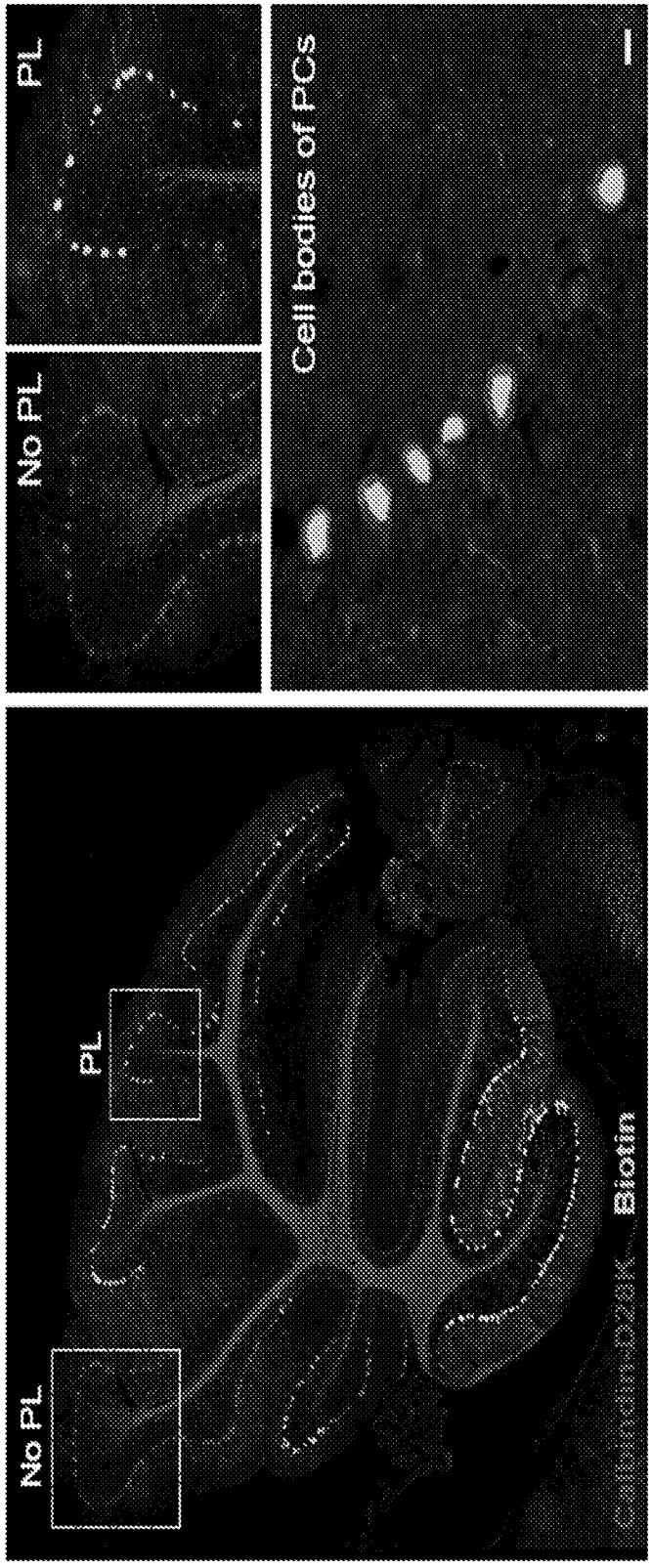


FIG. 6F

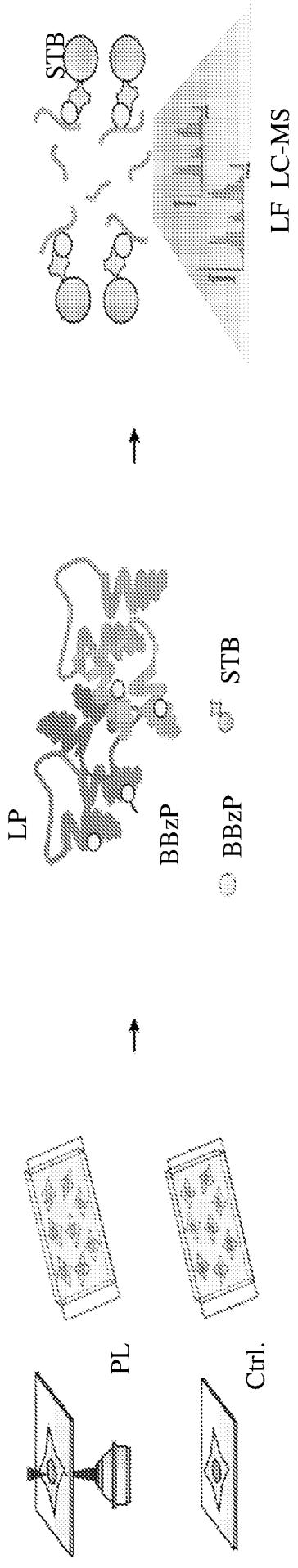


FIG. 7A

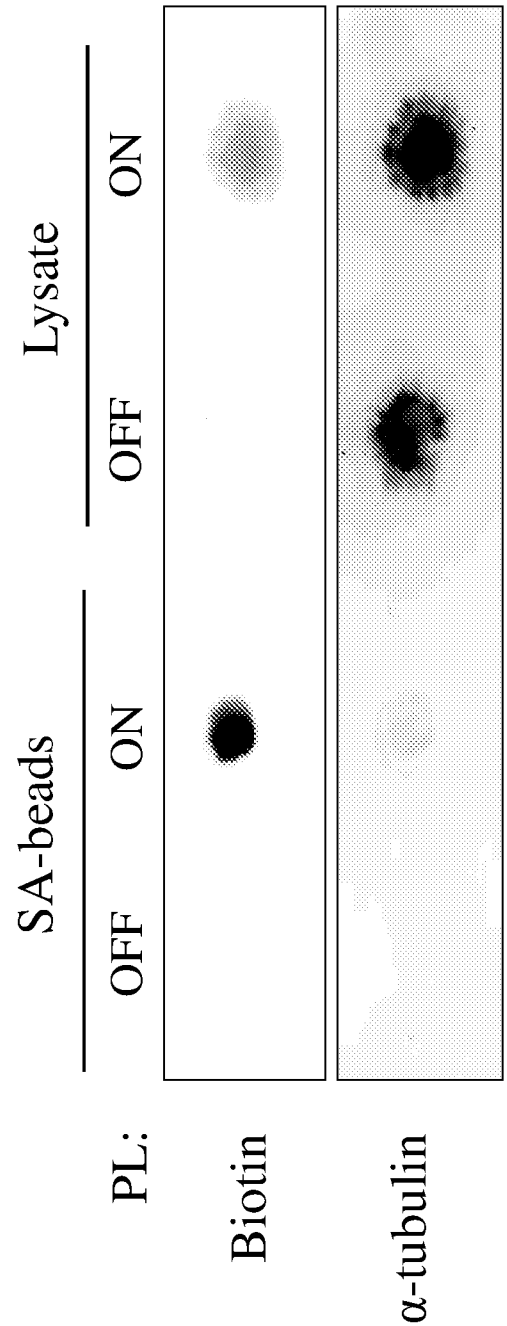


FIG. 7B

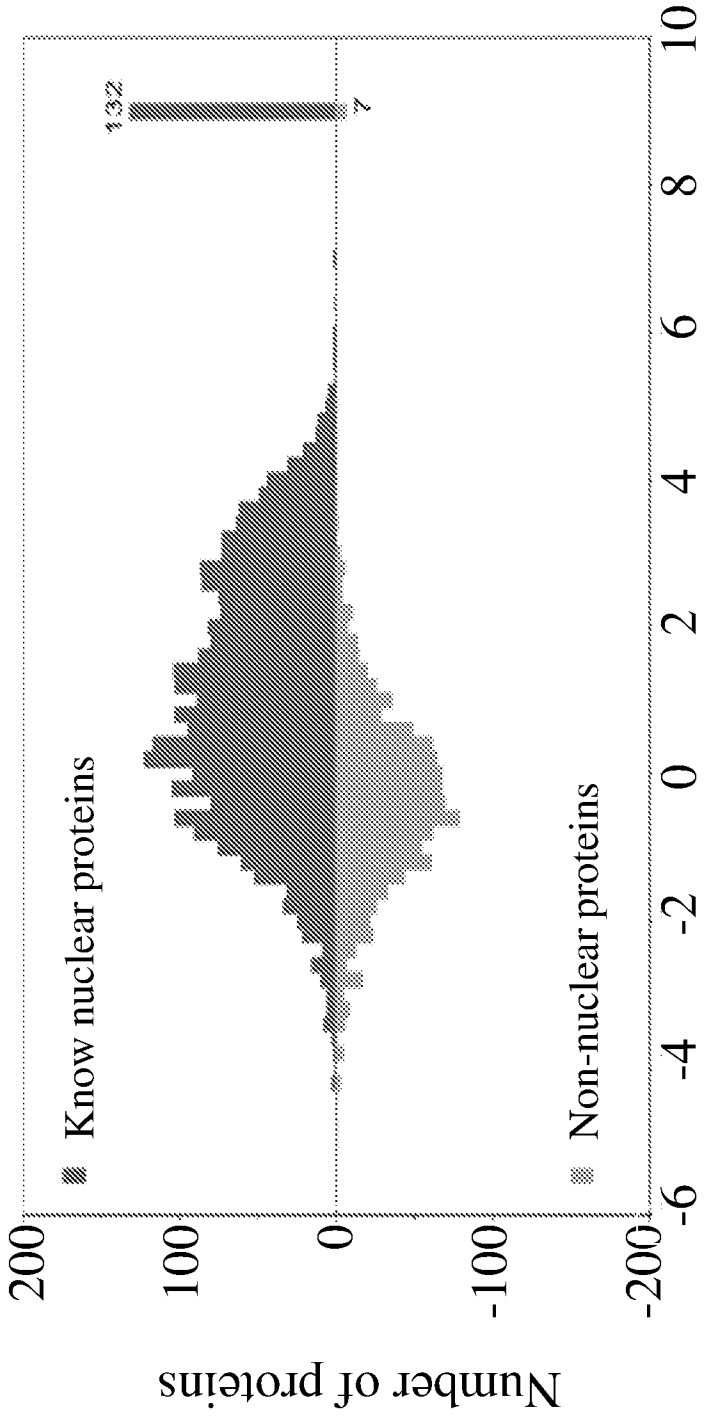

Know nuclear proteins
Non-nuclear proteins
132
7
200
100
0
-100
-200
-6
-4
-2
0
2
4
6
8
10
Number of proteins
LOG2 (PL/CTL)

FIG. 7C

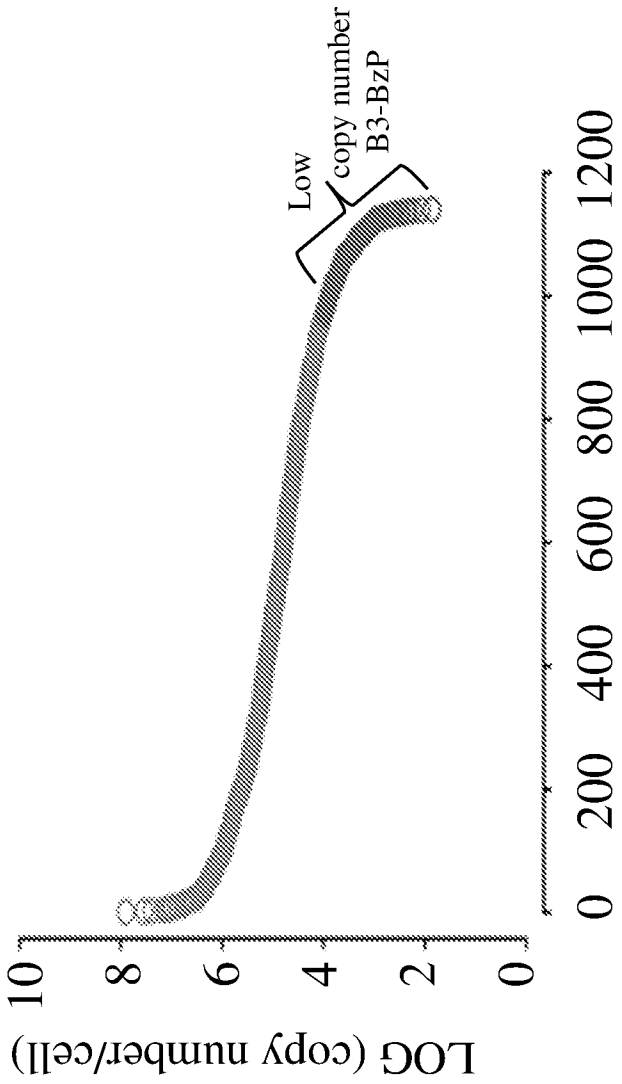

LOG (copy number/cell)
10
8
6
4
2
0
0
200
400
600
800
1000
1200
RANK
Low copy number B3-BzP

FIG. 7D

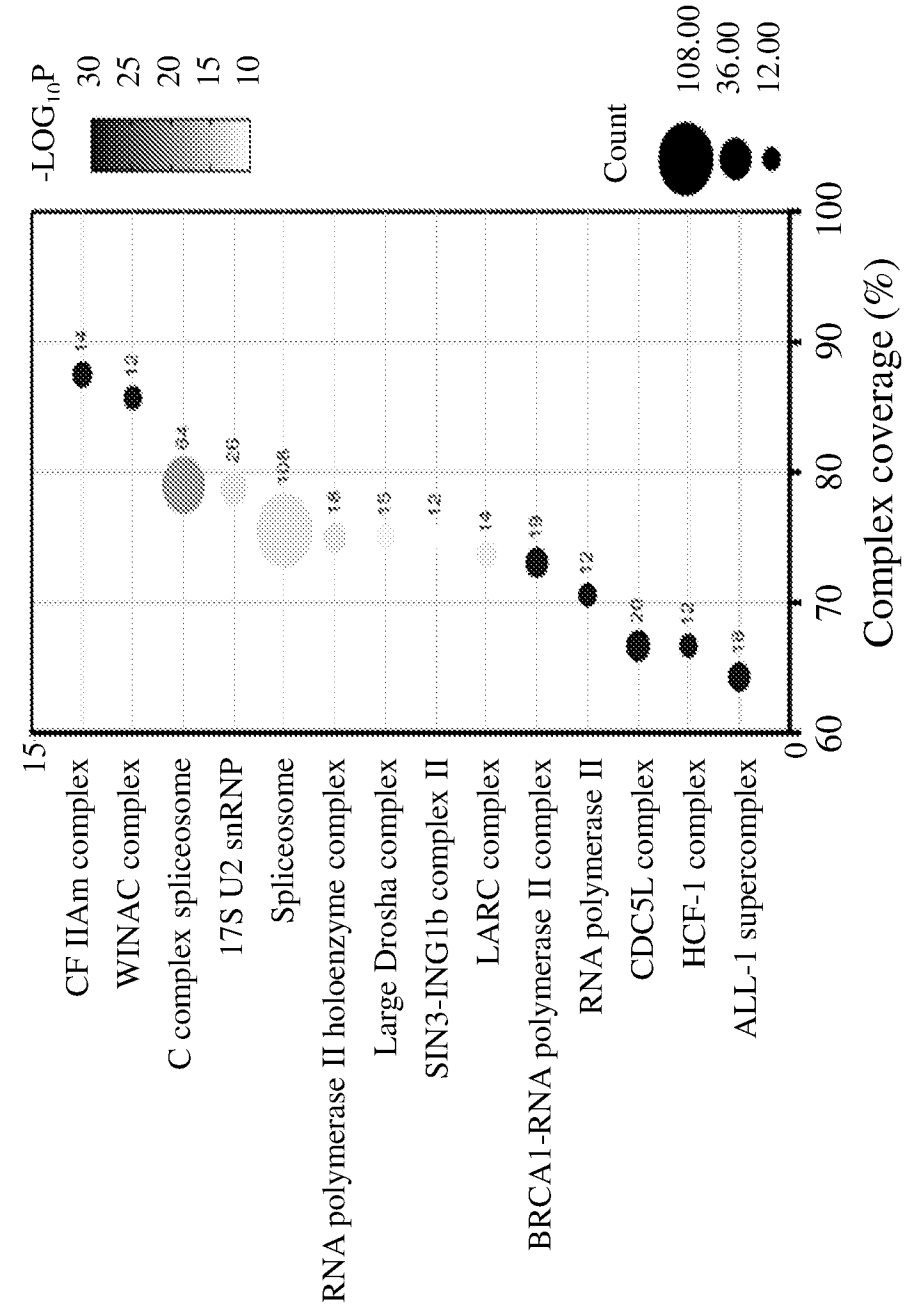

-LOG10P
30
25
20
15
10
Count
108.00
36.00
12.00
CF IIAm complex
WINAC complex
C complex spliceosome
17S U2 snRNP
Spliceosome
RNA polymerase II holoenzyme complex
Large Drosha complex
SIN3-ING1b complex II
LARC complex
BRCA1-RNA polymerase II complex
RNA polymerase II
CDC5L complex
HCF-1 complex
ALL-1 supercomplex
15
0
60
70
80
90
100
Complex coverage (%)

FIG. 7E

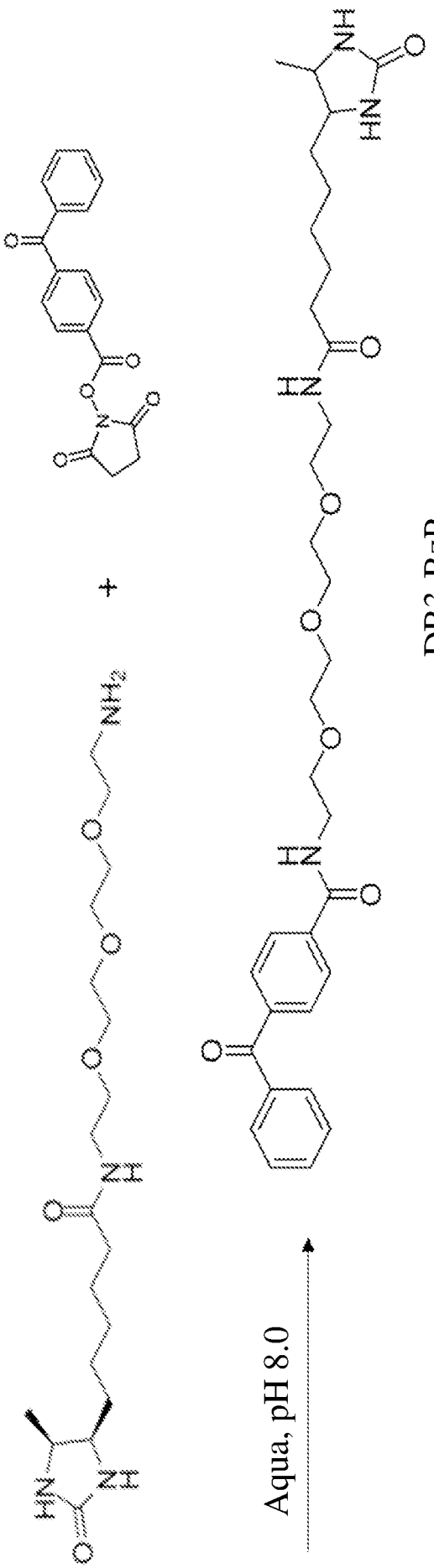
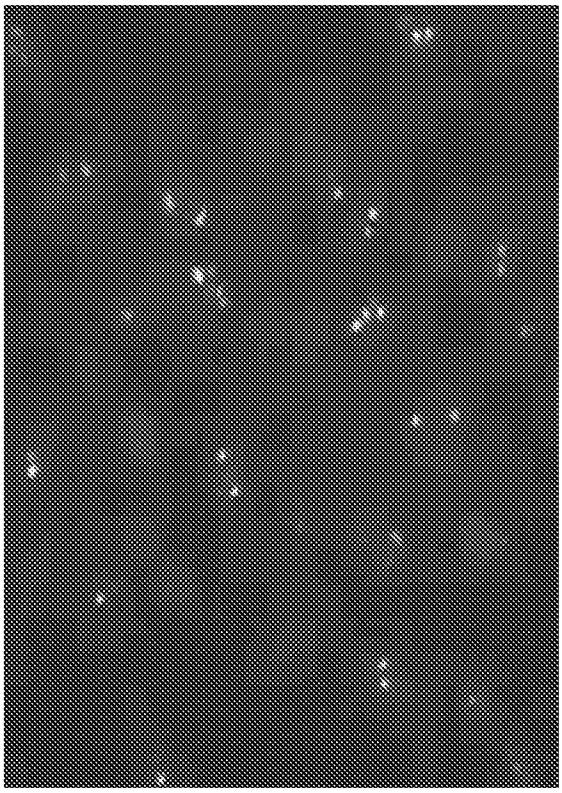
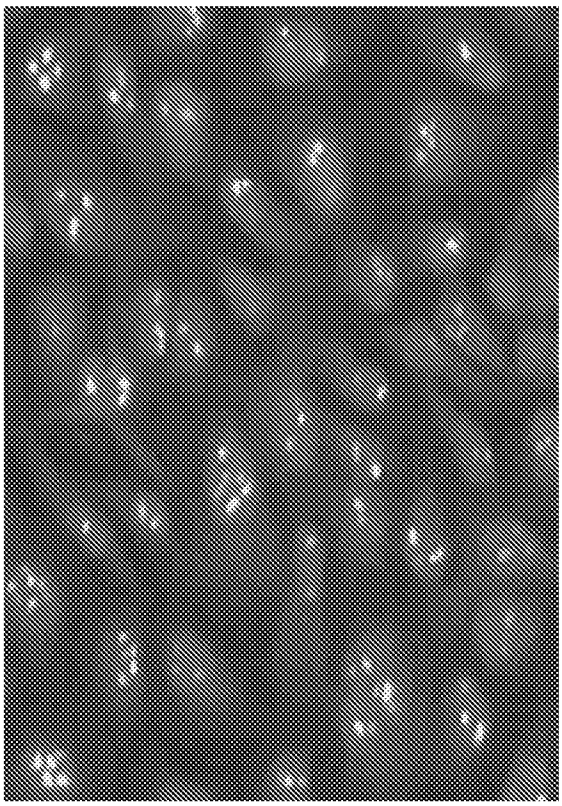


FIG. 8



DB3-BzP



B3-BzP

FIG. 9