

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

(19) World Intellectual Property

Organization

International Bureau

(43) International Publication Date

21 November 2019 (21.11.2019)



(10) International Publication Number

WO 2019/219002 A1

(51) International Patent Classification:

C07K 1/00 (2006.01)

C07D 217/00 (2006.01)

(21) International Application Number:

PCT/CN2019/086825

(22) International Filing Date:

14 May 2019 (14.05.2019)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/670,919

14 May 2018 (14.05.2018)

US

(71) Applicant: VISION GLOBAL HOLDINGS LIMITED

[CN/CN]; 18/F., Chevalier Commercial Centre, 8 Wang Hoi Road, Kowloon Bay, Hong Kong (CN).

(72) Inventors: KWOK, Sui-Yi; 18/F., Chevalier Commercial Centre, 8 Wang Hoi Road, Kowloon Bay, Hong Kong (CN).

WAI, Norman Fung-Man; 18/F., Chevalier Commercial Centre, 8 Wang Hoi Road, Kowloon Bay, Hong Kong (CN).

WONG, Man-Kin; Flat E, 37/F, Block 2, Cayman Rise, 29 Ka Wai Man Road, Kennedy Town, Hong Kong (CN).

(74) Agent: BEIJING PAT SF INTELLECTUAL PROPER-

TY AGENCY CO. LTD; 1110-11, A, Floor 11, Chengjian Plaza, 18 Beitaipingzhuang Road, Haidian District, Beijing 100088 (CN).

(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).

Published:

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

— with sequence listing part of description (Rule 5.2(a))

(54) Title: METHOD FOR N-TERMINAL FUNCTIONALIZATION OF POLYPEPTIDES AND PRODUCTS THEREOF

(57) Abstract: Provided herein are reagents and methods useful for N-terminal functionalization of polypeptides and intermediates and products thereof.



WO 2019/219002 A1

Method for N-terminal Functionalization of Polypeptides and Products Thereof

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the benefit of priority of United States Provisional Application Number 62/670,919, filed on May 14, 2018, the contents of which being hereby incorporated by reference in their entirety for all purposes.

TECHNICAL FIELD

[01] The present disclosure relates to methods for functionalization of polypeptides, such as proteins and antibodies, and products thereof.

BACKGROUND

[02] Covalent conjugation of a biomolecule to a therapeutic agent (e.g., chemotherapeutic/radiotherapeutic agent) opens the door to a new generation biotherapeutic agents with improved pharmaceutical properties.

[03] Among these, conjugation of a therapeutic agent to biomolecules as done with antibody-drug conjugates, polymer conjugates, immobilized biomaterials and controlled release of bioconjugates are excellent demonstrations of the potential of protein-based conjugates to improve efficacy, pharmacokinetics, biodistribution and safety of bioconjugates.

[04] Despite the vast therapeutic potential of bioconjugates, current approaches still suffer from unselective chemical modification, which results in heterogeneous chemical modification of the protein target. The resulting heterogeneous biotherapeutic agents lead to analytical complications as well as safety and efficacy concerns. Current antibody-drug conjugates with heterogeneously attached anti-cancer drugs can often times exhibit poor clinical efficacy.

[05] For example, Mylotarg® (Gemtuzumab ozogamicin), the first antibody-drug conjugate approved by the FDA in 2000, includes a non-specific linkage of the cytotoxic agent at one or more free amines present on the antibody. Mylotarg was withdrawn from the market in 2010, and then a post-marketing study for insufficient clinical benefit and adverse side-effects associated with its heterogeneity was conducted. It was found that Mylotarg's drug-antibody ratio strongly impacted Mylotarg's therapeutic efficacy with a large portion of the unlabeled antibody (~50%) serving as a competitive inhibitor to the antibody-drug conjugate for cellular uptake, which is of importance for its activity.

[06] It has been demonstrated that a high drug to antibody ratio generally enhances toxicity and hydrophobicity of the antibody-drug conjugate, thus inducing antibody precipitation, reducing drug stability, and accelerating plasma clearance.

[07] Particularly in the case of antibody-cytotoxic drug conjugates, the statistical conjugation of the cytotoxic drug can affect the therapeutic window and safety of the antibody-drug conjugate. Additional, multiple drug-loaded forms of antibody-drug conjugates (usually with drug-antibody ratio from 0 to 8) can have distinct pharmacokinetics *in vivo* and differ in clinical performance, immunogenicity, clearance and stability.

[08] Purifying antibody-drug conjugates can be particularly challenging in instances where there are multiple sites on the antibody capable of forming bonds with the drug. For example, the side chains of 70 out of 88 lysines in Kadcyla® (Trastuzumab emtansine) are conjugated, leading to the generation of a complex mixture of antibody-drug conjugate species with significant differences in clinical performance and pharmacokinetic properties.

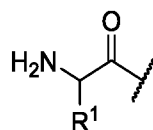
[009] The *N*-terminal α -amino group of polypeptide is a unique reactive site, which can render site-specific selectivity in polypeptide bioconjugation possible. Compared with the side chain amine of lysine with pK_a of 10, pK_a of *N*-terminal α -amino group ranges from 6 to 8. Additionally, the *N*-terminal residues are usually exposed on the surface of proteins over an estimated 80% of cases. *N*-terminal modification of a polypeptide with a therapeutic agent (e.g., chemotherapeutic/radiotherapeutic agent) *via* a highly *N*-terminal selective chemical linker is therefore a powerful approach to polypeptide modification as most proteins contain a unique *N*-terminus in their amino acid sequences and no extra bioorthogonal handles, such as ketones, aldehydes, azides and alkynes, are required to be genetically and/or chemically incorporated into the protein before coupling with a therapeutic agent (e.g., chemotherapeutic/radiotherapeutic agent).

[010] In view of at least the foregoing challenges, there exists a need to develop improved methods for the selective modification of polypeptides, which e.g., reduce product heterogeneity providing a reliable therapeutic tool with predictable properties and batch-to-batch consistency.

SUMMARY

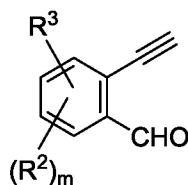
[011] Provided herein are compositions and methods useful for *N*-terminal selective functionalization of polypeptides and intermediates and products thereof. The methods provided herein are highly *N*-terminal selective and are capable of yielding *N*-terminal functionalized polypeptides with site selectivity of 99% or greater. Most protein-based therapeutic agents are heterogeneous mixtures. The method described herein provides a reliable tool to reduce the overall product heterogeneity, and hence deliver more consistent pharmaceutical and/or therapeutic properties.

[012] In a first aspect, provided herein is a method for preparing a functionalized polypeptide comprising the step of: contacting a polypeptide comprising an N-terminal amino acid represented by the structure shown below:



[013] or a conjugate salt or zwitterion thereof, wherein

[014] R^1 is the side chain of the N-terminal amino acid of the polypeptide and the N-terminal amino acid is a natural amino acid or a non-natural amino acid; with a compound of Formula 1:



1

[015] or a conjugate salt thereof, wherein:

[016] m is 1, 2, or 3; and

[017] each R^2 is independently for each instance selected from the group consisting of hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl, aralkyl, heterocycloalkyl, heteraryl, halide, cyano, nitro, hydroxyl, mercaptan, $-OR^5$, $-SR^5$, $-(C=O)OR^5$, $-O(C=O)R^5$, $-N(R^5)_2$, $-(C=O)N(R^5)_2$, $-N(R^5)(C=O)R^5$, $-N(R^5)(C=O)N(R^5)_2$, $-SO_2R^5$, $-N(R^5)SO_2R^5$, and $-SO_2N(R^5)_2$;

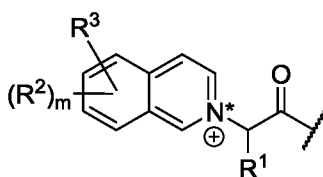
[018] R^3 is $-A-R^4$, wherein A is a linker or absent;

[019] R^4 is selected from the group consisting of hydroxyl, thiol, aldehyde, ketone, maleimide, *N*-hydroxysuccinimide ester, imidoester, haloacetyl, pyridyl disulfide,

alkyne, azide, nitron, nitrile oxide, norbornene, oxanorbornadiene, tetrazine, cyclooctene, isocyanide, quadricyclane, and bis(dithiobenzil)nickel(II); and

[020] R^5 for each instance is independently selected from the group consisting of hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl, aralkyl, heterocycloalkyl, and heteraryl; or two instances of R^5 taken together with the atoms to which they are attached form a 5-6 membered heterocycloalkyl;

[021] thereby forming the functionalized polypeptide, wherein the functionalized polypeptide is represented by the structure of Formula 2:



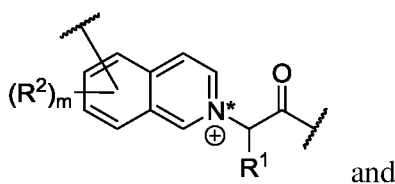
2,

[022] wherein N* is the N-terminal nitrogen of the functionalized polypeptide.

[023] In a first embodiment of the first aspect, provided herein is the method of the first aspect, wherein m is 1 and R^2 is hydrogen, $-OR^5$, or halide.

[024] In a second embodiment of the first aspect, provided herein is the method of the first aspect, wherein the linker is selected from the group consisting of $^{**}-(CR_2)_n-^{***}$, $^{**}-O(CR_2)_n-^{***}$, $^{**}-(CR_2)_nO-^{***}$, $^{**}-O(CR_2)_nO-^{***}$, $^{**}-(CR_2)_nC(=O)-^{***}$, $^{**}-C(=O)(CR_2)_n-^{***}$, $^{**}-C(=O)O(CR_2)_n-^{***}$, $^{**}-(CR_2)_nOC(=O)-^{***}$, $^{**}-O(CR_2)_nC(=O)-^{***}$, $^{**}-O(CR_2)_nC(=O)N(R)(CR_2)_p-^{***}$, $^{**}-O(CR_2)_nN(R)C(=O)N(R)(CR_2)_p-^{***}$, $^{**}-O(CR_2)_nC(=O)O(CR_2)_p-^{***}$, $^{**}-O(CR_2)_nOC(=O)N(R)(CR_2)_p-^{***}$, $^{**}-C(=O)(CR_2)_nO-^{***}$, $^{**}-O(CR_2)_nN(R)C(=O)-^{***}$, $^{**}-O(CR_2)_nN(R)C(=O)(CR_2)_p-^{***}$, $^{**}-C(=O)N(R)(CR_2)_nO-^{***}$, $^{**}-OC(=O)(CR_2)_n-^{***}$, $^{**}-(CR_2)_nC(=O)O-^{***}$, $^{**}-OC(=O)O(CR_2)_n-^{***}$, $^{**}-(CR_2)_nOC(=O)O-^{***}$, $^{**}-C(=O)N(R)(CR_2)_n-^{***}$, $^{**}-$

$(\text{CR}_2)_n(\text{R})\text{NC}(=\text{O})-$ ^{***}, $-\text{N}(\text{R})\text{C}(=\text{O})(\text{CR}_2)_n-$ ^{***}, $-(\text{CR}_2)_n\text{C}(=\text{O})\text{N}(\text{R})-$ ^{***}, $-\text{N}(\text{R})\text{C}(=\text{O})\text{O}(\text{CR}_2)_n-$ ^{***}, $-(\text{CR}_2)_n\text{OC}(=\text{O})\text{N}(\text{R})-$ ^{***}, $-\text{OC}(=\text{O})\text{N}(\text{R})(\text{CR}_2)_n-$ ^{***}, $-(\text{CR}_2)_n\text{N}(\text{R})\text{C}(=\text{O})\text{O}-$ ^{***}, $-(\text{OCR}_2\text{CR}_2)_n-$ ^{***}, $-(\text{CR}_2\text{CR}_2\text{O})_n-$ ^{***}, $-(\text{OCR}_2\text{CR}_2)_n(\text{C}=\text{O})-$ ^{***}, $-(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C}=\text{O})-$ ^{***}, $-(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})-$ ^{***}, $-(\text{OCR}_2\text{CR}_2)_n(\text{C}=\text{O})(\text{CR}_2)_p-$ ^{***}, $-(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C}=\text{O})(\text{CR}_2)_p-$ ^{***}, $-(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})(\text{CR}_2)_p-$ ^{***}, $-(\text{OCR}_2\text{CR}_2)_n(\text{C}=\text{O})\text{O}(\text{CR}_2)_p-$ ^{***}, $-(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C}=\text{O})\text{O}(\text{CR}_2)_p-$ ^{***}, $-(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})\text{O}(\text{CR}_2)_p-$ ^{***}, $-(\text{OCR}_2\text{CR}_2)_n(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p-$ ^{***}, $-(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p-$ ^{***}, $-(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p-$ ^{***}, $-\text{OCR}_2(\text{C}=\text{O})-$ ^{***}, $-\text{OCR}_2(\text{C}=\text{O})(\text{CR}_2)_p-$ ^{***}, $-\text{OCR}_2(\text{C}=\text{O})\text{O}(\text{CR}_2)_p-$ ^{***}, $-\text{OCR}_2(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p-$ ^{***}, $\text{C}(=\text{O})\text{O}(\text{CR}_2\text{CR}_2\text{O})_n\text{CR}_2\text{CR}_2-$ ^{***}, $-\text{C}(=\text{O})\text{N}(\text{R})(\text{CR}_2\text{CR}_2\text{O})_n\text{CR}_2\text{CR}_2-$ ^{***}, $\text{S}(\text{CR}_2)_n-$ ^{***}, $-(\text{CR}_2)_n\text{S}-$ ^{***}, $-(\text{CR}_2)_n\text{SS}(\text{CR}_2)_p-$ ^{***}, $-\text{SO}_2(\text{CR}_2)_n-$ ^{***}, $(\text{CR}_2)_n\text{SO}_2-$ ^{***}, $-\text{N}(\text{R})\text{SO}_2(\text{CR}_2)_n-$ ^{***}, $-(\text{CR}_2)_n\text{SO}_2\text{N}(\text{R})-$ ^{***}, $-\text{SO}_2\text{N}(\text{R})(\text{CR}_2)_n-$ ^{***}, $-(\text{CR}_2)_n\text{N}(\text{R})\text{SO}_2-$ ^{***}, $-(\text{CR}_2)_n\text{-Ar-(CR}_2)_p-$ ^{***}, $-(\text{CR}_2)_p\text{-Ar-(CR}_2)_n-$ ^{***}, $\text{O}(\text{CR}_2)_n\text{-Ar-(CR}_2)_p-$ ^{***}, and $-(\text{CR}_2)_p\text{-Ar-(CR}_2)_n\text{O}-$ ^{***}, wherein ^{***} indicates the position of a covalent bond with the moiety:



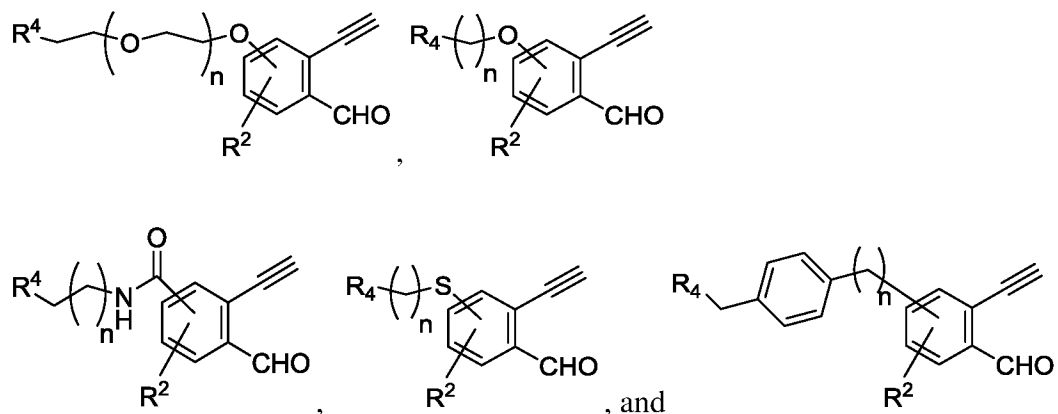
[025] ^{***} indicates the position of a covalent bond with R^4 ; each instance of n is independently a whole number selected from 1-20; each instance of p is independently an integer selected from 0-20; and R for each instance is independently selected from hydrogen, alkyl, cycloalkyl, and aryl; or two instances of R taken together with the carbons to which they are attached form a 3-6 membered

carbocyclic ring; or two instances of R taken together with the atoms to which they are attached form a 5-6 membered heterocyclic ring.

[026] In a third embodiment of the first aspect, provided herein is the method of the first aspect, wherein n is a whole number selected from 1-10 and p is a whole number selected from 1-4.

[027] In a fourth embodiment of the first aspect, provided herein is the method of the second embodiment of the first aspect, wherein R^4 is maleimide, azide, alkyne, hydroxyl, norbornene, or cyclooctene.

[028] In a fifth embodiment of the first aspect, provided herein is the method of the first aspect, wherein the compound of Formula 3 is selected from the group consisting of:



[029] wherein

[030] n is a whole number selected from 1-4;

[031] R^2 is hydrogen, alkyl, or $-OR^5$;

[032] R^4 is maleimide or alkyne; and

[033] R^5 is alkyl.

[034] In a sixth embodiment of the first aspect, provided herein is the method of the first aspect, wherein the functionalized polypeptide comprises at least one lysine residue and less than 5% of the side chains of the at least one lysine residue undergo a reaction with the compound of Formula 3.

[035] In a seventh embodiment of the first aspect, provided herein is the method of the first aspect, wherein the polypeptide is a protein, enzyme, antibody, glycoprotein, or lipoprotein.

[036] In an eighth embodiment of the first aspect, provided herein is the method of the seventh embodiment of the first aspect, wherein the protein is a hemoglobin protein.

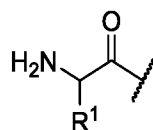
[037] In a ninth embodiment of the first aspect, provided herein is the method of the eighth embodiment of the first aspect, wherein the hemoglobin protein is a hemoglobin tetramer, hemoglobin dimer, or hemoglobin monomer.

[038] In an tenth embodiment of the first aspect, provided herein is the method of the ninth embodiment of the first aspect, wherein the hemoglobin tetramer is a cross-linked hemoglobin tetramer.

[039] In an eleventh embodiment of the first aspect, provided herein is the method of the first aspect further comprising the step of contacting the functionalized polypeptide with an active agent comprising a mercaptan or conjugate salt thereof.

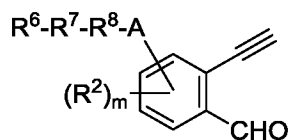
[040] In a twelfth embodiment of the first aspect, provided herein is the method of the eleventh embodiment of the first aspect, wherein the active agent is a chemotherapeutic agent, a radiotherapeutic agent, a fluorescent dye, or affinity tag.

[041] In a second aspect, provided herein is a method for preparing a functionalized polypeptide comprising the step of: contacting a polypeptide comprising an N-terminal amino acid represented by the structure shown below:



[042] or a conjugate salt or zwitterion thereof, wherein

[043] R^1 is the side chain of the N-terminal amino acid of the polypeptide and the N-terminal amino acid is a natural amino acid or a non-natural amino acid; with a compound of Formula 3:



3

[044] or a conjugate salt thereof, wherein

[045] m is 1, 2, or 3;

[046] A is a linker or is absent;

[047] R^2 is independently for each instance selected from the group consisting of hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl, aralkyl, heterocycloalkyl, heteraryl, halide, cyano, nitro, hydroxyl, mercaptan, $-OR^5$, $-SR^5$, $-(C=O)OR^5$, $-O(C=O)R^5$, $-N(R^5)_2$, $-(C=O)N(R^5)_2$, $-N(R^5)(C=O)R^5$, $-N(R^5)(C=O)N(R^5)_2$, $-SO_2R^5$, $-N(R^5)SO_2R^5$, and $-SO_2N(R^5)_2$;

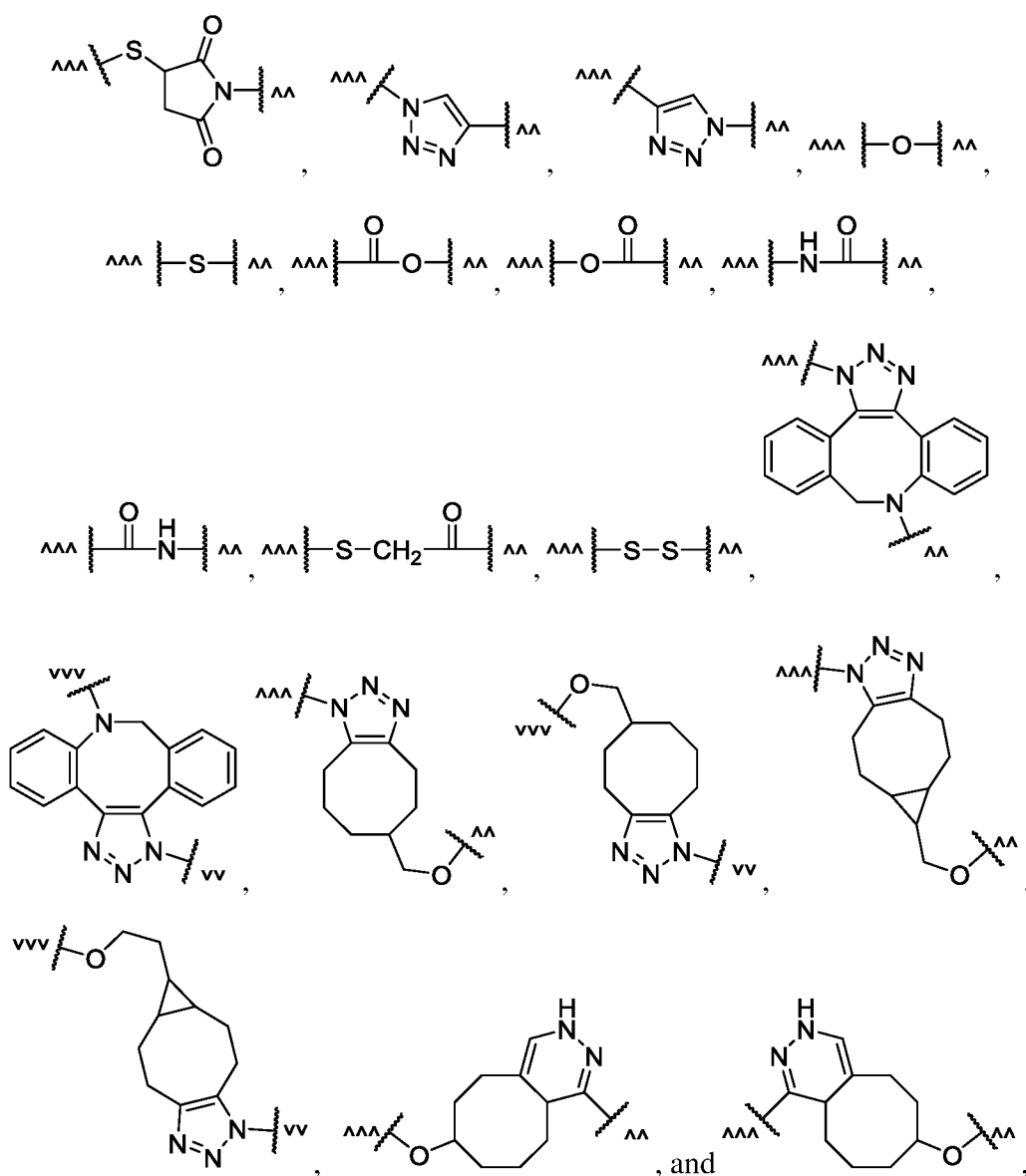
[048] R^5 for each instance is independently selected from the group consisting of hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl, aralkyl, heterocycloalkyl, and

heteraryl; or two instances of R⁵ taken together with the atoms to which they are attached form a 5-6 membered heterocycloalkyl;

[049] R^6 is an active agent;

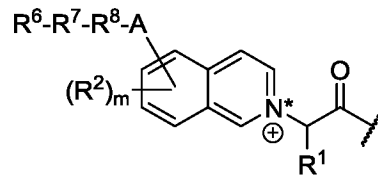
[050] R⁷ is a spacer or absent; and

[051] R^8 is selected from the group consisting of:



[052] wherein $\wedge\wedge$ indicates the position of a covalent bond with A and $\wedge\wedge\wedge$ indicates the position of a covalent bond with R⁷;

[053] thereby forming the functionalized polypeptide, wherein the functionalized polypeptide is represented by the structure of Formula 4:



4

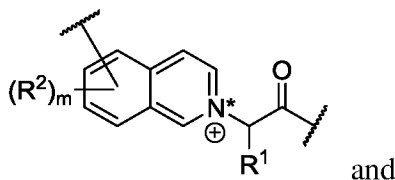
[054] wherein N* is the N-terminal nitrogen of the functionalized polypeptide.

[055] In a first embodiment of the second aspect, provided herein is the method of the second aspect, wherein the spacer is selected from the group consisting of $\text{---}(\text{CR}_2)_n\text{---}$, $\text{---O}(\text{CR}_2)_n\text{---}$, $\text{---}(\text{CR}_2)_n\text{O---}$, $\text{---O}(\text{CR}_2)_n\text{O---}$, $\text{---}(\text{CR}_2)_n\text{C(=O)---}$, $\text{---C(=O)(CR}_2)_n\text{---}$, $\text{---C(=O)O}(\text{CR}_2)_n\text{---}$, $\text{---}(\text{CR}_2)_n\text{OC(=O)---}$, $\text{---O}(\text{CR}_2)_n\text{C(=O)---}$, $\text{O}(\text{CR}_2)_n\text{C(=O)N(R)(CR}_2)_p\text{---}$, $\text{---O}(\text{CR}_2)_n\text{N(R)C(=O)N(R)(CR}_2)_p\text{---}$, $\text{---O}(\text{CR}_2)_n\text{C(=O)O}(\text{CR}_2)_p\text{---}$, $\text{---O}(\text{CR}_2)_n\text{OC(=O)N(R)(CR}_2)_p\text{---}$, $\text{---C(=O)(CR}_2)_n\text{O---}$, $\text{---O}(\text{CR}_2)_n\text{N(R)C(=O)---}$, $\text{---O}(\text{CR}_2)_n\text{N(R)C(=O)(CR}_2)_p\text{---}$, $\text{---C(=O)N(R)(CR}_2)_n\text{O---}$, $\text{---OC(=O)(CR}_2)_n\text{---}$, $\text{---}(\text{CR}_2)_n\text{C(=O)O---}$, $\text{---OC(=O)O}(\text{CR}_2)_n\text{---}$, $\text{---}(\text{CR}_2)_n\text{OC(=O)O---}$, $\text{---C(=O)N(R)(CR}_2)_n\text{---}$, $\text{---}(\text{CR}_2)_n\text{(R)NC(=O)---}$, $\text{---N(R)C(=O)(CR}_2)_n\text{---}$, $\text{---}(\text{CR}_2)_n\text{C(=O)N(R)---}$, $\text{---N(R)C(=O)O}(\text{CR}_2)_n\text{---}$, $\text{---}(\text{CR}_2)_n\text{OC(=O)N(R)---}$, $\text{---OC(=O)N(R)(CR}_2)_n\text{---}$, $\text{---}(\text{CR}_2)_n\text{N(R)C(=O)O---}$, $\text{---}(\text{OCR}_2\text{CR}_2)_n\text{---}$, $\text{---}(\text{CR}_2\text{CR}_2\text{O})_n\text{---}$, $\text{---}(\text{OCR}_2\text{CR}_2)_n\text{(C=O)---}$, $\text{---}(\text{OCR}_2\text{CR}_2)_n\text{O(C=O)---}$, $\text{---}(\text{OCR}_2\text{CR}_2)_n\text{N(R)(C=O)---}$, $\text{---}(\text{OCR}_2\text{CR}_2)_n\text{(C=O)(CR}_2)_p\text{---}$, $\text{---}(\text{OCR}_2\text{CR}_2)_n\text{O(C=O)(CR}_2)_p\text{---}$, $\text{---}(\text{OCR}_2\text{CR}_2)_n\text{N(R)(C=O)(CR}_2)_p\text{---}$, $\text{---}(\text{OCR}_2\text{CR}_2)_n\text{(C=O)O}(\text{CR}_2)_p\text{---}$, $\text{---}(\text{OCR}_2\text{CR}_2)_n\text{O(C=O)O}(\text{CR}_2)_p\text{---}$, $\text{---}(\text{OCR}_2\text{CR}_2)_n\text{N(R)(C=O)O}(\text{CR}_2)_p\text{---}$, $\text{---}(\text{OCR}_2\text{CR}_2)_n\text{(C=O)N(R)(CR}_2)_p\text{---}$, $\text{---}(\text{OCR}_2\text{CR}_2)_n\text{O(C=O)N(R)(CR}_2)_p\text{---}$, $\text{---}(\text{OCR}_2\text{CR}_2)_n\text{N(R)(C=O)N(R)(CR}_2)_p\text{---}$, $\text{---OCR}_2\text{(C=O)---}$, $\text{---OCR}_2\text{(C=O)(CR}_2)_p\text{---}$, $\text{---OCR}_2\text{(C=O)O}(\text{CR}_2)_p\text{---}$, $\text{---OCR}_2\text{(C=O)N(R)(CR}_2)_p\text{---}$, ---

$C(=O)O(CR_2CR_2O)_nCR_2CR_2-$ [‡], $-C(=O)N(R)(CR_2CR_2O)_nCR_2CR_2-$ [‡], $-S(CR_2)_n-$ [‡],
 $- (CR_2)_nS-$ [‡], $-SO_2(CR_2)_n-$ [‡], $-(CR_2)_nSO_2-$ [‡], $-N(R)SO_2(CR_2)_n-$ [‡], $-(CR_2)_nSO_2N(R)-$ [‡],
 $-SO_2N(R)(CR_2)_n-$ [‡], $-(CR_2)_nN(R)SO_2-$ [‡], $-(CR_2)_n-Ar-(CR_2)_p-$ [‡],
 $-(CR_2)_p-Ar-(CR_2)_n-$ [‡], $-O(CR_2)_n-Ar-(CR_2)_p-$ [‡], and $-(CR_2)_p-Ar-(CR_2)_nO-$ [‡],
 wherein [‡] indicates the position of a covalent bond with R⁶ and [‡] indicates the position of a covalent bond with R⁸; each instance of n is independently a whole number selected from 1-20; p is an integer selected from 0-20; and R for each instance is independently selected from hydrogen, alkyl, cycloalkyl, and aryl; or two instances of R taken together with the carbons to which they are attached form a 3-6 membered carbocyclic ring; or two instances of R taken together with the atoms to which they are attached form a 5-6 membered heterocyclic ring.

[056] In a second embodiment of the second aspect, provided herein is the method of the second aspect, wherein the linker is selected from the group consisting of $-Ar-$ [‡], $-(CR_2)_n-$ [‡], $-O(CR_2)_n-$ [‡], $-(CR_2)_nO-$ [‡], $-O(CR_2)_nO-$ [‡], $-(CR_2)_nC(=O)-$ [‡], $-C(=O)(CR_2)_n-$ [‡], $-C(=O)O(CR_2)_n-$ [‡], $-(CR_2)_nOC(=O)-$ [‡], $-O(CR_2)_nC(=O)-$ [‡], $-O(CR_2)_nC(=O)N(R)(CR_2)_p-$ [‡], $-O(CR_2)_nN(R)C(=O)N(R)(CR_2)_p-$ [‡], $-O(CR_2)_nOC(=O)N(R)(CR_2)_p-$ [‡], $-C(=O)(CR_2)_nO-$ [‡], $-O(CR_2)_nN(R)C(=O)-$ [‡], $-O(CR_2)_nN(R)C(=O)(CR_2)_p-$ [‡], $-C(=O)N(R)(CR_2)_nO-$ [‡], $-OC(=O)(CR_2)_n-$ [‡], $-(CR_2)_nC(=O)O-$ [‡], $-OC(=O)O(CR_2)_n-$ [‡], $-(CR_2)_nOC(=O)O-$ [‡], $-C(=O)N(R)(CR_2)_n-$ [‡], $-(CR_2)_n(R)NC(=O)-$ [‡], $-N(R)C(=O)(CR_2)_n-$ [‡], $-(CR_2)_nC(=O)N(R)-$ [‡], $-N(R)C(=O)O(CR_2)_n-$ [‡], $-(CR_2)_nOC(=O)N(R)-$ [‡], $-OC(=O)N(R)(CR_2)_n-$ [‡], $-(CR_2)_nN(R)C(=O)O-$ [‡], $-(OCR_2CR_2)_n-$ [‡], $-(CR_2CR_2O)_n-$ [‡], $-(OCR_2CR_2)_n(C=O)-$ [‡], $-(OCR_2CR_2)_nO(C=O)-$ [‡], $-(OCR_2CR_2)_nN(R)(C=O)-$ [‡], $-(OCR_2CR_2)_n(C=O)(CR_2)_p-$ [‡], $-$ [‡]

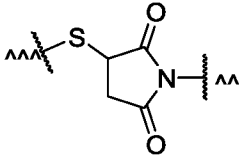
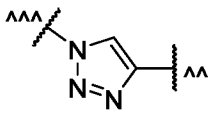
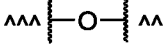
$(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C}=\text{O})$ $(\text{CR}_2)_p$ —^{<<<}, ^{<<}— $(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})(\text{CR}_2)_p$ —^{<<<}, ^{<<}—
 $(\text{OCR}_2\text{CR}_2)_n(\text{C}=\text{O})\text{O}(\text{CR}_2)_p$ —^{<<<}, ^{<<}— $(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C}=\text{O})\text{O}(\text{CR}_2)_p$ —^{<<<}, ^{<<}—
 $(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})\text{O}(\text{CR}_2)_p$ —^{<<<}, ^{<<}— $(\text{OCR}_2\text{CR}_2)_n(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p$ —^{<<<}, ^{<<}—
 $(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p$ —^{<<<}, ^{<<}— $(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p$ —^{<<<}, ^{<<}—
 $\text{OCR}_2(\text{C}=\text{O})$ —^{<<<}, ^{<<}— $\text{OCR}_2(\text{C}=\text{O})(\text{CR}_2)_p$ —^{<<<}, ^{<<}— $\text{OCR}_2(\text{C}=\text{O})\text{O}(\text{CR}_2)_p$ —^{<<<}, ^{<<}—
 $\text{OCR}_2(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p$ —^{<<<}, ^{<<}— $\text{C}(\text{O})\text{O}(\text{CR}_2\text{CR}_2\text{O})_n\text{CR}_2\text{CR}_2$ —^{<<<}, ^{<<}—
 $\text{C}(\text{O})\text{N}(\text{R})(\text{CR}_2\text{CR}_2\text{O})_n\text{CR}_2\text{CR}_2$ —^{<<<}, ^{<<}— $\text{S}(\text{CR}_2)_n$ —^{<<<}, ^{<<}— $(\text{CR}_2)_n\text{S}$ —^{<<<}, ^{<<}— $\text{SO}_2(\text{CR}_2)_n$ —
^{<<<}, ^{<<}— $(\text{CR}_2)_n\text{SO}_2$ —^{<<<}, ^{<<}— $\text{N}(\text{R})\text{SO}_2(\text{CR}_2)_n$ —^{<<<}, ^{<<}— $(\text{CR}_2)_n\text{SO}_2\text{N}(\text{R})$ —^{<<<}, ^{<<}—
 $\text{SO}_2\text{N}(\text{R})(\text{CR}_2)_n$ —^{<<<}, ^{<<}— $(\text{CR}_2)_n\text{N}(\text{R})\text{SO}_2$ —^{<<<}, ^{<<}— $(\text{CR}_2)_n\text{Ar}-(\text{CR}_2)_p$ —^{<<<}, ^{<<}— $(\text{CR}_2)_p$ —
 $\text{Ar}-(\text{CR}_2)_n$ —^{<<<}, ^{<<}— $\text{O}(\text{CR}_2)_n\text{Ar}-(\text{CR}_2)_p$ —^{<<<}, and ^{<<}— $(\text{CR}_2)_p\text{Ar}-(\text{CR}_2)_n\text{O}$ —^{<<<},
 wherein ^{<<} indicates the position of a covalent bond with the moiety:



[057] ^{<<<} indicates the position of a covalent bond with R⁸; each instance of n is independently a whole number selected from 1-20; p is an integer selected from 0-20; and R for each instance is independently selected from hydrogen, alkyl, cycloalkyl, and aryl; or two instances of R taken together with the carbons to which they are attached form a 3-6 membered carbocyclic ring; or two instances of R taken together with the atoms to which they are attached form a 5-6 membered heterocyclic ring.

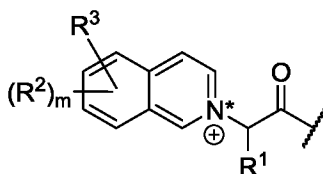
[058] In a third embodiment of the second aspect, provided herein is the method of the second aspect, wherein R² is hydrogen.

[059] In a fourth embodiment of the second aspect, provided herein is the method of the second aspect, wherein the linker is ^{<<}— $(\text{CR}_2)_n$ —^{<<<}, ^{<<}— $\text{O}(\text{CR}_2)_n$ —^{<<<}, or ^{<<}—

$(\text{OCR}_2\text{CR}_2)_n$; R^8 is , , ; and n is a whole number selected from 1-4.

[060] In a fifth embodiment of the second aspect, provided herein is the method of the second aspect, wherein the active agent is a chemotherapeutic agent, a radiotherapeutic agent, a fluorescent dye, or an affinity tag.

[061] In a third aspect, provided herein is a functionalized polypeptide comprising an N-terminal amino acid represented by the structure of Formula 2:



2

[062] m is 1, 2, or 3;

[063] N* is the N-terminal nitrogen of the functionalized polypeptide;

[064] R¹ is the side chain of the N-terminal amino acid of the functionalized polypeptide and the N-terminal amino acid is a natural amino acid or a non-natural amino acid;

[065] R² is independently for each instance selected from the group consisting of hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl, aralkyl, heterocycloalkyl, heteraryl, halide, cyano, nitro, hydroxyl, mercaptan, -OR⁵, -SR⁵, -(C=O)OR⁵, -O(C=O)R⁵, -N(R⁵)₂, -(C=O)N(R⁵)₂, -N(R⁵)(C=O)R⁵, -N(R⁵)(C=O)N(R⁵)₂, -SO₂R⁵, -N(R⁵)SO₂R⁵, and -SO₂N(R⁵)₂;

[066] R^3 is selected from the group consisting of $-R^4$ or $-A-R^4$, wherein A is a linker;

[067] R^4 is selected from the group consisting of hydroxyl, thiol, aldehyde, ketone, maleimide, *N*-hydroxysuccinimide ester, imidoester, haloacetyl, pyridyl disulfide, alkyne, azide, nitron, nitrile oxide, norbornene, oxanorbornadiene, tetrazine, cyclooctene, isocyanide, quadricyclane, and bis(dithiobenzil)nickel(II); and

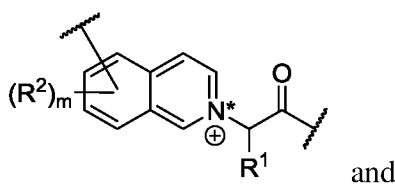
[068] R^5 for each instance is independently selected from the group consisting of hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl, aralkyl, heterocycloalkyl, and heteraryl; or two instances of R^5 taken together with the atoms to which they are attached form a 5-6 membered heterocycloalkyl.

[069] In a first embodiment of the third aspect, provided herein is the functionalized peptide of the third aspect, wherein the functionalized polypeptide is a protein, enzyme, antibody, glycoprotein, or lipoprotein.

[070] In a second embodiment of the third aspect, provided herein is the functionalized peptide of the third aspect, wherein the functionalized polypeptide is hemoglobin.

[071] In a third embodiment of the third aspect, provided herein is the functionalized peptide of the third aspect, wherein the linker is selected from the group consisting of $^{**}-(CR_2)_n-^{***}$, $^{**}-O(CR_2)_n-^{***}$, $^{**}-(CR_2)_nO-^{***}$, $^{**}-O(CR_2)_nO-^{***}$, $^{**}-(CR_2)_nC(=O)-^{***}$, $^{**}-C(=O)(CR_2)_n-^{***}$, $^{**}-C(=O)O(CR_2)_n-^{***}$, $^{**}-(CR_2)_nOC(=O)-^{***}$, $^{**}-O(CR_2)_nC(=O)-^{***}$, $^{**}-O(CR_2)_nC(=O)N(R)(CR_2)_p-^{***}$, $^{**}-O(CR_2)_nN(R)C(=O)N(R)(CR_2)_p-^{***}$, $^{**}-O(CR_2)_nC(=O)O(CR_2)_p-^{***}$, $^{**}-O(CR_2)_nOC(=O)N(R)(CR_2)_p-^{***}$, $^{**}-C(=O)(CR_2)_nO-^{***}$, $^{**}-O(CR_2)_nN(R)C(=O)-^{***}$, $^{**}-O(CR_2)_nN(R)C(=O)(CR_2)_p-^{***}$, $^{**}-C(=O)N(R)(CR_2)_nO-^{***}$, $^{**}-OC(=O)(CR_2)_n-^{***}$, $^{**}-$

$(\text{CR}_2)_n\text{C}(=\text{O})\text{O}-^{***}$, $^{**}\text{OC}(=\text{O})\text{O}(\text{CR}_2)_n-^{***}$, $^{**}-(\text{CR}_2)_n\text{OC}(=\text{O})\text{O}-^{***}$, $^{**}-$
 $\text{C}(=\text{O})\text{N}(\text{R})(\text{CR}_2)_n-^{***}$, $^{**}-(\text{CR}_2)_n(\text{R})\text{NC}(=\text{O})-^{***}$, $^{**}-\text{N}(\text{R})\text{C}(=\text{O})(\text{CR}_2)_n-^{***}$, $^{**}-$
 $(\text{CR}_2)_n\text{C}(=\text{O})\text{N}(\text{R})-^{***}$, $^{**}-\text{N}(\text{R})\text{C}(=\text{O})\text{O}(\text{CR}_2)_n-^{***}$, $^{**}-(\text{CR}_2)_n\text{OC}(=\text{O})\text{N}(\text{R})-^{***}$, $^{**}-$
 $\text{OC}(=\text{O})\text{N}(\text{R})(\text{CR}_2)_n-^{***}$, $^{**}-(\text{CR}_2)_n\text{N}(\text{R})\text{C}(=\text{O})\text{O}-^{***}$, $^{**}-(\text{OCR}_2\text{CR}_2)_n-^{***}$, $^{**}-$
 $(\text{CR}_2\text{CR}_2\text{O})_n-^{***}$, $^{**}-(\text{OCR}_2\text{CR}_2)_n(\text{C}=\text{O})-^{***}$, $^{**}-(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C}=\text{O})-^{***}$, $^{**}-$
 $(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})-^{***}$, $^{**}-(\text{OCR}_2\text{CR}_2)_n(\text{C}=\text{O})(\text{CR}_2)_p-^{***}$, $^{**}-(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C}=\text{O})$
 $(\text{CR}_2)_p-^{***}$, $^{**}-(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})(\text{CR}_2)_p-^{***}$, $^{**}-(\text{OCR}_2\text{CR}_2)_n(\text{C}=\text{O})\text{O}(\text{CR}_2)_p-^{***}$,
 $^{**}-(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C}=\text{O})\text{O}(\text{CR}_2)_p-^{***}$, $^{**}-(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})\text{O}(\text{CR}_2)_p-^{***}$, $^{**}-$
 $(\text{OCR}_2\text{CR}_2)_n(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p-^{***}$, $^{**}-(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p-^{***}$, $^{**}-$
 $(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p-^{***}$, $^{**}-\text{OCR}_2(\text{C}=\text{O})-^{***}$, $^{**}-\text{OCR}_2(\text{C}=\text{O})(\text{CR}_2)_p-$
 *** , $^{**}-\text{OCR}_2(\text{C}=\text{O})\text{O}(\text{CR}_2)_p-^{***}$, $^{**}-\text{OCR}_2(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p-^{***}$, $^{**}-$
 $\text{C}(=\text{O})\text{O}(\text{CR}_2\text{CR}_2\text{O})_n\text{CR}_2\text{CR}_2-^{***}$, $^{**}-\text{C}(=\text{O})\text{N}(\text{R})(\text{CR}_2\text{CR}_2\text{O})_n\text{CR}_2\text{CR}_2-^{***}$, $^{**}-$
 $\text{S}(\text{CR}_2)_n-^{***}$, $^{**}-(\text{CR}_2)_n\text{S}-^{***}$, $^{**}-\text{SO}_2(\text{CR}_2)_n-^{***}$, $^{**}-(\text{CR}_2)_n\text{SO}_2-^{***}$, $^{**}-\text{N}(\text{R})\text{SO}_2(\text{CR}_2)_n-$
 *** , $^{**}-(\text{CR}_2)_n\text{SO}_2\text{N}(\text{R})-^{***}$, $^{**}-\text{SO}_2\text{N}(\text{R})(\text{CR}_2)_n-^{***}$, $^{**}-(\text{CR}_2)_n\text{N}(\text{R})\text{SO}_2-^{***}$, $^{**}-(\text{CR}_2)_n-$
 $\text{Ar}-(\text{CR}_2)_p-^{***}$, $^{**}-(\text{CR}_2)_p-\text{Ar}-(\text{CR}_2)_n-^{***}$, $^{**}-\text{O}(\text{CR}_2)_n-\text{Ar}-(\text{CR}_2)_p-^{***}$, and $^{**}-(\text{CR}_2)_p-$
 $\text{Ar}-(\text{CR}_2)_n\text{O}-^{***}$, wherein ** indicates the position of a covalent bond with the moiety:



[072] *** indicates the position of a covalent bond with R^4 ; each instance of n is independently a whole number selected from 1-20; each instance of p is independently an integer selected from 0-20; and R for each instance is independently selected from hydrogen, alkyl, cycloalkyl, and aryl; or two instances of R taken together with the carbons to which they are attached form a 3-6 membered

carbocyclic ring; or two instances of R taken together with the atoms to which they are attached form a 5-6 membered heterocyclic ring.

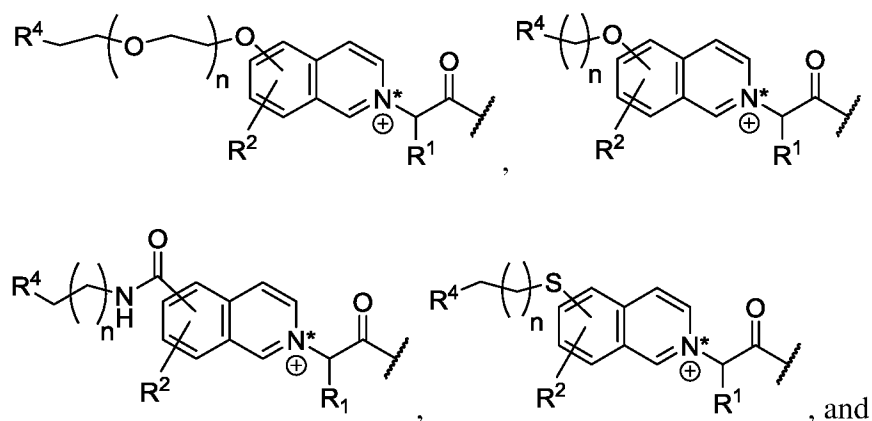
[073] In a fourth embodiment of the third aspect, provided herein is the functionalized peptide of the third aspect, wherein R^2 is hydrogen.

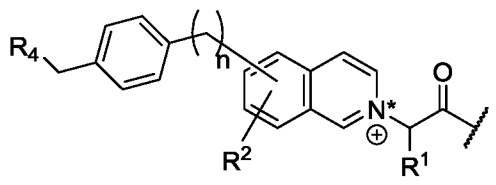
[074] In a fifth embodiment of the third aspect, provided herein is the functionalized peptide of the fourth embodiment of the third aspect, wherein the linker is $^{**}-(CR_2)_n-^{***}$, $^{**}-O(CR_2)_n-^{***}$, or $^{**}-(OCR_2CR_2)_n-^{***}$.

[075] In a sixth embodiment of the third aspect, provided herein is the functionalized peptide of the third aspect, wherein n is a whole number selected from 1-10 and p is a whole number selected from 1-4.

[076] In a seventh embodiment of the third aspect, provided herein is the functionalized peptide of the third aspect, wherein R^4 is maleimide, azide, alkyne, hydroxyl, norbornene, or cyclooctene.

[077] In an eighth embodiment of the third aspect, provided herein is the functionalized peptide of the third aspect, wherein the moiety of Formula 2 is selected from the group consisting of:





[078] wherein

[079] n is a whole number selected from 1-4;

[080] R^2 is hydrogen, alkyl, or $-OR^5$;

[081] R^4 is maleimide or alkyne; and

[082] R^5 is alkyl.

BRIEF DESCRIPTION OF THE DRAWINGS

[083] Figure 1 depicts the design approach for construction of an exemplary N-terminal selective hemoglobin-based therapeutic agent. One or more active agents (e.g., therapeutic drugs) can be linked to functionalized hemoglobin to form the hemoglobin active agent conjugate via an N-terminal selective bifunctional linker (cleavable or non-cleavable linkage).

[084] Figure 2 depicts the structure of the exemplary N-terminal selective bifunctional linkers comprising an N-terminal selective functional group and a functional group for active agent (e.g., therapeutic) attachment, such as maleimide, thiol, aldehyde, ketone, *N*-hydroxysuccinimide ester, imidoester, haloacetyl, pyridyl disulfide, alkyne, azide, nitron, nitrile oxide, norbornene, oxanorbornadiene, tetrazine, cyclooctene, isocyanide, quadricyclane and bis(dithiobenzil)nickel(II).

[085] Figure 3A depicts an exemplary synthetic sequences for two-step N-terminal modification of hemoglobin followed by reaction with a thiol containing active agent and one-step N-terminal modification with a reagent that includes the linked active agent.

[086] Figure 3B depicts exemplary functionalized polypeptides comprising (1) alkyl, (2) aryl, (3) polyethylene glycol (PEG), (4) ether, (5) thioether, (6) amide, (7) ester, (8) N-carbamate, (9) carbonate, (10) acylamino, (11) acyloxy, and (12) O-carbamate linkages.

[087] Figure 4 depicts the N-terminal modification of hemoglobin with 2-ethynyl-5-(prop-2-yn-1-yloxy)benzaldehyde thereby forming a functionalized hemoglobin for subsequent azide-alkyne [3+2] cycloaddition reaction with a fluorescent dye-, affinity-tag-, therapeutic drug- or active agent-linked azide.

[088] Figure 5 depicts an exemplary synthetic scheme for the synthesis of polyethylene (PEG) linked maleimide containing 2-ethynylbenzaldehyde according to certain embodiments of the present disclosure.

[089] Figure 6 depicts an exemplary synthetic scheme for an alkyl linked maleimide containing 2-ethynylbenzaldehyde according to certain embodiments of the present disclosure.

[090] Figure 7 depicts an exemplary synthetic sequence for the preparation of an exemplary active agent conjugate by Michael reaction of an active agent comprising a thiol or conjugate base thereof with a maleimide linked functionalized polypeptide according to certain embodiments of the present disclosure.

[091] Figure 8 depicts exemplary 2-ethynylbenzaldehydes analogs that can be used in connection with the methods described herein: (A) maleimide linked 2-ethynylbenzaldehyde (e.g., PEG and/or alkyl linked), (B) benzylmercaptan linked 2-ethynylbenzaldehyde and (C) *N*-Boc-L-cysteine methyl ester linked 2-ethynylbenzaldehyde, (D) alkyne linked 2-ethynylbenzaldehyde, and (E) fluorescein linked 2-ethynylbenzaldehyde.

[092] Figure 9A depicts the liquid chromatography–mass spectrometry (LC-MS) mass chromatogram of hemoglobin with an N-terminal amine depicted.

[093] Figure 9B depicts the liquid chromatography–mass spectrometry (LC-MS) mass chromatogram of hemoglobin N-terminal functionalized with 2-ethynylbenzaldehyde $\text{OCH}_2\text{CH}_2\text{CH}_2$ linked maleimide according to certain embodiments described herein.

[094] Figure 9C depicts the liquid chromatography–mass spectrometry (LC-MS) mass chromatogram of hemoglobin N-terminal functionalized with 2-ethynyl-5-(prop-2-yn-1-yloxy)benzaldehyde according to certain embodiments described herein.

[095] Figure 9D depicts the liquid chromatography–mass spectrometry (LC-MS) mass chromatogram of hemoglobin N-terminal functionalized with 2-ethynylbenzaldehyde $\text{OCH}_2\text{CH}_2\text{CH}_2\text{O}$ conjugated fluorescein according to certain embodiments described herein.

[096] Figure 9E depicts the liquid chromatography–mass spectrometry (LC-MS) mass chromatogram of the product of the [3+2] cycloaddition reaction of hemoglobin N-terminal functionalized with 2-ethynyl-5-(prop-2-yn-1-yloxy)benzaldehyde and a modified rhodamine comprising a linked azide according to certain embodiments described herein.

[097] Figure 9F depicts the liquid chromatography–mass spectrometry (LC-MS) mass chromatogram of hemoglobin N-terminal functionalized with 2-ethynylbenzaldehyde $(\text{OCH}_2\text{CH}_2)_2$ linked maleimide conjugated with benzylmercaptan according to certain embodiments described herein.

[098] Figure 9G depicts an expanded view of the liquid chromatography–mass

spectrometry (LC-MS) mass chromatogram of hemoglobin N-terminal functionalized with 2-ethynylbenzaldehyde $(\text{OCH}_2\text{CH}_2)_2$ linked maleimide conjugated with benzylmercaptan according to certain embodiments described herein.

[0099] Figure 10A depicts the liquid chromatography–mass spectrometry (LC-MS) mass chromatogram of lysozyme with the N-terminal amine depicted.

[0100] Figure 10B depicts the liquid chromatography–mass spectrometry (LC-MS) mass chromatogram of lysozyme N-terminal functionalized with 2-ethynylbenzaldehyde $(\text{OCH}_2\text{CH}_2)_2$ linked maleimide conjugated with benzylmercaptan according to certain embodiments described herein.

[0101] Figure 10C depicts the liquid chromatography–mass spectrometry (LC-MS) mass chromatogram of lysozyme N-terminal functionalized with 2-ethynyl-5-(prop-2-yn-1-yloxy)benzaldehyde according to certain embodiments described herein.

[0102] Figure 10D depicts the liquid chromatography–mass spectrometry (LC-MS) mass chromatogram of lysozyme N-terminal functionalized with 2-ethynylbenzaldehyde $\text{OCH}_2\text{CH}_2\text{CH}_2\text{O}$ conjugated fluorescein according to certain embodiments described herein.

[0103] Figure 11A depicts the liquid chromatography–mass spectrometry (LC-MS) mass chromatogram of ribonuclease A (RNase A) with the N-terminal amine depicted.

[0104] Figure 11B depicts the liquid chromatography–mass spectrometry (LC-MS) mass chromatogram of RNase A N-terminal functionalized with 2-ethynylbenzaldehyde $(\text{OCH}_2\text{CH}_2)_2$ linked maleimide conjugated with benzylmercaptan according to certain embodiments described herein.

[0105] Figure 11C depicts the liquid chromatography–mass spectrometry (LC-MS) mass chromatogram of RNase A N-terminal functionalized with 2-ethynyl-5-(prop-2-yn-1-yloxy)benzaldehyde according to certain embodiments described herein.

[0106] Figure 11D depicts the liquid chromatography–mass spectrometry (LC-MS) mass chromatogram of RNase A N-terminal functionalized with 2-ethynylbenzaldehyde $\text{OCH}_2\text{CH}_2\text{CH}_2\text{O}$ conjugated fluorescein according to certain embodiments described herein.

[0107] Figure 12A depicts the liquid chromatography–mass spectrometry (LC-MS) mass chromatogram of tryptic lysozyme N-terminal functionalized with 2-ethynylbenzaldehyde $(\text{OCH}_2\text{CH}_2)_2$ linked maleimide conjugated with benzylmercaptan according to certain embodiments described herein.

[0108] Figure 12B depicts the liquid chromatography–mass spectrometry (LC-MS) mass chromatogram of tryptic lysozyme N-terminal functionalized with 2-ethynyl-5-(prop-2-yn-1-yloxy)benzaldehyde according to certain embodiments described herein.

[0109] Figure 12C depicts liquid chromatography–mass spectrometry-mass spectrometry LC-MS-MS results for tryptic lysozyme N-terminal functionalized with 2-ethynylbenzaldehyde $\text{OCH}_2\text{CH}_2\text{CH}_2\text{O}$ conjugated fluorescein.

[0110] Figure 12D depicts the liquid chromatography–mass spectrometry (LC-MS) mass chromatogram of RNase A N-terminal functionalized with 2-ethynylbenzaldehyde $(\text{OCH}_2\text{CH}_2)_2$ linked maleimide conjugated with benzylmercaptan according to certain embodiments described herein.

[0111] Figure 12E depicts the liquid chromatography–mass spectrometry (LC-MS)

mass chromatogram of tryptic RNase A N-terminal functionalized with 2-ethynyl-5-(prop-2-yn-1-yloxy)benzaldehyde according to certain embodiments described herein.

[0112] Figure 12F depicts the liquid chromatography–mass spectrometry (LC-MS) mass chromatogram of tryptic RNase A N-terminal functionalized with 2-ethynylbenzaldehyde $\text{OCH}_2\text{CH}_2\text{CH}_2\text{O}$ conjugated fluorescein according to certain embodiments described herein.

[0113] Figure 13A depicts the liquid chromatography–mass spectrometry (LC-MS) mass chromatogram of YTSSSKNVVR (SEQ ID NO: 1).

[0114] Figure 13B depicts the liquid chromatography–mass spectrometry (LC-MS) mass chromatogram of YTSSSKNVVR (SEQ ID NO:1) N-terminal functionalized with 2-ethynylbenzaldehyde $(\text{OCH}_2\text{CH}_2)_2$ linked maleimide according to certain embodiments described herein.

[0115] Figure 13C depicts the liquid chromatography–mass spectrometry (LC-MS) mass chromatogram of YTSSSKNVVR (SEQ ID NO:1) N-terminal functionalized with 2-ethynylbenzaldehyde $(\text{OCH}_2\text{CH}_2)_2$ linked maleimide conjugated with *N*-Boc-L-cysteine methyl ester according to certain embodiments described herein.

[0116] Figure 13D depicts the liquid chromatography–mass spectrometry (LC-MS) mass chromatogram of YTSSSKNVVR (SEQ ID NO:1) N-terminal functionalized with 2-ethynylbenzaldehyde $(\text{OCH}_2\text{CH}_2)_2$ linked maleimide conjugated with benzylmercaptan according to certain embodiments described herein.

[0117] Figure 14A depicts the liquid chromatography (LC) results for the selectivity of the *N*-terminal functionalization of YTSSSKNVVR (SEQ ID NO:1) with 2-ethynylbenzaldehyde $(\text{OCH}_2\text{CH}_2)_2$ linked maleimide according to certain

embodiments described herein.

[0118] Figure 14B depicts the liquid chromatography (LC) results for the selectivity of the *N*-terminal functionalization of YTSSSKNVVR (SEQ ID NO:1) with 2-ethynylbenzaldehyde (OCH₂CH₂)₂ linked maleimide conjugated with benzylmercaptan according to certain embodiments described herein.

[0119] Figure 15A depicts the liquid chromatography–mass spectrometry (LC-MS) mass chromatogram of YTSSSKNVVR (SEQ ID NO:1) *N*-terminal functionalized with 2-ethynylbenzaldehyde OCH₂CH₂CH₂O linked maleimide (doubly charged precursor ion = *m/z* 712.35) according to certain embodiments described herein.

[0120] Figure 15B depicts the liquid chromatography–mass spectrometry (LC-MS) mass chromatogram of YTSSSKNVVR (SEQ ID NO:1) *N*-terminal functionalized with 2-ethynylbenzaldehyde OCH₂CH₂CH₂O linked maleimide (triply charged precursor ion = *m/z* 475.23) according to certain embodiments described herein.

[0121] Figure 15C depicts the liquid chromatography–mass spectrometry (LC-MS) mass chromatogram of YTSSSKNVVR (SEQ ID NO:1) *N*-terminal functionalized with 2-ethynylbenzaldehyde (OCH₂CH₂)₂ linked maleimide conjugated with benzylmercaptan (doubly charged precursor ion = *m/z* 780.23) according to certain embodiments described herein.

[0122] Figure 16 depicts the LC-MS results of stability tests of hemoglobin *N*-terminal functionalized with 2,5-diethynylbenzaldehyde (**2n**) stored for 1-8 weeks at 4 °C.

DETAILED DESCRIPTION

Definitions

[0123] The definitions of terms used herein are meant to incorporate the present

state-of-the-art definitions recognized for each term in the field of biotechnology. Where appropriate, exemplification is provided. The definitions apply to the terms as they are used throughout this specification, unless otherwise limited in specific instances, either individually or as part of a larger group.

[0124] When trade names are used herein, applicants intend to independently include the trade name product formulation, the generic drug, and the active pharmaceutical ingredient(s) of the trade name product.

[0125] The term “antibody” as used herein is used in the broadest sense and specifically covers intact monoclonal antibodies, polyclonal antibodies, multispecific antibodies (e.g., bispecific antibodies) formed from at least two intact antibodies, and antibody fragments, so long as they exhibit the desired biological activity. In certain embodiments, antibodies are polyclonal, monoclonal, bispecific, human, humanized or chimeric antibodies, single chain antibodies, Fv, Fab fragments, F(ab') fragments, F(ab')₂ fragments, anti-idiotypic (anti-Id) antibodies, CDR's, and epitope-binding fragments of any of the above which immunospecifically bind to an antigen, such as cancer cell antigens, viral antigens, or microbial antigens.

[0126] The term "amino acid" refers to naturally occurring and non-natural amino acids, as well as amino acid analogs and amino acid mimetics that function in a manner similar to the naturally occurring amino acids. Naturally encoded amino acids are the 20 common amino acids (alanine, arginine, asparagine, aspartic acid, cysteine, glutamine, glutamic acid, glycine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, and valine) and pyrrolysine and selenocysteine. Amino acid analogs refers to compounds that have the same basic chemical structure as a naturally occurring amino acid, i.e., an α carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group, such as,

homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium. Such analogs have modified R groups (such as, norleucine) or modified peptide backbones, but retain the same basic chemical structure as a naturally occurring amino acid.

[0127] A "non-natural amino acid" refers to an amino acid that is not one of the 20 common amino acids or pyrrolysine or selenocysteine; other terms that may be used synonymously with the term "non-natural amino acid" is "non-naturally encoded amino acid," "unnatural amino acid," "non-naturally-occurring amino acid," and the like. The term "non-natural amino acid" includes, but is not limited to, amino acids that have the same basic chemical structure as a naturally occurring amino acid, i.e., an α carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group, as well as amino acids that amino acids in which the amino group is attached at the β or γ carbon.

[0128] The term "hemoglobin" refers generally to the protein contained within red blood cells that transports oxygen. Each molecule of hemoglobin has 4 subunits, 2 α chains and 2 β chains, which are arranged in a tetrameric structure. Each subunit also contains one heme group, which is the iron-containing center that binds oxygen. Thus, each hemoglobin molecule can bind 4 oxygen molecules.

[0129] The term "modified hemoglobin" includes, but is not limited to, hemoglobin altered by a chemical reaction such as intra- and inter-molecular cross-linking, genetic manipulation and polymerization. As used herein, the term "hemoglobin" by itself refers both to native, unmodified, hemoglobin, as well as modified hemoglobin.

[0130] The term "protein", "polypeptide", or "peptide" as used herein indicates an organic polymer composed of two or more amino acid monomers and/or analogs thereof. The term "polypeptide" or "peptide" includes amino acid polymers of any length including full length proteins and peptides, as well as analogs and fragments

thereof. A polypeptide of three or more amino acids is also called an oligopeptide. As used herein, the term "amino acid", "amino acidic monomer", or "amino acid residue" refers to any of the twenty naturally occurring amino acids.

[0131] The term "heteroatom" is art-recognized and refers to an atom of any element other than carbon or hydrogen. Illustrative heteroatoms include boron, nitrogen, oxygen, phosphorus, sulfur and selenium.

[0132] The term "alkyl" is art-recognized, and includes saturated aliphatic groups, including straight-chain alkyl groups, branched-chain alkyl groups, cycloalkyl (alicyclic) groups, alkyl substituted cycloalkyl groups, and cycloalkyl substituted alkyl groups. In certain embodiments, a straight chain or branched chain alkyl has about 30 or fewer carbon atoms in its backbone (e.g., C₁-C₃₀ for straight chain, C₃-C₃₀ for branched chain), and alternatively, about 20 or fewer. Likewise, cycloalkyls have from about 3 to about 10 carbon atoms in their ring structure, and alternatively about 5, 6 or 7 carbons in the ring structure.

[0133] Unless the number of carbons is otherwise specified, "lower alkyl" refers to an alkyl group, as defined above, but having from one to about ten carbons, alternatively from one to about six carbon atoms in its backbone structure. Likewise, "lower alkenyl" and "lower alkynyl" have similar chain lengths.

[0134] The term "aralkyl" is art-recognized and refers to an alkyl group substituted with an aryl group (e.g., an aromatic or heteroaromatic group).

[0135] The terms "alkenyl" and "alkynyl" are art-recognized and refer to unsaturated aliphatic groups analogous in length and possible substitution to the alkyls described above, but that contain at least one double or triple bond respectively.

[0136] The term "aryl" is art-recognized and refers to 5-, 6- and 7-membered single-ring aromatic groups that may include from zero to four heteroatoms, for example, benzene, naphthalene, anthracene, pyrene, pyrrole, furan, thiophene, imidazole, oxazole, thiazole, triazole, pyrazole, pyridine, pyrazine, pyridazine and pyrimidine, and the like. Those aryl groups having heteroatoms in the ring structure may also be referred to as "aryl heterocycles" or "heteroaromatics." The aromatic ring may be substituted at one or more ring positions with such substituents as described above, for example, halogen, azide, alkyl, aralkyl, alkenyl, alkynyl, cycloalkyl, hydroxyl, alkoxyl, amino, nitro, sulfhydryl, imino, amido, phosphonate, phosphinate, carbonyl, carboxyl, silyl, ether, alkylthio, sulfonyl, sulfonamido, ketone, aldehyde, ester, heterocyclyl, aromatic or heteroaromatic moieties, -CF₃, -CN, or the like. The term "aryl" also includes polycyclic ring systems having two or more cyclic rings in which two or more carbons are common to two adjoining rings (the rings are "fused rings") wherein at least one of the rings is aromatic, e.g., the other cyclic rings may be cycloalkyls, cycloalkenyls, cycloalkynyls, aryls and/or heterocyclyls.

[0137] The terms ortho, meta and para are art-recognized and refer to 1,2-, 1,3- and 1,4-disubstituted benzenes, respectively. For example, the names 1,2-dimethylbenzene and ortho-dimethylbenzene are synonymous.

[0138] The terms "heterocyclyl", "heteroaryl", or "heterocyclic group" are art-recognized and refer to 3- to about 10-membered ring structures, alternatively 3- to about 7-membered rings, whose ring structures include one to four heteroatoms. Heterocycles may also be polycycles. Heterocyclyl groups include, for example, thiophene, thianthrene, furan, pyran, isobenzofuran, chromene, xanthene, phenoxanthene, pyrrole, imidazole, pyrazole, isothiazole, isoxazole, pyridine, pyrazine, pyrimidine, pyridazine, indolizine, isoindole, indole, indazole, purine,

quinolizine, isoquinoline, quinoline, phthalazine, naphthyridine, quinoxaline, quinazoline, cinnoline, pteridine, carbazole, carboline, phenanthridine, acridine, pyrimidine, phenanthroline, phenazine, phenarsazine, phenothiazine, furazan, phenoxazine, pyrrolidine, oxolane, thiolane, oxazole, piperidine, piperazine, morpholine, lactones, lactams such as azetidinones and pyrrolidinones, sultams, sultones, and the like. The heterocyclic ring may be substituted at one or more positions with such substituents as described above, as for example, halogen, alkyl, aralkyl, alkenyl, alkynyl, cycloalkyl, hydroxyl, amino, nitro, sulfhydryl, imino, amido, phosphonate, phosphinate, carbonyl, carboxyl, silyl, ether, alkylthio, sulfonyl, ketone, aldehyde, ester, a heterocyclyl, an aromatic or heteroaromatic moiety, -CF₃, -CN, or the like.

[0139] The term "optionally substituted" refers to a chemical group, such as alkyl, cycloalkyl aryl, and the like, wherein one or more hydrogen may be replaced with a with a substituent as described herein, for example, halogen, azide, alkyl, aralkyl, alkenyl, alkynyl, cycloalkyl, hydroxyl, alkoxy, amino, nitro, sulfhydryl, imino, amido, phosphonate, phosphinate, carbonyl, carboxyl, silyl, ether, alkylthio, sulfonyl, sulfonamido, ketone, aldehyde, ester, heterocyclyl, aromatic or heteroaromatic moieties, -CF₃, -CN, or the like

[0140] The terms "polycyclyl" or "polycyclic group" are art-recognized and refer to two or more rings (e.g., cycloalkyls, cycloalkenyls, cycloalkynyls, aryls and/or heterocyclyls) in which two or more carbons are common to two adjoining rings, e.g., the rings are "fused rings". Rings that are joined through non-adjacent atoms are termed "bridged" rings. Each of the rings of the polycycle may be substituted with such substituents as described above, as for example, halogen, alkyl, aralkyl, alkenyl, alkynyl, cycloalkyl, hydroxyl, amino, nitro, sulfhydryl, imino, amido, phosphonate,

phosphinate, carbonyl, carboxyl, silyl, ether, alkylthio, sulfonyl, ketone, aldehyde, ester, a heterocyclyl, an aromatic or heteroaromatic moiety, -CF₃, -CN, or the like.

[0141] The term "carbocycle" is art-recognized and refers to an aromatic or non-aromatic ring in which each atom of the ring is carbon.

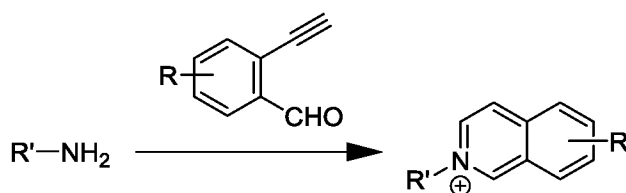
[0142] The term "nitro" is art-recognized and refers to -NO₂; the term "halogen" is art-recognized and refers to -F, -Cl, -Br or -I; the term "sulfhydryl" is art-recognized and refers to -SH; the term "hydroxyl" means -OH; and the term "sulfonyl" and "sulfone" is art-recognized and refers to -SO₂-. "Halide" designates the corresponding anion of the halogens.

[0143] A "chemotherapeutic agent" is a chemical compound useful in the treatment of cancer, regardless of mechanism of action. Classes of chemotherapeutic agents include, but are not limited to: alkylating agents, antimetabolites, spindle poison plant alkaloids, cytotoxic/antitumor antibiotics, topoisomerase inhibitors, antibodies, photosensitizers, and kinase inhibitors. Chemotherapeutic agents include compounds used in "targeted therapy" and conventional chemotherapy.

[0144] The term "fluorescent dye" refers interchangeably to molecules, groups or radicals which are fluorescent. The term "fluorescent" as applied to a molecule or compound is used to refer to the property of the compound of absorbing energy (such as UV, visible or IR radiation) and re-emitting at least a fraction of that energy as light over time. Fluorescent dyes include, but are not limited to small molecules, proteins and macromolecular complexes.

[0145] The methods described herein are capable of N-terminal selective functionalization, under mild conditions (e.g., pH 6.3 at 37 °C), of the N-terminal amine by reaction of a polypeptide with 2-ethynylbenzaldehyde or an analog thereof

thereby forming an N-terminal functionalized isoquinolinium with selectivity up to 99% or greater for the N-terminal amine as shown in the general reaction sequence below.



[0146] As demonstrated in the examples below, the methods described herein can afford the N-terminal functionalized polypeptide with yields of 86% or greater and moderate to excellent N-terminal selectivity (12:1 to >99:1). The resulting isoquinolinium is relatively stable and can be stored for extended periods of time without substantial degradation (See Figure 17, which shows LCMS data for N-terminal functionalized hemoglobin stored at 4 °C for 1-8 weeks storage). The methods described herein can thus be used as a means of selectively functionalizing a polypeptide and controlling the extent of functionalization (e.g., at least partially limiting the functionalization reaction to e.g., one per polypeptide).

[0147] By appropriate design and modification of the 2-ethynylbenzaldehyde, the methods described herein can be used as a means for selectively conjugating an active agent, such as chemotherapeutic agent, a radiotherapeutic agent, a fluorescent dye, or affinity tag, by covalent attachment either directly or indirectly (e.g., via a linker or linker and spacer) to the 2-ethynylbenzaldehyde, to the N-terminal of a polypeptide in a selective manner. As described in greater detail herein, N-terminal selective conjugation of an active agent to a polypeptide using the methods described herein can be accomplished using a one-step or two-step method.

[0148] The polypeptide can be a protein, enzyme, antibody, glycoprotein, or

lipoprotein.

[0149] In certain embodiments, the polypeptide is hemoglobin. In certain embodiments, the hemoglobin is derived from substantially any mammalian source. Exemplary sources of hemoglobin include common livestock animals, e.g., cows, pigs, sheep and the like and humans. In certain embodiments, the hemoglobin is recombinant hemoglobin.

[0150] In certain embodiments, the polypeptide is a hemoglobin tetramer, dimer, or subunit thereof.

[0151] In certain embodiments, the antibody is an intact antibody, single chain antibody, Fv, Fab fragment, F(ab') fragment, or F(ab')₂ fragment.

[0152] Chemotherapeutic agents suitable for use in connection with the methods described herein include any chemotherapeutic agent known to those skilled in the art including, but not limited to erlotinib (TARCEVAR®, Genentech/OSI Pharm.), docetaxel (TAXOTERE®, Sanofi-Aventis), 5-FU (fluorouracil, 5-fluorouracil, CAS No. 51-21-8), gemcitabine (GEMZAR®, Lilly), PD-0325901 (CAS No. 391210-10-9, Pfizer), cisplatin (cis-diamine, dichloroplatinum(II), CAS No. 15663-27-1), carboplatin (CAS No. 41575-94-4), paclitaxel (TAXOL®, Bristol-Myers Squibb Oncology, Princeton, N.J.), trastuzumab (HERCEPTIN®, Genentech), temozolomide (4-methyl-5-oxo-2,3,4,6,8-pentazabicyclo[4.3.0]nona-2,7,9-triene-9-carboxamide, CAS No. 85622-93-1, TEMODAR®, TEMODAL®, Schering Plough), tamoxifen ((Z)-2-[4-(1,2-diphenylbut-1-enyl)phenoxy]-N,N-dimethyl-ethanamine, NOLVADEX®, ISTUBAL®, VALODEX®), and doxorubicin (ADRIAMYCIN®), Akti-1/2, HPPD, and rapamycin.

[0153] Additional examples of suitable chemotherapeutic agents include:

oxaliplatin (ELOXATIN®, Sanofi), bortezomib (VELCADE®, Millennium Pharm.), sunitinib (SUNITINIB®, SU1 1248, Pfizer), letrozole (FEMARA®, Novartis), imatinib mesylate (GLEEVEC®, Novartis), XL-518 (Mek inhibitor, Exelixis, WO 2007/044515), ARRY-886 (Mek inhibitor, AZD6244, Array BioPharma, AstraZeneca), SF-1 126 (PI3K inhibitor, Semafore Pharmaceuticals), BEZ-235 (PI3K inhibitor, Novartis), XL-147 (PI3K inhibitor, Exelixis), PTK787/ZK 222584 (Novartis), fulvestrant (FASLODEX®, AstraZeneca), leucovorin (folinic acid), rapamycin (sirolimus, RAPAMUNE®, Wyeth), lapatinib (TYKERB®, GSK572016, Glaxo Smith Kline), lonafarnib (SARASAR™, SCH 66336, Schering Plough), sorafenib (NEXAVAR®, BAY43-9006, Bayer Labs), gefitinib (IRESSA®, AstraZeneca), irinotecan (CAMPTOSAR®, CPT-11, Pfizer), tipifarnib (ZARNESTRA™, Johnson & Johnson), ABRAXANE™ (Cremophor-free), albumin-engineered nanoparticle formulations of paclitaxel (American Pharmaceutical Partners, Schaumburg, IL), vandetanib (rINN, ZD6474, ZACTIMA®, AstraZeneca), chlorambucil, AG1478, AG1571 (SU 5271; Sugen), temsirolimus (TORISEL®, Wyeth), pazopanib (GlaxoSmithKline), canfosamide (TELCYTA®, Telik), thiotepa and cyclophosphamide (CYTOXAN®, NEOSAR®); alkyl sulfonates such as busulfan, improsulfan and piposulfan; aziridines such as benzodopa, carboquone, meturedopa, and uredopa; ethylenimines and methylenamines including altretamine, triethylenemelamine, triethylenephosphoramide, triethylenethiophosphoramide and trimethylmelamine; acetogenins (especially bullatacin and bullatacinone); a camptothecin (including the synthetic analog topotecan); bryostatin; callystatin; CC-1065 (including its adozelesin, carzelesin and bizelesin synthetic analogs); cryptophycins (particularly cryptophycin 1 and cryptophycin 8); dolastatin; duocarmycin (including the synthetic analogs, KW-2189 and CB1-TM1); eleutherobin; pancratistatin; a sarcodictyin;

spongistatin; nitrogen mustards such as chlorambucil, chlormaphazine, chlorophosphamide, estramustine, ifosfamide, mechlorethamine, mechlorethamine oxide hydrochloride, melphalan, novembichin, phenesterine, prednimustine, trofosfamide, uracil mustard; nitrosoureas such as carmustine, chlorozotocin, fotemustine, lomustine, nimustine, and ranimustine; antibiotics such as the enediyne antibiotics (e.g., calicheamicin, calicheamicin gammaII, calicheamicin omegallI (Angew Chem. Intl. Ed. Engl. (1994) 33:183-186); dynemicin, dynemicin A; bisphosphonates, such as clodronate; an esperamicin; as well as neocarzinostatin chromophore and related chromoprotein enediyne antibiotic chromophores), aclacinomysins, actinomycin, authramycin, azaserine, bleomycins, cactinomycin, carabycin, caminomycin, carzinophilin, chromomycinis, dactinomycin, daunorubicin, detorubicin, 6-diazo-5-oxo-L-norleucine, morpholino-doxorubicin, cyanomorpholino-doxorubicin, 2-pyrrolino-doxorubicin and deoxydoxorubicin), epirubicin, esorubicin, idarubicin, marcellomycin, mitomycins such as mitomycin C, mycophenolic acid, nogalamycin, olivomycins, peplomycin, porfiromycin, puromycin, quelamycin, rodorubicin, streptonigrin, streptozocin, tubercidin, ubenimex, zinostatin, zorubicin; anti-metabolites such as methotrexate and 5-fluorouracil (5-FU); folic acid analogs such as denopterin, methotrexate, pteropterin, trimetrexate; purine analogs such as fludarabine, 6-mercaptopurine, thiamiprine, thioguanine; pyrimidine analogs such as ancitabine, azacitidine, 6-azauridine, carmofur, cytarabine, dideoxyuridine, doxifluridine, enocitabine, floxuridine; androgens such as calusterone, dromostanolone propionate, epitio stanol, mepitio stanol, testolactone; anti-adrenals such as aminoglutethimide, mitotane, trilostane; folic acid replenisher such as frolinic acid; aceglatone; aldophosphamide glycoside; aminolevulinic acid; eniluracil; amsacrine; bestabucil; bisantrene; edatraxate; defofamine; demecolcine; diaziquone; elfomithine; elliptinium acetate; an

epothilone; etoglucid; gallium nitrate; hydroxyurea; lentinan; lonidainine; mitoguazone; mitoxantrone; mopidanmol; nitraerine; pentostatin; phenamet; pirarubicin; losoxantrone; podophyllinic acid; 2-ethylhydrazide; procarbazine; PSK® polysaccharide complex (JHS Natural Products, Eugene, Oreg.); razoxane; rhizoxin; sizofiran; spirogermanium; tenuazonic acid; triaziquone; 2,2',2''-trichlorotriethylamine; trichothecenes (especially T-2 toxin, verracurin A, roridin A and anguidine); urethan; vindesine; dacarbazine; mannomustine; mitobronitol; mitolactol; pipobroman; gacytosine; arabinoside ("Ara-C"); cyclophosphamide; thiotepa; 6-thioguanine; mercaptopurine; methotrexate; platinum analogs such as cisplatin and carboplatin; vinblastine; etoposide (VP-16); ifosfamide; mitoxantrone; vincristine; vinorelbine (NAVELBINE®); novantrone; teniposide; edatrexate; daunomycin; aminopterin; capecitabine (XELODA®, Roche); ibandronate; CPT-11; topoisomerase inhibitor RFS 2000; difluoromethylornithine (DMFO); retinoids such as retinoic acid; and pharmaceutically acceptable salts, acids and derivatives of any of the above.

[0154] Any fluorescent dye known to those skilled in the art can be used in connection with the methods described herein. Fluorescent dyes are well known in the art, and include, but are not limited to xanthene dyes, e.g., fluorescein and rhodamine dyes, such as fluorescein isothiocyanate (FITC), 6-carboxyfluorescein, 6-carboxy-2',4',7',4,7'-hexachlorofluorescein (HEX), 6-carboxy-4', 5'-dichloro-2', 7'-dimethoxyfluorescein, N,N,N',N'-tetramethyl-6-carboxyrhodamine, 6-carboxy-rhodamine, 5-carboxyrhodamine-6G, 6-carboxyrhodamine-6G, and rhodamine 110; cyanine dyes, e.g. Cy3, Cy5 and Cy7 dyes; coumarins, e.g. umbelliferone; benzimide dyes, e.g. Hoechst 33258; phenanthridine dyes, e.g. Texas Red; ethidium dyes; acridine dyes; carbazole dyes; phenoxazine dyes; porphyrin dyes;

polymethine dyes, e.g. cyanine dyes such as Cy3, Cy5, etc; BODIPY dyes and quinoline dyes.

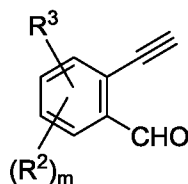
[0155] Affinity tags suitable for use in connection with the methods described herein include any affinity tag known to those skilled in the art including, but not limited to Biotin-tag, epitope-tag, FLAG-tag, poly(His) tag, streptag II, HA-tag, softag1, softag 3, c-myc, T7-tag, S-tag, elastin-like peptides, chitin-binding domain, thioredoxin, xylanase 10A, glutathione S-transferase, maltose binding protein, NusA, and the like.

[0156] Active agents suitable for use in connection with the methods described herein can also include crosslinking agents, which can be used to covalently bond with other chemical or biological moieties (N-hydroxysuccinimide (NHS) and sulfo-NHS, maleimide, iodoacetyl groups or pyridyl disulfides, primary amines, carbonyls, hydrazines and alkoxyamines.

[0157] In certain instances, chemical modifications may be required to introduce the necessary functionality in the chemotherapeutic agent, a radiotherapeutic agent, a fluorescent dye, or affinity tag in order to provide a functional group capable of conjugating the chemotherapeutic agent, a radiotherapeutic agent, a fluorescent dye, or affinity tag to the 2-ethynylbenzaldehyde analogs described herein (e.g., the compound of Formula **1** and the compound of Formula **3**). The choice of such functional group, the site of incorporation and the methods for incorporating the functional group in the molecule is well within the skill of a person of ordinary skill in the art.

[0158] In instances in which the methods described herein are used in a two-step method for selectively conjugating an active agent to the N-terminal amine of a

polypeptide, reactive functionality can be incorporated into the 2-ethynylbenzaldehyde analog. Such reactive functionality can have orthogonal reactivity with and/or a slower reaction rate than the 2-ethynylbenzaldehyde analog with the N-terminal amine. In such instances, the 2-ethynylbenzaldehyde analog can be represented by the compound of Formula 1:



1

[0159] or a conjugate salt thereof, wherein:

[0160] m is 1, 2, or 3; and

[0161] each R^2 is independently for each instance selected from the group consisting of hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl, aralkyl, heterocycloalkyl, heteraryl, halide, cyano, nitro, hydroxyl, mercaptan, $-OR^5$, $-SR^5$, $-(C=O)OR^5$, $-O(C=O)R^5$, $-N(R^5)_2$, $-(C=O)N(R^5)_2$, $-N(R^5)(C=O)R^5$, $-N(R^5)(C=O)N(R^5)_2$, $-SO_2R^5$, $-N(R^5)SO_2R^5$, and $-SO_2N(R^5)_2$;

[0162] R^3 is selected from the group consisting of $-R^4$ or $-A-R^4$, wherein A is a linker;

[0163] R^4 is selected from the group consisting of hydroxyl, thiol, aldehyde, ketone, maleimide, *N*-hydroxysuccinimide ester, imidoester, haloacetyl, pyridyl disulfide, alkyne, azide, nitron, nitrile oxide, norbornene, oxanorbornadiene, tetrazine, cyclooctene, isocyanide, quadricyclane, and bis(dithiobenzil)nickel(II); and

[0164] R^5 for each instance is independently selected from the group consisting of hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl, aralkyl, heterocycloalkyl, and heteraryl; or two instances of R^5 taken together with the atoms to which they are attached form a 5-6 membered heterocycloalkyl.

[0165] The moiety R^3 comprises reactive functional groups that can be further reacted with an active agent comprising at least one complimentary reactive functional group.

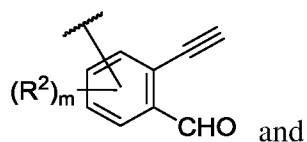
[0166] In certain embodiments of the compound of Formula **1**, m is 1.

[0167] In certain embodiments of the compound of Formula **1**, R^2 independently selected from is halide, cyano, nitro, hydrogen, or alkyl. In certain embodiments of the compound of Formula **1**, m is 1 and R^2 is halide, cyano, nitro, hydrogen, or alkyl. In certain embodiments of the compound of Formula **1**, R^2 is hydrogen.

[0168] R^4 can be covalently attached directly to the 2-ethynylbenzaldehyde analog or can be covalently attached via a linker. The linker can be any linker known to those of skill in the art. The selection of which is well within the skill of a person of skill in the art. In the broadest sense, a linker is a divalent chemical moiety comprising a covalent bond or a chain of atoms that covalently attaches two groups (e.g., R^4 and the 2-ethynylbenzaldehyde analog).

[0169] In certain embodiments of the compound of Formula **1**, the linker is at least one moiety selected from the group consisting of $^{**}-(CR_2)_n-^{***}$, $^{**}-O(CR_2)_n-^{***}$, $^{**}-(CR_2)_nO-^{***}$, $^{**}-O(CR_2)_nO-^{***}$, $^{**}-(CR_2)_nC(=O)-^{***}$, $^{**}-C(=O)(CR_2)_n-^{***}$, $^{**}-C(=O)O(CR_2)_n-^{***}$, $^{**}-(CR_2)_nOC(=O)-^{***}$, $^{**}-O(CR_2)_nC(=O)-^{***}$, $^{**}-O(CR_2)_nC(=O)N(R)(CR_2)_p-^{***}$, $^{**}-O(CR_2)_nN(R)C(=O)N(R)(CR_2)_p-^{***}$, $^{**}-O(CR_2)_nC(=O)O(CR_2)_p-^{***}$, $^{**}-O(CR_2)_nOC(=O)N(R)(CR_2)_p-^{***}$, $^{**}-C(=O)(CR_2)_nO-$

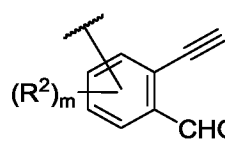
*** , $^{**}-O(CR_2)_nN(R)C(=O)-^{***}$, $^{**}-O(CR_2)_nN(R)C(=O)(CR_2)_p-^{***}$, $^{**}-$
 $C(=O)N(R)(CR_2)_nO-^{***}$, $^{**}-OC(=O)(CR_2)_n-^{***}$, $^{**}-(CR_2)_nC(=O)O-^{***}$, $^{**}-$
 $OC(=O)O(CR_2)_n-^{***}$, $^{**}-(CR_2)_nOC(=O)O-^{***}$, $^{**}-C(=O)N(R)(CR_2)_n-^{***}$, $^{**}-$
 $(CR_2)_n(R)NC(=O)-^{***}$, $^{**}-N(R)C(=O)(CR_2)_n-^{***}$, $^{**}-(CR_2)_nC(=O)N(R)-^{***}$, $^{**}-$
 $N(R)C(=O)O(CR_2)_n-^{***}$, $^{**}-(CR_2)_nOC(=O)N(R)-^{***}$, $^{**}-OC(=O)N(R)(CR_2)_n-^{***}$, $^{**}-$
 $(CR_2)_nN(R)C(=O)O-^{***}$, $^{**}-(OCR_2CR_2)_n-^{***}$, $^{**}-(CR_2CR_2O)_n-^{***}$, $^{**}-$
 $(OCR_2CR_2)_n(C=O)-^{***}$, $^{**}-(OCR_2CR_2)_nO(C=O)-^{***}$, $^{**}-(OCR_2CR_2)_nN(R)(C=O)-^{***}$,
 $^{**}-(OCR_2CR_2)_n(C=O)(CR_2)_p-^{***}$, $^{**}-(OCR_2CR_2)_nO(C=O)(CR_2)_p-^{***}$, $^{**}-$
 $(OCR_2CR_2)_nN(R)(C=O)(CR_2)_p-^{***}$, $^{**}-(OCR_2CR_2)_n(C=O)O(CR_2)_p-^{***}$, $^{**}-$
 $(OCR_2CR_2)_nO(C=O)O(CR_2)_p-^{***}$, $^{**}-(OCR_2CR_2)_nN(R)(C=O)O(CR_2)_p-^{***}$, $^{**}-$
 $(OCR_2CR_2)_n(C=O)N(R)(CR_2)_p-^{***}$, $^{**}-(OCR_2CR_2)_nO(C=O)N(R)(CR_2)_p-^{***}$, $^{**}-$
 $(OCR_2CR_2)_nN(R)(C=O)N(R)(CR_2)_p-^{***}$, $^{**}-OCR_2(C=O)-^{***}$, $^{**}-OCR_2(C=O)(CR_2)_p-^{***}$,
 *** , $^{**}-OCR_2(C=O)O(CR_2)_p-^{***}$, $^{**}-OCR_2(C=O)N(R)(CR_2)_p-^{***}$, $^{**}-$
 $C(=O)O(CR_2CR_2O)_nCR_2CR_2-^{***}$, $^{**}-C(=O)N(R)(CR_2CR_2O)_nCR_2CR_2-^{***}$, $^{**}-$
 $S(CR_2)_n-^{***}$, $^{**}-(CR_2)_nS-^{***}$, $^{**}-(CR_2)_nSS(CR_2)_p-^{***}$, $^{**}-SO_2(CR_2)_n-^{***}$, $^{**}-$
 $(CR_2)_nSO_2-^{***}$, $^{**}-N(R)SO_2(CR_2)_n-^{***}$, $^{**}-(CR_2)_nSO_2N(R)-^{***}$, $^{**}-SO_2N(R)(CR_2)_n-^{***}$,
 $^{**}-(CR_2)_nN(R)SO_2-^{***}$, $^{**}-(CR_2)_n-Ar-(CR_2)_p-^{***}$, $^{**}-(CR_2)_p-Ar-(CR_2)_n-^{***}$, $^{**}-$
 $O(CR_2)_n-Ar-(CR_2)_p-^{***}$, and $^{**}-(CR_2)_p-Ar-(CR_2)_nO-^{***}$, wherein ** indicates the
 position of a covalent bond with the moiety:



*** indicates the position of a covalent bond with R^4 ; each instance of n is independently a whole number selected from 1-20; each instance of p is independently an integer selected from 0-20; and R for each instance is independently selected from hydrogen, alkyl, cycloalkyl, and aryl; or two instances of R taken

together with the carbons to which they are attached form a 3-6 membered carbocyclic ring; or two instances of R taken together with the atoms to which they are attached form a 5-6 membered heterocyclic ring.

[0170] In certain embodiments of the compound of Formula **1**, the linker A comprises 2 or 3. In such instances, the linker can be represented by $-A^1-A^2-A^3-$, wherein each of A^1 , A^2 , and A^3 is independently selected from the linkers recited in the definition of A herein; or A^3 is absent. In instances in which A comprises 2 or 3 linkers, ** indicates the position of a covalent bond with the moiety:



** indicates the position of a covalent bond with A^1 , A^2 , or A^3 ; and *** indicates the position of a covalent bond with R^4 ; or *** indicates the position of a covalent bond with A^1 , A^2 , or A^3

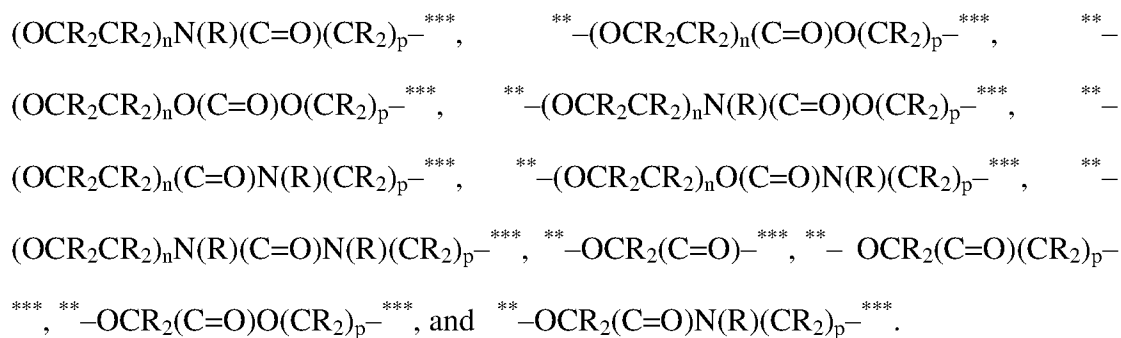
[0171] In certain embodiments of the compound of Formula **1**, the linker is at least one moiety selected from the group consisting of $^{**}-(CR_2)_n-^{***}$, $^{**}-O(CR_2)_n-^{***}$, $^{**}-O(CR_2)_nO-^{***}$, $^{**}-O(CR_2)_nC(=O)-^{***}$, $^{**}-O(CR_2)_nC(=O)N(R)(CR_2)_p-^{***}$, $^{**}-O(CR_2)_nN(R)C(=O)N(R)(CR_2)_p-^{***}$, $^{**}-O(CR_2)_nC(=O)O(CR_2)_p-^{***}$, $^{**}-O(CR_2)_nOC(=O)N(R)(CR_2)_p-^{***}$, $^{**}-O(CR_2)_nN(R)C(=O)-^{***}$, $^{**}-O(CR_2)_nN(R)C(=O)(CR_2)_p-^{***}$, $^{**}-OC(=O)(CR_2)_n-^{***}$, $^{**}-OC(=O)O(CR_2)_n-^{***}$, $^{**}-N(R)C(=O)(CR_2)_n-^{***}$, $^{**}-N(R)C(=O)O(CR_2)_n-^{***}$, $^{**}-(OCR_2CR_2)_n-^{***}$, $^{**}-(OCR_2CR_2)_n(C=O)-^{***}$, $^{**}-(OCR_2CR_2)_nO(C=O)-^{***}$, $^{**}-(OCR_2CR_2)_nN(R)(C=O)-^{***}$, $^{**}-(OCR_2CR_2)_n(C=O)(CR_2)_p-^{***}$, $^{**}-(OCR_2CR_2)_nO(C=O)(CR_2)_p-^{***}$, $^{**}-(OCR_2CR_2)_nN(R)(C=O)(CR_2)_p-^{***}$, $^{**}-(OCR_2CR_2)_n(C=O)O(CR_2)_p-^{***}$, $^{**}-(OCR_2CR_2)_nO(C=O)O(CR_2)_p-^{***}$, $^{**}-(OCR_2CR_2)_nN(R)(C=O)O(CR_2)_p-^{***}$, $^{**}-(OCR_2CR_2)_n(C=O)N(R)(CR_2)_p-^{***}$, $^{**}-(OCR_2CR_2)_nO(C=O)N(R)(CR_2)_p-^{***}$, $^{**}-(OCR_2CR_2)_nN(R)(C=O)N(R)(CR_2)_p-^{***}$, $^{**}-OCR_2(C=O)-^{***}$, $^{**}-OCR_2(C=O)(CR_2)_p-^{***}$

***, $^{**}\text{--OCR}_2(\text{C=O})\text{O}(\text{CR}_2)_p\text{--}^{***}$, $^{**}\text{--OCR}_2(\text{C=O})\text{N}(\text{R})(\text{CR}_2)_p\text{--}^{***}$, and $^{**}\text{--N}(\text{R})\text{SO}_2(\text{CR}_2)_n\text{--}^{***}$.

[0172] In certain embodiments of the compound of Formula **1**, n is a whole number selected from the group consisting of 1-10, 1-9, 1-8, 1-7, 1-6, 1-5, 1-4, and 1-3. In certain embodiments of the compound of Formula **1**, p is an integer selected from the group consisting of 0-10, 0-9, 0-8, 0-7, 0-6, 1-5, 1-4, and 1-3. In certain embodiments of the compound of Formula **1**, n is a whole number selected from the group consisting of 1-5, 1-4, and 1-3; and p is an integer selected from the group consisting of 0-6, 1-5, 1-4, and 1-3.

[0173] In certain embodiments of the compound of Formula **1**, R for each instance is independently selected from hydrogen and alkyl; or two instances of R taken together with the carbons to which they are attached form a 3-6 membered carbocyclic ring; or two instances of R taken together with the atoms to which they are attached form a 5-6 membered heterocyclic ring. In certain embodiments of the compound of Formula **1**, R for each instance is independently selected from hydrogen and C₁-C₆ alkyl. In certain embodiments of the compound of Formula **1**, R is hydrogen.

[0174] In certain embodiments of the compound of Formula **1**, the linker is at least one moiety selected from the group consisting of $^{**}\text{--}(\text{CR}_2)_n\text{--}^{***}$, $^{**}\text{--O}(\text{CR}_2)_n\text{--}^{***}$, $^{**}\text{--O}(\text{CR}_2)_n\text{O--}^{***}$, $^{**}\text{--O}(\text{CR}_2)_nC(=O)\text{--}^{***}$, $^{**}\text{--O}(\text{CR}_2)_nC(=O)\text{N}(\text{R})(\text{CR}_2)_p\text{--}^{***}$, $^{**}\text{--O}(\text{CR}_2)_n\text{N}(\text{R})\text{C}(=O)\text{N}(\text{R})(\text{CR}_2)_p\text{--}^{***}$, $^{**}\text{--O}(\text{CR}_2)_nC(=O)\text{O}(\text{CR}_2)_p\text{--}^{***}$, $^{**}\text{--O}(\text{CR}_2)_n\text{OC}(=O)\text{N}(\text{R})(\text{CR}_2)_p\text{--}^{***}$, $^{**}\text{--O}(\text{CR}_2)_n\text{N}(\text{R})\text{C}(=O)\text{--}^{***}$, $^{**}\text{--O}(\text{CR}_2)_n\text{N}(\text{R})\text{C}(=O)(\text{CR}_2)_p\text{--}^{***}$, $^{**}\text{--OC}(=O)(\text{CR}_2)_n\text{--}^{***}$, $^{**}\text{--OC}(=O)\text{O}(\text{CR}_2)_n\text{--}^{***}$, $^{**}\text{--}(\text{OCR}_2\text{CR}_2)_n\text{--}^{***}$, $^{**}\text{--}(\text{OCR}_2\text{CR}_2)_n(\text{C=O})\text{--}^{***}$, $^{**}\text{--}(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C=O})\text{--}^{***}$, $^{**}\text{--}(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C=O})\text{--}^{***}$, $^{**}\text{--}(\text{OCR}_2\text{CR}_2)_n(\text{C=O})(\text{CR}_2)_p\text{--}^{***}$, $^{**}\text{--}(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C=O})(\text{CR}_2)_p\text{--}^{***}$, and $^{**}\text{--}$.



[0175] In certain embodiments of the compound of Formula **1**, the linker is at least one moiety selected from the group consisting of $^{**}-(\text{CR}_2)_n^{***}$, $^{**}-\text{O}(\text{CR}_2)_n^{***}$, $^{**}-\text{O}(\text{CR}_2)_n\text{O}^{***}$, and $^{**}-(\text{OCR}_2\text{CR}_2)_n^{***}$.

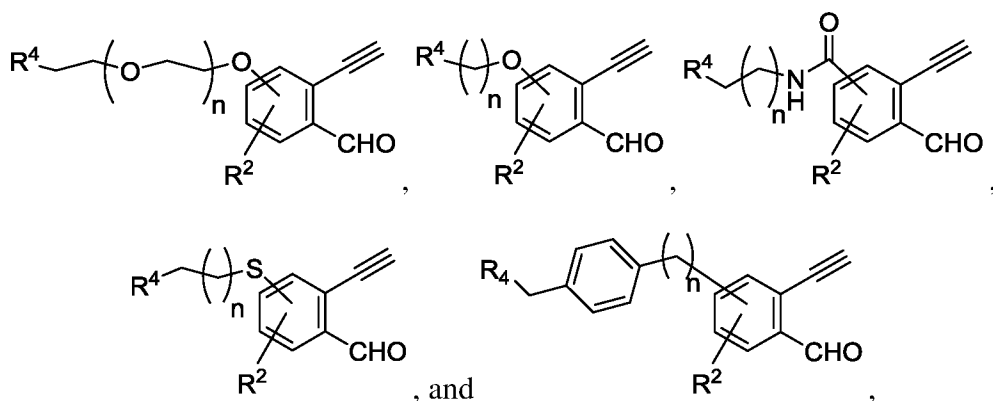
[0176] In certain embodiments of the compound of Formula **1**, R is hydrogen or C₁-C₆ alkyl. In certain embodiments of the compound of Formula **1**, R is hydrogen.

[0177] In certain embodiments of the compound of Formula **1**, R⁴ is maleimide, alkyne, azide, norbornene, or cyclooctene. In certain embodiments of the compound of Formula **1**, R⁴ is maleimide, alkyne, or azide.

[0178] In certain embodiments of the compound of Formula **1**, m is 1; R² is halide, cyano, nitro, hydrogen, or alkyl; the linker is absent, $^{**}-(\text{CR}_2)_n^{***}$, $^{**}-\text{O}(\text{CR}_2)_n^{***}$, $^{**}-\text{O}(\text{CR}_2)_n\text{O}^{***}$, or $^{**}-(\text{OCR}_2\text{CR}_2)_n^{***}$, wherein R is hydrogen and n is a whole number selected from 1-4; and R⁴ is maleimide, alkyne, azide, norbornene, or cyclooctene.

[0179] In certain embodiments of the compound of Formula **1**, m is 1; R² is hydrogen; the linker is absent, $^{**}-(\text{CR}_2)_n^{***}$, $^{**}-\text{O}(\text{CR}_2)_n^{***}$, $^{**}-\text{O}(\text{CR}_2)_n\text{O}^{***}$, or $^{**}-(\text{OCR}_2\text{CR}_2)_n^{***}$; R⁴ is maleimide, alkyne, or azide; R is hydrogen; and n for each instance is independently selected from 1, 2, 3, or 4.

[0180] In certain embodiments of the compound of Formula **1**, the compound of Formula **1** is selected from the group consisting of:



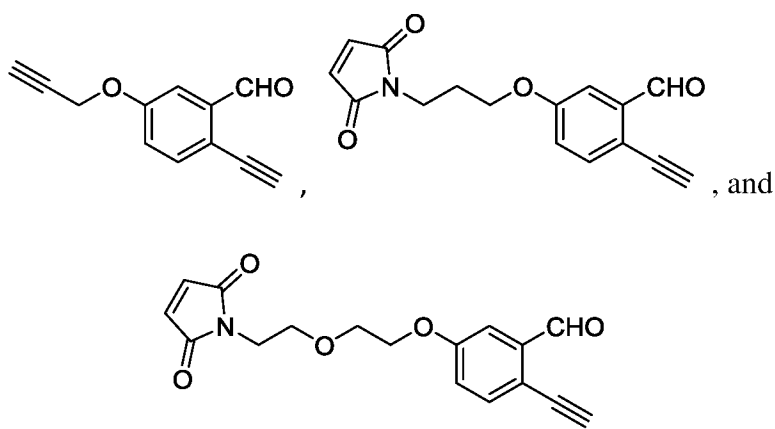
[0181] wherein n is a whole number selected from 1-4;

[0182] R^2 is hydrogen, alkyl, or $-OR^5$;

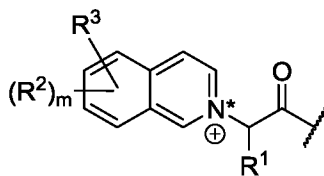
[0183] R^4 is maleimide or alkyne; and

[0184] R^5 is alkyl.

[0185] In certain embodiments of the compound of Formula **1**, the compound of Formula **1** is selected from the group consisting of:



[0186] In the first step, the compound of Formula **1** can be brought into contact with a polypeptide comprising an N-terminal amino acid under reaction conditions as described herein thereby forming an N-terminal functionalized polypeptide comprising an N-terminal amino acid represented by the structure of Formula **2**:



2

[0187] m is 1, 2, or 3;

[0188] N* is the N-terminal nitrogen of the functionalized polypeptide;

[0189] R¹ is the side chain of the N-terminal amino acid of the functionalized polypeptide and the N-terminal amino acid is a natural amino acid or a non-natural amino acid;

[0190] R² is independently for each instance selected from the group consisting of hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl, aralkyl, heterocycloalkyl, heteraryl, halide, cyano, nitro, hydroxyl, mercaptan, -OR⁵, -SR⁵, -(C=O)OR⁵, -O(C=O)R⁵, -N(R⁵)₂, -(C=O)N(R⁵)₂, -N(R⁵)(C=O)R⁵, -N(R⁵)(C=O)N(R⁵)₂, -SO₂R⁵, -N(R⁵)SO₂R⁵, and -SO₂N(R⁵)₂;

[0191] R³ is selected from the group consisting of -R⁴ or -A-R⁴, wherein A is a linker;

[0192] R⁴ is selected from the group consisting of hydroxyl, thiol, aldehyde, ketone, maleimide, *N*-hydroxysuccinimide ester, imidoester, haloacetyl, pyridyl disulfide, alkyne, azide, nitron, nitrile oxide, norbornene, oxanorbornadiene, tetrazine, cyclooctene, isocyanide, quadricyclane, and bis(dithiobenzil)nickel(II); and

[0193] R⁵ for each instance is independently selected from the group consisting of hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl, aralkyl, heterocycloalkyl, and heteraryl; or two instances of R⁵ taken together with the atoms to which they are attached form a 5-6 membered heterocycloalkyl.

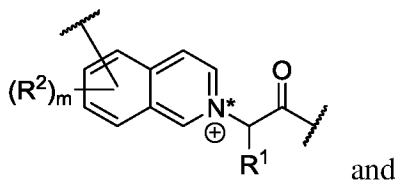
[0194] In certain embodiments of the compound of Formula **2**, m is 1.

[0195] In certain embodiments of the compound of Formula **2**, R² independently selected from is halide, cyano, nitro, hydrogen, or alkyl. In certain embodiments of the compound of Formula **2**, m is 1 and R² is halide, cyano, nitro, hydrogen, or alkyl. In certain embodiments of the compound of Formula **2**, R² is hydrogen.

[0196] R⁴ can be covalently attached directly to the 2-ethynylbenzaldehyde analog or can be covalently attached via a linker. The linker can be any linker known to those of skill in the art. The selection of which is well within the skill of a person of skill in the art. In the broadest sense, a linker is a divalent chemical moiety comprising a covalent bond or a chain of atoms that covalently attaches two groups (e.g., R⁴ and the 2-ethynylbenzaldehyde analog).

[0197] In certain embodiments of the compound of Formula **2**, the linker is at least one moiety selected from the group consisting of $^{**}-(CR_2)_n-^{***}$, $^{**}-O(CR_2)_n-^{***}$, $^{**}-(CR_2)_nO-^{***}$, $^{**}-O(CR_2)_nO-^{***}$, $^{**}-(CR_2)_nC(=O)-^{***}$, $^{**}-C(=O)(CR_2)_n-^{***}$, $^{**}-C(=O)O(CR_2)_n-^{***}$, $^{**}-(CR_2)_nOC(=O)-^{***}$, $^{**}-O(CR_2)_nC(=O)-^{***}$, $^{**}-O(CR_2)_nC(=O)N(R)(CR_2)_p-^{***}$, $^{**}-O(CR_2)_nN(R)C(=O)N(R)(CR_2)_p-^{***}$, $^{**}-O(CR_2)_nC(=O)O(CR_2)_p-^{***}$, $^{**}-O(CR_2)_nOC(=O)N(R)(CR_2)_p-^{***}$, $^{**}-C(=O)(CR_2)_nO-^{***}$, $^{**}-O(CR_2)_nN(R)C(=O)-^{***}$, $^{**}-O(CR_2)_nN(R)C(=O)(CR_2)_p-^{***}$, $^{**}-C(=O)N(R)(CR_2)_nO-^{***}$, $^{**}-OC(=O)(CR_2)_n-^{***}$, $^{**}-(CR_2)_nC(=O)O-^{***}$, $^{**}-OC(=O)O(CR_2)_n-^{***}$, $^{**}-(CR_2)_nOC(=O)O-^{***}$, $^{**}-C(=O)N(R)(CR_2)_n-^{***}$, $^{**}-(CR_2)_n(R)NC(=O)-^{***}$, $^{**}-N(R)C(=O)(CR_2)_n-^{***}$, $^{**}-(CR_2)_nC(=O)N(R)-^{***}$, $^{**}-N(R)C(=O)O(CR_2)_n-^{***}$, $^{**}-(CR_2)_nOC(=O)N(R)-^{***}$, $^{**}-OC(=O)N(R)(CR_2)_n-^{***}$, $^{**}-(CR_2)_nN(R)C(=O)O-^{***}$, $^{**}-(OCR_2CR_2)_n-^{***}$, $^{**}-(CR_2CR_2O)_n-^{***}$, $^{**}-(OCR_2CR_2)_n(C=O)-^{***}$, $^{**}-(OCR_2CR_2)_nO(C=O)-^{***}$, $^{**}-(OCR_2CR_2)_nN(R)(C=O)-^{***}$, $^{**}-(OCR_2CR_2)_n(C=O)(CR_2)_p-^{***}$, $^{**}-(OCR_2CR_2)_nO(C=O)(CR_2)_p-^{***}$, $^{**}-$

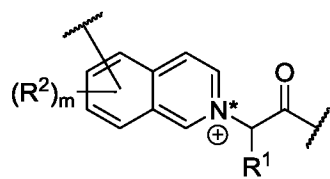
$(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})(\text{CR}_2)_p$ ***, ** $-(\text{OCR}_2\text{CR}_2)_n(\text{C}=\text{O})\text{O}(\text{CR}_2)_p$ ***, **
 $(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C}=\text{O})\text{O}(\text{CR}_2)_p$ ***, ** $-(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})\text{O}(\text{CR}_2)_p$ ***, **
 $(\text{OCR}_2\text{CR}_2)_n(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p$ ***, ** $-(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p$ ***, **
 $(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p$ ***, ** $-\text{OCR}_2(\text{C}=\text{O})$ ***, ** $-\text{OCR}_2(\text{C}=\text{O})(\text{CR}_2)_p$ ***,
***, ** $-\text{OCR}_2(\text{C}=\text{O})\text{O}(\text{CR}_2)_p$ ***, ** $-\text{OCR}_2(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p$ ***, **
 $\text{C}(\text{O})\text{O}(\text{CR}_2\text{CR}_2\text{O})_n\text{CR}_2\text{CR}_2$ ***, ** $-\text{C}(\text{O})\text{N}(\text{R})(\text{CR}_2\text{CR}_2\text{O})_n\text{CR}_2\text{CR}_2$ ***, **
 $\text{S}(\text{CR}_2)_n$ ***, ** $-(\text{CR}_2)_n\text{S}$ ***, ** $-(\text{CR}_2)_n\text{SS}(\text{CR}_2)_p$ ***, ** $-\text{SO}_2(\text{CR}_2)_n$ ***, **
 $(\text{CR}_2)_n\text{SO}_2$ ***, ** $-\text{N}(\text{R})\text{SO}_2(\text{CR}_2)_n$ ***, ** $-(\text{CR}_2)_n\text{SO}_2\text{N}(\text{R})$ ***, ** $-\text{SO}_2\text{N}(\text{R})(\text{CR}_2)_n$ ***,
** $-(\text{CR}_2)_n\text{N}(\text{R})\text{SO}_2$ ***, ** $-(\text{CR}_2)_n\text{-Ar}-(\text{CR}_2)_p$ ***, ** $-(\text{CR}_2)_p\text{-Ar}-(\text{CR}_2)_n$ ***, **
 $\text{O}(\text{CR}_2)_n\text{-Ar}-(\text{CR}_2)_p$ ***, and ** $-(\text{CR}_2)_p\text{-Ar}-(\text{CR}_2)_n\text{O}$ ***, wherein ** indicates the position of a covalent bond with the moiety:



*** indicates the position of a covalent bond with R^4 ; each instance of n is independently a whole number selected from 1-20; each instance of p is independently an integer selected from 0-20; and R for each instance is independently selected from hydrogen, alkyl, cycloalkyl, and aryl; or two instances of R taken together with the carbons to which they are attached form a 3-6 membered carbocyclic ring; or two instances of R taken together with the atoms to which they are attached form a 5-6 membered heterocyclic ring.

[0198] In certain embodiments of the compound of Formula 2, the linker A comprises 2 or 3. In such instances, the linker can be represented by $-\text{A}^1-\text{A}^2-\text{A}^3-$, wherein each of A^1 , A^2 , and A^3 is independently selected from the linkers recited in

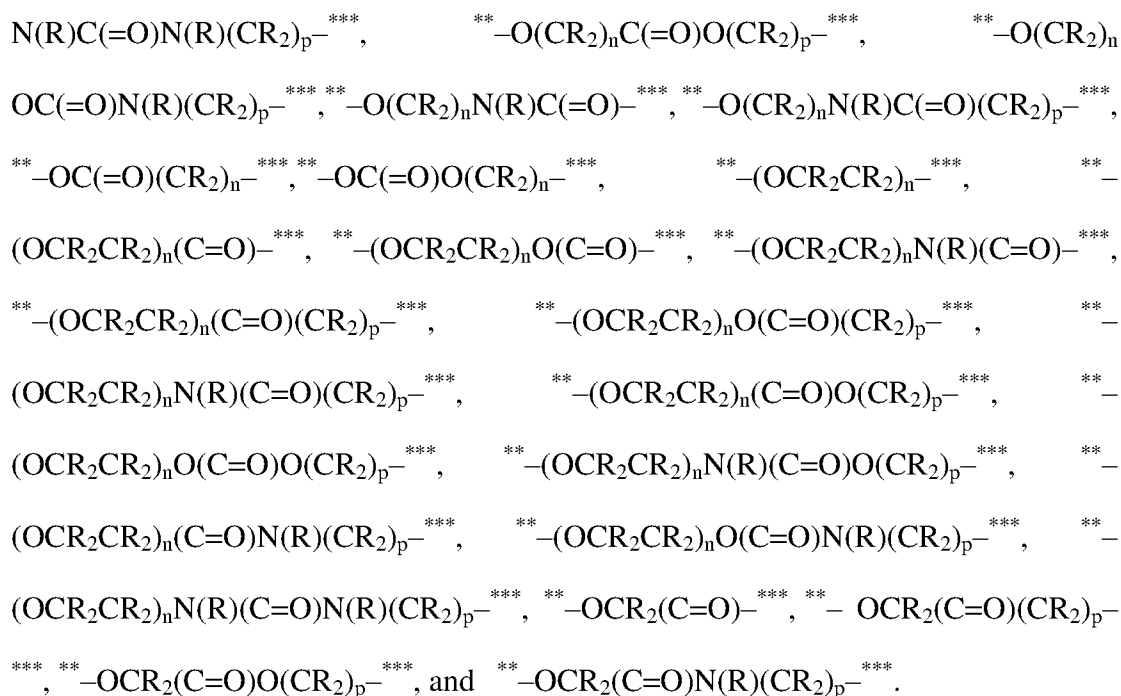
the definition of A herein; or A³ is absent. In instances in which A comprises 2 or 3 linkers, ** indicates the position of a covalent bond with the moiety:



; or ** indicates the position of a covalent bond with A¹, A², or A³; and *** indicates the position of a covalent bond with R⁴; or *** indicates the position of a covalent bond with A¹, A², or A³.

[0199] In certain embodiments of the compound of Formula 2, the linker is at least one moiety selected from the group consisting of $^{**}-(CR_2)_n^{***}$, $^{**}-O(CR_2)_n^{***}$, $^{**}-O(CR_2)_nO^{***}$, $^{**}-O(CR_2)_nC(=O)^{***}$, $^{**}-O(CR_2)_nC(=O)N(R)(CR_2)_p^{***}$, $^{**}-O(CR_2)_nN(R)C(=O)N(R)(CR_2)_p^{***}$, $^{**}-O(CR_2)_nC(=O)O(CR_2)_p^{***}$, $^{**}-O(CR_2)_nOC(=O)N(R)(CR_2)_p^{***}$, $^{**}-O(CR_2)_nN(R)C(=O)^{***}$, $^{**}-O(CR_2)_nN(R)C(=O)(CR_2)_p^{***}$, $^{**}-OC(=O)(CR_2)_n^{***}$, $^{**}-OC(=O)O(CR_2)_n^{***}$, $^{**}-N(R)C(=O)(CR_2)_n^{***}$, $^{**}-N(R)C(=O)O(CR_2)_n^{***}$, $^{**}-(OCR_2CR_2)_n^{***}$, $^{**}-(OCR_2CR_2)_n(C=O)^{***}$, $^{**}-(OCR_2CR_2)_nO(C=O)^{***}$, $^{**}-(OCR_2CR_2)_nN(R)(C=O)^{***}$, $^{**}-(OCR_2CR_2)_n(C=O)(CR_2)_p^{***}$, $^{**}-(OCR_2CR_2)_nO(C=O)(CR_2)_p^{***}$, $^{**}-(OCR_2CR_2)_nN(R)(C=O)(CR_2)_p^{***}$, $^{**}-(OCR_2CR_2)_n(C=O)O(CR_2)_p^{***}$, $^{**}-(OCR_2CR_2)_nO(C=O)O(CR_2)_p^{***}$, $^{**}-(OCR_2CR_2)_nN(R)(C=O)O(CR_2)_p^{***}$, $^{**}-(OCR_2CR_2)_n(C=O)N(R)(CR_2)_p^{***}$, $^{**}-(OCR_2CR_2)_nO(C=O)N(R)(CR_2)_p^{***}$, $^{**}-(OCR_2CR_2)_nN(R)(C=O)N(R)(CR_2)_p^{***}$, $^{**}-OCR_2(C=O)^{***}$, $^{**}-OCR_2(C=O)(CR_2)_p^{***}$, $^{**}-OCR_2(C=O)O(CR_2)_p^{***}$, $^{**}-OCR_2(C=O)N(R)(CR_2)_p^{***}$, and $^{**}-N(R)SO_2(CR_2)_n^{***}$.

[0200] In certain embodiments of the compound of Formula 2, the linker is at least one moiety selected from the group consisting of $^{**}-(CR_2)_n^{***}$, $^{**}-O(CR_2)_n^{***}$, $^{**}-O(CR_2)_nO^{***}$, $^{**}-O(CR_2)_nC(=O)^{***}$, $^{**}-O(CR_2)_nC(=O)N(R)(CR_2)_p^{***}$, $^{**}-O(CR_2)_n$



[0201] In certain embodiments of the compound of Formula **2**, the linker is at least one moiety selected from the group consisting of $\text{**}^{\text{***}}\text{-(CR}_2\text{)}_{\text{n}}^{\text{***}}$, $\text{**}^{\text{***}}\text{-O(CR}_2\text{)}_{\text{n}}^{\text{***}}$, $\text{**}^{\text{***}}\text{-O(CR}_2\text{)}_{\text{n}}\text{O}^{\text{***}}$, and $\text{**}^{\text{***}}\text{-(OCR}_2\text{CR}_2\text{)}_{\text{n}}^{\text{***}}$.

[0202] In certain embodiments of the compound of Formula **2**, R is hydrogen or C₁-C₆ alkyl. In certain embodiments of the compound of Formula **2**, R is hydrogen.

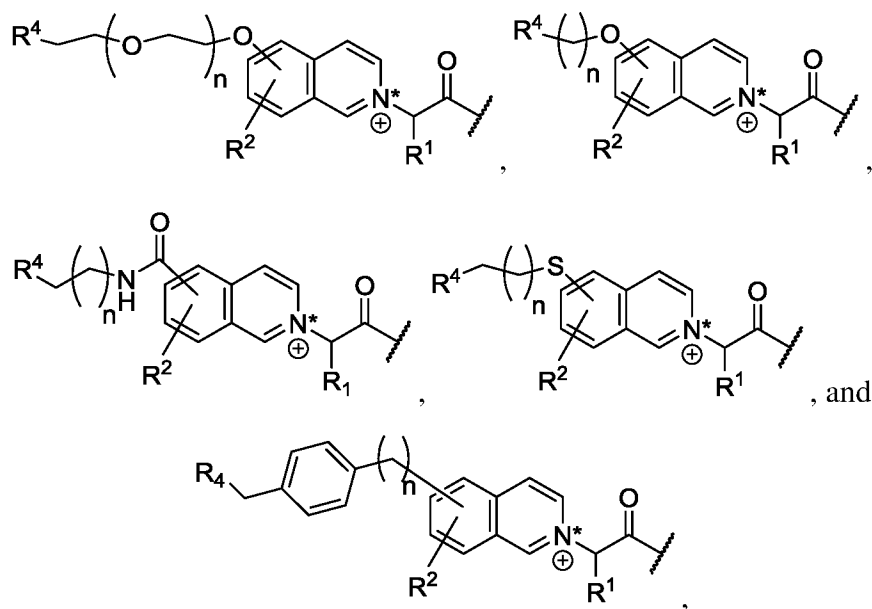
[0203] In certain embodiments of the compound of Formula **2**, R⁴ is maleimide, alkyne, azide, norbornene, or cyclooctene. In certain embodiments of the compound of Formula **2**, R⁴ is maleimide, alkyne, or azide.

[0204] In certain embodiments of the compound of Formula **2**, m is 1; R² is halide, cyano, nitro, hydrogen, or alkyl; the linker is absent, $\text{**}^{\text{***}}\text{-(CR}_2\text{)}_{\text{n}}^{\text{***}}$, $\text{**}^{\text{***}}\text{-O(CR}_2\text{)}_{\text{n}}^{\text{***}}$, $\text{**}^{\text{***}}\text{-O(CR}_2\text{)}_{\text{n}}\text{O}^{\text{***}}$, or $\text{**}^{\text{***}}\text{-(OCR}_2\text{CR}_2\text{)}_{\text{n}}^{\text{***}}$, wherein R is hydrogen and n is a whole number selected from 1-4; and R⁴ is maleimide, alkyne, azide, norbornene, or cyclooctene.

[0205] In certain embodiments of the compound of Formula **2**, m is 1; R² is hydrogen; the linker is absent, $\text{**}^{\text{***}}\text{-(CR}_2\text{)}_{\text{n}}^{\text{***}}$, $\text{**}^{\text{***}}\text{-O(CR}_2\text{)}_{\text{n}}^{\text{***}}$, $\text{**}^{\text{***}}\text{-O(CR}_2\text{)}_{\text{n}}\text{O}^{\text{***}}$, or $\text{**}^{\text{***}}\text{-}$

$(OCR_2CR_2)_{n-}^{***}$; R^4 is maleimide, alkyne, or azide; R is hydrogen; and n for each instance is independently selected from 1, 2, 3, or 4.

[0206] In certain embodiments of the compound of Formula 2, the compound of Formula 1 is selected from the group consisting of:



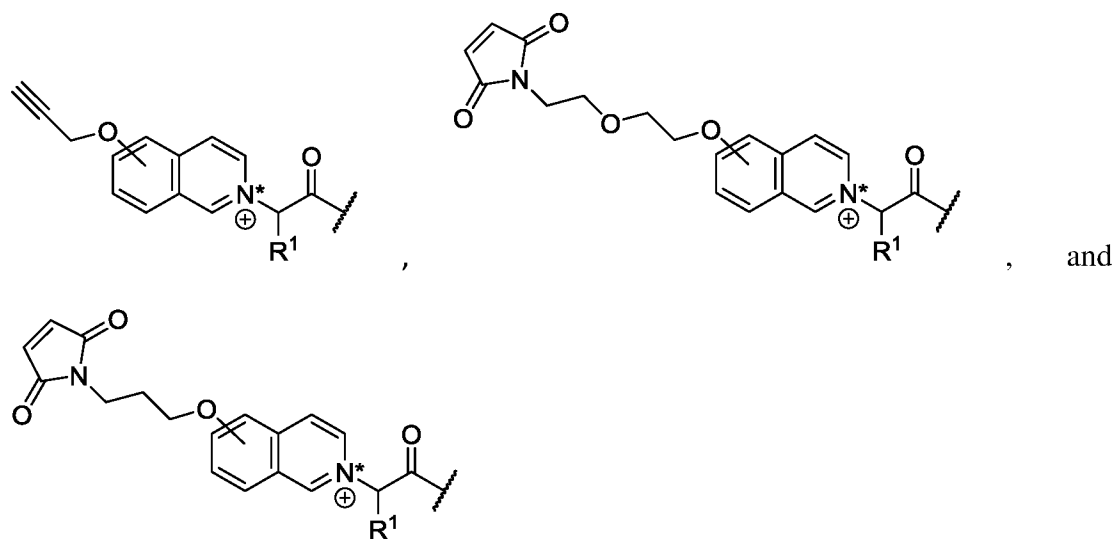
[0207] wherein n is a whole number selected from 1-4;

[0208] R^2 is hydrogen, alkyl, or $-OR^5$;

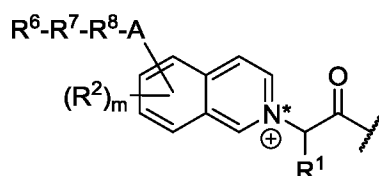
[0209] R^4 is maleimide or alkyne; and

[0210] R^5 is alkyl.

[0211] In certain embodiments of the compound of Formula 2, the compound of Formula 2 is selected from the group consisting of:



[0212] In the second step, the compound of Formula 2 can be brought into contact with an active agent comprising a complimentary reactive functional group capable of reacting with functionality present in R^4 thereby forming an active agent conjugate represented by the structure of Formula 4:



4

[0213] wherein m is 1, 2, or 3;

[0214] N^* is the N-terminal nitrogen of a polypeptide;

[0215] R^1 is the side chain of the N-terminal amino acid of the polypeptide and the N-terminal amino acid is a natural amino acid or a non-natural amino acid;

[0216] R^2 is independently for each instance selected from the group consisting of hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl, aralkyl, heterocycloalkyl, heteraryl, halide, cyano, nitro, hydroxyl, mercaptan, $-OR^5$, $-SR^5$, $-(C=O)OR^5$, $-O(C=O)R^5$,

-N(R⁵)₂, -(C=O)N(R⁵)₂, -N(R⁵)(C=O)R⁵, -N(R⁵)(C=O)N(R⁵)₂, -SO₂R⁵, -N(R⁵)SO₂R⁵,
and -SO₂N(R⁵)₂;

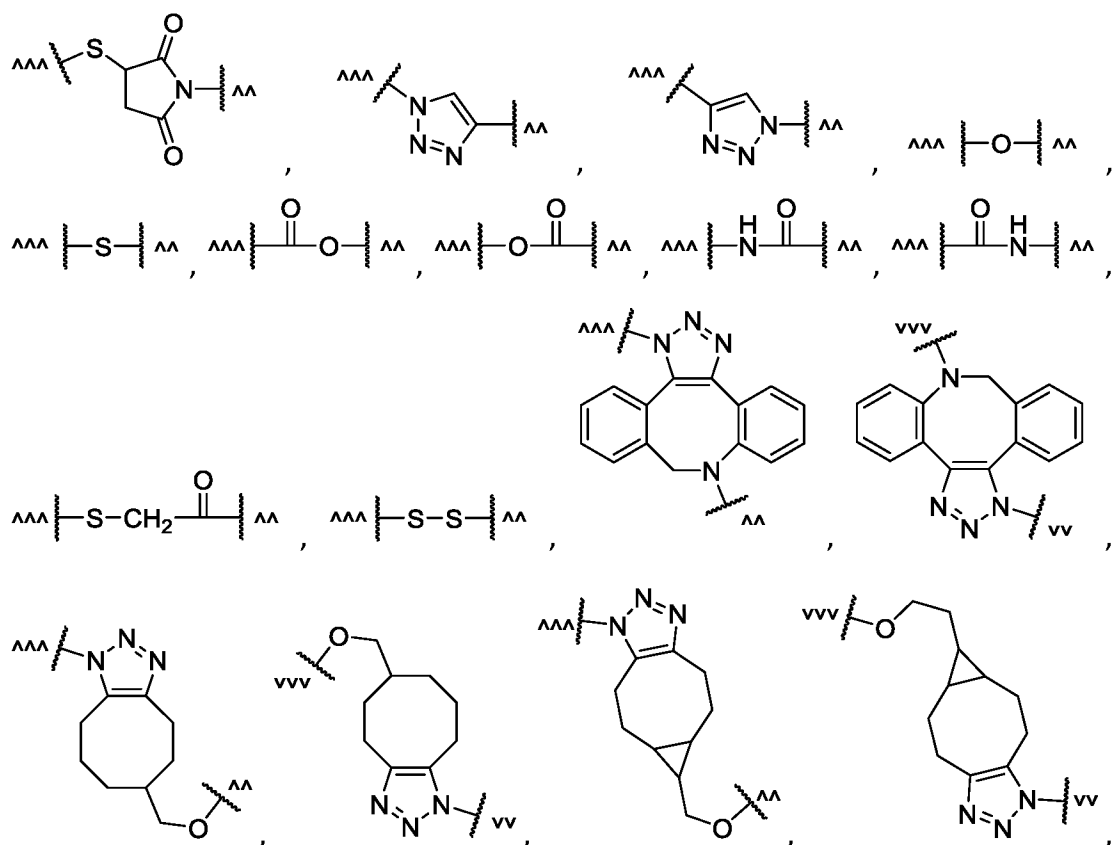
[0217] A is a linker or is absent;

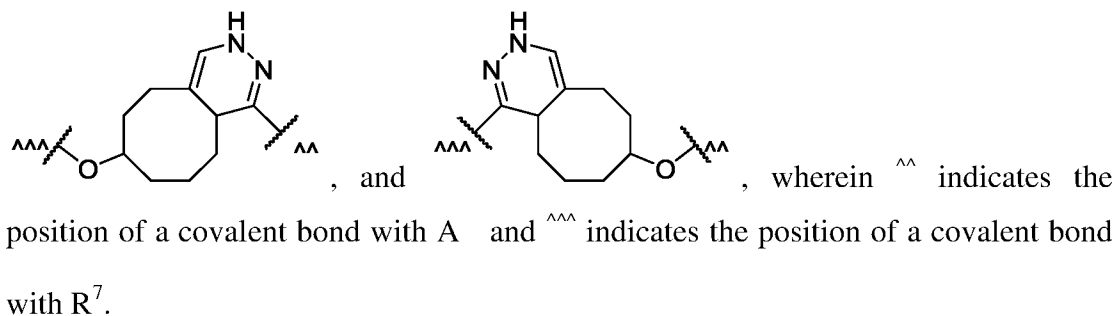
[0218] R⁵ for each instance is independently selected from the group consisting of hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl, aralkyl, heterocycloalkyl, and heteraryl; or two instances of R⁵ taken together with the atoms to which they are attached form a 5-6 membered heterocycloalkyl;

[0219] R^6 is an active agent;

[0220] R⁷ is a spacer or absent; and

[0221] R^8 is selected from the group consisting of:





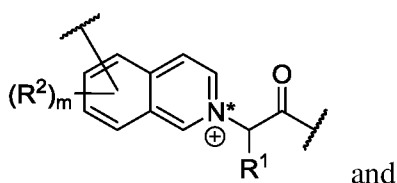
[0222] In certain embodiments of the compound of Formula **4**, R² independently selected from is halide, cyano, nitro, hydrogen, or alkyl. In certain embodiments of the compound of Formula **4**, m is 1 and R² is halide, cyano, nitro, hydrogen, or alkyl. In certain embodiments of the compound of Formula **4**, R² is hydrogen.

[0223] In certain embodiments of the compound of Formula **4**, the spacer is at least one moiety selected from the group consisting of a natural amino acid, a non-natural amino acid, piperazine, 4-amino piperidine, 4-(aminomethyl)piperidine, $\text{N(R)(CR}_2\text{)}_n$, $\text{-(CR}_2\text{)}_n$, $\text{-O(CR}_2\text{)}_n$, $\text{-(CR}_2\text{)}_n\text{O-}$, $\text{-O(CR}_2\text{)}_n\text{O-}$, $\text{-(CR}_2\text{)}_n\text{C(=O)-}$, $\text{-C(=O)(CR}_2\text{)}_n$, $\text{-C(=O)O(CR}_2\text{)}_n$, $\text{-(CR}_2\text{)}_n\text{OC(=O)-}$, $\text{-O(CR}_2\text{)}_n\text{C(=O)-}$, $\text{-O(CR}_2\text{)}_n\text{C(=O)N(R)(CR}_2\text{)}_p$, $\text{-O(CR}_2\text{)}_n\text{N(R)C(=O)N(R)(CR}_2\text{)}_p$, $\text{-O(CR}_2\text{)}_n\text{C(=O)O(CR}_2\text{)}_p$, $\text{-O(CR}_2\text{)}_n\text{OC(=O)N(R)(CR}_2\text{)}_p$, $\text{-C(=O)(CR}_2\text{)}_n\text{O-}$, $\text{-O(CR}_2\text{)}_n\text{N(R)C(=O)-}$, $\text{-O(CR}_2\text{)}_n\text{N(R)C(=O)(CR}_2\text{)}_p$, $\text{-C(=O)N(R)(CR}_2\text{)}_n\text{O-}$, $\text{-OC(=O)(CR}_2\text{)}_n$, $\text{-(CR}_2\text{)}_n\text{C(=O)O-}$, $\text{-OC(=O)O(CR}_2\text{)}_n$, $\text{-(CR}_2\text{)}_n\text{OC(=O)O-}$, $\text{C(=O)N(R)(CR}_2\text{)}_n$, $\text{-(CR}_2\text{)}_n\text{(R)NC(=O)-}$, $\text{-N(R)C(=O)(CR}_2\text{)}_n$, $\text{-(CR}_2\text{)}_n\text{C(=O)N(R)-}$, $\text{-N(R)C(=O)O(CR}_2\text{)}_n$, $\text{-(CR}_2\text{)}_n\text{OC(=O)N(R)-}$, $\text{-OC(=O)N(R)(CR}_2\text{)}_n$, $\text{-(CR}_2\text{)}_n\text{N(R)C(=O)O-}$, $\text{-(OCR}_2\text{CR}_2\text{)}_n$, $\text{-(CR}_2\text{CR}_2\text{O)}_n$, $\text{-(OCR}_2\text{CR}_2\text{)}_n\text{(C=O)-}$, $\text{-(OCR}_2\text{CR}_2\text{)}_n\text{O(C=O)-}$, $\text{-(OCR}_2\text{CR}_2\text{)}_n\text{N(R)(C=O)-}$, $\text{-(OCR}_2\text{CR}_2\text{)}_n\text{(C=O)(CR}_2\text{)}_p$, $\text{-(OCR}_2\text{CR}_2\text{)}_n\text{O(C=O)(CR}_2\text{)}_p$, $\text{(OCR}_2\text{CR}_2\text{)}_n\text{N(R)(C=O)(CR}_2\text{)}_p$, $\text{-(OCR}_2\text{CR}_2\text{)}_n\text{(C=O)O(CR}_2\text{)}_p$, -

$(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C}=\text{O})\text{O}(\text{CR}_2)_p^{\#}$, $^{\#}-(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})\text{O}(\text{CR}_2)_p^{\#}$, $^{\#}-$
 $(\text{OCR}_2\text{CR}_2)_n(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p^{\#}$, $^{\#}-(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p^{\#}$, $^{\#}-$
 $(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p^{\#}$, $^{\#}-\text{OCR}_2(\text{C}=\text{O})-$, $^{\#}-\text{OCR}_2(\text{C}=\text{O})(\text{CR}_2)_p^{\#}$, $^{\#}-$
 $\text{OCR}_2(\text{C}=\text{O})\text{O}(\text{CR}_2)_p^{\#}$, $^{\#}-\text{OCR}_2(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p^{\#}$, $^{\#}-$
 $\text{C}(\text{C}=\text{O})\text{O}(\text{CR}_2\text{CR}_2\text{O})_n\text{CR}_2\text{CR}_2-$, $^{\#}-\text{C}(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2\text{CR}_2\text{O})_n\text{CR}_2\text{CR}_2-$, $^{\#}-\text{S}(\text{CR}_2)_n-$,
 $^{\#}- (\text{CR}_2)_n\text{S}-$, $^{\#}-(\text{CR}_2)_n\text{SS}(\text{CR}_2)_p^{\#}$, $^{\#}-\text{SO}_2(\text{CR}_2)_n-$, $^{\#}-(\text{CR}_2)_n\text{SO}_2-$, $^{\#}-$
 $\text{N}(\text{R})\text{SO}_2(\text{CR}_2)_n-$, $^{\#}-(\text{CR}_2)_n\text{SO}_2\text{N}(\text{R})-$, $^{\#}-\text{SO}_2\text{N}(\text{R})(\text{CR}_2)_n-$, $^{\#}-(\text{CR}_2)_n\text{N}(\text{R})\text{SO}_2-$,
 $^{\#}-(\text{CR}_2)_n-\text{Ar}-(\text{CR}_2)_p^{\#}$, $^{\#}-(\text{CR}_2)_p-\text{Ar}-(\text{CR}_2)_n-$, $^{\#}-\text{O}(\text{CR}_2)_n-\text{Ar}-(\text{CR}_2)_p^{\#}$, and $^{\#}-$
 $(\text{CR}_2)_p-\text{Ar}-(\text{CR}_2)_n\text{O}-$, wherein $^{\#}$ indicates the position of a covalent bond with R^6
 and $^{\#}$ indicates the position of a covalent bond with R^8 ; each instance of n is
 independently a whole number selected from 1-20; p is an integer selected from 0-20;
 and R for each instance is independently selected from hydrogen, alkyl, cycloalkyl,
 and aryl; or two instances of R taken together with the carbons to which they are
 attached form a 3-6 membered carbocyclic ring; or two instances of R taken together
 with the atoms to which they are attached form a 5-6 membered heterocyclic ring.

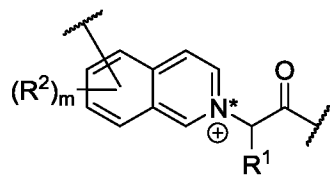
[0224] In certain embodiments of the compound of Formula 4, the linker is at least
 one moiety selected from the group consisting of $^{\#}-(\text{CR}_2)_n-$, $^{\#}-\text{O}(\text{CR}_2)_n-$, $^{\#}-$
 $(\text{CR}_2)_n\text{O}-$, $^{\#}-\text{O}(\text{CR}_2)_n\text{O}-$, $^{\#}-(\text{CR}_2)_n\text{C}(\text{C}=\text{O})-$, $^{\#}-\text{C}(\text{C}=\text{O})(\text{CR}_2)_n-$, $^{\#}-$
 $\text{C}(\text{C}=\text{O})\text{O}(\text{CR}_2)_n-$, $^{\#}-(\text{CR}_2)_n\text{OC}(\text{C}=\text{O})-$, $^{\#}-\text{O}(\text{CR}_2)_n\text{C}(\text{C}=\text{O})-$, $^{\#}-$
 $\text{O}(\text{CR}_2)_n\text{C}(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p^{\#}$, $^{\#}-\text{O}(\text{CR}_2)_n\text{N}(\text{R})\text{C}(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p^{\#}$, $^{\#}-$
 $\text{O}(\text{CR}_2)_n\text{C}(\text{C}=\text{O})\text{O}(\text{CR}_2)_p^{\#}$, $^{\#}-\text{O}(\text{CR}_2)_n \text{OC}(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p^{\#}$, $^{\#}-\text{C}(\text{C}=\text{O})(\text{CR}_2)_n\text{O}-$
 $^{\#}$, $^{\#}-\text{O}(\text{CR}_2)_n\text{N}(\text{R})\text{C}(\text{C}=\text{O})-$, $^{\#}-\text{O}(\text{CR}_2)_n\text{N}(\text{R})\text{C}(\text{C}=\text{O})(\text{CR}_2)_p^{\#}$, $^{\#}-$
 $\text{C}(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_n\text{O}-$, $^{\#}-\text{OC}(\text{C}=\text{O})(\text{CR}_2)_n-$, $^{\#}-(\text{CR}_2)_n\text{C}(\text{C}=\text{O})\text{O}-$, $^{\#}-$
 $\text{OC}(\text{C}=\text{O})\text{O}(\text{CR}_2)_n-$, $^{\#}-(\text{CR}_2)_n\text{OC}(\text{C}=\text{O})\text{O}-$, $^{\#}-\text{C}(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_n-$, $^{\#}-$
 $(\text{CR}_2)_n(\text{R})\text{NC}(\text{C}=\text{O})-$, $^{\#}-\text{N}(\text{R})\text{C}(\text{C}=\text{O})(\text{CR}_2)_n-$, $^{\#}-(\text{CR}_2)_n\text{C}(\text{C}=\text{O})\text{N}(\text{R})-$, $^{\#}-$

$N(R)C(=O)O(CR_2)_n-^{<<<}$, $^{<<}(CR_2)_nOC(=O)N(R)-^{<<<}$, $^{<<}OC(=O)N(R)(CR_2)_n-^{<<<}$, $^{<<}(CR_2)_nN(R)C(=O)O-^{<<<}$, $^{<<}(OCR_2CR_2)_n-^{<<<}$, $^{<<}(CR_2CR_2O)_n-^{<<<}$, $^{<<}(OCR_2CR_2)_n(C=O)-^{<<<}$, $^{<<}(OCR_2CR_2)_nO(C=O)-^{<<<}$, $^{<<}(OCR_2CR_2)_nN(R)(C=O)-^{<<<}$, $^{<<}(OCR_2CR_2)_n(C=O)(CR_2)_p-^{<<<}$, $^{<<}(OCR_2CR_2)_nO(C=O)(CR_2)_p-^{<<<}$, $^{<<}(OCR_2CR_2)_nN(R)(C=O)(CR_2)_p-^{<<<}$, $^{<<}(OCR_2CR_2)_n(C=O)O(CR_2)_p-^{<<<}$, $^{<<}(OCR_2CR_2)_nO(C=O)O(CR_2)_p-^{<<<}$, $^{<<}(OCR_2CR_2)_nN(R)(C=O)O(CR_2)_p-^{<<<}$, $^{<<}(OCR_2CR_2)_n(C=O)N(R)(CR_2)_p-^{<<<}$, $^{<<}(OCR_2CR_2)_nO(C=O)N(R)(CR_2)_p-^{<<<}$, $^{<<}(OCR_2CR_2)_nN(R)(C=O)N(R)(CR_2)_p-^{<<<}$, $^{<<}OCR_2(C=O)-^{<<<}$, $^{<<}OCR_2(C=O)(CR_2)_p-^{<<<}$, $^{<<}OCR_2(C=O)O(CR_2)_p-^{<<<}$, $^{<<}OCR_2(C=O)N(R)(CR_2)_p-^{<<<}$, $^{<<}C(=O)O(CR_2CR_2O)_nCR_2CR_2-^{<<<}$, $^{<<}C(=O)N(R)(CR_2CR_2O)_nCR_2CR_2-^{<<<}$, $^{<<}S(CR_2)_n-^{<<<}$, $^{<<}(CR_2)_nS-^{<<<}$, $^{<<}SO_2(CR_2)_n-^{<<<}$, $^{<<}(CR_2)_nSO_2-^{<<<}$, $^{<<}N(R)SO_2(CR_2)_n-^{<<<}$, $^{<<}(CR_2)_nSO_2N(R)-^{<<<}$, $^{<<}SO_2N(R)(CR_2)_n-^{<<<}$, $^{<<}(CR_2)_nN(R)SO_2-^{<<<}$, $^{<<}(CR_2)_n-Ar-(CR_2)_p-^{<<<}$, $^{<<}(CR_2)_p-Ar-(CR_2)_n-^{<<<}$, $^{<<}O(CR_2)_n-Ar-(CR_2)_p-^{<<<}$, and $^{<<}(CR_2)_p-Ar-(CR_2)_nO-^{<<<}$, wherein $^{<<}$ indicates the position of a covalent bond with the moiety:



$^{<<<}$ indicates the position of a covalent bond with R^8 ; each instance of n is independently a whole number selected from 1-20; each instance of p is independently an integer selected from 0-20; and R for each instance is independently selected from hydrogen, alkyl, cycloalkyl, and aryl; or two instances of R taken together with the carbons to which they are attached form a 3-6 membered carbocyclic ring; or two instances of R taken together with the atoms to which they are attached form a 5-6 membered heterocyclic ring.

[0225] In certain embodiments of the compound of Formula 4, the linker A comprises 2 or 3. In such instances, the linker can be represented by $-A^1-A^2-A^3-$, wherein each of A^1 , A^2 , and A^3 is independently selected from the linkers recited in the definition of A herein; or A^3 is absent. In instances in which A comprises 2 or 3 linkers, ** indicates the position of a covalent bond with the moiety:



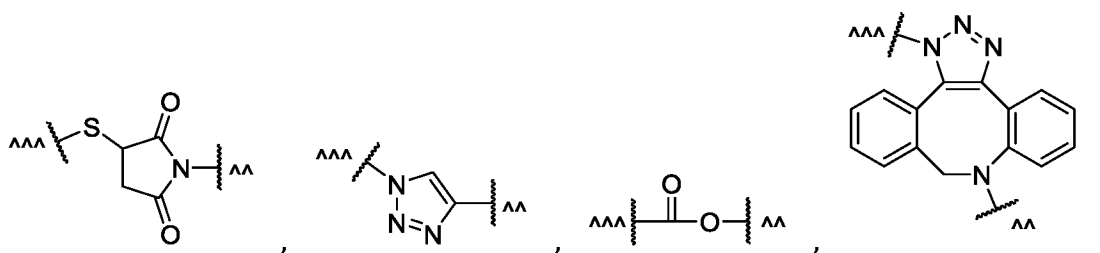
; or ** indicates the position of a covalent bond with A^1 , A^2 , or A^3 ; and *** indicates the position of a covalent bond with R^8 ; or *** indicates the position of a covalent bond with A^1 , A^2 , or A^3 .

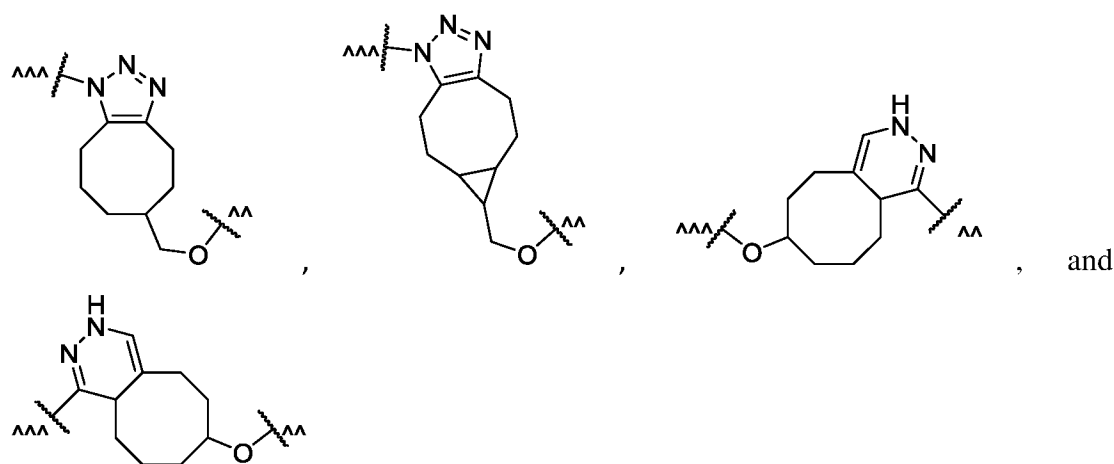
[0226] In certain embodiments of the compound of Formula 4, the linker is at least one moiety selected from the group consisting of $^{<<<}-(CR_2)_n-^{<<<}$, $^{<<<}-(O(CR_2)_n)-^{<<<}$, $^{<<<}-(O(CR_2)_nO)-^{<<<}$, $^{<<<}-(O(CR_2)_nC(=O))-^{<<<}$, $^{<<<}-(O(CR_2)_nC(=O)N(R)(CR_2)_p)-^{<<<}$, $^{<<<}-(O(CR_2)_nN(R)C(=O)N(R)(CR_2)_p)-^{<<<}$, $^{<<<}-(O(CR_2)_nC(=O)O(CR_2)_p)-^{<<<}$, $^{<<<}-(O(CR_2)_nOC(=O)N(R)(CR_2)_p)-^{<<<}$, $^{<<<}-(O(CR_2)_nN(R)C(=O))-^{<<<}$, $^{<<<}-(O(CR_2)_nN(R)C(=O)(CR_2)_p)-^{<<<}$, $^{<<<}-(OC(=O)(CR_2)_n)-^{<<<}$, $^{<<<}-(OC(=O)O(CR_2)_n)-^{<<<}$, $^{<<<}-(N(R)C(=O)(CR_2)_n)-^{<<<}$, $^{<<<}-(N(R)C(=O)O(CR_2)_n)-^{<<<}$, $^{<<<}-(OCR_2CR_2)_n-^{<<<}$, $^{<<<}-(OCR_2CR_2)_n(C=O)-^{<<<}$, $^{<<<}-(OCR_2CR_2)_nO(C=O)-^{<<<}$, $^{<<<}-(OCR_2CR_2)_n(C=O)(CR_2)_p-^{<<<}$, $^{<<<}-(OCR_2CR_2)_nO(C=O)(CR_2)_p-^{<<<}$, $^{<<<}-(OCR_2CR_2)_nN(R)(C=O)(CR_2)_p-^{<<<}$, $^{<<<}-(OCR_2CR_2)_n(C=O)O(CR_2)_p-^{<<<}$, $^{<<<}-(OCR_2CR_2)_nO(C=O)O(CR_2)_p-^{<<<}$, $^{<<<}-(OCR_2CR_2)_nN(R)(C=O)O(CR_2)_p-^{<<<}$, $^{<<<}-(OCR_2CR_2)_n(C=O)N(R)(CR_2)_p-^{<<<}$, $^{<<<}-(OCR_2CR_2)_nO(C=O)N(R)(CR_2)_p-^{<<<}$, $^{<<<}-(OCR_2CR_2)_nN(R)(C=O)N(R)(CR_2)_p-^{<<<}$, $^{<<<}-(OCR_2(C=O))-^{<<<}$, $^{<<<}-(OCR_2(C=O)(CR_2)_p)-^{<<<}$, $^{<<<}-(OCR_2(C=O)O(CR_2)_p)-^{<<<}$, $^{<<<}-(OCR_2(C=O)N(R)(CR_2)_p)-^{<<<}$, and $^{<<<}-(N(R)SO_2(CR_2)_n)-^{<<<}$.

[0227] In certain embodiments of the compound of Formula 4, the linker is at least one moiety selected from the group consisting of $\text{--}(\text{CR}_2)_n\text{--}$, $\text{--O}(\text{CR}_2)_n\text{--}$, $\text{--O}(\text{CR}_2)_n\text{O--}$, $\text{--O}(\text{CR}_2)_nC(=O)\text{--}$, $\text{--O}(\text{CR}_2)_nC(=O)N(R)(\text{CR}_2)_p\text{--}$, $\text{--O}(\text{CR}_2)_nN(R)C(=O)N(R)(\text{CR}_2)_p\text{--}$, $\text{--O}(\text{CR}_2)_nC(=O)O(\text{CR}_2)_p\text{--}$, $\text{--O}(\text{CR}_2)_nOC(=O)N(R)(\text{CR}_2)_p\text{--}$, $\text{--O}(\text{CR}_2)_nN(R)C(=O)\text{--}$, $\text{--O}(\text{CR}_2)_nN(R)C(=O)(\text{CR}_2)_p\text{--}$, $\text{--OC}(=O)(\text{CR}_2)_n\text{--}$, $\text{--OC}(=O)O(\text{CR}_2)_n\text{--}$, $\text{--(OCR}_2\text{CR}_2)_n\text{--}$, $\text{--(OCR}_2\text{CR}_2)_n(C=O)\text{--}$, $\text{--(OCR}_2\text{CR}_2)_nO(C=O)\text{--}$, $\text{--(OCR}_2\text{CR}_2)_nN(R)(C=O)\text{--}$, $\text{--(OCR}_2\text{CR}_2)_n(C=O)(\text{CR}_2)_p\text{--}$, $\text{--(OCR}_2\text{CR}_2)_nO(C=O)(\text{CR}_2)_p\text{--}$, $\text{--(OCR}_2\text{CR}_2)_nN(R)(C=O)(\text{CR}_2)_p\text{--}$, $\text{--(OCR}_2\text{CR}_2)_n(C=O)O(\text{CR}_2)_p\text{--}$, $\text{--(OCR}_2\text{CR}_2)_nO(C=O)O(\text{CR}_2)_p\text{--}$, $\text{--(OCR}_2\text{CR}_2)_nN(R)(C=O)O(\text{CR}_2)_p\text{--}$, $\text{--(OCR}_2\text{CR}_2)_n(C=O)N(R)(\text{CR}_2)_p\text{--}$, $\text{--(OCR}_2\text{CR}_2)_nO(C=O)N(R)(\text{CR}_2)_p\text{--}$, $\text{--(OCR}_2\text{CR}_2)_nN(R)(C=O)N(R)(\text{CR}_2)_p\text{--}$, $\text{--OCR}_2(C=O)\text{--}$, $\text{--OCR}_2(C=O)(\text{CR}_2)_p\text{--}$, $\text{--OCR}_2(C=O)O(\text{CR}_2)_p\text{--}$, and $\text{--OCR}_2(C=O)N(R)(\text{CR}_2)_p\text{--}$.

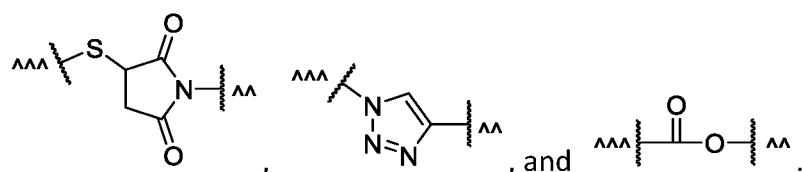
[0228] In certain embodiments of the compound of Formula 4, the linker is at least one moiety selected from the group consisting of $\text{--}(\text{CR}_2)_n\text{--}$, $\text{--O}(\text{CR}_2)_n\text{--}$, $\text{--O}(\text{CR}_2)_n\text{O--}$, and $\text{--(OCR}_2\text{CR}_2)_n\text{--}$.

[0229] In certain embodiments of the compound of Formula 4, R^8 is selected from the group consisting of:



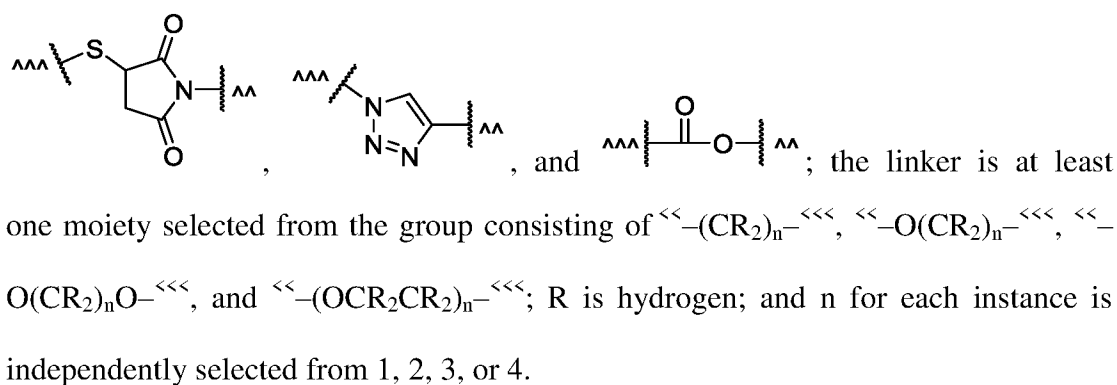


[0230] In certain embodiments of the compound of Formula 4, R^8 is selected from the group consisting of:

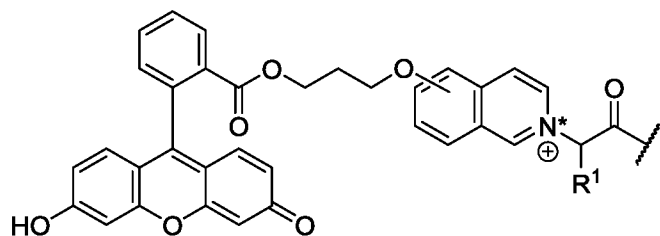


[0231] In certain embodiments of the compound of Formula 4, R is hydrogen or C_1 - C_6 alkyl. In certain embodiments of the compound of Formula 4, R is hydrogen.

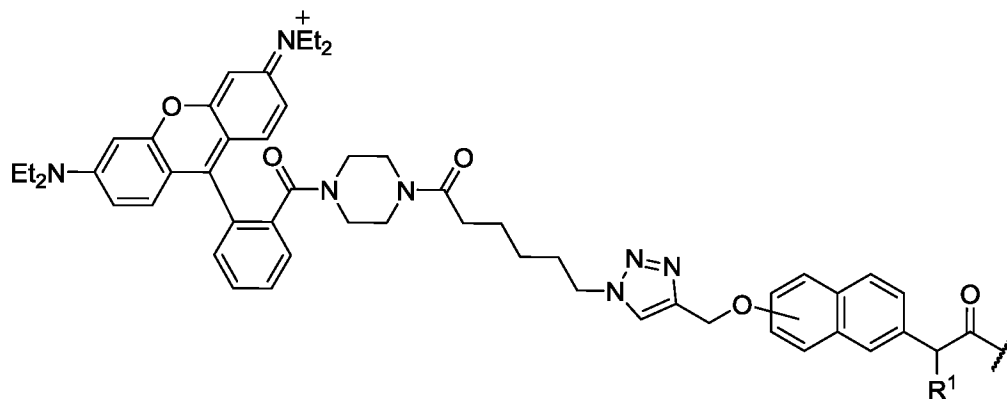
[0232] In certain embodiments of the compound of Formula 4, the spacer is at least one of piperazine, $^{\#}-(CR_2)_n-^{\#}$, $^{\#}-O(CR_2)_n-^{\#}$, $^{\#}-(CR_2)_nO-^{\#}$, $^{\#}-(CR_2)_nC(=O)-^{\#}$, $^{\#}-C(=O)O(CR_2)_n-^{\#}$, or $^{\#}-(CR_2)_nOC(=O)-^{\#}$; R^8 is selected from the group consisting of:



[0233] In certain embodiments of the compound of Formula 4, the compound is selected from the group consisting of:

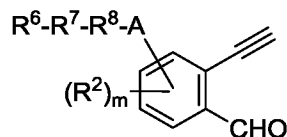


and



[0234] The selection of the appropriate complimentary reactive functional group can depend on the chemical reactivity of R^4 . For example, when R^4 is a maleimide, the complimentary reactive functional group can be a thiol or conjugate salt thereof; when R^4 is an alkyne, the complimentary reactive functional group can be an azide, nitrile oxide, or nitron; and when R^4 is an azide, the complimentary reactive functional group can be an alkyne. The selection of the appropriate complimentary reactive functional group in the active agent and reaction conditions for the reaction of the active agent comprising the complimentary reactive functional group is well within the skill of a person of ordinary skill in the art.

[0235] In instances in which the methods described herein are used in a one-step method for selectively conjugating an active agent to the N-terminal amine of a polypeptide, the 2-ethynylbenzaldehyde analog can comprise the active agent. In such instances, the 2-ethynylbenzaldehyde analog can be represented by the compound of Formula 3:



3

[0236] wherein m is 1, 2, or 3;

[0237] R^2 is independently for each instance selected from the group consisting of hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl, aralkyl, heterocycloalkyl, heteraryl, halide, cyano, nitro, hydroxyl, mercaptan, $-OR^5$, $-SR^5$, $-(C=O)OR^5$, $-O(C=O)R^5$, $-N(R^5)_2$, $-(C=O)N(R^5)_2$, $-N(R^5)(C=O)R^5$, $-N(R^5)(C=O)N(R^5)_2$, $-SO_2R^5$, $-N(R^5)SO_2R^5$, and $-SO_2N(R^5)_2$;

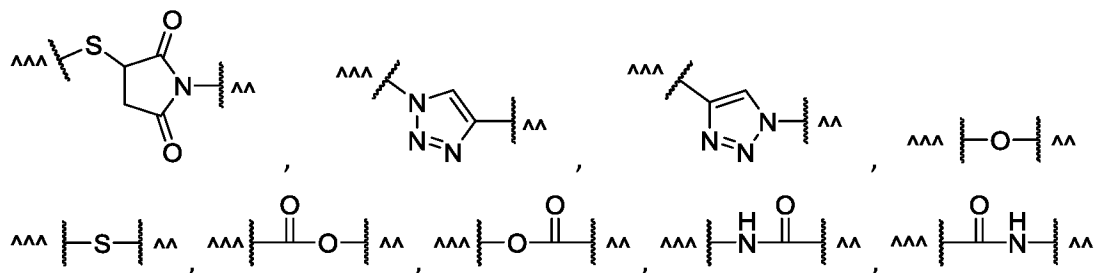
[0238] A is a linker or is absent;

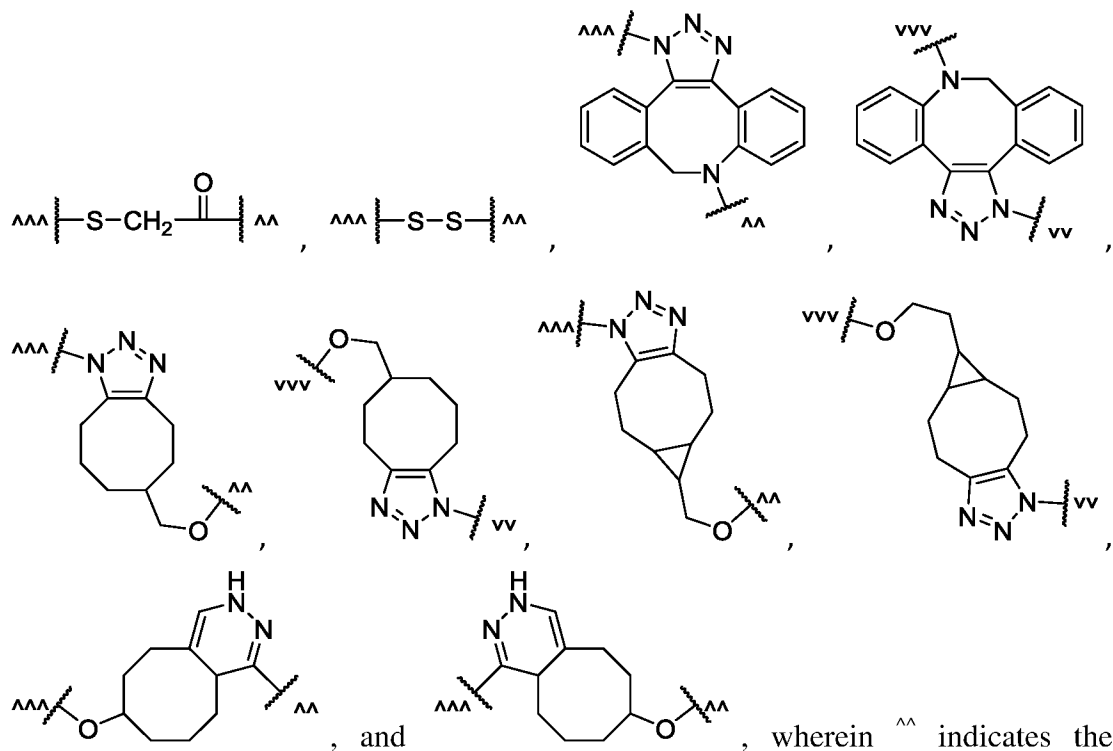
[0239] R^5 for each instance is independently selected from the group consisting of hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl, aralkyl, heterocycloalkyl, and heteraryl; or two instances of R^5 taken together with the atoms to which they are attached form a 5-6 membered heterocycloalkyl;

[0240] R^6 is an active agent;

[0241] R^7 is a spacer or absent; and

[0242] R^8 is selected from the group consisting of:







, and 

, wherein [^] indicates the position of a covalent bond with A and ^{^^} indicates the position of a covalent bond with R⁷.

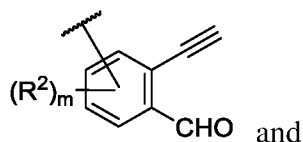
[0243] In certain embodiments of the compound of Formula **3**, R² independently selected from is halide, cyano, nitro, hydrogen, or alkyl. In certain embodiments of the compound of Formula **3**, m is 1 and R² is halide, cyano, nitro, hydrogen, or alkyl. In certain embodiments of the compound of Formula **3**, R² is hydrogen.

[0244] In certain embodiments of the compound of Formula **3**, the spacer is at least one moiety selected from the group consisting of a natural amino acid, a non-natural amino acid, piperazine, 4-amino piperidine, 4-(aminomethyl)piperidine, $^{\#}-\text{N(R)}(\text{CR}_2)_n-^{\#\#}$, $^{\#}-(\text{CR}_2)_n-^{\#\#}$, $^{\#}-\text{O}(\text{CR}_2)_n-^{\#\#}$, $^{\#}-(\text{CR}_2)_n\text{O}-^{\#\#}$, $^{\#}-\text{O}(\text{CR}_2)_n\text{O}-^{\#\#}$, $^{\#}-(\text{CR}_2)_n\text{C(=O)}-^{\#\#}$, $^{\#}-\text{C(=O)}(\text{CR}_2)_n-^{\#\#}$, $^{\#}-\text{C(=O)}\text{O}(\text{CR}_2)_n-^{\#\#}$, $^{\#}-(\text{CR}_2)_n\text{OC(=O)}-^{\#\#}$, $^{\#}-\text{O}(\text{CR}_2)_n\text{C(=O)}-^{\#\#}$, $^{\#}-\text{O}(\text{CR}_2)_n\text{C(=O)}\text{N(R)}(\text{CR}_2)_p-^{\#\#}$, $^{\#}-\text{O}(\text{CR}_2)_n\text{N(R)}\text{C(=O)}\text{N(R)}(\text{CR}_2)_p-^{\#\#}$, $^{\#}-\text{O}(\text{CR}_2)_n\text{C(=O)}\text{O}(\text{CR}_2)_p-^{\#\#}$, $^{\#}-\text{O}(\text{CR}_2)_n\text{OC(=O)}\text{N(R)}(\text{CR}_2)_p-^{\#\#}$, $^{\#}-\text{C(=O)}(\text{CR}_2)_n\text{O}-^{\#\#}$, $^{\#}-\text{O}(\text{CR}_2)_n\text{N(R)}\text{C(=O)}-^{\#\#}$, $^{\#}-$

$O(CR_2)_nN(R)C(=O)(CR_2)_p-^{\#}$, $^{\#}-C(=O)N(R)(CR_2)_nO-^{\#}$, $^{\#}-OC(=O)(CR_2)_n-^{\#}$, $^{\#}-$
 $(CR_2)_nC(=O)O-^{\#}$, $^{\#}-OC(=O)O(CR_2)_n-^{\#}$, $^{\#}-(CR_2)_nOC(=O)O-^{\#}$, $^{\#}-$
 $C(=O)N(R)(CR_2)_n-^{\#}$, $^{\#}-(CR_2)_n(R)NC(=O)-^{\#}$, $^{\#}-N(R)C(=O)(CR_2)_n-^{\#}$, $^{\#}-$
 $(CR_2)_nC(=O)N(R)-^{\#}$, $^{\#}-N(R)C(=O)O(CR_2)_n-^{\#}$, $^{\#}-(CR_2)_nOC(=O)N(R)-^{\#}$, $^{\#}-$
 $OC(=O)N(R)(CR_2)_n-^{\#}$, $^{\#}-(CR_2)_nN(R)C(=O)O-^{\#}$, $^{\#}-(OCR_2CR_2)_n-^{\#}$, $^{\#}-(CR_2CR_2O)_n-$
 $^{\#}$, $^{\#}-(OCR_2CR_2)_n(C=O)-^{\#}$, $^{\#}-(OCR_2CR_2)_nO(C=O)-^{\#}$, $^{\#}-(OCR_2CR_2)_nN(R)(C=O)-^{\#}$,
 $^{\#}-(OCR_2CR_2)_n(C=O)(CR_2)_p-^{\#}$, $^{\#}-(OCR_2CR_2)_nO(C=O)(CR_2)_p-^{\#}$, $^{\#}-$
 $(OCR_2CR_2)_nN(R)(C=O)(CR_2)_p-^{\#}$, $^{\#}-(OCR_2CR_2)_n(C=O)O(CR_2)_p-^{\#}$, $^{\#}-$
 $(OCR_2CR_2)_nO(C=O)O(CR_2)_p-^{\#}$, $^{\#}-(OCR_2CR_2)_nN(R)(C=O)O(CR_2)_p-^{\#}$, $^{\#}-$
 $(OCR_2CR_2)_n(C=O)N(R)(CR_2)_p-^{\#}$, $^{\#}-(OCR_2CR_2)_nO(C=O)N(R)(CR_2)_p-^{\#}$, $^{\#}-$
 $(OCR_2CR_2)_nN(R)(C=O)N(R)(CR_2)_p-^{\#}$, $^{\#}-OCR_2(C=O)-^{\#}$, $^{\#}-OCR_2(C=O)(CR_2)_p-^{\#}$, $^{\#}-$
 $OCR_2(C=O)O(CR_2)_p-^{\#}$, $^{\#}-OCR_2(C=O)N(R)(CR_2)_p-^{\#}$, $^{\#}-$
 $C(=O)O(CR_2CR_2O)_nCR_2CR_2-^{\#}$, $^{\#}-C(=O)N(R)(CR_2CR_2O)_nCR_2CR_2-^{\#}$, $^{\#}-S(CR_2)_n-^{\#}$,
 $^{\#}-(CR_2)_nS-^{\#}$, $^{\#}-(CR_2)_nSS(CR_2)_p-^{\#}$, $^{\#}-SO_2(CR_2)_n-^{\#}$, $^{\#}-(CR_2)_nSO_2-^{\#}$, $^{\#}-$
 $N(R)SO_2(CR_2)_n-^{\#}$, $^{\#}-(CR_2)_nSO_2N(R)-^{\#}$, $^{\#}-SO_2N(R)(CR_2)_n-^{\#}$, $^{\#}-(CR_2)_nN(R)SO_2-^{\#}$,
 $^{\#}-(CR_2)_n-Ar-(CR_2)_p-^{\#}$, $^{\#}-(CR_2)_p-Ar-(CR_2)_n-^{\#}$, $^{\#}-O(CR_2)_n-Ar-(CR_2)_p-^{\#}$, and $^{\#}-$
 $(CR_2)_p-Ar-(CR_2)_nO-^{\#}$, wherein $^{\#}$ indicates the position of a covalent bond with R^6
and $^{\#}$ indicates the position of a covalent bond with R^8 ; each instance of n is
independently a whole number selected from 1-20; p is an integer selected from 0-20;
and R for each instance is independently selected from hydrogen, alkyl, cycloalkyl,
and aryl; or two instances of R taken together with the carbons to which they are
attached form a 3-6 membered carbocyclic ring; or two instances of R taken together
with the atoms to which they are attached form a 5-6 membered heterocyclic ring.

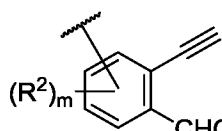
[0245] In certain embodiments of the compound of Formula 3, the linker is at least one moiety selected from the group consisting of $^{\#}- (CR_2)_n-^{\#}$, $^{\#}-O(CR_2)_n-^{\#}$, $^{\#}-$

$(\text{CR}_2)_n\text{O}-^{\lll}$, $^{\lll}\text{O}(\text{CR}_2)_n\text{O}-^{\lll}$, $^{\lll}-(\text{CR}_2)_n\text{C}(=\text{O})-^{\lll}$, $^{\lll}-\text{C}(=\text{O})(\text{CR}_2)_n-^{\lll}$, $^{\lll}-$
 $\text{C}(=\text{O})\text{O}(\text{CR}_2)_n-^{\lll}$, $^{\lll}-(\text{CR}_2)_n\text{OC}(=\text{O})-^{\lll}$, $^{\lll}-\text{O}(\text{CR}_2)_n\text{C}(=\text{O})-^{\lll}$, $^{\lll}-$
 $\text{O}(\text{CR}_2)_n\text{C}(=\text{O})\text{N}(\text{R})(\text{CR}_2)_p-^{\lll}$, $^{\lll}-\text{O}(\text{CR}_2)_n\text{N}(\text{R})\text{C}(=\text{O})\text{N}(\text{R})(\text{CR}_2)_p-^{\lll}$, $^{\lll}-$
 $\text{O}(\text{CR}_2)_n\text{C}(=\text{O})\text{O}(\text{CR}_2)_p-^{\lll}$, $^{\lll}-\text{O}(\text{CR}_2)_n\text{OC}(=\text{O})\text{N}(\text{R})(\text{CR}_2)_p-^{\lll}$, $^{\lll}-\text{C}(=\text{O})(\text{CR}_2)_n\text{O}-$
 $^{\lll}$, $^{\lll}-\text{O}(\text{CR}_2)_n\text{N}(\text{R})\text{C}(=\text{O})-^{\lll}$, $^{\lll}-\text{O}(\text{CR}_2)_n\text{N}(\text{R})\text{C}(=\text{O})(\text{CR}_2)_p-^{\lll}$, $^{\lll}-$
 $\text{C}(=\text{O})\text{N}(\text{R})(\text{CR}_2)_n\text{O}-^{\lll}$, $^{\lll}-\text{OC}(=\text{O})(\text{CR}_2)_n-^{\lll}$, $^{\lll}-(\text{CR}_2)_n\text{C}(=\text{O})\text{O}-^{\lll}$, $^{\lll}-$
 $\text{OC}(=\text{O})\text{O}(\text{CR}_2)_n-^{\lll}$, $^{\lll}-(\text{CR}_2)_n\text{OC}(=\text{O})\text{O}-^{\lll}$, $^{\lll}-\text{C}(=\text{O})\text{N}(\text{R})(\text{CR}_2)_n-^{\lll}$, $^{\lll}-$
 $(\text{CR}_2)_n(\text{R})\text{NC}(=\text{O})-^{\lll}$, $^{\lll}-\text{N}(\text{R})\text{C}(=\text{O})(\text{CR}_2)_n-^{\lll}$, $^{\lll}-(\text{CR}_2)_n\text{C}(=\text{O})\text{N}(\text{R})-^{\lll}$, $^{\lll}-$
 $\text{N}(\text{R})\text{C}(=\text{O})\text{O}(\text{CR}_2)_n-^{\lll}$, $^{\lll}-(\text{CR}_2)_n\text{OC}(=\text{O})\text{N}(\text{R})-^{\lll}$, $^{\lll}-\text{OC}(=\text{O})\text{N}(\text{R})(\text{CR}_2)_n-^{\lll}$, $^{\lll}-$
 $(\text{CR}_2)_n\text{N}(\text{R})\text{C}(=\text{O})\text{O}-^{\lll}$, $^{\lll}-(\text{OCR}_2\text{CR}_2)_n-^{\lll}$, $^{\lll}-(\text{CR}_2\text{CR}_2\text{O})_n-^{\lll}$, $^{\lll}-$
 $(\text{OCR}_2\text{CR}_2)_n(\text{C}=\text{O})-^{\lll}$, $^{\lll}-(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C}=\text{O})-^{\lll}$, $^{\lll}-(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})-^{\lll}$,
 $^{\lll}-(\text{OCR}_2\text{CR}_2)_n(\text{C}=\text{O})(\text{CR}_2)_p-^{\lll}$, $^{\lll}-(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C}=\text{O})(\text{CR}_2)_p-^{\lll}$, $^{\lll}-$
 $(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})(\text{CR}_2)_p-^{\lll}$, $^{\lll}-(\text{OCR}_2\text{CR}_2)_n(\text{C}=\text{O})\text{O}(\text{CR}_2)_p-^{\lll}$, $^{\lll}-$
 $(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C}=\text{O})\text{O}(\text{CR}_2)_p-^{\lll}$, $^{\lll}-(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})\text{O}(\text{CR}_2)_p-^{\lll}$, $^{\lll}-$
 $(\text{OCR}_2\text{CR}_2)_n(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p-^{\lll}$, $^{\lll}-(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p-^{\lll}$, $^{\lll}-$
 $(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p-^{\lll}$, $^{\lll}-\text{OCR}_2(\text{C}=\text{O})-^{\lll}$, $^{\lll}-\text{OCR}_2(\text{C}=\text{O})(\text{CR}_2)_p-$
 $^{\lll}$, $^{\lll}-\text{OCR}_2(\text{C}=\text{O})\text{O}(\text{CR}_2)_p-^{\lll}$, $^{\lll}-\text{OCR}_2(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p-^{\lll}$, $^{\lll}-$
 $\text{C}(=\text{O})\text{O}(\text{CR}_2\text{CR}_2\text{O})_n\text{CR}_2\text{CR}_2-^{\lll}$, $^{\lll}-\text{C}(=\text{O})\text{N}(\text{R})(\text{CR}_2\text{CR}_2\text{O})_n\text{CR}_2\text{CR}_2-^{\lll}$, $^{\lll}-$
 $\text{S}(\text{CR}_2)_n-^{\lll}$, $^{\lll}-(\text{CR}_2)_n\text{S}-^{\lll}$, $^{\lll}-\text{SO}_2(\text{CR}_2)_n-^{\lll}$, $^{\lll}-(\text{CR}_2)_n\text{SO}_2-^{\lll}$, $^{\lll}-$
 $\text{N}(\text{R})\text{SO}_2(\text{CR}_2)_n-^{\lll}$, $^{\lll}-(\text{CR}_2)_n\text{SO}_2\text{N}(\text{R})-^{\lll}$, $^{\lll}-\text{SO}_2\text{N}(\text{R})(\text{CR}_2)_n-^{\lll}$, $^{\lll}-$
 $(\text{CR}_2)_n\text{N}(\text{R})\text{SO}_2-^{\lll}$, $^{\lll}-(\text{CR}_2)_n-\text{Ar}-(\text{CR}_2)_p-^{\lll}$, $^{\lll}-(\text{CR}_2)_p-\text{Ar}-(\text{CR}_2)_n-^{\lll}$, $^{\lll}-$
 $\text{O}(\text{CR}_2)_n-\text{Ar}-(\text{CR}_2)_p-^{\lll}$, and $^{\lll}-(\text{CR}_2)_p-\text{Ar}-(\text{CR}_2)_n\text{O}-^{\lll}$, wherein $^{\lll}$ indicates the
 position of a covalent bond with the moiety:



“<<<” indicates the position of a covalent bond with R^8 ; each instance of n is independently a whole number selected from 1-20; each instance of p is independently an integer selected from 0-20; and R for each instance is independently selected from hydrogen, alkyl, cycloalkyl, and aryl; or two instances of R taken together with the carbons to which they are attached form a 3-6 membered carbocyclic ring; or two instances of R taken together with the atoms to which they are attached form a 5-6 membered heterocyclic ring.

[0246] In certain embodiments of the compound of Formula 3, the linker A comprises 2 or 3. In such instances, the linker can be represented by $-A^1-A^2-A^3-$, wherein each of A^1 , A^2 , and A^3 is independently selected from the linkers recited in the definition of A herein; or A^3 is absent. In instances in which A comprises 2 or 3 linkers, “*” indicates the position of a covalent bond with the moiety:



or “*” indicates the position of a covalent bond with A^1 , A^2 , or A^3 ; and “***” indicates the position of a covalent bond with R^8 ; or “***” indicates the position of a covalent bond with A^1 , A^2 , or A^3 .

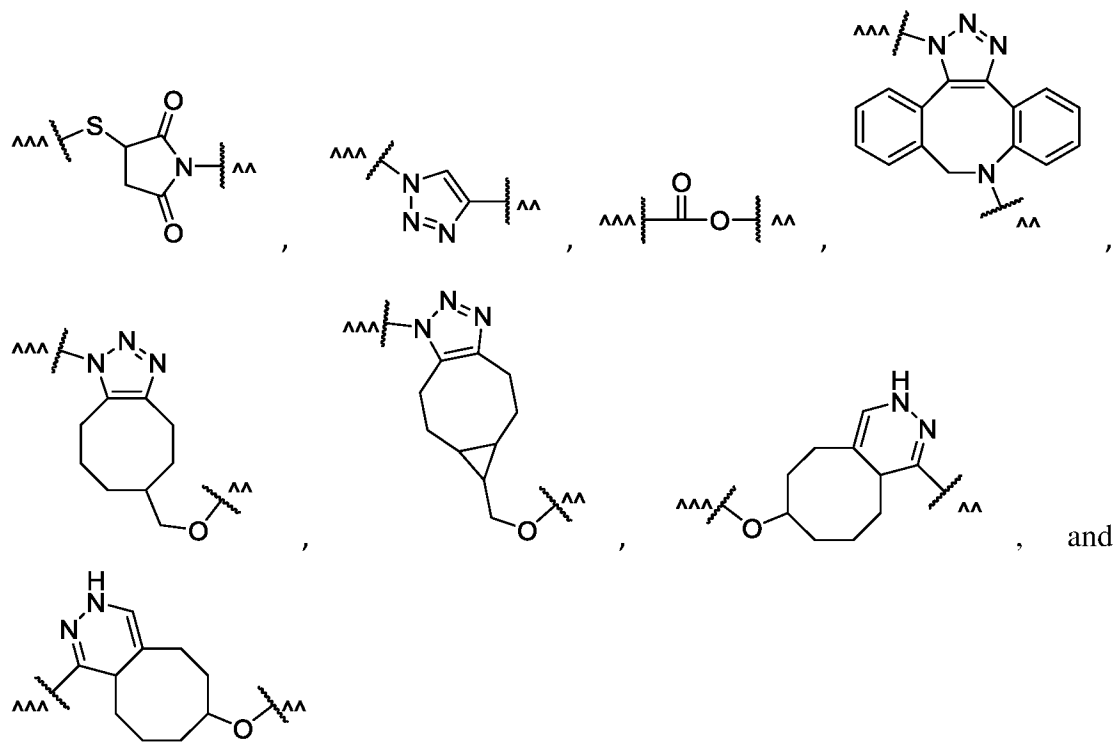
[0247] In certain embodiments of the compound of Formula 3, the linker is at least one moiety selected from the group consisting of $^{\<\<\<}\text{-(CR}_2\text{)}_n\text{-}^{\<\<\<}$, $^{\<\<\<}\text{-O(CR}_2\text{)}_n\text{-}^{\<\<\<}$, $^{\<\<\<}\text{-O(CR}_2\text{)}_n\text{O-}^{\<\<\<}$, $^{\<\<\<}\text{-O(CR}_2\text{)}_n\text{C(=O)-}^{\<\<\<}$, $^{\<\<\<}\text{-O(CR}_2\text{)}_n\text{C(=O)N(R)(CR}_2\text{)}_p\text{-}^{\<\<\<}$, $^{\<\<\<}\text{-O(CR}_2\text{)}_n\text{N(R)C(=O)N(R)(CR}_2\text{)}_p\text{-}^{\<\<\<}$, $^{\<\<\<}\text{-O(CR}_2\text{)}_n\text{C(=O)O(CR}_2\text{)}_p\text{-}^{\<\<\<}$, $^{\<\<\<}\text{-O(CR}_2\text{)}_n\text{OC(=O)N(R)(CR}_2\text{)}_p\text{-}^{\<\<\<}$, $^{\<\<\<}\text{-O(CR}_2\text{)}_n\text{N(R)C(=O)-}^{\<\<\<}$, $^{\<\<\<}\text{-O(CR}_2\text{)}_n\text{N(R)C(=O)(CR}_2\text{)}_p\text{-}^{\<\<\<}$, $^{\<\<\<}\text{-OC(=O)(CR}_2\text{)}_n\text{-}^{\<\<\<}$, $^{\<\<\<}\text{-OC(=O)O(CR}_2\text{)}_n\text{-}^{\<\<\<}$, $^{\<\<\<}\text{-N(R)C(=O)(CR}_2\text{)}_n\text{-}^{\<\<\<}$, $^{\<\<\<}\text{-N(R)C(=O)O(CR}_2\text{)}_n\text{-}^{\<\<\<}$, $^{\<\<\<}\text{-(OCR}_2\text{CR}_2\text{)}_n\text{-}^{\<\<\<}$, $^{\<\<\<}\text{-(OCR}_2\text{CR}_2\text{)}_n\text{(C=O)-}^{\<\<\<}$, $^{\<\<\<}\text{-(OCR}_2\text{CR}_2\text{)}_n\text{O(C=O)-}^{\<\<\<}$, $^{\<\<\<}\text{-(OCR}_2\text{CR}_2\text{)}_n\text{N(R)(C=O)-}^{\<\<\<}$, $^{\<\<\<}\text{-}$

$(\text{OCR}_2\text{CR}_2)_n(\text{C}=\text{O})(\text{CR}_2)_p-^{\lll}$, $^{\lll}-(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C}=\text{O})(\text{CR}_2)_p-^{\lll}$, $^{\lll}-$
 $(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})(\text{CR}_2)_p-^{\lll}$, $^{\lll}-(\text{OCR}_2\text{CR}_2)_n(\text{C}=\text{O})\text{O}(\text{CR}_2)_p-^{\lll}$, $^{\lll}-$
 $(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C}=\text{O})\text{O}(\text{CR}_2)_p-^{\lll}$, $^{\lll}-(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})\text{O}(\text{CR}_2)_p-^{\lll}$, $^{\lll}-$
 $(\text{OCR}_2\text{CR}_2)_n(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p-^{\lll}$, $^{\lll}-(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p-^{\lll}$, $^{\lll}-$
 $(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p-^{\lll}$, $^{\lll}-\text{OCR}_2(\text{C}=\text{O})-^{\lll}$, $^{\lll}-\text{OCR}_2(\text{C}=\text{O})(\text{CR}_2)_p-$
 $^{\lll}$, $^{\lll}-\text{OCR}_2(\text{C}=\text{O})\text{O}(\text{CR}_2)_p-^{\lll}$, $^{\lll}-\text{OCR}_2(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p-^{\lll}$, and $^{\lll}-$
 $\text{N}(\text{R})\text{SO}_2(\text{CR}_2)_n-^{\lll}$.

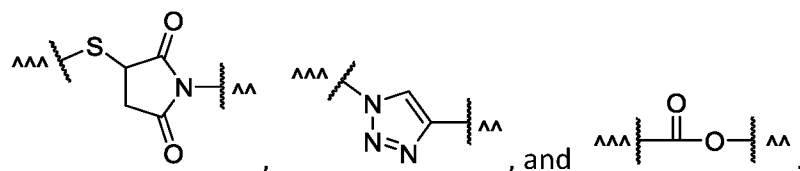
[0248] In certain embodiments of the compound of Formula 3, the linker is at least one moiety selected from the group consisting of $^{\lll}-(\text{CR}_2)_n-^{\lll}$, $^{\lll}-\text{O}(\text{CR}_2)_n-^{\lll}$, $^{\lll}-\text{O}(\text{CR}_2)_n\text{O}-^{\lll}$, $^{\lll}-\text{O}(\text{CR}_2)_n\text{C}(=\text{O})-^{\lll}$, $^{\lll}-\text{O}(\text{CR}_2)_n\text{C}(=\text{O})\text{N}(\text{R})(\text{CR}_2)_p-^{\lll}$, $^{\lll}-\text{O}(\text{CR}_2)_n\text{N}(\text{R})\text{C}(=\text{O})\text{N}(\text{R})(\text{CR}_2)_p-^{\lll}$, $^{\lll}-\text{O}(\text{CR}_2)_n\text{C}(=\text{O})\text{O}(\text{CR}_2)_p-^{\lll}$, $^{\lll}-\text{O}(\text{CR}_2)_n\text{OC}(=\text{O})\text{N}(\text{R})(\text{CR}_2)_p-^{\lll}$, $^{\lll}-\text{O}(\text{CR}_2)_n\text{N}(\text{R})\text{C}(=\text{O})-^{\lll}$, $^{\lll}-\text{O}(\text{CR}_2)_n\text{N}(\text{R})\text{C}(=\text{O})(\text{CR}_2)_p-^{\lll}$, $^{\lll}-\text{OC}(=\text{O})(\text{CR}_2)_n-^{\lll}$, $^{\lll}-\text{OC}(=\text{O})\text{O}(\text{CR}_2)_n-^{\lll}$, $^{\lll}-(\text{OCR}_2\text{CR}_2)_n-^{\lll}$, $^{\lll}-(\text{OCR}_2\text{CR}_2)_n(\text{C}=\text{O})-^{\lll}$, $^{\lll}-(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C}=\text{O})-^{\lll}$, $^{\lll}-(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})-^{\lll}$, $^{\lll}-(\text{OCR}_2\text{CR}_2)_n(\text{C}=\text{O})(\text{CR}_2)_p-^{\lll}$, $^{\lll}-(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C}=\text{O})(\text{CR}_2)_p-^{\lll}$, $^{\lll}-(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})(\text{CR}_2)_p-^{\lll}$, $^{\lll}-(\text{OCR}_2\text{CR}_2)_n(\text{C}=\text{O})\text{O}(\text{CR}_2)_p-^{\lll}$, $^{\lll}-(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C}=\text{O})\text{O}(\text{CR}_2)_p-^{\lll}$, $^{\lll}-(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})\text{O}(\text{CR}_2)_p-^{\lll}$, $^{\lll}-(\text{OCR}_2\text{CR}_2)_n(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p-^{\lll}$, $^{\lll}-(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p-^{\lll}$, $^{\lll}-(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p-^{\lll}$, $^{\lll}-\text{OCR}_2(\text{C}=\text{O})-^{\lll}$, $^{\lll}-\text{OCR}_2(\text{C}=\text{O})(\text{CR}_2)_p-^{\lll}$, $^{\lll}-\text{OCR}_2(\text{C}=\text{O})\text{O}(\text{CR}_2)_p-^{\lll}$, and $^{\lll}-\text{OCR}_2(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p-^{\lll}$.

[0249] In certain embodiments of the compound of Formula 3, the linker is at least one moiety selected from the group consisting of $^{\lll}-(\text{CR}_2)_n-^{\lll}$, $^{\lll}-\text{O}(\text{CR}_2)_n-^{\lll}$, $^{\lll}-\text{O}(\text{CR}_2)_n\text{O}-^{\lll}$, and $^{\lll}-(\text{OCR}_2\text{CR}_2)_n-^{\lll}$.

[0250] In certain embodiments of the compound of Formula 3, R^8 is selected from the group consisting of:

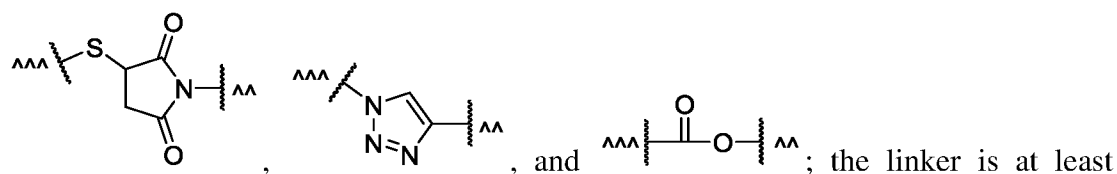


[0251] In certain embodiments of the compound of Formula 3, R⁸ is selected from the group consisting of:



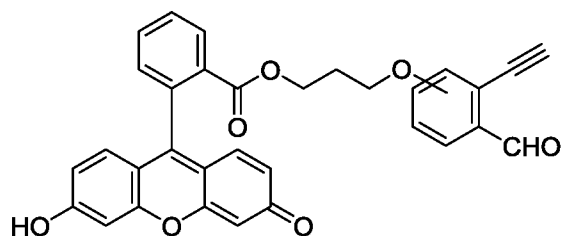
[0252] In certain embodiments of the compound of Formula **3**, R is hydrogen or C₁-C₆ alkyl. In certain embodiments of the compound of Formula **3**, R is hydrogen.

[0253] In certain embodiments of the compound of Formula **3**, the spacer is at least one of 4-amino piperidine, 4-(aminomethyl)piperidine, $^{\text{H}}\text{-N(R)}(\text{CR}_2)_n\text{-}^{\text{H}}$, $^{\text{H}}\text{-(CR}_2)_n\text{-}^{\text{H}}$, $^{\text{H}}\text{-O(CR}_2)_n\text{-}^{\text{H}}$, $^{\text{H}}\text{-(CR}_2)_n\text{O-}^{\text{H}}$, $^{\text{H}}\text{-(CR}_2)_n\text{C(=O)-}^{\text{H}}$, $^{\text{H}}\text{-C(=O)O(CR}_2)_n\text{-}^{\text{H}}$, or $^{\text{H}}\text{-(CR}_2)_n\text{OC(=O)-}^{\text{H}}$; or the spacer is absent; R^8 is selected from the group consisting of:

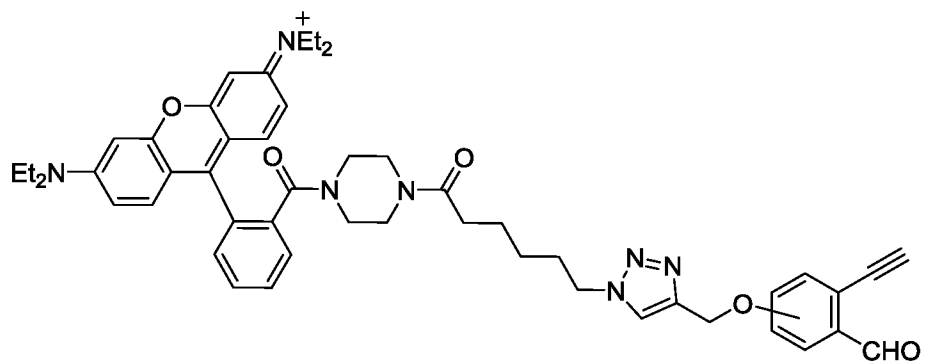


one moiety selected from the group consisting of $^{<<-}-(CR_2)_n-^{<<<}$, $^{<<-}O(CR_2)_n-^{<<<}$, $^{<<-}O(CR_2)_nO-^{<<<}$, and $^{<<-}(OCR_2CR_2)_n-^{<<<}$; R is hydrogen; and n for each instance is independently selected from 1, 2, 3, or 4.

[0254] In certain embodiments of the compound of Formula 3, the compound is selected from the group consisting of:



and



[0255] Conjugate salts include the conventional nontoxic salts or quaternary ammonium salts of the compounds, e.g., from nontoxic organic or inorganic acids. For example, such conventional nontoxic salts include those derived from inorganic acids such as hydrochloride, hydrobromic, sulfuric, sulfamic, phosphoric, nitric, and the like; and the salts prepared from organic acids such as acetic, propionic, succinic, glycolic, stearic, lactic, malic, tartaric, citric, ascorbic, palmitic, maleic, hydroxymaleic, phenylacetic, glutamic, benzoic, salicylic, sulfanilic, 2-acetoxybenzoic, fumaric, toluenesulfonic, methanesulfonic, ethane disulfonic, oxalic, isothionic, and the like.

[0256] In other cases, the compounds described herein may contain one or more

acidic functional groups and, thus, are capable of forming conjugate salts with pharmaceutically-acceptable bases. The term conjugate salts in these instances refer to the relatively non-toxic, inorganic and organic base addition salts of compounds of the present invention. These salts can likewise be prepared in situ in the administration vehicle or the dosage form manufacturing process, or by separately reacting the purified compound in its free acid form with a suitable base, such as the hydroxide, carbonate or bicarbonate of a pharmaceutically-acceptable metal cation, with ammonia, or with a pharmaceutically-acceptable organic primary, secondary or tertiary amine. Representative alkali or alkaline earth salts include the lithium, sodium, potassium, calcium, magnesium, and aluminum salts and the like. Representative organic amines useful for the formation of base addition salts include ethylamine, diethylamine, ethylenediamine, ethanolamine, diethanolamine, piperazine and the like.

[0257] The compounds described herein can comprise a cleavable linker and/or a cleavable spacer. A cleavable linker or spacer, is a linker or spacer that contains one or more covalent bonds and/or functional groups that are susceptible to being cleaved under physiological conditions, e.g., under the action of an enzyme, water, acidic conditions, basic conditions, oxygen, reducing agents, oxidizing agents, or the like. Examples of cleavable functional groups include, but are not limited to carbinoamine, disulfide, carbamide, amination, acetal, ester, carbonate, carbamate, phosphate, imine, oxime, and hydrazone groups. In certain embodiments, the linker and/or spacer comprises or further comprises at least one cleavable functional group selected from the group consisting of carbinoamine, disulfide, carbamide, amination, N,O amination, acetal, ester, carbonate, carbamate, phosphate, imine, oxime, and hydrazone. In certain embodiments, the linker and/or spacer comprises or further comprises at least

one cleavable functional group selected from the group consisting of carbinoamine, disulfide, aminor, N,O aminor, and acetal.

[0258] Advantageously, the methods provided herein can be conducted under mild conditions under which the risk of polypeptide denaturing, degradation, and/or isomerization is minimized.

[0259] The step of contacting a polypeptide comprising a N-terminal amine and a compound of Formula **1** or **3** can be conducted at a temperature of about 10 to about 60 °C. In certain embodiments, the step of contacting a polypeptide comprising a N-terminal amine and a compound of Formula **1** or **3** can be conducted at a temperature of about 10 to about 60, 20 to about 60, 20 to about 40, or 20 to about 30 °C. In certain embodiments, the step of contacting a polypeptide comprising an N-terminal amine and a compound of Formula **1** or **3** can be conducted at about 37 °C.

[0260] The step of contacting a polypeptide comprising an N-terminal amine and a compound of Formula **1** or **3** can be conducted at a pH of about 5 to about 10. In certain embodiments, the step of contacting a polypeptide comprising an N-terminal amine and a compound of Formula **1** or **3** can be conducted at a pH of about 5 to about 9; about 5 to about 8; about 5 to about 7; about 6 to about 7; or about 6 to about 6.5. In certain embodiments, the step of contacting a polypeptide comprising an N-terminal amine and a compound of Formula **1** or **3** can be conducted at a pH of about 6.3.

[0261] The step of contacting a polypeptide comprising an N-terminal amine and a compound of Formula **1** or **3** can be conducted in dimethylsulfoxide (DMSO), *N,N*-dimethylformamide (DMF), *N,N*-dimethylacetamide (DMA), methanol, ethanol, ethylene glycol, isopropanol, tetrahydrofuran, tetrahydropyran, water, and

combinations thereof. In certain embodiments, the step of contacting a polypeptide comprising an N-terminal amine and a compound of Formula **1** or **3** can be conducted in a mixture of DMSO and water. The ratio of DMSO to water can be about 1:1 to about 12:1. In certain embodiments, the ratio of DMSO to water is about 9:1.

[0262] The compound of Formula **1** or **3** can present in any amount relative to the polypeptide. However, to ensure higher yields of the desired product and/or higher rates of reaction, an excess of the compound of Formula **1** or **3** can be used. 1-500 molar equivalents of the compound of Formula **1** or **3** can be used relative to the polypeptide. In certain embodiments, 1-5, 5-10, 5-15, 5-20, or 5-25 equivalents of the compound of Formula **1** or **3** can be used relative to the polypeptide. In certain embodiments, about 10 equivalents of the compound of Formula **1** or **3** can be used relative to the polypeptide.

[0263] The methods for N-terminal amine functionalization are capable of selectively functionalizing the N-terminal amine of a polypeptide with a selectivity in favor of the N-terminal amine over amines present in side chains of amino acids in the polypeptide (such as lysine) of about 10:1, about 15:1, about 20:1, about 25:1, about 30:1, about 35:1, about 40:1, about 45:1, about 50:1, about 55:1, about 60:1, about 65:1, about 70:1, about 75:1, about 80:1, about 85:1, about 90:1, about 95:1, about 99:1 or greater. In certain embodiments, the methods for N-terminal amine functionalization are capable of selectively functionalizing the N-terminal amine of a polypeptide with a selectivity in favor of the N-terminal amine over amines present in side chains of the polypeptide of between about 10:1 and 99:1, 20:1 and 99:1, 30:1 and 99:1, 40:1 and 99:1, 50:1 and 99:1, 50:1 and 90:1, and 50:1 and 80:1. The N-terminal selectivity results are shown in Table 1.

[0264] Suitable reaction conditions for the reaction of the active agent comprising a

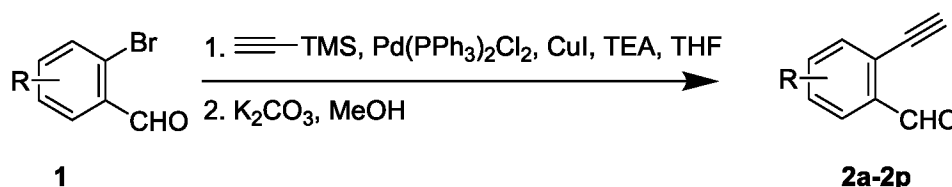
functional group capable of reacting with the functionalized polypeptide of Formula 2 depends on the chemical reactivity of the R⁴ and the selection of which is well within the skill of a person of ordinary skill in the art.

[0265] **Examples**

[0266] The following examples are provided by way of describing specific embodiments of this invention without intending to limit the scope of this invention in any way.

[0267] **Example 1**

[0268] **General procedure for synthesis of 2-ethynylbenzaldehydes**

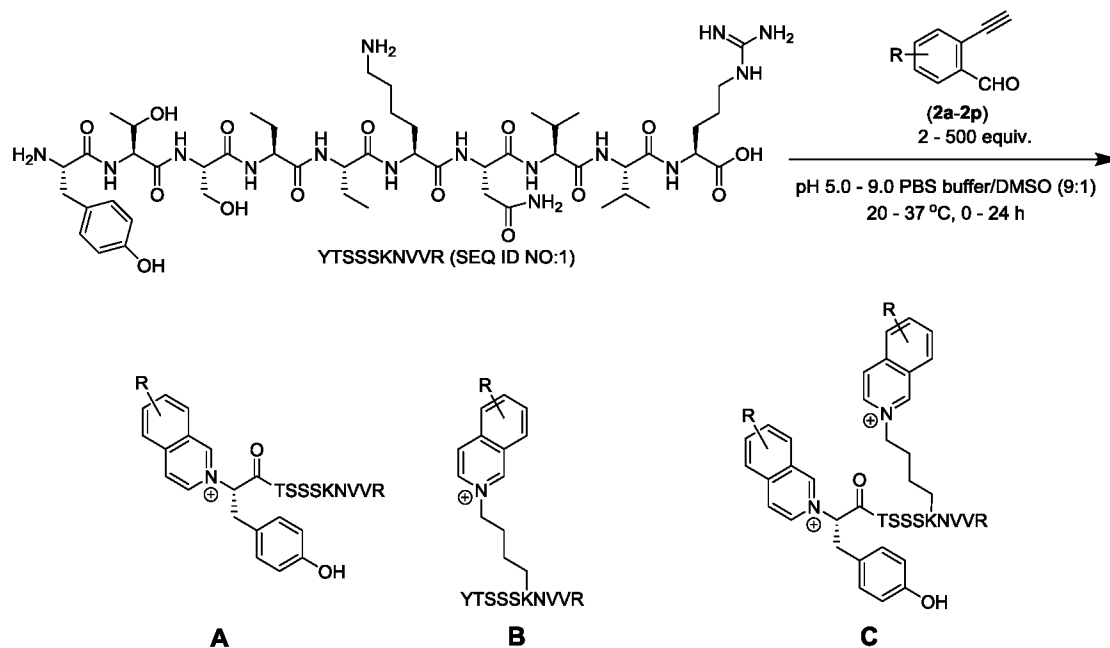


[0269] To a schlenk flask with bis(triphenylphosphine)palladium(II) dichloride (5 mol%) and copper(I) iodide (2.5 mol%) under nitrogen atmosphere, a solution of compound **1** (5 mmol) in tetrahydrofuran (50 mL) and triethylamine (0.4 M) was added at room temperature, followed by trimethylsilylacetylene (6 mmol, 1.2 equiv.), the mixture was then heated at 80-120 °C for overnight. After filtration of the mixture with celite, the filtrate was concentrated under reduced pressure and the residue was purified by column chromatography over silica gel with *n*-hexane/ethyl acetate (50:1) as the eluent to give reaction intermediate. The corresponding reaction intermediate was reacted with potassium carbonate (0.50 g, 3.64 mmol) in methanol (30 mL) for 30 min at room temperature, and the solvent was removed under the reduced pressure. The residue was extracted with dichloromethane and washed with saturated sodium carbonate, and dried over magnesium sulfate. The filtrate was concentrated under reduced pressure and purified by column chromatography over silica gel with

n-hexane/ethyl acetate (50:1) to give 2-ethynylbenzaldehydes **2a-2p**.

[0270] **Example 2**

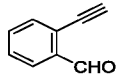
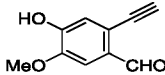
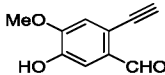
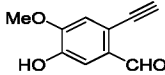
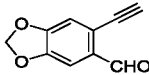
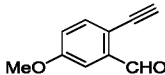
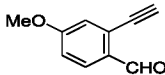
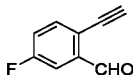
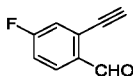
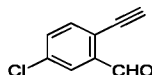
[0271] **N-Terminal modification of polypeptides using 2-ethynylbenzaldehydes **2a-2p****

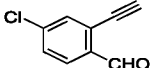
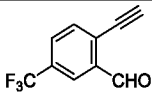
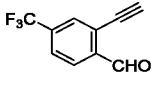
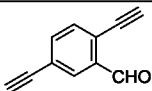
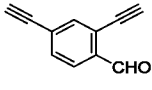
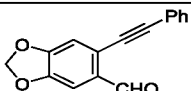


[0272] In a 1.0 mL eppendorf tube, a stock solution of polypeptides in water (1-10 mM, 10uL), **Compound 2a-2p** (2-500 equivalents, 10 uL of a 2-500 mM stock solution in dimethyl sulfoxide), and phosphate-buffered saline (pH 5.0–9.0, 80 μL) are mixed. The reaction mixture is kept at 20-37 °C for 0-24 h. 10 μL of the mixture was drawn, diluted with 10 μL water and subjected to LC/MS-MS analysis to analyze the product distribution of **A:B:C**. Conversion and selectivity data is presented in Table 1.

[0273] Table 1. Results of N-Terminal modification of model peptide YTSSSKNVVR (SEQ ID NO: 1) using 2-ethynylbenzaldehyde and analogs **2a-2p**.

Entry	Reagent	Conversion (%)			N-terminal selectivity of mono-modified
		Mono-	Di-modified	Total	

modified					peptide
					YTSSSKNVVR
					(SEQ ID NO: 1)
1		64	8	(72)	21:1
	2a				
2		73	13	(86)	>99:1
	2b				
3		28	0	(28)	20:1
	2c				
4		60	7	(67)	24:1
	2d				
5		40	1	(41)	50:1
	2e				
6		63	5	(68)	92:1
	2f				
7		65	6	(71)	42:1
	2g				
8		64	4	(68)	81:1
	2h				
9		71	7	(78)	62:1
	2i				
10		47	2	(49)	38:1

2j					
11		55	3	(58)	35:1
2k					
12		67	5	(72)	20:1
2l					
13		62	3	(65)	14:1
2m					
14		54	3	(41)	25:1
2n					
15		59	5	(64)	19:1
2o					
16		0	0	(0)	-
2p					

[0274] **Example 3**

[0275] **Preparation of benzylmercaptan-linked**

2-ethynylbenzaldehydes

[0276] To a 10 mL round bottom flask with maleimide-containing 2-ethynylbenzaldehyde (3-10 mg), a solution of benzylmercaptan (1.2 equivalent) in 500 uL of dimethyl sulfoxide or dimethylformamide was added at room temperature. The mixture was then stirred at room temperature for 5-120 min. Conversion of the reaction was monitored by analytical HPLC. The resulting product, benzylmercaptan-linked 2-ethynylbenzaldehyde, was purified by an isocratic elution of acetonitrile (or methanol)-water in semi-preparative HPLC equipped with a

fraction collector. Excess solvent in the collected fractions were then removed by lyophilisation to give a benzylmercaptan-linked 2-ethynylbenzaldehyde.

[0277] **Example 4**

[0278] **N-Terminal modification of polypeptides using benzylmercaptan-linked 2-ethynylbenzaldehyde as N-terminal selective linker**

[0279] In a 1.0 mL eppendorf tube, a stock solution of polypeptides in water (1-10 mM, 10uL), benzylmercaptan-linked 2-ethynylbenzaldehyde (2-500 equivalents, 10 uL of a 2-500 mM stock solution of benzylmercaptan-linked 2-ethynylbenzaldehyde (in dimethyl sulfoxide), and phosphate-buffered saline (pH 5.0–9.0, 80 μ L) are mixed. The reaction mixture is kept at 20-37 °C for 0-24 h. The conversion of the peptide is subjected to LC/MS and LC/MS-MS analysis to analyze the product distribution.

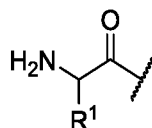
[0280] **Example 5**

[0281] **Stability of 2,5-diethynylbenzaldehyde (2n)-linked hemoglobin**

[0282] In a 1.0 mL eppendorf tube, a stock solution of hemoglobin in water (1-10 mM, 10uL), 2,5-diethynylbenzaldehyde (**2n**) (2-500 equivalents, 10 μ L of a 2-500 mM stock solution of **2n** in dimethyl sulfoxide), and phosphate-buffered saline (pH 5.0 – 9.0, 80 μ L) are mixed. The reaction mixture is kept at 20 - 37 °C for 0-24 h. After the reaction, the modified protein was purified and stored at 4 °C. The conversion and stability of the **2n**-modified protein is subjected to LC/MS analysis on a weekly basis over a period of 8 weeks. The results are shown in FIG. 16. It indicates that the **2n**-modified protein is still stable after 8 weeks.

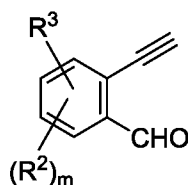
What is claimed is:

1. A method for preparing a functionalized polypeptide comprising the step of: contacting a polypeptide comprising an N-terminal amino acid represented by the structure shown below:



or a conjugate salt or zwitterion thereof, wherein

R^1 is the side chain of the N-terminal amino acid of the polypeptide and the N-terminal amino acid is a natural amino acid or a non-natural amino acid; with a compound of Formula **1**:



1

or a conjugate salt thereof, wherein:

m is 1, 2, or 3; and

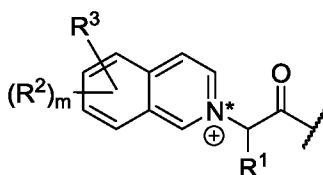
each R^2 is independently for each instance selected from the group consisting of hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl, aralkyl, heterocycloalkyl, heteraryl, halide, cyano, nitro, hydroxyl, mercaptan, $-OR^5$, $-SR^5$, $-(C=O)OR^5$, $-O(C=O)R^5$, $-N(R^5)_2$, $-(C=O)N(R^5)_2$, $-N(R^5)(C=O)R^5$, $-N(R^5)(C=O)N(R^5)_2$, $-SO_2R^5$, $-N(R^5)SO_2R^5$, and $-SO_2N(R^5)_2$;

R^3 is $-A-R^4$, wherein A is a linker or absent;

R^4 is selected from the group consisting of hydroxyl, thiol, aldehyde, ketone, maleimide, *N*-hydroxysuccinimide ester, imidoester, haloacetyl, pyridyl disulfide, alkyne, azide, nitron, nitrile oxide, norbornene, oxanorbornadiene, tetrazine, cyclooctene, isocyanide, quadricyclane, and bis(dithiobenzil)nickel(II); and

R^5 for each instance is independently selected from the group consisting of hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl, aralkyl, heterocycloalkyl, and heteraryl; or two instances of R^5 taken together with the atoms to which they are attached form a 5-6 membered heterocycloalkyl;

thereby forming the functionalized polypeptide, wherein the functionalized polypeptide is represented by the structure of Formula 2:

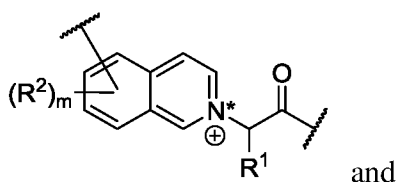


2

wherein N^* is the N-terminal nitrogen of the functionalized polypeptide.

- The method of claim 1, wherein m is 1 and R^2 is hydrogen, $-OR^5$, or halide.
- The method of claim 1, wherein the linker is selected from the group consisting of $^{**}-(CR_2)_n-^{***}$, $^{**}-O(CR_2)_n-^{***}$, $^{**}-(CR_2)_nO-^{***}$, $^{**}-O(CR_2)_nO-^{***}$, $^{**}-(CR_2)_nC(=O)-^{***}$, $^{**}-C(=O)(CR_2)_n-^{***}$, $^{**}-C(=O)O(CR_2)_n-^{***}$, $^{**}-(CR_2)_nOC(=O)-^{***}$, $^{**}-O(CR_2)_nC(=O)-^{***}$, $^{**}-O(CR_2)_nC(=O)N(R)(CR_2)_p-^{***}$, $^{**}-O(CR_2)_nN(R)C(=O)N(R)(CR_2)_p-^{***}$, $^{**}-O(CR_2)_nC(=O)O(CR_2)_p-^{***}$, $^{**}-O(CR_2)_nOC(=O)N(R)(CR_2)_p-^{***}$, $^{**}-C(=O)(CR_2)_nO-^{***}$, $^{**}-O(CR_2)_nN(R)C(=O)-^{***}$, $^{**}-O(CR_2)_nN(R)C(=O)(CR_2)_p-^{***}$, $^{**}-C(=O)N(R)(CR_2)_nO-^{***}$, $^{**}-OC(=O)(CR_2)_n-^{***}$, $^{**}-(CR_2)_nC(=O)O-^{***}$, $^{**}-OC(=O)O(CR_2)_n-^{***}$, $^{**}-(CR_2)_nOC(=O)O-^{***}$, $^{**}-$

$C(=O)N(R)(CR_2)_n^{***}$, $^{**}-(CR_2)_n(R)NC(=O)^{***}$, $^{**}-N(R)C(=O)(CR_2)_n^{***}$, $^{**}-(CR_2)_nC(=O)N(R)^{***}$, $^{**}-N(R)C(=O)O(CR_2)_n^{***}$, $^{**}-(CR_2)_nOC(=O)N(R)^{***}$, $^{**}-OC(=O)N(R)(CR_2)_n^{***}$, $^{**}-(CR_2)_nN(R)C(=O)O^{***}$, $^{**}-(OCR_2CR_2)_n^{***}$, $^{**}-(CR_2CR_2O)_n^{***}$, $^{**}-(OCR_2CR_2)_n(C=O)^{***}$, $^{**}-(OCR_2CR_2)_nO(C=O)^{***}$, $^{**}-(OCR_2CR_2)_nN(R)(C=O)^{***}$, $^{**}-(OCR_2CR_2)_n(C=O)(CR_2)_p^{***}$, $^{**}-(OCR_2CR_2)_nO(C=O)(CR_2)_p^{***}$, $^{**}-(OCR_2CR_2)_nN(R)(C=O)(CR_2)_p^{***}$, $^{**}-(OCR_2CR_2)_n(C=O)O(CR_2)_p^{***}$, $^{**}-(OCR_2CR_2)_nO(C=O)O(CR_2)_p^{***}$, $^{**}-(OCR_2CR_2)_nN(R)(C=O)O(CR_2)_p^{***}$, $^{**}-(OCR_2CR_2)_n(C=O)N(R)(CR_2)_p^{***}$, $^{**}-(OCR_2CR_2)_nO(C=O)N(R)(CR_2)_p^{***}$, $^{**}-(OCR_2CR_2)_nN(R)(C=O)N(R)(CR_2)_p^{***}$, $^{**}-OCR_2(C=O)^{***}$, $^{**}-OCR_2(C=O)(CR_2)_p^{***}$, $^{***}^{**}-OCR_2(C=O)O(CR_2)_p^{***}$, $^{**}-OCR_2(C=O)N(R)(CR_2)_p^{***}$, $^{**}-C(=O)O(CR_2CR_2O)_nCR_2CR_2^{***}$, $^{**}-C(=O)N(R)(CR_2CR_2O)_nCR_2CR_2^{***}$, $^{**}-S(CR_2)_n^{***}$, $^{**}-(CR_2)_nS^{***}$, $^{**}-(CR_2)_nSS(CR_2)_p^{***}$, $^{**}-SO_2(CR_2)_n^{***}$, $^{**}-(CR_2)_nSO_2^{***}$, $^{**}-N(R)SO_2(CR_2)_n^{***}$, $^{**}-(CR_2)_nSO_2N(R)^{***}$, $^{**}-SO_2N(R)(CR_2)_n^{***}$, $^{**}-(CR_2)_nN(R)SO_2^{***}$, $^{**}-(CR_2)_n-Ar-(CR_2)_p^{***}$, $^{**}-(CR_2)_p-Ar-(CR_2)_n^{***}$, $^{**}-O(CR_2)_n-Ar-(CR_2)_p^{***}$, and $^{**}-(CR_2)_p-Ar-(CR_2)_nO^{***}$, wherein ** indicates the position of a covalent bond with the moiety:



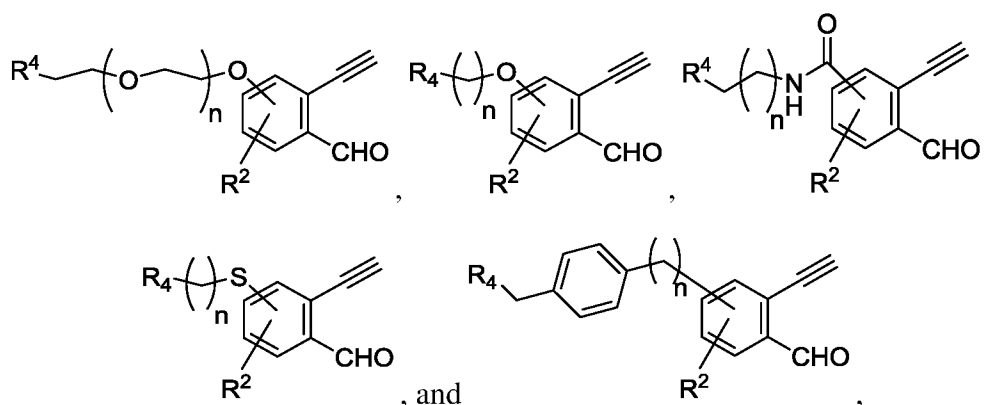
*** indicates the position of a covalent bond with R^4 ; each instance of n is independently a whole number selected from 1-20; each instance of p is independently an integer selected from 0-20; and R for each instance is independently selected from hydrogen, alkyl, cycloalkyl, and aryl; or two instances of R taken together with the carbons to which they are attached form a 3-6 membered

carbocyclic ring; or two instances of R taken together with the atoms to which they are attached form a 5-6 membered heterocyclic ring.

4. The method of claim 1, wherein n is a whole number selected from 1-10 and p is a whole number selected from 1-4.

5. The method of claim 3, wherein R⁴ is maleimide, azide, alkyne, hydroxyl, norbornene, or cyclooctene.

6. The method of claim 1, wherein the compound of Formula 3 is selected from the group consisting of:



wherein

n is a whole number selected from 1-4;

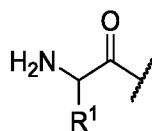
R² is hydrogen, alkyl, or -OR⁵;

R⁴ is maleimide or alkyne; and

R⁵ is alkyl.

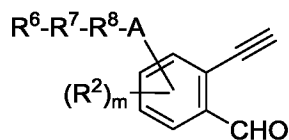
7. The method of claim 1, wherein the functionalized polypeptide comprises at least one lysine residue and less than 5% of the side chains of the at least one lysine residue undergo a reaction with the compound of Formula 3.

8. The method of claim 1, wherein the polypeptide is a protein, enzyme, antibody, glycoprotein, or lipoprotein.
9. The method of claim 8, wherein the protein is a hemoglobin protein.
10. The method of claim 9, wherein the hemoglobin protein is a hemoglobin tetramer, hemoglobin dimer, or hemoglobin monomer.
11. The method of claim 10, wherein the hemoglobin tetramer is a cross-linked hemoglobin tetramer.
12. The method of claim 1 further comprising the step of contacting the functionalized polypeptide with an active agent comprising a mercaptan or conjugate salt thereof.
13. The method of claim 12, wherein the active agent is a chemotherapeutic agent, a radiotherapeutic agent, a fluorescent dye, or affinity tag.
14. A method for preparing a functionalized polypeptide comprising the step of: contacting a polypeptide comprising an N-terminal amino acid represented by the structure shown below:



or a conjugate salt or zwitterion thereof, wherein

R^1 is the side chain of the N-terminal amino acid of the polypeptide and the N-terminal amino acid is a natural amino acid or a non-natural amino acid; with a compound of Formula 3:



3

or a conjugate salt thereof, wherein

m is 1, 2, or 3;

A is a linker or is absent;

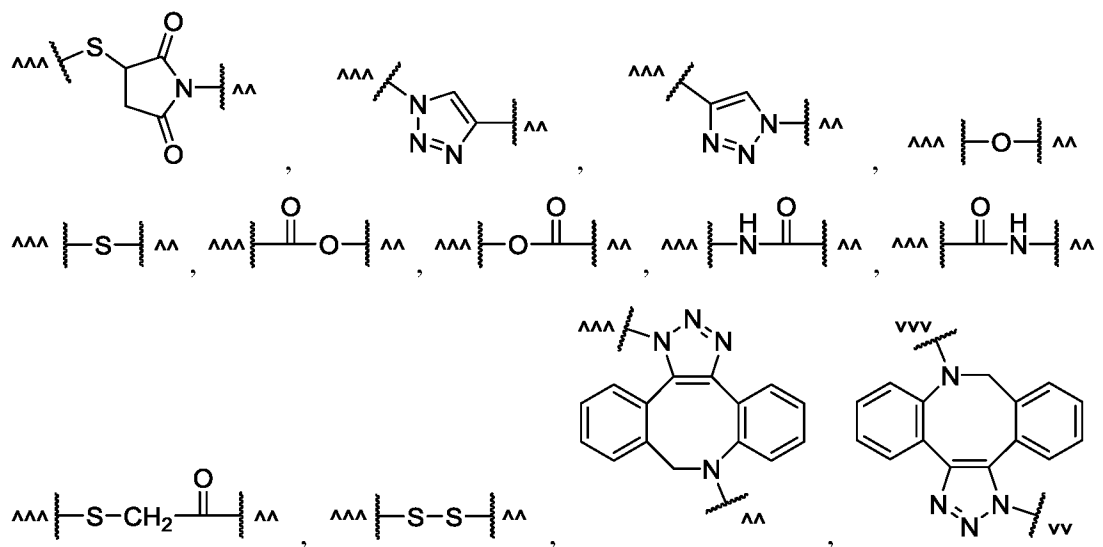
R^2 is independently for each instance selected from the group consisting of hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl, aralkyl, heterocycloalkyl, heteraryl, halide, cyano, nitro, hydroxyl, mercaptan, $-OR^5$, $-SR^5$, $-(C=O)OR^5$, $-O(C=O)R^5$, $-N(R^5)_2$, $-(C=O)N(R^5)_2$, $-N(R^5)(C=O)R^5$, $-N(R^5)(C=O)N(R^5)_2$, $-SO_2R^5$, $-N(R^5)SO_2R^5$, and $-SO_2N(R^5)_2$;

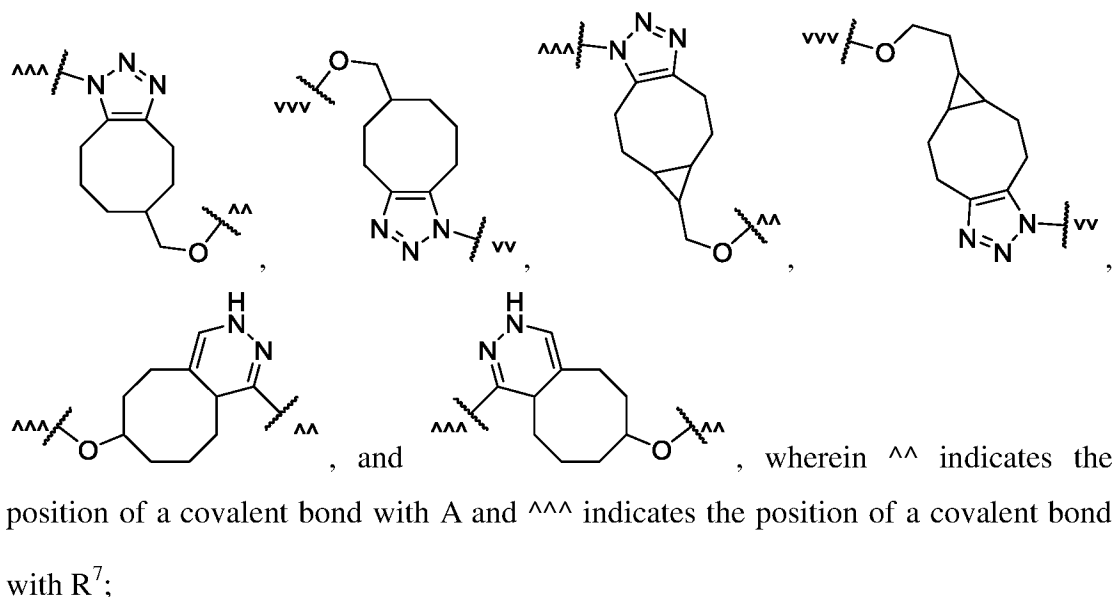
R^5 for each instance is independently selected from the group consisting of hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl, aralkyl, heterocycloalkyl, and heteraryl; or two instances of R^5 taken together with the atoms to which they are attached form a 5-6 membered heterocycloalkyl;

R^6 is an active agent;

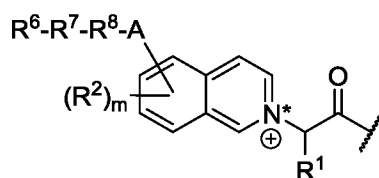
R^7 is a spacer or absent; and

R^8 is selected from the group consisting of:





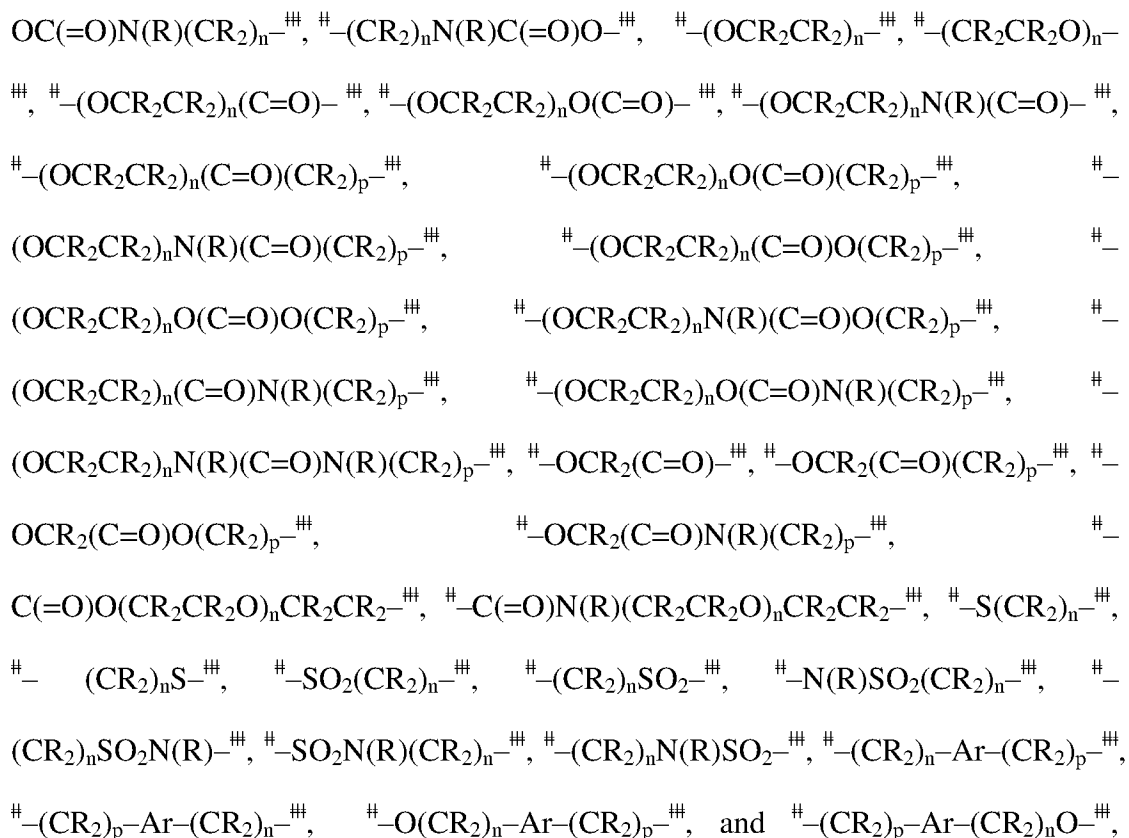
thereby forming the functionalized polypeptide, wherein the functionalized polypeptide is represented by the structure of Formula 4:



4

wherein N* is the N-terminal nitrogen of the functionalized polypeptide.

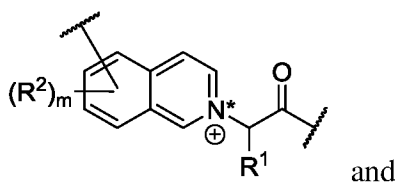
[illegible]



wherein [#] indicates the position of a covalent bond with R⁶ and ^{##} indicates the position of a covalent bond with R⁸; each instance of n is independently a whole number selected from 1-20; p is an integer selected from 0-20; and R for each instance is independently selected from hydrogen, alkyl, cycloalkyl, and aryl; or two instances of R taken together with the carbons to which they are attached form a 3-6 membered carbocyclic ring; or two instances of R taken together with the atoms to which they are attached form a 5-6 membered heterocyclic ring.

16. The method of claim 14, wherein the linker is selected from the group consisting of $^{\leftarrow\leftarrow}\text{--Ar--}^{\leftarrow\leftarrow\leftarrow}$, $^{\leftarrow\leftarrow}\text{--(CR}_2\text{)}_n\text{--}^{\leftarrow\leftarrow\leftarrow}$, $^{\leftarrow\leftarrow}\text{--O(CR}_2\text{)}_n\text{--}^{\leftarrow\leftarrow\leftarrow}$, $^{\leftarrow\leftarrow}\text{--(CR}_2\text{)}_n\text{O--}^{\leftarrow\leftarrow\leftarrow}$, $^{\leftarrow\leftarrow}\text{--O(CR}_2\text{)}_n\text{O--}^{\leftarrow\leftarrow\leftarrow}$, $^{\leftarrow\leftarrow}\text{--(CR}_2\text{)}_n\text{C(=O)--}^{\leftarrow\leftarrow\leftarrow}$, $^{\leftarrow\leftarrow}\text{--C(=O)(CR}_2\text{)}_n\text{--}^{\leftarrow\leftarrow\leftarrow}$, $^{\leftarrow\leftarrow}\text{--C(=O)O(CR}_2\text{)}_n\text{--}^{\leftarrow\leftarrow\leftarrow}$, $^{\leftarrow\leftarrow}\text{--(CR}_2\text{)}_n\text{OC(=O)--}^{\leftarrow\leftarrow\leftarrow}$, $^{\leftarrow\leftarrow}\text{--O(CR}_2\text{)}_n\text{C(=O)--}^{\leftarrow\leftarrow\leftarrow}$, $^{\leftarrow\leftarrow}\text{--O(CR}_2\text{)}_n\text{C(=O)N(R)(CR}_2\text{)}_p\text{--}^{\leftarrow\leftarrow\leftarrow}$, $^{\leftarrow\leftarrow}\text{--O(CR}_2\text{)}_n\text{N(R)C(=O)N(R)(CR}_2\text{)}_p\text{--}^{\leftarrow\leftarrow\leftarrow}$, $^{\leftarrow\leftarrow}\text{--O(CR}_2\text{)}_n\text{C(=O)O(CR}_2\text{)}_p\text{--}^{\leftarrow\leftarrow\leftarrow}$, $^{\leftarrow\leftarrow}\text{--O(CR}_2\text{)}_n\text{OC(=O)N(R)(CR}_2\text{)}_p\text{--}^{\leftarrow\leftarrow\leftarrow}$, $^{\leftarrow\leftarrow}\text{--C(=O)(CR}_2\text{)}_n\text{O--}^{\leftarrow\leftarrow\leftarrow}$, $^{\leftarrow\leftarrow}\text{--O(CR}_2\text{)}_n\text{N(R)C(=O)--}^{\leftarrow\leftarrow\leftarrow}$, $^{\leftarrow\leftarrow}\text{--}$

$O(CR_2)_nN(R)C(=O)(CR_2)_p-^{<<<}$, $^{<<-}C(=O)N(R)(CR_2)_nO-^{<<<}$, $^{<<-}OC(=O)(CR_2)_n-^{<<<}$,
 $^{<<-}(CR_2)_nC(=O)O-^{<<<}$, $^{<<-}OC(=O)O(CR_2)_n-^{<<<}$, $^{<<-}(CR_2)_nOC(=O)O-^{<<<}$, $^{<<-}$
 $C(=O)N(R)(CR_2)_n-^{<<<}$, $^{<<-}(CR_2)_n(R)NC(=O)-^{<<<}$, $^{<<-}N(R)C(=O)(CR_2)_n-^{<<<}$, $^{<<-}$
 $(CR_2)_nC(=O)N(R)-^{<<<}$, $^{<<-}N(R)C(=O)O(CR_2)_n-^{<<<}$, $^{<<-}(CR_2)_nOC(=O)N(R)-^{<<<}$, $^{<<-}$
 $OC(=O)N(R)(CR_2)_n-^{<<<}$, $^{<<-}(CR_2)_nN(R)C(=O)O-^{<<<}$, $^{<<-}(OCR_2CR_2)_n-^{<<<}$, $^{<<-}$
 $(CR_2CR_2O)_n-^{<<<}$, $^{<<-}(OCR_2CR_2)_n(C=O)-^{<<<}$, $^{<<-}(OCR_2CR_2)_nO(C=O)-^{<<<}$, $^{<<-}$
 $(OCR_2CR_2)_nN(R)(C=O)-^{<<<}$, $^{<<-}(OCR_2CR_2)_n(C=O)(CR_2)_p-^{<<<}$, $^{<<-}$
 $(OCR_2CR_2)_nO(C=O)(CR_2)_p-^{<<<}$, $^{<<-}(OCR_2CR_2)_nN(R)(C=O)(CR_2)_p-^{<<<}$, $^{<<-}$
 $(OCR_2CR_2)_n(C=O)O(CR_2)_p-^{<<<}$, $^{<<-}(OCR_2CR_2)_nO(C=O)O(CR_2)_p-^{<<<}$, $^{<<-}$
 $(OCR_2CR_2)_nN(R)(C=O)O(CR_2)_p-^{<<<}$, $^{<<-}(OCR_2CR_2)_n(C=O)N(R)(CR_2)_p-^{<<<}$, $^{<<-}$
 $(OCR_2CR_2)_nO(C=O)N(R)(CR_2)_p-^{<<<}$, $^{<<-}(OCR_2CR_2)_nN(R)(C=O)N(R)(CR_2)_p-^{<<<}$, $^{<<-}$
 $OCR_2(C=O)-^{<<<}$, $^{<<-}OCR_2(C=O)(CR_2)_p-^{<<<}$, $^{<<-}OCR_2(C=O)O(CR_2)_p-^{<<<}$, $^{<<-}$
 $OCR_2(C=O)N(R)(CR_2)_p-^{<<<}$, $^{<<-}C(=O)O(CR_2CR_2O)_nCR_2CR_2-^{<<<}$, $^{<<-}$
 $C(=O)N(R)(CR_2CR_2O)_nCR_2CR_2-^{<<<}$, $^{<<-}S(CR_2)_n-^{<<<}$, $^{<<-}(CR_2)_nS-^{<<<}$, $^{<<-}SO_2(CR_2)_n-^{<<<}$,
 $^{<<-}(CR_2)_nSO_2-^{<<<}$, $^{<<-}N(R)SO_2(CR_2)_n-^{<<<}$, $^{<<-}(CR_2)_nSO_2N(R)-^{<<<}$, $^{<<-}$
 $SO_2N(R)(CR_2)_n-^{<<<}$, $^{<<-}(CR_2)_nN(R)SO_2-^{<<<}$, $^{<<-}(CR_2)_n-Ar-(CR_2)_p-^{<<<}$, $^{<<-}(CR_2)_p-$
 $Ar-(CR_2)_n-^{<<<}$, $^{<<-}O(CR_2)_n-Ar-(CR_2)_p-^{<<<}$, and $^{<<-}(CR_2)_p-Ar-(CR_2)_nO-^{<<<}$,
 wherein $^{<<}$ indicates the position of a covalent bond with the moiety:

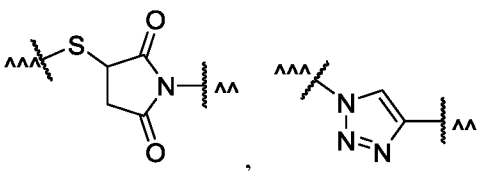


$^{<<<}$ indicates the position of a covalent bond with R^8 ; each instance of n is independently a whole number selected from 1-20; p is an integer selected from 0-20; and R for each instance is independently selected from hydrogen, alkyl, cycloalkyl, and aryl; or two instances of R taken together with the carbons to which they are

attached form a 3-6 membered carbocyclic ring; or two instances of R taken together with the atoms to which they are attached form a 5-6 membered heterocyclic ring.

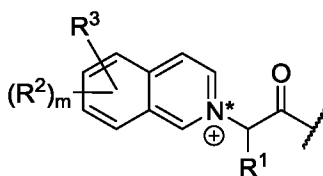
17. The method of claim 14, wherein R^2 is hydrogen.

18. The method of claim 14, wherein the linker is $^{<<}\text{--}(\text{CR}_2)_n\text{--}^{<<<}$, $^{<<}\text{--O}(\text{CR}_2)_n\text{--}^{<<<}$, or

$^{<<}\text{--}(\text{OCR}_2\text{CR}_2)_n\text{--}^{<<<}$; R^8 is , and n is a whole number selected from 1-4.

19. The method of claim 14, wherein the active agent is a chemotherapeutic agent, a radiotherapeutic agent, a fluorescent dye, or an affinity tag.

20. A functionalized polypeptide comprising an N-terminal amino acid represented by the structure of Formula 2:



2

m is 1, 2, or 3;

N^* is the N-terminal nitrogen of the functionalized polypeptide;

R^1 is the side chain of the N-terminal amino acid of the functionalized polypeptide and the N-terminal amino acid is a natural amino acid or a non-natural amino acid;

R^2 is independently for each instance selected from the group consisting of hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl, aralkyl, heterocycloalkyl, heteraryl, halide, cyano, nitro, hydroxyl, mercaptan, $-\text{OR}^5$, $-\text{SR}^5$, $-(\text{C}=\text{O})\text{OR}^5$, $-\text{O}(\text{C}=\text{O})\text{R}^5$, $-\text{N}(\text{R}^5)_2$,

$-(C=O)N(R^5)_2$, $-N(R^5)(C=O)R^5$, $-N(R^5)(C=O)N(R^5)_2$, $-SO_2R^5$, $-N(R^5)SO_2R^5$, and $-SO_2N(R^5)_2$;

R^3 is selected from the group consisting of $-R^4$ or $-A-R^4$, wherein A is a linker;

R^4 is selected from the group consisting of hydroxyl, thiol, aldehyde, ketone, maleimide, *N*-hydroxysuccinimide ester, imidoester, haloacetyl, pyridyl disulfide, alkyne, azide, nitron, nitrile oxide, norbornene, oxanorbornadiene, tetrazine, cyclooctene, isocyanide, quadricyclane, and bis(dithiobenzil)nickel(II); and

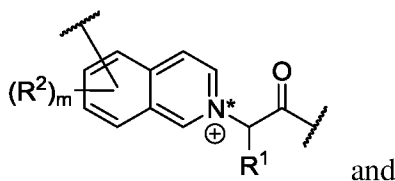
R^5 for each instance is independently selected from the group consisting of hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl, aralkyl, heterocycloalkyl, and heteraryl; or two instances of R^5 taken together with the atoms to which they are attached form a 5-6 membered heterocycloalkyl.

21. The functionalized peptide of claim 20, wherein the functionalized polypeptide is a protein, enzyme, antibody, glycoprotein, or lipoprotein.

22. The functionalized peptide of claim 20, wherein the functionalized polypeptide is hemoglobin.

23. The functionalized polypeptide of claim 20, wherein the linker is selected from the group consisting of $^{**}-(CR_2)_n^{***}$, $^{**}-O(CR_2)_n^{***}$, $^{**}-(CR_2)_nO^{***}$, $^{**}-O(CR_2)_nO^{***}$, $^{**}-(CR_2)_nC(=O)^{***}$, $^{**}-C(=O)(CR_2)_n^{***}$, $^{**}-C(=O)O(CR_2)_n^{***}$, $^{**}-(CR_2)_nOC(=O)^{***}$, $^{**}-O(CR_2)_nC(=O)^{***}$, $^{**}-O(CR_2)_nC(=O)N(R)(CR_2)_p^{***}$, $^{**}-O(CR_2)_nN(R)C(=O)N(R)(CR_2)_p^{***}$, $^{**}-O(CR_2)_nC(=O)O(CR_2)_p^{***}$, $^{**}-O(CR_2)_nOC(=O)N(R)(CR_2)_p^{***}$, $^{**}-C(=O)(CR_2)_nO^{***}$, $^{**}-O(CR_2)_nN(R)C(=O)^{***}$, $^{**}-O(CR_2)_nN(R)C(=O)(CR_2)_p^{***}$, $^{**}-C(=O)N(R)(CR_2)_nO^{***}$, $^{**}-OC(=O)(CR_2)_n^{***}$, $^{**}-(CR_2)_nC(=O)O^{***}$, $^{**}-OC(=O)O(CR_2)_n^{***}$, $^{**}-(CR_2)_nOC(=O)O^{***}$, $^{**}-C(=O)N(R)(CR_2)_n^{***}$, $^{**}-(CR_2)_n(R)NC(=O)^{***}$, $^{**}-N(R)C(=O)(CR_2)_n^{***}$, and $^{**}-$

$(\text{CR}_2)_n\text{C}(=\text{O})\text{N}(\text{R})-$ ^{***}, $-\text{N}(\text{R})\text{C}(=\text{O})\text{O}(\text{CR}_2)_n-$ ^{***}, $-(\text{CR}_2)_n\text{OC}(=\text{O})\text{N}(\text{R})-$ ^{***}, $-\text{OC}(=\text{O})\text{N}(\text{R})(\text{CR}_2)_n-$ ^{***}, $-(\text{CR}_2)_n\text{N}(\text{R})\text{C}(=\text{O})\text{O}-$ ^{***}, $-(\text{OCR}_2\text{CR}_2)_n-$ ^{***}, $-(\text{CR}_2\text{CR}_2\text{O})_n-$ ^{***}, $-(\text{OCR}_2\text{CR}_2)_n(\text{C}=\text{O})-$ ^{***}, $-(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C}=\text{O})-$ ^{***}, $-(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})-$ ^{***}, $-(\text{OCR}_2\text{CR}_2)_n(\text{C}=\text{O})(\text{CR}_2)_p-$ ^{***}, $-(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C}=\text{O})(\text{CR}_2)_p-$ ^{***}, $-(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})(\text{CR}_2)_p-$ ^{***}, $-(\text{OCR}_2\text{CR}_2)_n(\text{C}=\text{O})\text{O}(\text{CR}_2)_p-$ ^{***}, $-(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C}=\text{O})\text{O}(\text{CR}_2)_p-$ ^{***}, $-(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})\text{O}(\text{CR}_2)_p-$ ^{***}, $-(\text{OCR}_2\text{CR}_2)_n(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p-$ ^{***}, $-(\text{OCR}_2\text{CR}_2)_n\text{O}(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p-$ ^{***}, $-(\text{OCR}_2\text{CR}_2)_n\text{N}(\text{R})(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p-$ ^{***}, $-\text{OCR}_2(\text{C}=\text{O})-$ ^{***}, $-\text{OCR}_2(\text{C}=\text{O})(\text{CR}_2)_p-$ ^{***}, $-\text{OCR}_2(\text{C}=\text{O})\text{O}(\text{CR}_2)_p-$ ^{***}, $-\text{OCR}_2(\text{C}=\text{O})\text{N}(\text{R})(\text{CR}_2)_p-$ ^{***}, $-\text{C}(=\text{O})\text{O}(\text{CR}_2\text{CR}_2\text{O})_n\text{CR}_2\text{CR}_2-$ ^{***}, $-\text{C}(=\text{O})\text{N}(\text{R})(\text{CR}_2\text{CR}_2\text{O})_n\text{CR}_2\text{CR}_2-$ ^{***}, $-\text{S}(\text{CR}_2)_n-$ ^{***}, $-(\text{CR}_2)_n\text{S}-$ ^{***}, $-\text{SO}_2(\text{CR}_2)_n-$ ^{***}, $-(\text{CR}_2)_n\text{SO}_2-$ ^{***}, $-\text{N}(\text{R})\text{SO}_2(\text{CR}_2)_n-$ ^{***}, $-(\text{CR}_2)_n\text{SO}_2\text{N}(\text{R})-$ ^{***}, $-\text{SO}_2\text{N}(\text{R})(\text{CR}_2)_n-$ ^{***}, $-(\text{CR}_2)_n\text{N}(\text{R})\text{SO}_2-$ ^{***}, $-(\text{CR}_2)_n\text{Ar}-(\text{CR}_2)_p-$ ^{***}, $-(\text{CR}_2)_p\text{Ar}-(\text{CR}_2)_n-$ ^{***}, $-\text{O}(\text{CR}_2)_n\text{Ar}-(\text{CR}_2)_p-$ ^{***}, and $-(\text{CR}_2)_p\text{Ar}-(\text{CR}_2)_n\text{O}-$ ^{***}, wherein ^{***} indicates the position of a covalent bond with the moiety:



^{***} indicates the position of a covalent bond with R^4 ; each instance of n is independently a whole number selected from 1-20; each instance of p is independently an integer selected from 0-20; and R for each instance is independently selected from hydrogen, alkyl, cycloalkyl, and aryl; or two instances of R taken together with the carbons to which they are attached form a 3-6 membered carbocyclic ring; or two instances of R taken together with the atoms to which they are attached form a 5-6 membered heterocyclic ring.

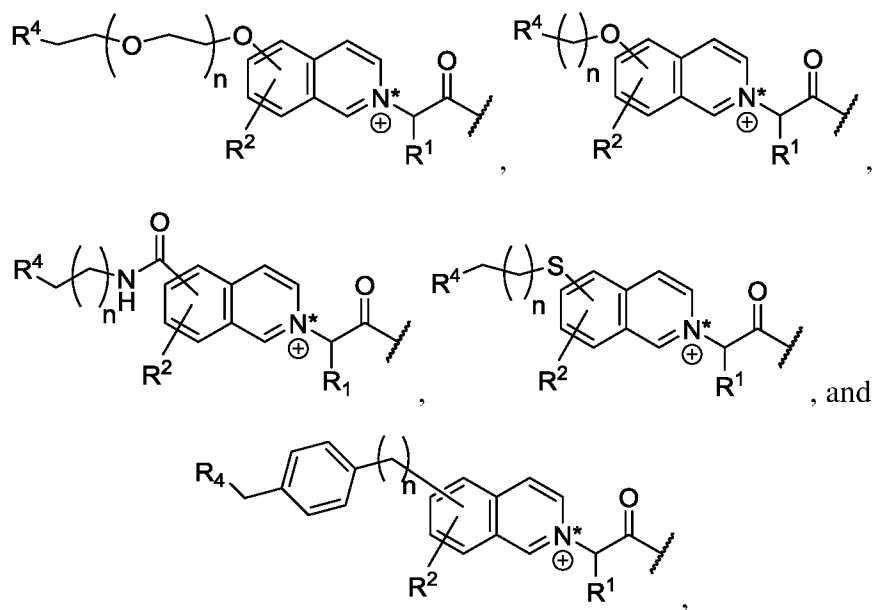
24. The functionalized polypeptide of claim 20, wherein R^2 is hydrogen.

25. The functionalized polypeptide of claim 20, wherein the linker is $^{**}-(CR_2)_n-^{***}$, $^{**}-O(CR_2)_n-^{***}$, or $^{**}-(OCR_2CR_2)_n-^{***}$.

26. The functionalized polypeptide of claim 25, wherein n is a whole number selected from 1-10 and p is a whole number selected from 1-4.

27. The functionalized polypeptide of claim 20, wherein R^4 is maleimide, azide, alkyne, hydroxyl, norbornene, or cyclooctene.

28. The functionalized polypeptide of claim 20, wherein the moiety of Formula 2 is selected from the group consisting of:



wherein

n is a whole number selected from 1-4;

R² is hydrogen, alkyl, or -OR⁵;

R⁴ is maleimide or alkyne; and

R⁵ is alkyl.

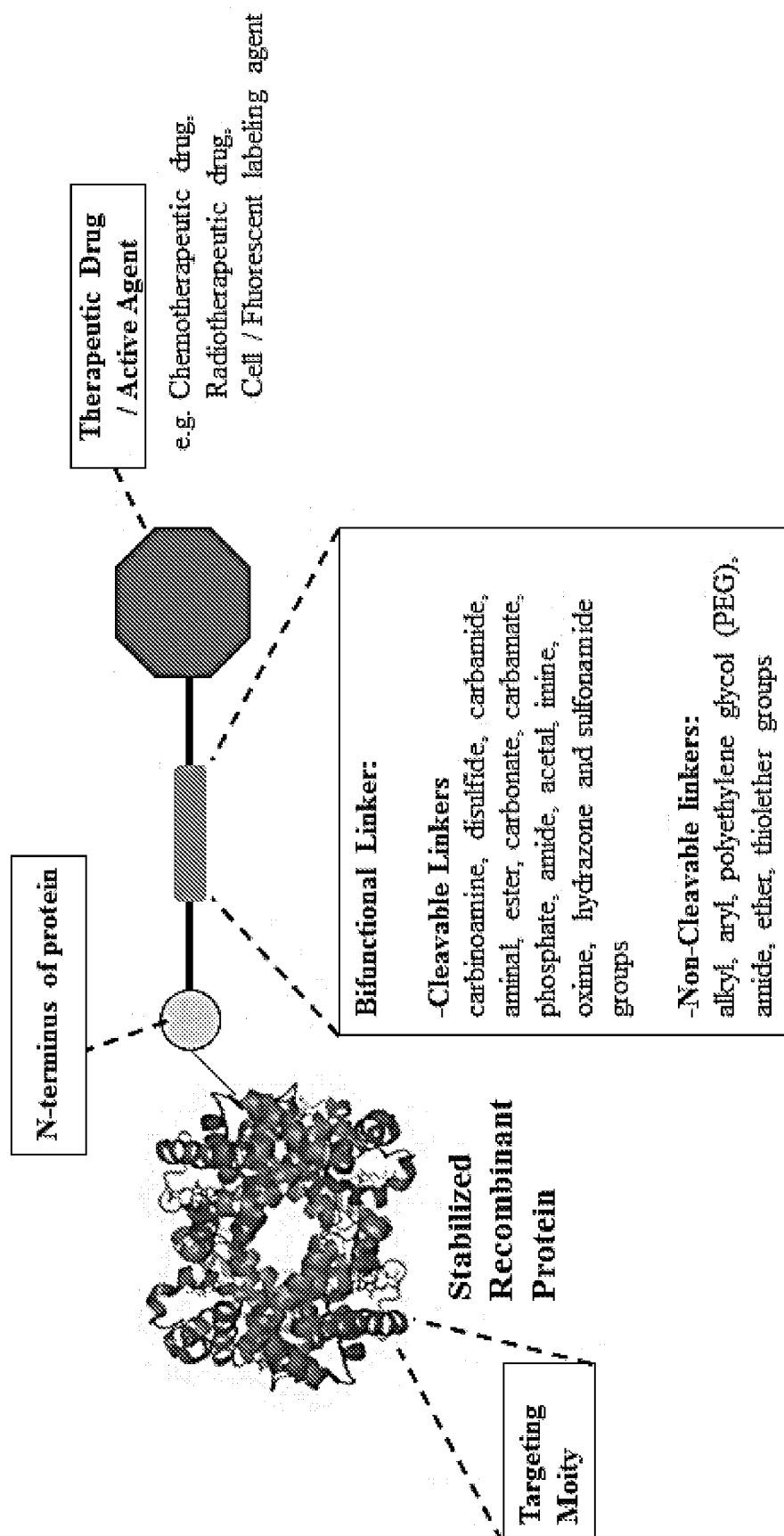


FIG. 1

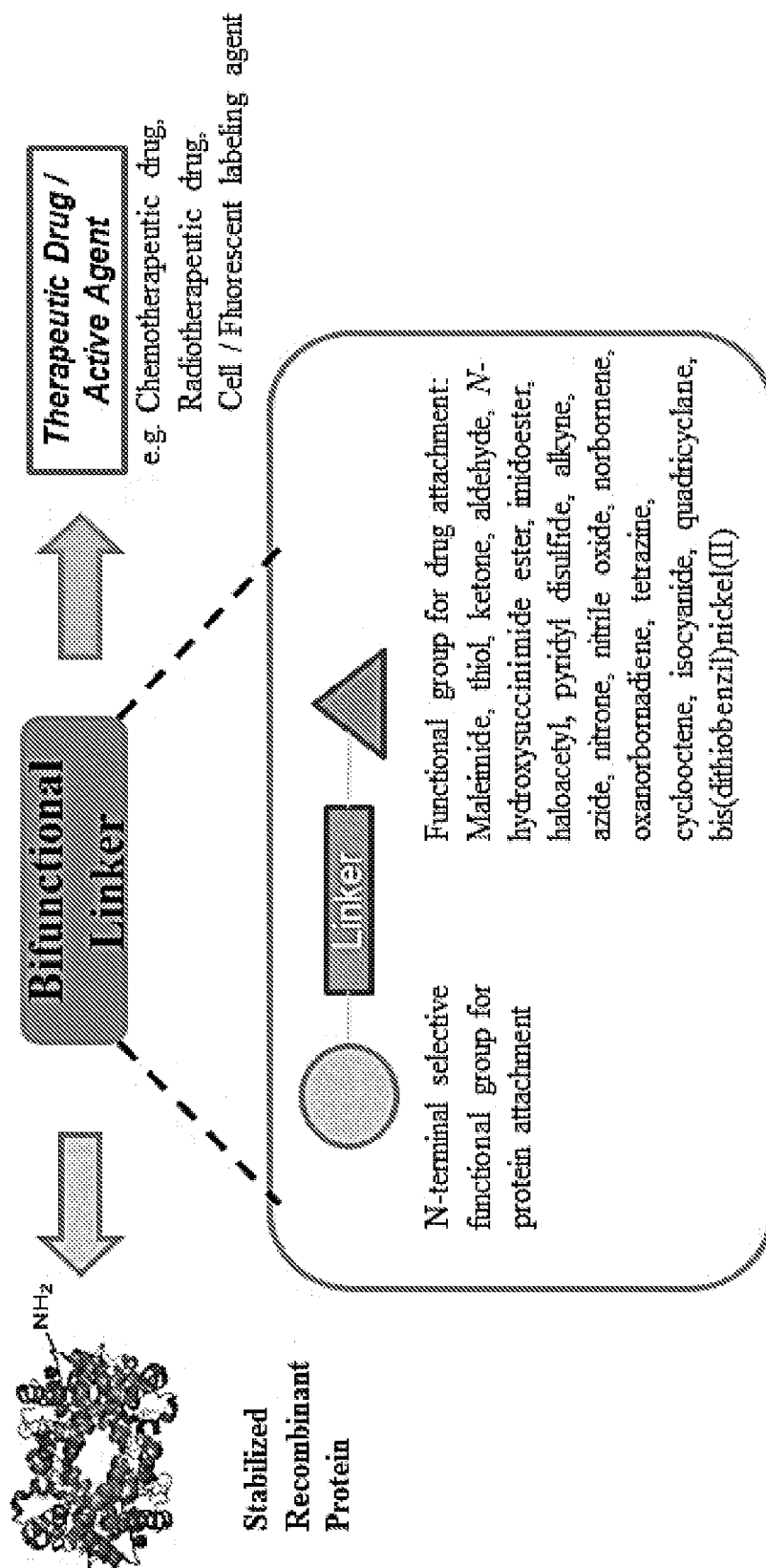


FIG. 2

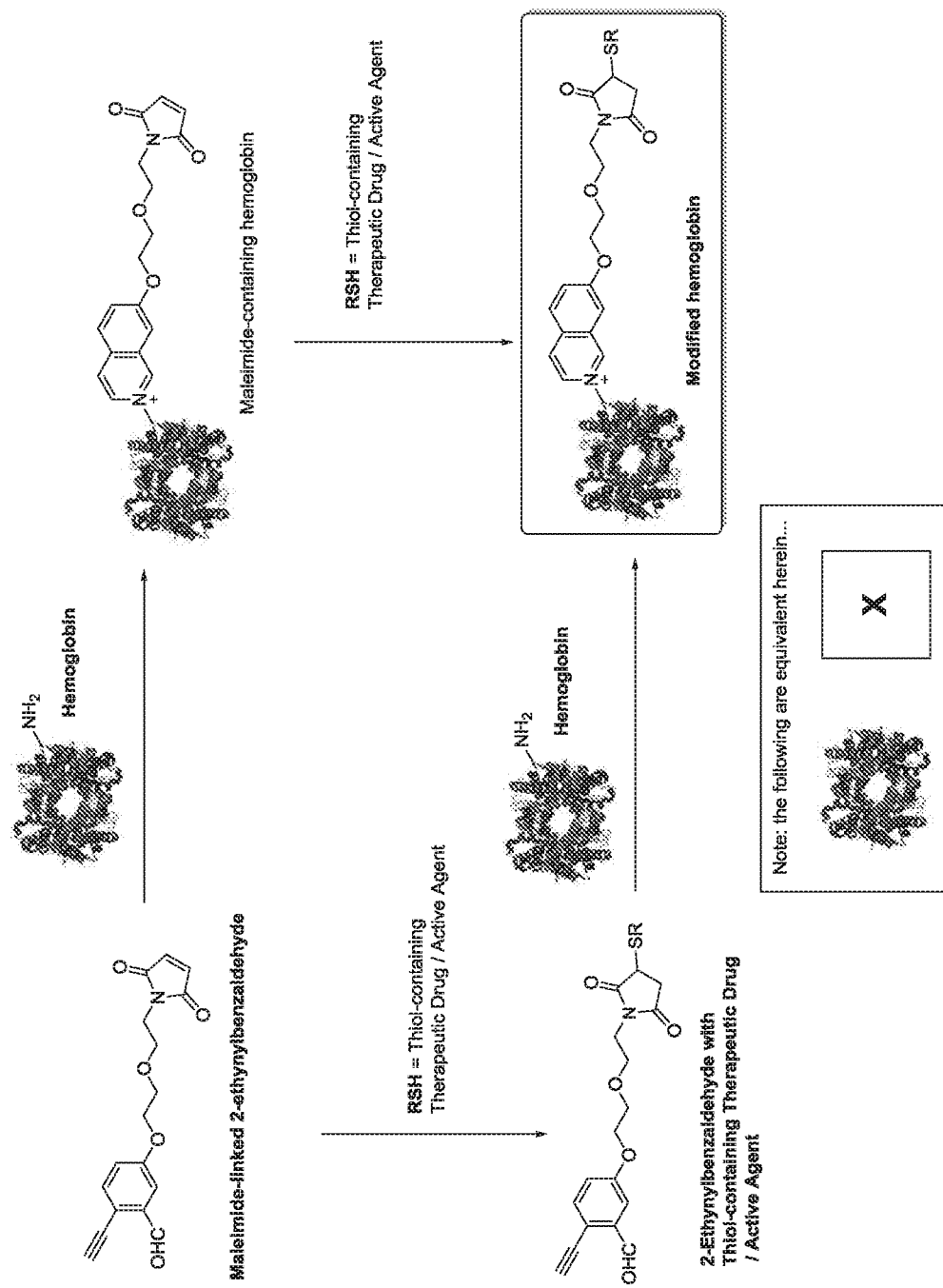


FIG. 3A

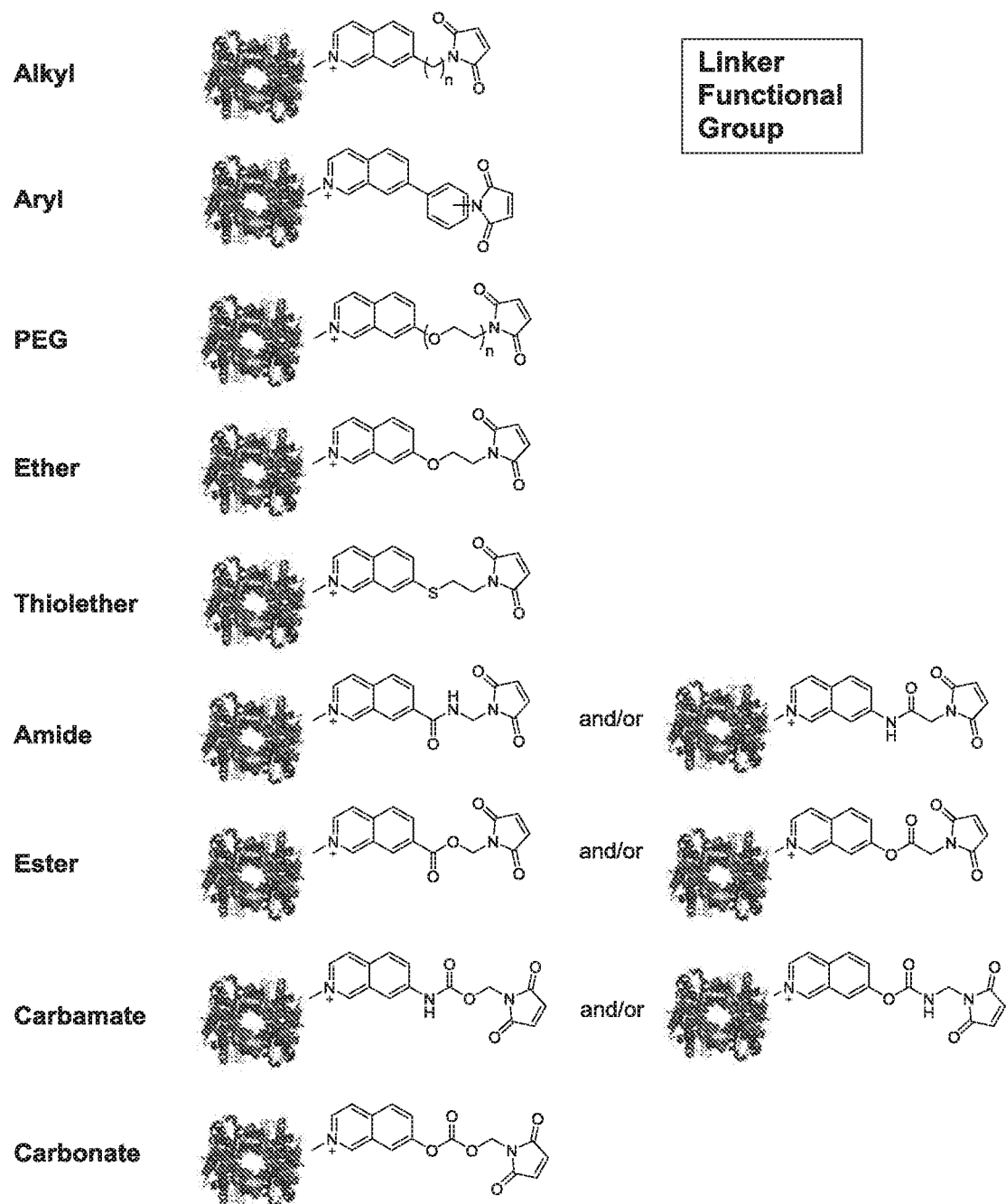


FIG. 3B

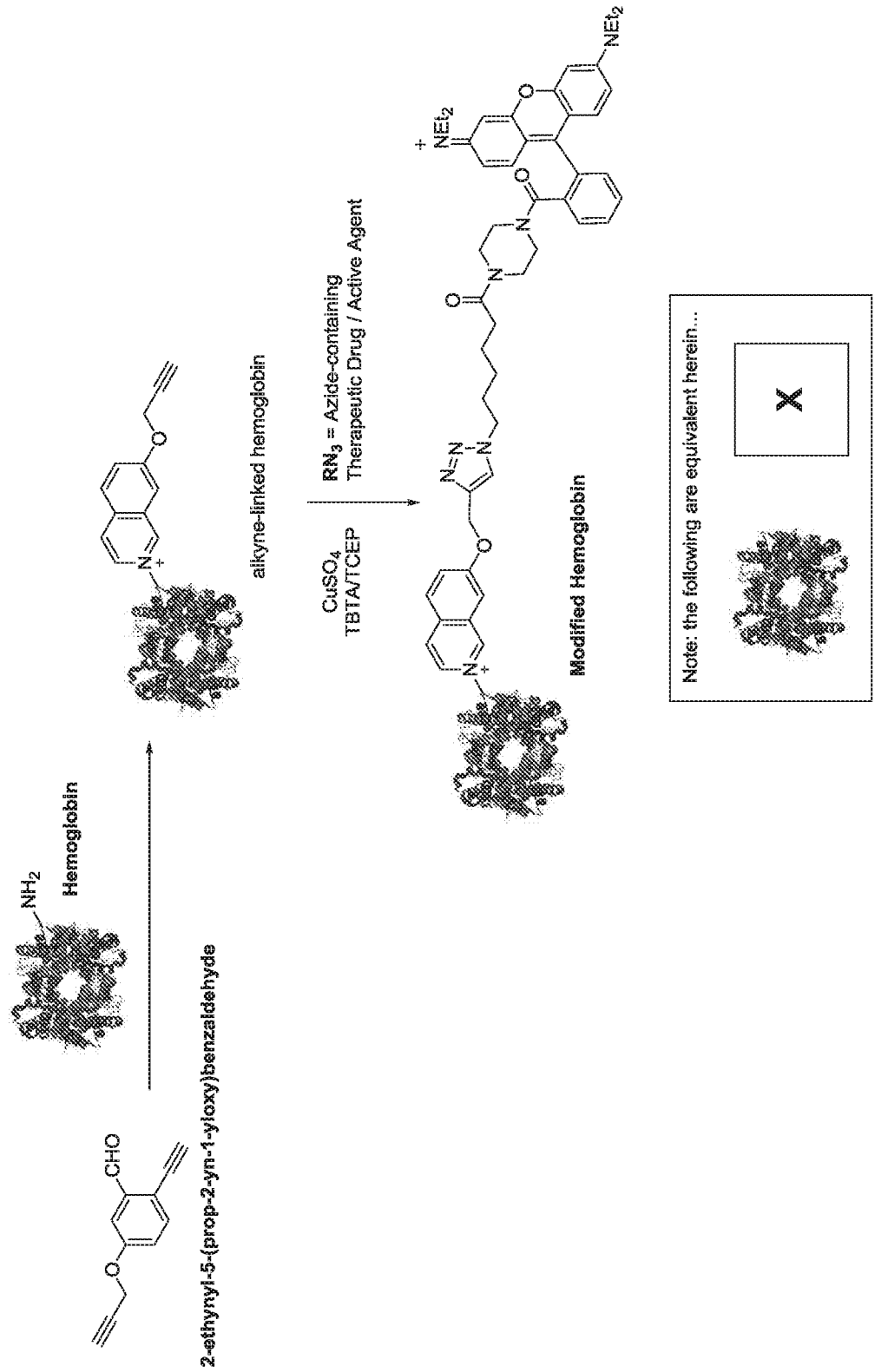


FIG. 4

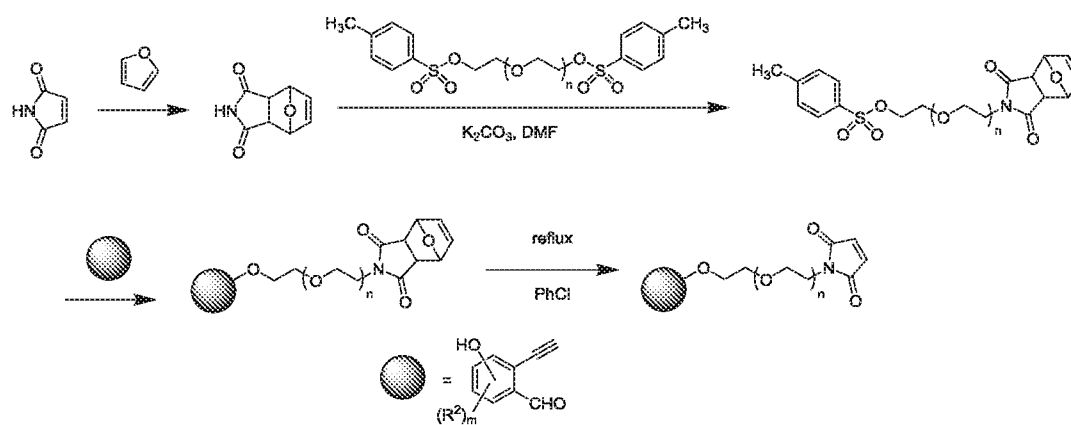


FIG. 5

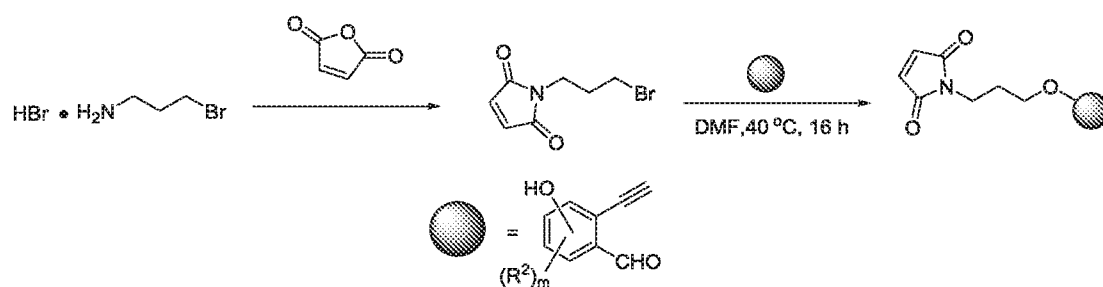


FIG. 6

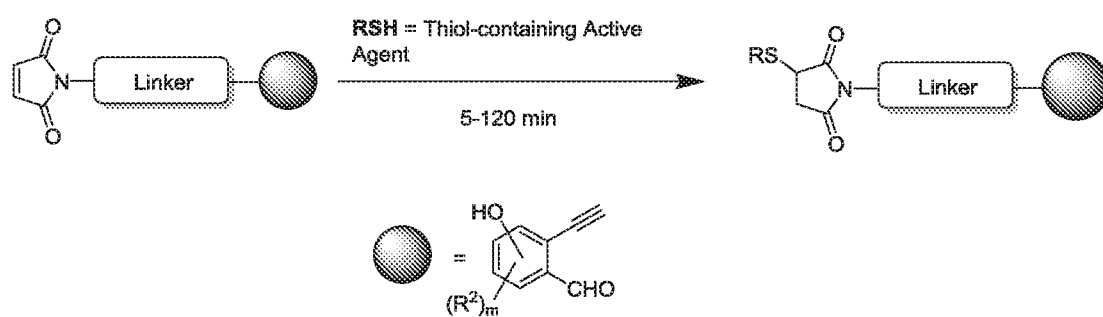
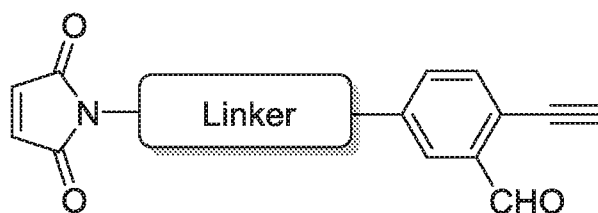


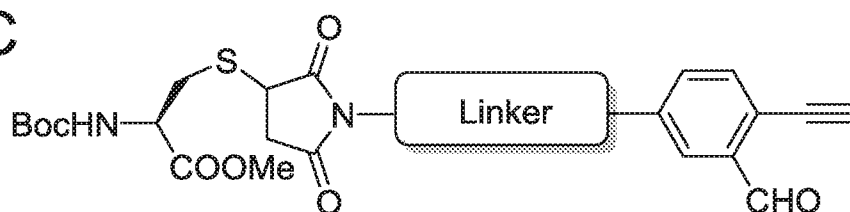
FIG. 7

A

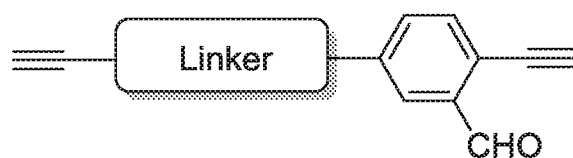
Maleimide-containing 2-ethynylbenzaldehyde (PEG and/or Alkyl)

B

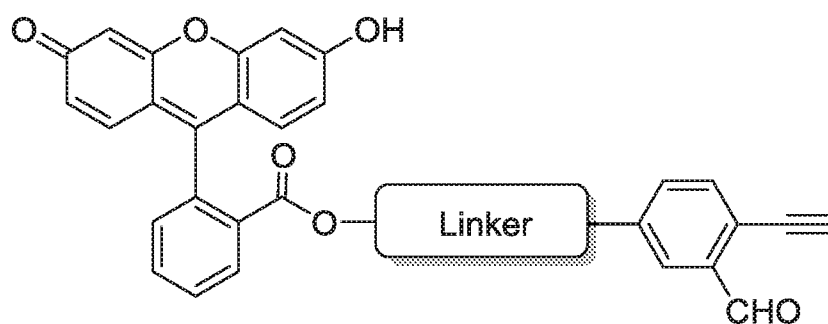
Benzylmercaptan-linked 2-ethynylbenzaldehyde

C

N-Boc-L-cysteine methyl ester-linked 2-ethynylbenzaldehyde

D

Alkyne-linked 2-ethynylbenzaldehyde

E

Fluorescein-linked 2-ethynylbenzaldehyde

FIG. 8

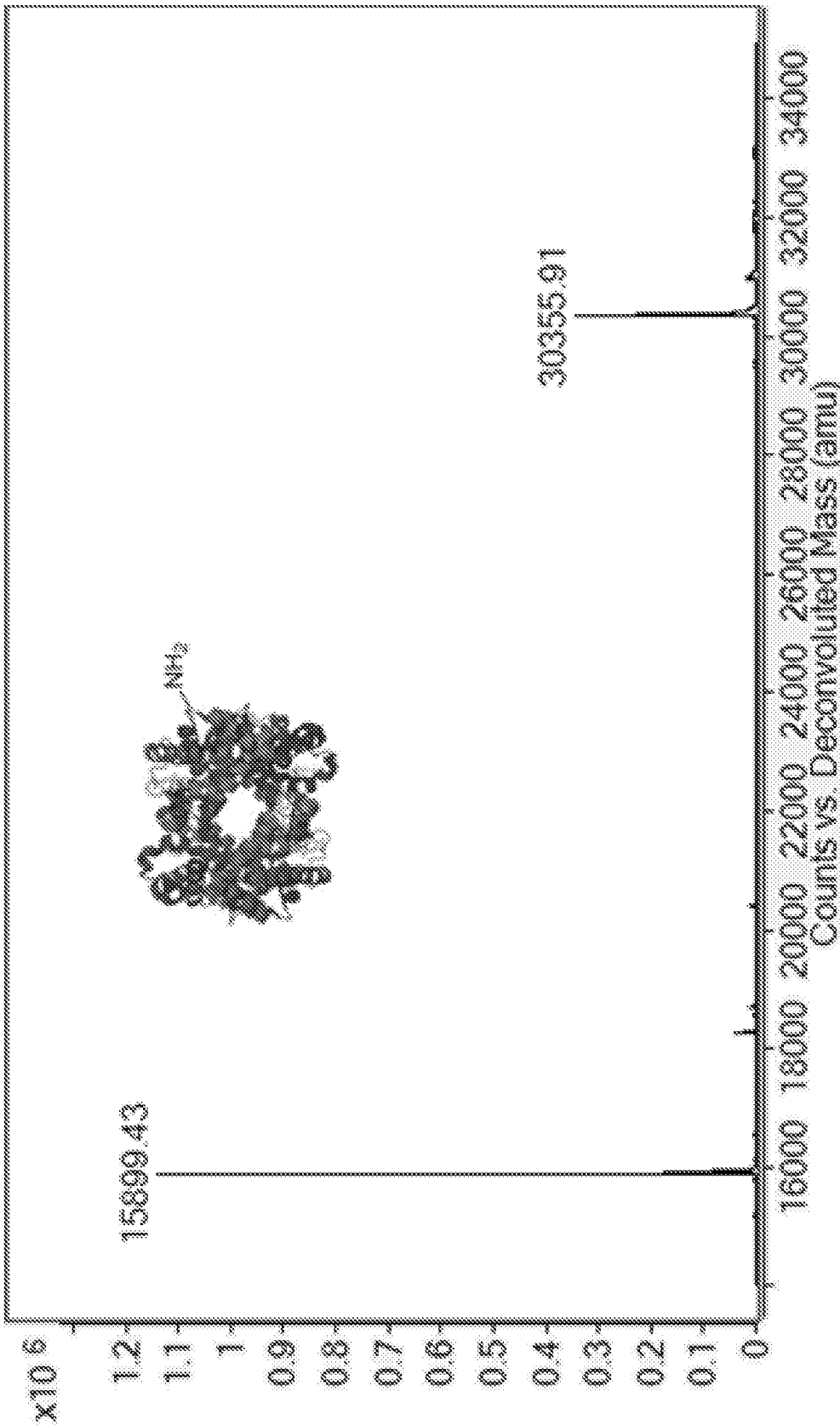


FIG. 9A

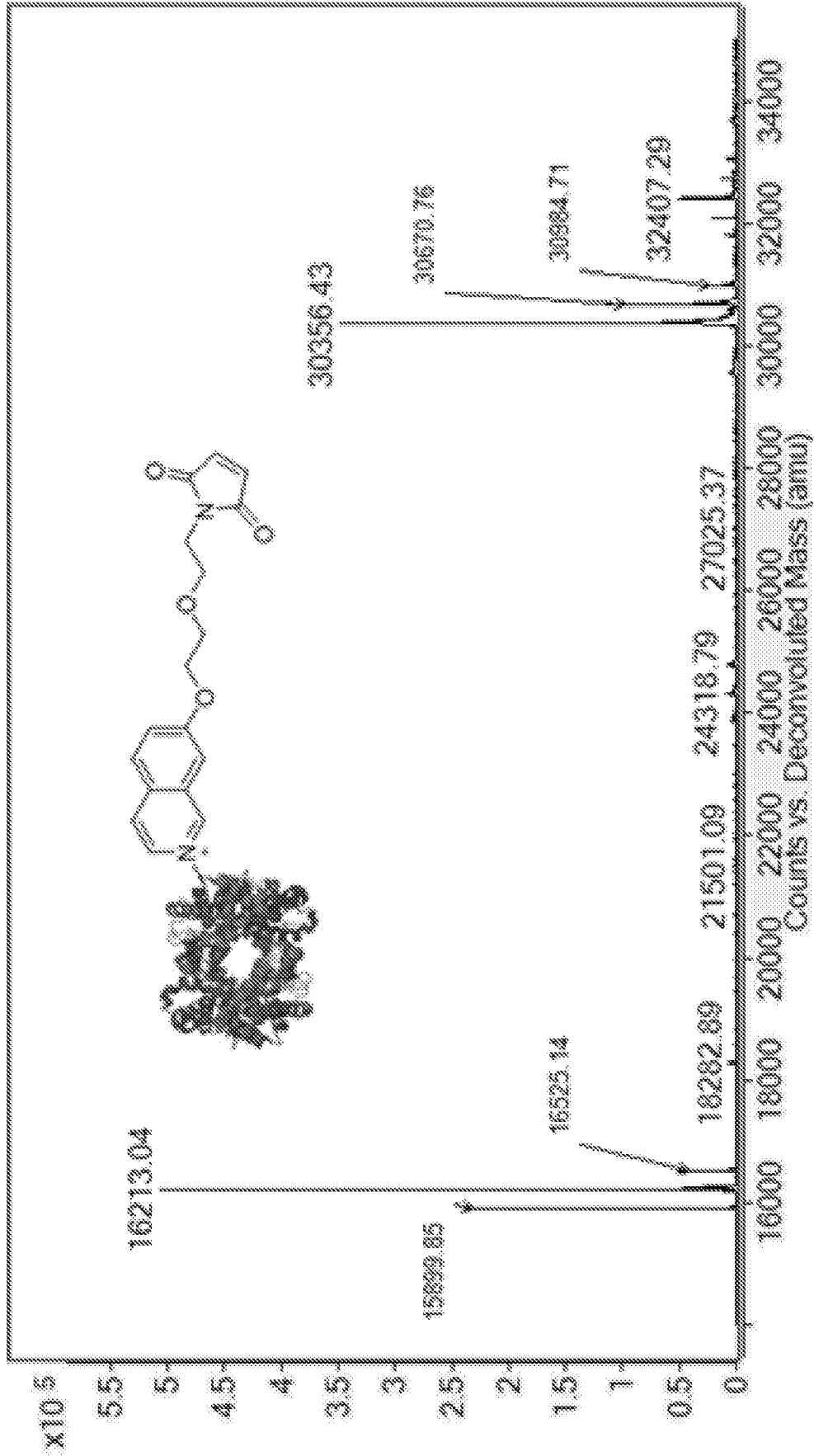


FIG. 9B

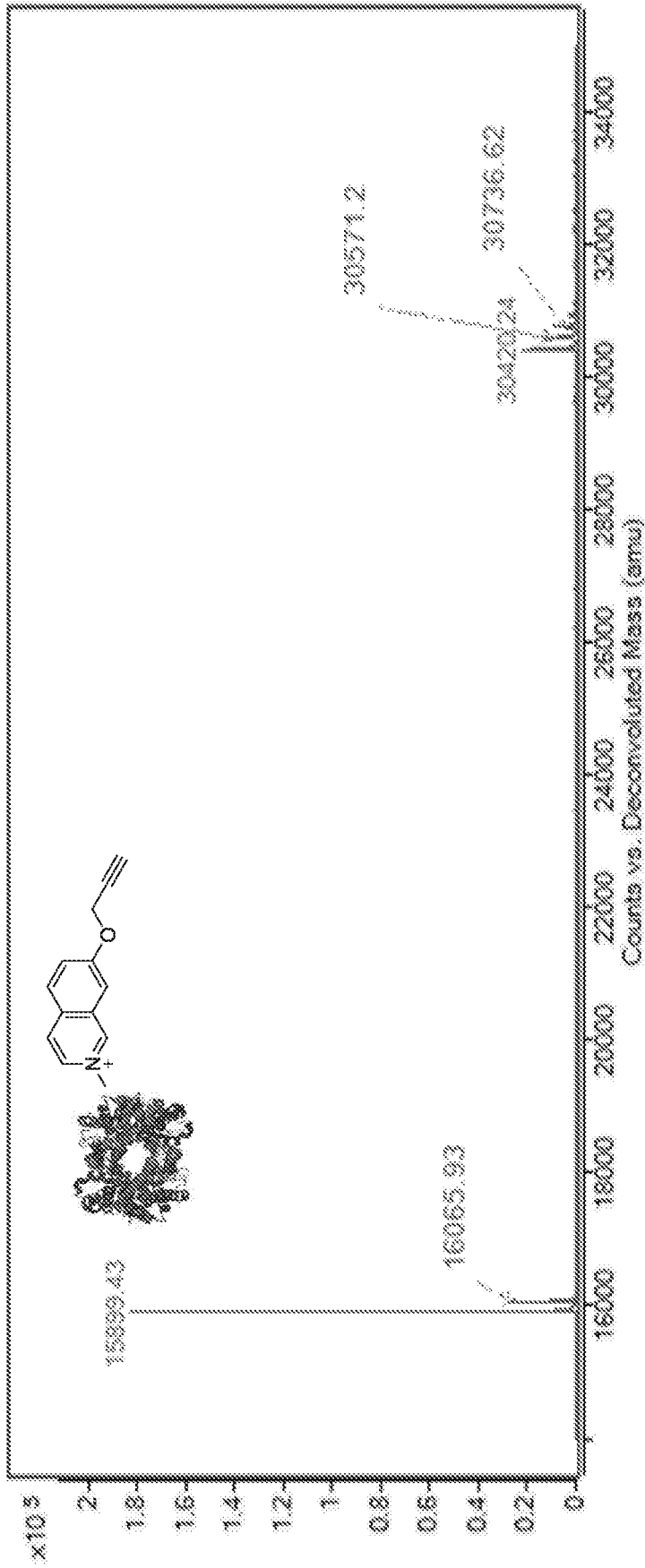


FIG. 9C

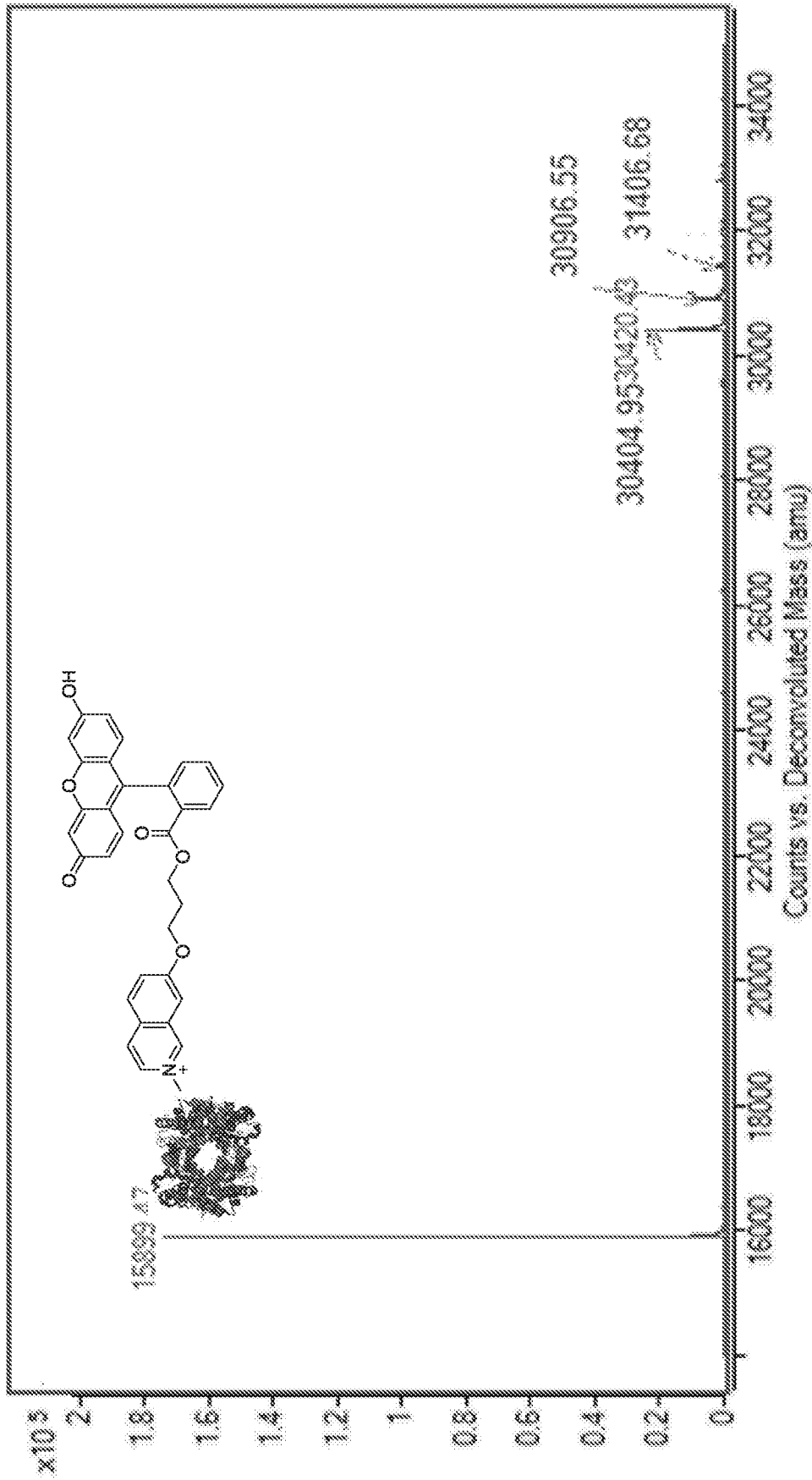


FIG. 9D

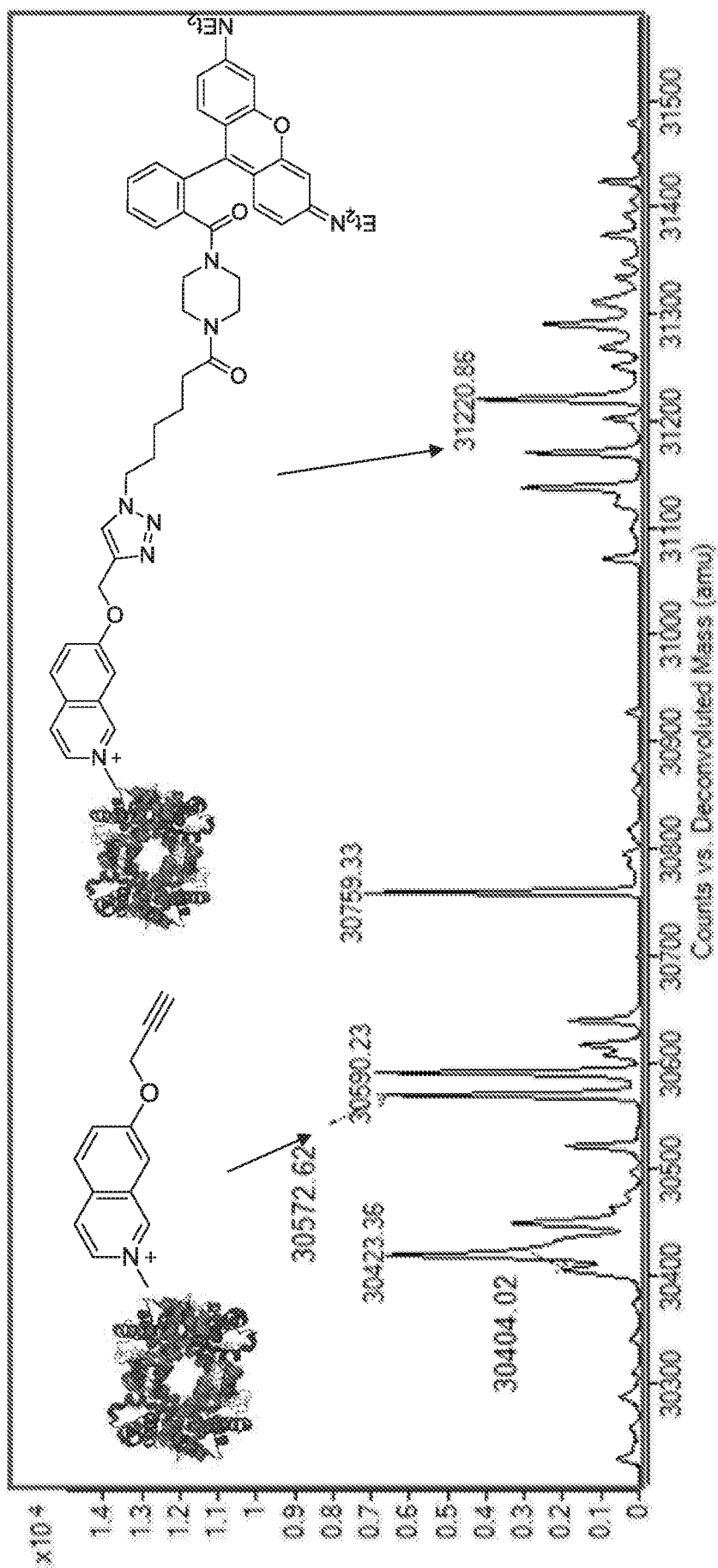


Fig. 9.

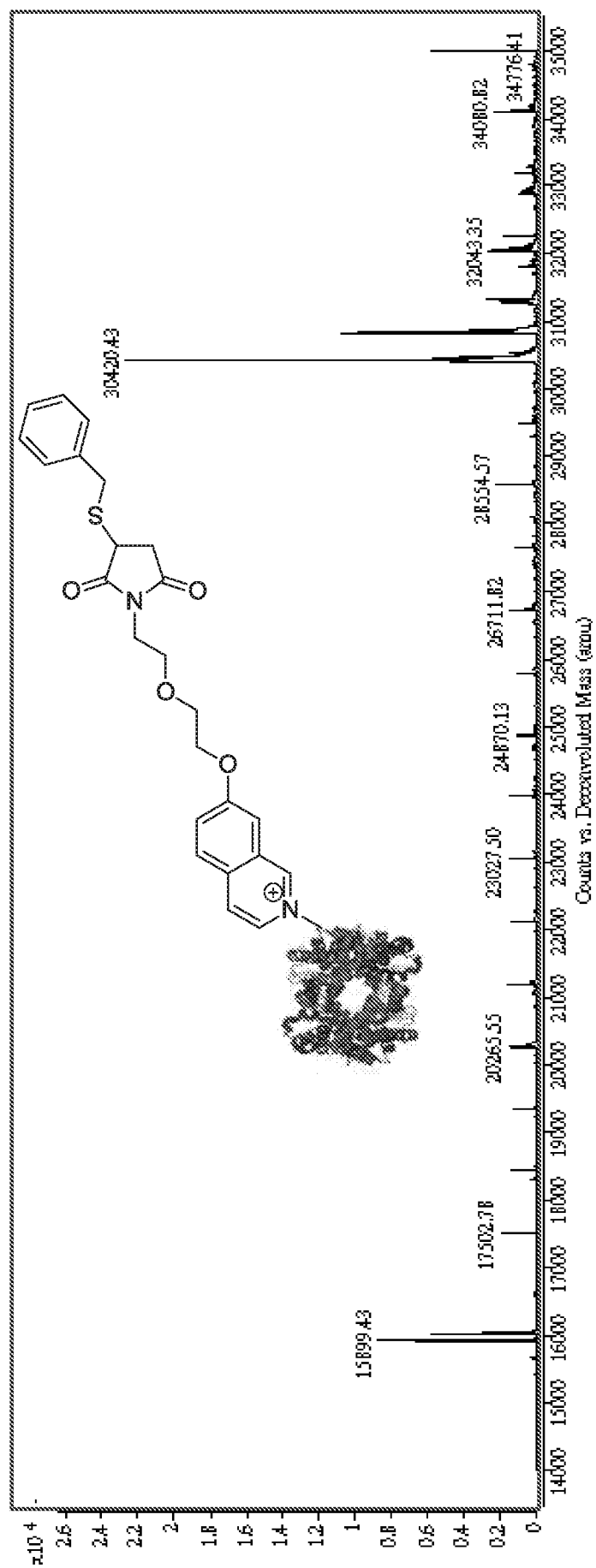


FIG. 9F

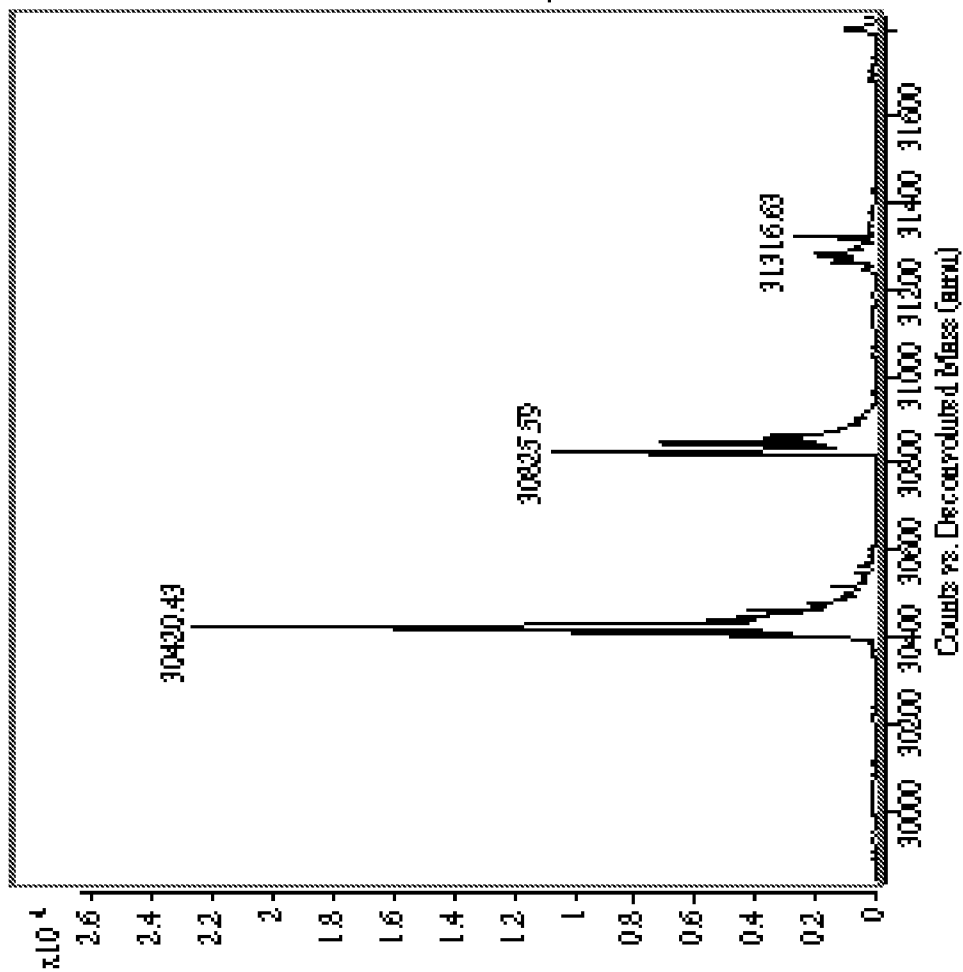


FIG. 9G

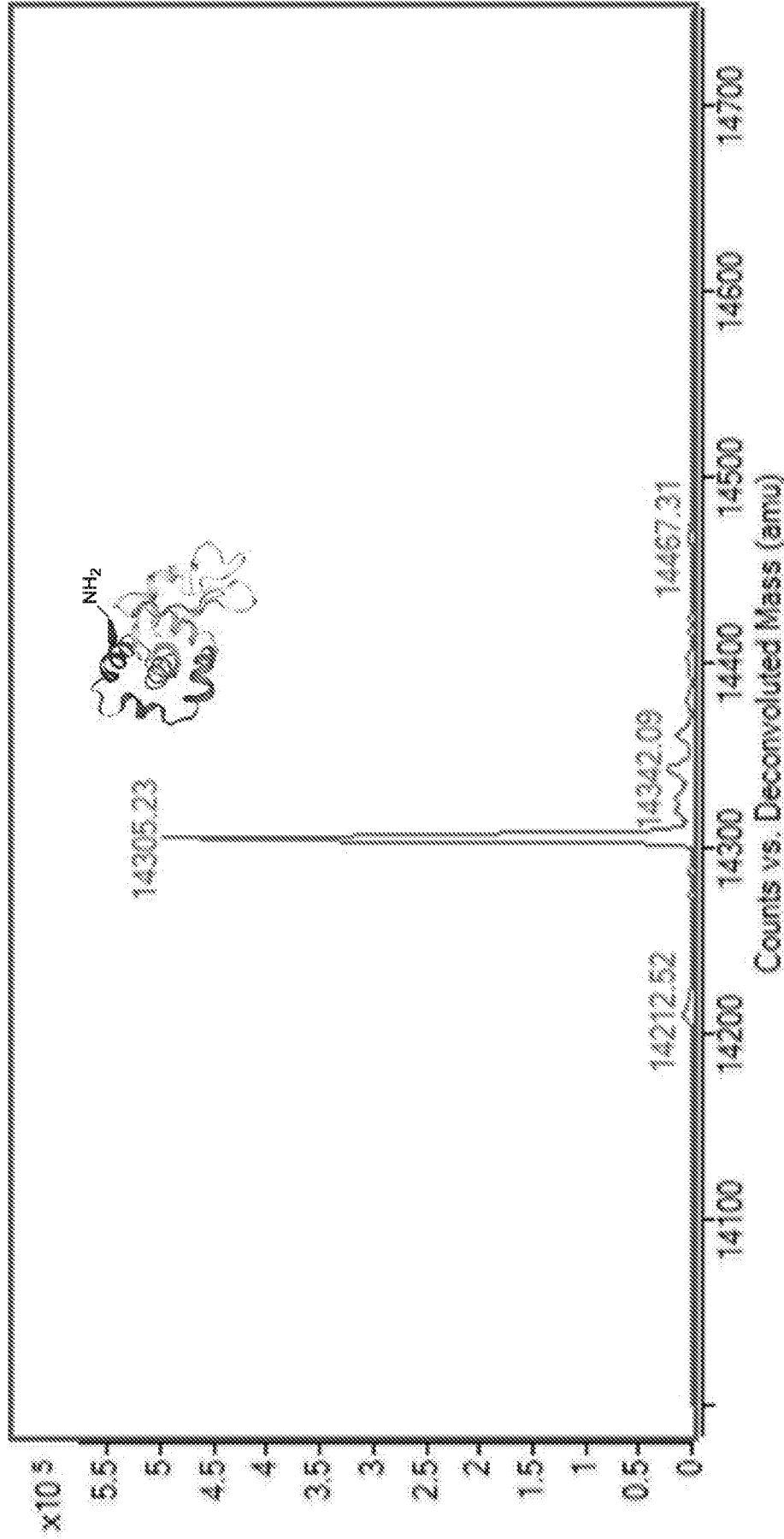


FIG. 10A

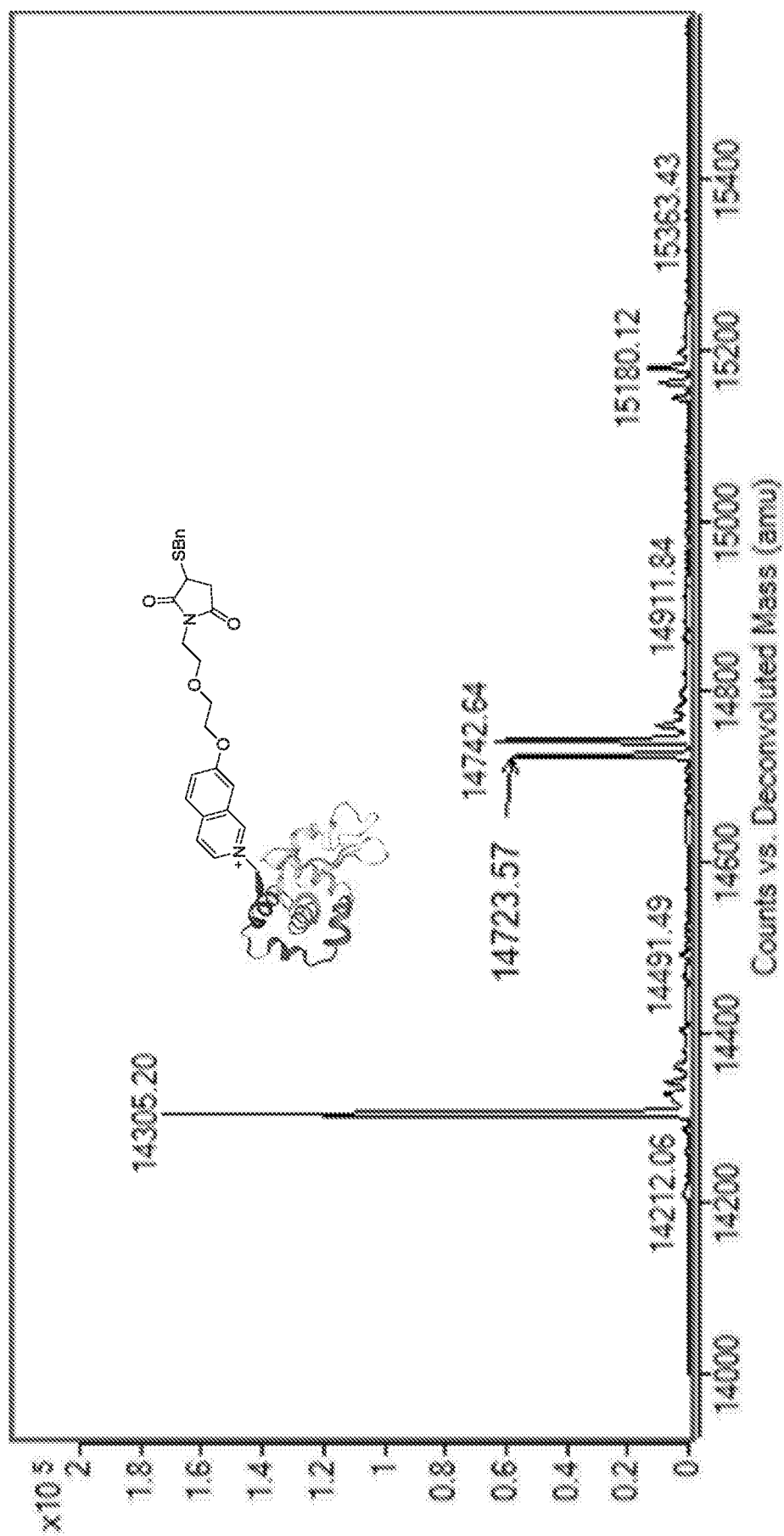


FIG. 10B

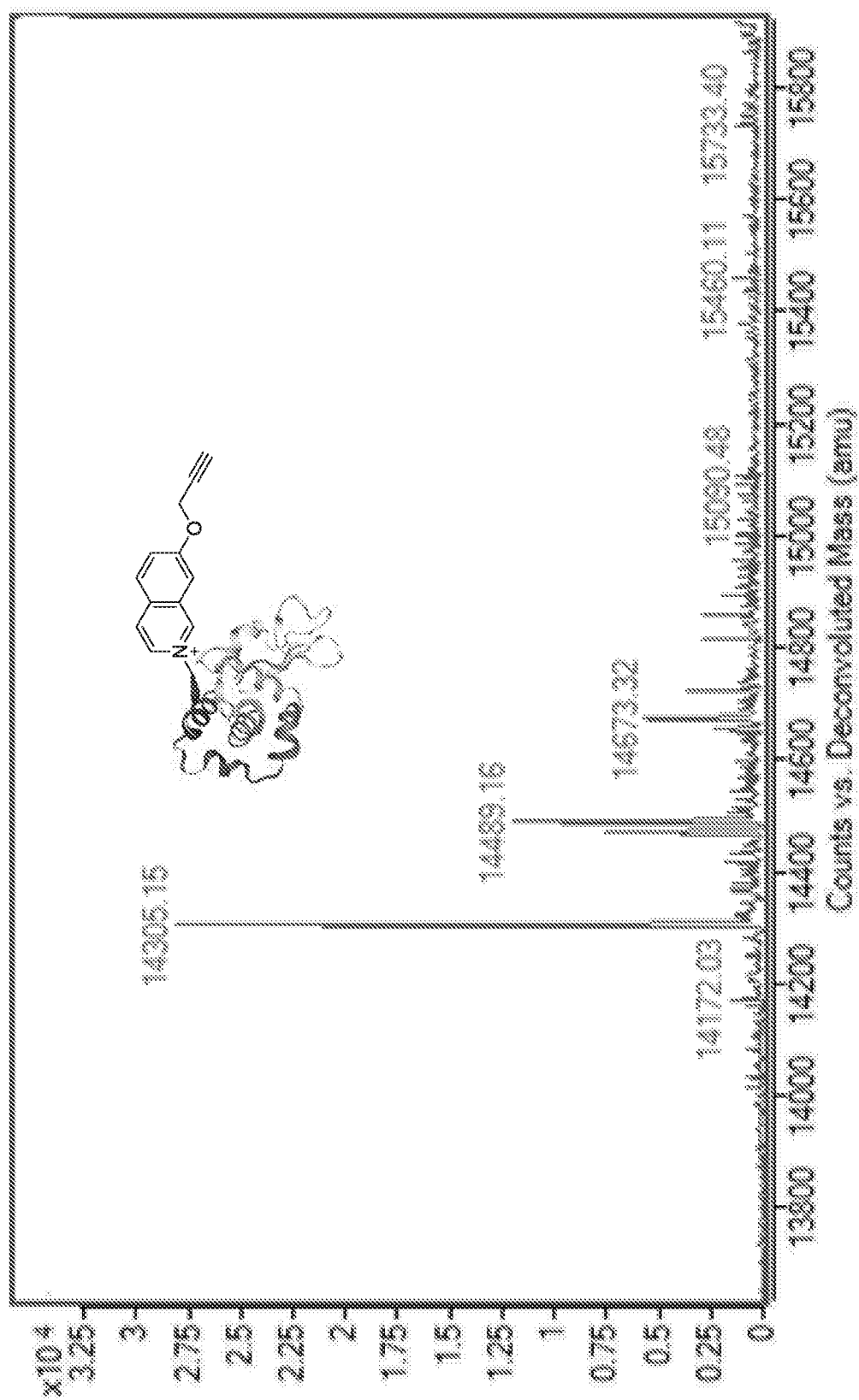


FIG. 10C

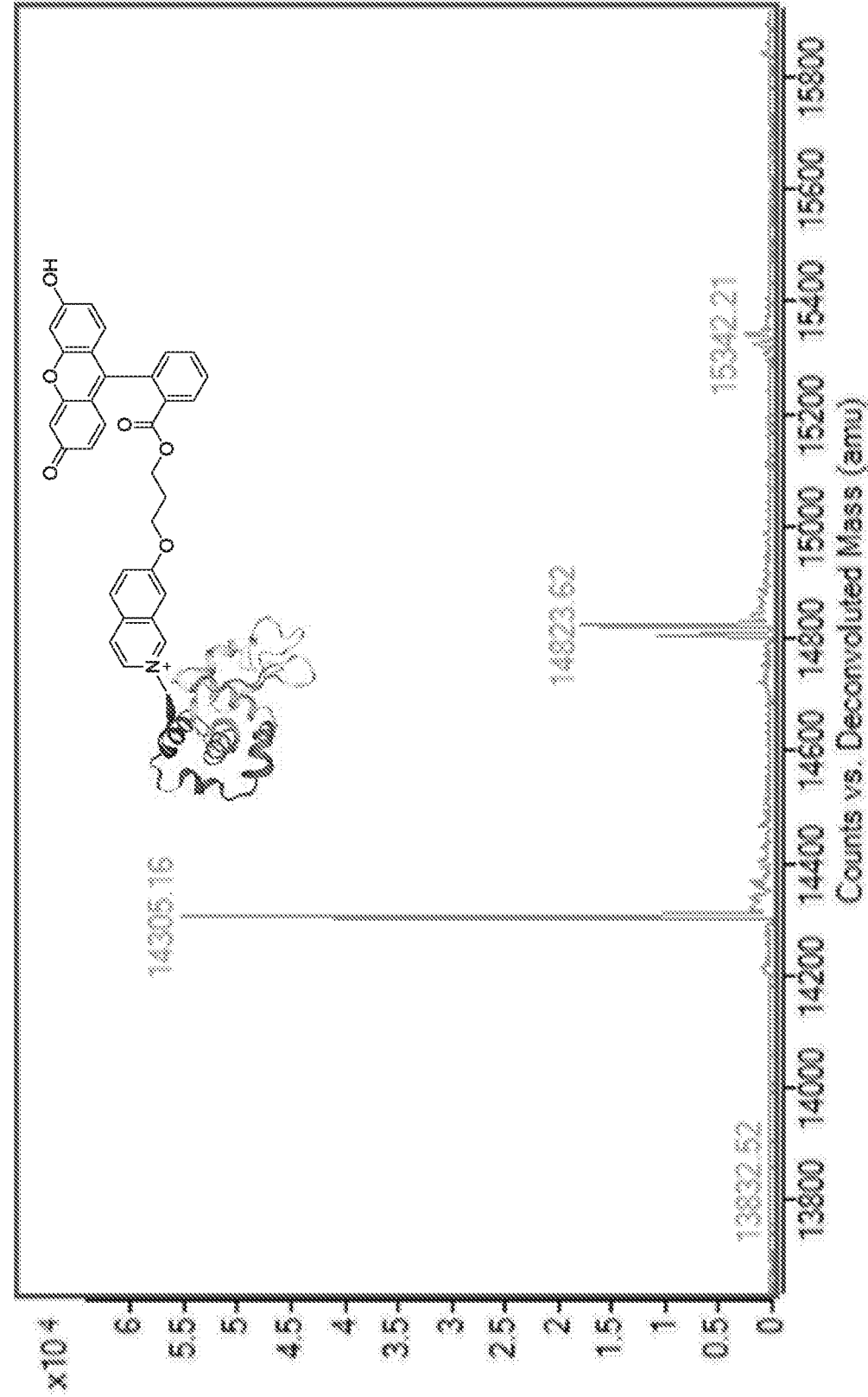


FIG. 10D

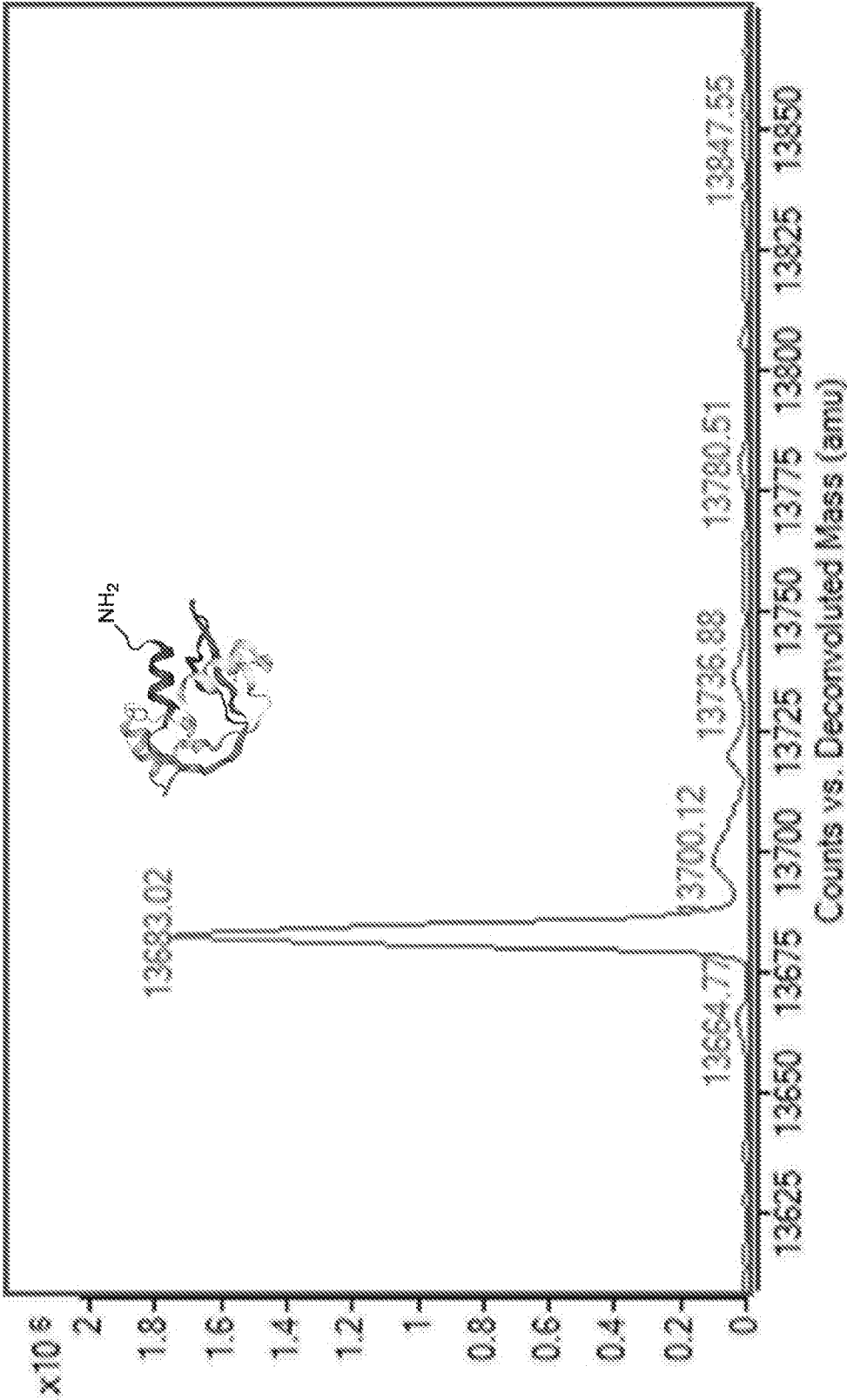


FIG. 11A

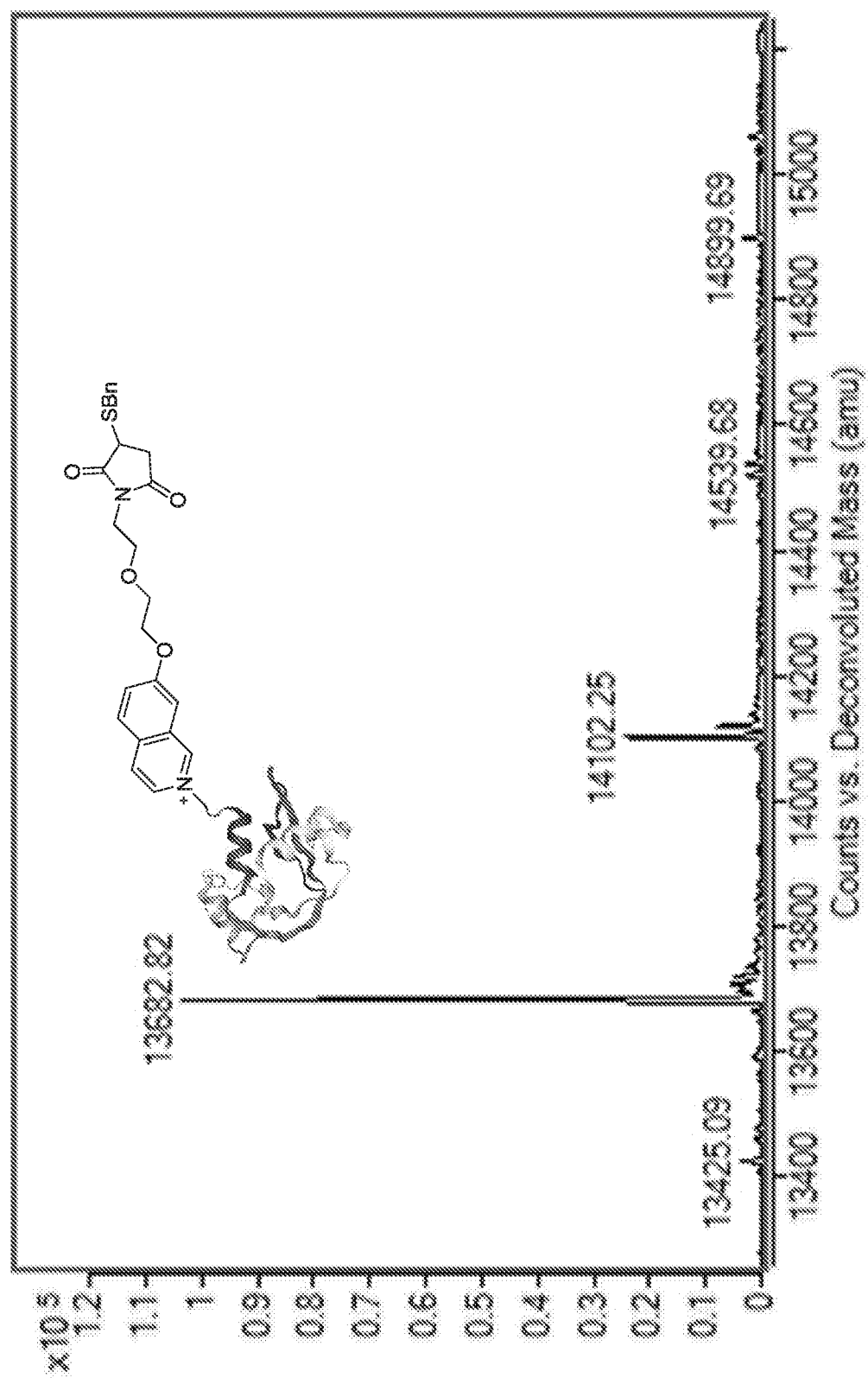


FIG. 11B

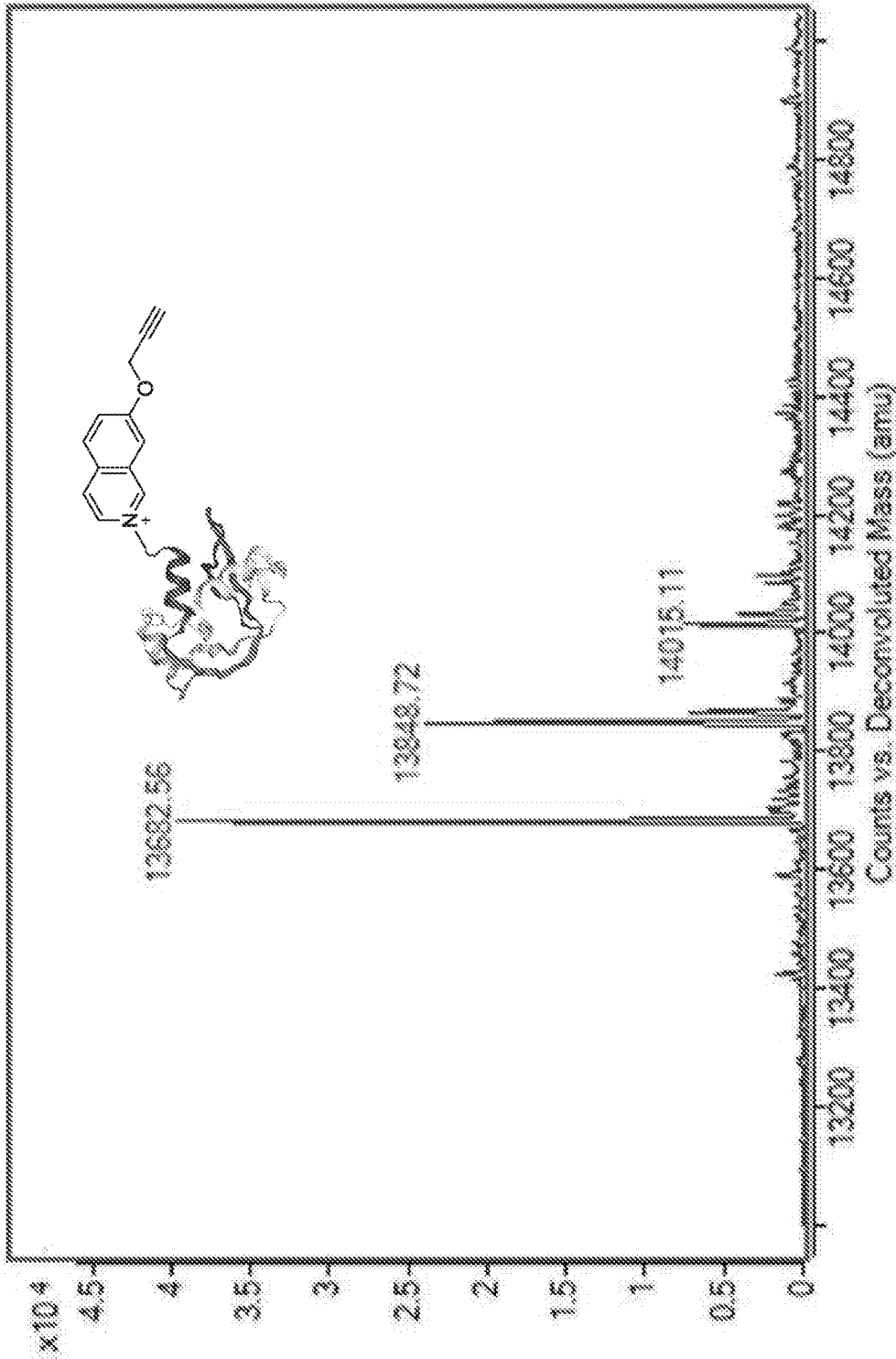


FIG. 11C

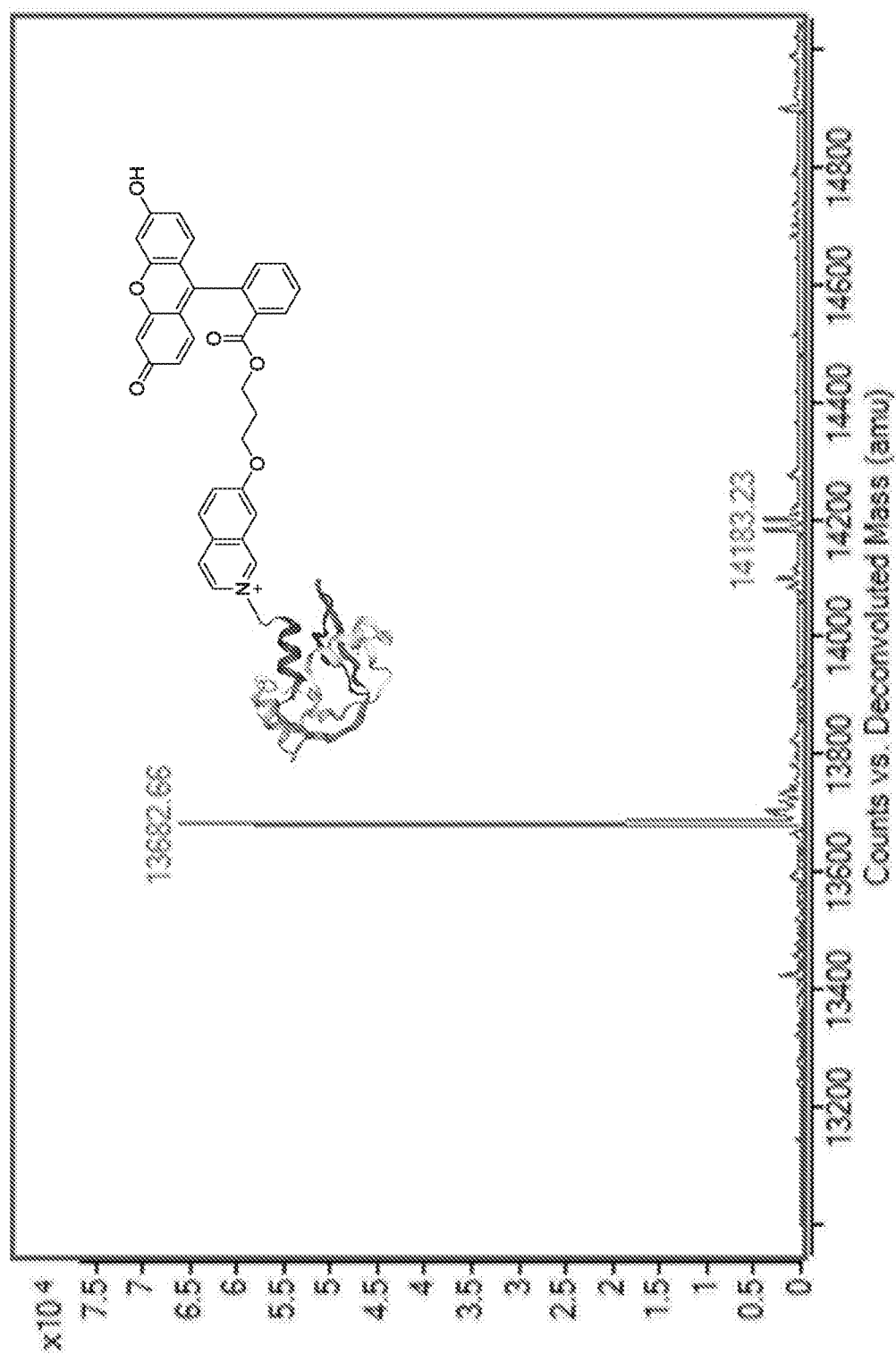


FIG. 11D

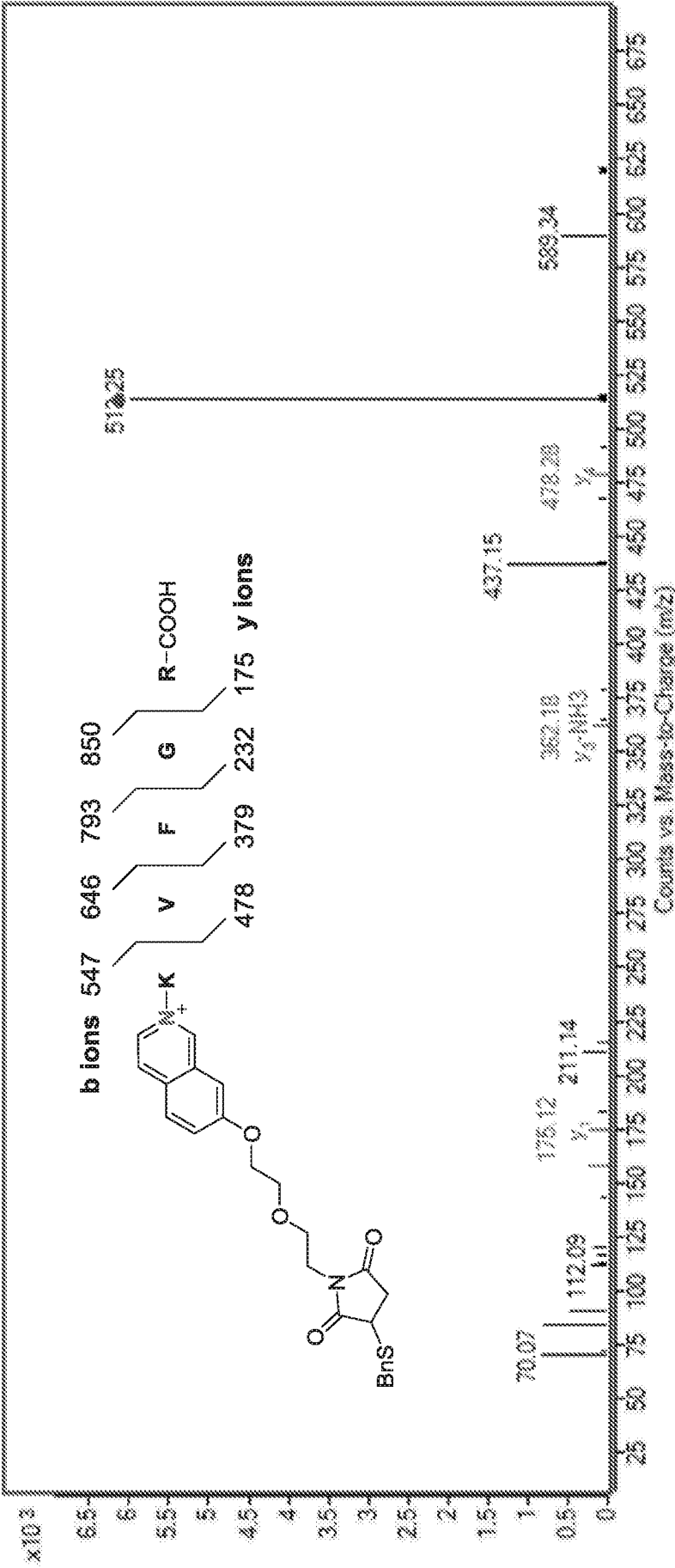


FIG. 12A

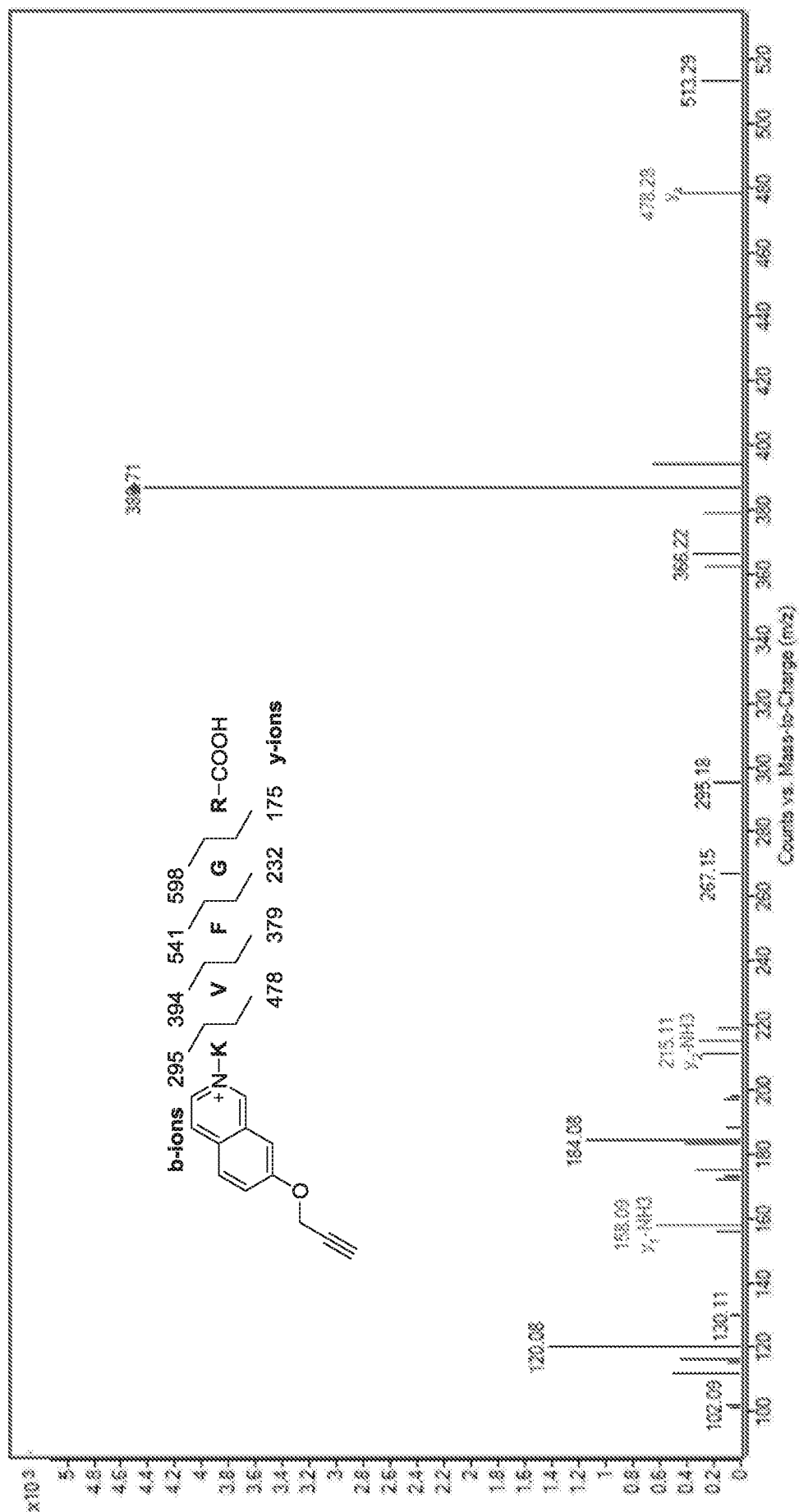


FIG. 12B

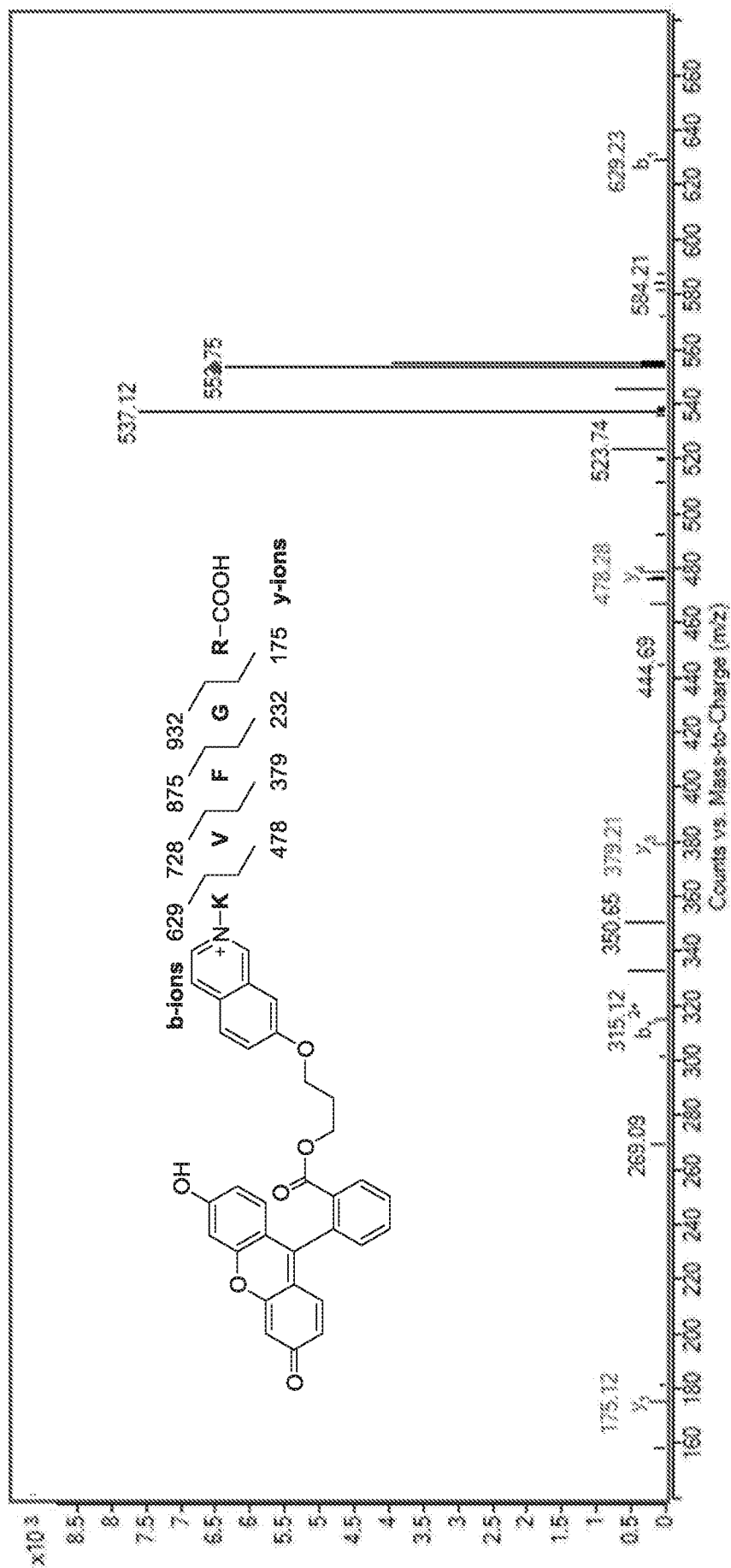


FIG. 12C

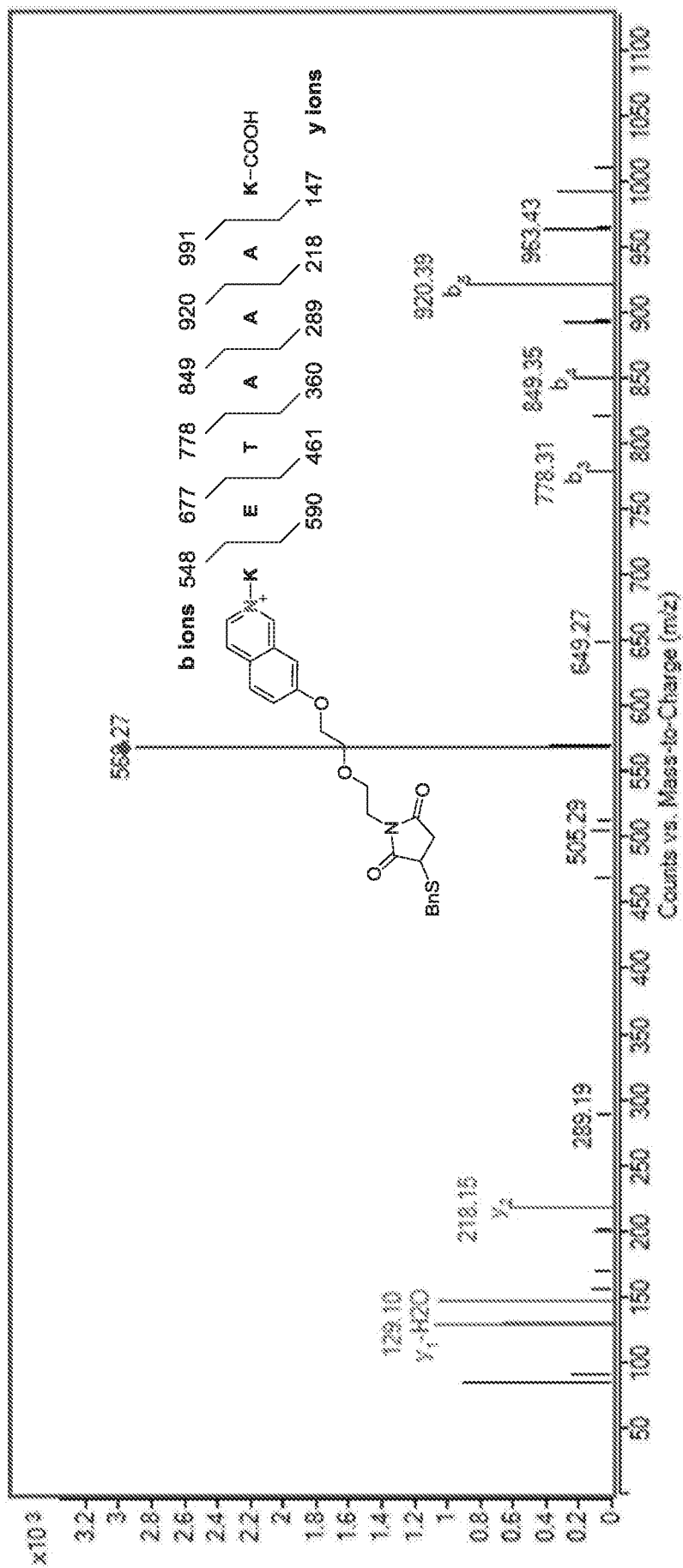


FIG. 12D

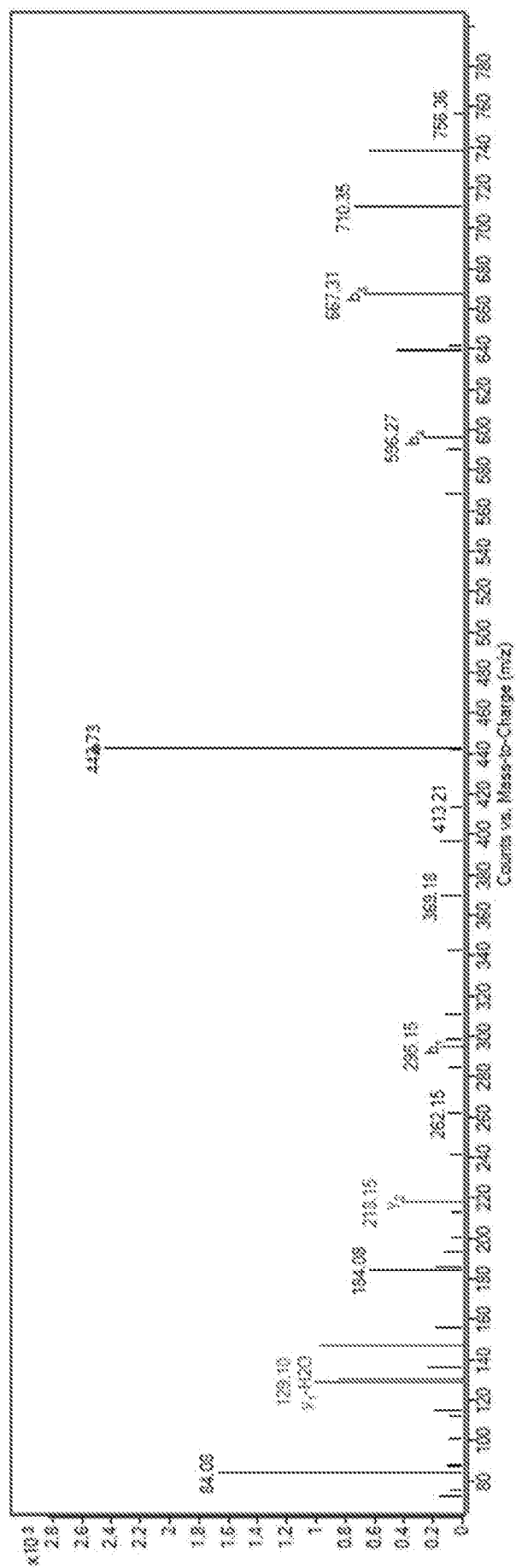
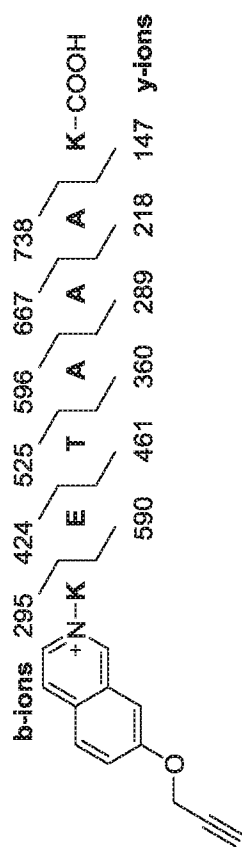


FIG. 12E

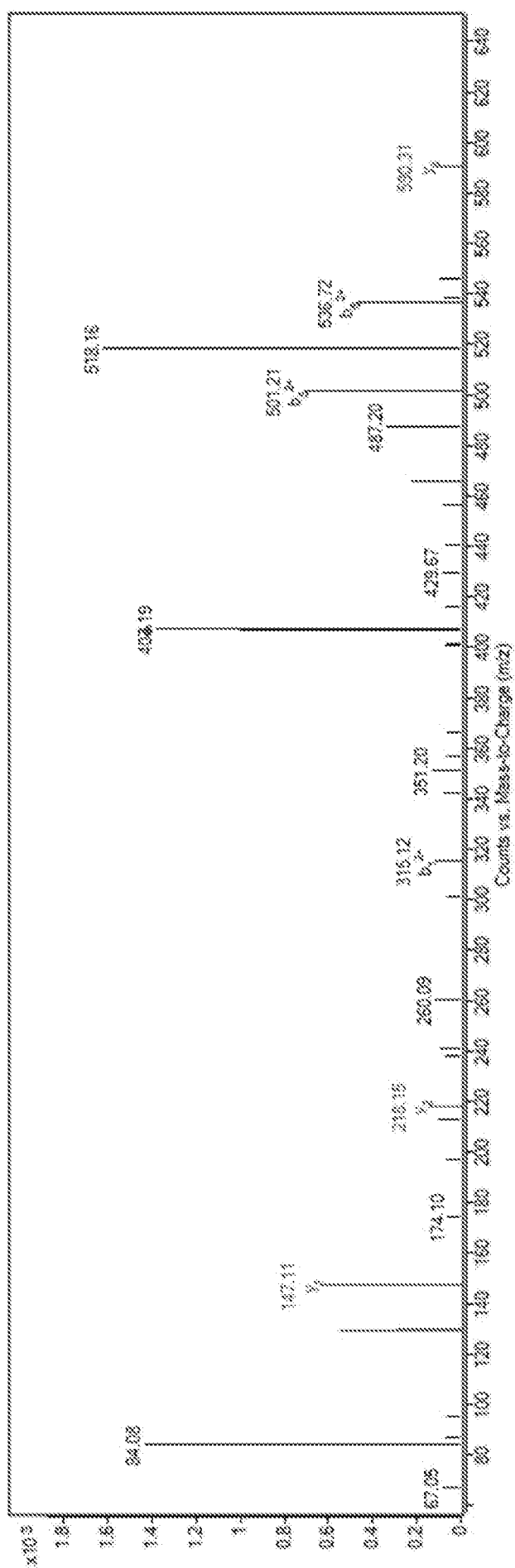
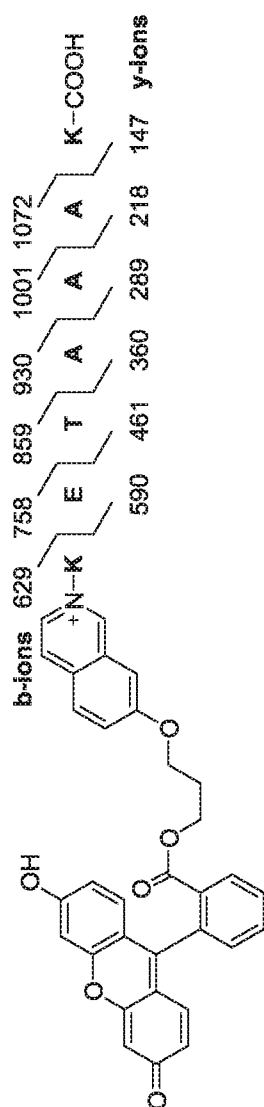


FIG. 12F

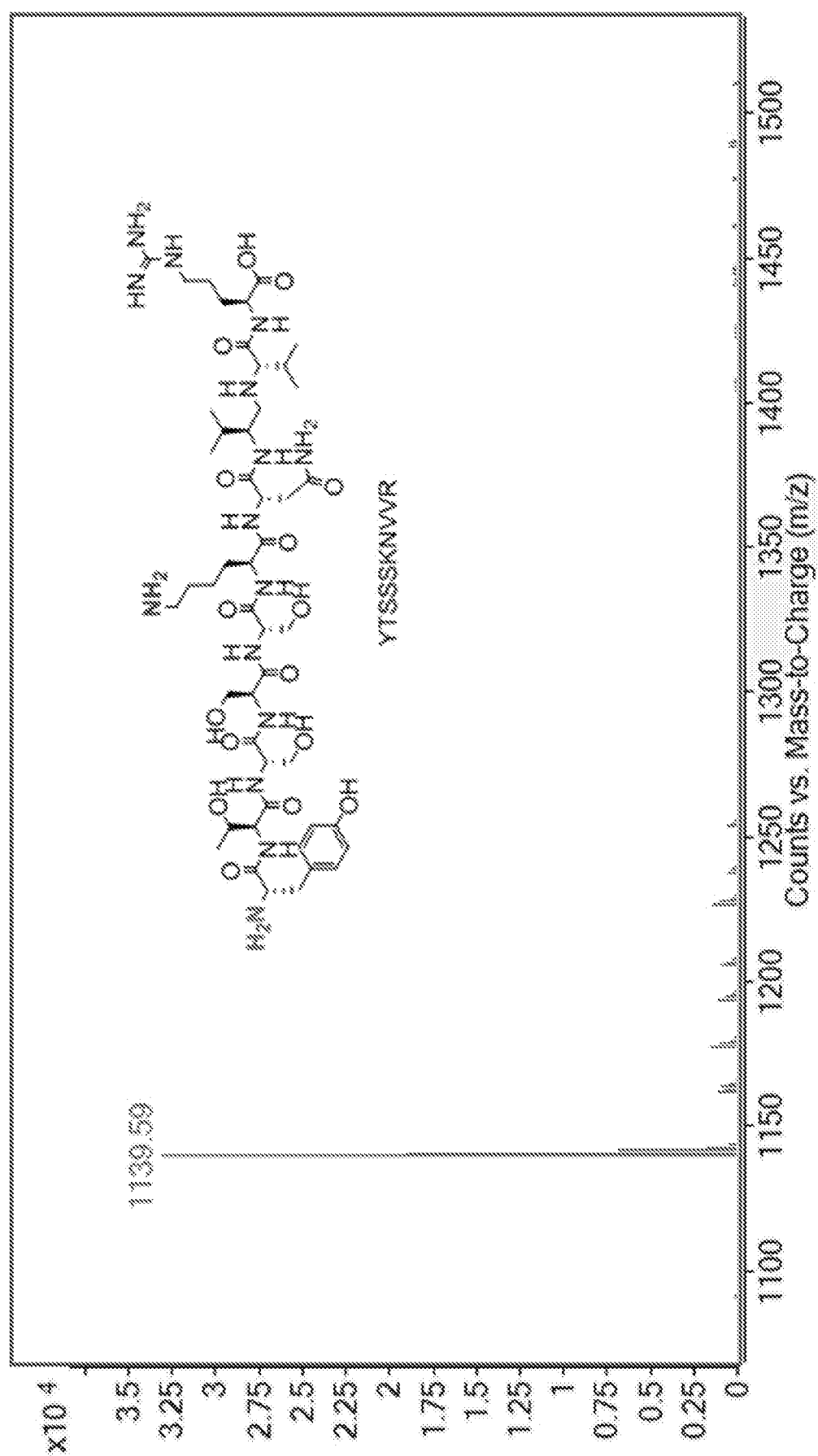


FIG. 13A

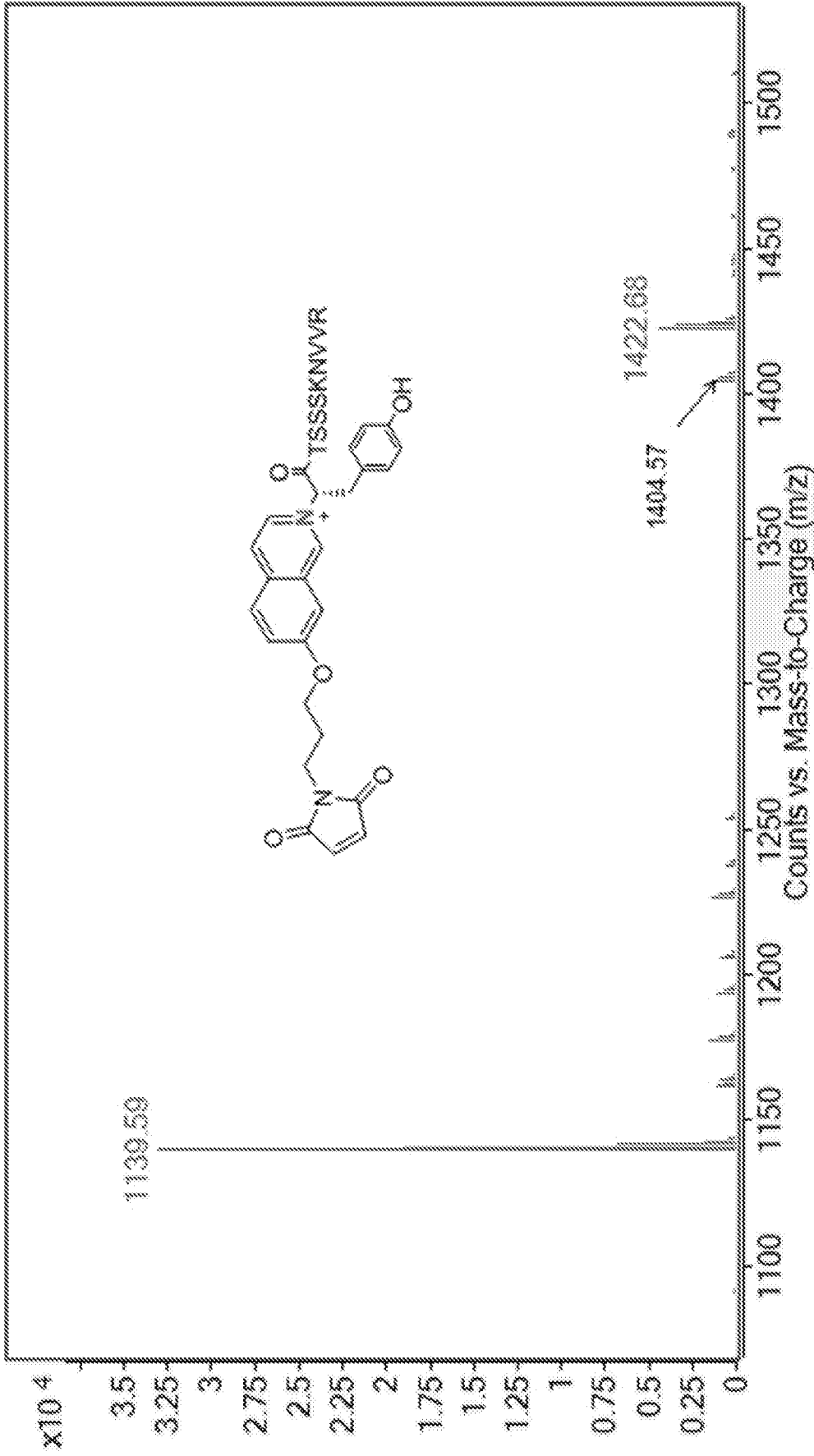


FIG. 13B

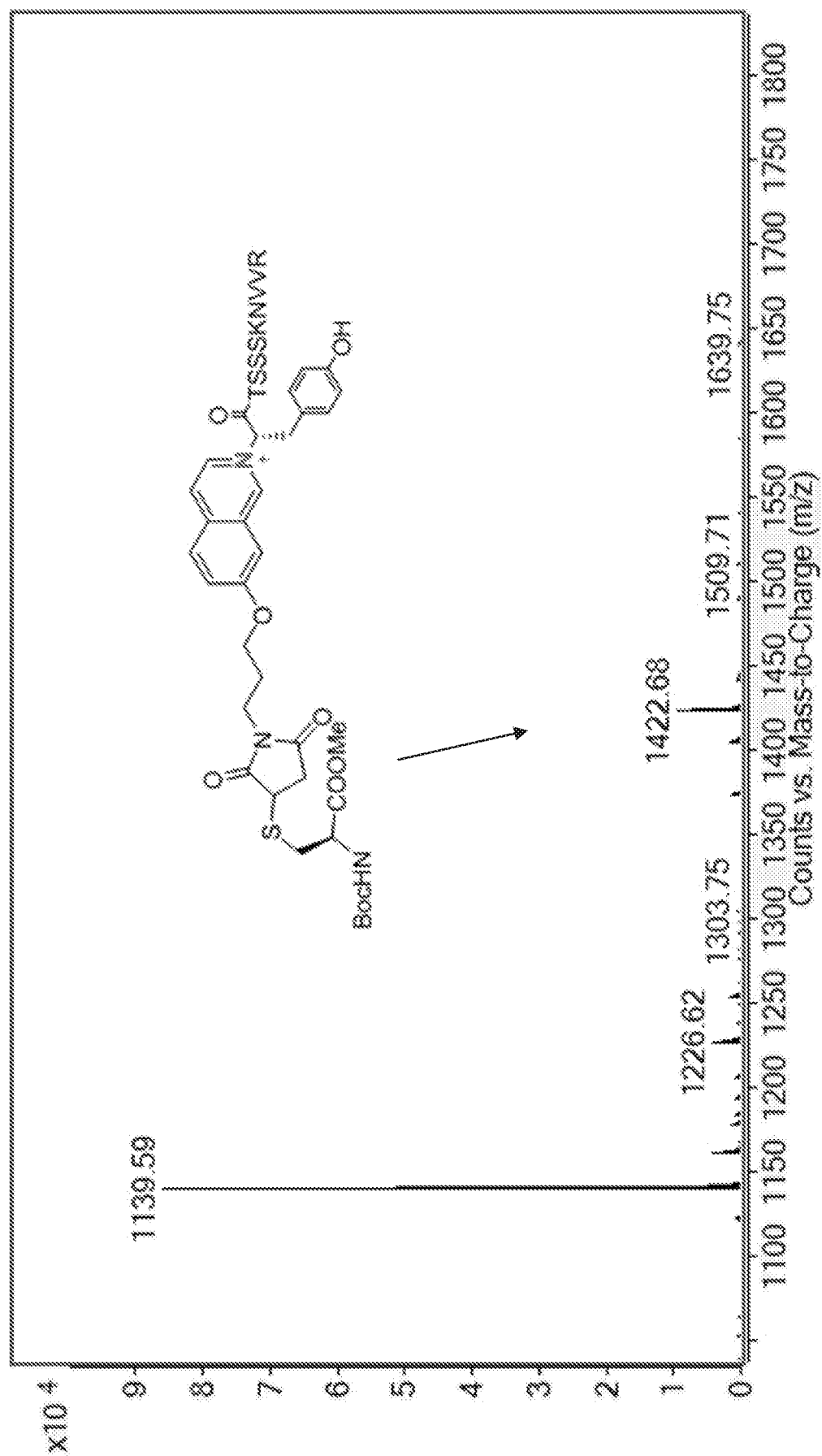


FIG. 13C

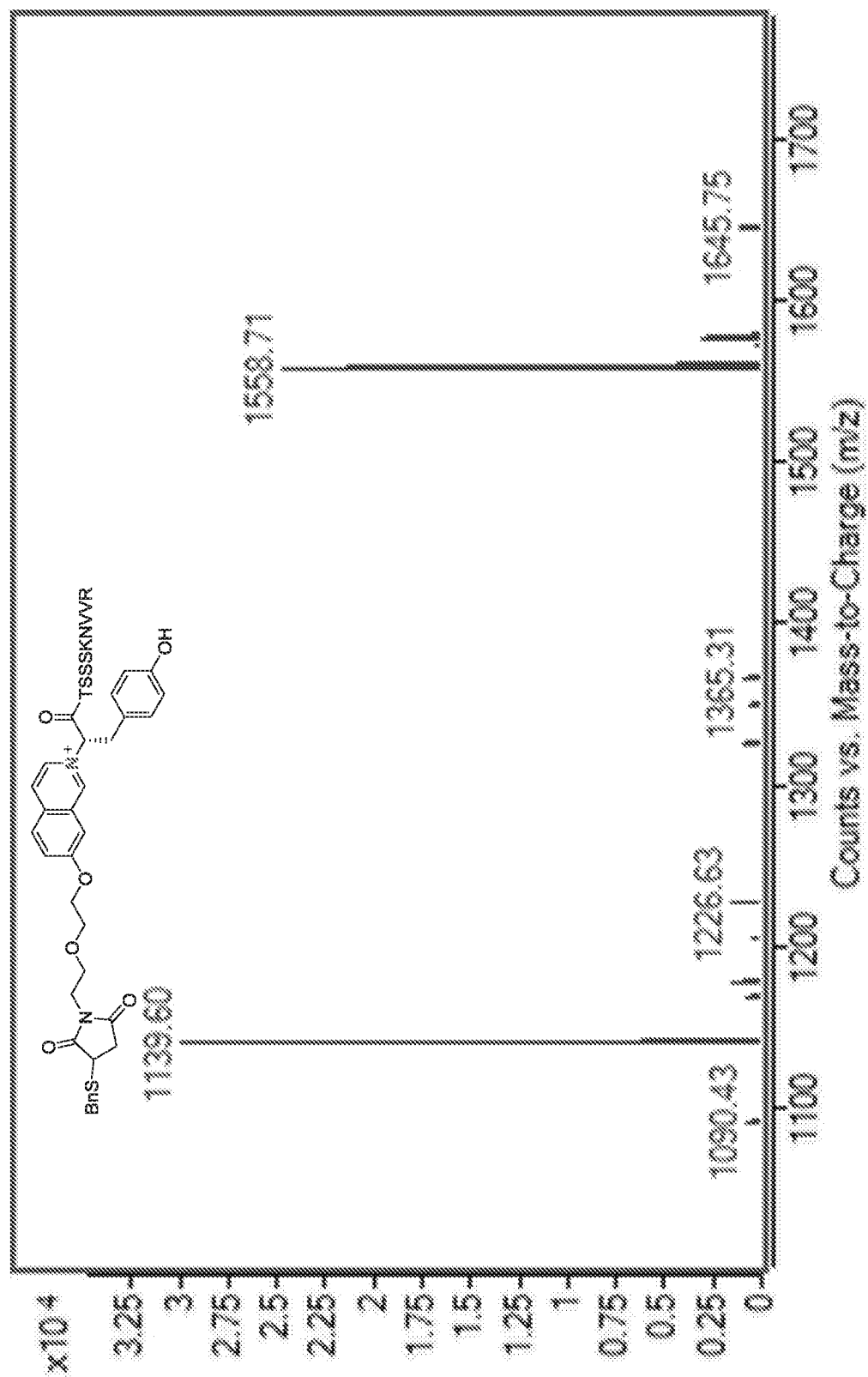


FIG. 13D

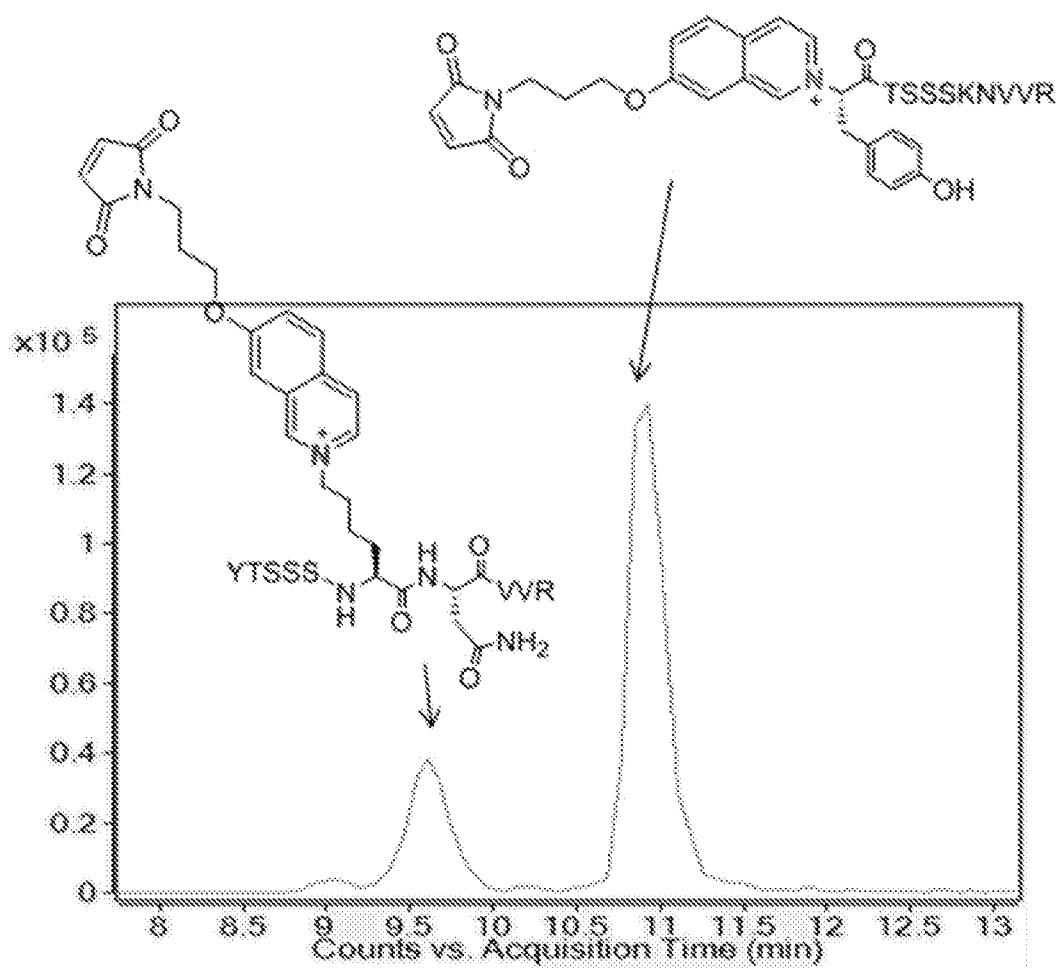


FIG. 14A

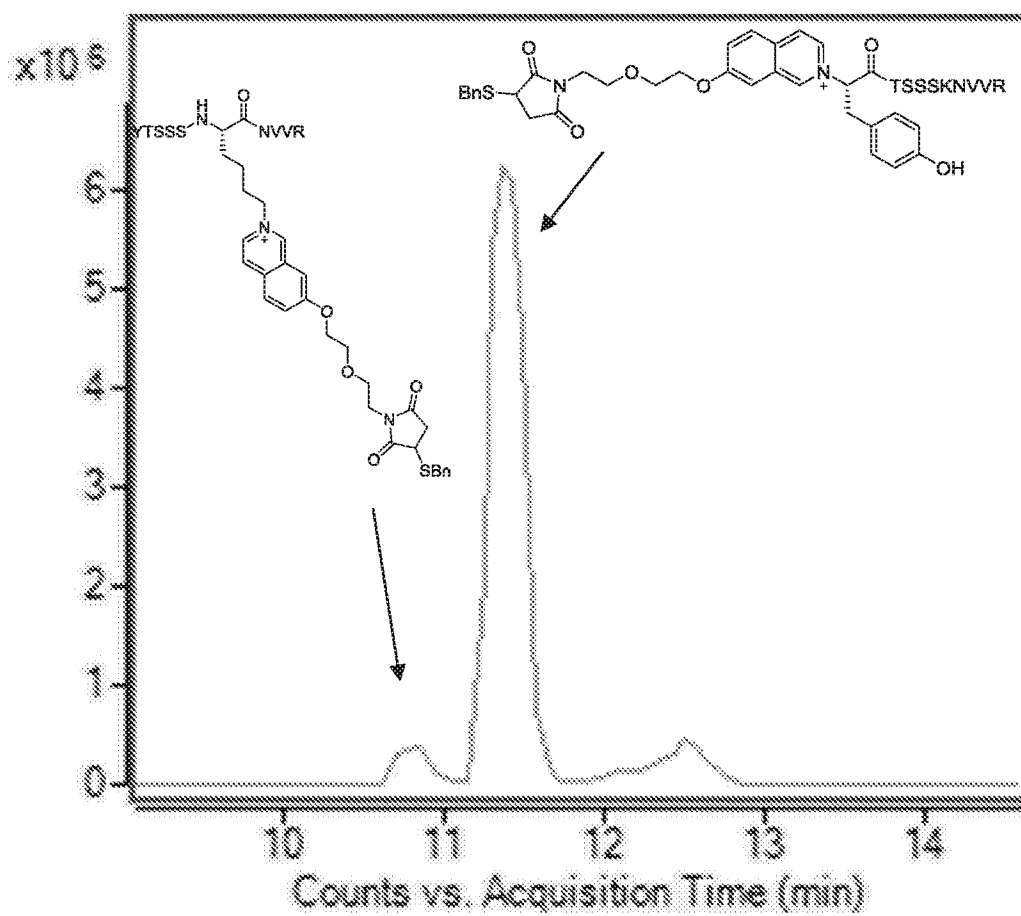


FIG. 14B

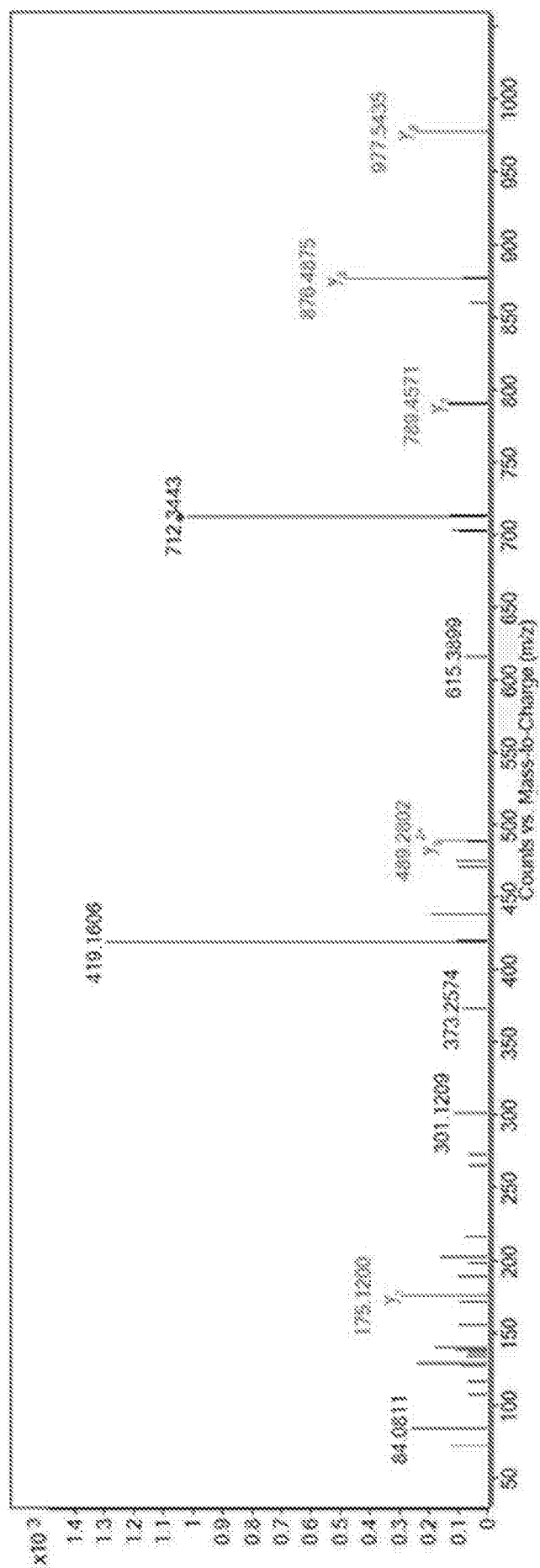
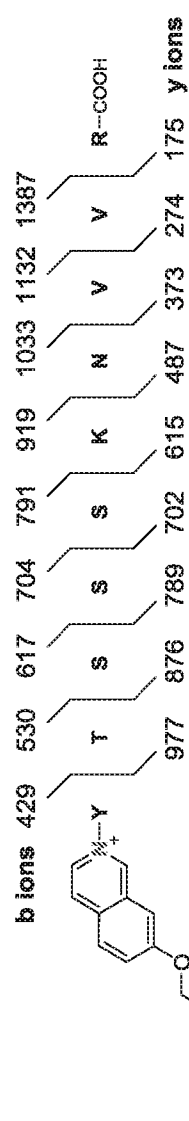


FIG. 15A

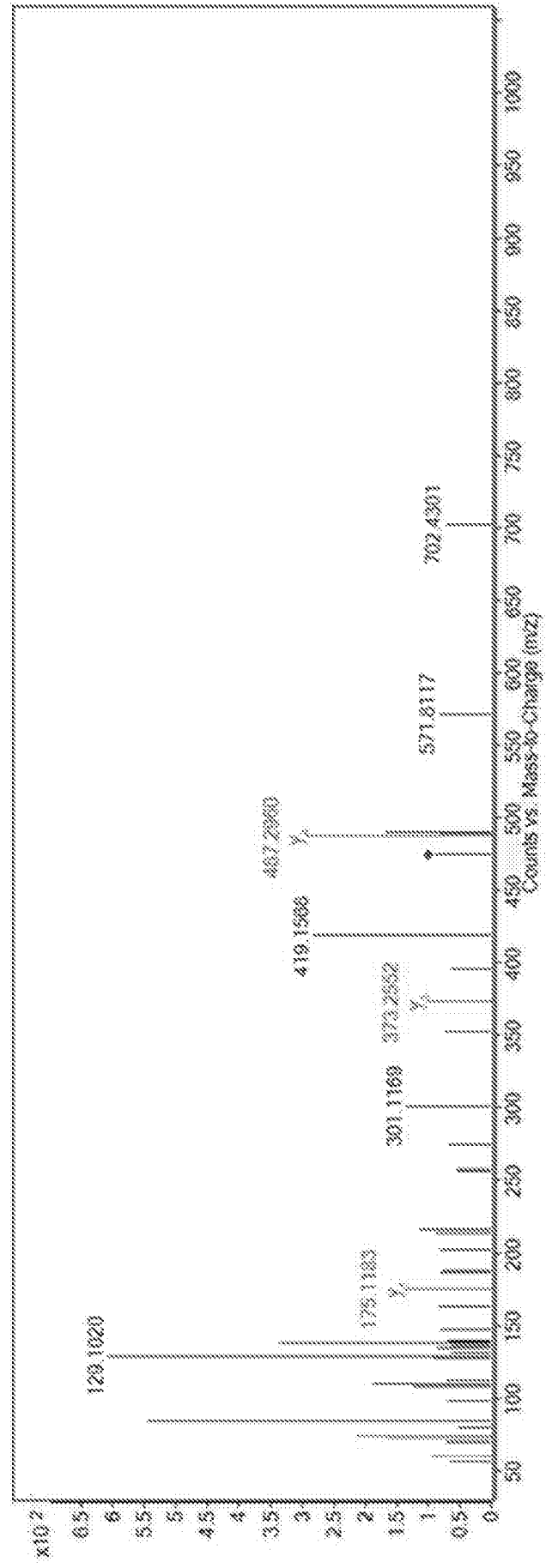
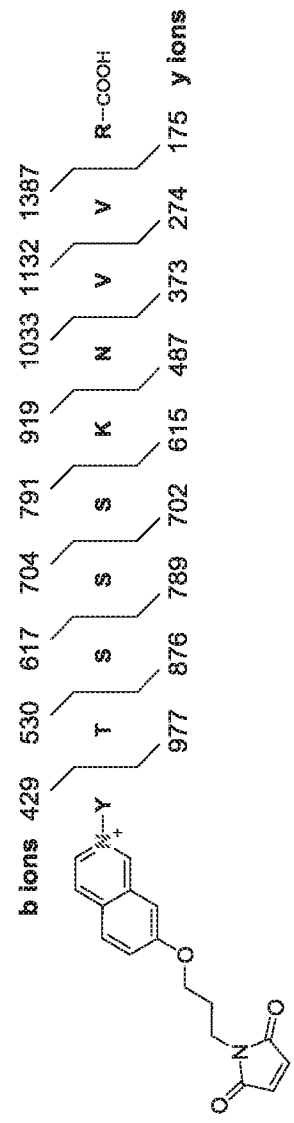


FIG. 15B

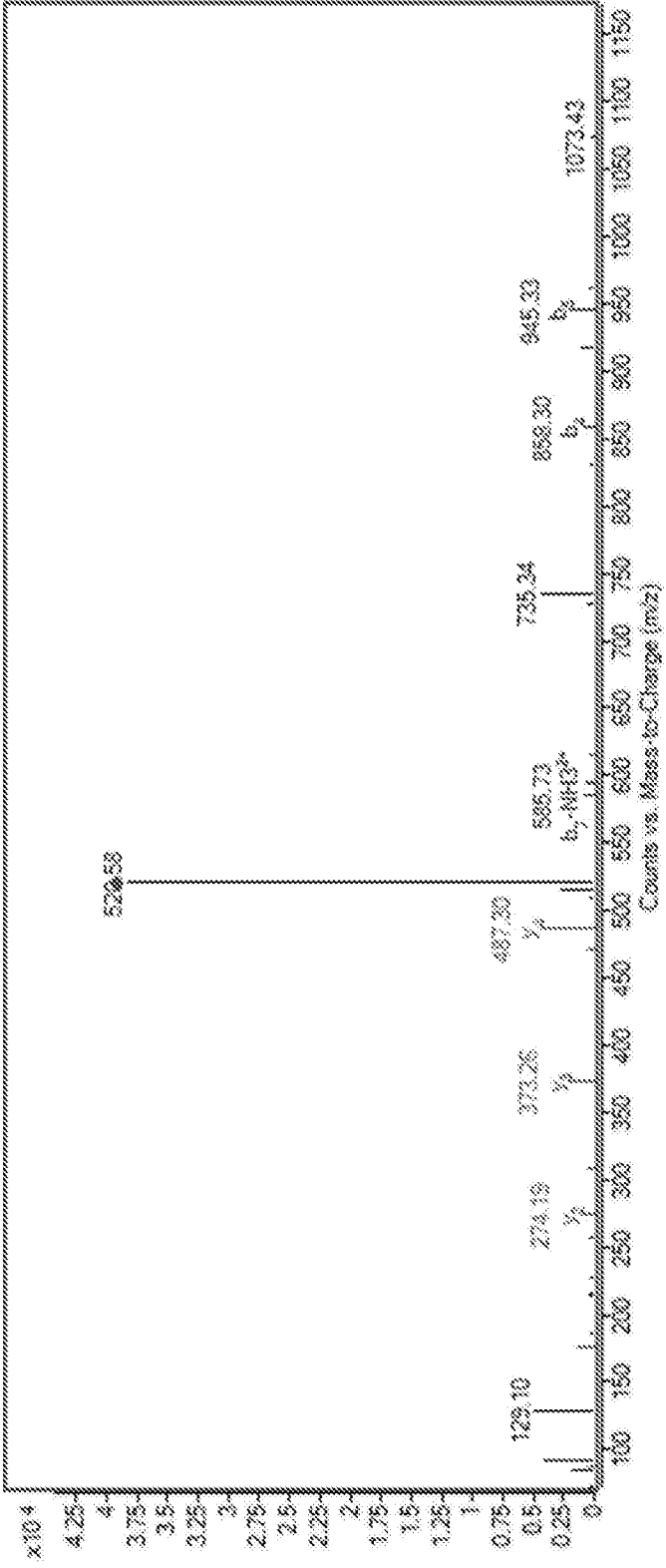
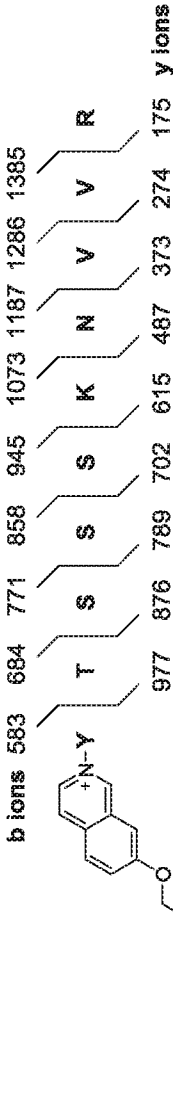


FIG. 15C

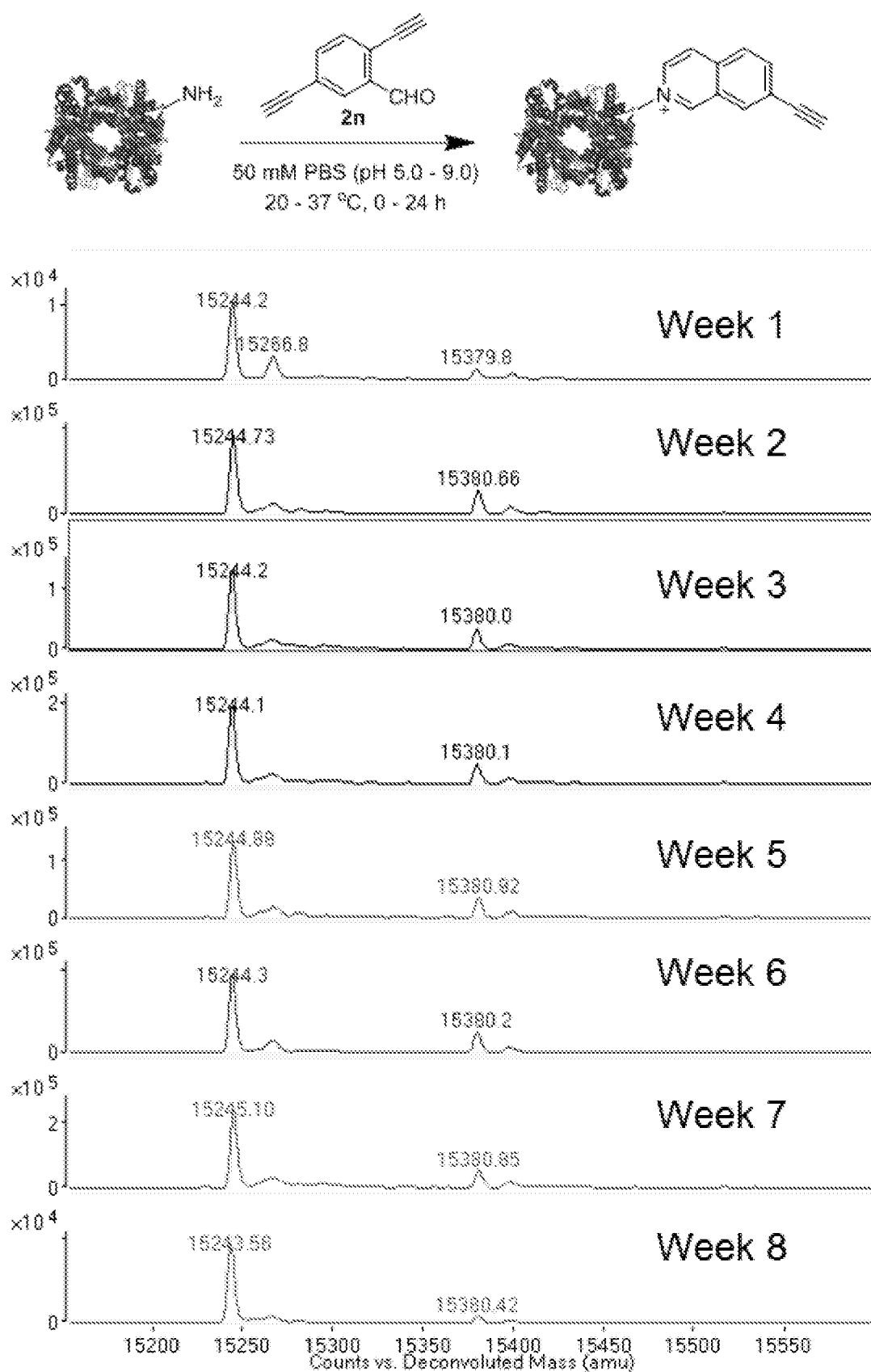


FIG. 16

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/086825

A. CLASSIFICATION OF SUBJECT MATTER

C07K 1/00(2006.01)i; C07D 217/00(2006.01)i

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)

C07K; C07D

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

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

CNABS、CPRS、DWPI、SIPOABS、CNKI、NCBI、ISI Web of Science: isoquinol+, benzo(c)pyridine, amino acid, peptide, polypeptide, protein

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 102477066 A (UNIV CAPITAL MEDICAL) 30 May 2012 (2012-05-30) claims 1 and 2	1-28
A	WO 2008106104 A2 (MASSACHUSETTS INST TECHNOLOGY ET AL.) 04 September 2008 (2008-09-04) formula (XII)	1-28

☐ 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

05 June 2019

Date of mailing of the international search report

03 July 2019

Name and mailing address of the ISA/CN

National Intellectual Property Administration, PRC
6, Xitucheng Rd., Jimen Bridge, Haidian District, Beijing
100088
China

Authorized officer

WU, Yongqing

Facsimile No. (86-10)62019451

Telephone No. +86-10-62089161

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2019/086825

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	102477066	A	30 May 2012	CN	102477066	B	04 December 2013
WO	2008106104	A2	04 September 2008	US	2008206885	A1	28 August 2008
				WO	2008106104	A3	24 December 2008