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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
22 February 2024 (22.02.2024)

WIPO | PCT

(10) International Publication Number
WO 2024/039743 A1

(51) International Patent Classification:
A61K 51/04 (2006.01) *A61K 51/08* (2006.01)

(21) International Application Number:
PCT/US2023/030385

(22) International Filing Date:
16 August 2023 (16.08.2023)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/398,654 17 August 2022 (17.08.2022) US

(71) Applicant: **EMORY UNIVERSITY** [US/US]; Office of Technology Transfer, 1599 Clifton Road NE, 4th Floor, Atlanta, Georgia 30322 (US).

(72) Inventors: **RAJ, Monika**; c/o Emory University Office of Technology Transfer, 1599 Clifton Road NE, 4th Floor, Atlanta, Georgia 30322 (US). **TANG, Kuei**; c/o Emory University Office of Technology Transfer, 1599 Clifton Road NE, 4th Floor, Atlanta, Georgia 30322 (US).

(74) Agent: **MASON, James C.** et al.; Emory University, Office of Technology Transfer, 1599 Clifton Road NE, 4th Floor, Atlanta, Georgia 30322 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report (Art. 21(3))*

(54) Title: COMPOSITIONS AND METHODS FOR SELECTIVE LABELING OF PRIMARY AMIDES WITH IODOVINYL COMPOUNDS

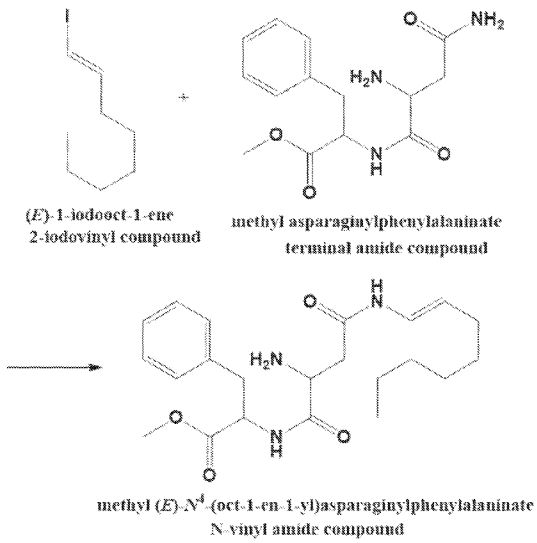


FIG. 1

(57) Abstract: This disclosure relates to compositions and methods for selective labeling of compounds containing primary amides. In certain embodiments, the primary amide is present in peptides containing the amino acids asparagine and/or glutamine. In certain embodiments, the primary amide is present on aliphatic or aromatic compounds or nucleic acids. In certain embodiments, this disclosure relates to methods of forming N-vinyl amide labeled peptides or compounds, comprising contacting a compound containing a primary amide group with iodovinyl compounds disclosed herein.

WO 2024/039743 A1

COMPOSITIONS AND METHODS FOR SELECTIVE LABELING OF PRIMARY AMIDES WITH IODOVINYL COMPOUNDS

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the benefit of U.S. Provisional Application No. 63/398,654 filed August 17, 2022. The entirety of this application is hereby incorporated by reference for all purposes.

BACKGROUND

Amino acid sequencing of naturally occurring proteins and antibodies may be accomplished by reconstructing segments of degraded and labeled protein fragments including the utilization of unique amino acid markers and size exclusion chromatography. Mass spectrometry can be helpful to identify sequences of smaller segments, but the technique becomes inherently less reliable as the protein segments become larger. Thus, there is a need to identify improved methods for identifying and sequencing proteins using selective amino acid labels.

Cao et al. report selective (15)N-labeling of the side-chain amide groups of asparagine and glutamine for applications in paramagnetic NMR spectroscopy. J Biomol NMR, 2014, 59:251–261.

Feng et al. report the stereoselective synthesis of vinyl iodides through copper(I)-catalyzed Finkelstein-type halide-exchange reaction. Synthesis, 2017, 49, 2727–2732.

Tang et al. report a one-step azolation strategy for site- and chemo-selective labeling of proteins with mass-sensitive probes. Angew Chem Int Ed, 2021, 60, 1797–1805.

Tang et al. report tunable amine-reactive electrophiles for selective profiling of lysine. Angew Chem Int Ed, 2022, 61, e202112107.

References cited herein are not an admission of prior art.

SUMMARY

This disclosure relates to compositions and methods for selective labeling of compounds containing primary amides. In certain embodiments, the primary amide is present in peptides containing the amino acids asparagine and/or glutamine. In certain embodiments, the primary amide is present on aliphatic or aromatic compounds or nucleic acids. In certain embodiments, this disclosure relates to methods of forming N-vinyl amide labeled peptides or compounds,

comprising contacting a compound containing a primary amide group with iodovinyl compounds disclosed herein.

In certain embodiments, this disclosure relates to methods of labeling a primary amide compound comprising contacting a primary amide compound with an iodovinyl compound, and a copper salt such that the primary amide compound reacts with the iodovinyl compound providing a labeled compound with a N-vinyl amide compound formed in place of the iodo group on the iodovinyl compound. In certain embodiments, the copper salt is copper iodine. In certain embodiments, the reaction further comprises a copper chelating ligand. In certain embodiments, the copper chelating ligand is N,N′-dimethylethylenediamine. In certain embodiments, the reaction further comprises a cesium salt. In certain embodiments, the cesium salt is cesium carbonate.

In certain embodiments, contacting a primary amide compound with an iodovinyl compound is in an aqueous pH buffered solution. In certain embodiments, the aqueous pH buffered solution contains is a phosphate buffer.

In certain embodiments, the iodovinyl compound comprises a label. In certain embodiments, the label is an aromatic molecule, a fluorescent dye, a ligand, biotin, a receptor, an antibody, or an antigen.

In certain embodiments, this disclosure relates to compounds, compositions, and materials comprising or coated with compounds disclosed herein.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWINGS

Figure 1 illustrates a method of labeling a primary amide compound comprising contacting a primary amide compound with an iodovinyl compound and copper salt such that the primary amide compound reacts with the iodovinyl compound providing a labeled compound with a N-vinyl amide compound formed in place of the iodo group. Benzene acts as a label as benzene is aromatic thus detectable using ultraviolet light. The dipeptide asparagine-phenylalanine has a terminal amine on the N-terminal asparagine and a methyl ester on the C-terminal phenylalanine. The amino acid side chain of asparagine contains a primary amide which is substituted with a vinyl group in place of the iodo group providing a N-vinyl amide.

Figure 2 illustrates a method wherein an iodovinyl compound is used to label a peptide with a primary amide providing a N-vinyl amide compound with an alkynyl group that allows for

reaction with solid surface coated with an azide providing a heterocyclic linking group attached the solid surface due to an azide-alkyne cycloaddition reaction.

DETAILED DISCUSSION

Before the present disclosure is described in greater detail, it is to be understood that this disclosure is not limited to particular embodiments described, and as such may, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to be limiting, since the scope of the present disclosure will be limited only by the appended claims.

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure belongs. Although any methods and materials similar or equivalent to those described herein can also be used in the practice or testing of the present disclosure, the preferred methods and materials are now described.

All publications and patents cited in this specification are herein incorporated by reference as if each individual publication or patent were specifically and individually indicated to be incorporated by reference and are incorporated herein by reference to disclose and describe the methods and/or materials in connection with which the publications are cited.

As will be apparent to those of skill in the art upon reading this disclosure, each of the individual embodiments described and illustrated herein has discrete components and features which may be readily separated from or combined with the features of any of the other several embodiments without departing from the scope or spirit of the present disclosure. Any recited method can be carried out in the order of events recited or in any other order that is logically possible.

An "embodiment" of this disclosure indicates that it is an example and not necessarily limited to such example. Embodiments of the present disclosure will employ, unless otherwise indicated, techniques of medicine, organic chemistry, biochemistry, molecular biology, pharmacology, and the like, which are within the skill of the art. Such techniques are explained fully in the literature.

It must be noted that, as used in the specification and the appended claims, the singular forms “a,” “an,” and “the” include plural referents unless the context clearly dictates otherwise. In

this specification and in the claims that follow reference will be made to a number of terms that shall be defined to have the following meanings unless a contrary intention is apparent.

As used in this disclosure and claim(s), the words "comprising" (and any form of comprising, such as "comprise" and "comprises"), "having" (and any form of having, such as "have" and "has"), "including" (and any form of including, such as "includes" and "include") or "containing" (and any form of containing, such as "contains" and "contain") have the meaning ascribed to them in U.S. Patent law in that they are inclusive or open-ended and do not exclude additional, unrecited elements or method steps.

"Consisting essentially of" or "consists of" or the like, when applied to methods and compositions encompassed by the present disclosure refers to compositions like those disclosed herein that exclude certain prior art elements to provide an inventive feature of a claim, but which may contain additional composition components or method steps, etc., that do not materially affect the basic and novel characteristic(s) of the compositions or methods.

The terms "protein," "peptide," and "polypeptide" refer to compounds comprising amino acids joined via peptide bonds and are used interchangeably. As used herein, where "amino acid sequence" is recited herein to refer to an amino acid sequence of a protein molecule. An "amino acid sequence" can be deduced from the nucleic acid sequence encoding the protein. However, terms such as "peptide" or "protein" are not meant to limit be limited to natural amino acids. The term includes naturally and non-naturally derived material optionally having naturally or non-naturally occurring amino acids and modifications such as, substitutions, glycosylations, and addition of hydrophilic or lipophilic moieties. In certain embodiments, the protein/peptide/polypeptide comprises more than three, four, five, six, seven, eight, nine, or ten amino acids.

The term "nucleobase polymer" refers to a polymer comprising nitrogen containing aromatic or heterocyclic bases that bind to naturally occurring nucleic acids through hydrogen bonding otherwise known as base pairing. A typical nucleobase polymer is a nucleic acid, RNA, DNA, or chemically modified form thereof. A nucleobase polymer may contain DNA or RNA or a combination of DNA or RNA nucleotides or may be single or double stranded or both, e.g., they may contain overhangs, hairpins, bends, etc. Nucleobase polymers may contain naturally occurring or synthetically modified bases and backbones.

As used herein, the term “conjugated” refers to linking molecular entities through covalent bonds, or by other specific binding interactions, such as due to hydrogen bonding and other van der Walls forces. The force to break a covalent bond is high, e.g., about 1500 pN for a carbon-to-carbon bond. The force to break a combination of strong protein interactions is typically a magnitude less, e.g., biotin to streptavidin is about 150 pN. Thus, a skilled artisan would understand that conjugation must be strong enough to bind molecular entities in order to implement the intended results.

A "linking group" refers to any variety of molecular arrangements that can be used to bridge or conjugate molecular moieties together. An example formula may be -R_n- wherein R is selected individually and independently at each occurrence as: -CR_nR_n-, -CHR_n-, -CH-, -C-, -CH_2-, -$C(OH)R_n$, -C(OH)(OH)-, -C(OH)H, -$C(Hal)R_n$-, -C(Hal)(Hal)-, -C(Hal)H-, -$C(N_3)R_n$-, -$C(CN)R_n$-, -C(CN)(CN)-, -C(CN)H-, -$C(N_3)(N_3)$-, -$C(N_3)H$-, -O-, -S-, -N-, -NH-, -NR_n-, -(C=O)-, -(C=NH)-, -(C=S)-, -$(C=CH_2)$-, which may contain single, double, or triple bonds individually and independently between the R groups. If an R is branched with an R_n it may be terminated with a group such as -CH_3, -H, -$CH=CH_2$, -CCH, -OH, -SH, -NH_2, -N_3, -CN, or -Hal, or two branched Rs may form an aromatic or non-aromatic cyclic structure. It is contemplated that in certain instances, the total Rs or “n” may be less than 100 or 50 or 25 or 10. Examples of linking groups include bridging alkyl groups, alkoxyalkyl, and aromatic groups.

The term "specific binding agent" refers to a molecule, such as a proteinaceous molecule, that binds a target molecule with a greater affinity than other random molecules or proteins. Examples of specific binding agents include antibodies that bind an epitope of an antigen or a receptor which binds a ligand. "Specifically binds" refers to the ability of a specific binding agent (such as a ligand, receptor, enzyme, antibody or binding region/fragment thereof) to recognize and bind a target molecule or polypeptide, such that its affinity (as determined by, e.g., affinity ELISA or other assays) is at least 10 times as great, but optionally 50 times as great, 100, 250 or 500 times as great, or even at least 1000 times as great as the affinity of the same for any other or other random molecule or polypeptide.

As used herein, the term “ligand” refers to any organic molecule, i.e., substantially comprised of carbon, hydrogen, and oxygen, that specifically binds to a “receptor.” Receptors are organic molecules typically found on the surface of a cell. Through binding a ligand to a receptor, the cell has a signal of the extra cellular environment which may cause changes inside the cell. As

a convention, a ligand is usually used to refer to the smaller of the binding partners from a size standpoint, and a receptor is usually used to refer to a molecule that spatially surrounds the ligand or portion thereof. However as used herein, the terms can be used interchangeably as they generally refer to molecules that are specific binding partners. For example, a glycan may be expressed on a cell surface glycoprotein and a lectin protein may bind the glycan. As the glycan is typically smaller and surrounded by the lectin protein during binding, it may be considered a ligand even though it is a receptor of the lectin binding signal on the cell surface. An antibody may be a receptor, and the epitope may be considered the ligand. In certain embodiments, a ligand is contemplated to be a compound that has a molecular weight of less than 500 or 1,000. In certain embodiments, a receptor is contemplated to be a protein-based compound that has a molecular weight of greater than 1,000, 2,000 or 5,000. In any of the embodiments disclosed herein, the position of a ligand and a receptor may be switched.

A "label" refers to a detectable compound or composition that is conjugated directly or indirectly to another molecule, such as an antibody or a protein, to facilitate detection of that molecule. Specific, non-limiting examples of labels include fluorescent tags, enzymatic linkages, and radioactive isotopes. In one example, a peptide "label" refers to incorporation of a heterologous polypeptide in the peptide, wherein the heterologous sequence can be identified by a specific binding agent, antibody, or bind to a metal such as nickel/ nitrilotriacetic acid, e.g., a poly-histidine sequence. Specific binding agents and metals can be conjugated to solid surfaces to facilitate purification methods. A label includes the incorporation of a radiolabeled amino acid or the covalent attachment of biotinyl moieties to a polypeptide that can be detected by marked avidin (for example, streptavidin containing a fluorescent marker or enzymatic activity that can be detected by optical or colorimetric methods). Various methods of labeling polypeptides and glycoproteins are known in the art and may be used. Examples of labels for polypeptides include, but are not limited to, the following: radioisotopes or radionucleotides (such as ^{35}S or ^{131}I), fluorescent labels (such as fluorescein isothiocyanate (FITC), rhodamine, lanthanide phosphors), enzymatic labels (such as horseradish peroxidase, beta-galactosidase, luciferase, alkaline phosphatase), chemiluminescent markers, biotinyl groups, predetermined polypeptide epitopes recognized by a secondary reporter (such as a leucine zipper pair sequences, binding sites for secondary antibodies, metal binding domains, epitope tags), or magnetic agents, such as

gadolinium chelates. In some embodiments, labels may be attached by spacer arms of various lengths to reduce potential steric hindrance.

In certain contexts, an "antibody" refers to a protein-based molecule that is naturally produced by animals in response to the presence of a protein or other molecule or that is not recognized by the animal's immune system to be a "self" molecule, i.e., recognized by the animal to be a foreign molecule and an antigen to the antibody. The immune system of the animal will create an antibody to specifically bind the antigen, and thereby targeting the antigen for elimination or degradation. It is well recognized by skilled artisans that the molecular structure of a natural antibody can be synthesized and altered by laboratory techniques. Recombinant engineering can be used to generate fully synthetic antibodies or fragments thereof providing control over variations of the amino acid sequences of the antibody. Thus, as used herein the term "antibody" is intended to include natural antibodies, monoclonal antibody, or non-naturally produced synthetic antibodies, and binding fragments, such as single chain binding fragments. These antibodies may have chemical modifications. The term "monoclonal antibodies" refers to a collection of antibodies encoded by the same nucleic acid molecule that are optionally produced by a single hybridoma (or clone thereof) or other cell line, or by a transgenic mammal such that each monoclonal antibody will typically recognize the same antigen. The term "monoclonal" is not limited to any particular method for making the antibody, nor is the term limited to antibodies produced in a particular species, e.g., mouse, rat, etc.

Hydrophilic polymers contain polar or charged functional groups, rendering them soluble in water. Examples include polyethylene glycol, polylactides, polyglycolide, poly(ε-caprolactone), poly(2-methoxyethyl acrylate), poly(tetrahydrofurfuryl acrylate), poly(2-methacryloyloxyethyl phosphorylcholine), poly(p-dioxanone), poly(serine methacrylate), poly[oligo(ethylene glycol) vinyl ether], poly{[2-(methacryloyloxy)ethyl], copolymers of ethylene glycol and propylene glycol, poly(oxyethylated polyol), poly(olefmic alcohol), poly(vinylpyrrolidone), poly(hydroxyalkylmethacrylamide), poly(hydroxyalkylmethacrylate), poly(saccharides), poly(alpha-hydroxy acid), and poly(vinyl alcohol). "PEG," "polyethylene glycol" and "poly(ethylene glycol)" refers to water-soluble poly(ethylene oxide). Typically, PEGs comprise the following structure "-(OCH_2CH_2)n-" where (n) is 2 to 4000.

As used herein, the term "derivative" refers to a structurally similar compound that retains sufficient functional attributes of the identified analogue. The derivative may be structurally

similar because it is lacking one or more atoms, substituted, a salt, in different hydration/oxidation states, or because one or more atoms within the molecule are switched, such as, but not limited to, replacing an oxygen atom with a sulfur atom, replacing an amino group with a hydroxyl group, replacing a nitrogen with a protonated carbon (CH) in an aromatic ring, replacing a bridging amino group (-NH-) with an oxy group (-O-), or vice versa. The derivative may be a prodrug. A derivative may be a polypeptide variant. Derivatives may be prepared by any variety of synthetic methods or appropriate adaptations presented in synthetic or organic chemistry textbooks, such as those provide in March's Advanced Organic Chemistry: Reactions, Mechanisms, and Structure, Wiley, 6th Edition (2007) Michael B. Smith or Domino Reactions in Organic Synthesis, Wiley (2006) Lutz F. Tietze hereby incorporated by reference.

The term "substituted" refers to a molecule wherein at least one hydrogen atom is replaced with a substituent. When substituted, one or more of the groups are "substituents." The molecule may be multiply substituted. In the case of an oxo substituent ("=O"), two hydrogen atoms are replaced. Example substituents within this context may include halogen, hydroxy, alkyl, alkoxy, nitro, cyano, carbocyclyl, carbocycloalkyl, heterocarbocyclyl, heterocarbocycloalkyl, aryl, arylalkyl, heteroaryl, heteroarylalkyl, $-NR_aR_b$, $-NR_aC(=O)R_b$, $-NR_aC(=O)NR_aNR_b$, $-NR_aC(=O)OR_b$, $-NR_aSO_2R_b$, $-C(=O)R_a$, $-C(=O)OR_a$, $-C(=O)NR_aR_b$, $-OC(=O)NR_aR_b$, $-OR_a$, $-SR_a$, $-SOR_a$, $-S(=O)_2R_a$, $-OS(=O)_2R_a$ and $-S(=O)_2OR_a$. R_a and R_b in this context may be the same or different and independently hydrogen, halogen hydroxyl, alkyl, alkoxy, alkyl, amino, alkylamino, dialkylamino, carbocyclyl, carbocycloalkyl, heterocarbocyclyl, heterocarbocycloalkyl, aryl, arylalkyl, heteroaryl, and heteroarylalkyl.

As used herein, "amide" refers to a carbonyl group covalently bonded to an amine, e.g., $-C(=O)NH_2$ or $-C(=O)NH-CH=CH_2$.

A "primary amide" refers to $-C(=O)NH_2$. Amino acids asparagine and glutamine contain primary amide side chains, i.e., $-CH_2C(=O)NH_2$ and $-CH_2CH_2C(=O)NH_2$ respectively. The carboxylic acid group on the C-terminus of a peptide may be converted by synthetic means to contain a primary amide referred to as a "C-terminal" amide.

A "bridging amide" refers to an amide that is not a primary amide, i.e., because the amide contains a substitution of one of the hydrogens connected to the amide nitrogen. Such is the case when amides link amino acids together in a peptide chain.

A "N-vinyl amide" refers to a bridging amide wherein a double bonded vinyl group is substituted on the amide at the nitrogen, e.g., -C(=O)NH-CH=CH_2.

An "iodovinyl " refers to a vinyl group substituted with the halogen, iodine, e.g., -CH=CHI.

As used herein, "alkyl" means a noncyclic straight chain or branched, unsaturated or saturated hydrocarbon such as those containing from 1 to 10 carbon atoms. Representative saturated straight chain alkyls include methyl, ethyl, n-propyl, n-butyl, n-pentyl, n-hexyl, n-septyl, n-octyl, n-nonyl, and the like; while saturated branched alkyls include isopropyl, sec-butyl, isobutyl, tert-butyl, isopentyl, and the like. Unsaturated alkyls contain at least one double or triple bond between adjacent carbon atoms (referred to as an "alkenyl" or "alkynyl", respectively). Representative straight chain and branched alkenyls include ethylenyl, propylenyl, 1-butenyl, 2-butenyl, isobutylenyl, 1-pentenyl, 2-pentenyl, 3 -methyl- 1-butenyl, 2-methyl-2-butenyl, 2,3-dimethyl-2-butenyl, and the like; while representative straight chain and branched alkynyls include acetylenyl, propynyl, 1-butynyl, 2-butynyl, 1-pentynyl, 2-pentynyl, 3- methyl-1-butynyl, and the like.

"Haloalkyl" refers to an alkyl group wherein one or more or all of the hydrogens are substituted with a halogen(s), e.g., -CH_2CH_2Cl or -CF_3.

"Alkylthio" refers to an alkyl group as defined above with the indicated number of carbon atoms attached through a sulfur bridge. An example of an alkylthio is methylthio, (e.g., -S-CH_3).

"Alkoxy" refers to an alkyl group as defined above with the indicated number of carbon atoms attached through an oxygen bridge. Examples of alkoxy include, but are not limited to, methoxy, ethoxy, n-propoxy, i-propoxy, n-butoxy, s-butoxy, t-butoxy, n- pentoxy, and s-pentoxy. Preferred alkoxy groups are methoxy, ethoxy, n-propoxy, i- propoxy, n-butoxy, s-butoxy, t-butoxy.

"Alkylamino" refers an alkyl group as defined above with the indicated number of carbon atoms attached through an amino bridge. An example of an alkylamino is methylamino, (e.g., -NH-CH_3).

"Alkanoyl" refers to an alkyl as defined above with the indicated number of carbon atoms attached through a carbonyl bride (e.g., -(C=O)alkyl).

"Alkylthio" refers to an alkyl group as defined above with the indicated number of carbon atoms attached through a sulfur bridge. An example of an alkylthio is methylthio, (e.g., -S-CH_3).

"Alkoxy" refers to an alkyl group as defined above with the indicated number of carbon atoms attached through an oxygen bridge. Examples of alkoxy include, but are not limited to, methoxy, ethoxy, n-propoxy, i-propoxy, n-butoxy, s-butoxy, t-butoxy, n- pentoxy, and s-pentoxy. Preferred alkoxy groups are methoxy, ethoxy, n-propoxy, i- propoxy, n-butoxy, s-butoxy, t-butoxy.

"Carboxyl" refers to a carboxylic acid and a "carboxyl ester" refers to an ester of the acid (e.g., -(C=O)O-alkyl).

"Alkylamino" refers an alkyl group as defined above with the indicated number of carbon atoms attached through an amino bridge. An example of an alkylamino is methylamino, (e.g., -NH-CH_3).

"Aryl" means an aromatic carbocyclic monocyclic or polycyclic ring such as phenyl or naphthyl. Polycyclic ring systems may, but are not required to, contain one or more non-aromatic rings, as long as one of the rings is aromatic.

Non-aromatic mono or polycyclic alkyls are referred to herein as "carbocycles" or "carbocyclyl" groups. Representative saturated carbocycles include cyclopropyl, cyclobutyl, cyclopentyl, cyclohexyl, and the like; while unsaturated carbocycles include cyclopentenyl and cyclohexenyl, and the like.

As used herein, "heterocycle" or "heterocyclyl" refers to mono- and polycyclic ring systems having 1 to 4 heteroatoms selected from nitrogen, oxygen, and sulfur, and containing at least 1 carbon atom. The mono- and polycyclic ring systems may be aromatic, non-aromatic or mixtures of aromatic and non-aromatic rings. Heterocycle includes heterocarbocycles, heteroaryls, and the like.

As used herein, "heteroaryl" or “heteroaromatic” refers an aromatic heterocarbocycle having 1 to 4 heteroatoms selected from nitrogen, oxygen and sulfur, and containing at least 1 carbon atom, including both mono- and polycyclic ring systems. Polycyclic ring systems may, but are not required to, contain one or more non-aromatic rings, as long as one of the rings is aromatic. Representative heteroaryls are furyl, benzofuranyl, thiophenyl, benzothiophenyl, pyrrolyl, indolyl, isoindolyl, azaindolyl, pyridyl, quinolinyl, isoquinolinyl, oxazolyl, isooxazolyl, benzoxazolyl, pyrazolyl, imidazolyl, benzimidazolyl, thiazolyl, benzothiazolyl, isothiazolyl, pyridazinyl, pyrimidinyl, pyrazinyl, triazinyl, cinnolinyl, phthalazinyl, and quinazolinyl. It is contemplated that

the use of the term "heteroaryl" includes N-alkylated derivatives such as a 1-methylimidazol- 5-yl substituent.

Methods and compositions for detecting primary amide groups on compounds, peptides containing asparagine and glutamine, and other biological molecules comprising primary amide groups

In certain embodiments, this disclosure relates to methods of forming N-vinyl amide labeled compounds, comprising contacting a compound containing a primary amide group with iodovinyl compounds.

In certain embodiments, this disclosure relates to methods of labeling a primary amide compound comprising contacting a primary amide compound with an iodovinyl compound, and copper salt such that the primary amide compound reacts with the iodovinyl compound providing a labeled compound with a N-vinyl amide compound formed in place of the iodo group.

In certain embodiments, the copper salt is a copper iodine. In certain embodiments, the reaction further comprises a copper chelating ligand. In certain embodiments, the copper chelating ligand is N,N′-dimethylethylenediamine. In certain embodiments, the copper chelating ligand is a bidentate ligand of nitrogen, oxygen (carboxylate), or a combination thereof.

In certain embodiments, reaction further comprises a cesium salt. In certain embodiments, cesium salt is cesium carbonate, cesium chloride, or cesium sulphate.

In certain embodiments, contacting a primary amide compound with an iodovinyl compound is in an aqueous pH buffered solution. In certain embodiments, the aqueous pH buffered solution contains is a phosphate buffer, sulfate buffer, or citrate buffer.

In certain embodiments, for any of the methods disclosed herein, reactions are in an aqueous solution with a pH of between 6.5 to 8.5. In certain embodiments, for any of the methods disclosed herein, reactions are in an aqueous solution with a pH of between 7.0 to 8.5. In certain embodiments, for any of the methods disclosed herein reactions are in an aqueous solution with a pH of between 7.0 to 8.0.

In certain embodiments, the primary amide compound is a peptide comprising an asparagine or glutamine or C-terminal amide of a peptide and methods provide a labeled N-vinyl (amide) asparagine, labeled N-vinyl (amide) glutamine, and/or a labeled N-vinyl C-terminal peptide.

In certain embodiments, the iodovinyl compound comprises a label. In certain embodiments, the label is an aromatic molecule, a fluorescent dye, a ligand, biotin, a receptor, an antibody, or an antigen.

In certain embodiments, this disclosure relates to compounds, compositions, and materials comprising or coated with compounds disclosed herein.

In certain embodiments, this disclosure relates to methods and compositions for labeling, isolating, detecting, measuring, and purifying compounds containing primary amides made by processes disclosed herein. In certain embodiments, the primary amide is asparagine, glutamine amino acid, or a C-terminal amide of a peptide. In certain embodiments, this disclosure relates to methods of labeling, isolating, detecting, measuring, purifying, quantifying, and amino acid sequencing compounds/peptides comprising an asparagine, glutamine amino acid, or a C-terminal amide of a peptide containing the same from a sample optionally utilizing solid supports. In certain embodiments, the sample is from a human subject. In certain embodiments, the sample is a blood sample, urine sample, stool sample, saliva sample, or sputum sample.

In certain embodiments, the peptide or compound for labeling having a primary amide is asparagine, glutamine, synthetic amino acid, C-terminal amide of a peptide, a nucleobase, nucleic acid, DNA, RNA, nucleobase polymer, alkyl, aryl, carbocyclyl, or heteroaryl optionally substituted with one or more, the same or different substituents such as a halogen, hydroxy, amino, thiol, alkyl, alkoxy, alkylamino, alkylthio, dialkylamino, acetamido, formyl, alkanoyl, carboxyl, carbonyl ester, carbamoyl, aryl, carbocyclyl or heterocyclyl, or N-substituted carbamoyl group, which is optionally further substituted or conjugated to a label or a solid support through a linking group.

In certain embodiments, for any of the methods disclosed herein, the methods further comprise determining the molecular weight or exact mass of the peptide or N-vinyl amide compound.

In certain embodiments, for any of the methods disclosed herein reactions are at about room temperature, e.g., between 10 and 30 degrees Celsius. In certain embodiments, for any of the methods disclosed herein reactions are at between about 5 and 40 degrees Celsius. In certain embodiments, for any of the methods disclosed herein reactions are at between about 35 and 55 degrees Celsius.

In certain embodiments, this disclosure relates to methods of contacting a compound or protein containing a primary amide, asparagine, or glutamine wherein the compound or protein containing a primary amide, asparagine or glutamine is in a sample such as a biological sample, (e.g., cell, tissue, etc.) or environmental sample. Biological samples may be obtained from animals (including humans) and encompass fluids, blood, solids, tissues, and gases. Environmental samples include environmental material such as surface matter, soil, water, and industrial samples.

In certain embodiments, for any of the method disclosed herein, the conditions for providing a N-vinyl amide labeled peptide or compound does not result in the labeling of bridging amides and/or terminal amines, including primary, secondary, nor tertiary amines, in the peptide compound or sample. In certain embodiments, for any of the methods disclosed herein, the conditions for providing a N-vinyl amide labeled peptide or compound does not result in the labeling of bridging amides in the peptide compound or sample.

In certain embodiments, the compound having a N-vinyl amide labeled peptide or compound is a benzene para substituted with an alkynyl group.

In certain embodiments, the compound having a N-vinyl amide labeled peptide or compound is further substituted with a label.

In certain embodiments, the label is biotin, aromatic molecule, a fluorescent dye, a second alkynyl group, ligand, receptor, antibody, or antigen.

In certain embodiments, the methods further comprise contacting the alkynyl labeled compound with a solid surface conjugated to a triazene under conditions such that a triazole compound is conjugated to the solid surface.

In certain embodiments, methods further comprise contacting the triazole compound conjugated to the solid surface with an acid solution such that the purified compound is cleaved from the solid surface providing purified compound.

In certain embodiments, the label is a ligand providing ligand labeled compound or peptide. In certain embodiments, the method further comprises contacting the ligand labeled or peptide with a solid surface conjugated to a receptor under conditions such that the ligand labeled compound or peptide is conjugated to the solid surface.

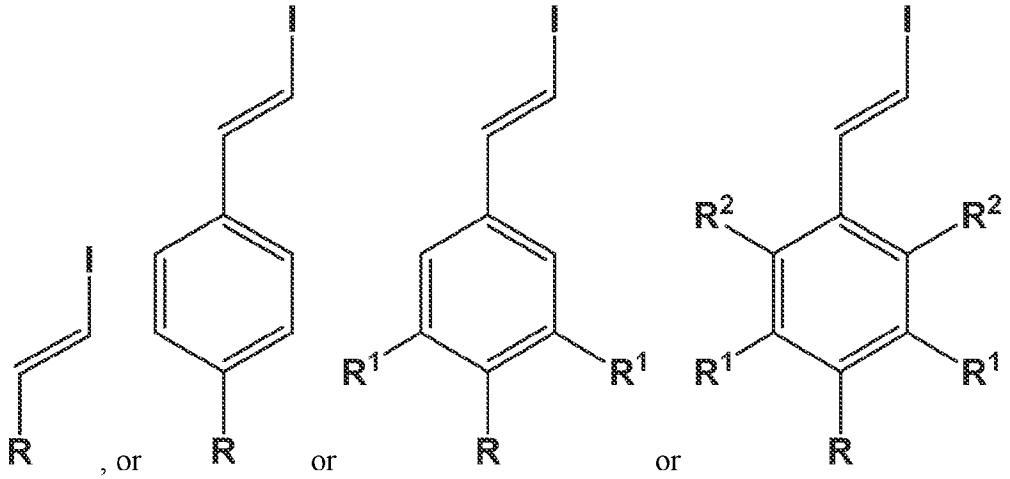
In certain embodiments, the label is biotin providing biotin labeled compound or peptide. In certain embodiments, the method further comprises contacting the biotin labeled compound or

peptide with a solid surface conjugated to avidin or streptavidin under conditions such that the biotin labeled compound or peptide is conjugated to the solid surface.

In certain embodiments, this disclosure relates to methods of forming N-vinyl amide labeled compounds comprising contacting a compound or peptide containing a primary amide with a compound comprising an aromatic group with iodovinyl group substituted with an alkynyl group under conditions such that a N-vinyl amide labeled compound or peptide is formed at the amide group. In certain embodiments, the aromatic group is benzene substitute with an iodovinyl group.

In certain embodiments, the N-vinyl amide labeled compound is an aliphatic amide, aromatic amide, phenyl amide, benzyl amide, asparagine, glutamine, peptide, other compound containing the same.

In certain embodiments, the iodovinyl compound comprising an aromatic group has the following formula,



or derivatives thereof, wherein R is optionally substituted to a label or conjugated to a solid support through a linking group. In certain embodiments, the linking group is a hydrophilic polymer. In certain embodiments, R is alkyl, aryl, carbocyclyl, or heterocyclyl, wherein R is optionally substituted with one or more substituents, or conjugated to a label or a solid support through a linking group. In certain embodiments, R^1 is hydrogen, alkyl, halogen, haloalkyl, alkoxy, alkylthio, dialkylamino, acetamido, nitrile, nitro, formyl, carboxyl, carbonyl ester, carbamoyl, or N-substituted carbamoyl group, wherein R^1 is optionally substituted. In certain embodiments, R^2 is hydrogen, alkyl, halogen, haloalkyl, alkoxy, alkylthio, dialkylamino, acetamido, nitrile, nitro, formyl, carboxyl, carbonyl ester, carbamoyl, or N-substituted carbamoyl group, wherein R^2 is

optionally substituted. In certain embodiments, R, R^1 or R^2 are individually and independently at each occurrence hydrogen, alkyl, halogen, haloalkyl, alkoxy, alkylthio, dialkylamino, acetamido, nitrile, nitro, formyl, carboxyl, carbonyl ester, carbamoyl, or N-substituted carbamoyl group, wherein R, R^1 or R^2 are optionally substituted e.g., conjugated to a label or a solid support through a linking group, or R^1 or R, or R^1 or R^2, and the attached atoms together form an aromatic or non-aromatic ring optionally substituted e.g., conjugated to a label or a solid support through a linking group.

In certain embodiments, R, R^1 or R^2 is a linking group comprising an alkynyl group or R is $-(C=O)NHR^3$, $-(C=O)OR^3$,$-(C=O)SR^3$, or $-(C=O)R^3$ wherein R^3 is an alkynyl group. In certain embodiments, R, R^1 or R^2 is a linking group comprising an alkynyl group or R, R^1 or R^2 is $-(C=O)NHR^3$, $-(C=O)OR^3$,$-(C=O)SR^3$, or $-(C=O)R^3$, wherein R^3 is an alkynyl group.

In certain embodiments, this disclosure relates to methods of purifying a compound comprising a primary amide in a sample comprising contacting the sample with a compound with an iodovinyl group substituted with an alkynyl group (e.g., 2-ethynylaniline) with a copper salt, providing a N-vinyl amide labeled compound at the iodovinyl group.

In certain embodiments, the N-vinyl amide compound substituted with an alkynyl group is formed by the process of contacting an iodovinyl compound a copper salt, a cesium salt, and a phosphate salt at pH of between 6.5 to 8.5.

In certain embodiments, this disclosure relates to methods of purifying a compound comprising a primary amide in a sample comprising contacting the sample with an iodovinyl compound, a copper salt, cesium salt, and a phosphate salt at pH of about 7 providing a compound with a N-vinyl bridging amide.

In certain embodiments, the N-vinyl amide (aromatic or phenyl) compound comprises a meta or para substituted group comprising an alkynyl group (e.g., N-(prop-2-yn-1-yl)benzamide) and the method further comprises contacting the N-vinyl amide labeled compound comprising an alkynyl group with a solid support conjugated to a triazene group under conditions such that the meta or para alkynyl group reacts with the triazene group on the solid support to form a triazole linkage providing a solid support conjugated to the N-vinyl labeled compound; washing the solid support to remove unreacted material from the sample providing a purified solid support conjugated to the N-vinyl labeled compound; and contacting the purified solid support conjugated

to the N-vinyl amide labeled compound to conditions providing a composition with a purified compound comprising a N-vinyl amide released from the solid support.

In certain embodiments, the conditions are acid conditions, e.g., exposure to trifluoroacetic acid.

In certain embodiments, the methods further comprise determining the molecular weight or exact mass of the N-vinyl amide labeled compound or peptide comprising an asparagine or glutamine amino acid.

In certain embodiments, the methods further comprise separating the composition with a purified N-vinyl amide into two or more compounds with a N-vinyl amide. In certain embodiments, separating is by chromatography.

In certain embodiments, for any of the methods disclosed herein the methods further comprise separating a purified, cleaved, or isolated peptide composition or N-vinyl amide compound composition into two or more peptides or labels compounds. In certain embodiments, separating is by chromatography, e.g., adsorption and/or size-based separations by chromatography or filtration through a gel.

In certain embodiments, methods comprise obtaining a sample from a subject comprising a protein, contacting the sample comprising a protein in a sample with a protease providing a cleaved protein, contacting the cleaved protein with a labeled iodovinyl compound providing a N-vinyl amide labeled cleaved protein in the sample, and purifying the N-vinyl amide labeled cleaved protein, e.g., by chromatography.

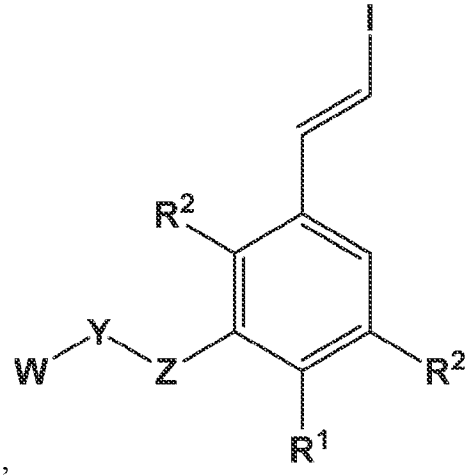
In certain embodiments, methods comprise obtaining a sample comprising a protein from a subject, contacting the sample comprising a protein in a sample with a labeled iodovinyl compound providing a N-vinyl amide labeled protein in the sample, contacting N-vinyl amide labeled protein with a protease providing a cleaved N-vinyl amide labeled protein, and purifying the cleaved N-vinyl amide labeled protein by chromatography.

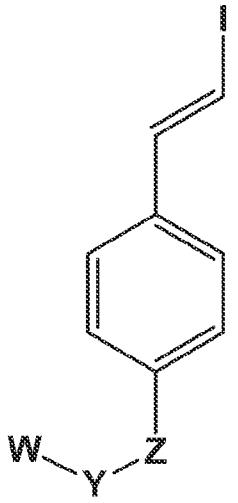
In certain embodiments, this disclosure relates to methods of purifying a compound with a primary amide in a sample comprising, contacting the sample with a solid support conjugated to a compound comprising an iodovinyl such that the compounds with a primary amide in the sample reacts with the iodovinyl providing a N-vinyl amide compound immobilized to the solid support; washing the solid support to remove unreacted material from the sample providing a purified N-vinyl compound immobilized to the sold support. In certain embodiments, the method further

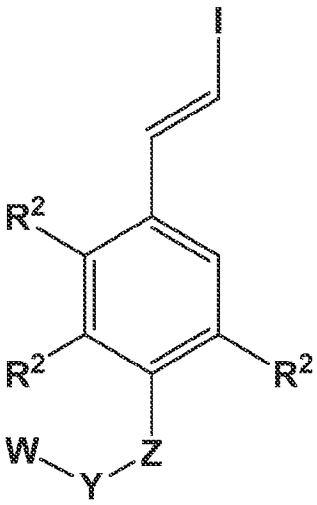
comprises exposing the purified compound immobilized to the solid support to conditions releasing a purified N-vinyl amide compound.

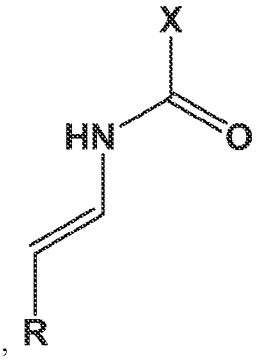
In certain embodiments, this disclosure relates to a labeled compound, peptide, or solid support such as a particle conjugated with a compound having the following formula:

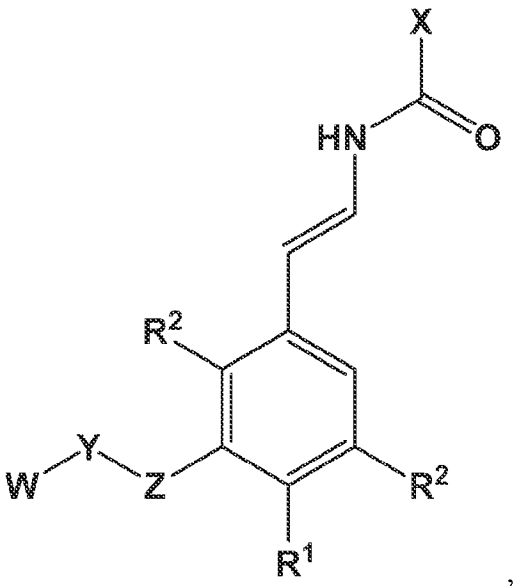


,


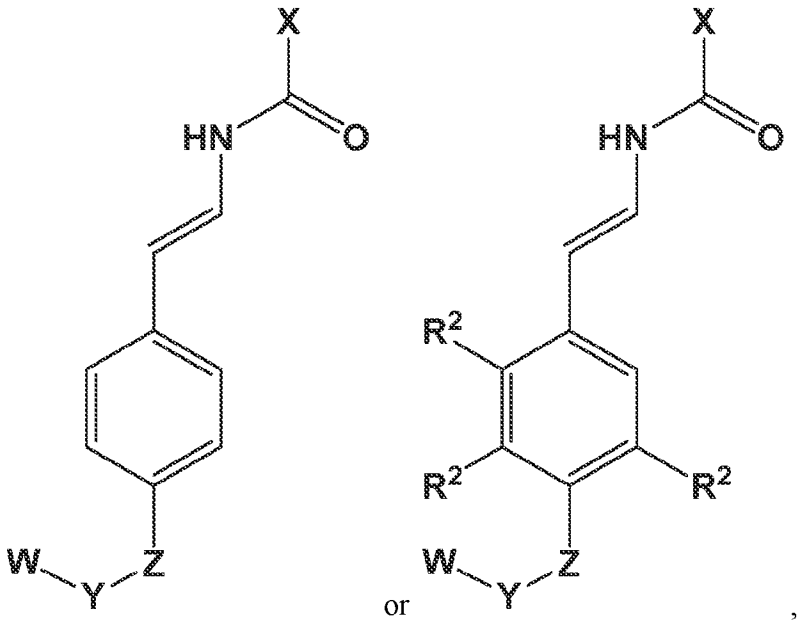
,


or


,


,


,



or derivative or salt thereof wherein,

W is a solid support such as a particle or a label;

X is a peptide;

Y is O, NH, or S;

Z is a linking group;

R is alkyl, aryl, carbocyclyl, or heterocyclyl optionally substituted with one or more substituents, or conjugated to a label or a solid support through a linking group;

R^1 is hydrogen, alkyl, halogen, haloalkyl, alkoxy, alkylthio, dialkylamino, acetamido, nitrile, nitro, formyl, carboxyl, carbonyl ester, carbamoyl, or N-substituted carbamoyl group, wherein R^1 is optionally substituted with one or more substituents, or conjugated to a label or a solid support through a linking group;

R^2 is individually and independently at each occurrence hydrogen, alkyl, halogen, haloalkyl, alkoxy, alkylthio, dialkylamino, acetamido, nitrile, nitro, formyl, carboxyl, carbonyl ester, carbamoyl, or N-substituted carbamoyl group, wherein R^2 is optionally substituted with one or more substituents, -Z-W, or -Z-Y-W, or two R^2s and the attached atoms come together to form an aromatic or non-aromatic ring that is optionally substituted with one or more substituents or -Z-W, or -Z-Y-W.

In certain embodiments, this disclosure relates to a compound or material comprising or coated with or conjugate to chemical arrangements disclosed herein.

In certain embodiments, the material, solid support, or particle further comprises or is coated with or conjugated to a hydrophilic polymer.

In certain embodiments, this disclosure contemplates that the solid support is a magnetic material, e.g., magnetic bead, and methods of purifying entail capturing the magnetic bead with a magnetic field thereby separating the magnetic bead and contents thereof from a solution or mixture.

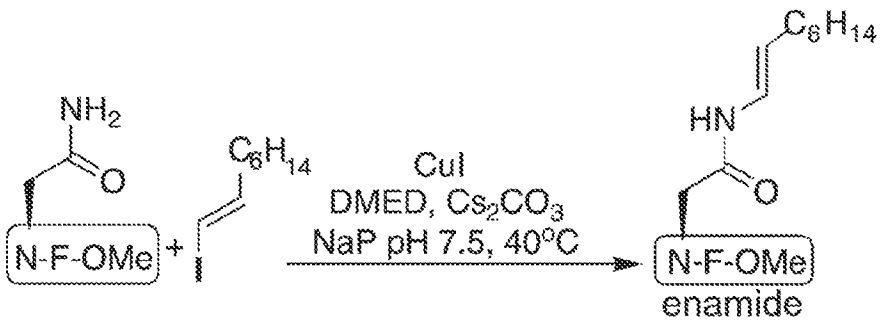
EXAMPLES

Methods for selective labeling of asparagine and glutamine

Methods for selective labeling of naturally occurring proteins containing asparagine and glutamine amino acid have been challenging. Asparagine and glutamine contain a primary amide side chain that is not reactive in typical nucleophilic reactions. Experiments were performed to identify a method for selective and robust labeling of asparagine and glutamine using vinyl iodide analogs and Cu catalyst. It is contemplated that method can be used with other protein-labeling techniques, in sequencing methods of peptides and proteins, and for activity-based protein profiling. The products produced by this method appear to be stable in acidic conditions.

The preparation of iodovinyl compounds can be accomplished by halide-exchange reactions of vinyl bromides with potassium iodide in the presence of a copper catalyst. See, e.g., Feng et al. Synthesis, 2017, 49, 2727-2732.

A general procedure for labeling primary amides



To 1 equiv. of peptide (N-F-OMe) in sodium phosphate buffer (10 mM, pH 7.5), was added 2 equiv. of vinyl iodide, 10 mol% of CuI and DMED and 1 equiv. of Cs_2O_3. The reaction was

stirred at 40 °C for 10h. The reaction was then purified by HPLC or silica gel column and % conversion to the product was determined by NMR % yield or by HPLC and LCMS.

It is contemplated that the C_6H_{14} alkyl group could be any R group (aliphatic or aromatic, alkyne and others).

5

CLAIMS

1. A method of labeling a primary amide compound comprising contacting a primary amide compound with an iodovinyl compound, and copper salt such that the primary amide compound reacts with the iodovinyl compound providing a labeled compound with a N-vinyl amide compound formed in place of the iodo group on the iodovinyl compound.

2. The method of claim 1, wherein the copper salt is a copper iodine.

3. The method of claim 1, wherein the reaction further comprises a cesium salt.

4. The method of claim 3, wherein the cesium salt is cesium carbonate.

5. The method of claim 4, wherein the reaction further comprises a copper chelating ligand.

6. The method of claim 5, wherein the copper chelating ligand is N,N′-dimethylethylenediamine.

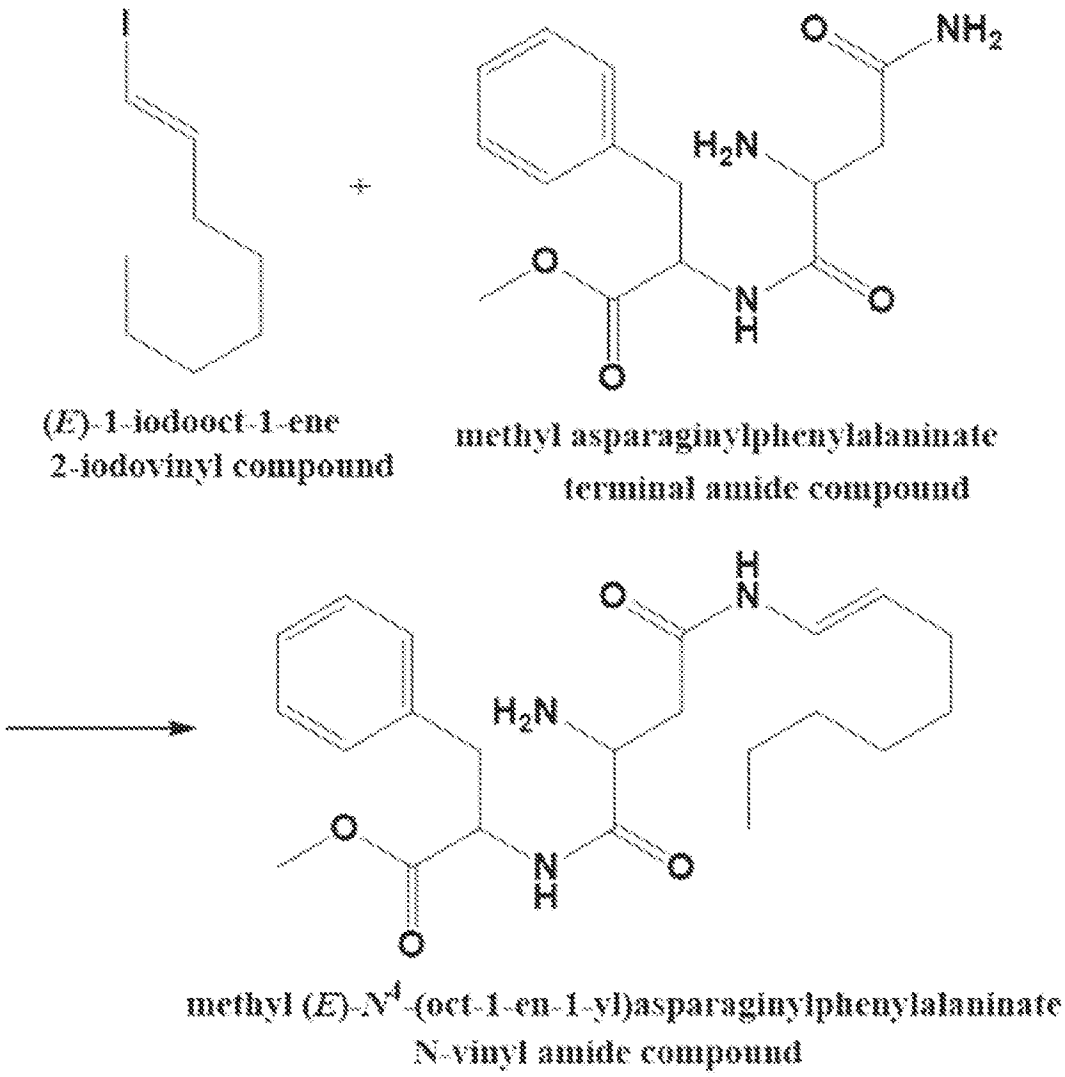
7. The method of claim 1, wherein contacting the primary amide compound with the iodovinyl compound is in an aqueous pH buffered solution.

8. The method of claim 7, wherein the aqueous pH buffered solution contains is a phosphate buffer.

9. The method of claim 7, wherein the primary amide compound is a peptide comprising an asparagine or glutamine amino acid.

10. The method of claim 7, wherein iodovinyl compound comprises a label.

11. The method of claim 10, wherein the label is an aromatic molecule, a fluorescent dye, a ligand, biotin, a receptor, an antibody, or an antigen.

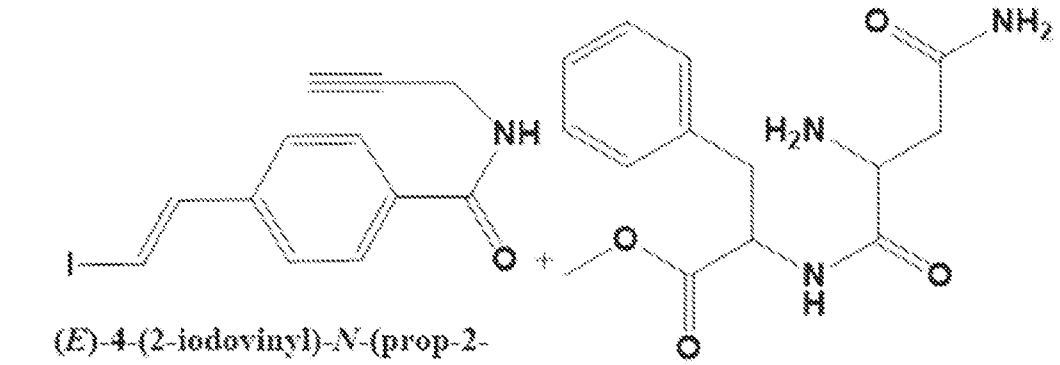


(*E*)-1-iodooct-1-ene
2-iodovinyl compound

methyl asparaginylphenylalaninate
terminal amide compound

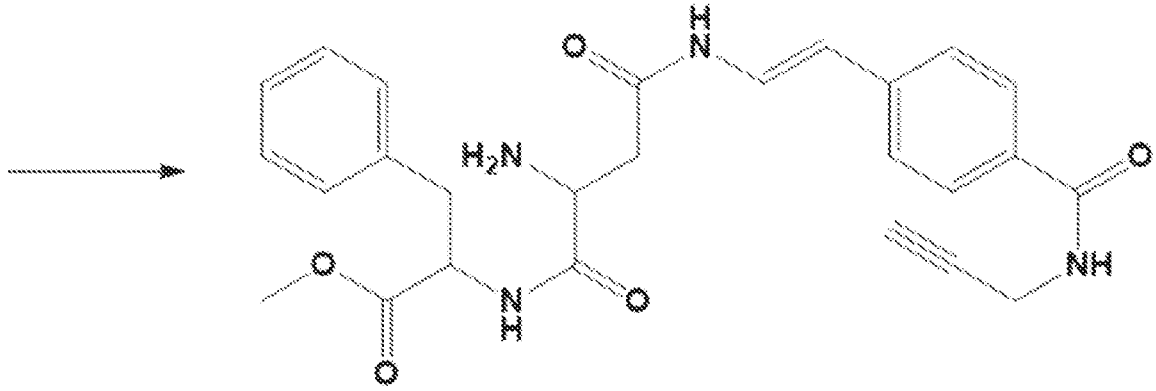
methyl (*E*)-N^4-(oct-1-en-1-yl)asparaginylphenylalaninate
N-vinyl amide compound

FIG. 1



(*E*)-4-(2-iodovinyl)-*N*-(prop-2-yn-1-yl)benzamide
2-iodovinyl compound

methyl asparaginylphenylalaninate
terminal amide compound



methyl (*E*)-N^4-(4-(prop-2-yn-1-ylcarbamoyl)styryl)asparaginylphenylalaninate
N-vinyl amide compound

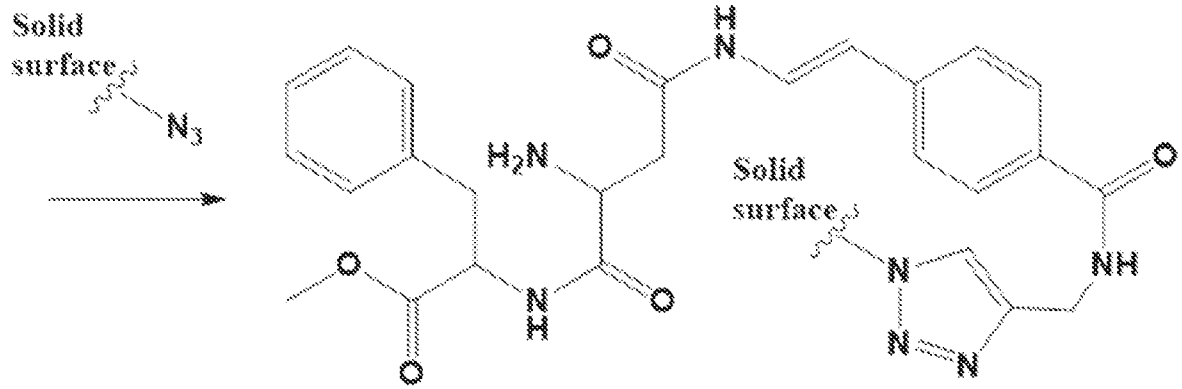


FIG. 2

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 23/30385

A. CLASSIFICATION OF SUBJECT MATTER

IPC - INV. A61K 51/04 (2023.01)
ADD. A61K 51/08 (2023.01)

CPC - INV. A61K 51/0482, A61K 51/0497, A61K 51/08

ADD. A61K 51/088

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
See Search History document

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

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	Wang et al. Copper-catalyzed cross-coupling of amino acid-derived amides with (Z)-vinyl iodides: Unexpected solvent effect and preparation of plocabulin. Tetrahedron 83 (2021) 131953, pg1-10; title; pg2, Table 1; pg3, Table 2	1-6 ------- 7-11
Y	Baqi. Recent Advances in Microwave-Assisted Copper-Catalyzed Cross-Coupling Reactions. Catalysts 2021, 11, 46; pg1-35; title; pg2, Scheme 1	7-11
A	Gille et al. Studies on the synthesis of peptides containing dehydrovaline and dehydroisoleucine based on copper-mediated enamide formation. Beilstein J. Org. Chem. 2016, 12, 564–570.	1-11
A	Ricard. Copper-Catalyzed ß-Iodovinylation of Carbamates: Expedient Access to Highly Functionalized Vinyl-Carbamates. ChemistrySelect 2018, 3, 4923 – 492	1-11
A	Theunissen et al. Synthesis of Ynamides by Copper-Mediated Coupling of 1,1-Dibromo-1-alkenes with Nitrogen Nucleophiles. Preparation of 4-Methyl-N-(2-phenylethynyl)-N-(phenylmethyl)benzenesulfonamide. Org. Synth. 2019, 96, 195-213	1-11

☐ 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
"D" document cited by the applicant in the international application
"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	Date of mailing of the international search report
31 October 2023	NOV 14 2023
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-8300	Authorized officer Kari Rodriquez Telephone No. PCT Helpdesk: 571-272-4300