



US 20230111954A1

(19) **United States**

(12) **Patent Application Publication**
Kwong et al.

(10) **Pub. No.: US 2023/0111954 A1**

(43) **Pub. Date: Apr. 13, 2023**

(54) **MULTIPLEXED DETECTION WITH ISOTOPE-CODED REPORTERS**

(71) Applicant: **Massachusetts Institute of Technology**, Cambridge, MA (US)

(72) Inventors: **Gabriel A. Kwong**, Atlanta, GA (US);
Sangeeta N. Bhatia, Lexington, MA (US)

(73) Assignee: **Massachusetts Institute of Technology**, Cambridge, MA (US)

(21) Appl. No.: **18/063,162**

(22) Filed: **Dec. 8, 2022**

Related U.S. Application Data

- (63) Continuation of application No. 15/842,162, filed on Dec. 14, 2017, now Pat. No. 11,549,947, which is a continuation of application No. 14/005,416, filed on Jun. 12, 2014, now Pat. No. 10,006,916, filed as application No. PCT/US2012/029200 on Mar. 15, 2012.
- (60) Provisional application No. 61/452,908, filed on Mar. 15, 2011.

Publication Classification

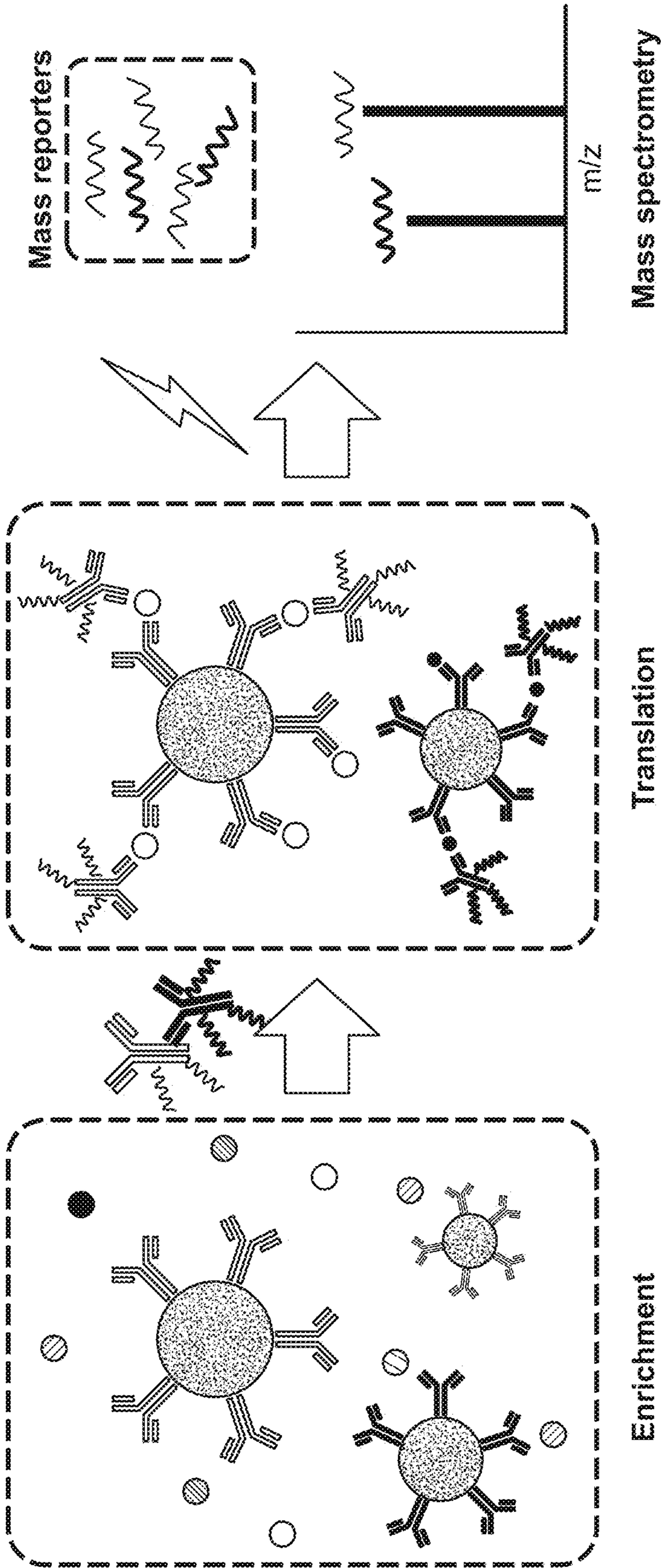
- (51) **Int. Cl.**
G01N 33/58 (2006.01)
C07K 7/08 (2006.01)
C07K 14/00 (2006.01)
G01N 33/68 (2006.01)
- (52) **U.S. Cl.**
CPC **G01N 33/58** (2013.01); **C07K 7/08** (2013.01); **C07K 14/001** (2013.01); **G01N 33/6848** (2013.01); **G01N 2458/15** (2013.01); **A61K 38/00** (2013.01)

(57) **ABSTRACT**

Some aspects of this invention provide reagents and methods for the sensitive, quantitative and simultaneous detection of target analytes in complex biological samples by liquid chromatography tandem mass spectrometry (LC MS/MS). Some aspects of this invention provide affinity reagents encoded with mass reporters for the sensitive and quantitative translation of an analyte of interest into a mass tag. The reagents and methods provided herein have general utility in analyte detection and encoding, for example, in biomolecular profiling, molecular diagnostics, and biochemical encoding.

Specification includes a Sequence Listing.

FIG. 1a



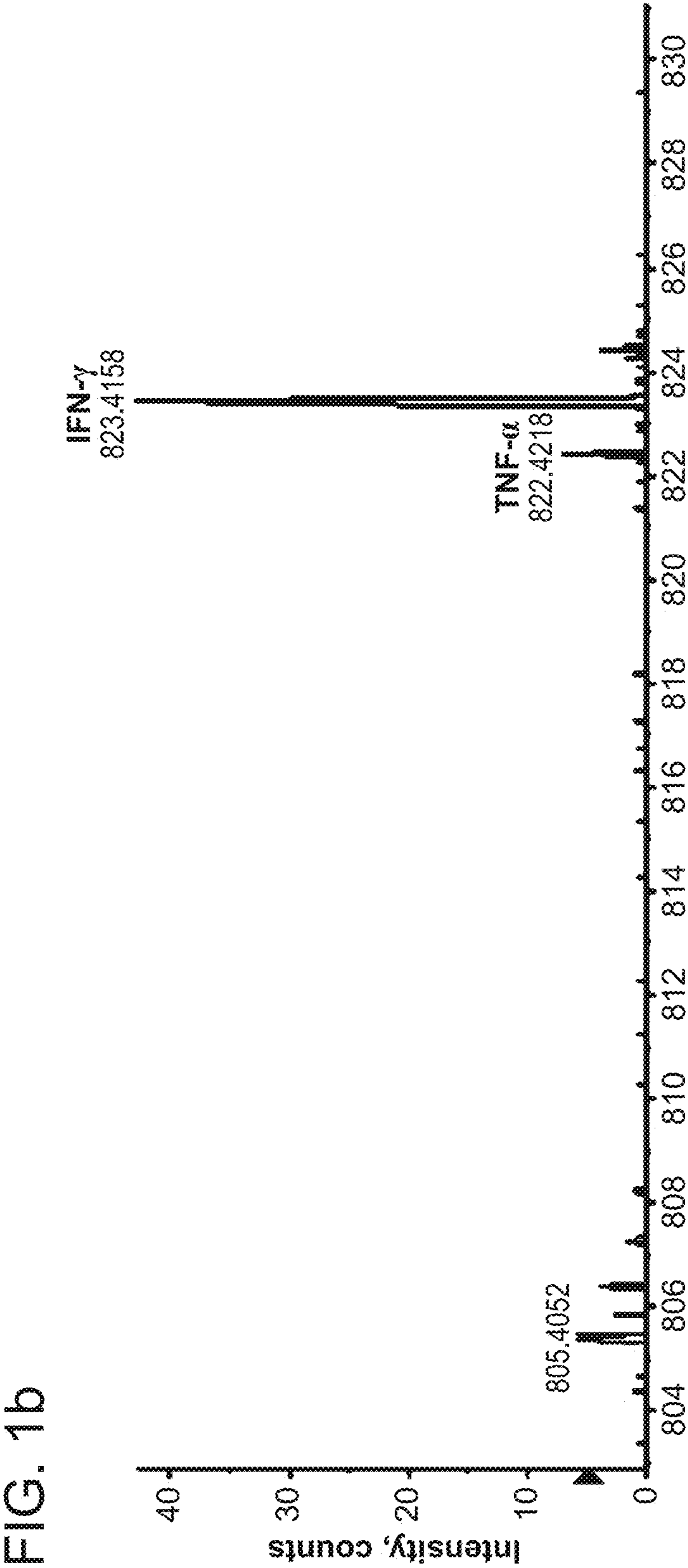
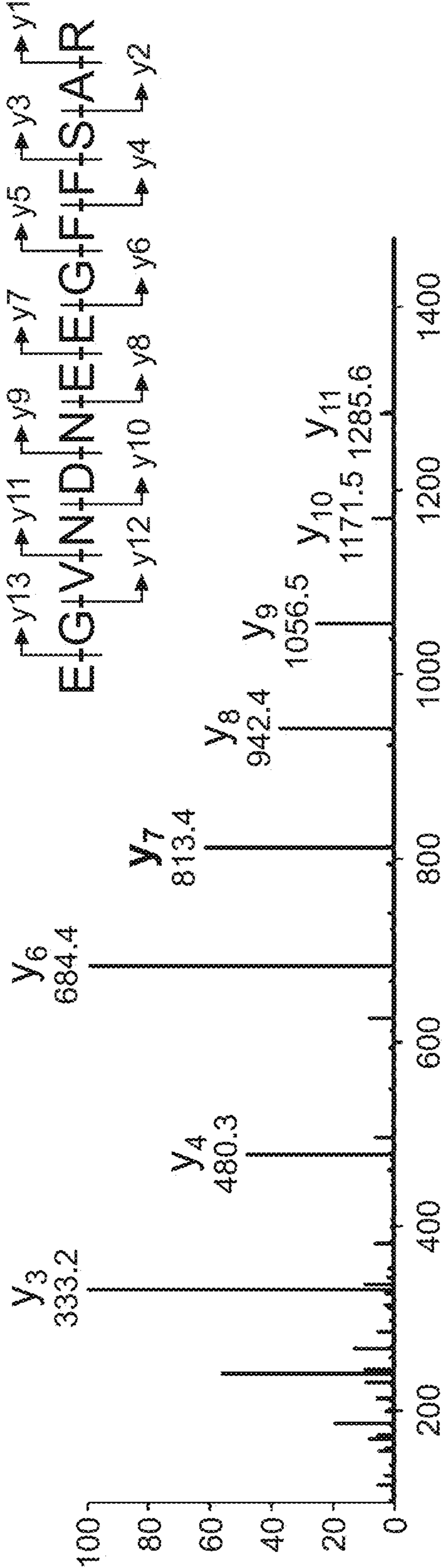


FIG. 2

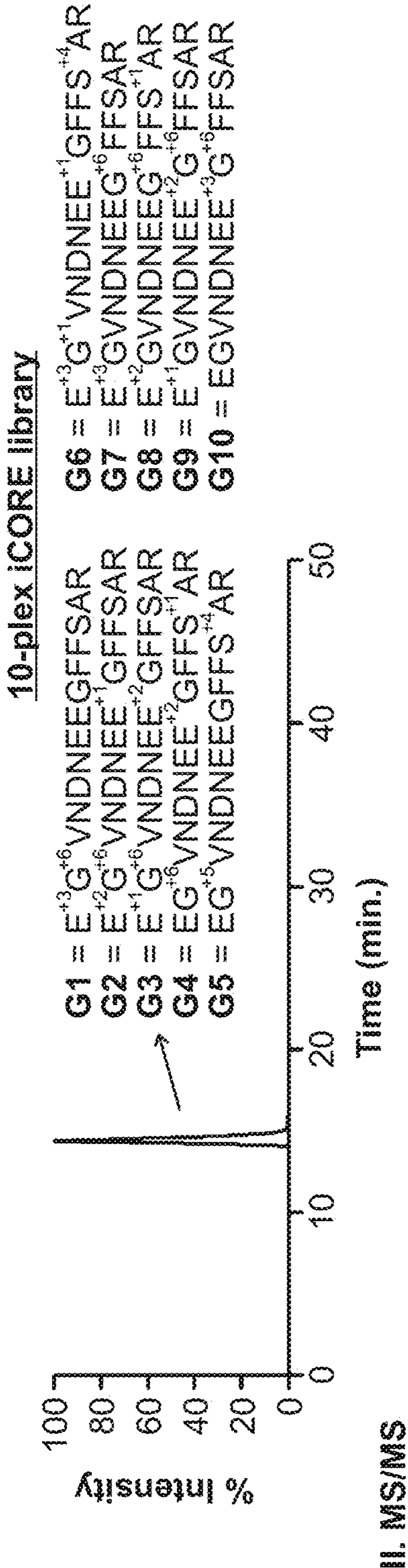
Utilization of a canonical peak from Glu-fib MS/MS spectra as reporter



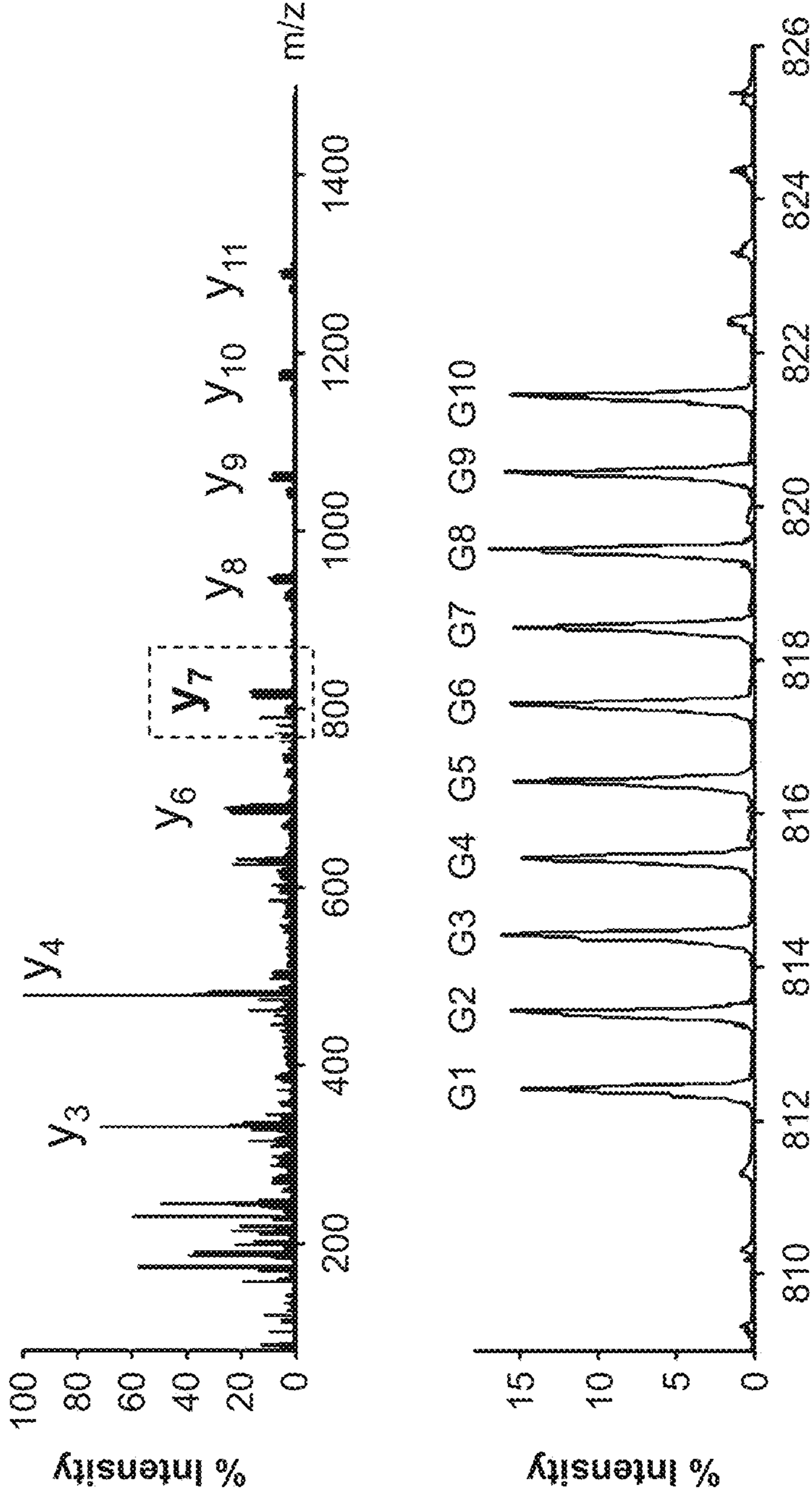
Mass reporter (y ₇ ion)	
Balance	Reporter
E- ⁺³ G- ⁺⁶ V-N-D-N-E-E-G-F-F-S-A-R	812.38
E- ⁺² G- ⁺⁶ V-N-D-N-E-E- ⁺¹ G-F-F-S-A-R	813.38
E- ⁺¹ G- ⁺⁶ V-N-D-N-E-E- ⁺² G-F-F-S-A-R	814.38
⋮	⋮
⋮	⋮
⋮	⋮
E-G-V-N-D-N-E-E- ⁺³ G- ⁺⁶ F-F-S-A-R	821.38
758.3	1579.68

FIG. 3

I. Extracted Ion Chromatograph (MS)



II. MS/MS



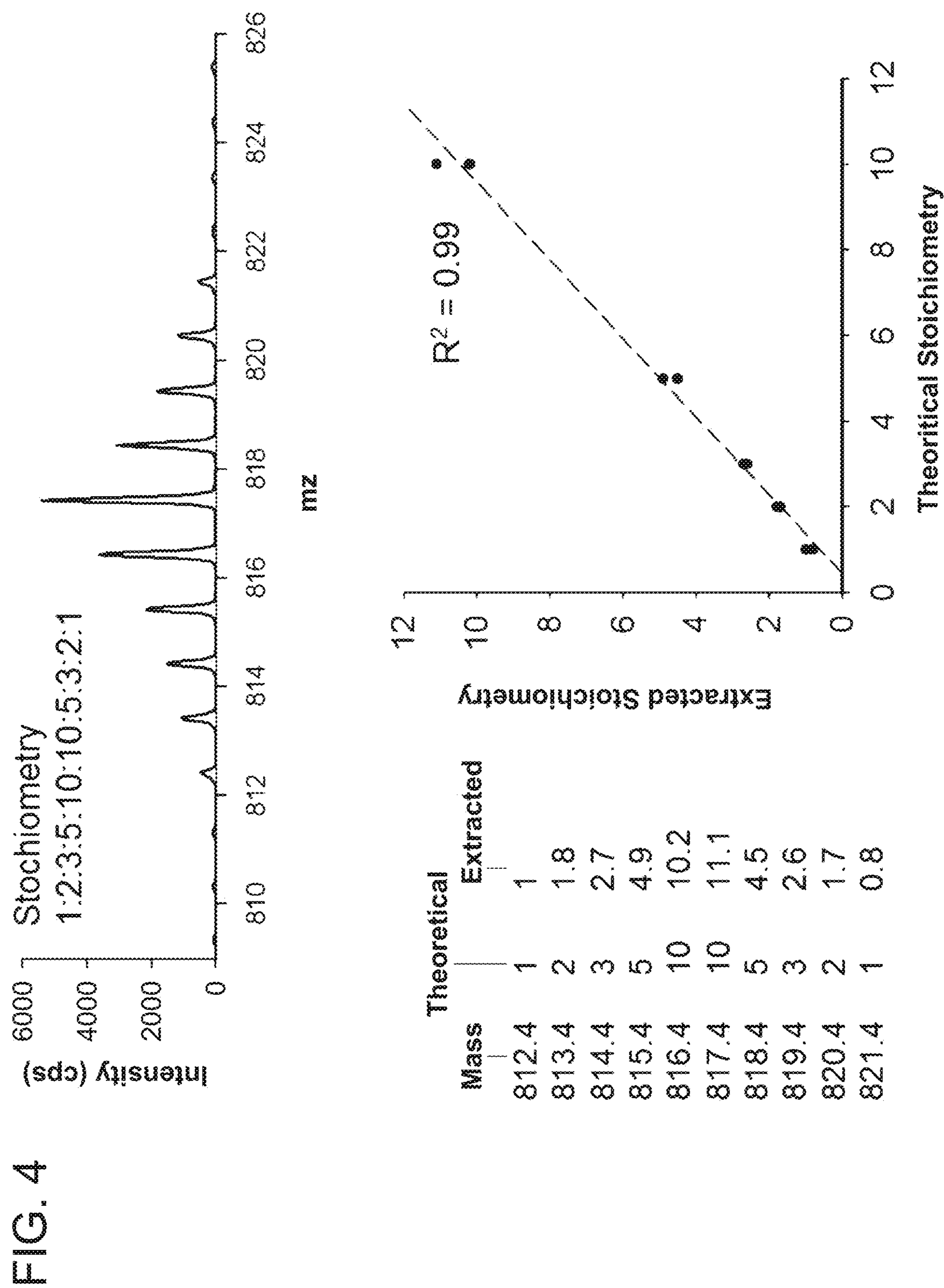
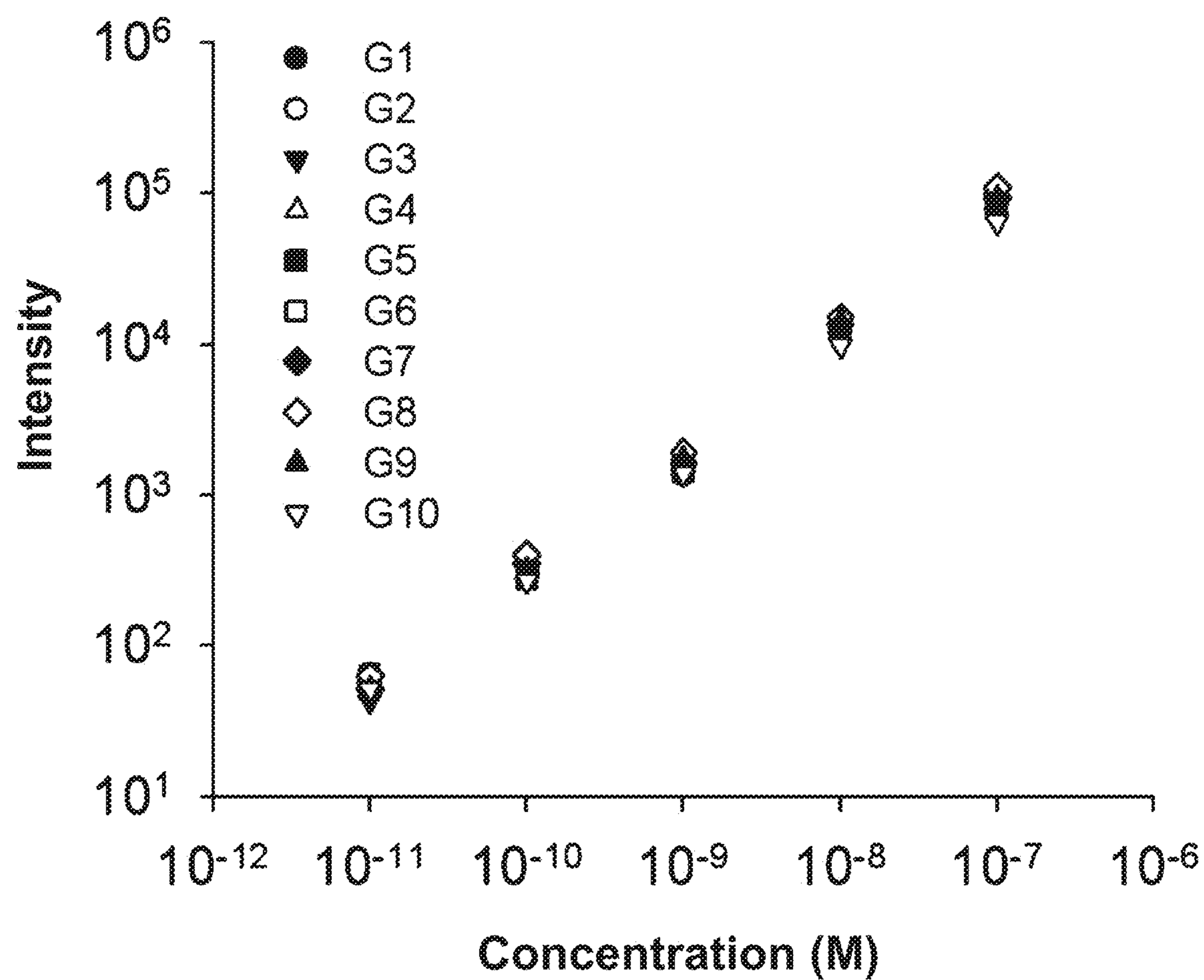
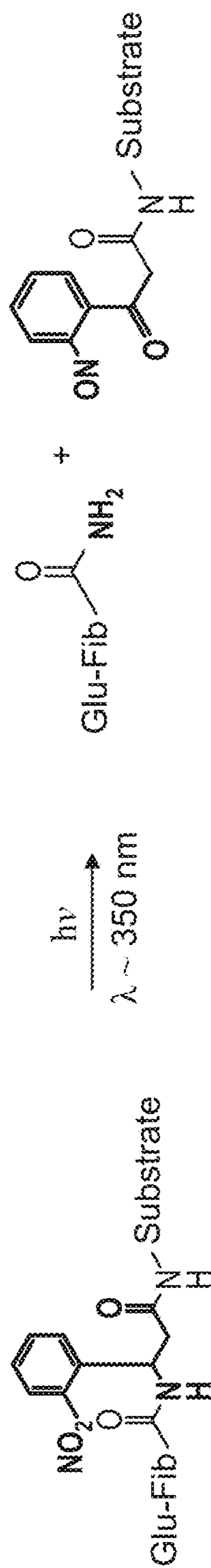


FIG. 5



1. Photo-caged reporter-substrate tandem peptides



iii. Photochemical cleavage monitored by LC MS

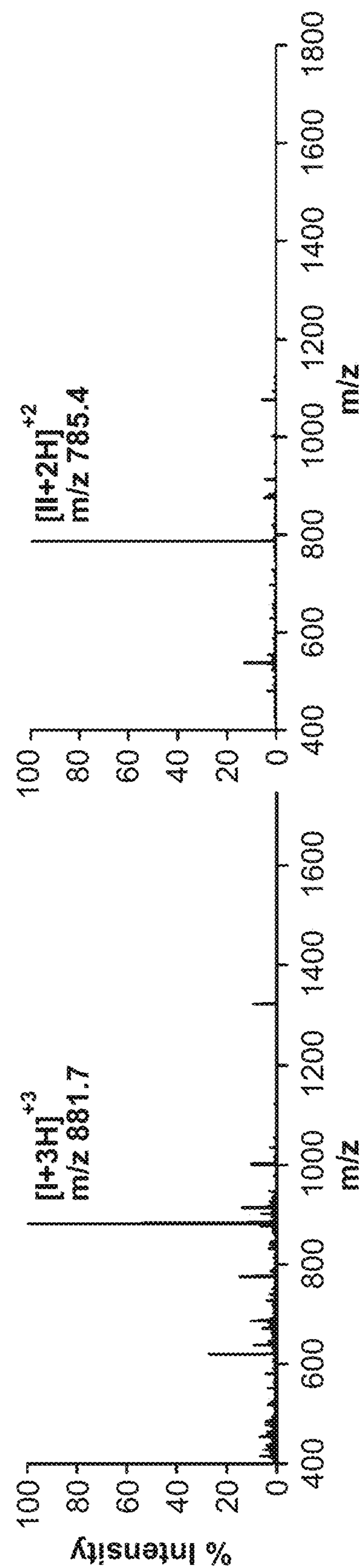
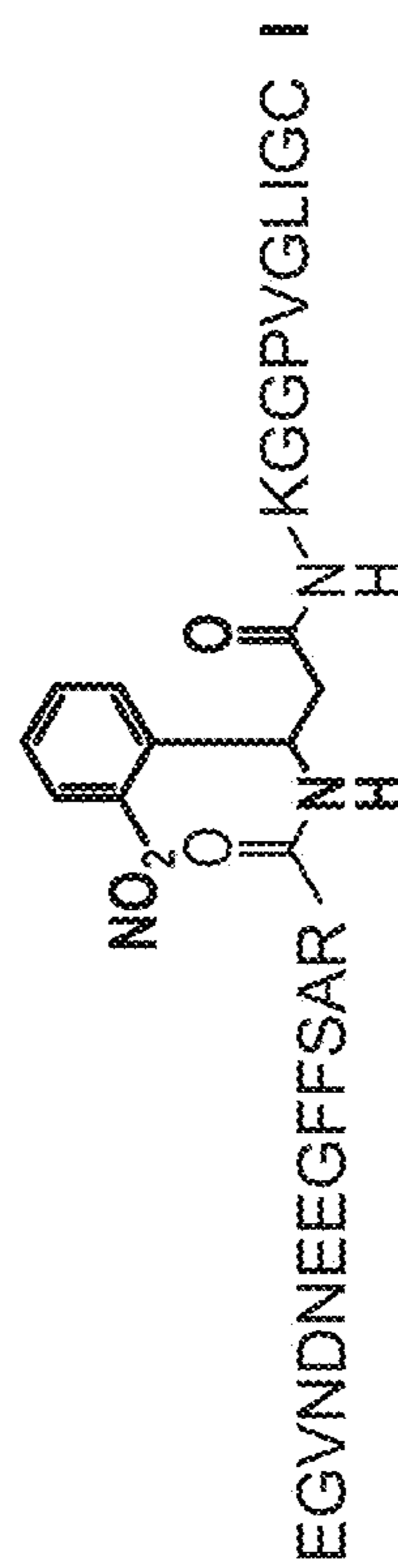
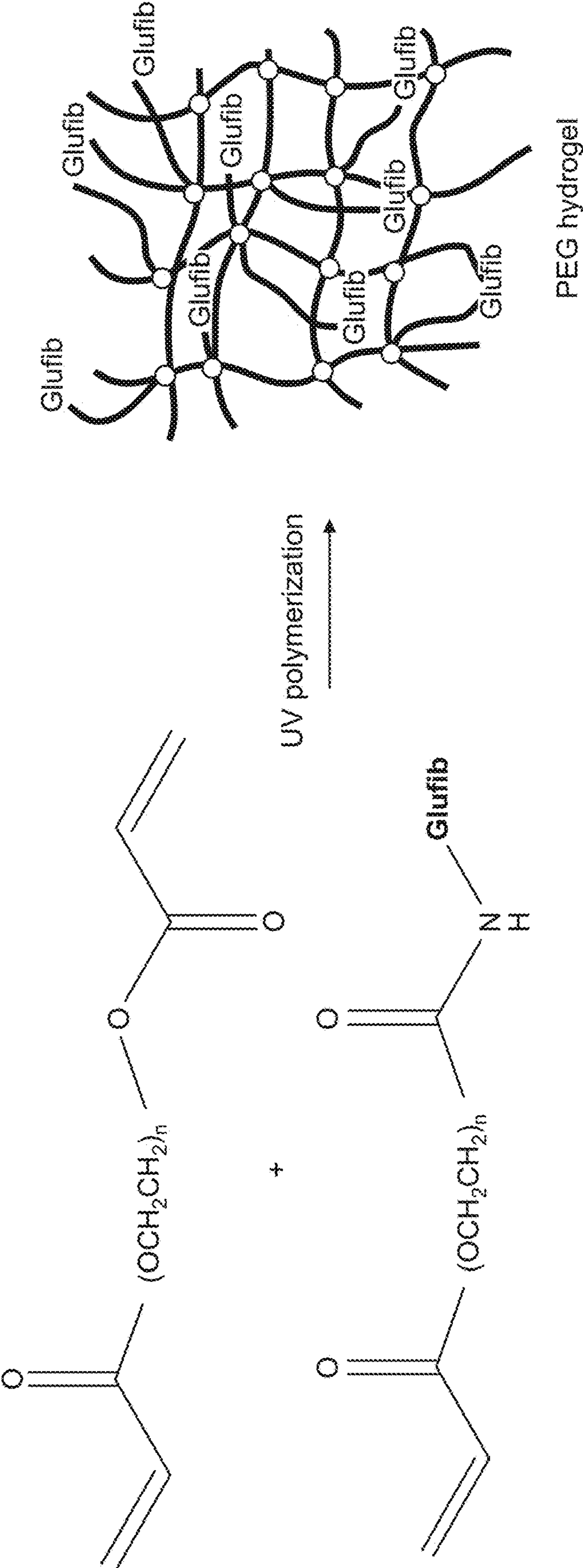
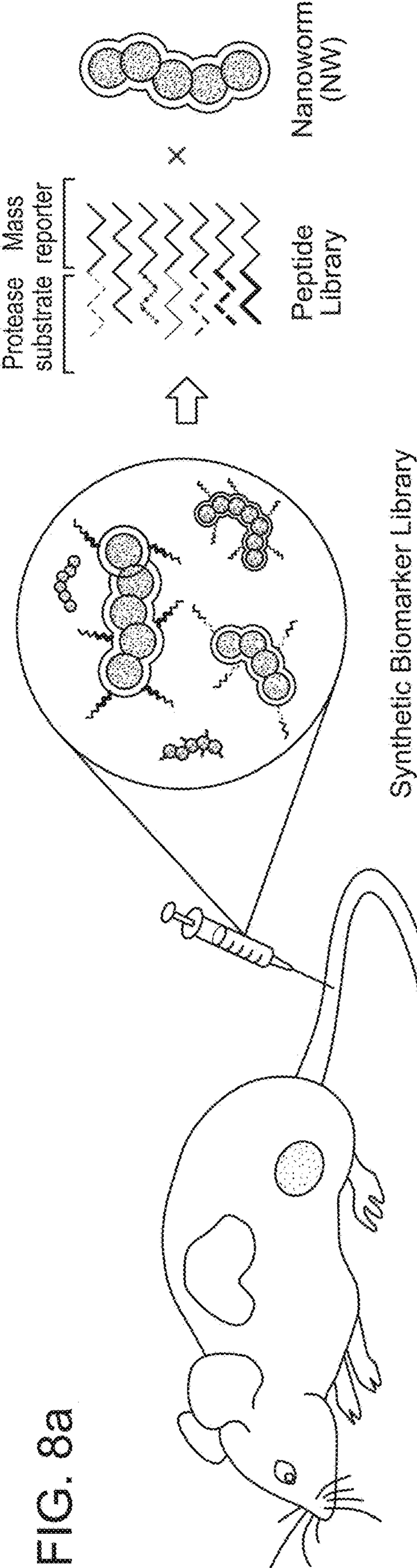


FIG. 7





Delivery to
disease sites

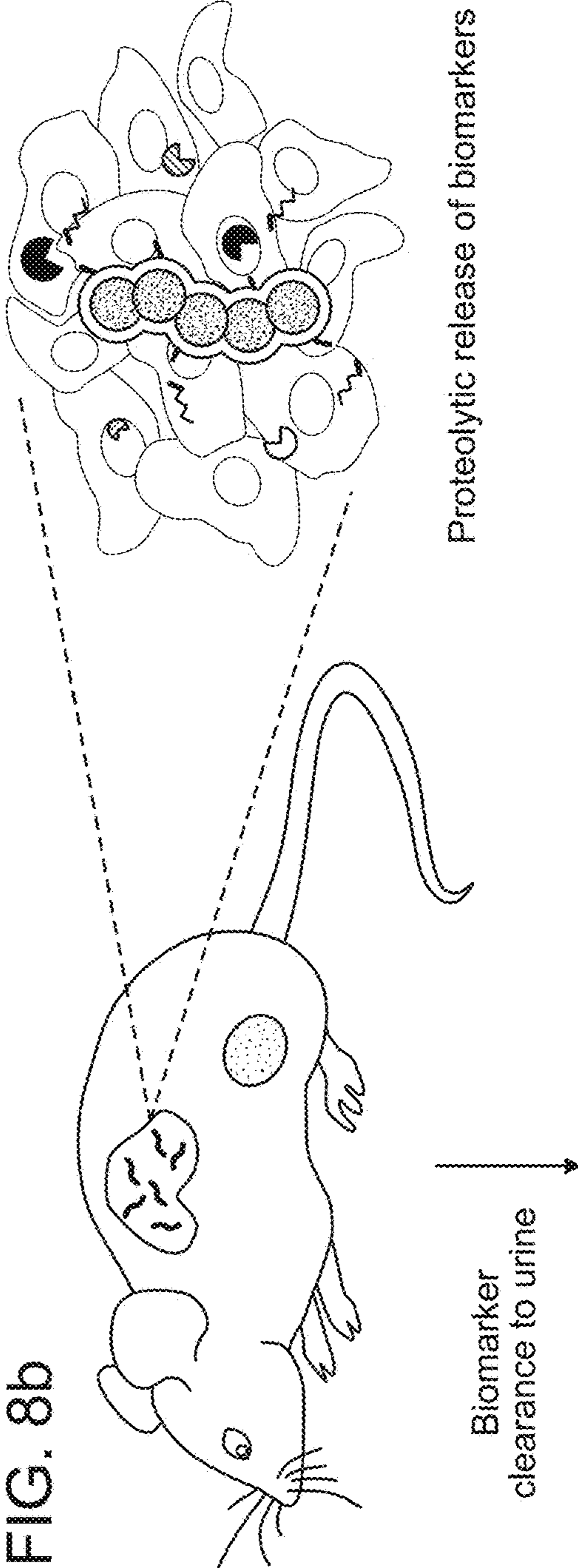


FIG. 8c

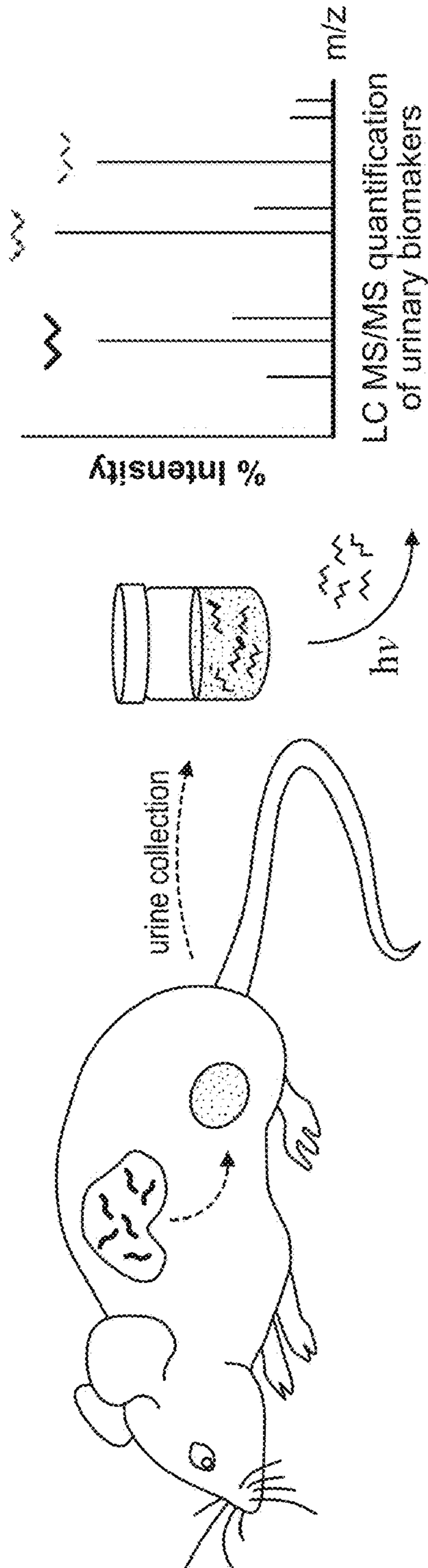


FIG. 9a

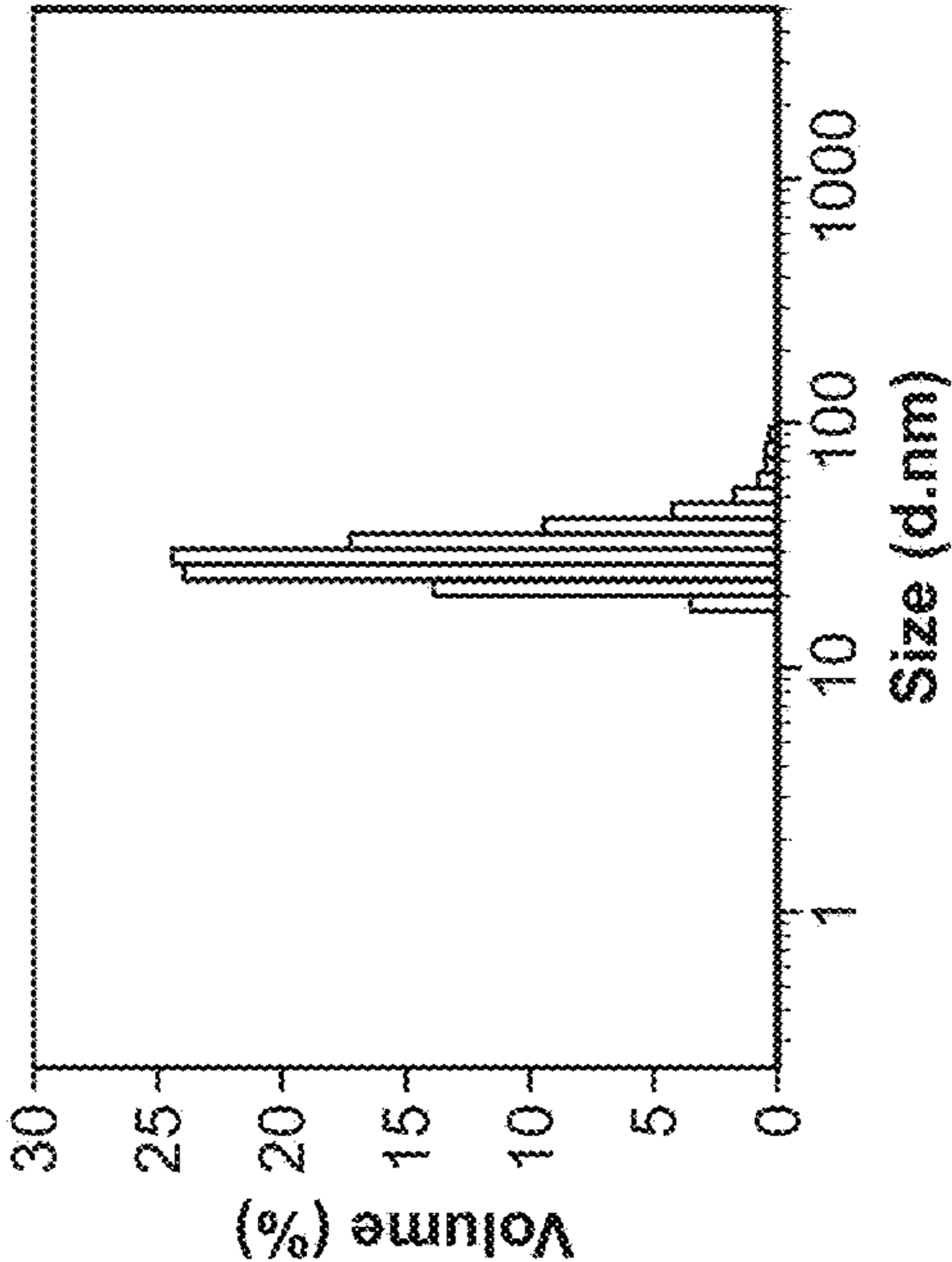


FIG. 9b

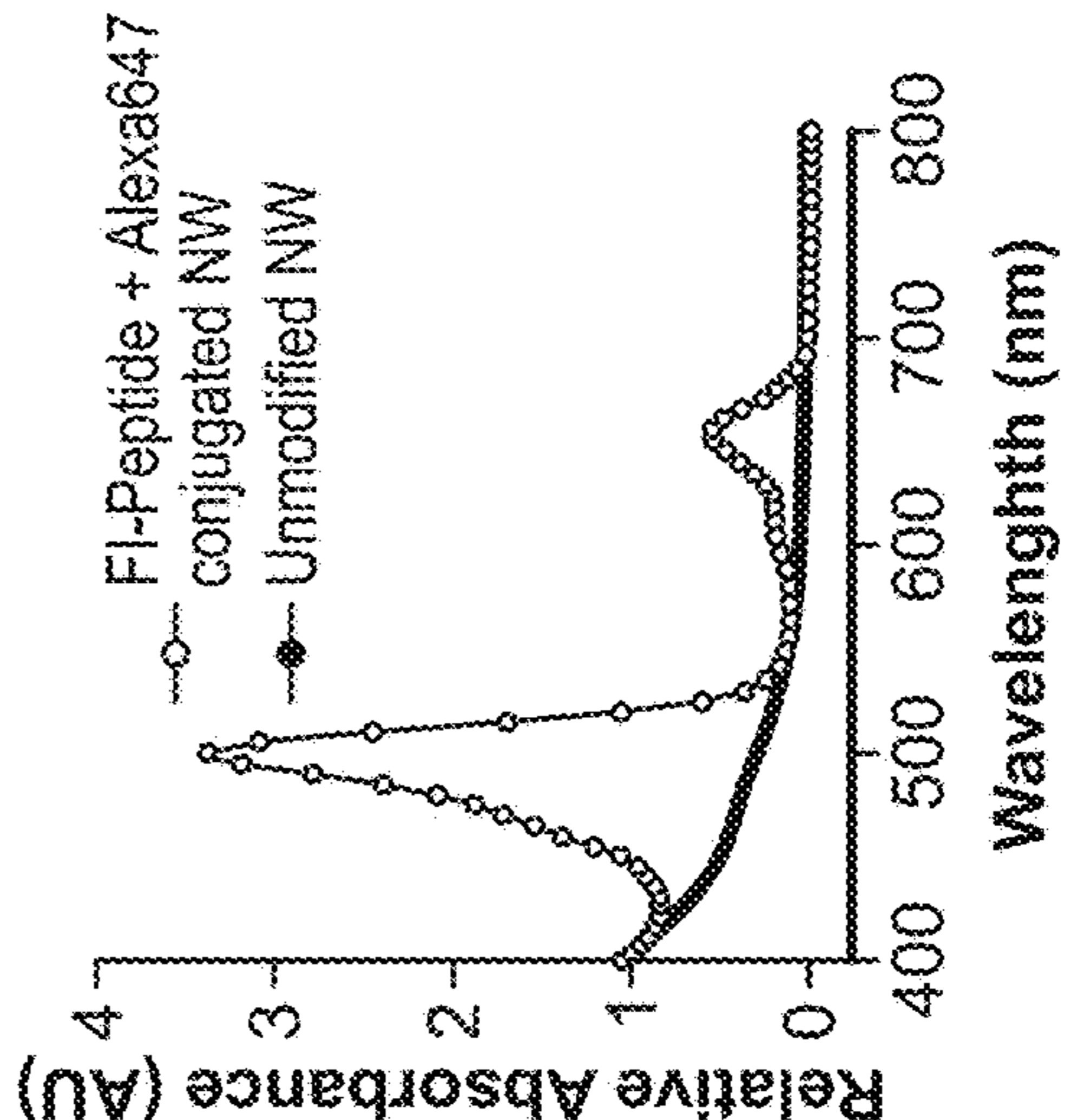
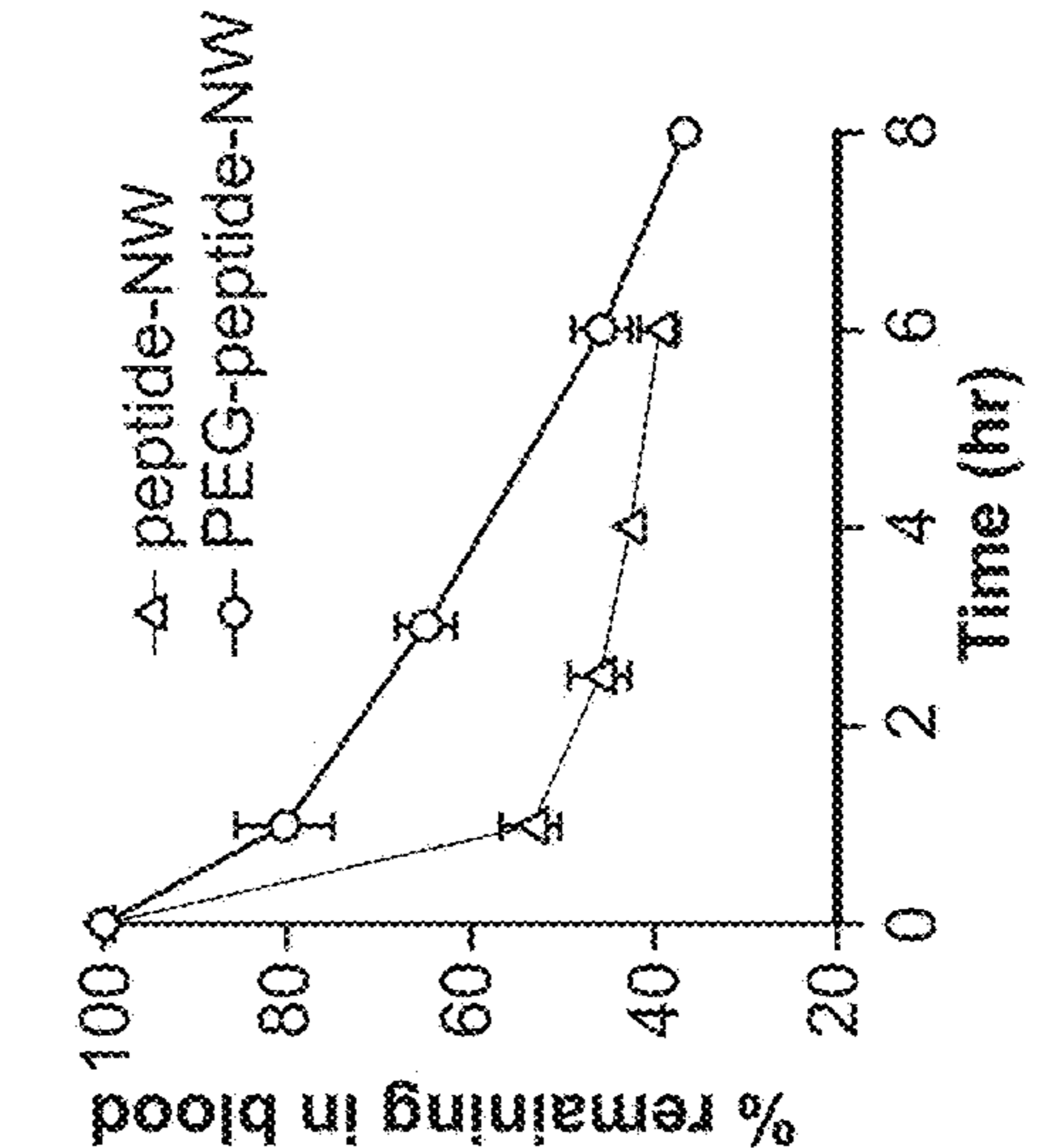
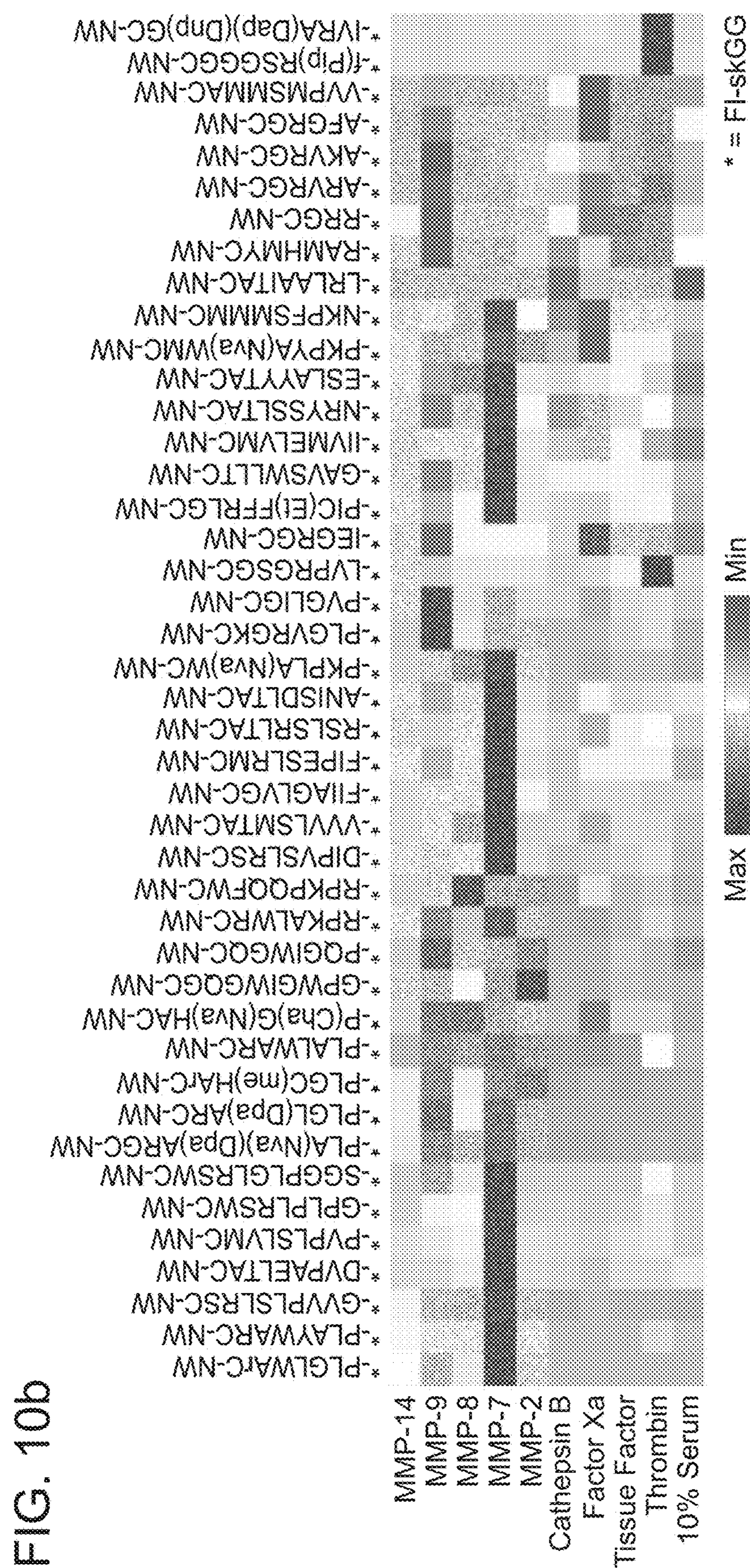


FIG. 9c





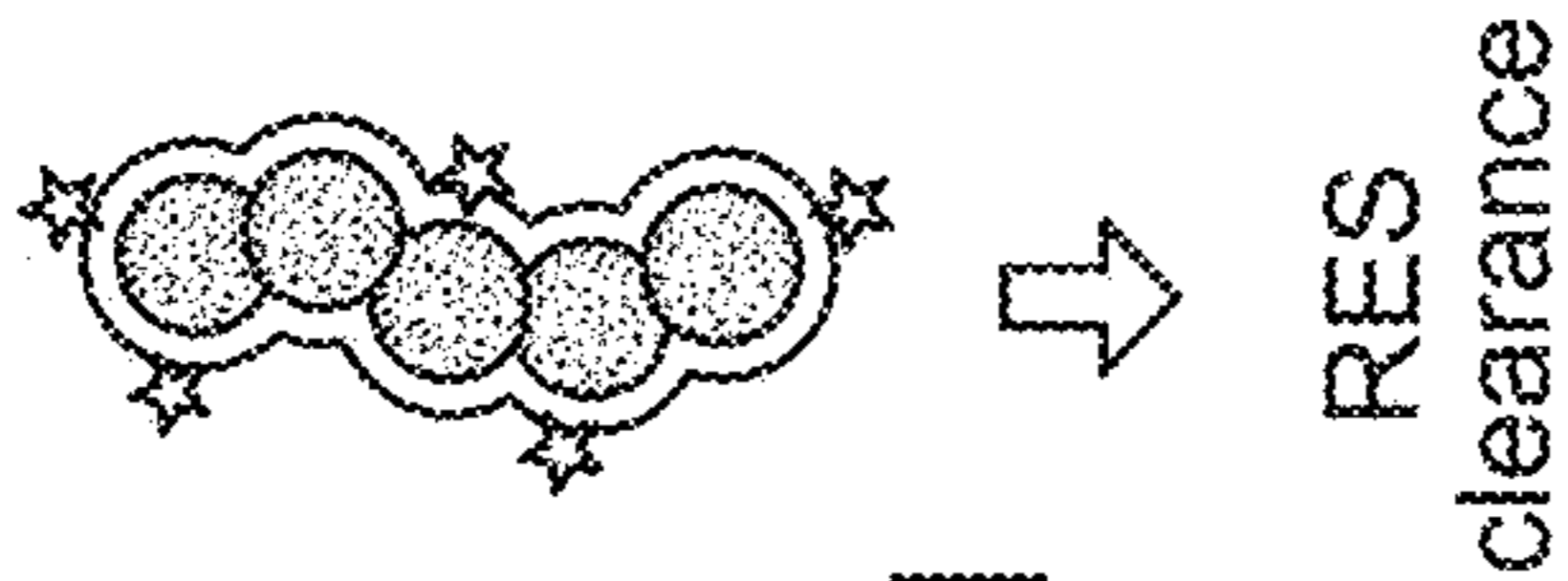
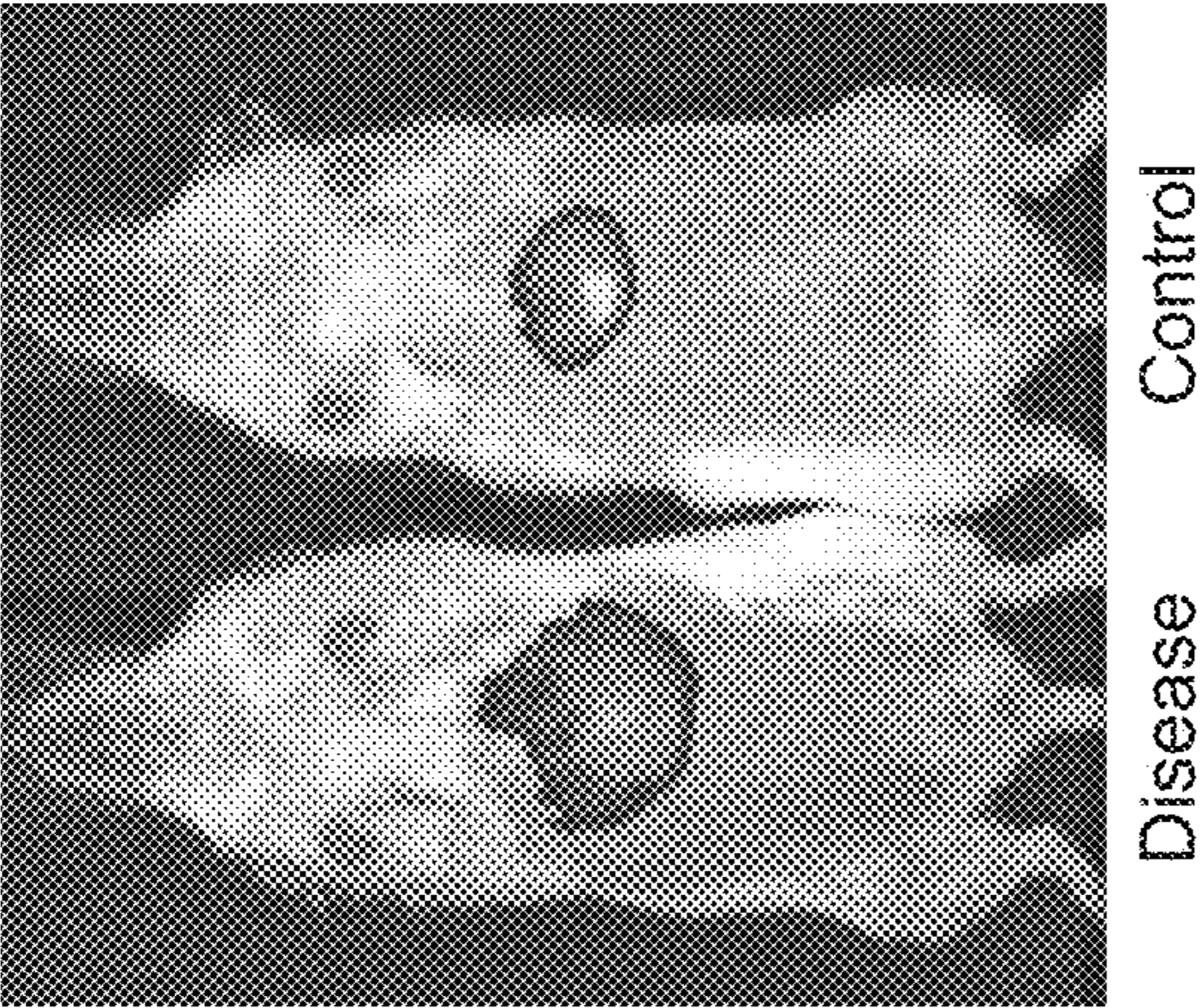


FIG. 10d

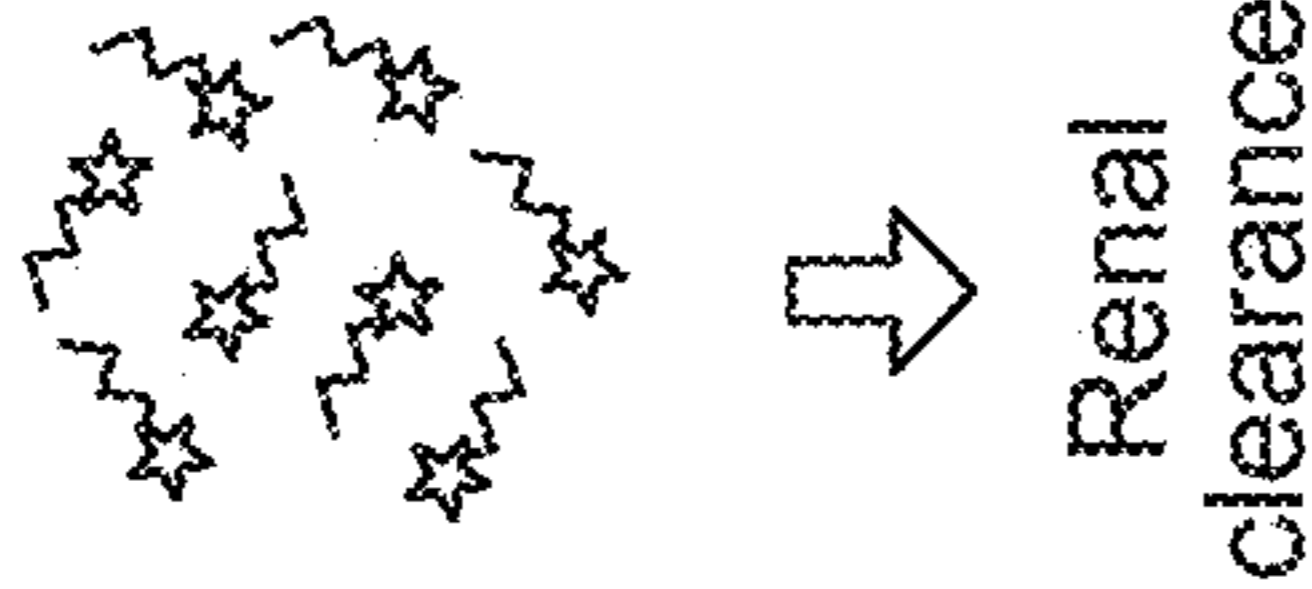
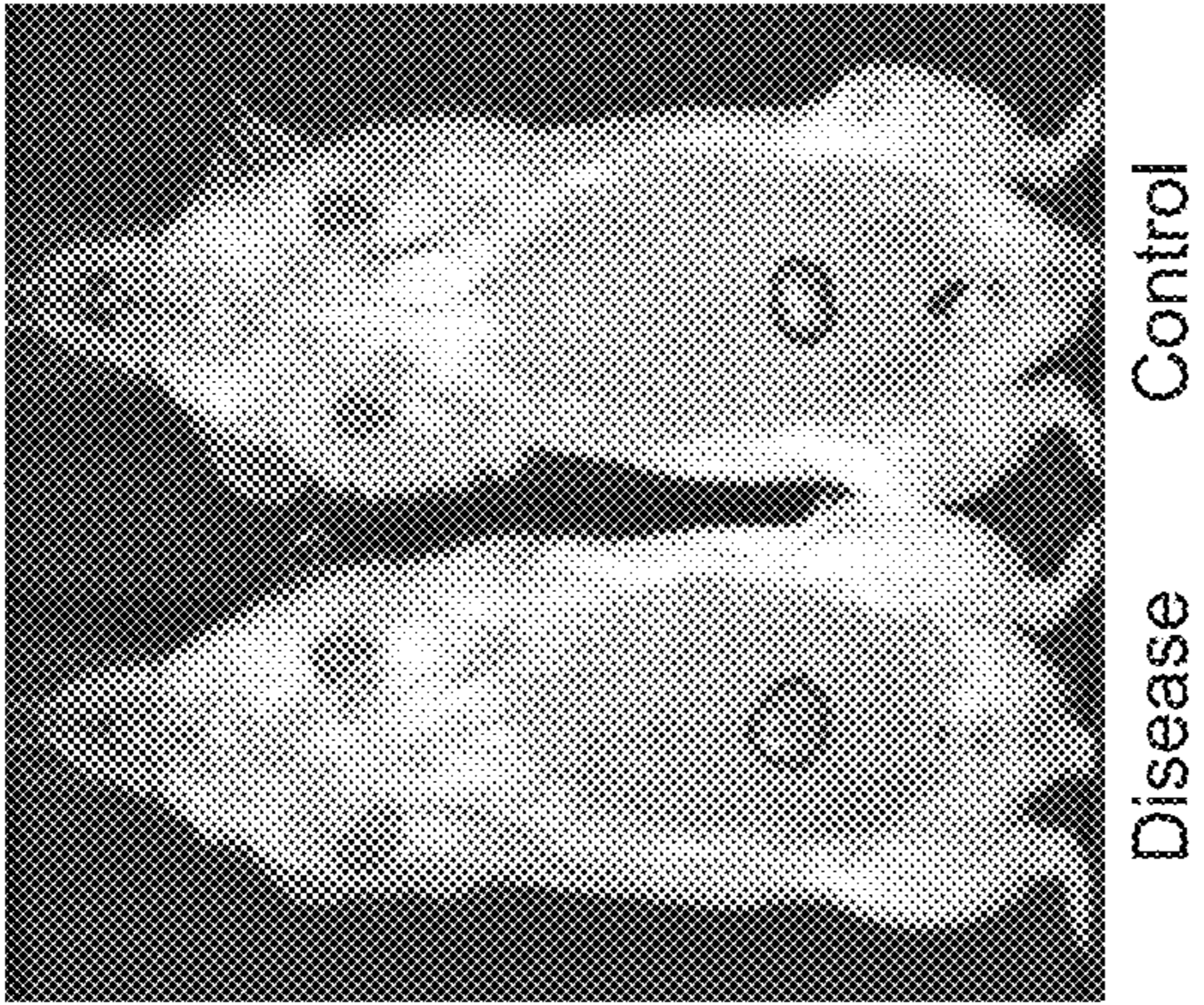


FIG. 10c

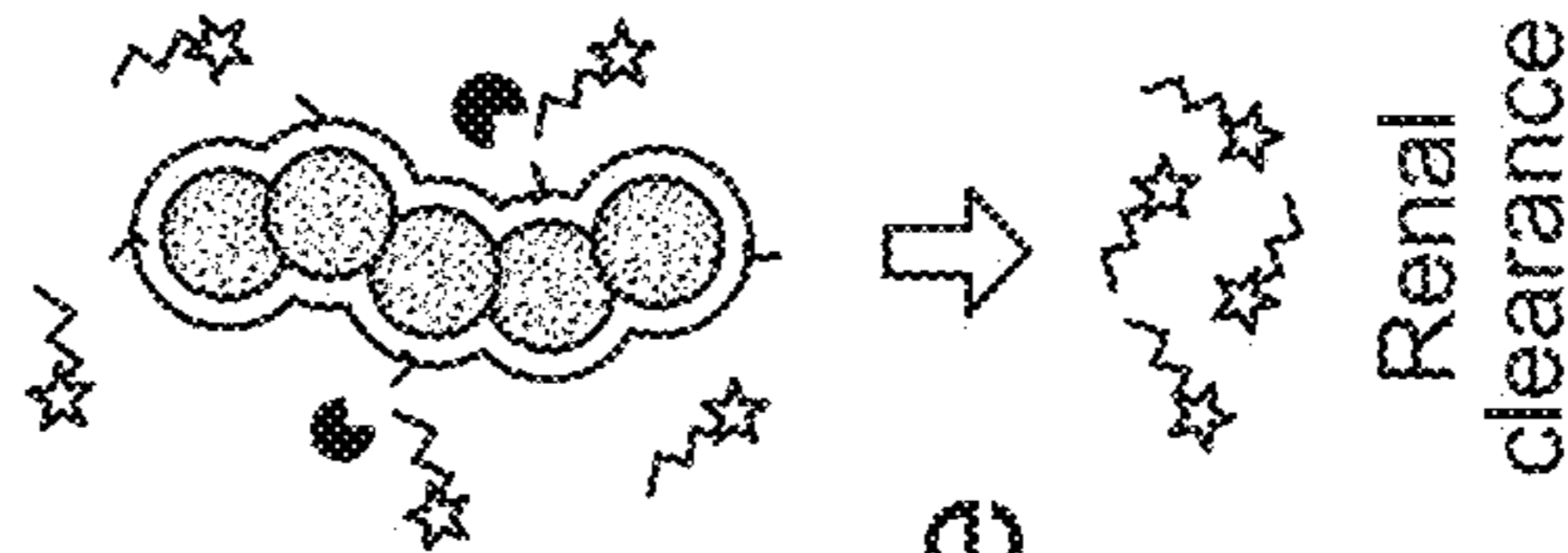
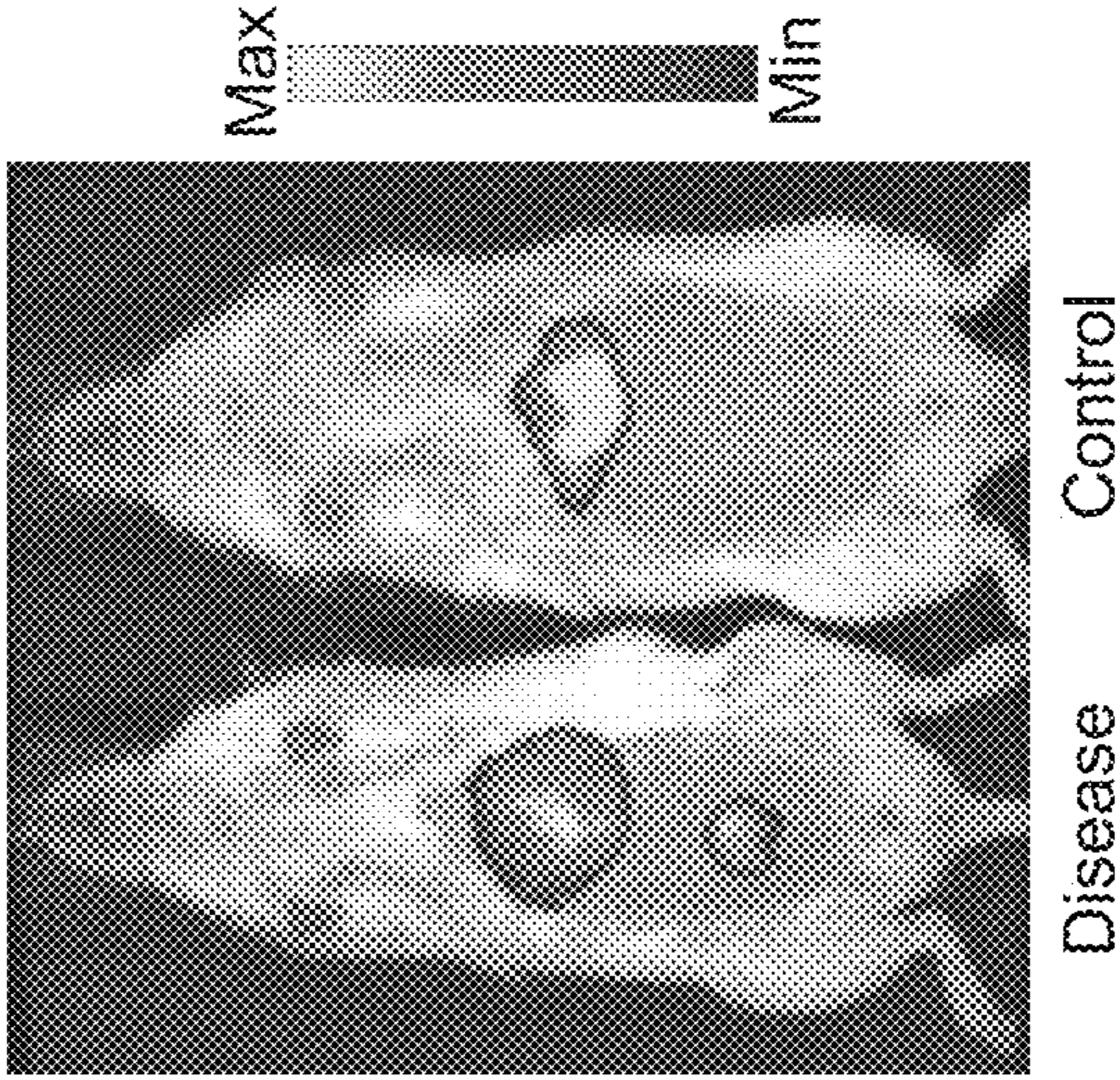


FIG. 10e

FIG. 11a

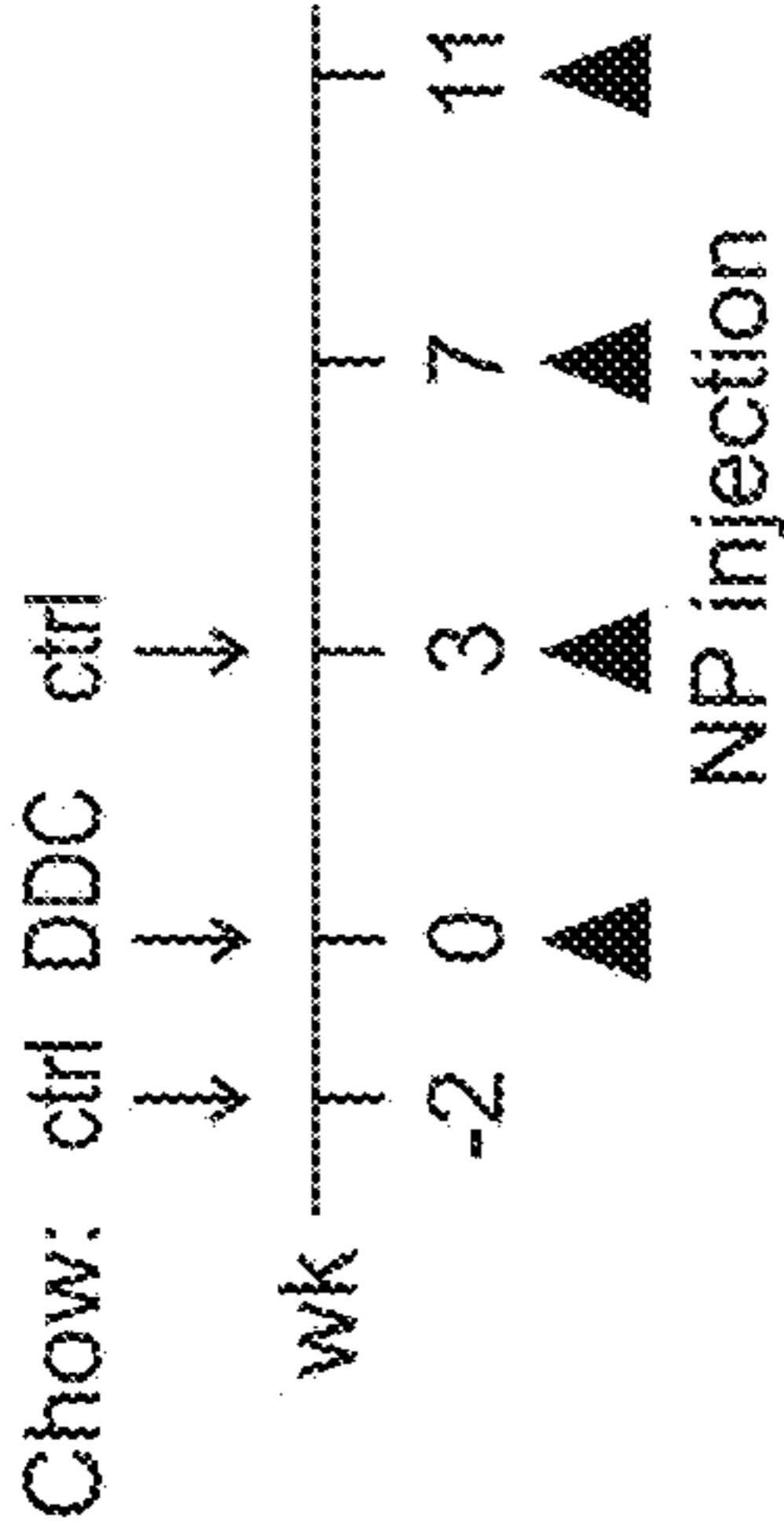


FIG. 11b

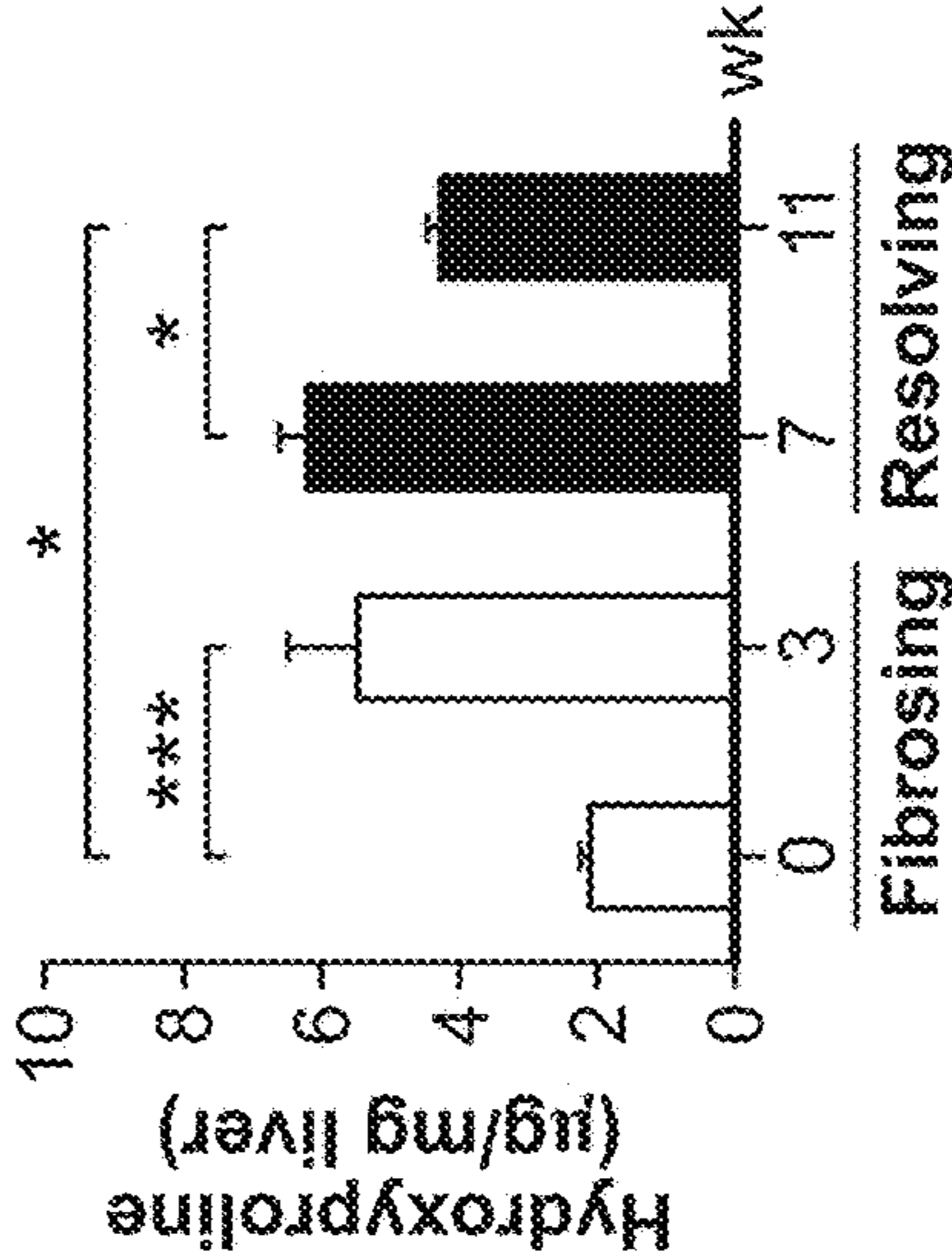


FIG. 11c

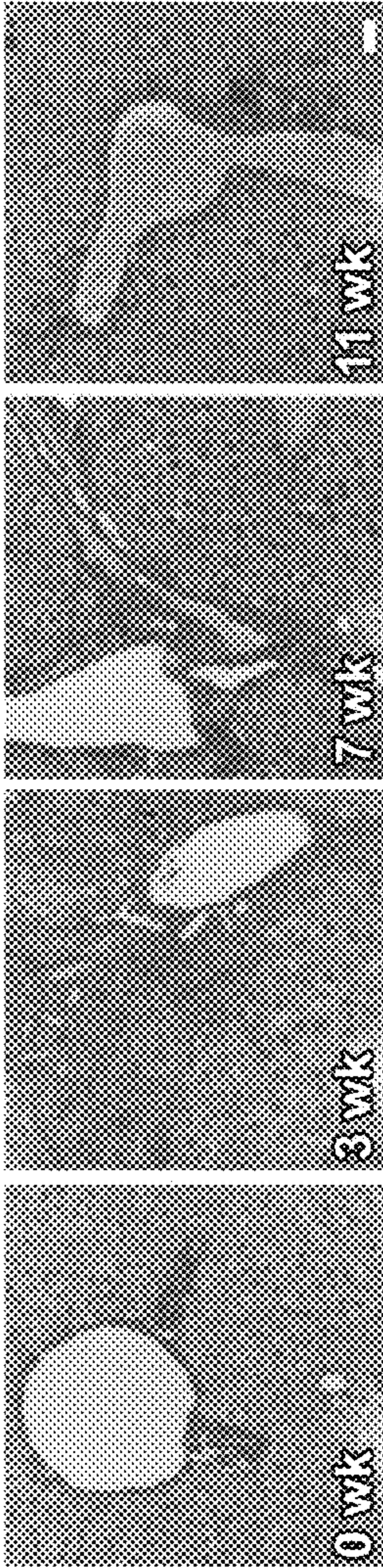


FIG. 11d

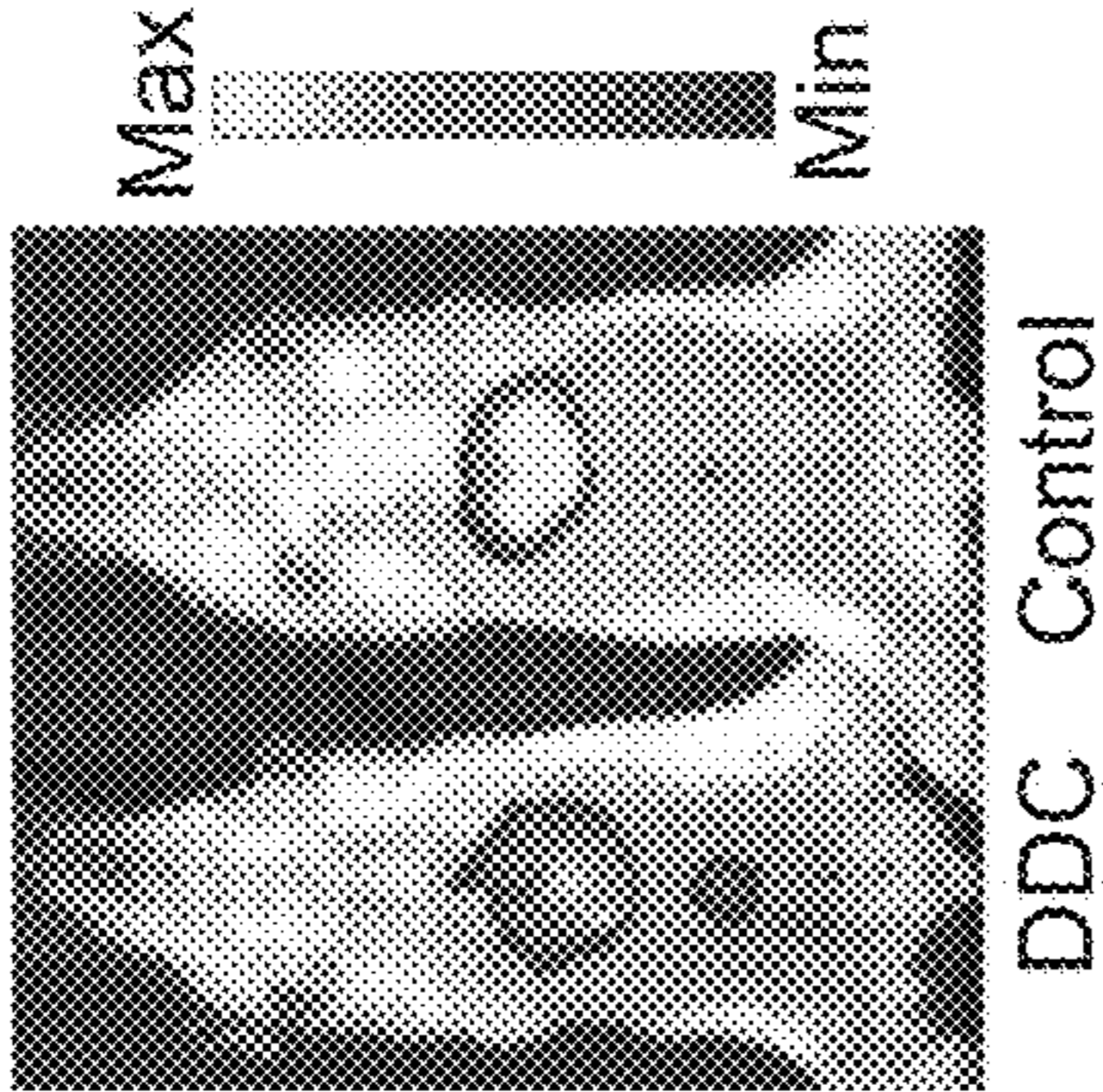


FIG. 11e

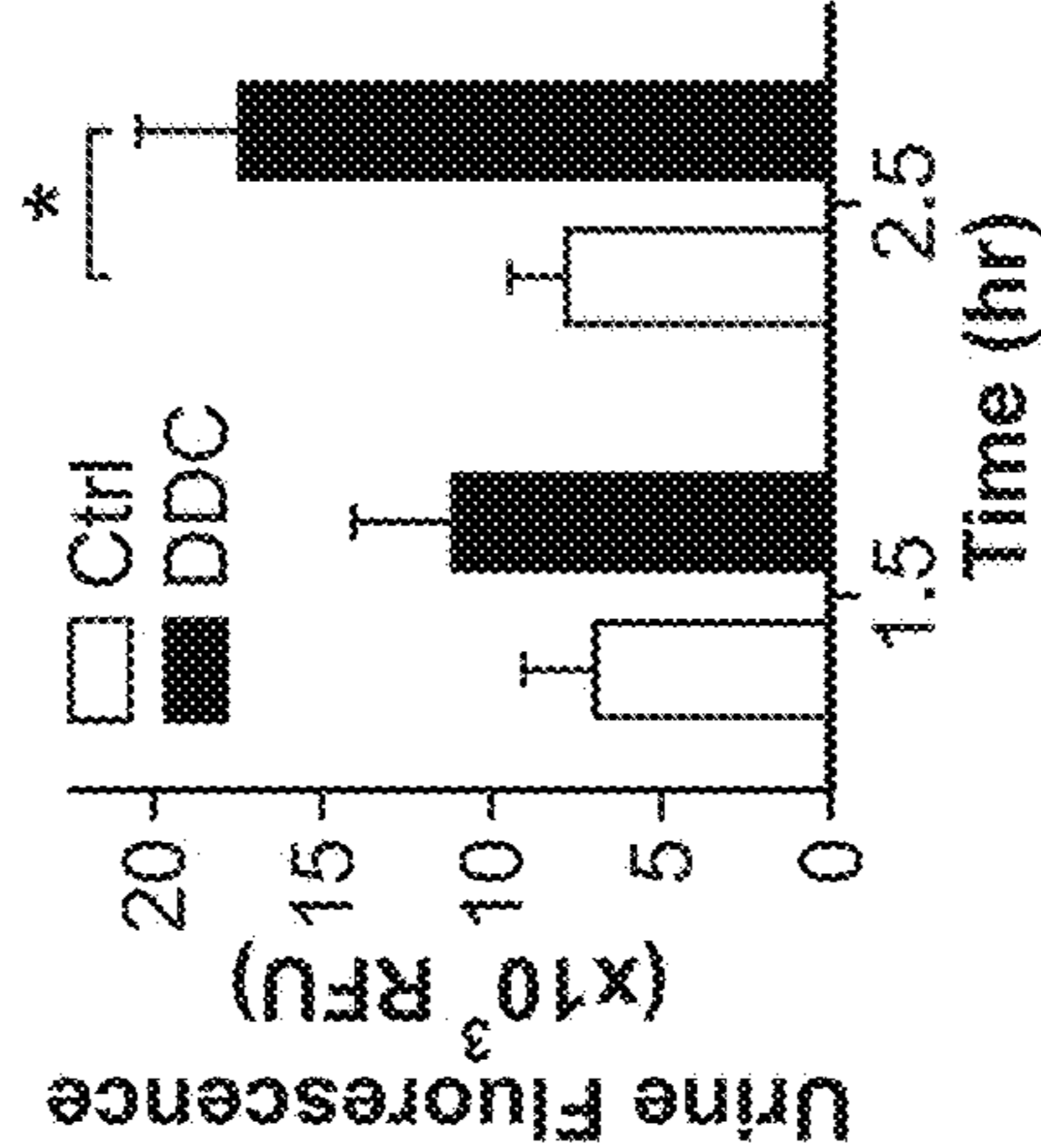


FIG. 11g

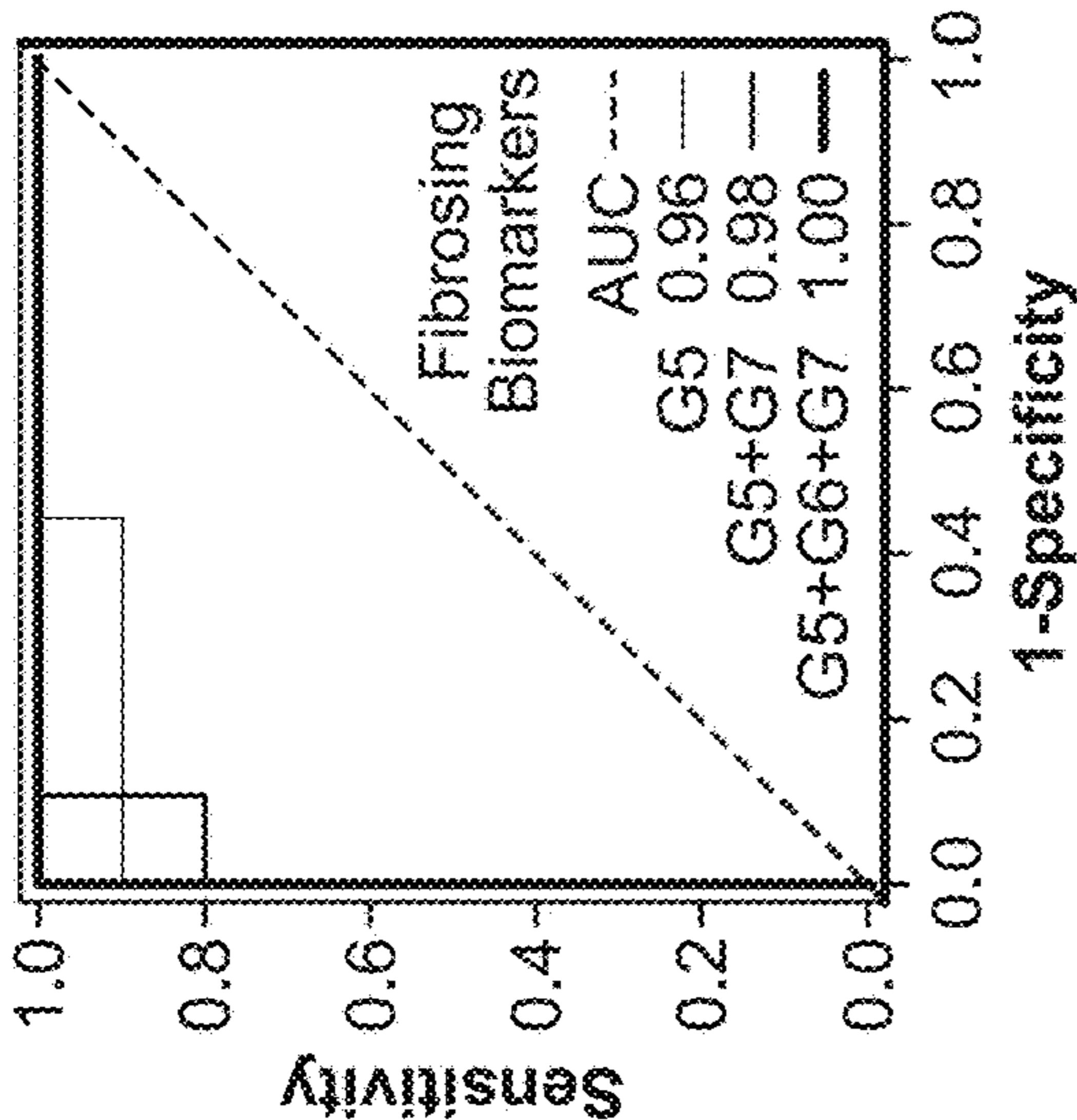


FIG. 11h

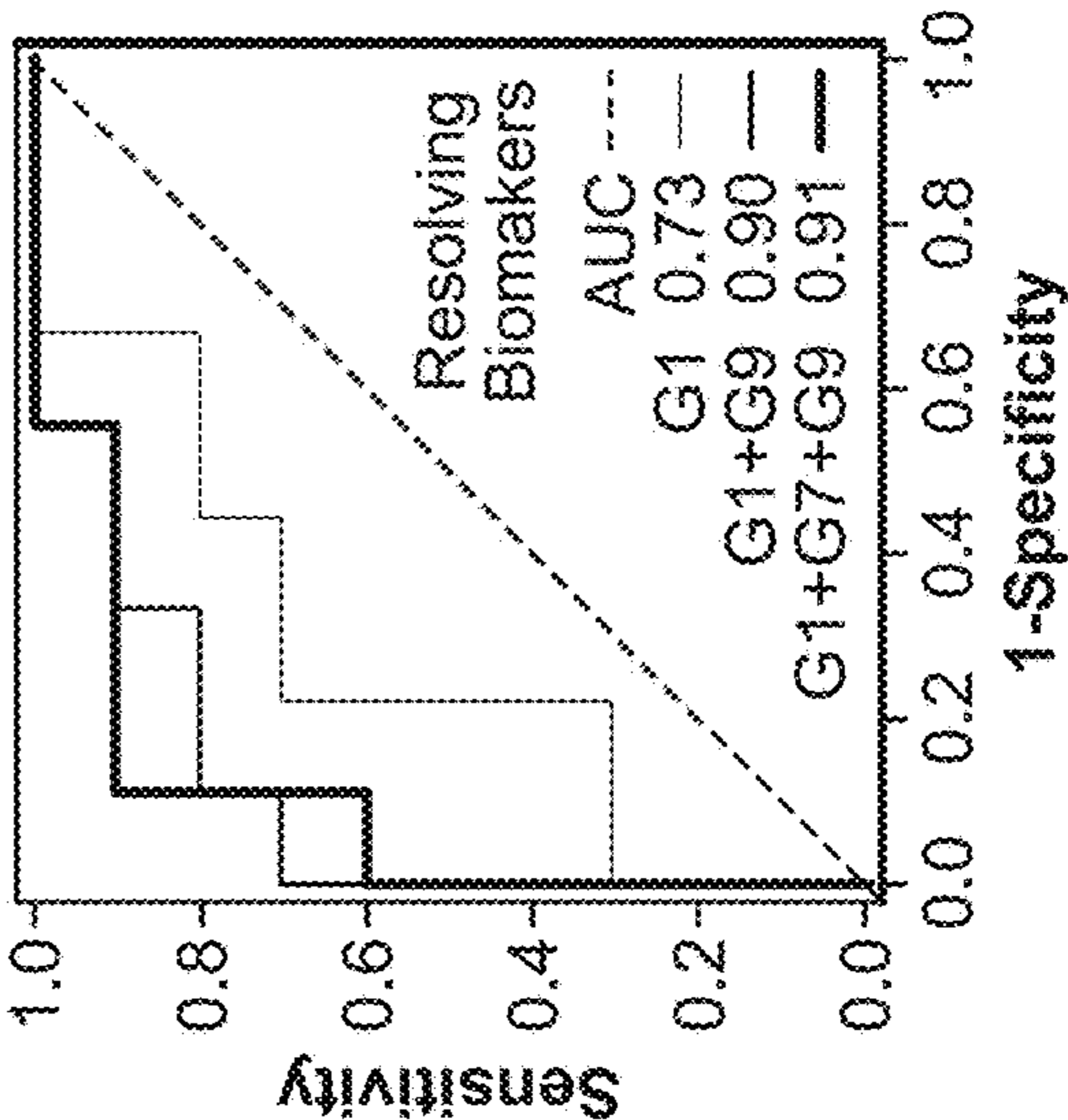


FIG. 11f

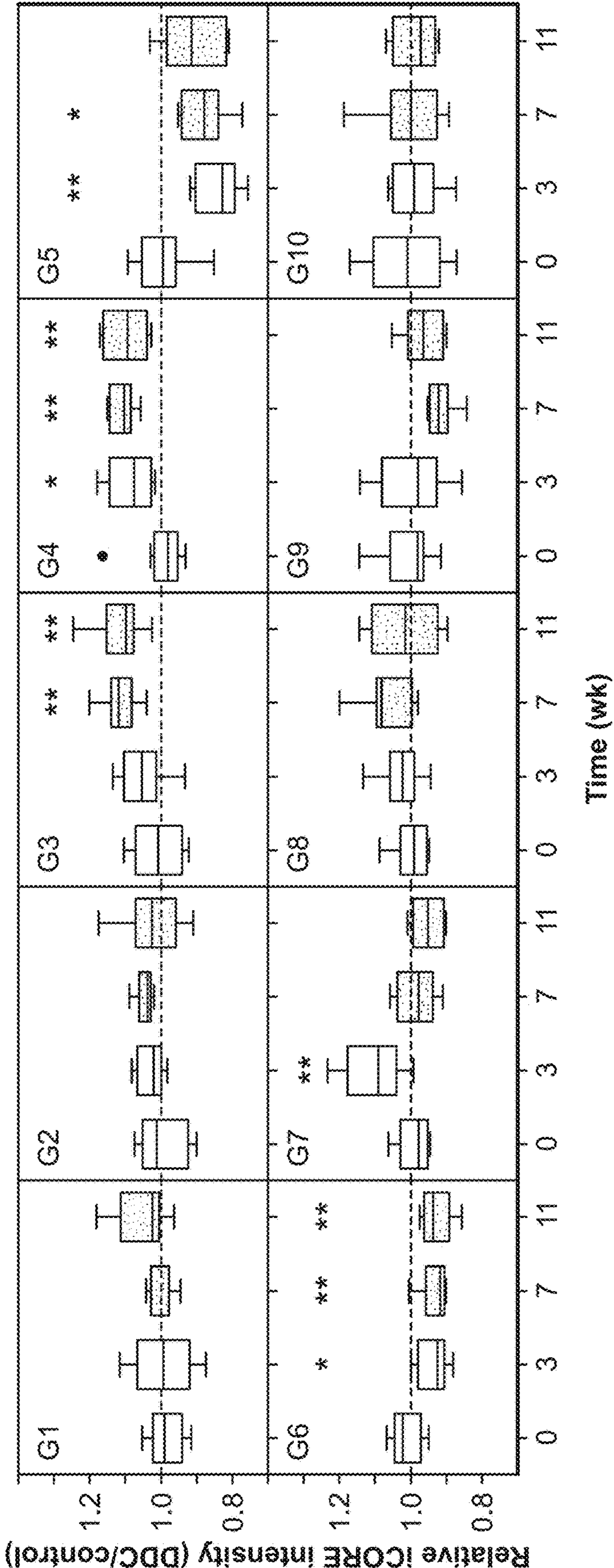


FIG. 12a

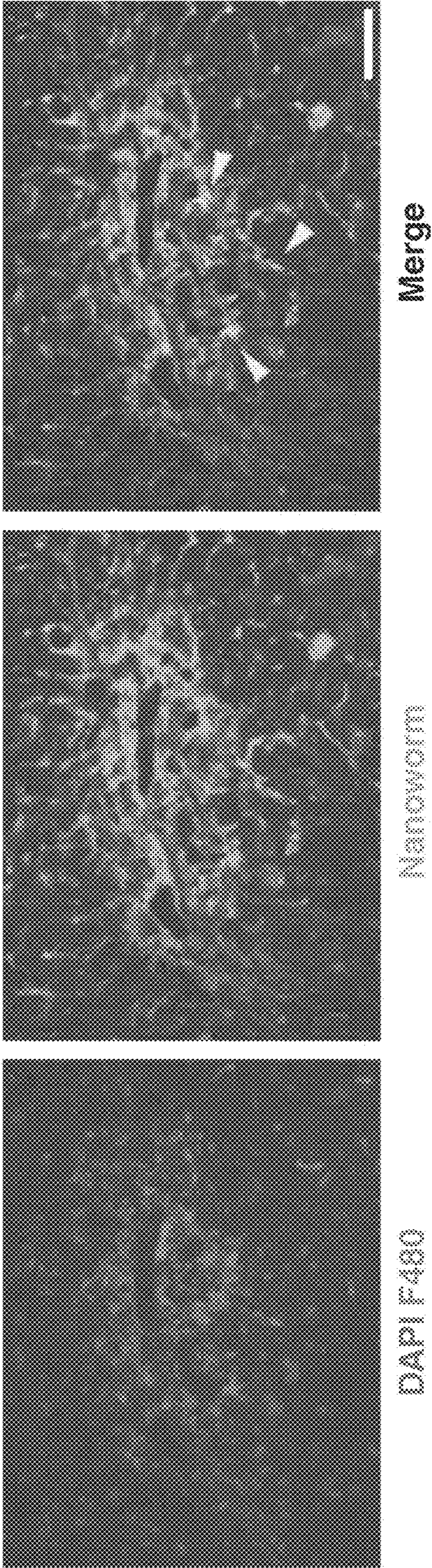


FIG. 12b

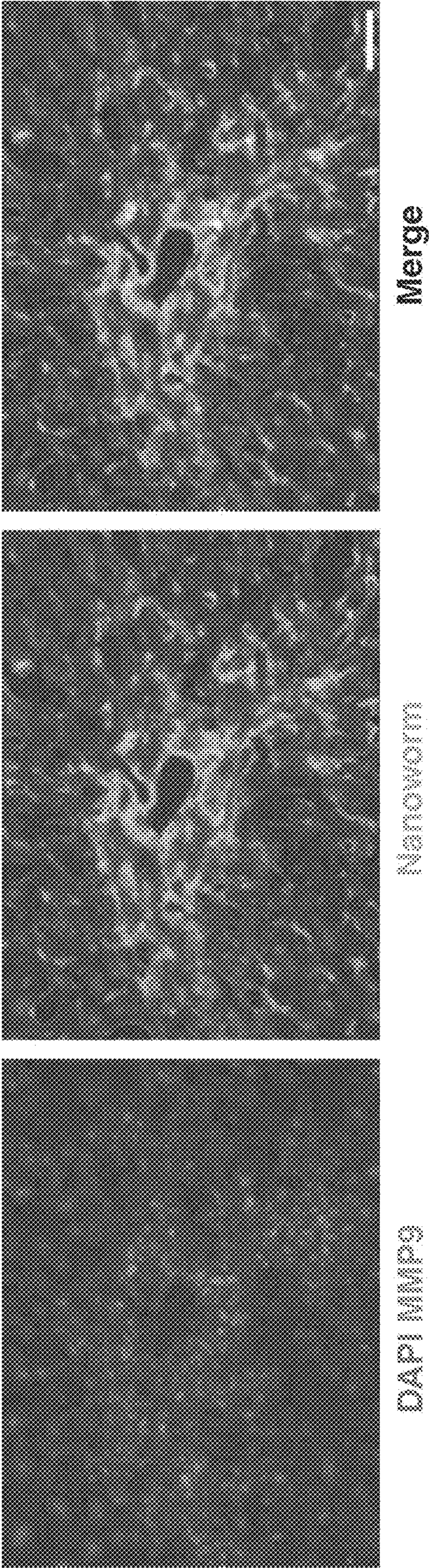
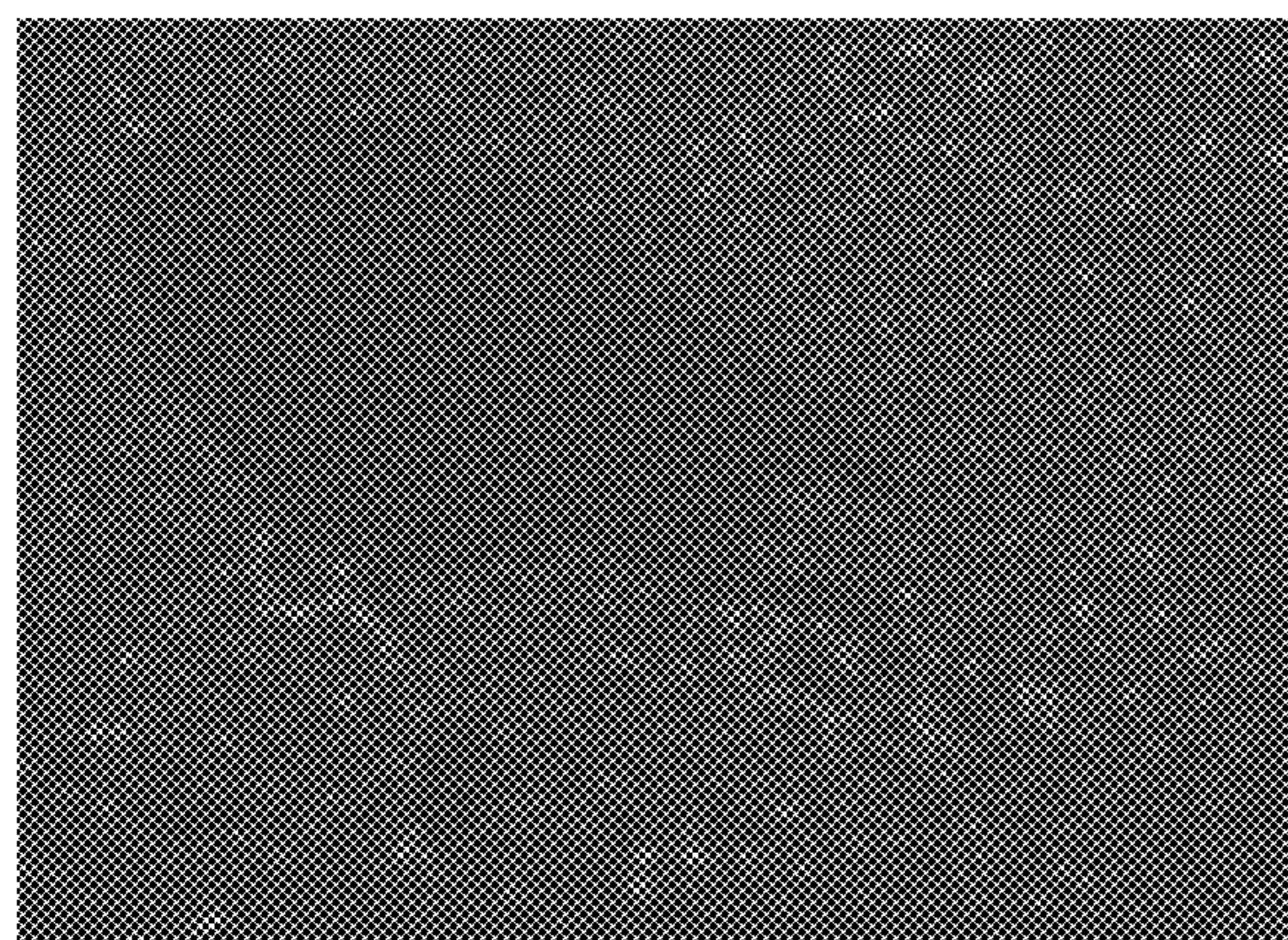
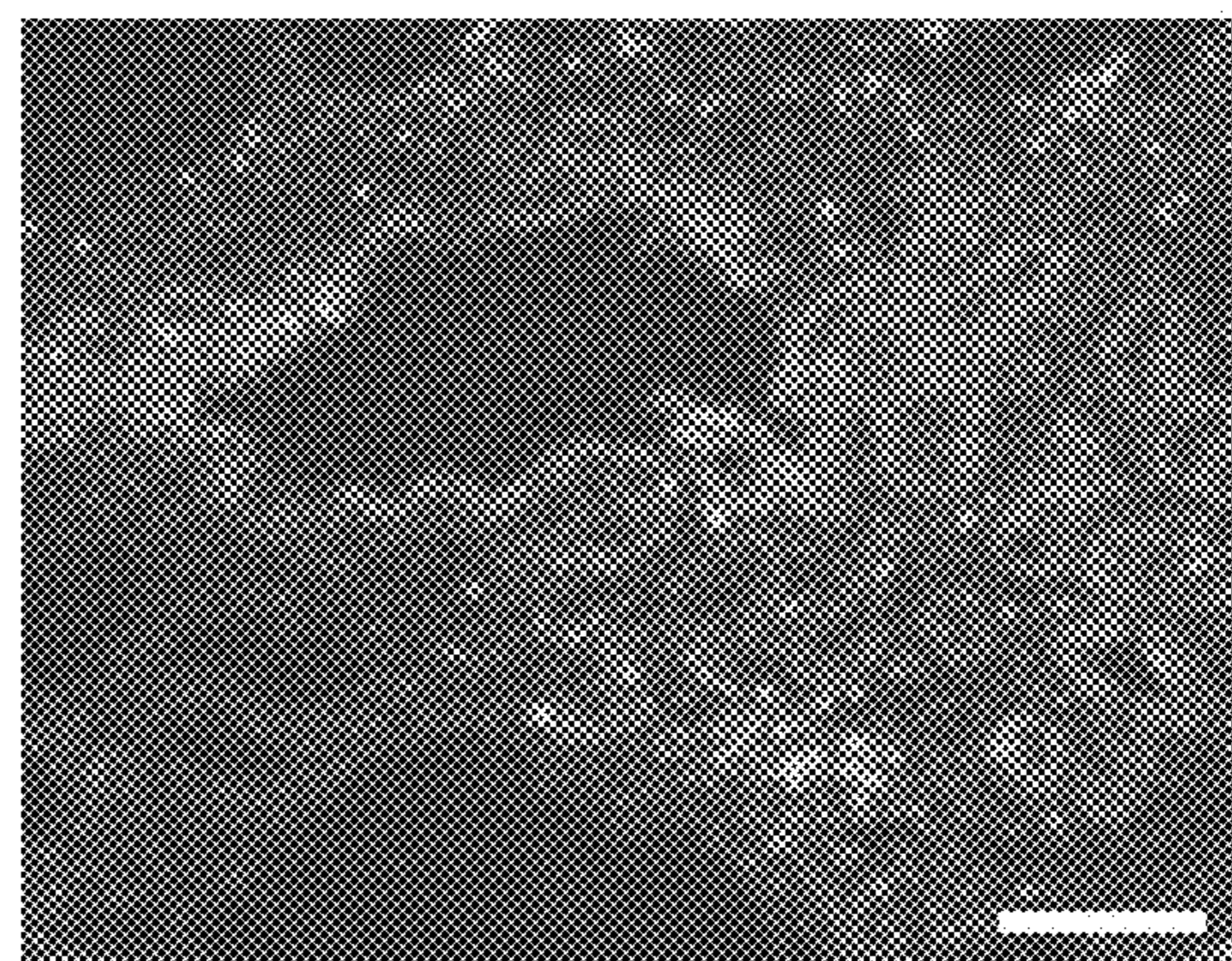


FIG. 12c



DAPI MMP9

FIG. 12d



DAPI DQ-Gelatin

FIG. 13a

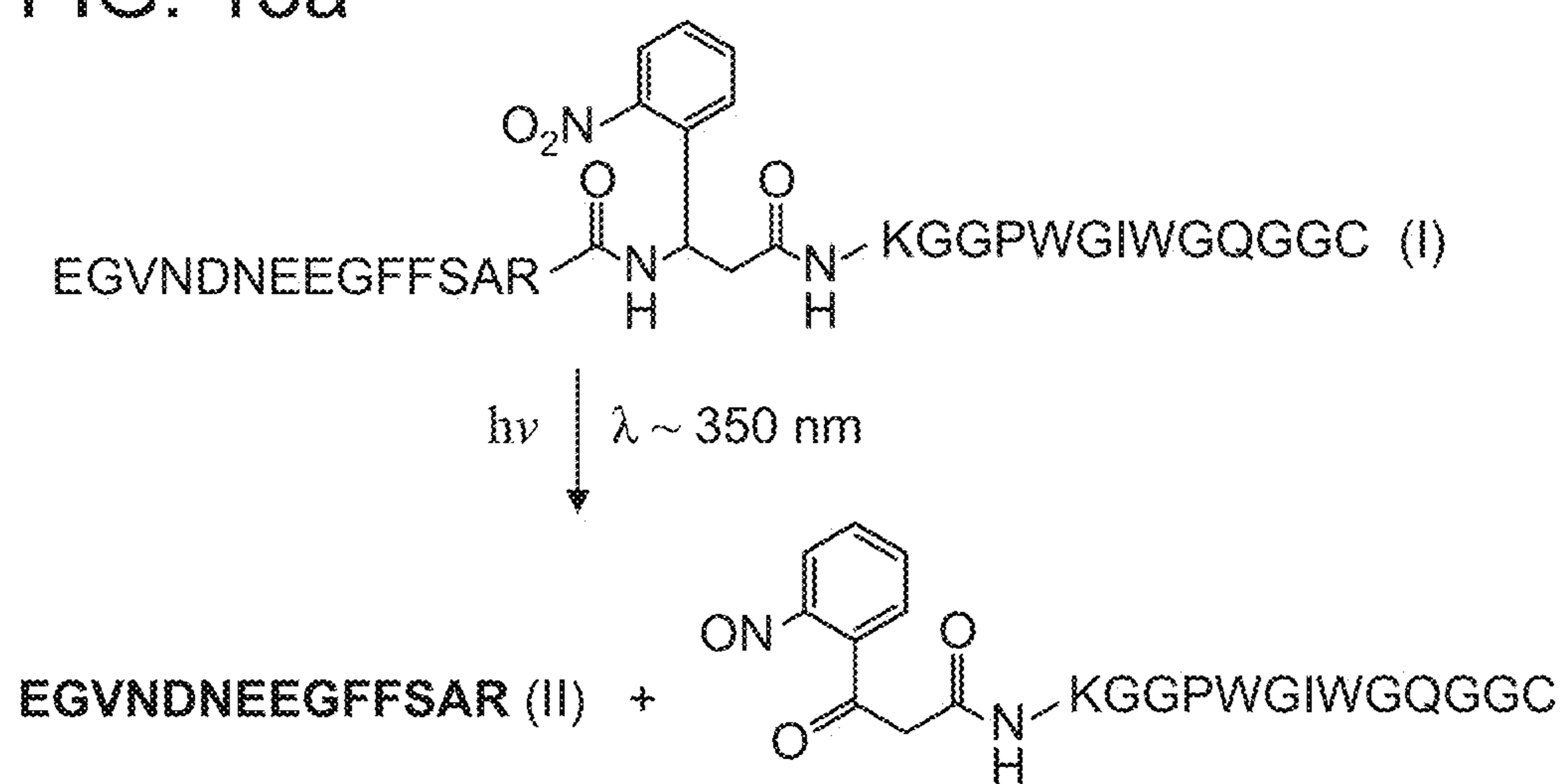


FIG. 13b

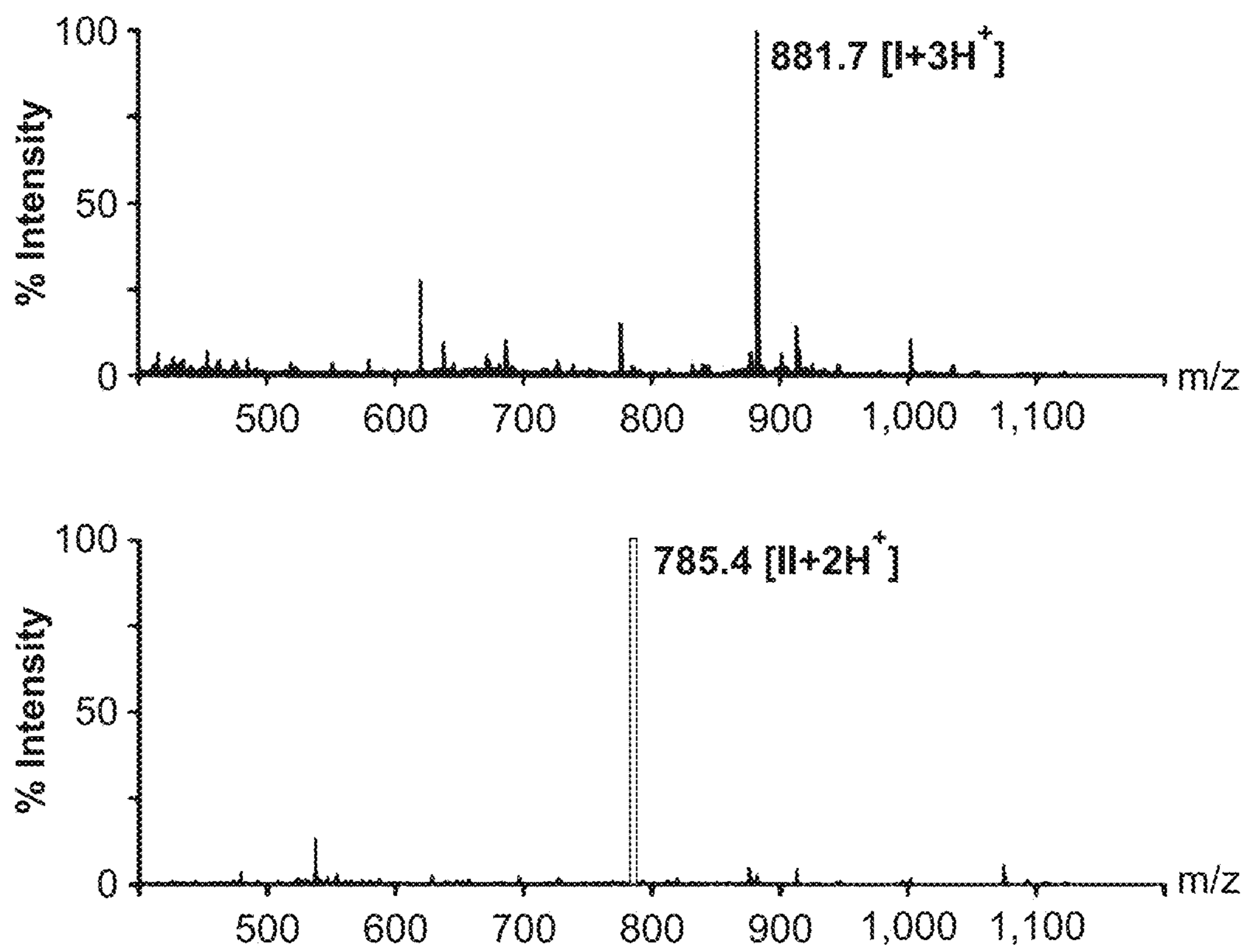


FIG. 13c

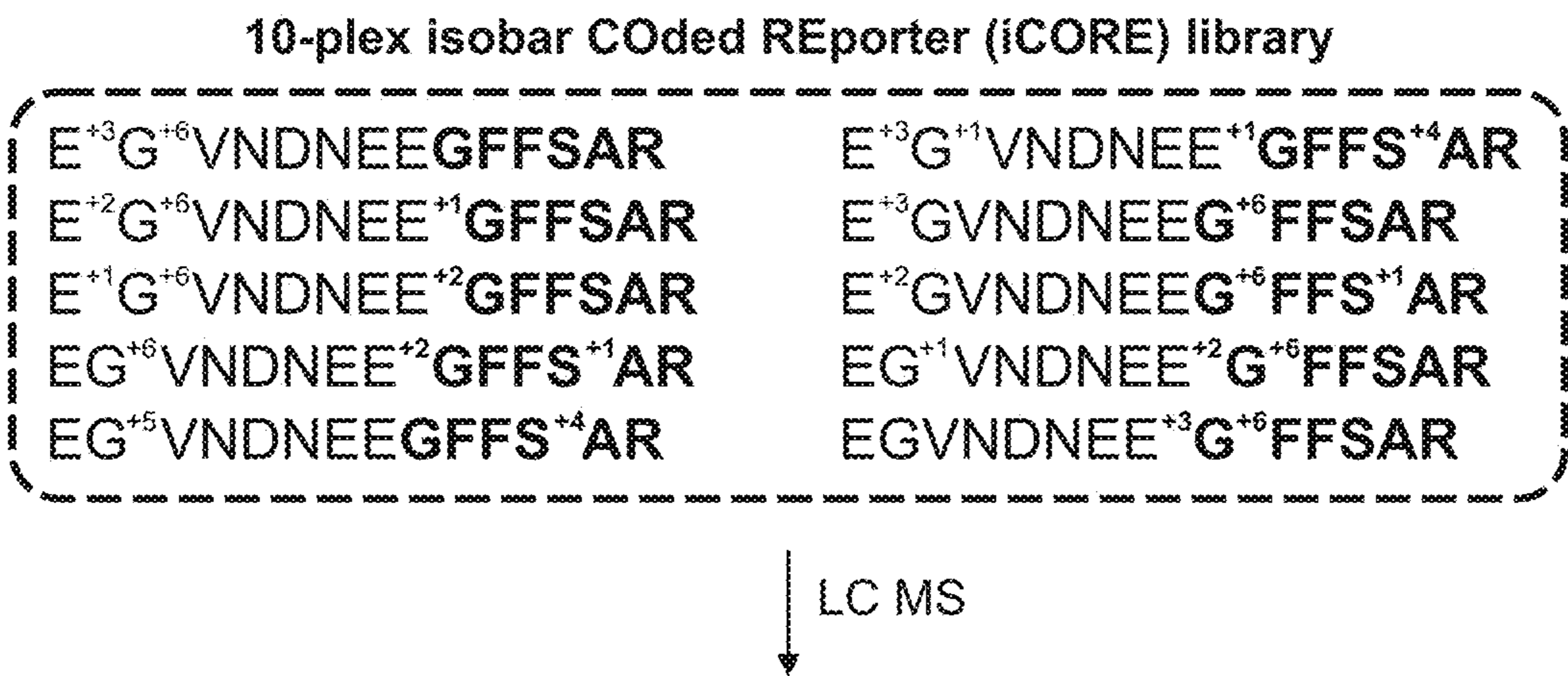


FIG. 13d

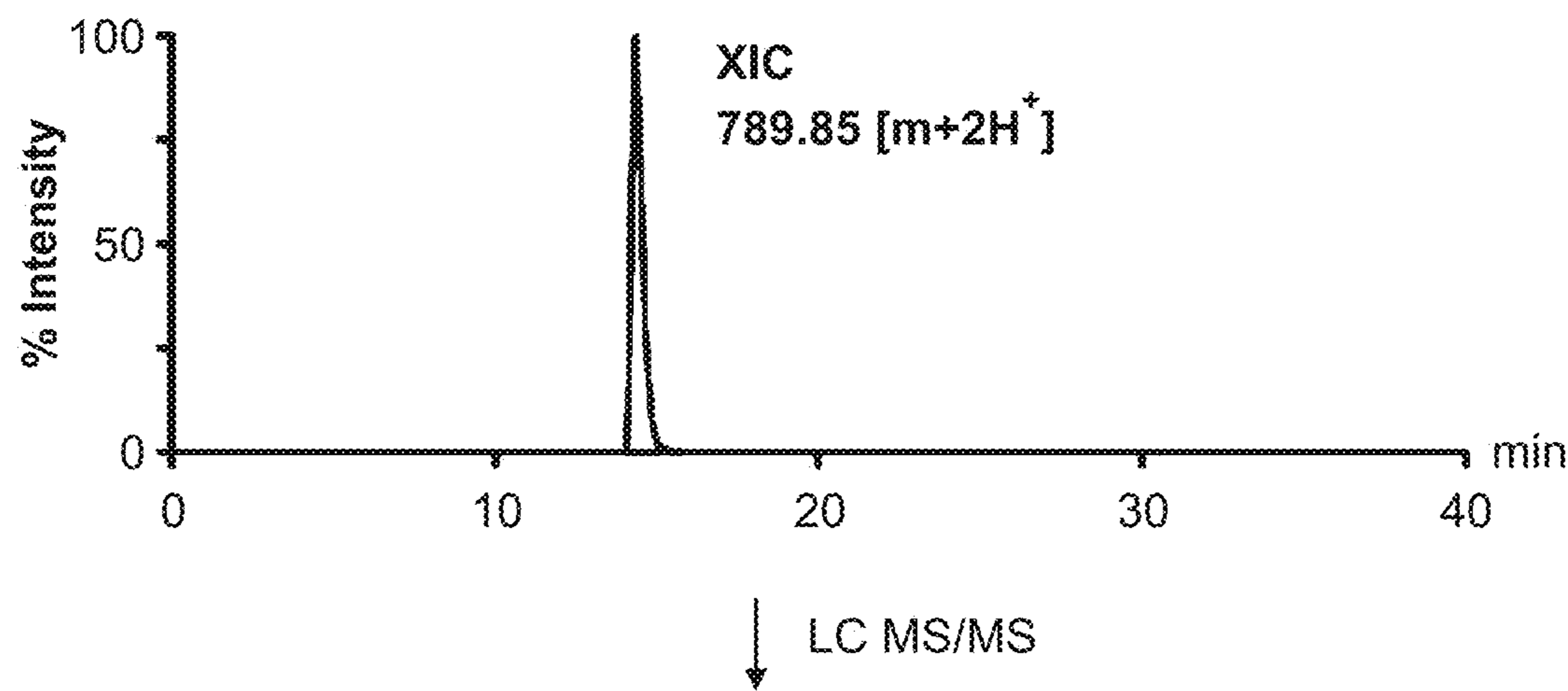


FIG. 13e

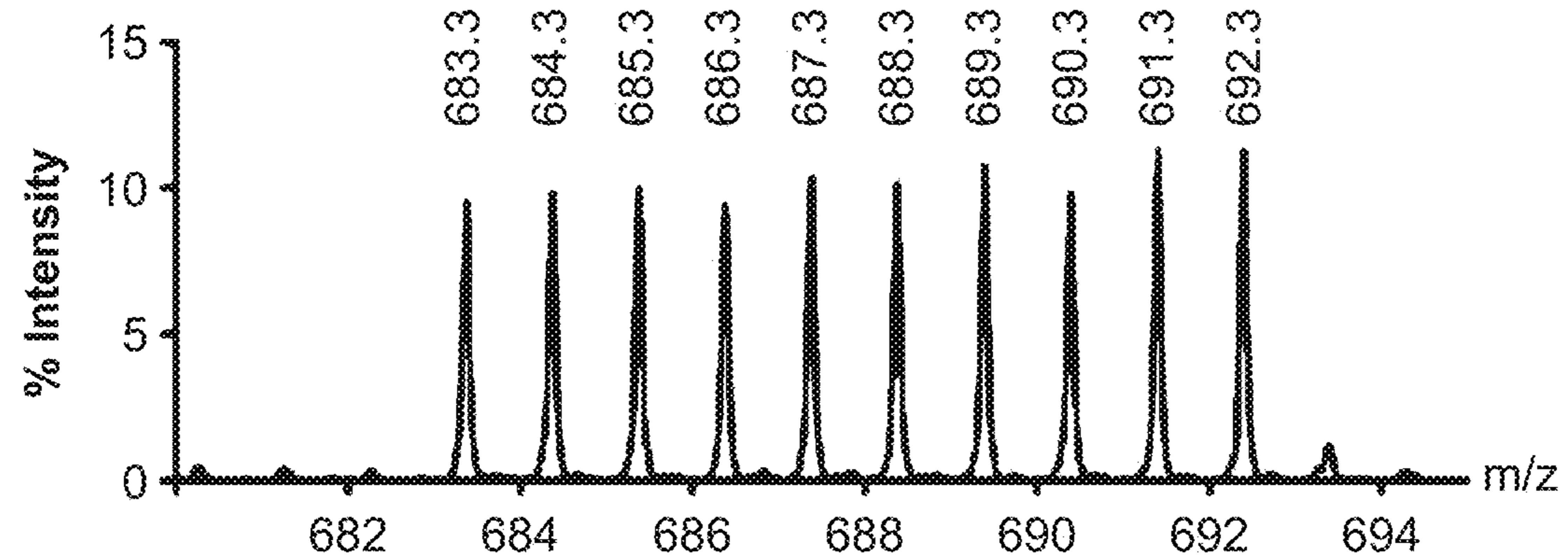


FIG. 13f

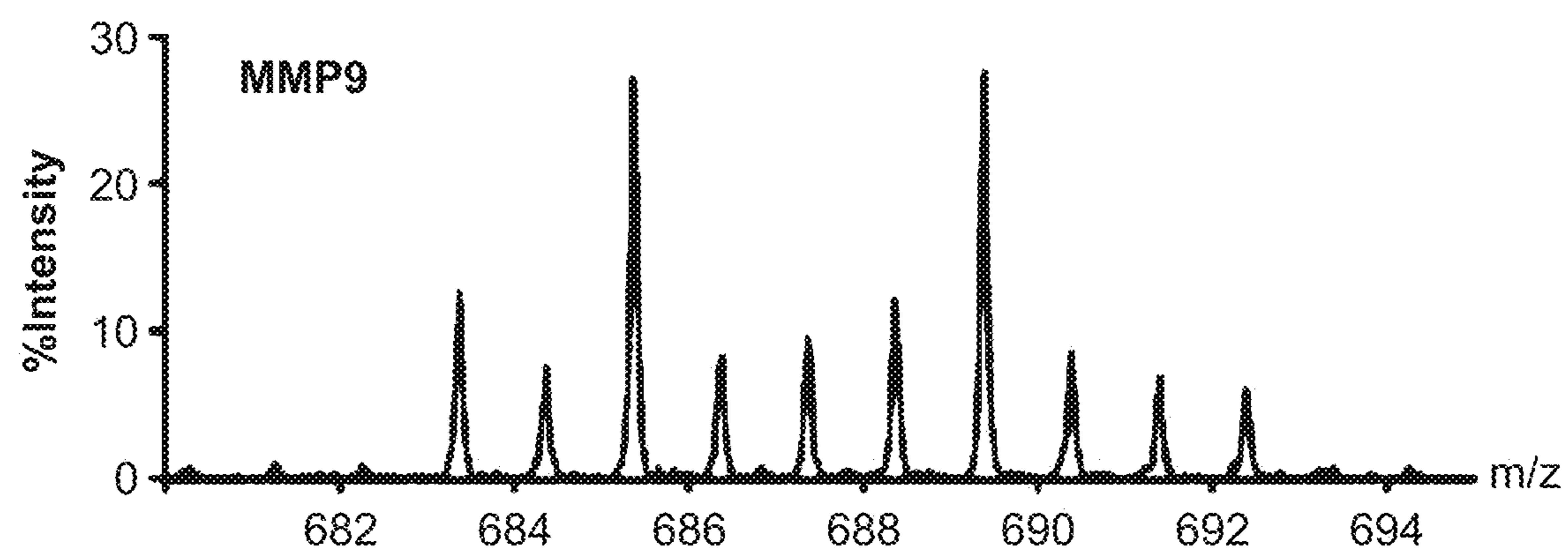
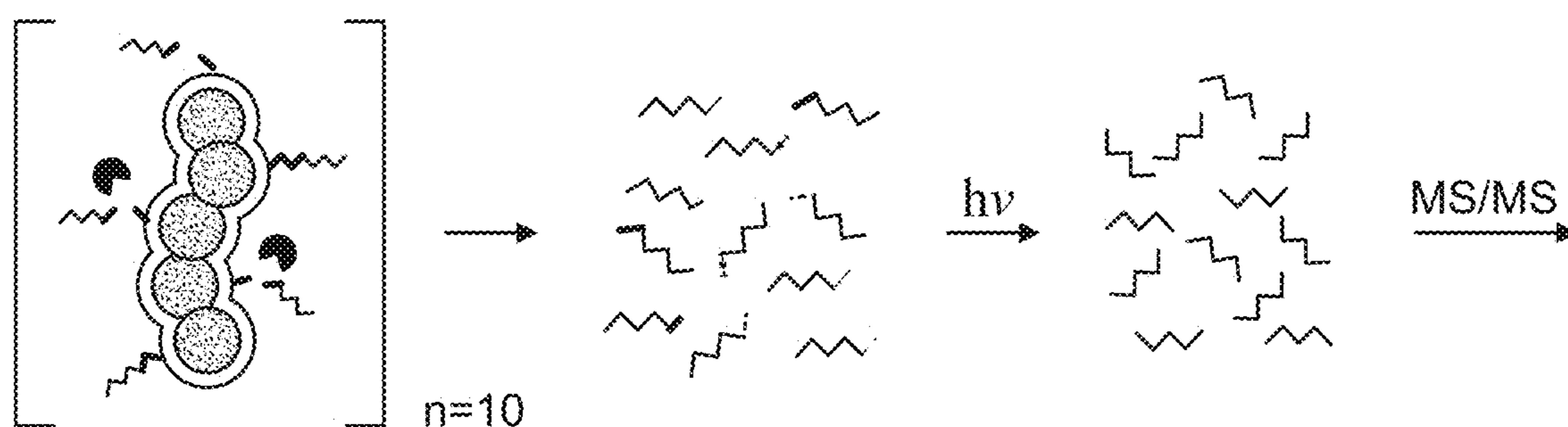


FIG. 14a

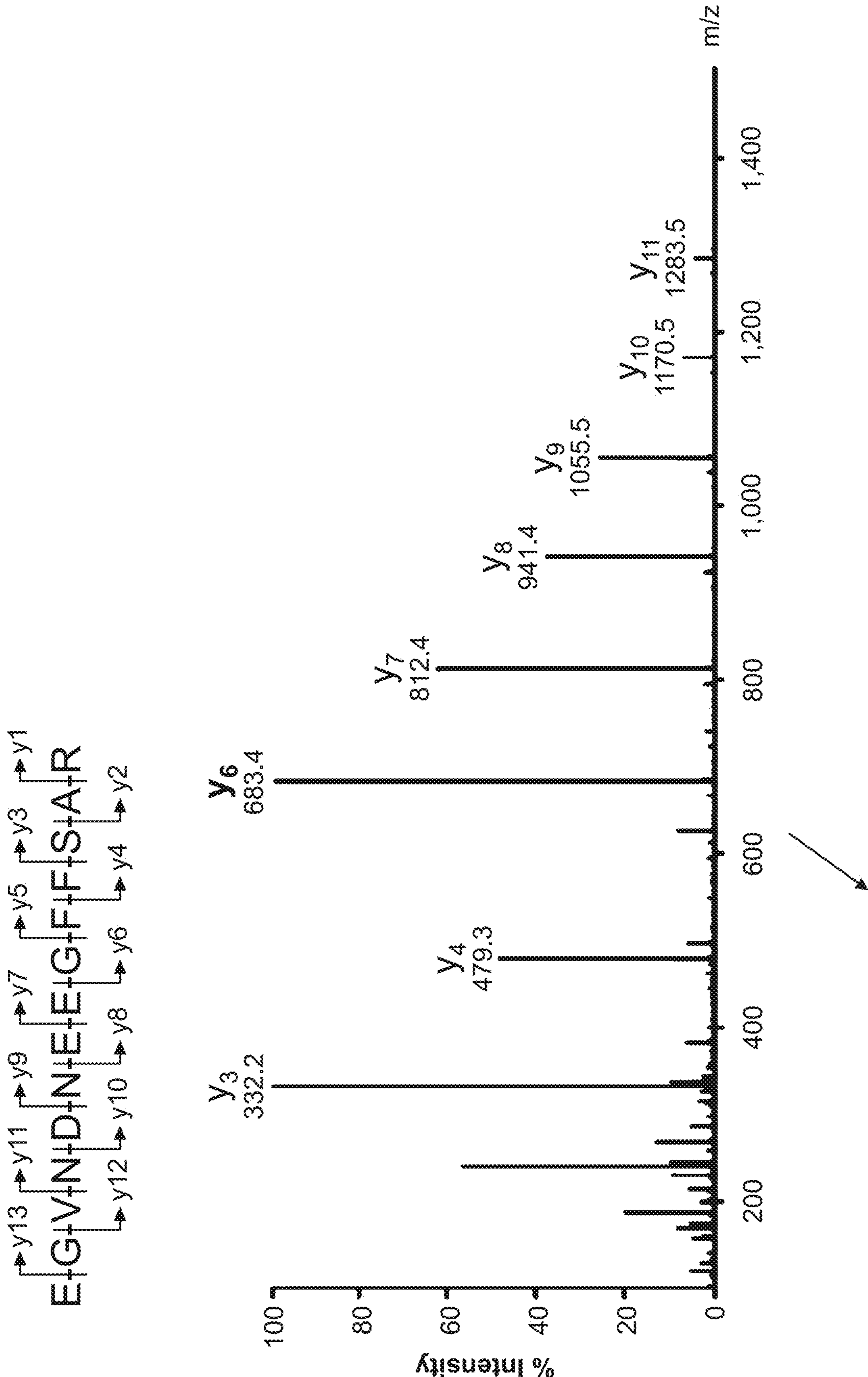


FIG. 14b

[X+H ⁺] mass (Da)		
Balance	Reporter	Total
E ⁺³ G ⁺⁶ VNDNEE-GFFSAR	683.3	1578.7
E ⁺² G ⁺⁶ VNDNEE- ⁺¹ GFFSAR	684.3	1578.7
E ⁺¹ G ⁺⁶ VNDNEE- ⁺² GFFSAR	685.3	1578.7
EG ⁺⁶ VNDNEE- ⁺² GFFS ⁺¹ AR	686.3	1578.7
EG ⁺⁵ VNDNEE-GFFS ⁺⁴ AR	687.3	1578.7
E ⁺³ G ⁺¹ VNDNEE- ⁺¹ GFFS ⁺⁴ AR	688.3	1578.7
E ⁺³ GVNDNEE-G ⁺⁶ FFSAR	689.3	1578.7
E ⁺² GVNDNEE-G ⁺⁶ FFS ⁺¹ AR	690.3	1578.7
EG ⁺¹ VNDNEE- ⁺² G ⁺⁶ FFSAR	691.3	1578.7
EGVNDNEE- ⁺³ G ⁺⁶ FFSAR	692.3	1578.7

FIG. 15

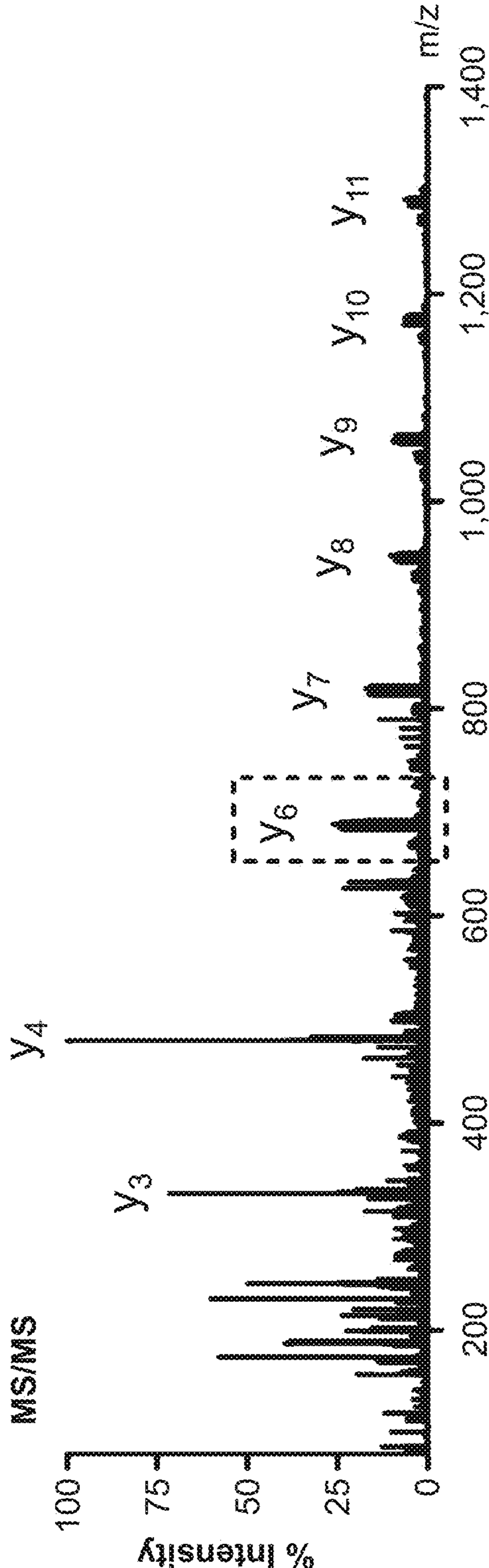


FIG. 16a

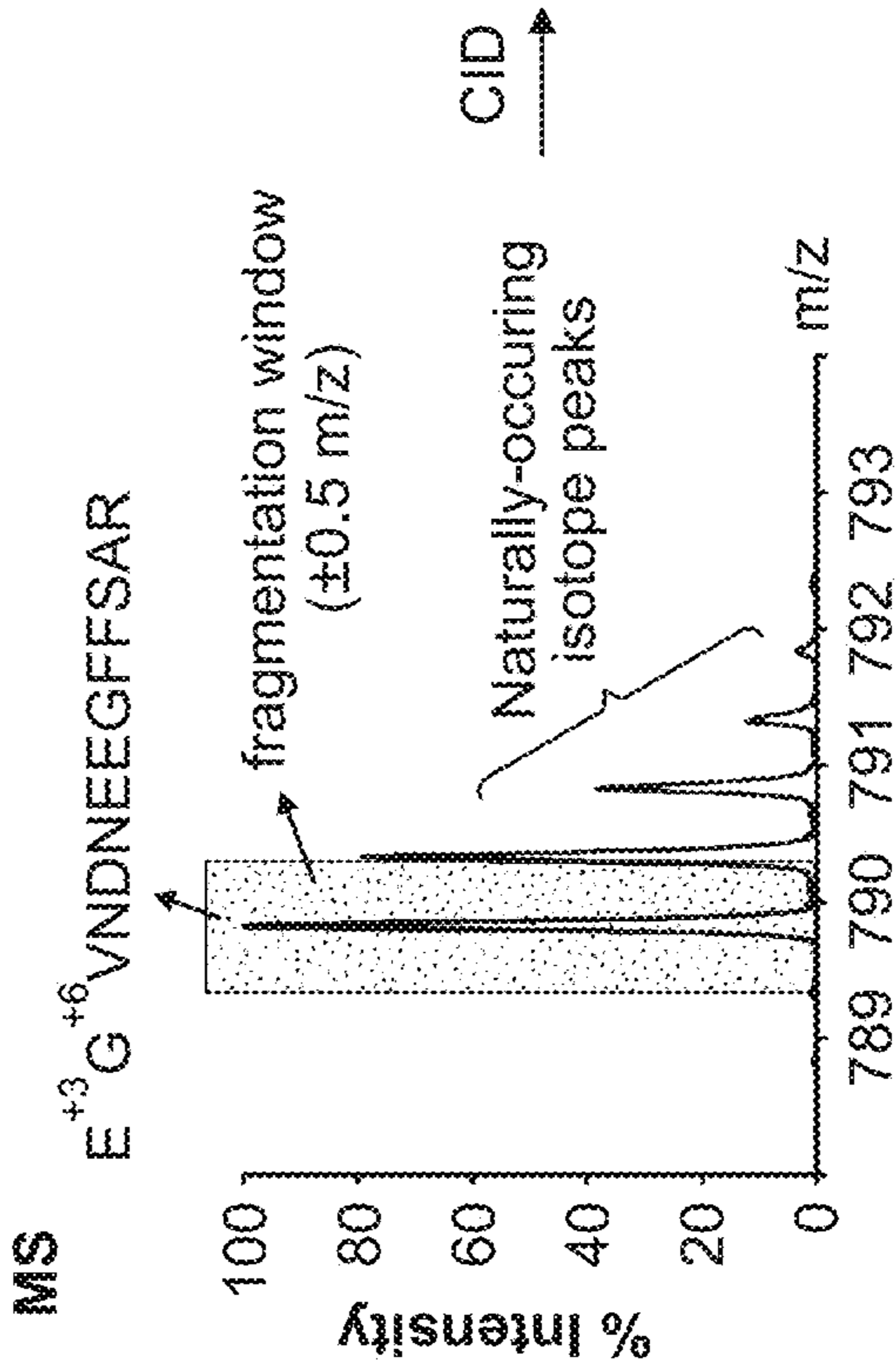


FIG. 16b

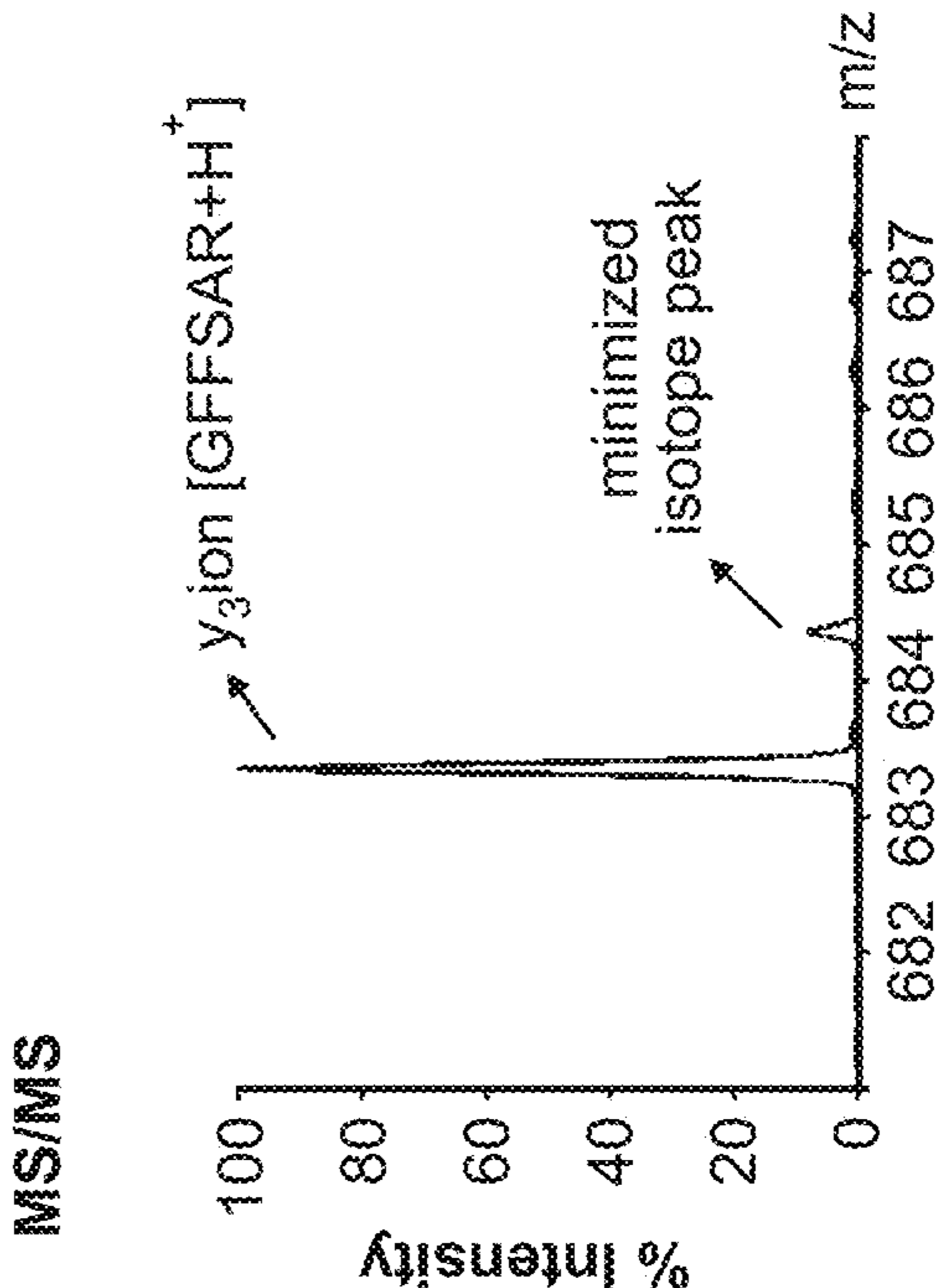


FIG. 17a

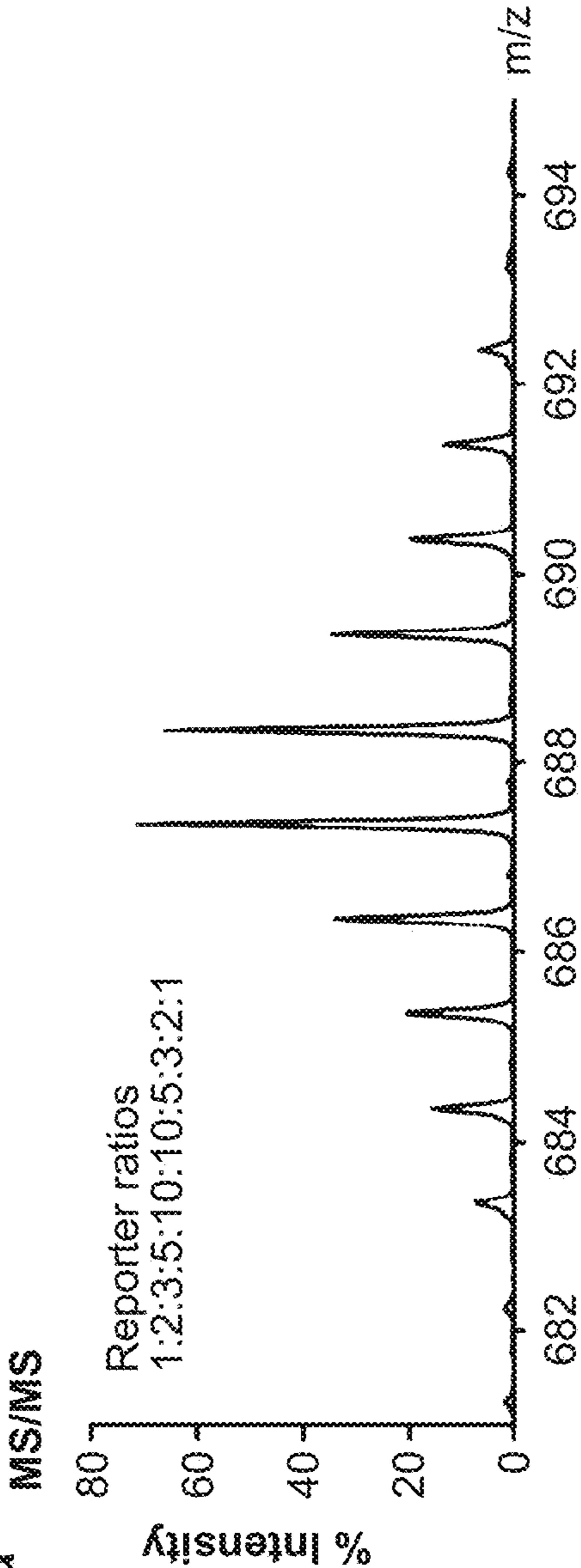


FIG. 17b

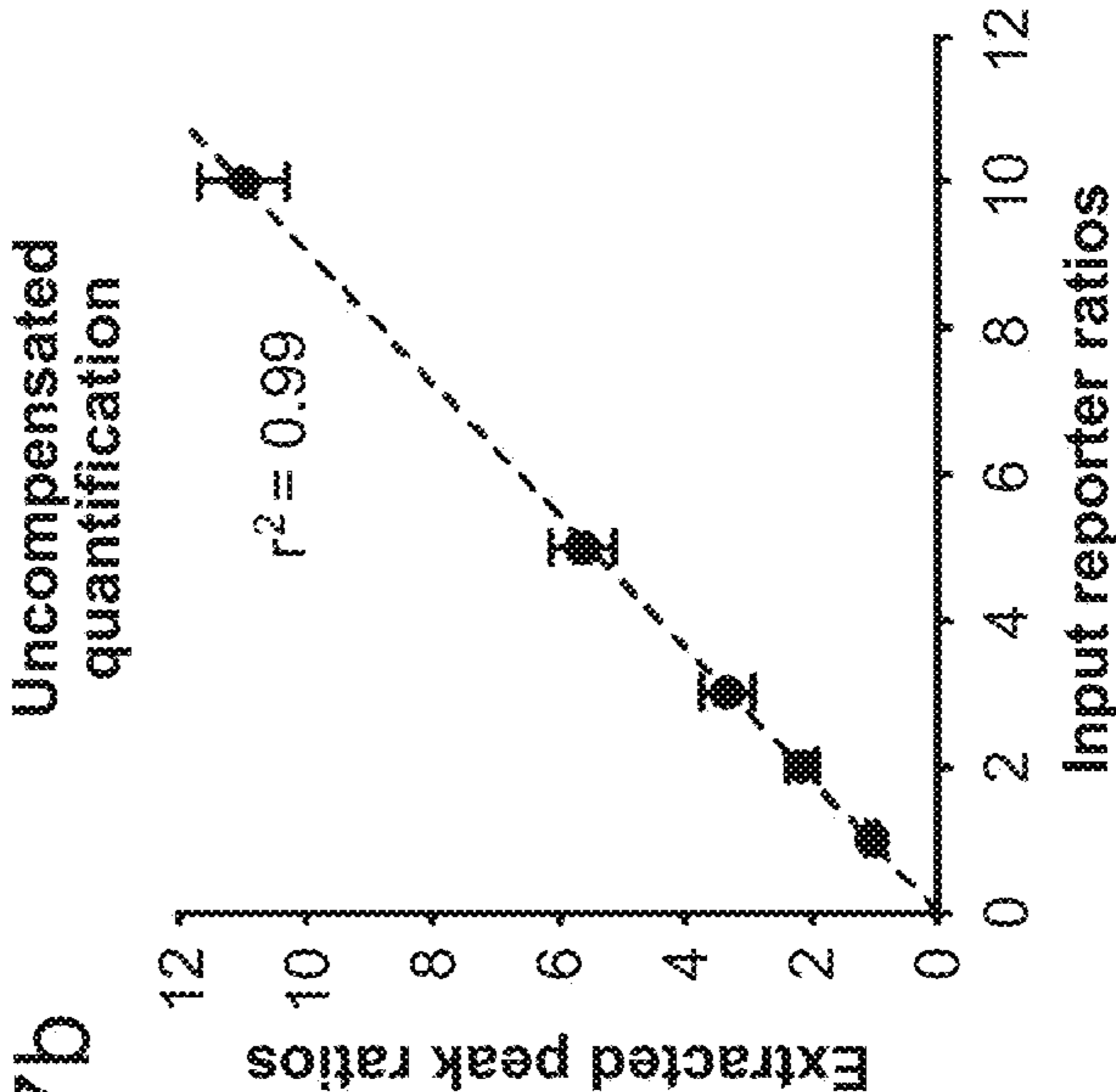


FIG. 17c

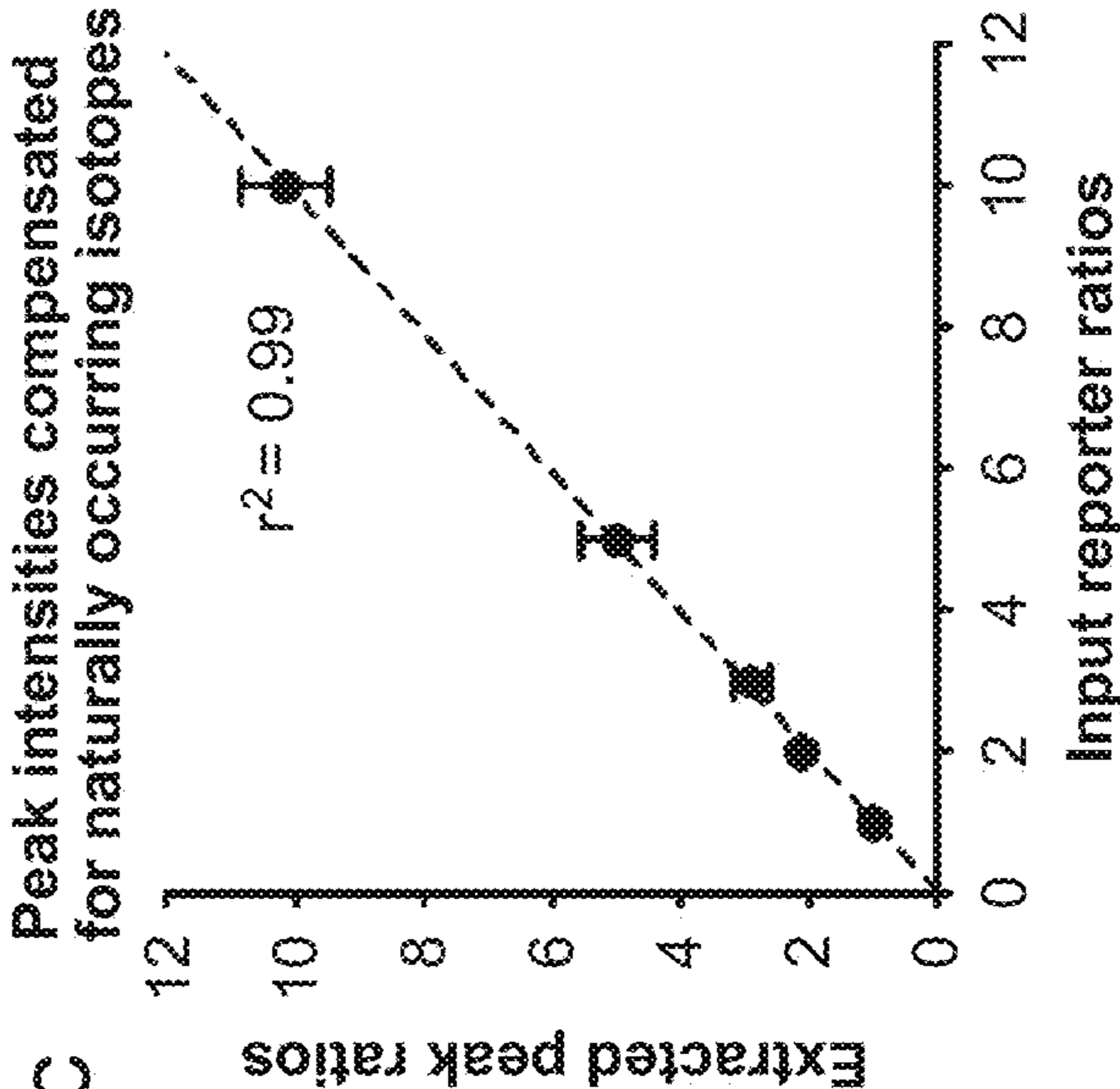


FIG. 18a

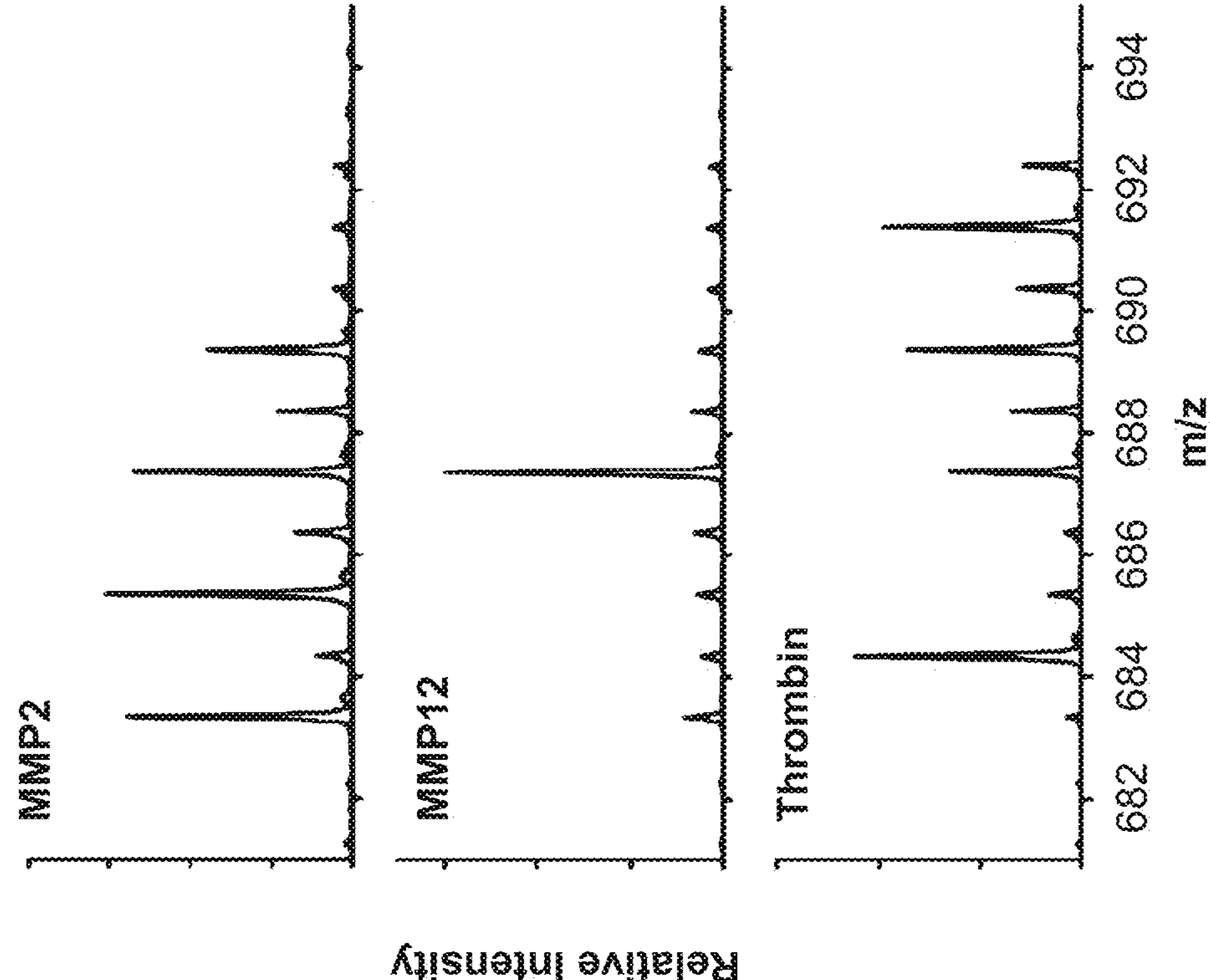


FIG. 18b

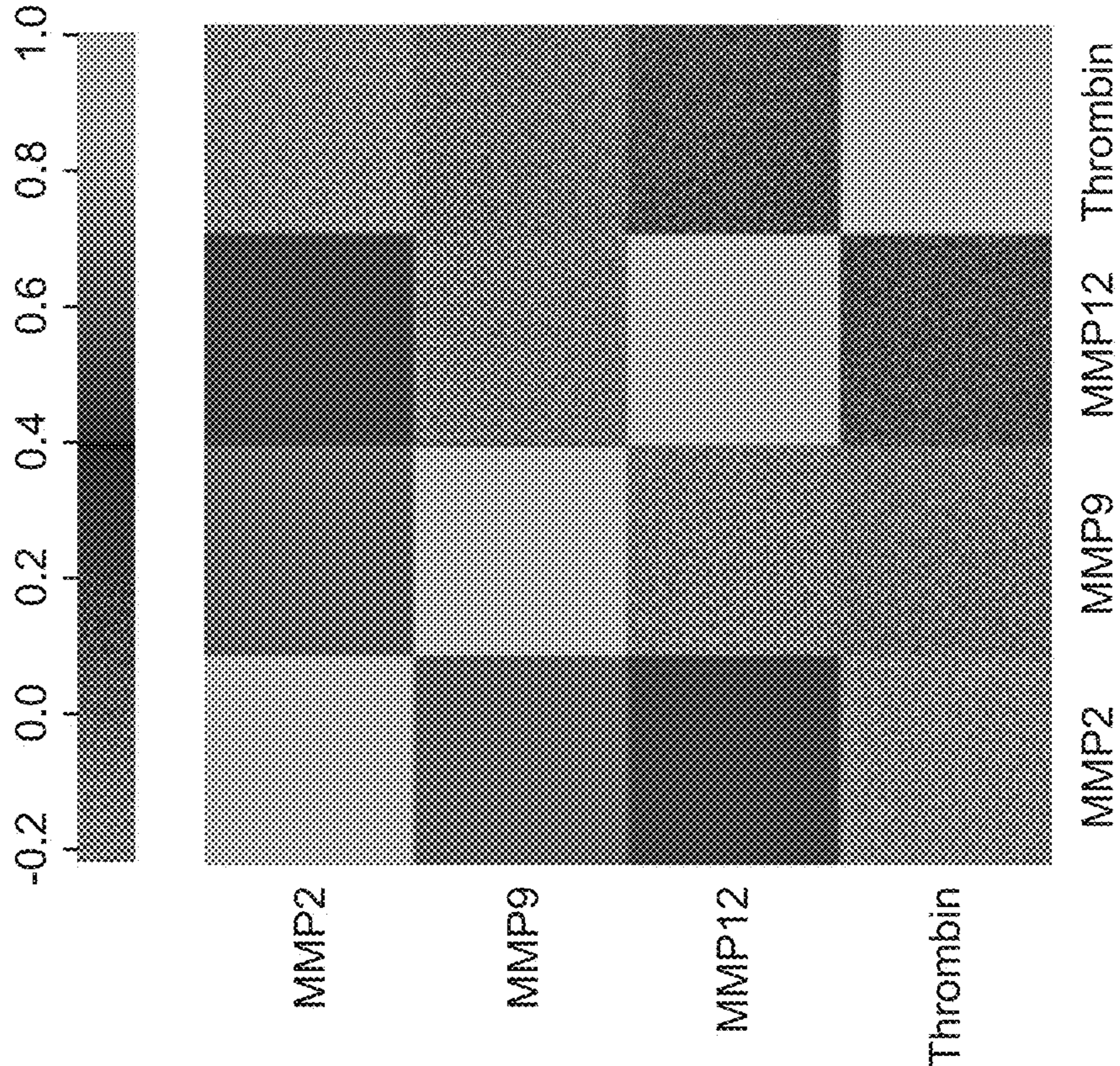


FIG. 19

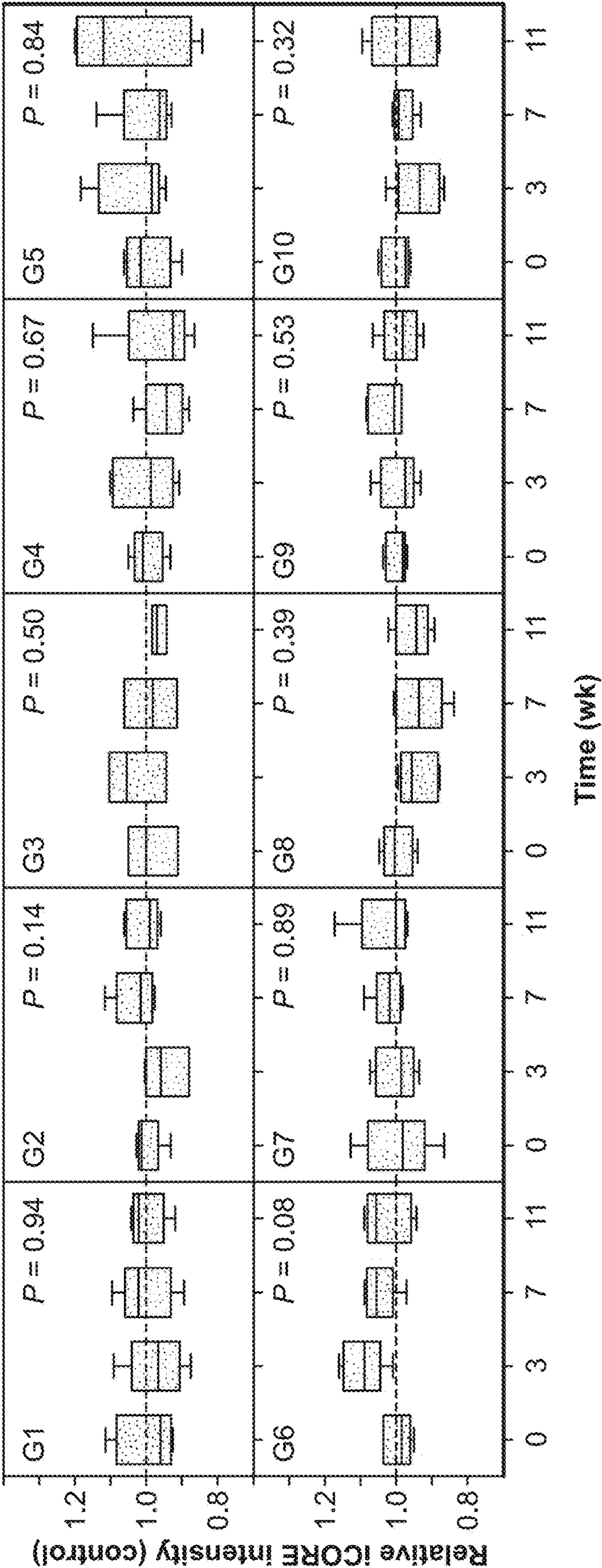


FIG. 20A

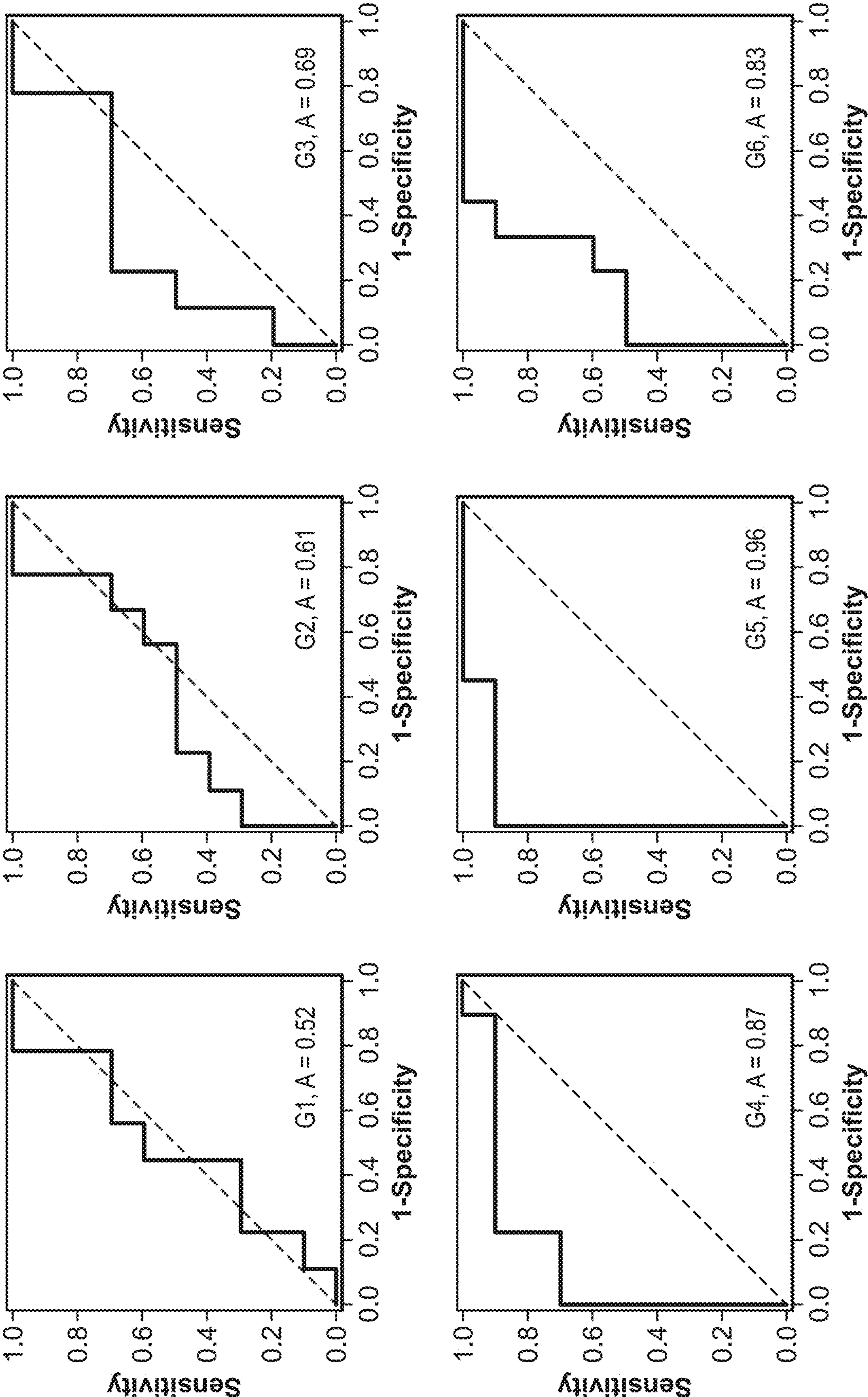


FIG. 20B

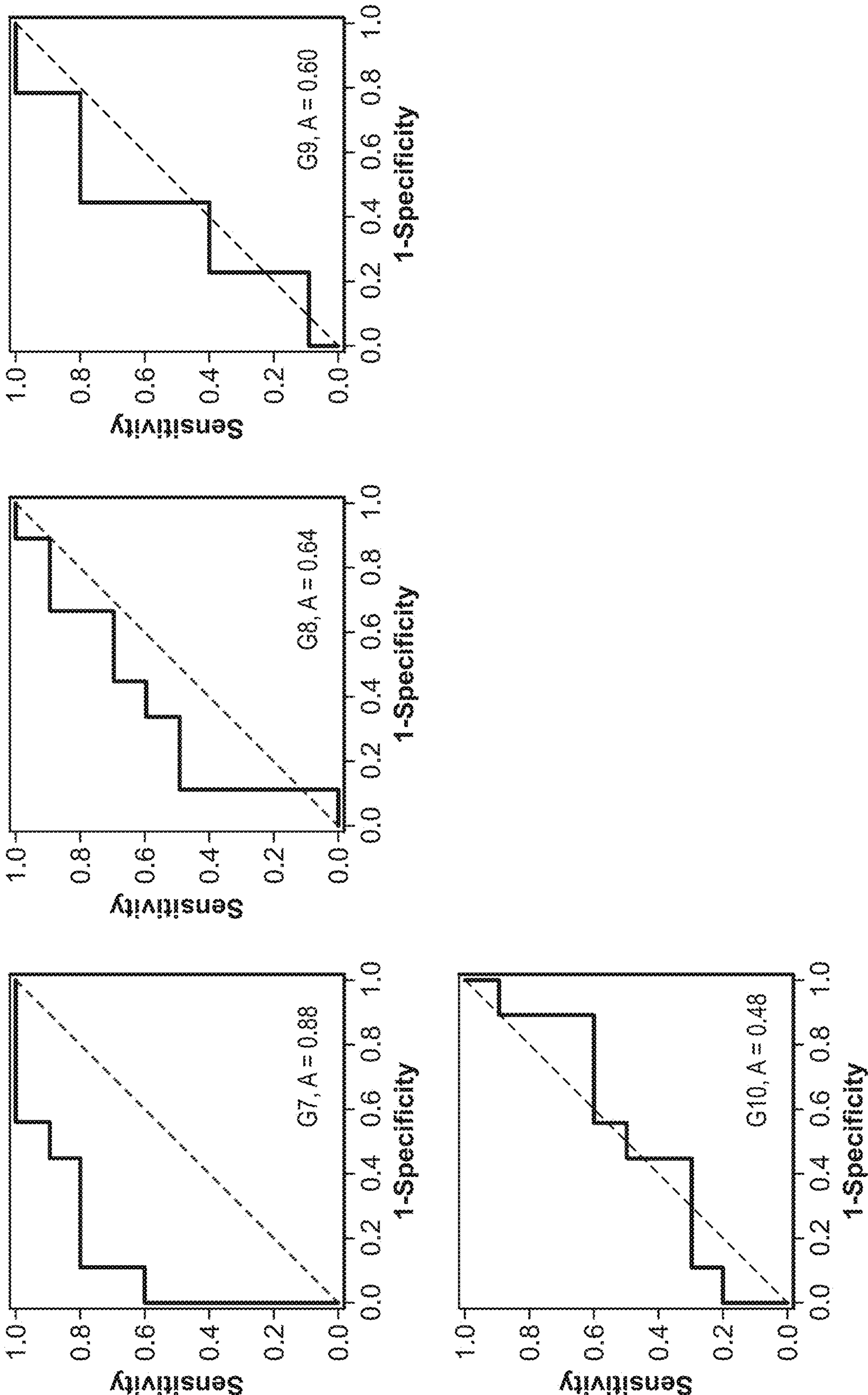


FIG. 21A

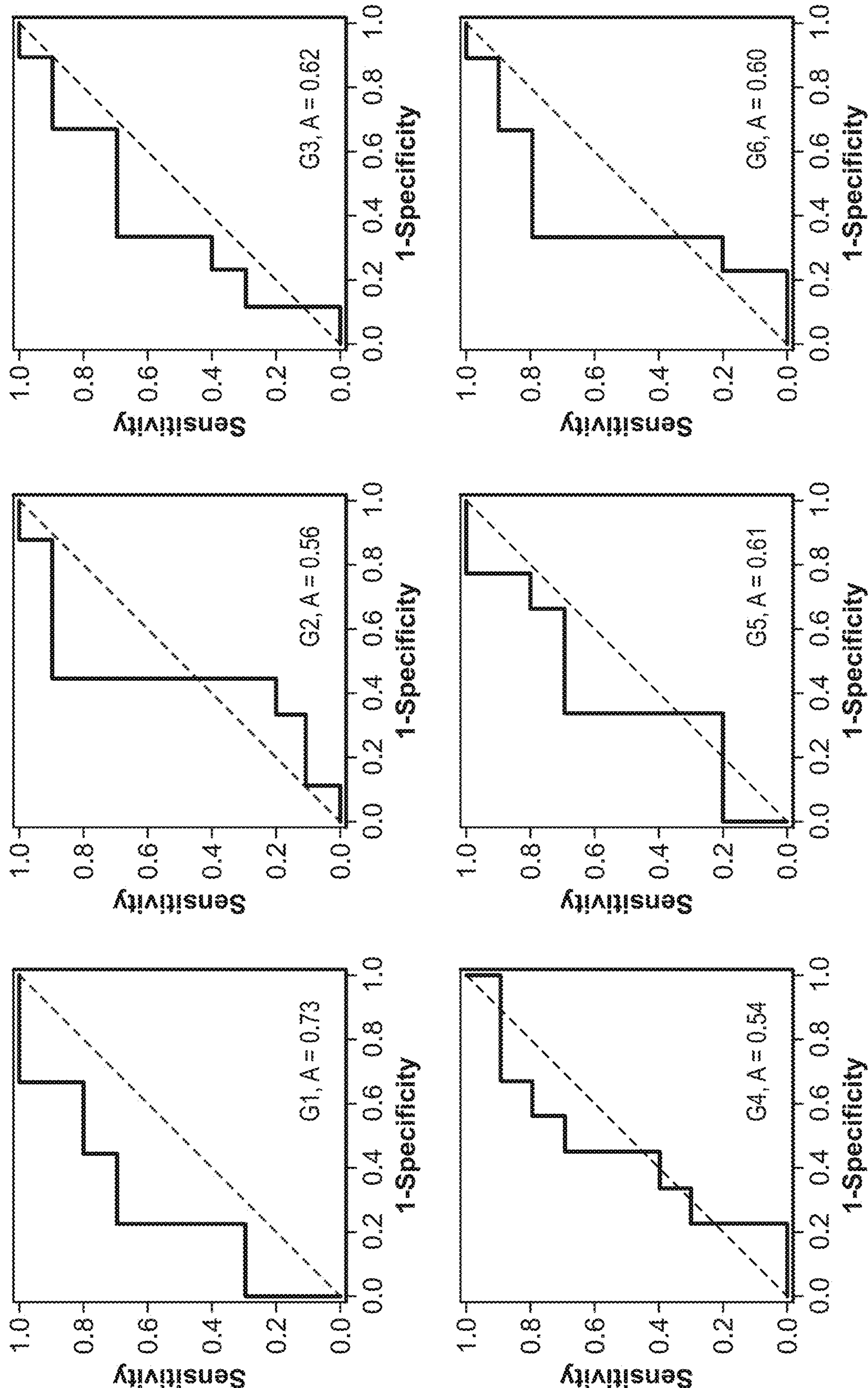


FIG. 21B

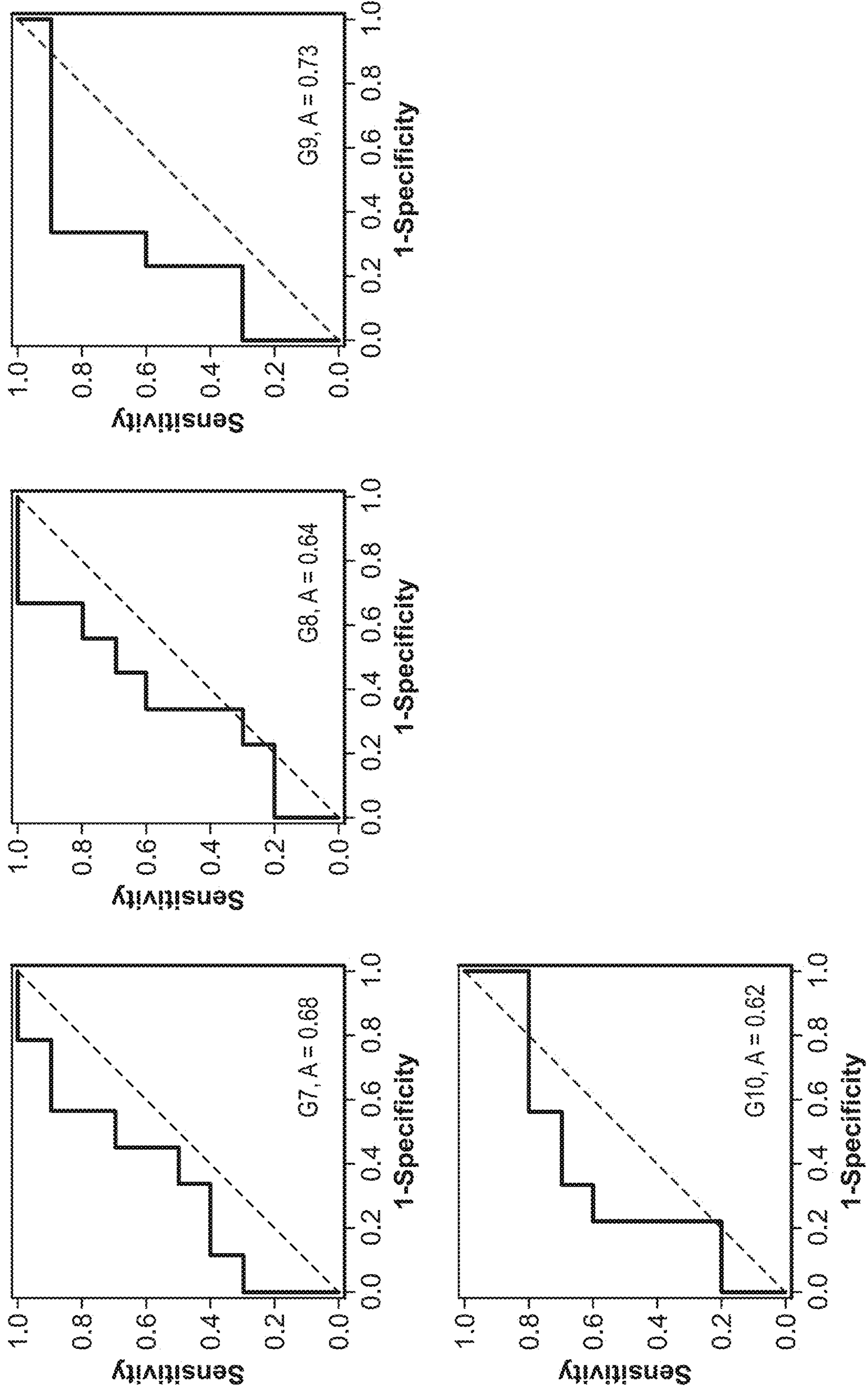


FIG. 22

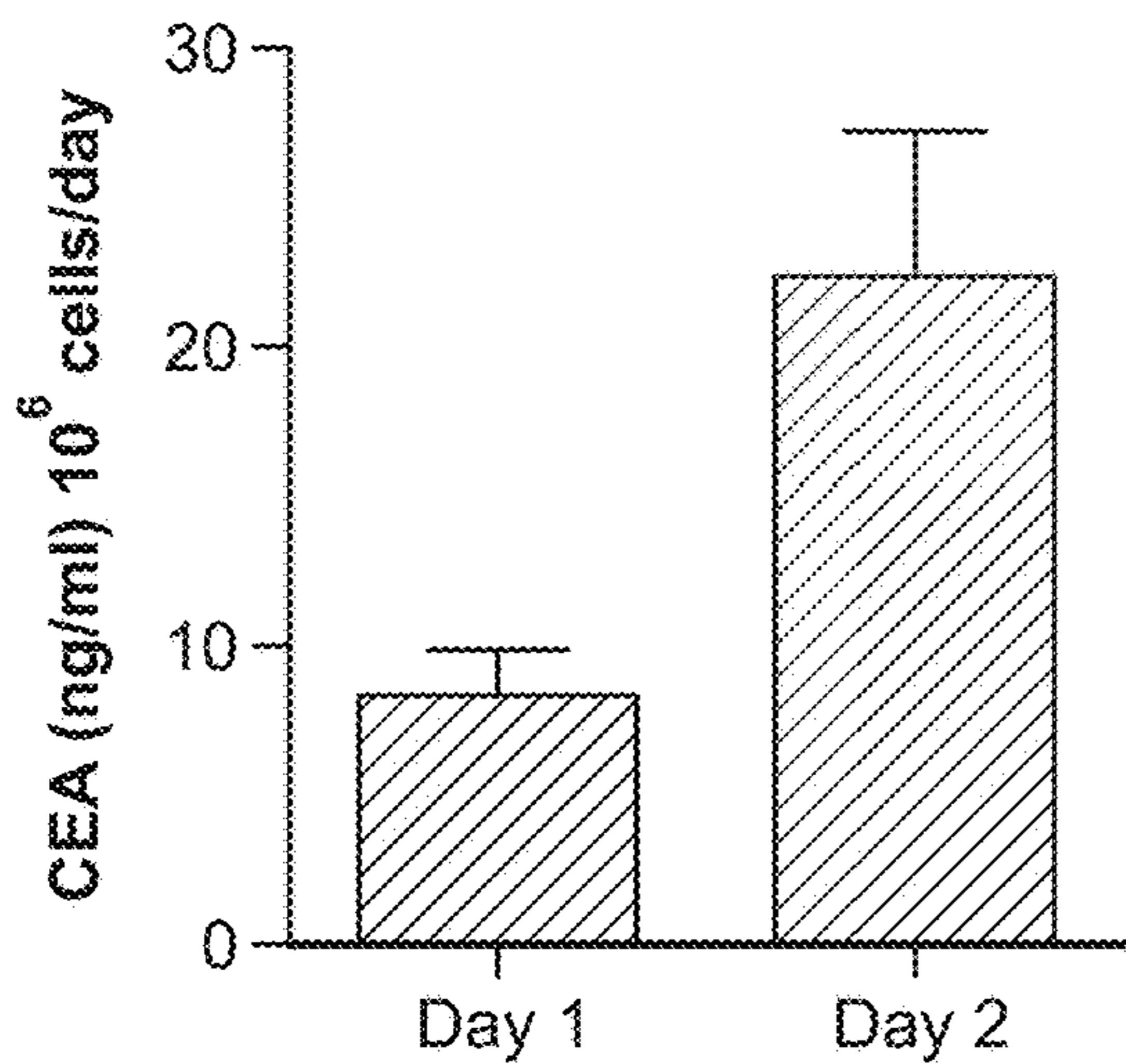


FIG. 23d

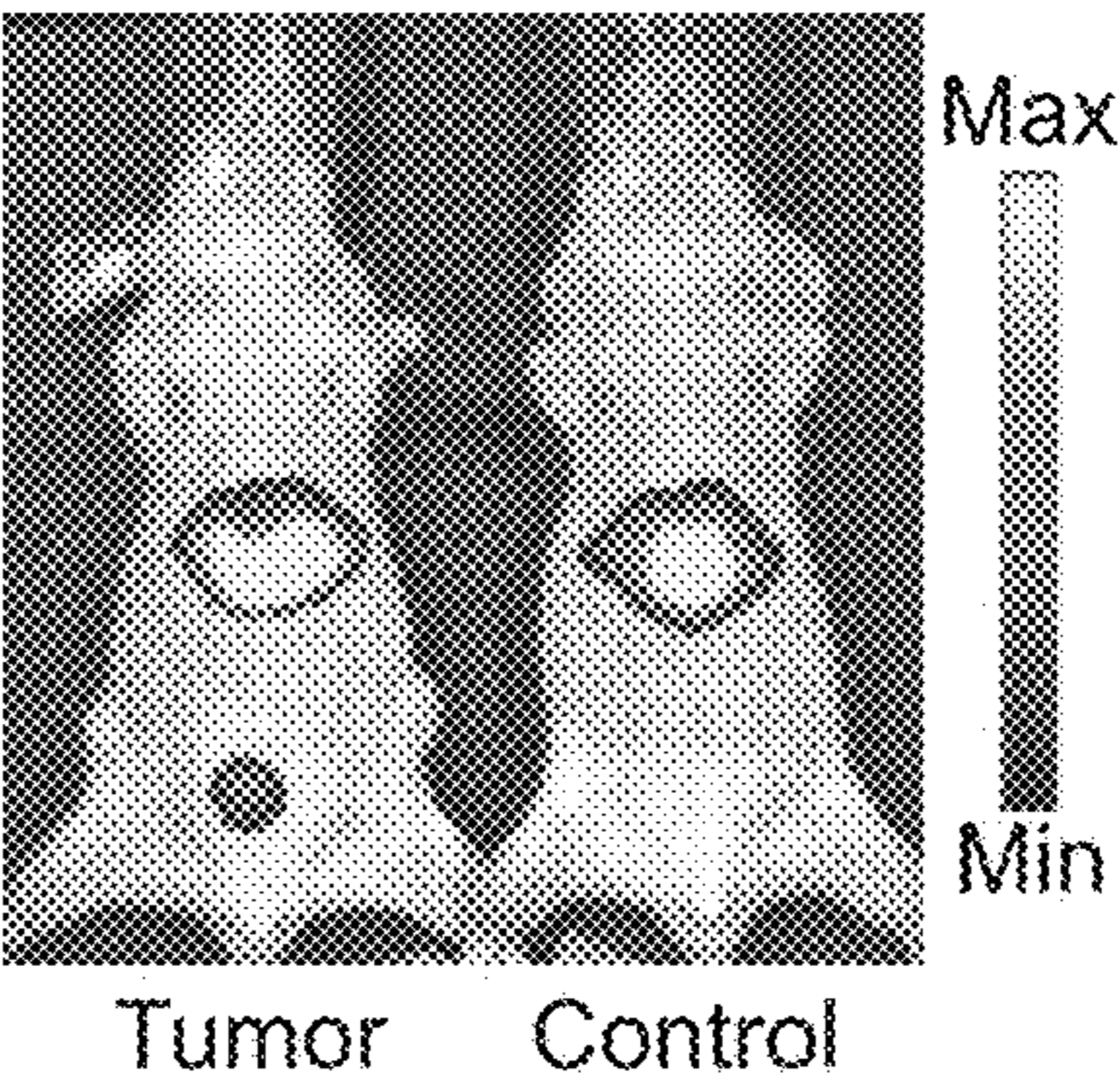


FIG. 23a

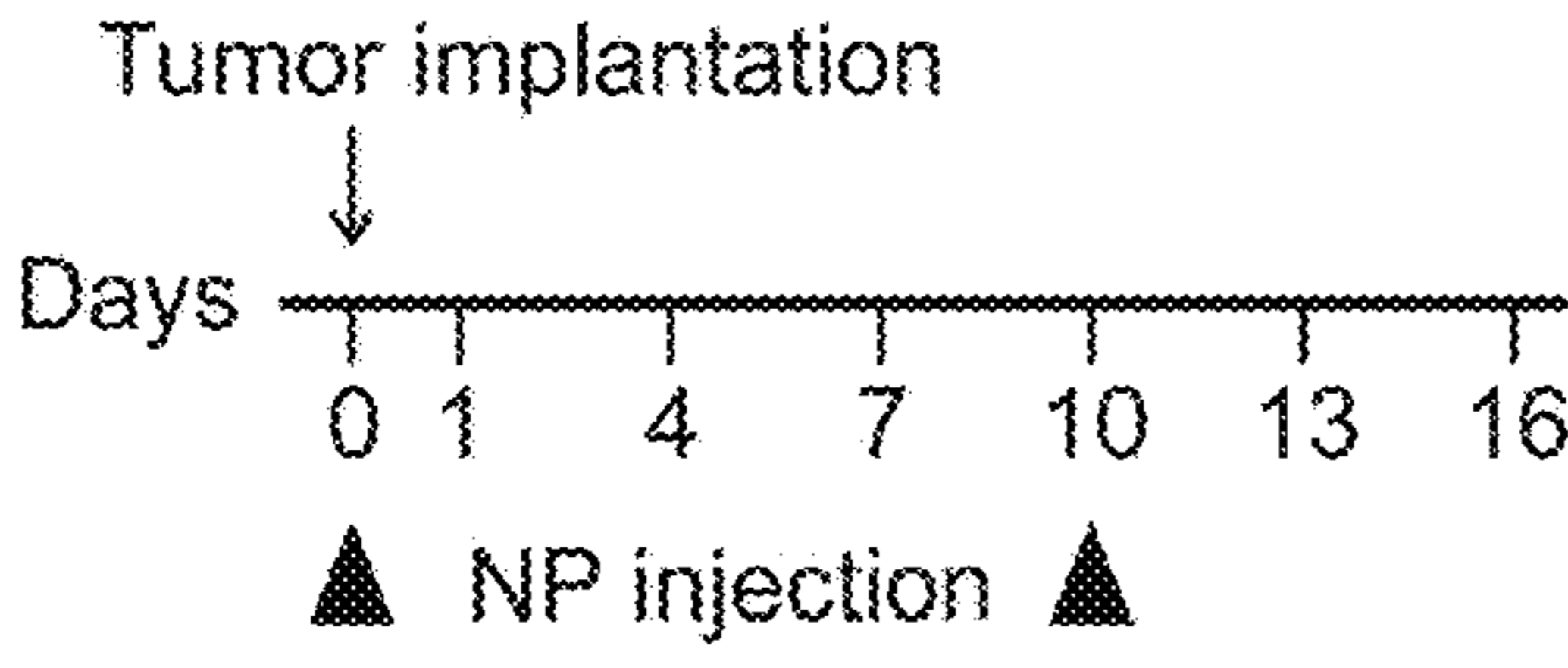


FIG. 23e

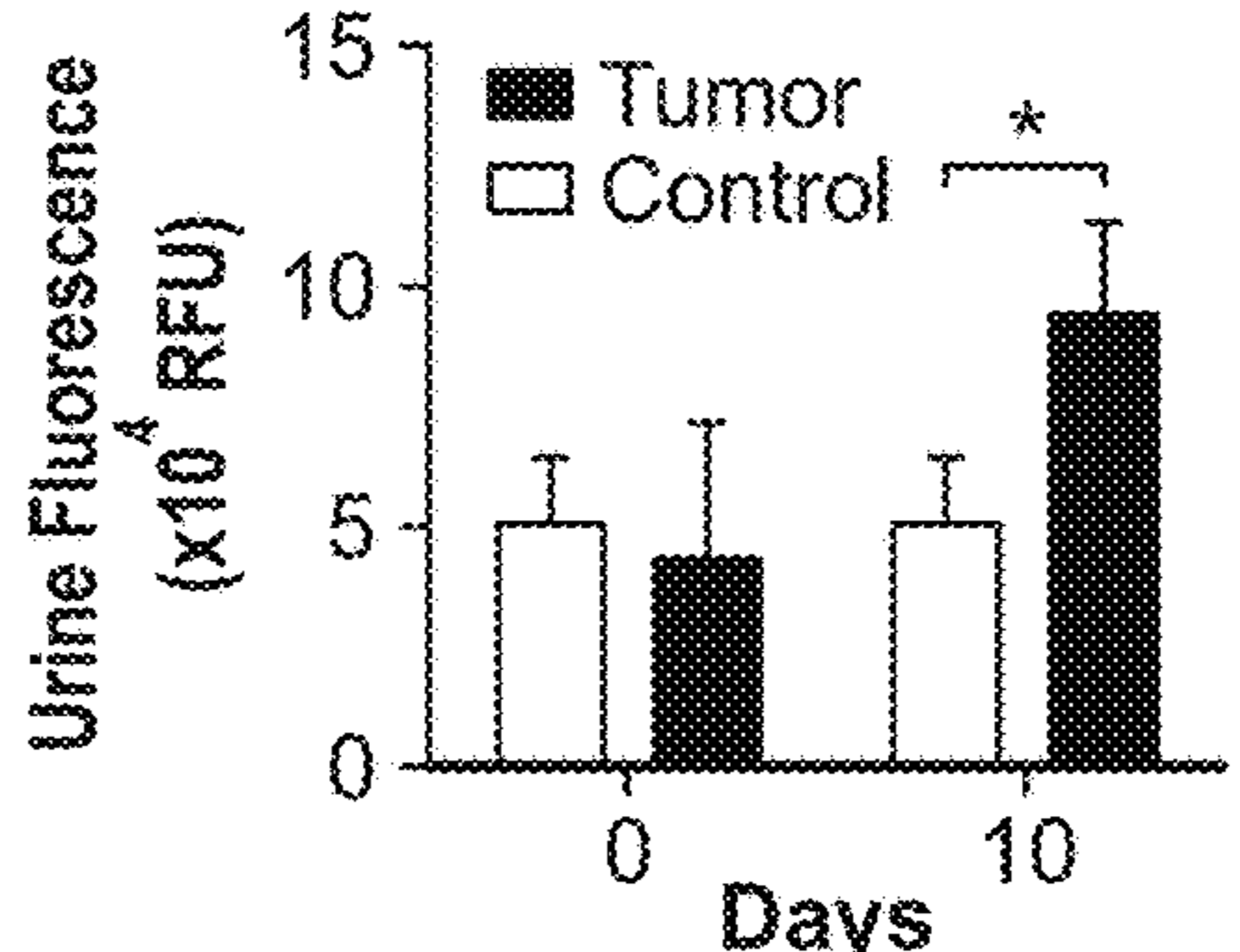


FIG. 23b

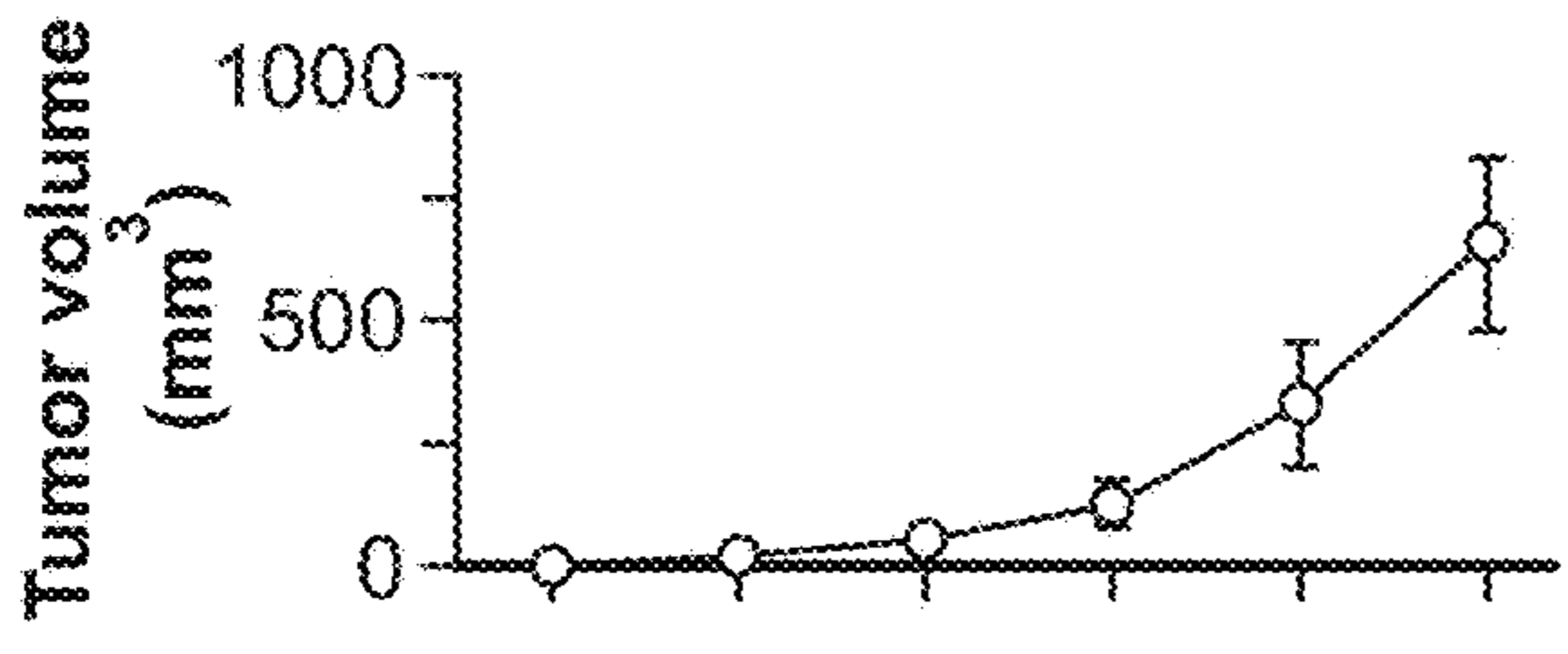


FIG. 23f

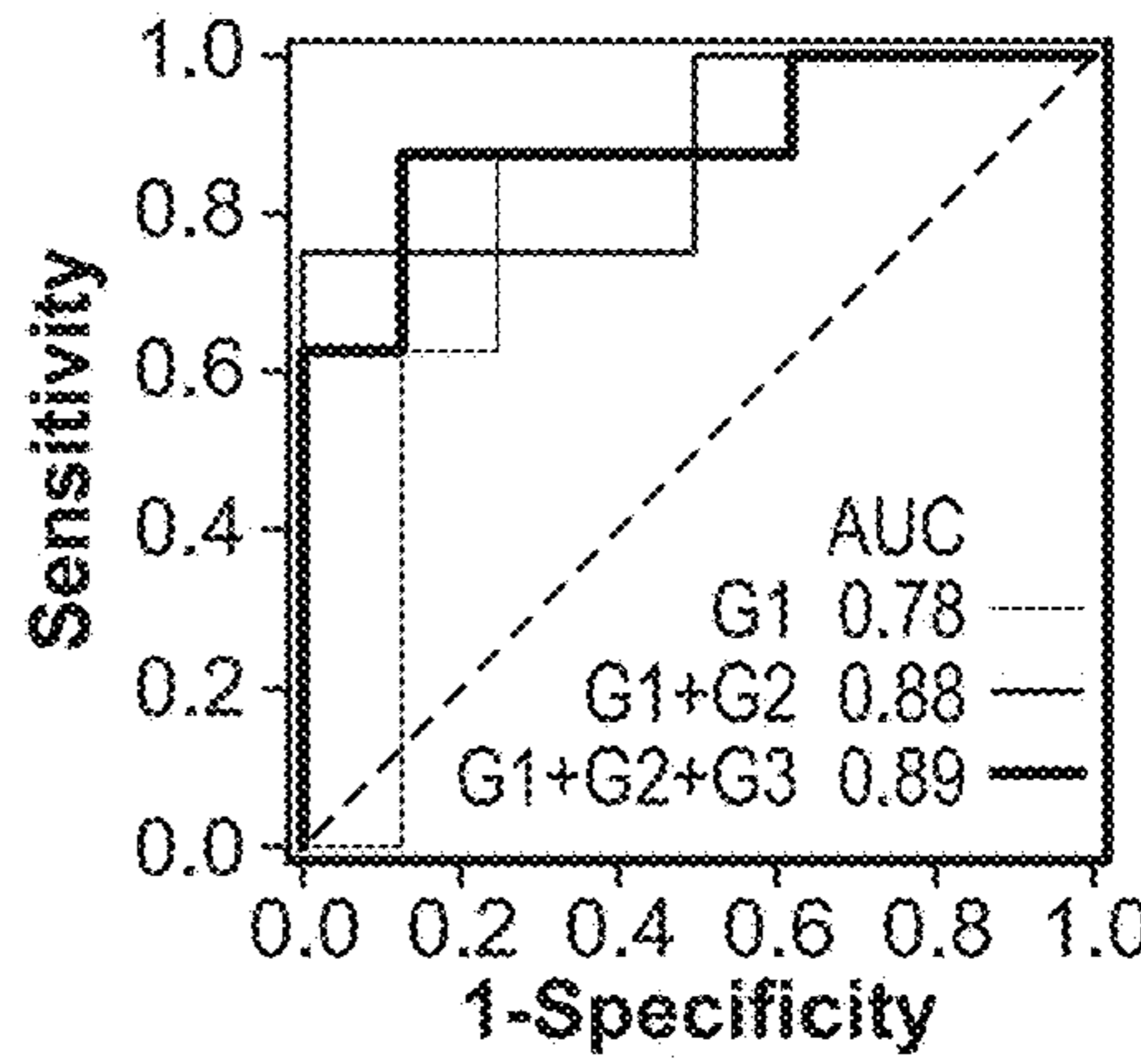


FIG. 23c

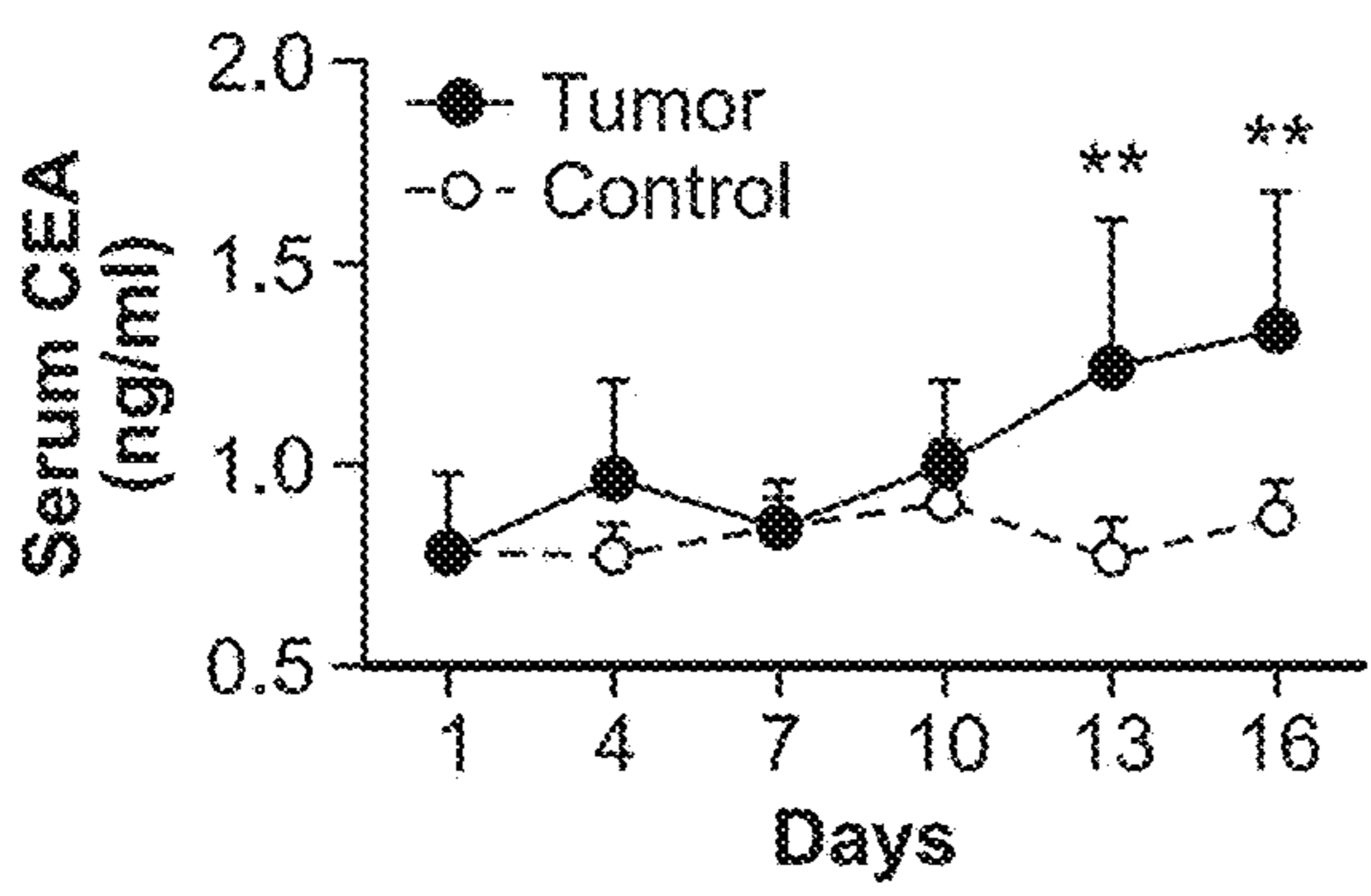


FIG. 23g

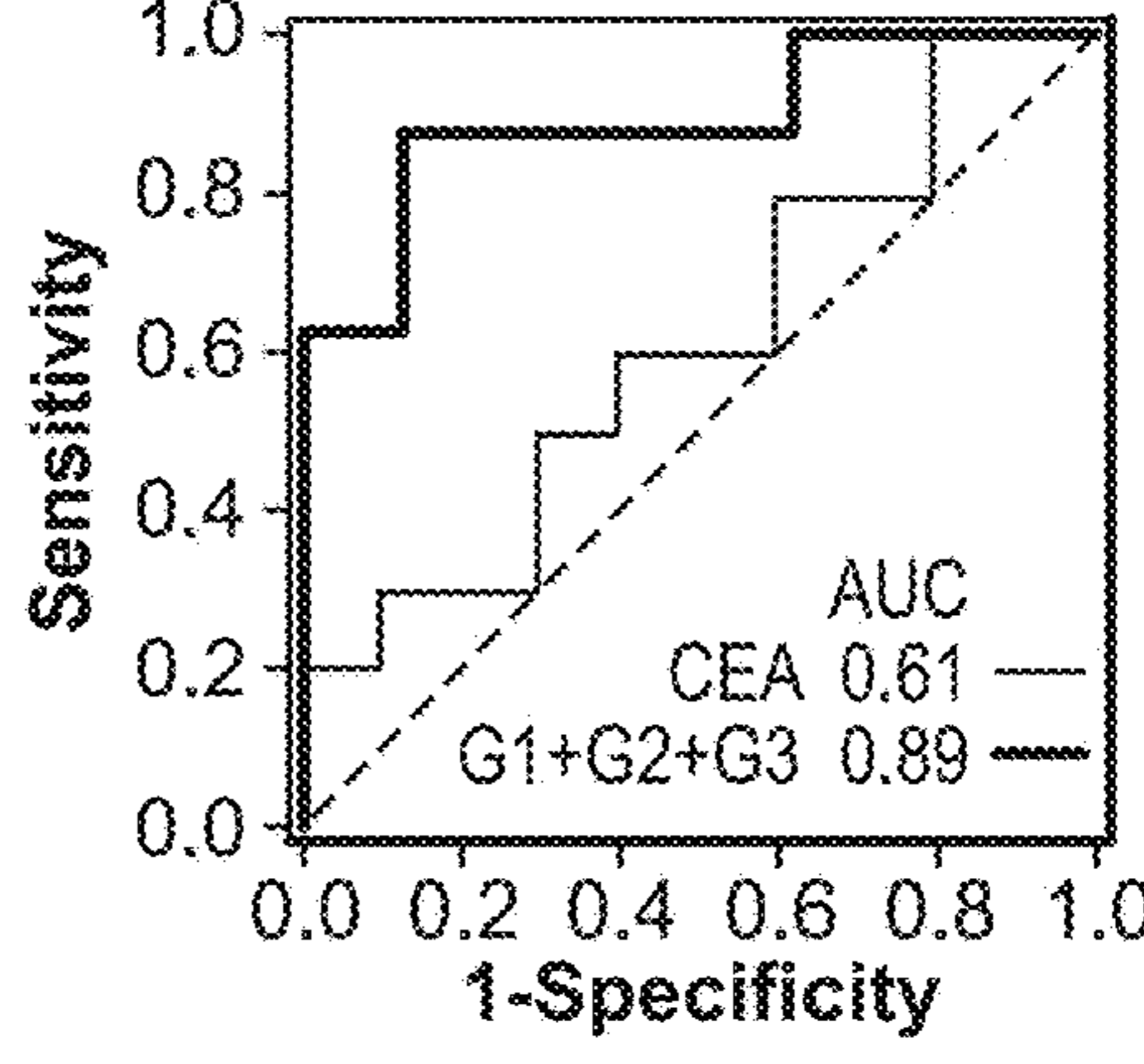


FIG. 24a

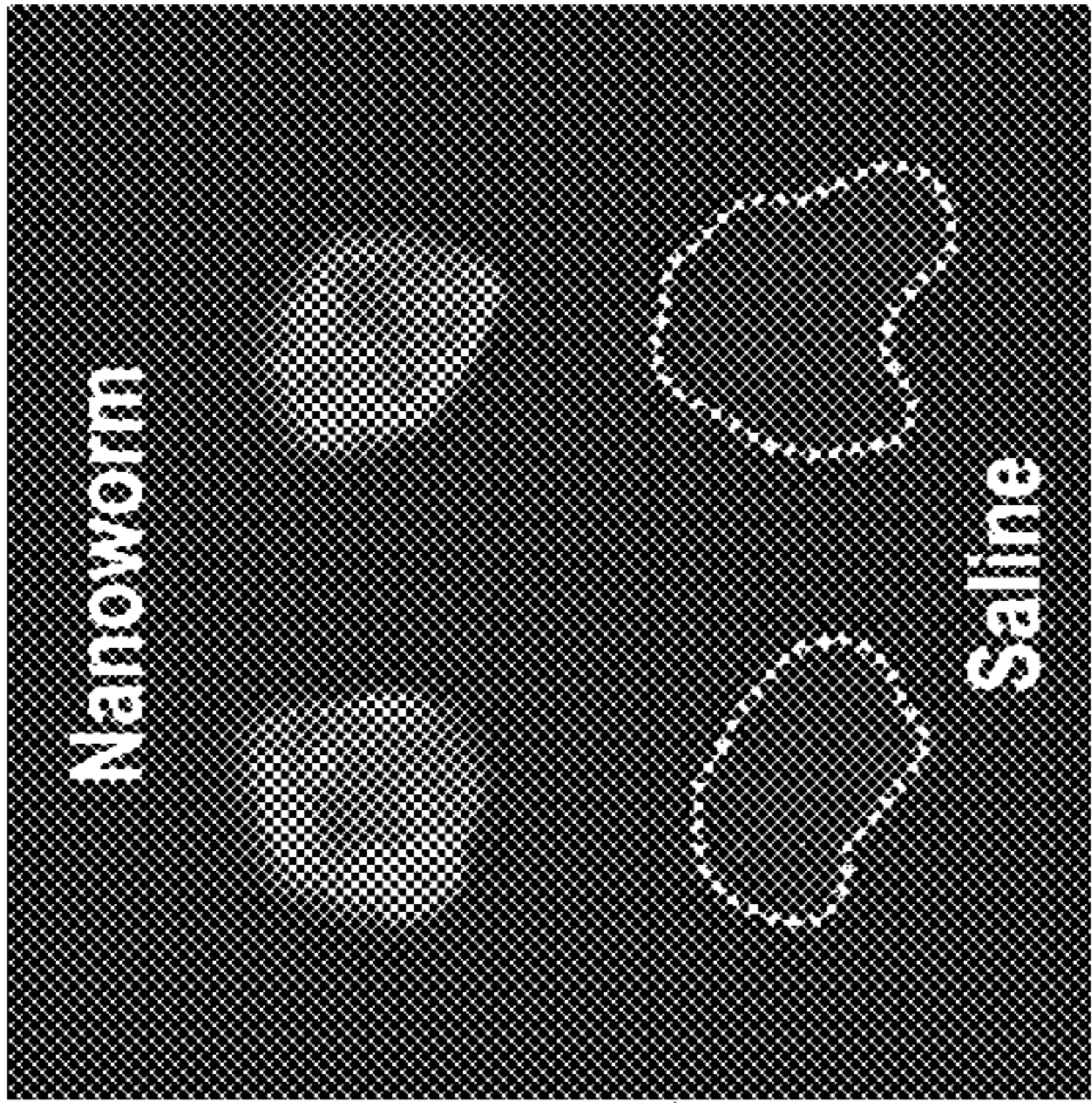


FIG. 24b

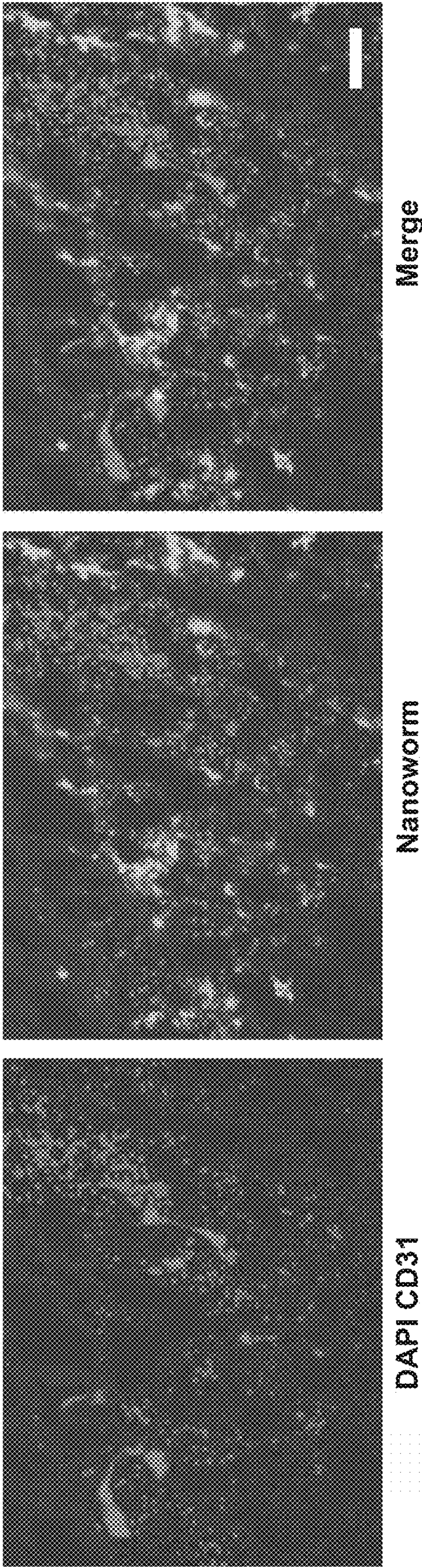


FIG. 25A

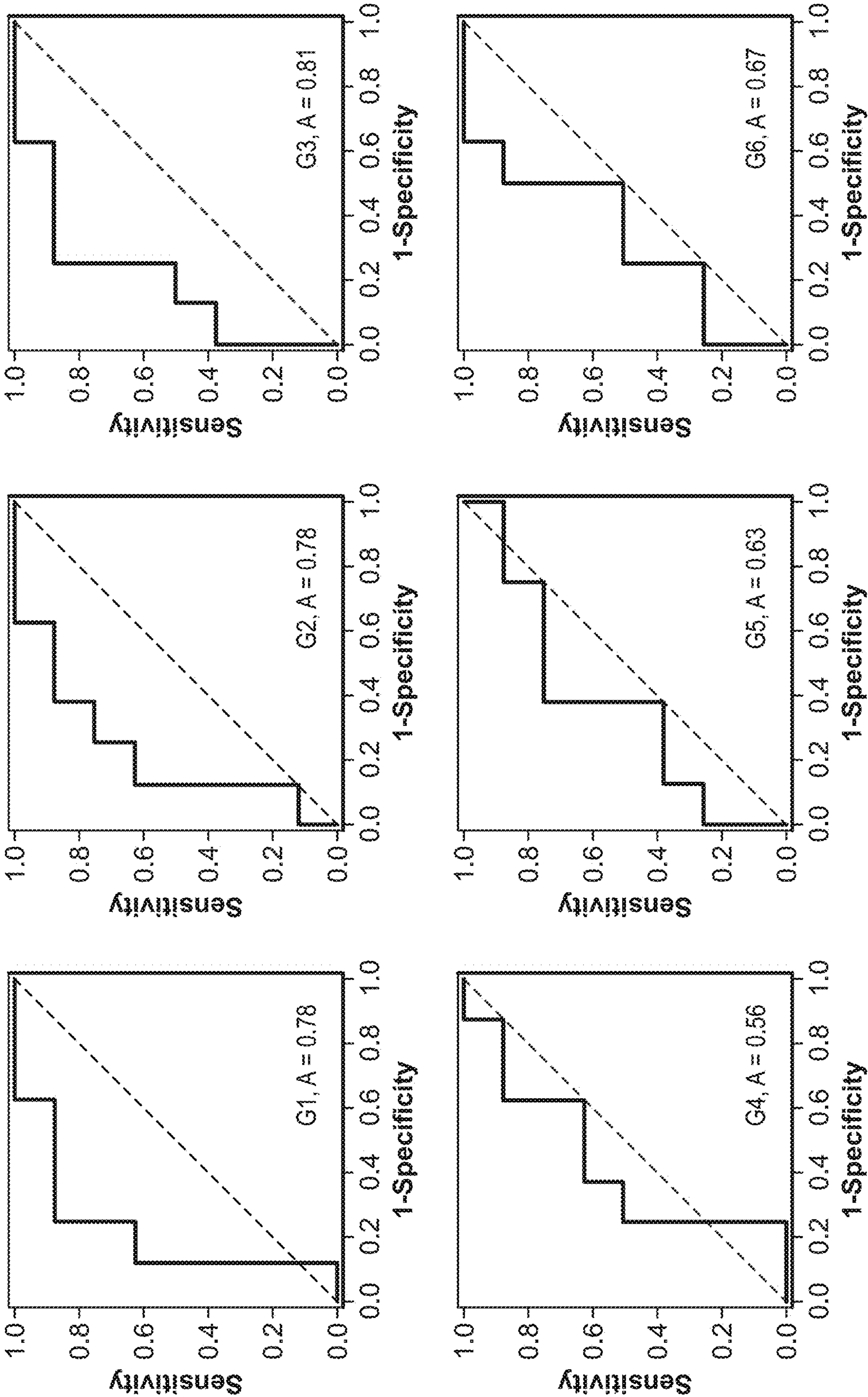
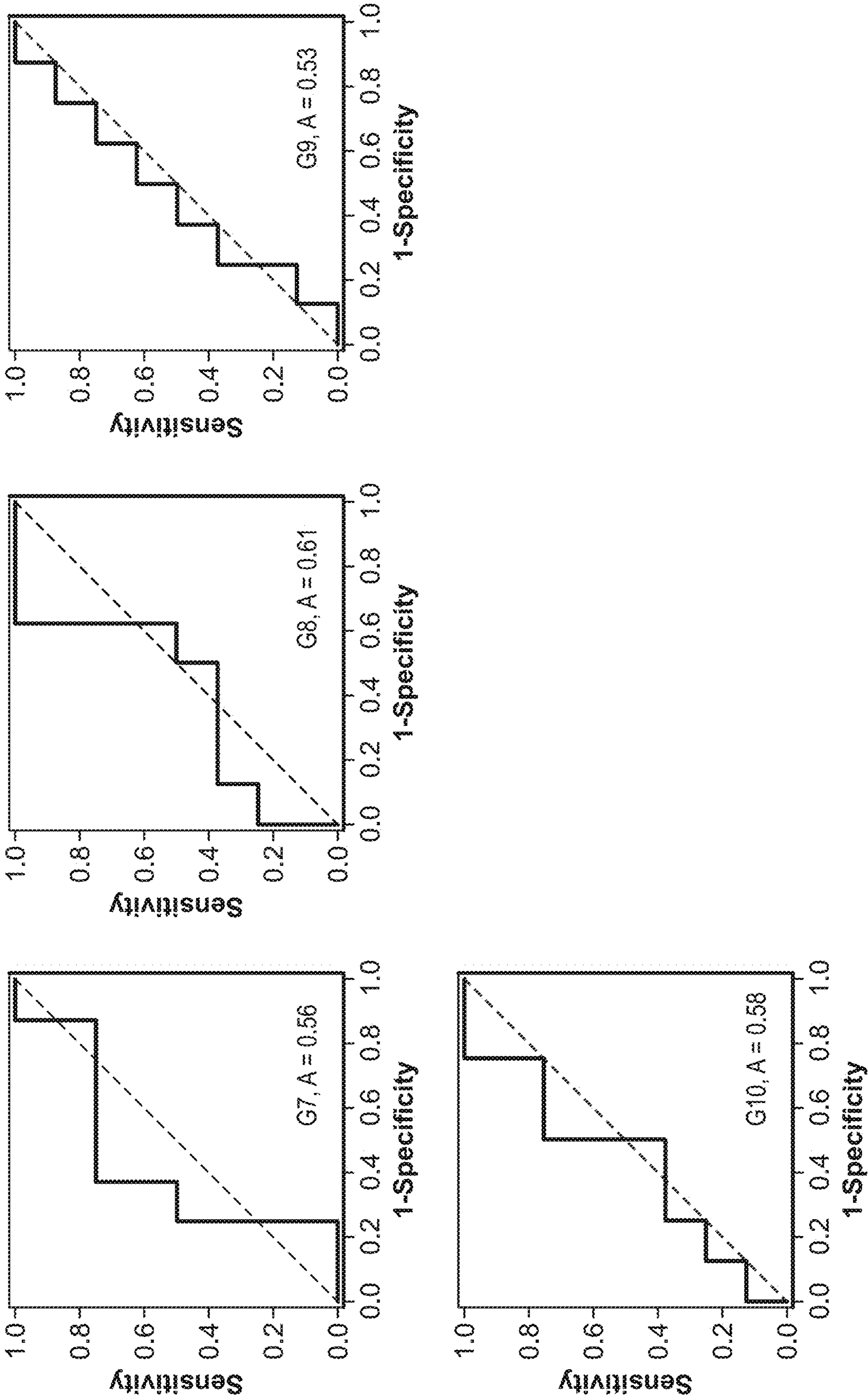


FIG. 25B



MULTIPLEXED DETECTION WITH ISOTOPE-CODED REPORTERS

RELATED APPLICATIONS

[0001] This application is a continuation of U.S. application Ser. No. 15/842,162, entitled “MULTIPLEXED DETECTION WITH ISOTOPE-CODED REPORTERS” filed on Dec. 14, 2017, which is a continuation of U.S. application Ser. No. 14/005,416, entitled “MULTIPLEXED DETECTION WITH ISOTOPE-CODED REPORTERS” filed on Jun. 12, 2014, which is a National Phase application under 35 U.S.C. § 371 of International Application No. PCT/US2012/029200, filed Mar. 15, 2012, which claims priority under 35 U.S.C. § 119(e) to U.S. Provisional Application Ser. No. 61/452,908, entitled “MULTIPLEXED DETECTION OF ANALYTES WITH ISOTOPE-CODED REPORTERS” filed on Mar. 15, 2011, which are herein incorporated by reference in their entirety.

GOVERNMENT SUPPORT

[0002] This invention was made with government support under Grant R01-CA124427 awarded by the National Institutes of Health. The U.S. government has certain rights in this invention.

REFERENCE TO AN ELECTRONIC SEQUENCE LISTING

[0003] The contents of the electronic sequence listing (M065670233US03-SEQ-FL.xml; Size: 111,002 bytes; and Date of Creation: Dec. 6, 2022) is herein incorporated by reference in its entirety.

BACKGROUND OF THE INVENTION

[0004] Mass spectrometry (MS) is an analytical technique that is useful for detecting analytes in both a quantitative and a qualitative way. During the MS process, detectable molecules, called mass tags, are ionized to generate charged molecules or molecule fragments and subsequently the mass-to-charge ratio of these molecules is measured.

[0005] MS is a highly versatile technology that can be used in various scenarios of analyte detection, for example, in biomedical diagnostics, and environmental analysis. For example, MS can be used to determine the presence or absence of an analyte, for example, a protein, nucleic acid, biomolecules, or small molecule compound, in a sample (for example, a biological sample), and/or to quantify an analyte in a sample.

[0006] While conventional MS assays are capable of detecting multiple analytes, e.g., peptides, simultaneously, quantitative analysis of multiple samples, e.g. from different individuals or time points, is currently limited to 4-plex, 6-plex, and 8-plex formats. Further, different analytes often have widely varying physiochemical properties, and, thus, MS properties, which results in similarly wide variability in sensitivity, specificity, and accuracy of MS based detection of different analytes. This variability in physiochemical properties of different analytes limits accurate simultaneous quantification of such different analytes in multiplex MS assays. The detection of multiple naturally-occurring or endogenous analytes, the chemical structures, of which often widely vary, also often necessitate monitoring a wide mass window in a single MS assay, which can be very time-intensive. Further, the detection of target analytes in com-

plex samples, such as biological samples (e.g., blood, serum, or tissue biopsies), is often difficult and challenging without extensive prior front-end processing prior to MS assays.

SUMMARY OF THE INVENTION

[0007] Some aspects of this invention address the shortcomings of conventional MS technology by providing novel reagents and methods for the simultaneous detection of a virtually unlimited number of analytes in a sample, for example, a complex biological sample (e.g., a blood or tissue sample obtained from a subject) using MS methodology.

[0008] In some embodiments, this invention provides isotope coded reporter molecules, referred to herein as iCOREs, which are useful as mass tags in multiplexed MS based analyte detection. In some embodiments, sets, or libraries of iCOREs are provided that are useful in multiplexed detection of a virtually unlimited number of analytes in a sample. In some embodiments, the iCOREs in such a set or library are isobaric and different iCOREs within such a set or library are distinguished by their unique fragmentation ion signature. In some embodiments, different fragmentation signatures are conferred to different iCOREs by differential isotope labeling. For example, in some embodiments, this invention provides a set or library of iCOREs, e.g., a set of isobaric, isotope-labeled, peptide mass tags, that are useful in the multiplexed MS based analyte detection methods provided herein. Methods for the use of iCOREs and sets or libraries of iCOREs are also provided herein.

[0009] Some aspects of this invention provide reagents and methods that are useful for the translation of a parameter, for example, the presence of a target analyte in a sample, the presence of a target activity in a sample, or the identity of a sample, cell, or tissue, into an iCORE for MS detection, allowing for biochemical encoding of the parameter. For example, in some embodiments, this invention provides reagents and methods for the simultaneous translation, or biochemical encoding, of a plurality of analytes, for example, analytes of varying a chemical structure, in a complex sample (e.g., a blood or tissue sample obtained from a subject) into a plurality of iCOREs, for example, a plurality of isobaric, isotope-labeled, peptide mass tags, so that each analyte is represented by a different iCORE having a characteristic fragmentation signature. Accordingly, each analyte can be identified by detecting the fragmentation signature associated with the specific iCORE encoding the analyte, for example, in an MS/MS assay. To give another example, in some embodiments, this invention provides reagents and methods for the simultaneous translation, or biochemical encoding, of a plurality of biological activities, for example, enzymatic activities (e.g., protease, kinase, or phosphatase activities), in a complex sample (e.g., a blood or tissue sample obtained from a subject) into a plurality of iCOREs, for example, a plurality of isobaric, isotope-labeled, peptide mass tags, so that each activity is represented by a different iCORE having a characteristic fragmentation signature. Accordingly, each activity can be identified by detecting the fragmentation signature associated with the specific iCORE encoding the activity, for example, in an MS/MS assay.

[0010] Each of the embodiments of the invention can encompass various recitations made herein. It is, therefore, anticipated that each of the recitations of the invention involving any one element or combinations of elements can,

optionally, be included in each aspect of the invention. Similarly, each aspect or embodiment of the invention can be excluded from any other aspect or embodiment, or any combination of aspects or embodiments of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

[0011] FIGS. 1A-1B. Schematic of an exemplary method for multiplexed translational analyte detection. The displayed method includes an optional enrichment step.

[0012] FIG. 2. Utilization of a canonical peak from glu-fib MS/MS spectra as reporter. The fragmentation spectrum of glu-fib and an exemplary encoding strategy is shown. EGVNDNEEGFFSAR corresponds to SEQ ID NO: 1.

[0013] FIG. 3. MS and MS/MS data obtained from a 10-plex iCORE library experiment. EGVNDNEEGFFSAR corresponds to SEQ ID NO: 1.

[0014] FIG. 4. Multiplex, quantitative detection of iCOREs in library.

[0015] FIG. 5. Evaluation of the dynamic range of multiplexed iCORE assays.

[0016] FIG. 6. Photocleavage and photocleavage efficiency. EGVNDNEEGFFSAR corresponds to SEQ ID NO: 1. KGGPVGLIGC corresponds to SEQ ID NO: 2.

[0017] FIG. 7. Exemplary chemistry for the generation of PEG hydrogel-embedded iCORE tags.

[0018] FIGS. 8A-8C. Schematic of iCORE approach for urinary diagnostics.

[0019] FIGS. 9A-9C. Long-circulating iron oxide nanoworm chaperones.

[0020] FIGS. 10A-10E. Selecting protease-sensitive substrates for NW-chaperoned urinary trafficking.

[0021] FIGS. 11A-11H. Urinary biomarkers of hepatic fibrosis and resolution in DDC-treated mice.

[0022] FIGS. 12A-12D. Immunofluorescence of liver sections.

[0023] FIGS. 13A-13F. Photo-caged iCORE libraries for multiplexed profiling of protease activities by LC MS/MS. EGVNDNEEGFFSAR corresponds to SEQ ID NO: 1. KGGPWGIWGQGGC corresponds to SEQ ID NO: 3.

[0024] FIGS. 14A-14B. Isobaric Coded REporter (iCORE) mass encoding. EGVNDNEEGFFSAR corresponds to SEQ ID NO: 1.

[0025] FIG. 15. MS/MS spectrum of 10-plex iCORE library. iCORE peak clusters centered on y-type ions. The y6 region outlined by a box is presented as FIG. 13e.

[0026] FIGS. 16A-16B. Unit collection window for peptide fragmentation minimizes peak overlap arising from naturally occurring isotopes. (a) A typical MS spectrum of an isotope-coded Glu-fib peptide. The parent precursor peptide was collected for fragmentation via a unit mass window (gray), excluding the naturally occurring isotope peaks. (b) Resulting MS/MS spectrum. Isotope peak was minimized, comprising ~5% of the original peak intensity. EGVNDNEEGFFSAR corresponds to SEQ ID NO: 1. GFFSAR corresponds to SEQ ID NO: 4.

[0027] FIGS. 17A-17C. iCORE LC MS/MS analysis is quantitative.

[0028] FIGS. 18A-18B. Protease-specific iCORE mass signatures. (a) iCORE MS/MS profiles of recombinant proteases MMP2, MMP12, and thrombin. (b) Graphical representation of Pearson's correlation coefficients between proteases.

[0029] FIG. 19. iCORE profiles of control animals. Box-and-whisker plots of individual iCORE peak intensities (repeated measures ANOVA, n=5)

[0030] FIGS. 20A-20B. ROC curves of fibrosing biomarkers.

[0031] FIGS. 21A-21B. ROC curves of resolving biomarkers.

[0032] FIG. 22. In vitro CEA secretion by LS 174T human colorectal cancer cells. Quantification of CEA from media at days 1 and 2 by ELISA.

[0033] FIGS. 23A-23G. Synthetic biomarkers outperform serum CEA for early cancer detection.

[0034] FIGS. 24A-24B. NW accumulation in tumor tissue. (a) VivoTag-680-labeled NWs or saline solutions were injected into LS 174T xenograft animals. Following excision, the tumors were scanned for NW accumulation. (b) Immunofluorescence analysis of tumor sections for blood vessels and NW. Scale bar=50 μ m.

[0035] FIGS. 25A-25B. ROC curves of tumor biomarkers.

DEFINITIONS

[0036] The term activity, as used herein, refers to a biological activity. The term includes, in some embodiments, enzyme activity, for example, hydrolase, transferase, lyase, isomerase, ligase, or oxidoreductase activity. In some embodiments, the activity is a protease activity. In some embodiments, the activity is a phosphatase activity. In some embodiments, the activity is a kinase activity. Enzymatic activity can be encoded in iCOREs, for example, by providing a substrate of the target enzyme comprised in an iCORE, exposing the iCORE to a sample comprising the enzyme, and detecting the enzyme-modified substrate, for example, by methods described herein. Typically, an iCORE for encoding an enzyme activity comprises a target substrate of that enzyme (e.g., a protease recognition site, or a phosphorylation site, etc.), and the modification of the target site by the enzyme (e.g., cleavage, phosphorylation or de-phosphorylation) can be detected, for example, by methods described herein.

[0037] The term analyte, as used herein, refers to a molecule the presence or absence or the quantity of which is subject to analysis. Typically, an analyte is a molecule of interest, for example, a protein or peptide, a nucleic acid molecule, a carbohydrate, a lipid, a metabolite, a small organic molecule, a drug, or a drug derivative (e.g. a drug metabolite), a cell surface marker, or a secreted molecule, the detection or quantification of which is of interest to a researcher or clinician, for example, for research or diagnostic purposes. A target analyte is an analyte the presence, absence, or the quantity of which in a sample, for example, in a biological, experimental, or environmental sample is subject to analysis. An analyte may be a biomarker, for example, a biomarker the presence, absence, or quantity of which in a sample indicates a particular condition of the sample or the subject, experiment, or environment the sample was obtained from. In some embodiments, an analyte is a biomedical biomarker, for example, a protein, peptide, polysaccharide, small molecule, or metabolite in a sample obtained from a subject diagnosed with or suspected to have a disease or condition, wherein the presence, absence, or quantity of the biomarker in the sample is indicative of the presence, absence, or state of the disease or condition in the subject. For example, in some embodiments, a diagnostic assay provided herein comprises the

detection of a panel of protein and metabolite biomarkers the presence, absence, or quantity of which in a blood or serum sample obtained from a subject is indicative of the presence, absence, or state of a cancer, or a plurality of cancers in the subject. The protein and metabolite biomarkers investigated in such an assay would be the target analytes of that assay.

[0038] The term antibody, as used herein, refers to an immunoglobulin, whether natural or wholly or partially synthetically produced. Antibody derivatives which maintain specific binding ability, for example, antigen-binding antibody fragments, such as Fab, Fab', or F(ab')₂ fragments, or engineered antibodies, such as scFvs, are also included referred to by the term antibody. The term also refers to any protein having a binding domain which is homologous or largely homologous to an immunoglobulin binding domain. These proteins may be derived from natural sources, or partly or wholly synthetically produced. An antibody may be monoclonal or polyclonal. An antibody may be a member of any immunoglobulin class, including, but not limited to, any of the human classes: IgG, IgM, IgA, IgD, and IgE.

[0039] The term antibody fragment, as used herein, refers to any derivative of an antibody which is less than full-length. Preferably, the antibody fragment retains at least a significant portion of the full-length antibody's specific binding ability. Examples of antibody fragments include, but are not limited to, Fab, Fab', F(ab')₂, scFv, Fv, dsFv, diabody, single variable domain, and Fd fragments. The antibody fragment may be produced by any means. For instance, the antibody fragment may be enzymatically or chemically produced by fragmentation of an intact antibody or it may be recombinantly produced from a gene encoding the partial antibody sequence. Alternatively, the antibody fragment may be wholly or partially synthetically produced. The antibody fragment may optionally be a single chain antibody fragment. Alternatively, the fragment may comprise multiple chains which are linked together, for instance, by disulfide linkages. The fragment may also optionally be a multi-molecular complex. A functional antibody fragment will typically comprise at least about 50 amino acids and more typically will comprise at least about 200 amino acids. Single-chain Fvs (scFvs) are recombinant antibody fragments consisting of only the variable light chain (VL) and variable heavy chain (VH) covalently connected to one another by a polypeptide linker. Either VL or VH may be the NH₂-terminal domain. The polypeptide linker may be of variable length and composition so long as the two variable domains are bridged without serious steric interference. Typically, the linkers are comprised primarily of stretches of glycine and serine residues with some glutamic acid or lysine residues interspersed for solubility. Diabodies are dimeric scFvs. The components of diabodies typically have shorter peptide linkers than most scFvs, and they show a preference for associating as dimers. An Fv fragment is an antibody fragment which consists of one VH and one VL domain held together by non-covalent interactions. The term dsFv is used herein to refer to an Fv with an engineered intermolecular disulfide bond to stabilize the VH-VL pair. An F(ab')₂ fragment is an antibody fragment essentially equivalent to that obtained from immunoglobulins (typically IgG) by digestion with an enzyme pepsin at pH 4.0-4.5. The fragment may be recombinantly produced. A Fab' fragment is an antibody fragment essentially equivalent to that obtained by reduction of the disulfide bridge or bridges joining the two heavy chain pieces in the F(ab')₂ fragment.

The Fab' fragment may be recombinantly produced. A Fab fragment is an antibody fragment essentially equivalent to that obtained by digestion of immunoglobulins (typically IgG) with the enzyme papain. The heavy chain segment of the Fab fragment is the Fd piece.

[0040] The term binding agent, as used herein, refers to a molecule that binds to another molecule with high affinity. In some embodiments, the binding is through non-covalent interaction. In some embodiments, the binding is specific, meaning that the binding agent binds only one particular type of molecule, or a narrow class of highly similar molecules with high affinity. Non-limiting examples of binding agents are antibodies, antibody fragments, aptamers, and adnectins.

[0041] The term body fluid, as used herein, refers to any body fluid including, without limitation, serum, plasma, lymph fluid, synovial fluid, follicular fluid, seminal fluid, amniotic fluid, milk, whole blood, sweat, urine, cerebrospinal fluid, saliva, semen, sputum, tears, perspiration, mucus, tissue culture medium, tissue extracts, and cellular extracts. It may also apply to fractions and dilutions of body fluids. The source of a body fluid can be a human body, an animal body, an experimental animal, a plant, or other organism.

[0042] The term conjugated, as used herein, refers to a state of relatively stable association between two entities, for example, between an iCORE and a binding agent. In some embodiments, conjugated entities are linked by a direct or indirect covalent or non-covalent interaction. Preferably, the association is via covalent bond. In some embodiments, two peptides are conjugated via protein fusion, e.g., a peptide iCORE as provided herein, may be conjugated to a peptidic binding agent, e.g., an antibody or antibody fragment, by fusing the iCORE to the binding agent, e.g., by expression of a recombinant iCORE-binding agent fusion protein. Non-covalent interactions that result in conjugation include, but are not limited to hydrogen bonding, van der Waals interactions, hydrophobic interactions, magnetic interactions, and electrostatic interactions. Typically, two conjugated entities are associated with each other in a manner stable enough to withstand the conditions typically encountered during an iCORE experiment, for example, an iCORE conjugated to a binding agent is associated with the binding agent in a manner sufficient for the bond between the two to endure the binding and washing steps typically comprised in the methods they are used in, for example, before the two entities are intentionally separated (e.g., by cleaving a linker connecting the iCORE to the binding agent).

[0043] The term enriched, as used herein, refers to a sample or composition in which the proportion of a material of interest in a mixture of materials comprising both the material of interest and at least one additional material is increased as compared to the original proportions in the sample. Such an increase can be achieved by methods of physical separation, chemical interaction or reaction, and other methods well known to those of skill in the art, or provided herein. For example, an original sample comprising a population of different iCOREs conjugated to different binding agents may be enriched to include predominantly those iCOREs conjugated to binding agents that bind to their specific target analytes by immobilizing the analyte-bound iCOREs on a solid support and washing away all or most of the unbound iCOREs. Isolation or purification are within the scope of the term, but such steps are not required in order to enrich a compound, e.g., a desired iCORE.

[0044] The term fragmentation signature, or fragmentation ion signature, as used herein, refers to the pattern of ions that an iCORE can fragment into, for example, during an MS assay, e.g., an MS/MS assay. In some embodiments, differential isotope labeling of iCOREs of the same base sequence, e.g., of the same amino acid sequence in the case of peptide iCOREs, is used to produce a set of different isobaric iCOREs, each of which produces at least one ion, for example, a y_7 ion, that can be distinguished from other ions of the same type, e.g., y_7 ions, produced by the other iCOREs of the set, e.g., in an MS assay, e.g. an MS/MS assay. For example, the ten exemplary iCOREs G1-G10 described in FIG. 3 each exhibit a unique fragmentation signature, each producing a different mass peak representing the y_7 fragment of the iCOREs in an MS/MS assay. A unique fragmentation signature is a signature of a given polymeric molecule that results in a unique fragmentation ion, for example, a unique y_7 fragmentation ion, which can be unambiguously identified, or which can be distinguished from any other fragmentation ion, e.g., any other y_7 ion produced by a library of iCOREs. In some embodiments, different fragmentation signatures are conferred to different iCOREs by differential isotope labeling. For example, in some embodiments, an isobaric iCORE library is produced by distributing heavy isotopes across the iCORE molecule, e.g., across the amino acid residues of a peptide iCORE, in a way that each different iCORE produces a reporter fragmentation ion (e.g., a y_7 ion) of a different mass, while the whole sequence, (e.g., the peptide sequence comprising the reporter and the balance) of all iCOREs is of the same mass.

[0045] The term iCORE, as used herein, refers to an isotope-coded reporter molecule that can be used as a mass tag in an MS assay, for example, an MS/MS assay. Typically, an iCORE is an isotope-labeled polymer, for example, a polypeptide, polynucleotide, or polysaccharide that comprises at least 5 monomeric residues, e.g. amino acid, nucleotide, or monosaccharide residues. In some embodiments, an iCORE comprises more than 5 monomeric residues. In some preferred embodiments, the structure of an iCORE allows for the generation of different fragmentation signatures, for example, by differential isotope labeling of a monomeric residue, or a combination of monomeric residues. In some embodiments, the polymer structure, e.g., the amino acid, nucleotide, or monosaccharide sequence of an iCORE allows for the generation of at least about 10 iCORE molecules, for example, isobaric iCORE molecules, having different fragmentation signatures that can be distinguished in an MS assay, e.g., an MS/MS assay. In some embodiments, the polymer structure allows for the generation of more than 10 different iCOREs having different fragmentation signatures.

[0046] The term isobaric, as used herein, refers to a group of molecules having the same molecular weight. For example, in some embodiments, a set of isobaric peptide iCOREs may be a set of peptides that all have the same weight, but different fragmentation signatures.

[0047] The term parameter, as used herein in the context of iCORE encoding, refers to a characteristic, feature, or measurable factor in a sample, for example, in a biological sample. In some embodiments, the term includes the presence, absence, or quantity of an analyte. In some embodiments, the term includes the presence, absence or quantity of a biological activity, for example, an enzymatic activity or a binding activity.

[0048] The term plurality, as used herein, refers to two or more of the elements so qualified.

[0049] The term polynucleotide, which is used herein interchangeably with the terms oligonucleotide and nucleic acid molecule herein, refers to a polymer of nucleotides. The polymer may include natural nucleosides (i.e., adenosine, thymidine, guanosine, cytidine, uridine, deoxyadenosine, deoxythymidine, deoxyguanosine, and deoxycytidine), nucleoside analogs (e.g., 2-aminoadenosine, 2-thiothymidine, inosine, pyrrolo-pyrimidine, 3-methyl adenosine, 5-methylcytidine, C5-bromouridine, C5-fluorouridine, C5-iodouridine, C5-propynyl-uridine, C5-propynyl-cytidine, C5-methylcytidine, 7-deazaadenosine, 7-deazaguanosine, 8-oxoadenosine, 8-oxoguanosine, 0(6)-methylguanine, 4-acetylcytidine, 5-(carboxyhydroxymethyl)uridine, dihydrouridine, methylpseudouridine, 1-methyl adenosine, 1-methyl guanosine, N6-methyl adenosine, and 2-thiocytidine), chemically modified bases, biologically modified bases (e.g., methylated bases), intercalated bases, modified sugars (e.g., 2'-fluororibose, ribose, 2'-deoxyribose, 2'-O-methylcytidine, arabinose, and hexose), or modified phosphate groups (e.g., phosphorothioates and 5'-N-phosphoramidite linkages).

[0050] The term polypeptide is used herein interchangeably with the terms peptide, oligopeptide, and protein, and refers to a polymer of amino acid residues linked together by peptide bonds. The term, as used herein, refers to proteins, polypeptides, and peptides of any size, structure, or function. Typically, a polypeptide is at least three amino acids long. In some embodiments, polypeptides, for example, peptide iCOREs, comprise naturally-occurring amino acids, although non-naturally-occurring amino acids (e.g., compounds that do not occur in nature but that can be incorporated into a polypeptide chain and/or amino acid analogs as are known in the art may alternatively be employed. Also, one or more of the amino acids in an inventive polypeptide may be modified, for example, by the addition of a chemical entity such as a carbohydrate group, a hydroxyl group, a phosphate group, a farnesyl group, an isofarnesyl group, a fatty acid group, a linker for conjugation, functionalization, or other modification, etc. A peptide may be isotope-labeled. A peptide may comprise D-amino acids, L-amino acids, or a mixture of D-amino acids and L-amino acids. In some embodiments, a D-amino acid or a plurality of D-amino acids in a polypeptide confers increased protease resistance to the respective polypeptide as compared to a polypeptide of the same sequence but consisting of L-amino acids. A polypeptide may also be a single molecule or may be a multi-molecular complex. A polypeptide may be a fragment of a naturally occurring protein or peptide. A polypeptide may be naturally occurring, recombinant, or synthetic, or any combination of these.

[0051] The term polysaccharide, as used herein, refers to a polymer of sugars. Typically, a polysaccharide comprises at least two sugars. The polymer may include natural sugars (e.g., glucose, fructose, galactose, mannose, arabinose, ribose, and xylose) and/or modified sugars (e.g., 2'-fluororibose, 2'-deoxyribose, and hexose).

[0052] The term sample, as used herein, refers to a composition of matter representative of a biological, clinical, or experimental environment. For example, a biological sample may be a sample obtained from a subject, such as a body fluid sample, or a cell or tissue sample, or to a sample obtained from an experimental environment, such as a

composition comprising a small molecule compound, a cell culture supernatant, a composition comprising an engineered organ, and so forth. A complex sample is a sample comprising a plurality of different analytes and/or of non-analyte molecules in addition to an analyte. Non-limiting examples of complex samples are a serum sample, a blood sample, a urine sample, and a tissue sample. In some embodiments, a complex sample is a sample comprising such a large number of molecules (e.g. analytes or non-analyte molecules) that the detection of a single target analyte, for example, a peptide or metabolite, cannot readily be achieved by an MS/MS assay, for example, because of interference of other molecules with similar MS signatures. In some embodiments, a complex sample may require extensive front-end processing, for example, target analyte enrichment, to allow detection of a target analyte in a conventional MS assay.

[0053] The term small molecule, which is used herein interchangeably with the terms small molecule compound, and drug, refers to a compound either synthesized in the laboratory or found in nature, which is typically characterized in that it contains several carbon-carbon bonds, and has a molecular weight of less than 1500, although this characterization is not intended to be limiting for the purposes of the present invention. Examples of small molecules that occur in nature include, but are not limited to, taxol, dyne-micin, and rapamycin. Examples of small molecules that are synthesized in the laboratory include, but are not limited to, compounds described in Tan et al., (“Stereoselective Synthesis of over Two Million Compounds Having Structural Features Both Reminiscent of Natural Products and Compatible with Miniaturized Cell-Based Assays” J. Am. Chem. Soc. 1998, 120, 8565) and U.S. Pat. No. 7,109,377, entitled “Synthesis of Combinatorial Libraries of Compounds Reminiscent of Natural Products”, the entire contents of which are incorporated herein by reference. In certain other preferred embodiments, natural-product-like small molecules are utilized.

[0054] The term subject, as used herein, refers to a human, a non-human primate, a non-human mammal (e.g., a cow, a horse, a pig, a sheep, a goat, a dog, a cat, or a rodent), a vertebrate, an arthropod, a chordate, an annelid, a mollusk, a nematode, an echinoderm. In some embodiments, a subject is a laboratory animal, e.g., a mouse, rat, cat, dog, pig, cow, hamster, gerbil, frog, fish, worm (e.g., *C. elegans*), or fly (e.g., fruit fly, *D. melanogaster*). In some embodiments, a subject is a microorganism, for example, a yeast, bacteria, or fungus. In some embodiments, for example, in some embodiments involving a clinical application of an aspect of this invention, the subject is diagnosed with or suspected to have a disease or condition.

DETAILED DESCRIPTION OF THE INVENTION

[0055] Some aspects of this invention provide methods and reagents for encoding a plurality of biological parameters, for example, the presence or absence of an analyte, or of an enzymatic activity in a sample, into mass-encoded reporters, termed iCOREs. This allows for, e.g., the multiplexed detection of analytes or enzymatic activities in complex biological samples. The exogenous reporters can be analyzed in a single assay, e.g., an MS/MS assay, thus avoiding the need for multiple assays of different types, as

often required when assessing different analytes or enzymatic activities in biological samples using endogenous reporters.

[0056] The dependence on endogenous reporters to indicate disease in biological samples is a major limitation to diagnostic approaches. Herein, the development of iCOREs, exogenous mass-encoded reporters, as ‘synthetic biomarkers’ of disease is described. In some embodiments, iCOREs described herein are used for in vivo diagnostics. In some embodiments, iCOREs are designed to perform three functions in vivo (upon administration to a subject): target sites of disease, sample dysregulated protease activities, and emit mass-encoded reporters into host urine for multiplexed detection by mass spectrometry. To demonstrate exemplary applications of this technology, it was applied to a xenobiotic model of liver fibrosis as a noninvasive alternative to biopsy-based monitoring, and sensitive and specific synthetic biomarkers were identified that report on both actively fibrosing and resolving stages of liver disease. Different iCORE panels were also identified that markedly lowered the threshold for early cancer detection when compared with blood biomarkers in a mouse model of colorectal cancer. The ability to rapidly design, screen and identify iCOREs as synthetic biomarkers for precise, multiplexed monitoring of disease apart from endogenous biomarkers is broadly amenable to distinct pathophysiological processes, with additional applications in systems biology, drug development and point-of-care diagnostics.

[0057] Biomarker discovery is motivated by the desire to identify reliable indicators of disease for risk assessment, early detection, predicting patient responses to therapies, and surveillance of recurrent disease.^{1,2} To date, a broad range of distinct biological species such as metabolites,³ peptides,⁴ proteins,^{2,5} cell-free nucleic acids,⁶ exosomes,⁷ and circulating tumor cells,⁸ have been developed into biomarkers with varying degrees of performance. However, the reliance on native components to indicate disease is limited by fundamental technical and biological challenges because biomarkers are frequently found in low levels in circulation,^{8,9} are difficult to resolve in complex biological fluids,¹⁰ and can be rapidly degraded both in vivo and ex vivo.^{4,11}

[0058] An alternative to endogenous biomarkers is the systemic administration of exogenous reporter agents to interrogate biological states as described herein. These approaches offer the potential to tailor agents to exploit host physiology or interface with disease-specific molecular processes as alternative indicators of disease. Examples include the polysaccharide inulin to assess glomerulus filtration rates, FDG-PET to unveil regions of increased glucose metabolism, and a suite of molecular and activity-based probes for imaging biological activities in vivo.¹²⁻¹⁴ Because these agents can be designed and tested in vitro and in preclinical models, they can be iteratively optimized and can be administered at concentrations significantly above biological background. The limitations with these approaches include the inability to monitor large family of probes simultaneously due to limited multiplexing capabilities and substantial infrastructure for in vivo analysis requiring patients to be on-site (e.g. PET, MRI) precluding remote data or sample collection.

[0059] Some aspects of this disclosure provide a framework for engineering nanoscale, isobaric mass-encoded reporter agents that passively accumulate in diseased tissues

from host circulation via organ- or disease-specific vascular fenestrations (e.g. liver sinusoid endothelium or angiogenic tumor vessels respectively).^{15, 16} In an exemplary embodiment, iCOREs are provided that are designed to interrogate protease activity in tumor microenvironments. Upon arrival in the diseased microenvironment, the iCORE agents interface with aberrantly active proteases to direct the cleavage and release of surface-conjugated, mass-encoded peptide substrates into host urine for detection by mass spectrometry (MS) as synthetic biomarkers of disease. Because dysregulated protease activities are implicated in a wide range of human diseases including cancer, fibrosis, atherosclerosis, inflammation, Alzheimer's and many others,¹⁷ highly multiplexed monitoring of aberrant protease activities has the potential to distinguish diverse disease states through combinatorial analysis. While the methods and reagents provided herein are widely applicable to a wide variety of diseases, described herein is the exemplary application of this technology to address two unmet clinical challenges: the need for a noninvasive alternative to biopsy-based monitoring for liver fibrosis,¹⁸ and the inability of current clinically-utilized blood biomarkers to reliably detect early stage cancers.¹⁹

[0060] Other exemplary parameters that can be encoded and measured with iCORE technology are analytes. Current tools for analyte detection include gel electrophoresis, western blot, ELISA, PCR, immunofluorescence, microarray, and MS-based platforms like MALDI and Liquid Chromatography-MS technologies, such as LC MS/MS. LC MS is an analytical chemistry technique that combines the physical separation capabilities of liquid chromatography (e.g., High Performance LC, HPLC) with the mass analysis capabilities of mass spectrometry. LC-MS has very high sensitivity and selectivity and is, thus broadly applicable to a variety of analytes, for example, to the specific detection and/or identification of analytes in the presence of other chemicals (e.g., in a complex biological, experimental, or environmental). MS/MS (or tandem MS) involves two or more steps of MS, with some form of fragmentation occurring in between the stages. MS/MS is a technique commonly used to identify sequence information, for example, sequence information of individual peptides. Modern mass spectrometers, in-line with extensive chromatographic fractionation, can detect and assess hundreds of analytes from a single sample. MS, however, is not without limitations. Current challenges associated with mass spectrometric analysis can be simplified into two major contributing factors.

[0061] First, many samples, for example, many biological samples (e.g., blood, serum, tissue (e.g., tumor) biopsy, cell lysate, urine, cerebral spinal fluid), experimental samples (e.g., samples of interest in combinatorial drug screens), and environmental samples (e.g., soil or water samples) containing or suspected to contain an analyte of interest are highly complex, comprising different types of analytes, e.g., biomolecules (e.g., proteins, nucleic acids, lipids, carbohydrates, metabolites), small molecules, drugs and drug metabolites, inorganic and organic matter, cells, or cell debris, which are present in concentrations spanning many orders of magnitude (e.g. pg/ml to mg/ml in blood). The task of finding and detecting target analytes within this such complex samples is difficult because of high background signal and more importantly, the suppression of target analyte ionization, an integral process for mass spectrometric analysis, by bystander molecules (e.g. lipids).

[0062] Second, many analytes themselves are difficult to detect reliably and robustly via MS base methods, or, for some analytes, any conventional method for that matter, because of their suboptimal physiochemical properties (e.g. mass, charge, hydrophilicity). Variability in physiochemical properties amongst analytes leads to poor experimental repeatability and high variance between assays. This is particularly relevant for protein quantification. Often, protein-comprising samples are digested with a protease, e.g., trypsin, prior to MS analysis, and typically, the majority of peptides produced from trypsin digests do not contain sequences that are optimal for MS analysis. Such peptides are not detected or only poorly detected in a typical MS experiment, resulting in gross under-sampling and decreased overall sensitivity of current MS assays.

[0063] To address these challenges, technologies and methodologies have focused on affinity enrichment of the target analyte prior to MS to increase signal intensity, selective depletion of high abundance biomolecules to reduce background, and targeted chemistries to isolate sub-proteomes (see, e.g., Anderson, N. L., Anderson, N. G., Haines, L. R., Hardie, D. B., Olafson, R. W., Pearson, T. W., 2004. *Mass spectrometric quantitation of peptides and proteins using Stable Isotope Standards and Capture by Anti-Peptide Antibodies* (SISCAPA). *J. Proteome Res.* 3, 235.; Whiteaker, J. R., Zhao, L., Zhang, H. Y., Feng, L. C., Piening, B. D., Anderson, L., Paulovich, A. G., 2007b. *Antibody-based enrichment of peptides on magnetic beads for mass-spectrometry-based quantification of serum biomarkers*. *Anal. Biochem.* 362, 44; Kuhn, E., Addona, T., Keshishian, H., Burgess, M., Mani, D. R., Lee, R. T., Sabatine, M. S., Gerszten, R. E., Carr, S. A., 2009. *Developing multiplexed assays for Troponin I and Interleukin-33 in plasma by peptide immunoaffinity enrichment and targeted mass spectrometry*. *Clin. Chem.* 55, 1108; Wollscheid B, Bausch-Fluck D, Henderson C, O'Brien R, Bibel M, Schiess R, Aebersold R, Watts J D., *Mass-spectrometric identification and relative quantification of N-linked cell surface glycoproteins*. *Nat Biotechnol.* 2009 April; 27(4): 378-86.; Hui Zhang, Xiao-jun Li, Daniel B Martin, Ruedi Aebersold, *Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry*. *Nature Biotechnology* 21, 660-666 (2003); the entire contents of all of which are incorporated herein by reference). Some efforts have also been directed at developing computational algorithms to predict high-responding peptides from a given protein, and to reduce the resources required to develop robust proteomic assays (see, e.g., Mallick, P., Schirle, M., Chen, S. S., Flory, M. R., Lee, H., Martin, D., Ranish, J., Raught, B., Schmitt, R., Werner, T., Kuster, B., Aebersold, R., *Computational prediction of proteotypic peptides for quantitative proteomics*. *Nat. Biotechnol.* 25: 125-131, 2007; Fusaro, V. A., Mani, D. R., Mesirov, J. P., Carr, S. A., *Prediction of high-responding peptides for targeted protein assays by mass spectrometry*. *Nat. Biotechnol.* 27: 190-198, 2009).

[0064] In general, conventional approaches to address the shortcomings of current MS technology focus on developing novel chemistries and technologies to improve quantitative mass analysis, accepting the underlying condition that certain biological compounds are more difficult to ionize and detect by virtue of their chemical structure.

[0065] Some aspects of this invention, in contrast, provide molecules, compositions, and methods for the translation of

target analytes into biochemical moieties with high ionization efficiency and, thus, optimized for MS-based detection and quantification. In some embodiments, this translation of analyte identity and quantity into biochemical structures optimized for MS allows for the simultaneous assessment of multiple analytes, for example, analytes of different physiochemical properties, and, thus, of different detectability in MS based assays. In some embodiments, multiple analytes of different physiochemical properties are translated into a set of isobaric mass reporters, for example, isobaric peptide mass tags that are optimized for MS based detection, for example, in MS/MS assays. In some embodiments, the translation allows for the qualitative and/or quantitative analysis of a plurality of target analytes by LC MS/MS. In some embodiments, biochemical encoding of analytes of different physiochemical properties into isobaric mass tags as provided herein allows to focus the MS analysis on the specific mass window of the isobaric mass tags, which increases sensitivity and accuracy of the analysis and/or reduces the time required to monitor the relevant mass window. In some embodiments, the translation comprises a step of enriching target analytes, for example, target analytes in a complex biological sample.

[0066] Some aspects of this invention provide isotope-encoded reporters (iCOREs), which can be used as mass tags. In some embodiments, iCOREs are provided that are attached to binding agents, e.g., antibodies or antibody fragments, for example, via photo-labile linkers. In some embodiments, iCOREs are used for the qualitative or quantitative detection of a target analyte by an MS assay, for example, by LC MS/MS. In some embodiments, iCOREs are used for the qualitative or quantitative detection of an activity, for example, of a target enzyme activity (e.g., a protease, kinase, or phosphatase activity) by an MS assay, for example, by LC MS/MS.

[0067] In some embodiments, complex samples, for example, biological or clinical samples, containing target analytes are first selectively enriched by capture antibodies coated onto magnetic microspheres (see, e.g., FIG. 1). The immobilized analytes are then contacted with iCORE-labeled binding agents. Following binding and removal of unbound binding reagents, individual iCOREs are cleaved from the binding agents through UV irradiation and the pool of these “rescued” iCOREs is analyzed by an MS assay, for example, by LC MS/MS. In some embodiments, the presence and/or abundance of individual iCOREs, as determined by the presence and/or the signal strength obtained from each iCORE, is used to determine the presence or absence and/or the abundance (e.g., the concentration) of a target analyte or activity.

[0068] In some embodiments, one or more iCOREs are administered to a subject. For example, in some embodiments, iCOREs may be formulated, administered to a subject, and collected according to the methods described herein, or those described in PCT Application PCT/US2010/000633, filed on Mar. 2, 2010, and entitled *Methods And Products For In Vivo Enzyme Profiling*, the entire contents of which are incorporated herein by reference. For example, iCOREs designed to interrogate analytes or enzyme activity in a subject suspected of having a disease may be administered to the subject, and analyte- or enzyme-modified iCOREs may be collected at a time sufficient for the iCOREs to be exposed to the analyte or enzyme activity in the subject. In some embodiments, a sample (e.g., a urine,

blood, serum, or plasma sample) is collected from the subject and “rescued” iCOREs are detected within the urine sample, e.g., as described herein or in PCT Application PCT/US2010/000633. In some embodiments, the presence and/or abundance of individual iCOREs, as determined by the presence and/or the signal strength obtained from each iCORE, is used to determine the presence or absence and/or the abundance (e.g., the concentration) of a target analyte or activity.

[0069] Some aspects of this invention also provide a method for producing sets, or libraries, of mass codes from a single peptide-based reporter to enable multiplexed experiments in which multiple analytes are simultaneously detected and/or quantified in a single sample. In some embodiments, an iCORE with favorable MS properties is selected. Some non-limiting examples of suitable iCOREs are described herein and additional iCOREs useful according to some aspects of this invention will be apparent to those of skill in the art based on this disclosure. For example, in some embodiments, a glu-fib peptide (EGVNDNEEGFF-SAR, SEQ ID NO: 1) is used as the parent iCORE. Starting from this parent sequence, a set of isotopic analogs (e.g., glu-fib iCOREs with unique fragmentation signature) are designed, thus creating an iCORE library comprising a number of iCOREs of identical mass (isobaric iCOREs), but with distinct fragmentation reporter ions, e.g., y-ions of different masses. Such isobaric iCOREs are indistinguishable during MS analysis, but following peptide fragmentation, e.g., by collision induced disassociation (CID), infrared multiphoton dissociation (IRMPD), or any other suitable fragmentation method known to those of skill in the art), each member of the set, can be distinguished by its unique fragmentation reporter ion in a tandem MS (MS/MS) assay (FIG. 2). The generation of unique, distinguishable fragmentation signatures is accomplished, in some embodiments, by the strategic substitution of stable-isotope enriched amino acids in the parent peptide sequence. In some embodiments, a set of binding agents is conjugated to a set of iCOREs having a set of unique, distinguishable fragmentation signatures, for example, a set of iCOREs differentially labeled with isotopes that give rise to fragmentation ions that can be identified by MS/MS, so that each unique isotope-labeling pattern, and associated fragmentation ion signature, is associated only with binding agents specifically binding one analyte. In such embodiments, the cognate specificity of each binding agent is encoded by the isotope-labeling pattern, and the associated fragmentation signature of the associated iCORE. In some embodiments, such a set of binding agent-conjugated iCOREs is used for the multiplexed detection and/or quantification of analytes or activities in a sample, for example, a biological or clinical sample, or in vivo, according to methods provided herein.

[0070] There are several advantages associated with the reagents and methods provided herein as compared to conventional MS-based analyte detection methodology. For example, in some embodiments, target analytes are enriched by binding agents, e.g., by capture antibodies specifically binding an analyte of interest, immobilized on a solid substrate, e.g., a magnetic bead or microsphere, or a membrane or resin. In some embodiments, this enrichment allows for more sensitive and/or specific analyte detection, and/or reduces background signal, particularly in complex samples, such as samples of body fluids (e.g. blood or serum), tissues, or cells.

[0071] Further, in some embodiments, target analytes are translated into surrogate isotope-coded reporters (iCOREs) that are pre-designed for facile detection and/or quantification by MS/MS. In some embodiments, this translation, or biochemical encoding, of analytes, e.g., analytes which are challenging to detect by MS assays, into MS tags that are facile to detect, allows for the circumvention of challenges of detecting endogenous chemical structures or peptides directly. This is of particular advantage in the context of multiplexed assays, in which a plurality of analytes of different structure are assayed. Direct detection of different structures in one multiplex MS assay typically requires the screening of a large mass window and while some analytes can readily be detected via MS, many analytes are hard or impossible to detect without extensive pre-processing, or cannot be identified unambiguously when analyzing complex samples.

[0072] In some embodiments, a single analyte molecule is translated into a plurality of iCORE molecules, resulting in a higher sensitivity of analyte detection, for example, by a signal amplification of about 3-20 fold. For example, in some embodiments, a binding agent that specifically binds a single molecule or two molecules of a target analyte, e.g., an antibody or fragment thereof, is conjugated to a plurality of iCORE molecules, for example, to about 3, about 5, about 10, or about 20 iCORE molecules, providing an about 3, about 5, about 10, or about 20 fold amplification of the number of molecules available for MS as compared to direct MS detection of the target analyte, respectively. It will be appreciated by those of skill in the art that higher amplification rates can be achieved by attaching more iCORE molecules to a binding agent molecule. Methods and reagents for the attachment of iCOREs to binding agents are provided herein and additional methods will be apparent to those of skill in the art. The use of photo-labile linkers during the translation process, for example, in embodiments where iCOREs bound to binding agents, e.g. antibodies or antibody fragments, are employed, enables ultra-violet light-triggered release of iCOREs from the binding agents. The high efficiency of this photochemical process is in marked contrast with the derivation of peptides from proteins by enzymatic digestion (necessary for multiplexed protein detection in conventional LC MS/MS assays) which is limited by the biophysical properties of the enzyme employed (e.g. KD, kcat, substrate specificity) and the requirement for optimal sample conditions for enzymatic activity (e.g. pH, salt concentration). Both of these constraints lead to increased sample processing requirements and/or decreased detection efficiency.

[0073] Further, the use of pre-determined mass tags, e.g., isobaric iCOREs, with known masses greatly simplifies the collection and analysis of data since it is unnecessary to query a large mass window (e.g. 50-2000 m/z) for analytes of diverse molecular weight. Rather, with isobaric codes, narrow mass windows can be centered on the parent mass (e.g. ± 0.5 m/z) to efficiently collect signal. In some embodiments, this targeted approach, combined with analyte signal amplification during translation, increases detection sensitivity by ~ 30 -300 fold. Moreover, in some embodiments, the simplification of data collection and the resulting decrease in spectrometer operating time reduces overall costs by ~ 10 fold.

[0074] Further, the method of generating isobaric iCORE mass tag libraries, for example, isobaric iCORE peptide

libraries, is an improvement over current technologies because of the large degree of encoding that can be achieved, which translates into multiplexing capabilities far beyond the current limitations. The current state-of-the-art in isobaric MS multiplexing is iTRAQ mass tag technology (Ross P L, Huang Y N, Marchese J N, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Bartlett-Jones M, He F, Jacobson A, Pappin D J., *Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents*, Mol Cell Proteomics. 2004 December; 3(12):1154-69). ITRAQ labeling reagents are amine-reactive small molecules providing for 4 or maximally 8 unique mass codes. A larger library is precluded by the number of atoms available in small molecules for isotopic substitution. With iCOREs, e.g., peptide iCOREs, as provided by some aspects of this invention, however, the number of atoms available for isotopic substitution is increased by the number of monomers, e.g., amino acids, that can be labeled with heavy isotopes. Accordingly, it is feasible to construct, for example, at least ~ 30 -40 unique codes from an average peptide length of ~ 15 amino acids. The library size, and, thus, the number of unique encoding tags can be increased further by using longer peptides and/or by combining distinct sets of isobaric iCORE peptides, providing for ~ 100 -1000, or more, elements.

[0075] Some aspects of this invention provide isotope-coded reporter molecules (iCORE) and methods of their use. In some embodiments, iCOREs are useful as mass tags in MS assays, for example, in multiplex LC MS/MS assays, as described in more detail elsewhere herein. In some embodiments, iCOREs are useful in the simultaneous qualitative and/or quantitative detection of tens, hundreds, or thousands of analytes in a sample, for example, a complex biological sample, or in the simultaneous tracking of tens, hundreds, or thousands of cells, tissues, or samples using an MS readout.

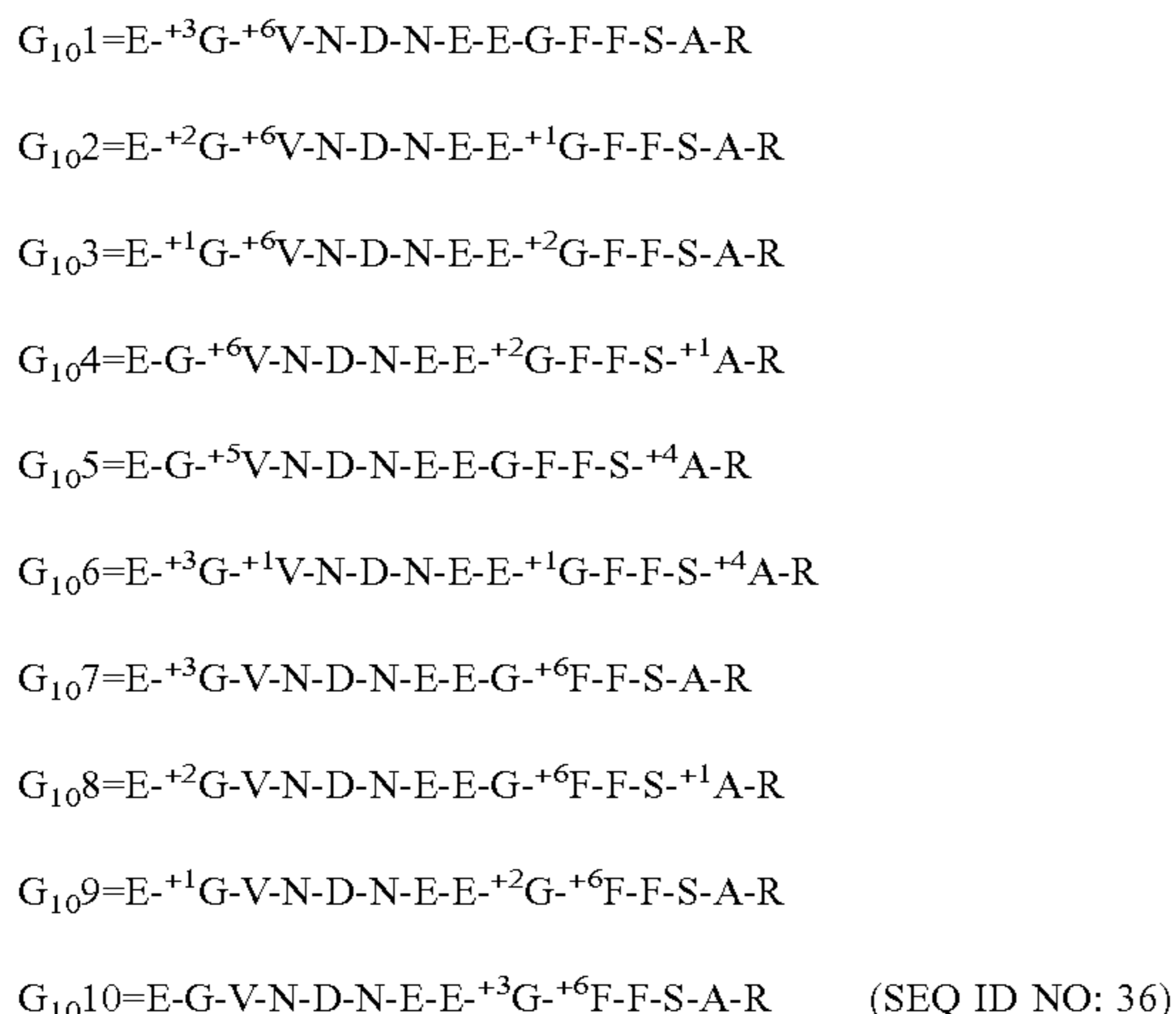
[0076] In some embodiments, a set, plurality, or library of different iCOREs is provided that is useful in multiplex MS assays, for example, in multiplex LC MS/MS assays. In some embodiments, the iCOREs of a set, plurality, or library of iCOREs are isobaric. For example, in some embodiments, a set, plurality, or library of isobaric iCOREs is provided in which all iCOREs are polymers, e.g., peptides of the same amino acid sequence (e.g., glu-fib peptides (EGVND-NEEGFFSAR, SEQ ID NO: 1)), and have, accordingly, the same physiochemical properties as relevant for MS analysis. Accordingly, all iCOREs of such a set, plurality, or library are detected with substantially the same specificity, sensitivity, and accuracy in an LC MS/MS assay. In some embodiments, different iCOREs in such a set, plurality, or library of isobaric iCOREs have different fragmentation signatures, which can readily be distinguished in an MS/MS assay as described in more detail elsewhere herein. Because even short polymers, e.g., short polypeptides, allow for the generation of a large number of unique fragmentation signatures in a set, plurality or library of isobaric iCOREs that can be distinguished via MS/MS assays, for example, by differential isotope-labeling of the iCOREs, this technology can be used for multiplex qualitative or quantitative MS analysis far beyond the limitations of current multiplex MS technologies, such as iTRAQ (4- or 8-plex max.) or tandem mass tag (TMT, 6-plex max.) technologies.

[0077] A non-limiting, exemplary set of 10 isobaric peptide iCOREs of the amino acid sequence EGVNDNEEGFF-

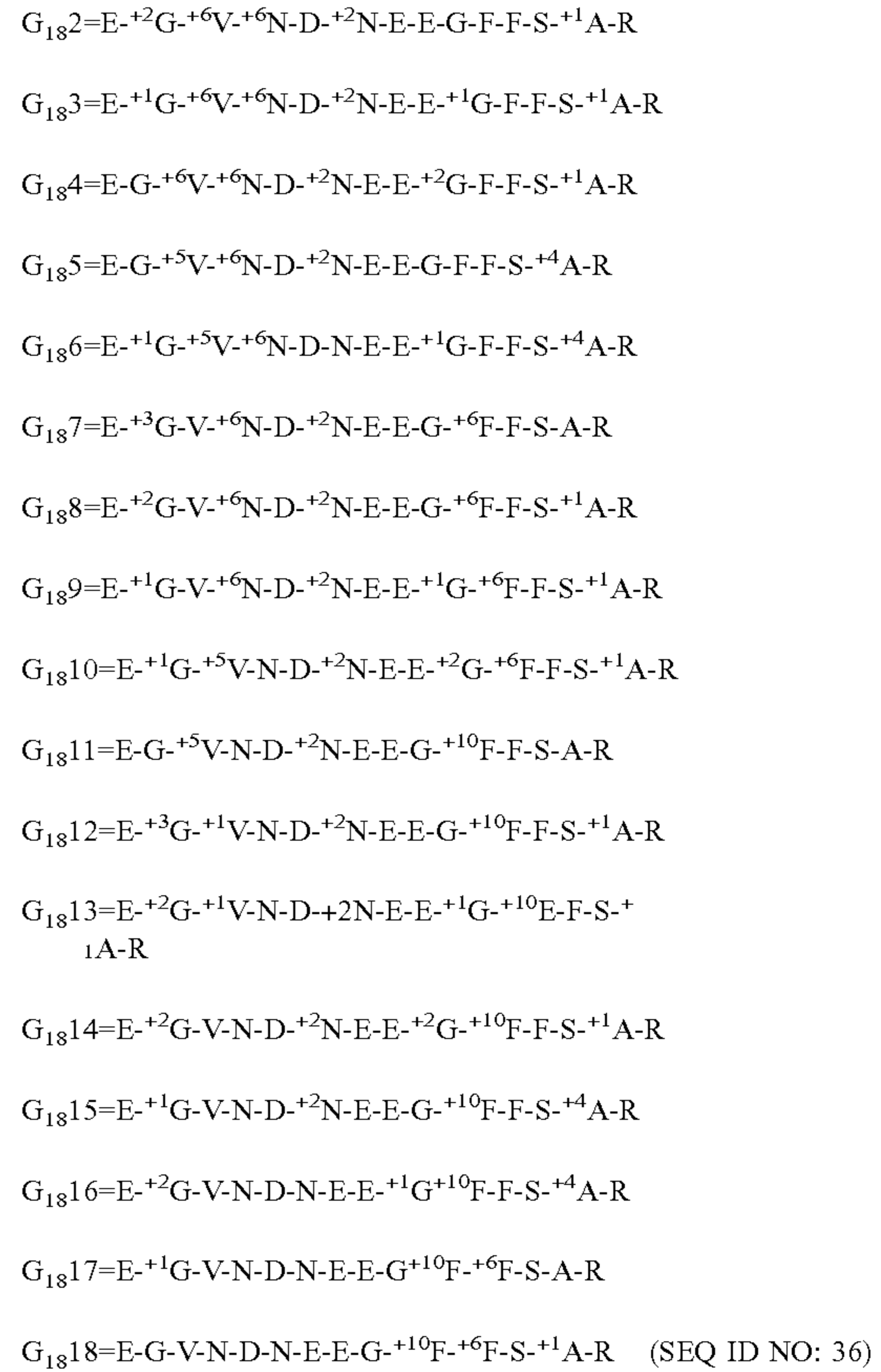
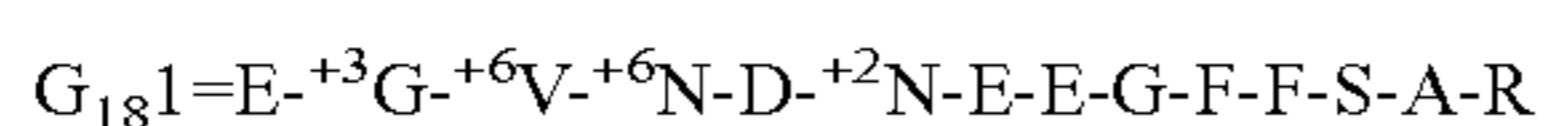
SAR (SEQ ID NO: 1) is described in FIG. 3. The different iCOREs G1-G10 have different isotope patterns, and, thus, different fragmentation signatures, e.g., G1 has the isotope pattern $E^{+3}G_{+6}VNDNEEGFFSAR$ (SEQ ID NO: 1), G2 has the isotope pattern $E^{+2}G^{+6}VNDNEE^{+1}GFFSAR$ (SEQ ID NO: 1), G3 has the isotope pattern $E^{+1}G^{+6}VNDNEE^{+2}GFFSAR$ (SEQ ID NO: 1), and so forth. In the first round of MS, only a single peak is observed, reflecting the isobaric nature of iCOREs G1-10 (FIG. 3 I.). After fragmentation, however, the unique fragmentation signatures of G1-10 result in resolution of the fragments as 10 distinct peaks (FIG. 3 I.). The lower panel of the figure shows a close-up of the peaks obtained from the y7 ion (EGFFSAR, SEQ ID NO: 6, see FIG. 2 for y ion nomenclature), allowing qualitative (present/absent) as well as quantitative analysis of each unique iCORE in the set. Those of skill in the art will appreciate that the glu-fib peptide sequence allows for the generation of more than 10 unique fragmentation signatures via differential isotope labeling, for example, by further encoding the remaining amino acids within the sequence, and that the exemplary set of ten iCOREs is not limiting in this respect. The multiplexing capabilities of iCORE technology are limited only by the number of unique fragmentation patterns that can be created in iCOREs. Even a small peptide sequence, such as the glu-fib peptide sequence exemplified in FIG. 3 allows for the generation of at least 30-40 unique isotope signatures, if commercially available isotope-labeled amino acids are used, and more if such amino acids are custom-synthesized.

[0078] Three non-limiting, exemplary differential isotope-labeling strategies for generating multiplex glu-fib based iCORE libraries are provided below. In these examples, iCOREs are isotope-labeled by substitution of one or more atoms in the indicated amino acids with one or more heavy isotopes, resulting in a change in mass of the labeled amino acid. The resulting iCOREs in each library are isobaric, with different iCOREs having a different fragmentation signature, e.g., with respect to the y_7 ion. The examples are intended for illustration only and do not limit this aspect of the invention. Additional useful differential isotope-labeling strategies and schemes will be apparent to those of skill in the art.

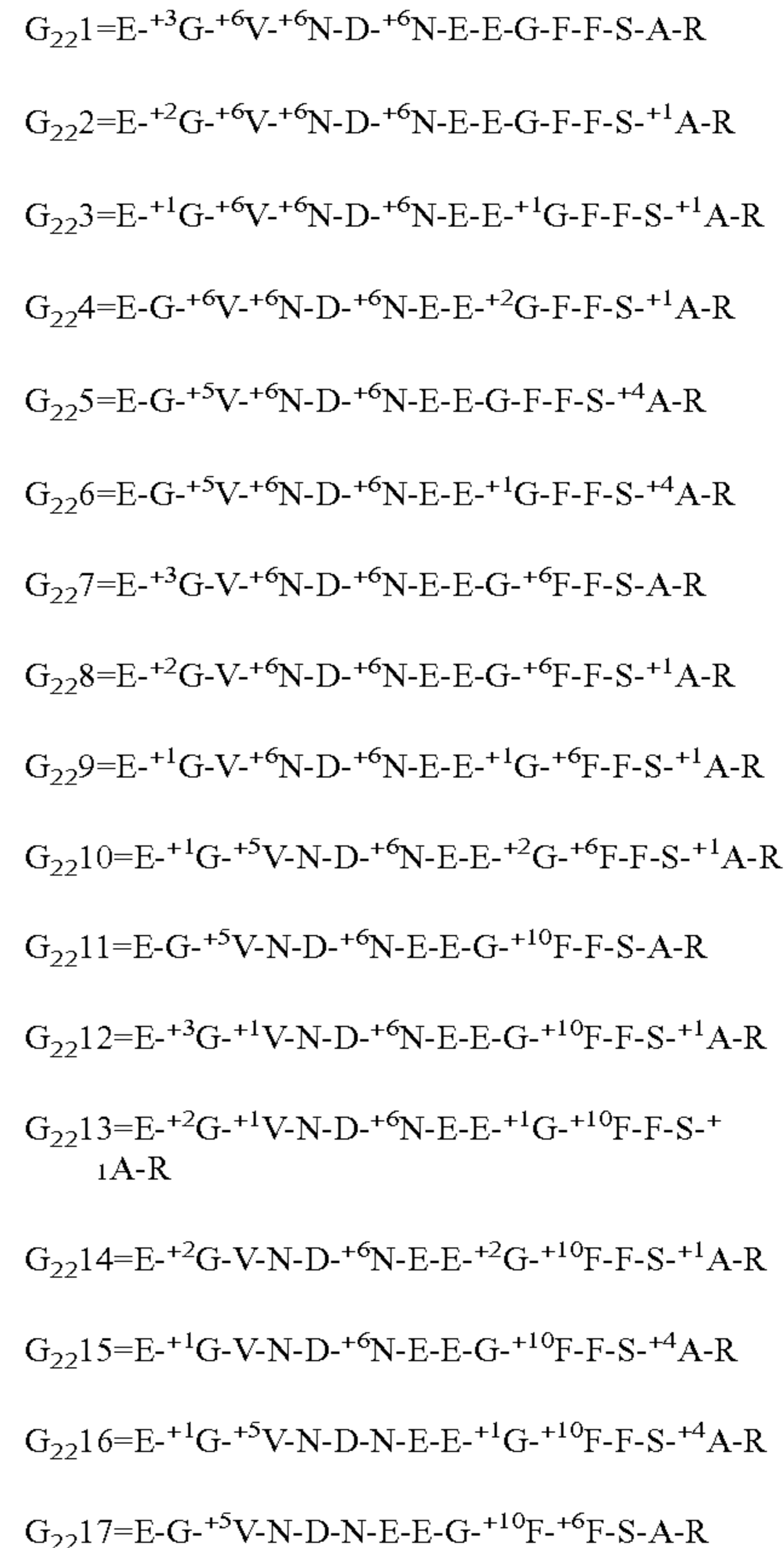
10-Plex Glu-Fib iCORE Library (G_{10}):



18-Plex Glu-Fib iCORE Library (G_{18})



22-Plex Glu-Fib iCORE Library (G_{22})



G₂₂18=E-⁺³G-⁺¹V-N-D-N-E-E-G-⁺¹⁰F-⁺⁶F-S-⁺¹A-R

G₂₂19=E-⁺²G-⁺¹V-N-D-N-E-E-⁺¹G-⁺¹⁰F-⁺⁶F-S-⁺¹A-R

G₂₂20=E-⁺²G-V-N-D-N-E-E-⁺²G-⁺¹⁰F-⁺⁶F-S-⁺¹A-R

G₂₂21=E-⁺¹G-V-N-D-N-E-E-G-⁺¹⁰F-⁺¹⁰F-S-A-R

G₂₂22=E-G-V-N-D-N-E-E-G-⁺¹⁰F-⁺¹⁰F-S-⁺¹A-R (SEQ ID NO: 36)

[0079] In some embodiments, the isotope-labeled amino acids are D-amino acids. In some embodiments, the isotope-labeled amino acids are L-amino acids. In some embodiments, the isotope-labeled amino acids are a mix of D- and L-amino acids.

[0080] Those of skill in the art will appreciate that longer polymers, for example, longer polypeptides, polysaccharides, or polynucleotides, allow for even more unique fragmentation signatures to be generated by differential isotope labeling, thus further expanding the multiplexing capabilities of iCORE technology.

[0081] In some embodiments, a set, plurality, or library of iCOREs (e.g. polypeptide, polynucleotide, or polysaccharide iCOREs) is provided that comprises at least 2, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 25, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 125, at least 150, at least 200, at least 250, at least 500, at least 1000, or more than 1000 different iCOREs, e.g. iCOREs that have a unique fragmentation signature that is distinguishable by an MS assay, e.g. an MS/MS assay. In some embodiments, the iCOREs in the set, plurality, or library are isobaric iCOREs, for example, isobaric polypeptide, polynucleotide, or polysaccharide iCOREs. In some embodiments, each unique fragmentation signature in a set of iCOREs, for example, a set of isobaric iCOREs, is associated with, or represents, a specific analyte, or parameter. Methods for the use of iCOREs, for example, in analyte detection, as well as cell, tissue, sample, and liquid tracking, are also provided.

[0082] The term iCORE refers to isotope-coded reporter molecules. In some embodiments, an iCORE is a molecule that is readily detectable by MS technology. In some embodiments, an iCORE is a peptide, for example, a peptide known to be readily detectable by MS technology. In other embodiments, an iCORE is a polynucleotide or a polysaccharide. Typically, an iCORE is a molecule that allows for differential isotope-labeling, e.g., for the generation of unique fragmentation signatures that are distinguishable in a detection assay, such as MS/MS. This allows for the generation of libraries of isobaric iCOREs which share the same physiochemical properties for the purpose of MS detection, for example, peptide iCOREs having the same amino acid sequence, and are readily distinguishable by their unique fragmentation signature, for example, in LC MS/MS assays. Typically, an iCORE is a polymer comprised of monomers that exhibit different MS signatures and, thus, comprises a sequence of monomers. Such polymers are well known to those of skill in the art and include, but are not limited to polymers of amino acids, nucleotides, and monosaccharides, for example, polypeptides, polynucleotides, and polysaccharides, respectively.

[0083] The length of iCOREs is typically chosen to produce a mass signature that is useful in MS assays, for

example, a mass signature that can readily be detected in an MS/MS assay. For example, in some embodiments, an iCORE is a polymeric molecule (e.g., a polypeptide, polynucleotide, or polysaccharide) that comprises at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 25, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, or at least 100 monomer residues, for example, amino acid, nucleotide, or monosaccharide residues. In some embodiments, an iCORE is a polymeric molecule, for example, a polypeptide, polynucleotide, or polysaccharide, that comprises 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 50, 51-60, 61-70, 71-80, 81-90, or 91-100 monomers, e.g., amino acid or nucleotide residues. In some embodiments, an iCORE is a polymer comprising more than 100 monomers, for example, more than 100 amino acid, nucleotide, or monosaccharide residues.

[0084] Non-limiting examples of polymeric molecules that are useful for the generation of iCOREs are the polypeptides glu-fib (EGVNDNEEGFFSAR (SEQ ID NO: 1)), bradykinin (PPGFSPFR (SEQ ID NO: 30)), angiotensin I (DRVYIHPFHL (SEQ ID NO: 31)), ACTH1-17 (SYSMEHFRWGKPVGKKR (SEQ ID NO: 32)), ACTH18-39 (RPVKVYPNGAEDESAEAFPLEF (SEQ ID NO: 33)), and ACTH7-38 (FRWGKPVGKKRRPVKVYPNGAEDESAEAFPLE (SEQ ID NO: 34)). Additional polymeric molecules and sequences that can be used for the generation of iCOREs will be readily evident to those of skill in the art and the invention is not limited in this respect.

[0085] In order to generate an iCORE from a non-labeled polymeric molecule, for example, a polypeptide, polynucleotide, or polysaccharide, the polymeric molecule is isotope-labeled. Methods and reagents for isotope-labeling of polymeric molecules are well known to those of skill in the art. In some embodiments, such methods include the generation of a polymeric molecule, for example, a peptide, from monomers, for example, amino acids, at least some of which are isotope-labeled to obtain an isotope-labeled polymeric molecule. In some embodiments, isotope-labeling is effected by the introduction of heavy isotopes. Isotope-labeling with radioactive isotopes or with stable isotopes is possible. In some embodiments, labeling with stable isotopes is preferred.

[0086] For example, while non-labeled monomers, e.g., non-labeled amino acids, nucleotides, or monosaccharides, predominantly comprise C¹², O¹⁶, N¹⁴, and S³² atoms, isotope-labeled monomers, e.g., isotope-labeled amino acids, may comprise one or more heavy isotope(s), for example, a C¹³, N¹⁵, O¹⁷, O¹⁸, H², S³³, S³⁴, or S³⁶ isotope. A monomer comprising a heavy isotope has a different mass than a non-labeled amino acid, and even a single isotope comprised in a monomer can readily be measured by MS assays. This difference in mass is conferred to any polymer comprising such an isotope-labeled monomer. For example, the difference in mass resulting from the substitution of a C¹² atom of an amino acid with a C¹³ isotope can be measured in a peptide comprising such an isotope-labeled amino acid.

[0087] To give a non-limiting example of monomer isotope labeling, the phenyl ring of non-labeled phenylalanine (F) typically comprises six C¹² atoms. In some embodi-

ments, an isotope-labeled phenylalanine residue comprises a phenyl ring comprising five C^{12} atoms and one C^{13} isotope, increasing the mass of the labeled amino acid by the mass of one neutron ($+1F$). In some embodiments, an isotope-labeled phenylalanine residue comprises a phenyl ring comprising four C^{12} atoms and two C^{13} isotopes, increasing the mass of the labeled amino acid by the mass of two neutrons ($+2F$). In some embodiments, an isotope-labeled phenylalanine residue comprises a phenyl ring comprising three C^{12} atoms and three C^{13} isotopes ($+3F$), two C^{12} atoms and four C^{13} isotopes ($+4F$), one C^{12} atom and five C^{13} isotopes ($+5F$), or six C^{13} isotopes ($+6F$). The remaining three C^{12} atoms of non-labeled phenylalanine and as well as the N^{14} atom of the amino group can also be isotope-substituted, yielding $+7F$, $+8F$, $+9F$, and $+10F$. It will be appreciated by those of skill in the art that any combination of isotope substitutions are possible, for example, substitution of the N^{14} of the amino group, of the C^{12} of the carboxy group and of two C^{12} of the phenyl ring would produce $+4F$, as would a substitution of three phenyl ring C^{12} with C^{13} and of the N^{14} with N^{15} , or substitution of four C^{12} of the phenyl ring with 4 C^{13} . In some embodiments, the isotope-labeled phenylalanine is used to generate an isotope-labeled peptide. In some embodiments, this involves peptide synthesis *in vivo*, for example, by cells that are incubated in the presence of isotope-labeled phenylalanine and incorporate the labeled amino acid into any peptide they synthesize. In some embodiments, isotope-labeled peptides are synthesized *in vitro*, for example, via fmoc synthesis using isotope-labeled amino acids as building blocks. It will be appreciated by those of skill in the art that amino acids other than phenylalanine can be isotope-labeled according using similar strategies and methods and reagents known to those of skill in the art and the invention is not limited in this respect. It will further be appreciated that monomers other than amino acids can be isotope-labeled, for example, nucleotides and monosaccharides. Further, other methods for the generation of isotope-labeled polymers that are useful as iCOREs, for example, methods in which a polymer is directly labeled, are also known to those of skill in the art, and the invention is not limited in this respect.

[0088] Different fragmentation signatures can be created by differential isotope labeling of polymeric molecules, for example, by adding different amounts of heavy isotopes to the same monomeric residue, or by generating different combinations of labeled monomeric residues within the polymeric molecule. For an example of different iCORE fragmentation signatures created by differential isotope labeling, see iCOREs G1-G10 described in FIG. 3. Isotope-labeling of a polypeptide, polynucleotide, or polysaccharide at a specific residue, for example, a specific amino acid, nucleotide, or monosaccharide residue, produces a specific isotope labeling signature that can readily be identified, and distinguished from other polypeptides, polynucleotides, or polysaccharides of the same monomeric sequence that are labeled at a different residue, via MS assays, for example, via MS/MS assays. This allows for the generation of multiple iCOREs comprising the same monomeric sequence, for example, multiple polynucleotide iCOREs comprising the same amino acid sequence, but isotope-labeled at different residues. Such differentially labeled, isobaric iCOREs exhibit identical physiochemical properties for the purpose of MS assays, and are, thus, detectable at an equal level of sensitivity, specificity, and accuracy in MS assays, while still

being distinguishable by their unique fragmentation signature, for example, in LC MS/MS assays as described herein.

[0089] In some embodiments, iCOREs are polypeptides. Typically, polypeptide iCOREs comprise a sequence that is optimized for facile detection in MS assays, as exemplified by the iCORE sequences provided herein. In some embodiments, a library of isobaric iCOREs is provided, in which all iCOREs comprise the same amino acid sequence, but different iCOREs comprise different isotope-labeling signatures. Even short polypeptides, e.g. of about 5-20 amino acids in length, allow for the generation of tens or hundreds of unique isotope-labeling signatures, which translates into the possibility to detect and analyze tens or hundreds of analytes in a multiplexed MS assay using isobaric iCOREs of such lengths. In some embodiments, a library of peptide iCOREs is provided that comprises iCOREs of different amino acid sequences, further increasing the number of unique isotope-labeling signatures and, thus, the number of analytes that can be assayed in a single multiplexed MS experiment. Such heterogeneous iCORE libraries can be generated, for example, by combining a first iCORE library comprising isobaric, differentially labeled iCOREs with an additional library comprising isobaric, differentially labeled iCOREs. For example, such a library could be generated by combining a library of glu-fib based iCOREs with a library of bradykinin based iCOREs and a library of angiotensin I-based iCOREs, thus creating a larger, combinatorial library of iCOREs.

[0090] Some aspects of this invention relate to the recognition that even a relatively small number of relatively short polypeptide tags that are useful in MS assays provide virtually unlimited multiplexing capabilities, since polypeptides are comprised of amino acid residues, and there are 20 naturally occurring and a large number of artificial amino acids that are useful as building blocks for iCOREs. For example, an iCORE library based on isobaric, differentially labeled glu-fib (EGVNDNEEGFFSAR (SEQ ID NO: 1)) iCOREs allows for the detection of about 10-100 analytes in a multiplexed MS assay, while an iCORE library based on isobaric, differentially labeled ACTH7-38 (FRWGKPVGKKRRPVKVYPNGAEDESAEAFPLE (SEQ ID NO: 34)) iCOREs allows for the detection of about 10-400 analytes in a multiplex MS assay. Combinatorial iCORE libraries, accordingly, allow for the assessment of thousands of analytes simultaneously.

[0091] Some aspects of this invention provide sets or libraries of iCOREs. In some embodiments, the iCOREs comprise or are conjugated to a reactive moiety that reacts with another molecule. For example, in some embodiments, the iCOREs comprise a reactive moiety that forms a covalent bond to a molecule of interest, for example, to an analyte, or a set of analytes. In some embodiments, the molecule of interest is a peptide or protein, a carbohydrate, a nucleic acid, or a lipid. In some embodiments, the molecule of interest to which the iCORE is aimed to be conjugated is a binding agent, for example, an antibody, antibody fragment, aptamer, ligand, receptor, or adnectin. For example, in some embodiments, a set of iCOREs is provided that is conjugated to a reactive chemical moiety that forms a covalent bond to a protein if contacted with a protein. Reactive chemical moieties are well known to those of skill in the art. One exemplary embodiment is depicted in FIG. 7 to illustrate this point. Here, the iCORE is a glu-fib peptide, covalently bound to a chemical moiety that polym-

erizes under UV light with PEG monomers to form a PEG hydrogel. Polyethylene glycol (PEG), also known as poly (oxyethylene) glycol, is a condensation polymer of ethylene oxide and water having the general chemical formula $\text{HO}(\text{CH}_2\text{CH}_2\text{O})[\text{n}]\text{H}$. Such hydrogels are useful, for example, in the generation of engineered tissues and the conjugation of an iCORE to such a hydrogel allows for tracking a tissue grown on such a tagged hydrogel in in vitro or in vivo experiments as described in more detail elsewhere herein. Other reactive chemical moieties and methods for conjugation of such moieties to iCOREs provided herein are well known to those of skill in the art and the invention is not limited in this respect. It will be apparent to those of skill in the art that the type of reactive chemical moiety can be selected to be reactive with a moiety comprised in the molecule of interest to which the iCORE is aimed to be conjugated. In some embodiments, iCOREs are provided that are conjugated to a reactive chemical moiety via a covalent bond. In some embodiments, iCOREs are provided that are conjugated to a reactive chemical moiety via a cleavable linker, for example, a photocleavable linker.

[0092] Providing iCOREs conjugated to reactive chemical moieties allows for the customized labeling of molecules of interest by the end-user, for example, by a scientist, thus extending the versatility of iCORE technology to custom-made experimental designs. For example, in some embodiments, iCOREs conjugated with a reactive moiety that forms a covalent bond with peptides may be used to label proteins or peptides obtained from a biological sample for comparison with proteins or peptides obtained from a different sample, much in the same way as ITRAQ tags are commonly used. In contrast to ITRAQ technology, the virtually unlimited number of different fragmentation signatures available for a given iCORE set, for example, a set of isobaric peptide iCOREs as described herein, allows for the simultaneous assessment of a virtually unlimited number of analytes, for example, for the simultaneous assessment and comparison of a virtually unlimited number of protein or peptide samples.

[0093] Scalability of iCORE technology depends on the specific binding agents used (for embodiments, in which iCOREs conjugated to specific binding agents are used or provided) and the specific, distinguishable mass reporters, e.g., iCOREs, having a unique, distinguishable fragmentation signature. Regarding specific binding agents, iCOREs can be conjugated to virtually all commercially available binding agents, including, but not limited to, antibodies, antibody fragments, aptamers, ligand-binding proteins or proteins domains, and adnectins by methods well known to those of skill in the art. The number of iCOREs with unique fragmentation signatures scales with the peptide space, which scales as 20^n , where n =length of peptide, for peptide iCOREs comprising only naturally occurring amino acids. Accordingly, even very short peptide iCOREs provide a large number of possible unique fragmentation signatures, while longer peptides provide virtually limitless unique signatures. In some embodiments, peptide iCOREs that are readily detected by MS are preferable, but even the exemplary, readily detectable peptide iCOREs described herein alone allow for hundreds or thousands of unique fragmentation signatures. Some peptide sequences suitable for the generation of iCOREs are provided herein and additional suitable peptide sequences will be apparent to those of skill in the art based on this disclosure.

[0094] In some embodiments, iCOREs are provided that are conjugated to a binding agent. A binding agent is an agent that specifically binds a molecule, for example, an analyte. In some embodiments, the binding agent is an antibody or antibody fragment. In some embodiments, the binding agent is a peptide or protein. In some embodiments, the binding agent is an aptamer or adnectin. In some embodiments, the binding agent is a ligand or a receptor, or comprises a ligand binding domain.

[0095] Accordingly, in some embodiments, iCORE-binding agent conjugates are provided that specifically bind to an analyte. In some embodiments, the conjugation of iCORE and binding agent is via a covalent bond, for example, a covalent peptide bond in an iCORE-binding agent fusion protein. In some embodiments, the conjugation of iCORE and binding agent is via a cleavable linker, for example, a protease cleavable linker or a photocleavable linker. In some embodiments, the linker is a peptide linker comprising a protease cleavage site. Some exemplary photocleavable linkers and protease cleavage sites are described herein, and additional linkers and cleavage sites will be apparent to those of skill in the art, and the invention is not limited in this respect. For example, in some embodiments, an iCORE is provided that is conjugated to a binding agent, for example, an antibody, via a photocleavable linker, as described herein. In some embodiments, a set of iCOREs is provided that is conjugated to a set of binding agents, for example, a set of antibodies, or antibody fragments, or ligand binding peptides and proteins, or aptamers or adnectins, or any combination of such agents. Typically, the iCOREs in such a set are conjugated to the binding agents in a manner that allows for the identification of a particular binding agent, and, thus, the analyte bound by a binding agent, by determining the identity of the iCORE.

[0096] For example, in a set of 10 different iCOREs (e.g., iCORE G1-G10 in FIG. 2) that are conjugated to a set of 10 different binding agents (B1-B10), which, in turn, specifically bind a set of analytes (A1-A10), the iCOREs are conjugated to the binding agents so that each iCORE of a unique fragmentation signature is conjugated to particular binding agent, e.g. G1 to B1, G2 to B2, G3 to B3, and so forth. The unique fragmentation signature of each iCORE can be identified and distinguished from the fragmentation signatures of the other iCOREs in the library in an MS assay, e.g. the MS/MS assay depicted in FIG. 3. Accordingly, the presence or absence of a specific binding agent in a sample can be inferred from the presence or absence of the unique fragmentation signature of the iCORE associated with that particular binding agent. In some embodiments, a sample is assayed for the presence or absence of a set of analytes by contacting the sample with the set of iCOREs conjugated to the binding agents under conditions suitable for the binding agents to bind to their respective analytes. Subsequently, those iCORE-conjugated binding agents that are specifically bound to their respective analytes are enriched or isolated. In some embodiments, the iCOREs are then released, for example, by cleaving the linker connecting the iCORE to the binding agent. In some embodiments, the iCOREs are then subjected to MS analysis, for example, in an MS/MS assay, and the presence or absence, and/or the quantity of each iCORE is determined. In some embodiments, the presence or absence, and/or the quantity of each analyte is determined from the result of the MS assay. For example, in some embodiments, if an iCORE MS assay results in the detection

of iCOREs G1, G3, G5-7 and G10, the presence of analytes A1, A3, A5-7, and A10, and/or the absence of analytes A2, A4-6, and A8-9 can be determined. In some embodiments, a comparison of the signal obtained for a given iCORE in the MS assay can be used to determine the quantity of the respective analyte in the sample, and/or to compare the level of the respective analytes in the sample to the level of a different analyte in the sample via MS signal comparison. For example, if the MS result obtained from a sample contacted with 10 iCOREs conjugated to 10 binding agents specifically binding to 10 analytes would represent the MS result as shown in FIG. 3, the presence of all unique fragmentation signatures of iCOREs G1-10 would indicate that all analytes A1-10 are present in the sample. In some embodiments, the similar signal level of all unique fragmentation signatures of iCOREs G1-10 may further indicate that all analytes A1-10 are present at similar levels in the sample, if the binding agent-analyte interactions are substantially similar across the analytes.

[0097] In some embodiments, a library of isobaric iCOREs is provided in which the iCOREs comprise the same sequence, for example, a set of peptide or nucleic acid iCOREs that comprise the same amino acid or nucleotide sequence, respectively, but in which the different iCOREs are isotope-labeled at different positions, for example, at different amino acid or nucleotide residues within the amino acid or nucleotide sequence.

[0098] Some aspects of iCORE technology rely on or use proven, existing technologies, for example, bead-based immuno-enrichment, photo-labile bioconjugation chemistry, direct labeling of analytes, for example, of proteins by reactive chemical moiety conjugation, and qualitative and/or quantitative mass spectrometry, for example, LC MS/MS assays. These technologies are well known to those of skill in the art and many variations and equivalents of the exemplary embodiments described herein will be apparent to the skilled artisan based on this disclosure. It will be appreciated that iCORE technology can be applied to many different analyte detection scenarios, e.g. diagnostic scenarios, as well as many different tracking scenarios. The invention is not limited in this respect.

[0099] For example, some embodiments of iCORE technology as described herein provide molecular diagnostic assays (e.g. biomarker detection, monitoring of vaccines, and HLA screening) as well as qualitative and quantitative assays for basic science research (e.g. protein expression profiling, cell profiling, tissue profiling, and metabolic profiling). For example, certain embodiments provide a library of iCORE-labeled binding agents (e.g. antibodies or antibody fragments) targeted against a panel of serum biomarkers (e.g. organ- and/or disease-specific markers) that can be used for monitoring biological processes by analyzing body fluid or tissue samples obtained from a subject (e.g. tumor evolution, host responses after drug administration, or drug metabolism). Some embodiments of iCORE technology are particularly suited for such multi-parameter studies, particularly in the diagnosis of diseases or conditions, because iCORE technology allows for increases in detection sensitivity and resolution through multiplexing, which, in turn, allows for the efficient detection of multi-parameter diagnostic signatures.

[0100] Some embodiments provide a library of chemically-active iCOREs, for example, isobaric peptide iCOREs conjugated to a reactive chemical moiety that can be utilized

by an end user to encode user-defined experiments. For example, in case-and-control comparative proteomics studies, proteins from each unique condition, experiment, or subject, can be encoded (e.g., labeled) with a particular iCORE fragmentation signature. All samples can then be combined and simultaneously analyzed by LC MS/MS to determine the relative effects of the experimental conditions. The use of iCORE technology enables comparative studies at a scale precluded by current technologies and methods.

[0101] It will be appreciated by the skilled artisan that the use of iCORE technology, as described herein, is not limited to analyte detection, encoding and translation. Because iCOREs are versatile, novel MS tags that have multiple advantages over previously known MS tags, iCORE technology is amenable to adaptation to various MS tag detection technologies and diagnostic strategies. For example, iCORE technology can be used in the context of multiplexed enzyme (e.g., protease) activity profiling using MS/MS strategies, as described in PCT application PCT/US2010/000633, published as WO/2010/101628 on Oct. 9, 2010, the entire contents of which are incorporated herein by reference.

[0102] For example, iCOREs can be conjugated to a carrier, for example, a nanoparticle (e.g., a nanoworm (NW)), via a linker comprising an amino acid sequence that is cleaved by an enzyme of interest, e.g., by a protease associated with a disease. Such iCORE NWs, also sometimes referred to as pro-diagnostic reagents, can then be administered to a subject to interrogate the activity of the enzyme in the subject. In some embodiments, a set of iCOREs is provided that is linked to a carrier via different linkers targeted by different enzymes (e.g., different proteases) in a way that the activity of each protease releases one specific iCORE tag. Used in this manner, iCORE technology can be employed to interrogate the activities of a plurality of disease-associated enzymes (e.g., disease-associated proteases) in parallel in one multiplexed MS assay.

[0103] In some embodiments, a set of pro-diagnostic iCORE reagents is administered to a subject suspected to exhibit an aberrant (e.g., pathogenic) level of activity of one or more enzymes (e.g., one or more proteases) to probe the activity levels of numerous proteases in one multiplexed assay. While the carrier of the prodiagnostic reagent is typically not secreted into the urine of the subject, when exposed to protease activity, the respective iCORE is cleaved from the NW and released into the urine of the subject. Urine samples are collected after a sufficient time after administration of the pro-diagnostic reagent to the subject has passed for the reagent to be exposed to enzyme activity in the subject. Urine samples are then subjected to MS/MS assays as described in more detail elsewhere herein, or in PCT/US2010/000633, published as WO/2010/101628, to detect the cleaved iCOREs and determine whether or not the subject exhibits an aberrant protease activity signature.

[0104] Protease activities that are associated with disease are well known to those of skill in the art (see, e.g., Table 1 of PCT/US2010/000633, published as WO/2010/101628, for some non-limiting examples of disease-associated enzyme activities, incorporated herein in its entirety by reference). Protease target sequences that are useful as linker sequences for iCOREs to carriers are also well known to those of skill in the art. Some exemplary iCOREs, protease-cleavable linkers, and methods for the diagnosis of diseases

in a subject are described in detail herein. Additional iCOREs, linkers, and methods will be apparent to those of skill in the art based on this disclosure, and the disclosure is not limited in these aspects. As will be apparent to the skilled artisan, the use of iCORE technology in the context of multiplex protease activity assays allows for an improvement of multiplexing capabilities and increased ease of detection as compared to previously described methods.

[0105] Some embodiments provide kits of reagents described herein. For example, some embodiments provide a kit comprising a set of iCOREs, for example, isobaric peptide iCOREs, that comprises iCOREs having different fragmentation signatures. In some embodiments, the iCOREs are conjugated to a reactive chemical moiety, for example, a peptide-reactive moiety that forms a covalent bond to a peptide when contacted with such a peptide. In some embodiments, the different iCOREs in a set of iCOREs so provided are separated, thus allowing the end-user to label a particular sample, for example, a protein sample, with a specific iCORE and a different sample with a different iCORE, thus allowing subsequent mixing and simultaneous analysis if the samples in an MS assay, e.g. in an LC MS/MS assay.

[0106] In some embodiments, a kit is provided that comprises a set or library of binding agents that are conjugated to iCOREs, as described herein. In some embodiments, each binding agent binding a particular analyte is conjugated to an iCORE of a particular fragmentation signature, so that the analyte bound by the binding agent can be identified by the fragmentation signature of the iCORE.

[0107] Some aspects of this invention relate to biochemical encoding of parameters, for example, of analytes, into mass tags, for example, iCOREs. In some embodiments, the biochemical encoding of an analyte parameter, e.g., the presence of an analyte in a sample or the amount of an analyte in a sample, involves contacting the sample with a binding agent (e.g., an antibody or antibody fragment) specifically binding the analyte and bound to an iCORE the fragmentation signature is assigned to the specific analyte bound by the binding agent, isolating or enriching for those binding agent molecules that have bound the analyte, optionally, releasing the iCORE from the isolated or enriched binding agent molecules and subjecting the iCORE (released or not) to an MS assay. In other embodiments, the biochemical encoding is achieved by translating a parameter, for example, the presence of an analyte in a sample, or the identity of a cell or tissue, into an iCORE signature by tagging the sample, cell, or tissue with a unique iCORE, for example, by spiking an iCORE into the sample, or by attaching (e.g., covalently via a cleavable linker) an iCORE to a cell or tissue. In some embodiments, the biochemical encoding is carried out for multiple parameters in parallel, for example, by contacting a complex sample (e.g. a blood or tissue sample) containing or suspected to contain a plurality of analytes with a plurality of iCOREs bound to a plurality of binding agents, wherein each binding agent specifically binds a particular target analyte and is conjugated to a specific iCORE with a unique fragmentation signature. After isolation or enrichment of analyte-bound binding agents and rescue of the conjugated iCOREs from the isolated or enriched binding agents, the presence or absence as well as the relative or absolute amount of the

target analytes can be determined by an MS assay, e.g., an LC MS/MS assay as described in more detail elsewhere herein.

[0108] A schematic of an exemplary biochemical encoding strategy in which the presence of a plurality of analytes in a sample is translated into iCOREs is described in FIG. 1. In some embodiments, the translation of a parameter into an iCORE comprises a step of deconvolution, for example, a step of analyte enrichment from a complex biological sample. Deconvolution can increase assay sensitivity, specificity, and/or accuracy. In some embodiments, a sample to be assessed is enriched for an analyte by contacting with an affinity agent that binds the analyte, specifically or non-specifically, and allows for enrichment of the analyte, for example, by physical separation of the bound analyte from unbound material. An affinity agent may be an agent that non-specifically binds the analyte based on a non-specific interaction, for example, a non-specific interaction based on the surface charge of the analyte. Such non-specific binding will typically result in a lower level of deconvolution of the sample than the use of an affinity agent that specifically binds the analyte, since non-specific binding is generally not restricted to the specific analyte, but also extends to other molecules of similar physiochemical properties. Alternatively, an affinity agent may also be an agent that specifically binds the analyte. For example, an affinity agent may be a binding agent, for example, an antibody or antibody fragment, that is immobilized on a solid support, such as a bead or membrane surface. Solid supports useful for the separation of materials bound to their surface from biological samples are well known to those of skill in the art, and include, but are not limited to, membranes, resins, beads, and the surface of plates, dishes and tubes.

[0109] In some embodiments, a complex sample is contacted for deconvolution with an affinity agent under conditions suitable for the affinity agents to bind the analyte(s). Typically, the sample comprises a liquid phase that is contacted with the affinity agent. In some embodiments, the affinity agent, and any analyte(s) bounds to it are subsequently removed from the sample, for example, by physical separation, or aspiration of liquid supernatant. In some embodiments, the affinity agent is washed after separation to remove remaining residue of the biological sample on the affinity agent for further deconvolution.

[0110] In some embodiments of biochemical encoding of analytes, a sample is contacted with an affinity agent or a set of affinity agents that bind(s) a set of analytes of interest, under conditions suitable for the affinity agents to bind the respective analytes, and for the analytes to become immobilized on the affinity agent. In some embodiments, the sample is then de-convoluted, for example, by washing away any unbound material. In some embodiments, the sample, deconvoluted or not, is then contacted with a set of iCOREs conjugated to binding agents that specifically bind a set of analytes under conditions suitable for the binding agents to bind their respective analytes. In some embodiments, those iCOREs conjugated to binding agents that have actually bound to an analyte are subsequently isolated or enriched, for example, by removing any unbound iCOREs and binding agents from the sample. This “rescued” set of iCOREs represents a biochemical encoding of the analytes present in the sample. In some embodiments, the iCOREs are conjugated to the binding agents via a cleavable linker, and the process of biochemical encoding of the analytes in

the sample comprises a step of cleaving the linker and releasing the iCOREs from the binding agents. In some embodiments, the rescued iCOREs are subjected to an MS assay, for example, an MS/MS assay as described herein, to determine their identity and/or quantity, and thus the identity and/or quantity of the respective analytes in the sample.

[0111] It will be appreciated, that translation for biochemical encoding is not restricted to analytes, but can be applied to other parameters as well. For example, iCORE technology provides a possibility to track samples, cells, tissues, reagents, or molecules. As one example of biochemical encoding for tracking purposes, iCORE technology can be used to encode small molecule compound identities, e.g., in the context of small molecule screens. Biochemical encoding of small molecule compound identities is particularly useful in the context of large combinatorial small molecule screens, where thousands, tens of thousands, or hundreds of thousands of small molecule combinations can be tracked without the logistic problems commonly imposed by such screens. In some embodiments, each small molecule in such a screen is associated with an iCORE having a unique fragmentation signature, for example, by simply adding the iCORE to the small molecule, thus generating a composition comprising the iCORE and the small molecule. In some embodiments, the composition comprises the iCORE in the small molecule at a specific, known ratio. In some embodiments, the composition is then used in a combinatorial chemical screening assay, for example, in an assay in which multiple combinations of multiple small molecule compounds in multiple dilution ratios are tested for a desired effect. The desired effect may be any effect for which small molecule libraries are currently being screened, including, but not limited to, a biological effect, such as the induction of cell death in a cancer cell line, the inhibition of aberrant cell proliferation, the induction or repression of the expression of a specific gene, or the induction of epigenetic reprogramming. Many other desired effects for which small molecule compound libraries can be screened will be apparent to those skilled in the art, and the invention is not limited in this respect.

[0112] Biochemical encoding of each small molecule with a unique iCORE allows for the identification and/or the quantification of each small molecule compound in a mixture or dilution series of small molecule compounds by subjecting the mixture or dilution to an MS assay. Any specific iCORE detected in the MS assay can then be traced back to the particular small molecule compound it encodes. If the concentration of the small molecule compound was also encoded, for example, by generating the original iCORE/small molecule compound mixture used in the screen to comprise a particular ratio of the two components, the intensity of signal detected for the specific iCORE can be used to determine the concentration of the small molecule compound in the mixture or dilution of interest.

[0113] For example, a simple combinatorial small molecule screening assay may comprise a combinatorial dilution series screen of 10 small molecule compounds (SMC 1-10) at 10 different dilutions (D 1-10), in which all possible mixtures of any or all of the ten compounds at any possible dilution ratios of the compounds are to be tested. For example, ten potentially synergistic cancer drug candidates might be tested this way for their efficiency to kill cancer cells, e.g., to investigate which combination of these candidates at which ratio effects the greatest synergy or the

greatest biological effect. Even such a simple screen poses a significant logistic challenge in view of the necessity to map the contents and drug concentrations of multiple thousands of samples. Biochemical encoding of the drug candidates with only 10 iCOREs (e.g., iCORE G1-10, assigned to CD1-10, respectively) circumvents the problems associated with this logistical undertaking. For example, each drug candidate could be spiked with one of the iCOREs at a specific, known ratio of drug candidates and iCORE at the beginning of the experiment. The generation of dilution series and drug candidate mixtures can then be carried out as usual, but without the necessity to keep track of all pipetting steps and the resulting sample contents. After screen read-out, for example after the detection of a sample showing the desired biological effect, the supernatant of the sample can be obtained, and subjected to an MS assay. The MS assay will not only identify the presence or absence of each of the drug candidates, but can also be used to determine the ratios of the drug candidates identified to be present in the sample based on the known iCORE/drug candidate ratios in the original spiked mixture at the beginning of the experiment.

[0114] For example, if in the hypothetical screen described above, the MS assay would detect the presence of iCOREs 1, 4, and 7 at a ratio of 1:10:2 in a sample of interest, the presence of drug candidates 1, 4, and 7 can be inferred. It can further be inferred that drug candidate 1 was present at the lowest concentration (e.g., D10), drug candidate 4 at the highest concentration (e.g., D1), and drug candidate 7 at the second lowest concentration (e.g., D9) of the dilution series. If molar ratios of iCORE and drug candidate were known in the original spiked sample at the beginning of the experiment, a determination of the concentration of each drug candidate may be possible directly from the MS data, without the need to track dilutions.

[0115] Aside from circumventing the necessity to track thousands of pipetting steps, drug candidate tracking with iCORE biochemical encoding during combinatorial screening allows for an assessment of the actual conditions in the sample of interest, which, for example, avoids erroneous results caused by pipetting errors. The particular iCOREs used for biochemical encoding and drug screening should be chosen to not interfere with the process being probed. Peptide iCOREs are suitable for a variety of biochemical encoding applications during drug screening, since the peptides are inert to most biochemical processes. It will be apparent to those of skill in the art, that, depending on the length and the conditions encountered by the iCOREs during the drug screening process, useful iCOREs will have to exhibit sufficient stability to withstand the conditions encountered. In some embodiments, peptide iCOREs with improved stability are generated by using non-naturally occurring amino acid residues and/or peptides comprising non-hydrolyzable bonds. Peptides exhibiting increased stability, for example, peptides comprising non-hydrolyzable bonds, are well known to those of skill in the art, and the skilled artisan will be able to readily ascertain methods and reagents for the generation.

[0116] In some embodiments, a method for tracking a cell, cell population, tissue, or compound by biochemical encoding via iCOREs is provided. In some such embodiments, the method comprises conjugating a unique iCORE to a cell, cell population, or tissue, for example, by contacting the cell, population of cells, or tissue with an iCORE under conditions suitable for the iCORE to bind to the cell,

population of cells, or tissue. In some embodiments, an iCORE covalently bound to a binding agent is used to achieve this conjugation. For example, in some embodiments, the iCORE is bound to a binding agent, for example, an antibody or antibody fragment, that specifically binds a surface marker, for example, a surface protein or polysaccharide marker of the cell, population of cells, or tissue. In other embodiments, a cell, cell population, tissue, or compound is contacted with an iCORE comprising or conjugated to a reactive chemical moiety under conditions suitable for the reactive chemical moiety to form a covalent bond with a surface marker of the cell, cell population, tissue, or with the compound. In some embodiments, the compound is a compound that is used in the generation of or can be embedded into a hydrogel, for example, a PEG hydrogel as described in FIG. 7. In some embodiments, the hydrogel is used in the generation of beads or in the production of a scaffold for tissue engineering, e.g., as a substrate for growing artificial, engineered tissues, for example, engineered organs, such as micro-livers. The result of this approach is the production of iCORE-tagged beads or scaffolds and, if such scaffolds are used for the generation of engineered organs, iCORE-tagged organs. In some embodiments, a plurality of different iCORE-tagged beads or engineered organs is subjected to different experimental conditions, e.g., the exposure to different growth factors and/or small molecule compounds, or growth factor and/or small molecule compound combinations at different dilutions, and those beads or engineered organs exhibiting a desired phenotype or parameter at the end of the experiment is identified by rescuing the iCORE tag from the desired bead or organ and subjecting it to an MS assay, e.g., an LC MS/MS assay.

[0117] In some embodiments, a set or library of iCOREs is provided that comprise or is conjugated to a reactive chemical group that can form a covalent bond to a peptide. Such reactive chemical groups are known to those of skill in the art and include, but are not limited to those reactive chemical groups conjugated or comprised in mass tags used in iTRAQ assays. In some embodiments, methods for the use of such peptide-binding iCOREs are provided, for example for the simultaneous analysis of proteins in samples of different origin, for example, blood or tissue samples from different subjects. In some embodiments, protein samples obtained from different subjects, for example, from different experimental mice, or from different tumor biopsies, are conjugated to a unique iCORE each. In some embodiments, a plurality of such iCORE-labeled samples is then pooled and simultaneously analyzed in an MS assay, for example, an LC MS/MS assay. In some embodiments, the data obtained from such simultaneous proteomics experiments is used to compare protein expression levels across a large number of samples, for example, across at least 10, at least 15, at least 20, at least 25, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 150, at least 200, at least 250, at least 500, at least 750, at least 1000, or more than 1000 samples. The large number of possible unique fragmentation signatures that can be obtained from a single iCORE base sequence, for example, from a glu-fib amino acid sequence, allows for the unique labeling and subsequent simultaneous analysis of such sample numbers, which is far beyond the multiplexing capabilities of the currently available technologies, e.g., iTRAQ.

[0118] Some aspects of this invention provide multiplexed diagnostic assays using iCORE technology. Currently, many molecular biomarkers are known that are correlated with, and, thus, indicative of a disease or condition in a subject. For example, a plethora of molecular biomarkers have been reported for diseases such as cancer, infectious disease, cardiovascular disease, renal disease, autoimmune disease, toxic states, and for the detection of certain side effects of drugs. Current diagnostic assays typically focus on one or only a few biomarkers to detect a specific disease or condition. For example, some cancers, such as prostate and ovarian cancer are monitored by use of single biomarkers in the blood of subjects diagnosed or suspected to have such a cancer. Such diagnostic techniques are achieved, for instance using fluorescence detection of molecular markers which are activated in a particular disease state. Other diagnostic techniques to monitor molecules include the use of gene arrays for stratifying breast cancer patients via unique signatures (Sotiriou C., Piccart, M. J., Taking gene-expression profiling to the clinic: when will molecular signatures become relevant to patient care?, *Nat. Rev. Cancer*, 7, 534-553, 2007) and sequencing target gene mutations to uncover the molecular lesions that predict glioblastoma response to EGFR kinase inhibitors (Mellinghoff, I. et al., Molecular Determinants of the response of glioblastomas to EGFR Kinase Inhibitors, *NEJM*, 353, 16, 2005). Both of the foregoing references are incorporated herein in their entirety by reference for their disclosure of diagnostic biomarkers.

[0119] Some of the reagents and methods for multiplex analyte detection and quantification in complex samples, for example, in a body fluid or tissue sample obtained from a subject, allow for the sensitive and accurate assessment of multiple biomarkers simultaneously, thus providing an avenue for the generation of clinically relevant multi-parameter biomarker profiles of a single disease or condition or of a plurality of diseases or conditions from such a sample.

[0120] In some embodiments, iCORE technology as described herein is used to assess a plurality of biomarkers, e.g., the presence or absence or the quantity of an analyte in a biological sample obtained from a subject, such as a body fluid (e.g. blood, serum, urine, cerebrospinal fluid, or lymph) or tissue (e.g., tumor, malignant, neoplastic, hyperplastic, liver, kidney, lung, muscle, lymph node, or brain tissue). In some embodiments, the plurality of biomarkers comprises a panel of biomarkers known to be associated with a particular disease or condition. In some embodiments, a diagnostic iCORE method is provided. In some embodiments, the method comprises obtaining a sample from a subject. In some embodiments, the subject is a human subject. In some embodiments, the subject is a non-human mammal. In some embodiments, the method comprises a deconvolution step in which target analytes are isolated or enriched, for example, by a deconvolution method as described elsewhere herein. In some embodiments, the method comprises contacting the sample, whether deconvoluted or not, with a set or library of iCOREs conjugated to a set of binding agents specifically binding to the target analytes, as described herein, under conditions suitable for the iCORE-conjugated binding agents to bind their respective target ligands. In some embodiments, the method comprises a step of isolating or enriching the bound iCORE-conjugated binding agents, for example, by washing off unbound binding agents. In some embodiments, the method comprises rescuing the iCOREs conjugated to the analyte-binding binding agents, for

example, by cleavage of a linker connecting the iCOREs to the binding agents, as described elsewhere herein. In some embodiments, the method comprises subjecting the rescued iCOREs to an MS/MS assay, for example, an MS/MS assay as described herein. In some embodiments, the method further comprises determining the signal obtained from each specific iCORE in the sample based on the results of the MS/MS assay. In some embodiments, the method comprises determining the presence or absence of the analyte represented or identified by the specific iCORE based on whether or not a signal was obtained in the MS/MS assay that identified the specific iCORE. In some embodiments, the method comprises quantifying the amount or concentration of the analyte in the sample, or relative to the amount or concentration of other analytes in the sample, or relative to the amount or concentration of the same analyte in a different sample, for example, a reference or control sample, by quantifying the signal obtained from the specific iCORE in the MS/MS assay.

[0121] In some embodiments, the assessment of multiple biomarkers in a single diagnostic assay increases the specificity, sensitivity, and/or accuracy of the assessment of the biomarkers and/or the diagnosis inferred from the assay. In some embodiments, the iCORE MS/MS assay also comprises the assessment of a control marker, for example, of an analyte or a set of analytes that are known to be virtually constant across healthy and diseased states or across individuals. There are many protein biomarkers present in blood that are useful for assessment by iCORE MS/MS analyses. For example, in certain embodiments for ovarian cancer, 5-20 protein or peptide biomarkers may be monitored (see, e.g., Petricoin, E., et. Al., Use of proteomic patterns in serum to identify ovarian cancer, *Lancet*, 359, 572, 2002, incorporated in its entirety by reference for disclosure of ovarian cancer biomarkers). Serum cytokines, of which IFN-gamma and TNF-alpha are normally classified, including, but not limited to, IL-1, IL-2, IL-4, IL-5, IL-6, IL-10, IL-12, IL-13, IL-17, IL-21, IL-23, MCP-1, TGF-beta, TNF-beta, TWEAK, GM-CSF, and Granzyme B, are useful for monitoring inflammatory processes involved in cancer, as well as host responses to vaccines, and infectious diseases. It will be apparent to those of skill in the art that the biomarkers described herein are for illustration only and not meant to limit the invention in this respect. Any biomarker, and particularly, any peptide or protein biomarker can be assessed by iCORE MS/MS assays provided herein, and the invention is not limited in this respect.

[0122] In some embodiments, iCORE technology is used in the context of personalized medicine. For example, in some embodiments, a clinical intervention, e.g., a drug, a treatment schedule, or a surgical intervention, is chosen from a group of such interventions, based on the results of an iCORE assay performed on a sample obtained from the subject, for example, a blood or tissue sample assessed for a set of biomarkers, or based on analyte or enzymatic activity signatures produced from an in vivo iCORE assay, by methods described herein.

[0123] In some embodiments, iCORE technology is used to monitor a subject's response to a treatment. For example, in some embodiments, iCORE assays are performed repeatedly during the treatment of a diseased subject, for example, in order to determine whether the assessed biomarkers are approaching desired values, for example, values typically associated with a state of health. In some embodiments, a

treatment regimen may be adjusted based on the results of an iCORE assay performed during the course of the treatment, for example, in order to increase the efficacy of the treatment or to decrease the dosage of a drug used to the minimal effective dose in order to avoid undesired side effects.

[0124] In some embodiments, iCORE technology is used in clinical trials of novel drugs and drug candidates. Multiplex iCORE technology allows for the assessment of a variety of disease-associated biomarkers, and, in some embodiments, also for the monitoring of a plurality of base parameters, e.g. blood cholesterol, blood triacylglycerides, ketone bodies, and so forth, in a subject during a clinical trial. This allows for the comprehensive monitoring of multiple metabolic pathways during a clinical trial and the detection of side effects that may be non-symptomatic during the trial.

[0125] Mass spectrometry (MS) assays are well known to those of skill in the art. Of particular use in the context of some embodiments of iCORE technology and biochemical encoding described herein are tandem mass spectrometry (MS/MS) assays.

[0126] Tandem mass spectrometry (MS/MS) is used to produce structural information about a compound by fragmenting specific sample ions inside the mass spectrometer and identifying the resulting fragment ions. Tandem mass spectrometry allows for specific iCOREs to be detected in complex mixtures of iCOREs, for example, in complex sets or libraries of iCOREs on account of their unique fragmentation signature, which results in a specific and characteristic fragmentation pattern.

[0127] MS/MS assays typically comprise two or more MS steps with some form of fragmentation taking place between the steps. A tandem mass spectrometer typically comprises more than one analyzer, for example, two analyzers. In some embodiments, the analyzers are separated by a collision cell into which an inert gas (e.g. argon, xenon) is admitted to collide with the selected sample ions and bring about their fragmentation. However, some MS/MS assays can also be performed on certain single analyzer mass spectrometers such as ion trap and time-of-flight instruments, for example, by using a post-source decay experiment to effect the fragmentation of sample ions.

[0128] In some embodiments, because both the masses of the parent peptide as well as the reporter ions of individual iCOREs are known, the detection and quantification of the iCOREs in a complex sample can be achieved by selected reaction monitoring (SRM), also referred to as multiple reaction monitoring (MRM) (see, e.g., Lange, V., Picotti, P., Domon, B., Aebersold, R., *Selected reaction monitoring for quantitative proteomics: a tutorial*, *Molecular Systems Biology* 4:222, 2008, the entire contents of which are incorporated herein by reference). A typical SRM experiment is enabled by the unique capability of a triple quadrupole (QQQ) MS for quantitative analysis in which the first and third quadrupoles act as filters to specifically select pre-defined m/z values corresponding to the peptide ion and a specific fragment ion of the peptide.

[0129] For example, in some embodiments, for iCOREs based on 10-plex glu-fib iCORE library (G10), the first quadrupole serves to collect the parent Glu-fib peptide (e.g. 789.85 m/z doubly charged), the second quadrupole fragments, while the third quadrupole acts to filter an individual reporter ion (e.g. 683.3 m/z) (FIG. 14). In some embodiments, such precursor/fragment ion 'transitions' (e.g. 789.85

to 683.3 m/z) are monitored over time, which results in high selectivity and sensitivity, as co-eluting background ions are filtered out effectively with this approach.

[0130] In some QQQ embodiments, unlike other MS techniques (e.g. Q-TOF), no full mass spectra are recorded and the non-scanning nature of QQQ analysis translates into 1-2 orders of magnitude increase in sensitivity over full scan techniques. Because isotopic amino acids can be selectively incorporated by the user to generate unique fragment ions with iCORE encoding, multiple unique precursor/fragment ion transitions can be designed into a family of iCOREs. For example, in FIGS. 2 and 3, the transitions are defined from the precursor peptide to the y7 family of fragment ions, while in FIG. 13c,d,e and FIG. 14, the transitions are defined from the precursor peptide to the y6 family of fragment ions.

[0131] In some embodiments, the first analyzer is used to select user-specified sample ions arising from a particular iCORE. These chosen iCORE ions, for example, the ions of the glu-fib iCOREs described herein, pass into the collision cell, are bombarded by the gas molecules which cause fragment ions to be formed, and these fragment ions are then separated according to their mass to charge ratios, by the second analyzer. The fragment ions arise directly from the parent iCORE ions, and thus produce a fingerprint pattern specific to the compound under investigation. (see FIG. 3 I. for data obtained in the first round of MS from a library of glu-fib iCOREs; and FIG. 3 II. for data obtained in the second round (after fragmentation) from the same library).

[0132] In some embodiments, the first analyzer allows the transmission of all iCORE ions, whilst the second analyzer is set to monitor specific iCORE fragment ions, for example, the glu-fib iCORE y₇ ions, as described in FIGS. 2 and 3, which are generated by bombardment of the sample ions with the collision gas in the collision cell. This strategy is particularly useful in embodiments, where a set of isobaric iCOREs is used, which fragment to produce distinct, but similar fragment ions. Using this strategy allows for monitoring only the part of the spectrum spanning the particular masses of the iCORE ions, e.g. the spectrum from about 812 to about 822 Da in the example using iCOREs G1-G10 as described in FIG. 3. This, in turn allows for a drastic reduction in scanning time, because the rest of the spectrum can be ignored.

[0133] In some embodiments, both analyzers are static and only specific parent ions are transmitted through the first analyzer and only specific fragments arising from these parent ions are measured by the second analyzer. The parent ion properties of any iCORE provided herein can be calculated according to methods well known to those of skill in the art. Additionally, any iCORE can be tested for its specific ionization and fragmentation signature in an MS/MS test run. Since the structure and MS/MS behavior of each iCORE is either known or can readily be calculated or established, this type of MS/MS assay can be employed to determine the presence or absence of a specific iCORE ion in a matrix or fragment ions, for example, a matrix of fragment ions originating from a complex isobaric iCORE library in a multiplex iCORE assay as provided by aspects of this invention.

[0134] In some embodiments, peptide iCOREs, sets or libraries of peptide iCOREs, and methods for the use of peptide iCOREs are provided. It is known to those of skill in the art that peptides fragment in a reasonably well-documented manner in MS/MS assays (see, e.g., P. Roep-

storff, J. Fohlmann, *Biomed. Mass Spectrom.*, 1984, 11, 601; and R. S. Johnson, K. Biemann, *Biomed. Environ. Mass Spectrom.*, 1989, 18, 945; the entire contents of which are incorporated herein by reference). In some embodiments, the peptides fragment along the peptide backbone (see, e.g., A. E. Ashcroft, P. J. Derrick in "Mass Spectrometry of Peptides" ed. D. M. Desiderio, CRC Press, Florida, 1990; the entire contents of which are incorporated herein by reference). There are three different types of bonds that can fragment along the amino acid backbone: the NH—CH, CH—CO, and CO—NH bonds. Each bond breakage gives rise to two species, one neutral and the other one charged, and only the charged species is monitored by the mass spectrometer. The charge can stay on either of the two fragments depending on the chemistry and relative proton affinity of the two species. Hence there are six possible fragment ions for each amino acid residue and these are labeled as a, b, and c ions, and x, y, and z ions, with the a, b, and c ions having the charge retained on the N-terminal fragment, and the x, y, and z ions having the charge retained on the C-terminal fragment. The most common fragmentation occurs at the CO—NH bonds of the peptide backbone, which give rise to the b and/or the y ions. FIG. 2 shows an exemplary nomenclature of 13 y-ions originating from a parental glu-fib iCORE sequence. Since the mass difference of adjacent fragment ions correlates to the mass of the respective amino acid residue missing in the shorter ion, e.g., the difference in the mass of the y₉ and y₁₀ ions is the mass of the D residue present in the y₁₀ but not the y₉ ion, the peaks representing each y ion can be identified. In the exemplary MS/MS experiment shown in FIG. 2, y ions y₃-y₁₁ could unambiguously be identified.

[0135] An example of an MS/MS spectrum obtained from a set of 10 isobaric iCOREs is illustrated in FIG. 3. The first round of MS generated one single peak, consistent with the isobaric character of the iCOREs (FIG. 3 I.). The iCORE ions were fragmented and the size of the fragments analyzed by the second analyzer to produce the MS/MS spectrum depicted in FIG. 3. II.). The iCOREs fragmented predominantly at the CO—NH bonds ions y ions y₃-y₁₁ could unambiguously be identified. A close-up of the region spanning the mass of the possible y₇ ions shows that all ten iCOREs were detected with highly similar signal strength, indicating that all fragment ions were present at about the same concentration. This indicates that there is no fragmentation bias in iCOREs of different fragmentation signature, which allows iCORE technology to be used in sensitive analyte quantification MS/MS assays.

[0136] In some embodiments, the iCOREs are polynucleotides or polysaccharides. MS/MS-based assays oligonucleotide and oligosaccharide fragment identification are well known to those of skill in the art and are similar to those used in peptide MS/MS assays. For a general overview of MS/MS assays useful for the detection and/or quantification of iCOREs provided herein, for example, of peptide, oligonucleotide, and/or oligosaccharide iCOREs, see, e.g., S. Pomerantz, J. A. Kowalak, J. A. McClosky, *J. Amer. Soc. Mass Spectrom.*, 1993, 4, 204; "An Introduction to Biological Mass Spectrometry", C. Dass, Wiley, USA, 2002; "The Expanding Role of Mass Spectrometry in Biotechnology", G. Siuzdak, MCC Press, San Diego, 2004; "Ionization Methods in Organic Mass Spectrometry", A. E. Ashcroft, Analytical Monograph, Royal Society of Chemistry, U K, 1997; www.astbury.leeds.ac.uk (A. E. Ashcroft's MS web

pages and tutorial); Chapter 9, pages 415-452 of “Mass Spectrometry: A Textbook” by Jurgen H. Gross, Springer; 2nd ed. edition (Mar. 1, 2011), ISBN-10: 9783642107092; U.S. Pat. Nos. 5,885,775; 7,412,332; 6,597,996; and “Mass Spectrometry: Clinical and Biomedical Applications Volume 1 (Modern Analytical Chemistry)” by Dominic M. Desiderio, Springer; 1 edition (Jan. 31, 1993), ISBN-13: 978-3642107092; the entire contents of all of which are incorporated herein by reference).

[0137] Some aspects of this invention provides kits of reagents useful for carrying out iCORE-based MS assays. Typically, a kit comprises a container housing an iCORE reagent, or a set or library of such reagents, as described herein. In some embodiments, a kit also includes instructions or labels describing the use of the iCORE reagents.

[0138] In some embodiments, iCORE reagents described herein, for example, iCOREs conjugated to a reactive chemical moiety, or iCOREs conjugated to a binding agent specifically binding a target ligand, or sets or libraries of such iCORE reagents, are assembled into kits for clinical and/or research applications. A kit may be designed to facilitate use of the methods described herein by clinicians and/or researches and can take many forms. Each of the iCORE reagents comprised in the kit, where applicable, may be provided in liquid form (e.g., in solution), or in solid form, (e.g., a dry powder). A kit may have a variety of forms, such as a blister pouch, a shrink wrapped pouch, a vacuum sealable pouch, a sealable thermoformed tray, or a similar pouch or tray form, with the accessories loosely packed within the pouch, one or more tubes, containers, a box or a bag. In some embodiments, a kit comprising iCORE reagents as provided herein further comprises additional components that are useful in performing the iCORE-based assays provided herein. Such additional components may include, but are not limited to deconvolution reagents (e.g., binding agents specifically binding the set of analytes targeted by the iCORE reagents comprised in the kit, for example, immobilized to a solid support, such as beads (e.g., magnetic beads), a membrane, glass slide, resin, or plastic vessel surface), buffers (e.g., wash buffers, binding buffers), and enzymes (e.g., proteases cutting a cleavable linker connecting iCOREs to binding reagents).

EXAMPLES

Example 1: Materials and Methods

[0139] Peptide and Ab peptide synthesis. All peptides were synthesized by Fmoc chemistry in house by the biopolymers facility at the Swanson Biotechnology Center (MIT). Isotope-labeled amino acids were purchased from Cambridge Isotopes. The photosensitive linker (3-Na-Fmoc-Amino-3-(2-nitrophenyl)propionic acid) was purchased from Advanced Chemtech. Detection antibodies against human IFN-gamma and TNF-alpha (eBioscience, clones 4S.B3 and Mab11) were first reacted with SIA (Pierce) at a 50:1 mole ratio. Following removal of excess SIA, peptides were incubated with Abs at a 20:1 mole ratio. Excess peptides were removed by size filtration spin filters (30 k mwco, Amicon).

[0140] Ab-bead synthesis. Tosyl-activated magnetic beads (Myone, Invitrogen) were reacted with anti-human IFN-gamma and anti-human TNF-alpha (eBioscience, clones NIB42 and Mab1 respectively) according to manufacturer's instructions. Briefly, 1 mg of either antibody was buffered

exchange into borate buffer (0.1M Sodium borate, pH 9.0) prior to incubation with the beads in conjugation buffer (0.1M Borate, 1M ammonium sulphate, pH 9.0). After overnight incubation at 37° C., remaining tosyl groups were saturated by incubating the beads in 1M tris. The Ab-labeled beads were stored in PBS prior to experiments.

[0141] Serum assays. Recombinant human IFN-gamma (eBioscience) was spiked into 10% bovine serum in PBS. This solution was then incubated with anti-IFN-gamma-beads for 1 hour at 37° C. The beads were then washed twice with PBS before incubation with an equimolar solution of iCORE-labeled IFN-g and TNF-a for another hour at 37° C. After excess antibodies were removed by washing with 0.1% BSA PBS, the bound antigens and antibodies were eluted in 5% acetic acid. The solution was then exposed to UV light (365 nm) for 30 min. (CL-1000, UVP) to liberate the photocaged iCORE codes into solution. The sample was then analyzed by LC MS/MS.

[0142] Nanomaterial synthesis. Nanoworms were synthesized according to previously published protocols.²⁵ Peptides were synthesized at MIT (Swanson Biotechnology Center); isotopically labeled Fmoc amino acids were purchased from Cambridge Isotopes and 3-Na-Fmoc-Amino-3-(2-nitrophenyl)propionic acid from Advanced Chemtech. Amine-terminated NWs were first reacted with Vivotag 680 (Perkin Elmer) to enable in vivo imaging, before reacted with SIA (Pierce) to introduce sulfhydryl-reactive handles. Cysteine-peptides and PEG-SH were then mixed with NWs overnight at room temperature (95:20:1 molar ratio respectively) and excess peptides were removed by size filtration. Peptide-NW stock solutions were stored in PBS at 4° C.

[0143] In vitro protease assays. For substrate screening, FI-peptide-NWs (2.5 μM by peptide) were mixed with recombinant MMP-2/8/9 (R&D Systems), MMP-7/14 (AnaSpec, Inc.), Thrombin, Tissue Factor, Factor Xa, or Cathepsin B (Haematologic Technologies, Inc.) in a 96-well plate at 37° C. in activity buffers according to manufacturer's instructions and monitored with a microplate reader (SpectroMax Gemini EM). For MS analysis, equimolar iCORE-encoded NWs (Table) were incubated with proteases for 2.5 hrs at 37° C. Cleavage fragments were purified from NWs by size filtration before UV treatment (365 nm, CL-1000 UV crosslinker, UVP). Reporters were then dried by speed vacuum centrifuge and stored at 4° C.

[0144] In vivo imaging. All animal work was approved by the committee on animal care (MIT, protocol #0408-038-11). FVB/NJ mice (Jackson Labs) were fed with 0.1% w/w DDC (Sigma) rodent chow for three weeks (Research Diets). Fibrotic and age control animals were i.v. infused with VivoTag-680 labeled reagents and visualized by IVIS imaging (Xenogen). For tumor xenografts, LS 174T cancer cells lines were maintained in 10% FBS EMEM and inoculated s.c. (5×10⁶/flank) in NU/NU mice (Charles River) prior to imaging.

[0145] Characterization of models. For in situ zymography, fibrotic sections were covered with 90 μl solution of 0.5% w/v low melt agarose (Sigma) in MMP activation buffer (50 mM Tris, 150 NaCl, 5 mM CaCl₂, 0.025% Brij 35, pH 7.5) with 10 μl of DQ-gelation (1 mg/ml, Invitrogen) and Hoechst dye at 37° C. Slides were solidified at 4° C. and then incubated at room temperature overnight to promote gelation proteolysis by tissue proteases. To quantify hepatic collagen, tissue from the right and left lobes (250-300 mg) were hydrolyzed in 5 ml of 6 N HCl at 110° C. for 16 hours

followed by hydroxyproline quantification as previously described.⁴³ To quantify CEA, blood was collected from tumor animals into Capiject microtubes (Terumo) to isolate serum before ELISA analysis (Calbiotech). For immunofluorescence analysis, equimolar NW cocktails (5 μ M/peptide) were administered in Fibrotic FVB/NJ or tumor bearing Nude mice. After perfusion, livers or tumors were fixed in 4% PFA, froze for sectioning, and stained for F4/80 (AbD Serotec), MMP-9 (R&D Biosystems), CD31 (Santa Cruz Biotechnologies) and/or FITC (Genetex) before analyzed by fluorescence microscopy (Nikon Eclipse Ti).

[0146] Collection of urinary peptides. Mice were intravenously infused with 200 μ L of PBS containing equimolar NW cocktails (5 μ M per peptide) with EDTA-free protease inhibitor tablets (Roche) to isolate MMP activity. Mice were placed over 96-well plates surrounded by cylindrical sleeves for urine collection. To prevent further reporter degradation, voided samples were spiked with EDTA+ complete protease inhibitors (Roche) immediately after collection. To quantify urinary fluorescence, 2 μ L of each sample was incubated with magnetic beads (Dyna) coated with α -FITC antibodies (Genetex) in 50 μ L binding buffer (100 mM NH_4OAc , 0.01% CHAPS) for 1 hr at 37° C., washed twice with 100 mM NH_4OAc , and eluted twice with 15 μ L of 5% acetic acid. Samples were neutralized with 2 M Tris and quantified by microplate fluorimetry. For iCORE analysis, samples were irradiated with UV for 30 min. before TCA precipitation (20% final volume) to remove proteins. Soluble fractions were applied to C18 reverse phase columns (Nest Group) and eluted via step gradients of 20% ACN increments in 0.1% formic acid. 60% ACN fractions containing Glu-fib peptides were collected and dried by vacuum centrifuge.

[0147] LC MS/MS analyses. Peptide samples were reconstituted in 5% ACN, 0.1% formic acid and analyzed at MIT or the Taplin MS facility (Harvard Medical School). At MIT, peptides were captured and eluted from a C18 nanoflow HPLC column (75 μ m internal diameter Magic C18 AQ, Michrom BioResources, Inc.) at a flow rate of 300 nL/min using water-acetonitrile solvent system with 0.1% formic acid. ESI mass spectrometry was carried out on a QSTAR Elite QTOF mass spectrometer (AB Sciex). At Harvard, samples are reconstituted in 2.5% ACN, 0.1% formic acid. Samples are injected using a Famos autosampler (LC Packings) into an Agilent 1100 HPLC prior to mass analysis on a LTQ-Orbitrap (Thermo Electron). To account for discrepancies in urine volumes and concentration, peak intensities of individual reporters were scaled relative to its respective total iCORE ion current before normalization against control samples to account for technical and age-related variations.

[0148] Statistical Analyses. Pearson's correlation coefficients between different protease profiles were calculated with MatLAB. ANOVA analyses were calculated with GraphPad 5.0 (Prism). For ROC analysis, risk score functions were first estimated by logistic regression on individual biomarkers followed by ROC curve analyses of single or biomarker combinations (SigmaPlot).

Example 2: Detection of TNF-alpha and IFN-gamma in a Liquid Sample

[0149] A liquid sample comprising of recombinant human IFN-gamma spiked into 10% bovine serum to simulate a complex biological sample was made. As shown in FIG. 1a, the sample was first deconvoluted by contacting the sample with antibodies conjugated to the surface of beads. Specific

binding to anti-human-TNF-alpha and anti-human-IFN-gamma antibodies bound to the surface of the beads resulted in immobilization of the target analyte IFN-gamma on the bead surface. The beads were then separated from the liquid sample by a magnet to remove the liquid phase and any unbound materials comprised in any sample residue remaining on the beads were washed away from the beads by washing the beads with a wash buffer. The immobilized analytes bound to the binding agents on the bead surface were then contacted with another set of antibodies against TNF- α and IFN-g that were conjugated to a set of iCOREs via a photocleavable linker (dE-⁺³G-⁺⁶V-dN-dD-dN-dE-dE-G-F-F-dS-A-dR-(ANP)-K(FAM)-G-G-C (SEQ ID NO: 37), dE-⁺²G-⁺⁶V-dN-dD-dN-dE-dE-G-F-F-dS-⁺¹A-dR-(ANP)-K(FAM)-G-G-C (SEQ ID NO: 37), respectively).

[0150] In this experiment, the set comprised two antibodies, one specifically binding IFN-gamma and the other specifically binding TNF-alpha. Both antibodies were conjugated to isobaric iCOREs, with the iCOREs bound to the antibodies binding IFN-gamma having a different fragmentation signature than the iCOREs bound to the antibodies binding TNF-alpha. Any unbound antibodies were then removed from the beads by washing the beads with wash buffer. The beads were then incubated in a stripping solution to elute the bound material, and the resulting solution collected following removal of beads by a magnet. The photocleavable linkers were then cleaved by exposure to UV light. The iCOREs rescued in this manner were then subjected to an LC MS/MS assay. FIG. 1b shows the resulting peaks of the second round MS (after fragmentation). The target analyte IFN-gamma was detected in the sample while endogenous levels of TNF-alpha, or non-specific binding contributed to the first peak. Because the iCOREs were isobaric, the signal intensity of each iCORE can directly be translated into a relative abundance of the target analyte. Accordingly, IFN-gamma was far more abundant than TNF-alpha in the sample analyzed, representing a robust signal to noise of approximately 10.

Example 3: Utilization of a Canonical Peak from Glu-Fib MS/MS Spectra as a Reporter

[0151] Glu-fib was evaluated for its usefulness as a base sequence for the generation of an isobaric iCORE library. Glu-fib is a 13 amino acid peptide of the sequence EGVND-NEEGFFSAR (SEQ ID NO: 1). The fragmentation spectrum resulting from collision induced disassociation results in the generation of y-type ions. FIG. 2 is a typical spectrum to illustrate the resulting y-type ions (y1-y13) and their relative intensities.

[0152] The propensity of glu-fib to fragment along certain peptide bonds was exploited to construct 10 mass codes centered on the y₇ 'reporter' ion (EGFFSAR (SEQ ID NO: 6)) by enriching the truncated sequence with heavy amino acids to produce variants differentiated by 1 Da each. This introduced mass shift is then compensated by isotope enrichment within the 'balance' fragment (EGVNDNE (SEQ ID NO: 7)). As a result, each peptide is characterized by an identical nominal mass and a distinct y₇ fragment ion. Thus this iCORE set is isobaric, with a shared parent mass of 1579.68 Da. The mass of the b₇ fragment (balance) was calculated for each balance as $M_{b7(i)} = (M_{b7} + n) - i$ and the mass of each y₇-fragment as $M_{y7(i)} = (M_{y7} + i)$, with $M_{b7} = 767.3$ and $M_{y7} = 812.38$. The calculated masses of four iCOREs of the set are given in the table on the lower right

of FIG. 2. The set of isobaric iCOREs was then subjected to an MS/MS assay. The data obtained is shown in the MS/MS graph in FIG. 3.

Example 4: MS/MS Data from Isobaric iCORE Libraries

[0153] A 10-plex library of glu-fib iCOREs was subjected to an MS/MS assay. The results are depicted in FIG. 3. The upper panel of the figure provides the sequences and fragmentation signatures of the iCOREs in the set (iCOREs G1, G2, G3, G4, G5, G6, G7, G8, G9, and G10). The upper panel also shows the extracted ion chromatograph obtained from the 10-plex iCORE library (G1-G10) after the first round of MS. Only one peak was observed ($m/z=789.7-790.0$).

[0154] Accordingly, each iCORE is mass indistinguishable from the multiplexed set as well as undifferentiable by chromatography, which, in some embodiments, is beneficial for quantification, which is consistent with the calculations in Example 3. The middle panel of FIG. 3 shows MS/MS results following fragmentation of the peak observed in the upper panel. The results demonstrate that, also consistent with the observations of Example 3, y -ions y_3 - y_{11} could be identified. Importantly, magnification of the m/z window around the predicted m/z values of the y_7 fragmentation ions generated revealed individual, distinguishable peaks corresponding to individual iCORE fragmentation signatures (lower panel). Relative quantification of the signals obtained for G1-G10 revealed that the relative representation of the iCOREs in the original sample (equal amounts of each iCORE) was closely mirrored in the peak pattern obtained, indicating that the method is not only fit for qualitative detection (absence/presence of an analyte), but also for quantitative and comparative quantitative studies.

Example 5: Quantitative iCORE MS/MS Assays

[0155] In order to investigate whether relative quantities of iCORE abundances in a sample are accurately reflected in MS/MS assays such libraries are subjected to, a sample comprising a 10-plex iCORE library (G1-G10), as described in Example 4, was created, in which the different iCOREs were comprised at different relative abundances in defined ratios. The original sample comprised the following iCORE stoichiometry: G1:G2:G3:G4:G5:G6:G7:G8:G9:G10 were at a ratio of 1:2:3:4:5:10:10:5:3:2:1. The sample was subjected to an MS/MS analysis and the results are shown in FIG. 4. The bar graph shows a good correlation of the original stoichiometry and the measured abundances of the iCOREs. Further analysis of expected (theoretical) and observed (extracted) abundances of each iCORE demonstrate that the MS/MS results closely resemble the actual abundances in the sample. The correlation coefficient of the observed and the expected data set was calculated to be $R^2=0.99$. These results indicate that multiplex iCORE assays are suitable for quantitative analysis and that iCORE abundance is maintained and accurately reflected after multiplex MS/MS assays.

Example 6: Linear Dynamic Range of iCORE MS/MS Assays

[0156] In order to investigate the dynamic range in which iCOREs could be reliably detected and quantified, a dilution series of a 10-plex iCORE library, as described in Example 4, was generated. The library comprise each iCORE (G1-

G10) at the same abundance. Five diluted samples, with a concentration of $10^{-7}M$ (100 nM), $10^{-8}M$ (10 nM), $10^{-9}M$ (1 nM), $10^{-10}M$ (100 pM), and $10^{-11}M$ (10 pM), respectively, were then subjected to an MS/MS assay. The observed intensities for each iCORE at each concentration were plotted in relation to the concentration of the iCOREs in the dilution sample. FIG. 5 shows that all iCOREs were accurately detected at all concentrations, indicating that sensitivity of iCORE MS/MS assays and the linear dynamic range, in which actual iCORE abundances in the sample are accurately reflected in the observed signal intensity, extends to at least from the picomolar range to the nanomolar range, spanning at least five orders of magnitude. These results indicate that accurate detection of iCOREs in a multiplex MS/MS assay is possible at high sensitivity, with even picomolar levels of iCOREs being detectable at an accuracy allowing for quantitative analysis.

Example 7: Efficiency of iCORE Release from Covalently Bound Substrate Via Photocleavage

[0157] In order to evaluate the efficiency of release of iCOREs covalently bound to a substrate, for example, a binding agent specifically binding a target analyte, a glu-fib-based iCORE was generated and conjugated to a substrate via a covalent bond. The iCORE was connected to the substrate, another peptide via a photocleavable linker, thus generating a photo-caged, or photocleavable, iCORE-substrate tandem peptide. The photo-caged tandem peptide was exposed to UV light of ~ 350 nm wave length for a time (30 min) sufficient for photocleavage of the linker, which resulted in a release of the iCORE from the tandem peptide as shown in the upper panel of FIG. 6. The process of photocleavage was monitored by LC MS (lower panel). The data obtained demonstrate that the photo-caged tandem peptide (I, $m/z=881.7$) was quantitatively cleaved and close to 100% of the theoretically recoverable amount of the released iCORE (II, $m/z=785.4$) was rescued. These results indicate that photocleavage of iCOREs covalently bound to other molecules, e.g., other peptides, such as binding agents, or hydrogel-forming molecules, via a photocleavable linker, can be efficiently released by exposure to UV, and can be quantitatively rescued and accurately quantified in LC MS assays.

Example 8: Labeling of Molecules with iCOREs by Conjugation to Reactive Moieties

[0158] In order to investigate whether iCOREs could be used to covalently label other molecules, glu-fib iCOREs are conjugated to a reactive chemical group able to form a covalent bond to a compatible chemical moiety under certain conditions. For this, a glu-fib iCORE is attached to a reactive chemical moiety able to form a covalent bond to a PEG monomer when exposed to UV light in the presence of PEG monomers. As depicted in FIG. 7, UV exposure results in the formation of a glu-fib iCORE-labeled PEG polymer, in this case a hydrogel, in which the PEG scaffold is covalently bound by iCOREs.

[0159] This proof-of principle experiment demonstrates that substrates can be covalently labeled with iCOREs and that iCORE-labeled hydrogels, useful, for example, as scaffolds for tissue engineering, can readily be achieved.

Example 9: Mass-Encoded Synthetic Biomarkers
for Multiplexed Urinary Monitoring of Disease

[0160] A protease sensing platform comprising multiplexed synthetic biomarkers for the detection of disease was developed. FIG. 8 shows a schematic of this approach. FIG. 8a shows a synthetic biomarker library comprised of mass encoded substrate peptide library conjugated onto nano-worm nanoparticles. Administration of NW cocktails in whole animals leads to accumulation in disease tissues (b). Local proteases cleave peptide fragments that subsequently filter into urine. Photo-caged mass reporters are released upon exposure to UV-light and quantified by liquid chromatography tandem mass spectrometry (c).

[0161] To establish a protease sensing platform, peptide substrates were identified that would be susceptible to

velocities were extracted for each protease-substrate combination and compiled for comparative analysis (FIG. 10a, b). FIG. 10 illustrates selecting protease-sensitive substrates for NW-chaperoned urinary trafficking. Kinetics of peptide-NW cleavage monitored by fluorimetry are shown in (a). Specific protease-substrate combinations led to rapid activation. FIG. 10b shows a heat map comparison of initial cleavage velocities for different substrate-protease combinations grouped according to activity and specificity. IVIS in vivo imaging of DDC treated and control animals following intravenous injection of VivoTag-680 labeled Glu-fib peptides, (d) peptide-free NWs, or (e) peptide-conjugated NWs.

[0163] From this set, 10 peptides with robust protease susceptibility were selected to serve as a peptide-NW library (Table 1).

TABLE 1

iCORE-encoded peptide biomarker library. Individual probe substrates encoded with photo-caged isobaric mass codes for quantification by LC MS/MS.					
	Synthetic Biomarker Library ^②	Substrate	Isobaric mass code ^②	y ^② reporter	[y ^② +H ^②]
G1	e ^② c ^② VndneeGFfsAr-X-K(FAM)	PQGIWGQ	e ^② c ^② VndneeGFfsAr	GPSAE	683.4
G2	e ^② c ^② Vndnee ^② GFfsAr-X-K(FAM)	LVPRGSG	e ^② c ^② Vndnee ^② GFfsAr	^② GFIsA	884.4
G3	e ^② c ^② Vndnee ^② GFfsAr-X-K(FAM)	PVGLIG	e ^② c ^② Vndnee ^② GFfsAr	^② GFsAr	685.4
G4	e ^② c ^② Vndnee ^② GFfs ^② Ar-X-K(FAM)	PWGIWGQG	e ^② c ^② Vndnee ^② GFfs ^② Ar	^② GFfs ^② Ar	686.4
G5	e ^② c ^② VndneeGFfs ^② Ar-X-K(FAM)	PVPLSLVM	e ^② c ^② VndneeGFfs ^② Ar	GFfs ^② Ar	687.4
G6	e ^② c ^② Vndnee ^② GFfs ^② Ar-X-K(FAM)	PLGLRSW	e ^② c ^② Vndnee ^② GFfs ^② Ar	^② GFfs ^② Ar	688.4
G7	e ^② GVndneec ^② FfsAr-X-K(FAM)	PLGVRGK	e ^② GVndneec ^② FfsAr	c ^② FfsAr	889.4
G8	e ^② GVndneec ^② Ffs ^② Ar-X-K(FAM)	f(Pip)RSGGG	e ^② GVndneec ^② Ffs ^② Ar	c ^② Ffs ^② Ar	690.4
G9	e ^② GVndnee ^② c ^② FfsAr-X-K(FAM)	fPRSGGG	e ^② GVndnee ^② c ^② FfsAr	^② c ^② FfsAr	691.4
G10	eGVndnee ^② c ^② FfsAr-X-K(FAM)	f(Pip)KSGGG	eGVndnee ^② c ^② FfsAr	^② c ^② FfsAr	692.4

②

② indicates text missing or illegible when filed

cleavage by extracellular proteases associated with liver fibrosis and cancer. Fluorescein-labeled derivatives of ~50 candidate peptide sequences²⁰⁻²⁴ were synthesized and conjugated to PEG-coated, long-circulating iron oxide nano-worm²⁵ (NW) nanoparticles (FIG. 9a,b,c) and incubated with recombinant proteases commonly overexpressed in these diseases (e.g. Matrixmetalloproteases (MMPs), cathepsins) as well as proteases normally present in blood (FXa, Tissue factor (TF), thrombin) to evaluate substrate specificity and accessibility of surface-conjugated peptides to proteases. FIG. 9 shows long-circulating iron oxide nanoworm chaperones. NW size distribution as determined by dynamic light scattering of dextran-coated, iron oxide nanoworms is shown in (a). Absorbance spectra of free NWs (red) and NWs conjugated with Fluorescein-labeled peptides (~500 nm) and Alexa 647 (blue) are shown in (b). In vivo clearance kinetics of PEG-coated (blue) and PEG-free (red) NWs conjugated with substrate peptides are shown in (c). Pegylated NWs extended circulation half-life by ~4 hours.

[0162] Using this assay, increases in sample fluorescence were observed upon proteolytic cleavage of the tethered fluorescent peptides resulting from abrogation of homo-quenching between adjacent fluorophores. Initial reaction

[0164] Sequences of G1-G10 correspond, from top to bottom, to SEQ ID NO: 8-SEQ ID NO: 17, respectively. Substrate sequences correspond, from top to bottom, to SEQ ID NO: 18-SEQ ID NO: 27, respectively. The isobaric mass codes correspond to SEQ ID NO: 28, and the y6 reporters correspond to SEQ ID NO: 29.

[0165] In order to design a system that would probe diseased microenvironments, the biodistribution of each system component (i.e. peptide and nanoparticle) was investigated in disease models in vivo. Here, a mouse model of liver fibrosis was selected in which FVB/NJ mice fed with 3,5-diethoxycarbonyl-1,4-dihydrocollidine (DDC) develop progressive liver disease as a result of chronic bile duct injury²⁶ (FIG. 11a,b,c), leading to fibrotic and proteolytically active livers (FIG. 12b,c,d).

[0166] FIG. 11 illustrates urinary biomarkers of hepatic fibrosis and resolution in DDC-treated mice. FIG. 11a shows the induction of liver fibrosis and NW administration timeline. FIG. 11b shows quantification of total liver collagen by hydroxyproline analysis. DDC treatment led to ~3 fold increase in liver collagen by week 3 (**P<0.001) and a ~30% drop between week 7 and 11 (*P<0.05) that remained above pre-treatment values (*P<0.05) (One-way ANOVA,

Tukey post test, $n=3$, s.e.m.). Sirius red histochemistry of liver sections (scale bar=50 μm) indicated the presence of fibrotic extensions emanating from portal triads at week 3, persisting to week 7, and resolving by 11 weeks (c). IVIS in vivo imaging showed urinary accumulation of ensemble reporter library in DDC-treated animals (d). Kinetics of urinary accumulation quantified by α -FITC immunoprecipitation (3 wk timepoint, $*P<0.05$, Two-way ANOVA, Tukey post test; error bars=s.e.m.) are shown in (e). FIG. 11f shows box and whisker plots of individual iCORE peak intensities plotted as DDC over control at 0, 3, 7, and 11 weeks ($*P<0.05$, $**P<0.01$; repeated measures ANOVA, Tukey post test; $n=10$). FIG. 11g shows ROC curves of single, double and triple combinations of fibrosing biomarkers with associated area under curves (AUC). (h) ROC curves of single, double and triple combinations of resolving biomarkers with associated area under curves (AUC).

[0167] FIG. 12 illustrates immunofluorescence of liver sections. FIG. 12a shows periportal images of macrophage (red) and NW (green) infiltrate into fibrosing zones. Arrows highlight areas of colocalization. FIG. 12b shows periportal images of MMP9 (red) and NW (green). MMP9 expression was not found in healthy livers (c). FIG. 12d shows Fluorescence image of periportal zone after gelatin in situ zymography. Scale bars=100 μm .

[0168] To evaluate peptide trafficking, the peptide glutamatefibrinopeptide B (Glu-fib, EGVNDNEEGFFSAR, SEQ ID NO: 1) was selected as a prototypic urinary marker because its endogenous derivative (fibrinopeptide B) is biologically inert and normally filters freely into urine when released during coagulation.²⁷ Fluorophore-labeled Glu-fib administered intravenously (i.v.) efficiently filtered into urine in both fibrosing and healthy animals with undetectable liver homing (FIG. 10c). In contrast, larger peptide-free NWs (~40 nm hydrodynamic radius, FIG. 9a) homed to the liver (FIG. 10d), consistent with the size-dependent trafficking of nanomaterials and a renal clearance threshold of ~5 nm for inorganic nanoparticles.²⁸ Importantly, in animals treated with fluorophore-labeled peptides conjugated to NWs, significant NW-mediated peptide homing was observed to both fibrosing and healthy livers leading to a strong increase in urinary fluorescence in diseased animals resulting from renal filtration of cleaved peptide fragments (FIG. 10e).

[0169] Profiling Protease Activities by Mass Spectrometry

[0170] Despite the multiplexing advantages of mass-encoding, one challenge of detecting protease activity by MS is that peptide substrates in complex proteolytic environments can be cleaved at multiple sites by promiscuous proteases and truncated by exoproteases^{4, 29} to produce diverse pools of poorly defined fragments that confound mass analysis. Here, well-defined mass reporters to encode a substrate library were designed. In light of the favorable renal clearance properties of Glu-fib, d-isomer rich derivatives of Glu-fib were appended to the N-termini of each protease substrate to serve as protease resistant mass reporters and to promote renal filtration upon substrate cleavage and release from NWs (Table 1). These tandem peptides were further modified with internal photo-labile residues³⁰ to enable the recovery of well-defined Glufib peptides by photolysis from complex urinary cleavage fragments following in vivo proteolysis. To test this construct, a model photo-caged tandem peptide was synthesized (compound I, FIG. 13a). Consistent with previously published reports on

nitrophenyl groups, exposure of compound I (triply charged, 881.7 m/z; FIG. 13b, top panel) to UV light triggered peptide cleavage, resulting in the appearance of doubly charged, acetamide-terminated Glu-fib (785.4 m/z; FIG. 13b, bottom panel).

[0171] FIG. 13 shows photo-caged iCORE libraries for multiplexed profiling of protease activities by LC MS/MS. FIG. 13a illustrates the structure of tandem peptide (compound I) containing an internal UV-sensitive linker. Shown here is the structure of free Glu-fib (compound II) generated after photolysis (~350 nm). FIG. 13b shows LC MS spectra of compound I before (top, triply-charged m/z: 881.7) and after (bottom, doubly-charged m/z: 785.4) exposure to UV light. FIG. 13c shows a 10-plex isobaric peptide library derived from Glu-fib. FIG. 13d shows an extracted ion chromatogram of an equimolar 10-plex iCORE mixture (789.80-789.90 m/z). The entire multiplexed set was chromatographically indistinguishable. FIG. 13e shows an iCORE MS/MS spectrum following collision induced dissociation. Individual reporters were identified via unique y6 reporter ions (683.3-692.3 m/z) each differentiable by a single mass unit. FIG. 13f shows an iCORE MS/MS spectrum following incubation of a 10-plex, iCORE-encoded peptide-NW cocktail with recombinant MMP9.

[0172] FIG. 14 illustrates Isobaric COded REporter (iCORE) mass encoding. FIG. 14a shows an MS/MS spectrum following collision induced disassociation of Glu-fib. Peaks correspond to c-terminal, y-type peptide fragments. FIG. 14b lists 10 isotopic analogs and corresponding masses produced via isobaric encoding. Each sequence was constructed by selectively enriching the balance or reporter regions with 'heavy' amino acids to produce unique y6 reporter ions while maintaining a uniform total mass.

[0173] In order to design an extensible encoding strategy for the library of protease substrates, the ability of isobaric mass encoding^{31, 32} was investigated to be extended to peptide scaffolds such as the urinary reporter Glu-fib to produce a family of mass reporters. The distinguishing feature of an isobaric encoding strategy is that individual members of a family of reporters share a parent mass to facilitate efficient peptide collection by MS, but can be subsequently identified via unique MS/MS ions upon fragmentation. It was first determined that Glu-fib fragments into C-terminal y-type ions (FIG. 10a) and 10 mass codes were constructed centered on the y6 ion (GFFSAR, SEQ ID NO: 4) by enriching the hexamer with heavy amino acids to produce variants differentiated by 1 Da each (FIG. 10b). This introduced mass shift was then balanced by isotope enrichment within the remainder of the peptide (EGVNDNEE, SEQ ID NO: 5). As a result, each peptide was characterized by an identical nominal mass and a distinct y6 fragment ion. This encoding method was termed "isobar COded REporters" (iCORE). To validate this approach, an equimolar 10-plex iCORE mixture (FIG. 13c) was analyzed by LC MS/MS and the entire peptide library was found to initially appear as a single, unresolved peak (extracted ion chromatogram, 789.95 ± 0.5 m/z, FIG. 13d) but following fragmentation, to resolve into a 10 peak spectrum with no fragmentation bias (683.4-692.4 m/z, FIG. 15, FIG. 13e). To remove confounding peak overlap arising from naturally occurring isotopes (e.g. 13C), iCORE peptides were selectively fragmented via a unit mass window centered on the precursor ion (FIG. 16a) to minimize the signal from naturally occurring isotopes to ~5% of the parent peak (FIG.

16b). Thus, in samples spiked with reporters at defined ratios (1:2:3:5:10:10:5:3:2:1), a linear correlation was observed between peak intensity and stoichiometry in both unmodified and peak-subtracted analysis ($n=3$, $R^2=0.99$ and $R^2=0.99$ respectively, FIG. 17a,b,c). FIG. 17 demonstrates that iCORE LC MS/MS analysis is quantitative. FIG. 17a shows an MS/MS spectrum of a 10-plex mixture of iCORE reporters combined in a 1:2:3:5:10:10:5:3:2:1 ratio. Extracted peak intensities were highly correlated with the input reporter stoichiometry ($r^2=0.99$; $n=3$; error bars=s.e.m.) (b). To account for naturally occurring isotopes (FIG. 16), individual reporter intensities were modified by subtracting 5% of the prior reporter intensity (c). Compensated reporter ratios also strongly correlated with reporter ratios ($r^2=0.99$; $n=3$; error bars=s.e.m.). All subsequent samples were peak-adjusted to reflect contributions from naturally occurring isotopes.

[0174] To evaluate the ability of iCORE multiplexing to simultaneously report on the activities of many protease-substrate combinations, a 10-plex, equimolar cocktail of iCORE encoded peptide-NWs was treated with recombinant MMP9 (Table 1). Following incubation, cleavage products were isolated by size filtration and exposed to UV-light prior to MS/MS analysis.

[0175] Notably, collective substrate activities were translated into distinct iCORE landscapes characterized by markedly different y6 reporter intensities (FIG. 13F). This library was applied to several other recombinant proteases (FIG. 18a) and no strong correlation was found between different iCORE protease profiles (i.e. MMP2, MMP9, MMP12, and thrombin) as determined from Pearson's correlation analysis (FIG. 18b), illustrating the ability of iCORE-encoded NWs to monitor many protease and protease-substrate combinations uniquely.

Example 10: Monitoring Hepatic Fibrogenesis and Resolution

[0176] Liver fibrosis is a wound healing response to chronic liver injury and results in the buildup of scar tissue that can lead to cirrhosis, liver failure and cancer.¹⁸ The dynamics of extracellular matrix accumulation such as collagen is largely driven by fibrogenic hepatic stellate cells and myfibroblasts, and matrix remodeling proteases such as MMPs and their inhibitors. The current gold standard for monitoring is a needle biopsy followed by histological analysis; however, this technique is invasive, confounded by high sampling heterogeneity, carries a finite risk of complications and cannot be performed frequently as needed (e.g. monitoring antifibrotic therapies).³³ Noninvasive assays including ultrasound imaging, elastography, and serum biomarkers are limited by their low accuracies and limited prognostic utility.³⁴ Thus, there remains an urgent need for noninvasive biomarkers to replace biopsy-based monitoring, identify and validate new antifibrotic agents, and to support clinical decision making.³⁵ Here, it was sought to identify synthetic biomarkers with the capacity to monitor liver fibrosis, and to extend the DDC-induced model to include both fibrosing and resolving disease (FIG. 11a,b,c).

[0177] The in vivo trafficking and ability of iCORE-encoded synthetic biomarkers to produce an ensemble urinary signal during fibrotic progression was determined. Intravenous administration of a fluorophore-labeled, 10-plex peptide-NW cocktail resulted in a 2-fold increase in urinary fluorescence in animals treated with DDC for 3 weeks

(* $P<0.05$, Twoway ANOVA; FIG. 11d,e). Immunofluorescence imaging of liver sections revealed that most NWs escaped sequestration by resident macrophages (FIG. 12a, arrow), infiltrating freely into the parenchyma and into actively fibrosing periportal zones characterized by punctate MMP9 expression (FIG. 12b,c). Similar patterns appeared in sections treated with DQ-gelatin substrates (FIG. 12d) which permit in situ visualization of active collagen-degrading proteases.

[0178] Collectively, these results indicated that NWs home to fibrosing liver microenvironments containing active proteases. In order to determine the behavior of individual reporters to identify sensitive and specific synthetic biomarkers, the processes of fibrosis and resolution were probed by iCORE mass analysis. Mice treated transiently with DDC for 3 weeks followed by restoration of DDC-free chow develop distinct fibrosing and resolving windows (0-3 and 7-11 weeks respectively, FIG. 11a) as verified macroscopically by Sirius red collagen staining of liver sections (FIG. 11c), and by hydroxyproline analysis which quantifies total tissue collagen (FIG. 11b). With this treatment regime, liver collagen increased ~3-fold compared to pretreatment levels after 3 weeks on DDC, persisted from week 3-7 after initial removal of DDC, and significantly decreased from week 7-11 after sustained DDC withdrawal (* $P<0.05$, *** $P<0.005$, $n=3$).

[0179] Thus, to monitor the transitions between fibrosing and resolving disease, NWs were administered at 0, 3, 7, and 11 weeks into DDC-treated and age-matched control animals followed by iCORE MS/MS analysis. The resulting activities of individual biomarkers displayed markedly divergent kinetics (FIG. 11f). Biomarkers G3 and G4 both strongly increased relative to pretreatment baselines, reaching a plateau by week 11 despite staggered onset at week 7 and 3 respectively. G5 and G6 showed opposing kinetics, significantly decreasing at week 3 before either gradually returning to pretreatment intensities (G5) or persisting to week 11 (G6). Interestingly, G7 tracked with the kinetics of DDC treatment, elevating sharply at week 3 followed by a rapid reversal at week 7. All remaining biomarkers (G1, G2, G8, G9 and G10) did not deviate from initial pretreatment activities (G1-G10; * $P<0.05$, ** $P<0.01$, repeated measures ANOVA, Tukey post test, $n=10$). Importantly, all biomarkers in control animals also did not significantly depart from baseline (FIG. 19).

[0180] The diagnostic performance of these biomarkers was determined by performing receiver operating characteristic (ROC) analyses for individual as well as biomarker combinations. ROC curves characterize the sensitivity and specificity of a biomarker as a function of the discrimination threshold by returning the area under the curve (AUC) as a performance metric with a baseline AUC of 0.5 representing a random biomarker classifier (dashed line, FIG. 11g,h). Within the fibrosing window of 0-3 weeks, G5 displayed the highest AUC amongst the 10 biomarkers (0.96, FIG. 20), which was further improved by including G7 in a double biomarker combination (0.98), or G6 and G7 in a triple combination (1.00), resulting in a perfect synthetic biomarker classifier for fibrosing disease in this model (FIG. 11g). Within the resolving time frame of 7-11 weeks, G1 led with AUC=0.73 (FIG. 21), which was significantly improved via the dual combination of G1+G9 (AUC=0.9) and finally the triple combination G1+G7+G9 (AUC=0.91) (FIG. 11h).

[0181] Collectively, these experiments demonstrate that liver fibrosis and resolution are revealed by distinct collections of synthetic biomarkers, and that multiplexed combinations allow the highest diagnostic performance—illustrating the ability of this platform to noninvasively illuminate otherwise inaccessible aspects of liver disease evolution.

Example 11: Early Detection of Colorectal Tumors

[0182] When diagnosed prior to systemic dissemination, many primary tumors can be effectively treated with conventional clinical interventions.³⁶ However, most clinically-utilized biomarkers lack the diagnostic accuracy and sensitivity to discriminate small tumors, and it remains unclear whether current endogenous blood biomarker strategies targeting shed or processed byproducts of cancerous tissue can be sufficiently improved for early detection.¹⁹ Moreover, it is becoming apparent that individual serum biomarkers (e.g. CA-125 for ovarian, PSA for prostate cancer) do not possess the necessary sensitivity and specificity required for early detection, and that panels of biomarkers are most likely required.³⁷

[0183] Here, it was investigated whether panels of synthetic urinary biomarkers could be readily adapted to allow earlier and more accurate detection of cancer compared with single clinical blood biomarkers, e.g., because nanoparticles can be passively targeted to tumors to sample proteases through fenestrated angiogenic tumor vessels.¹⁶ In order to explore this concept, athymic nude mice were used bearing LS174T xenograft tumors, a human colorectal cancer (CRC) cell line³⁸ that secretes the blood biomarker carcinoembryonic antigen (CEA) as a model system (FIG. 22). FIG. 23a shows a timeline of LS 174T colorectal cancer cell inoculation and NW administration in Nude mice. FIG. 23b illustrates macroscopic quantification of tumor growth (n=5, s.d.), and FIG. 23c shows circulating levels of CEA in tumor and control animals analyzed every third day post tumor implantation by ELISA (**P<0.01, Two-way ANOVA, Tukey post test). IVIS in vivo imaging showed urinary accumulation of ensemble reporter library in tumor-bearing animals (d). FIG. 23e displays quantification of urinary fluorescence by FITC-immunoprecipitation (*P<0.05, Two-way ANOVA, Tukey post test; n=5, s.d.). FIG. 23f shows ROC curves of a single, double and triple biomarker combination with associated AUC (n=16; 8 ctrl, 8 tumor), and FIG. 23g shows ROC comparison between triple biomarker combination (G1+G2+G3) with serum CEA at day 10. Following implantation, tumor engraftment was monitored by sampling serum every 3 days and analyzing for CEA by ELISA (FIG. 23b,c). At day 10, CEA levels were insufficiently elevated to distinguish tumor from control animals, corresponding to an average tumor burden of ~130 mm³ (spherical diameter d ~6.3 mm). Tumors permitted to grow further were readily detected by CEA analysis (day 13 and 16), representing a limit of detection of ~330 mm³ (d ~8.6 mm) (**P<0.01, Two way ANOVA, n=5). In order to determine whether synthetic urinary biomarkers could outperform serum CEA, the fluorophore-labeled, 10-plex peptide-NW library described herein was administered at day 0 and 10. Ex vivo imaging of excised tumors and immunofluorescence analysis of corresponding sections indicated that NWs readily homed to the extravascular milieu following i.v. administration (FIG. 24). Despite the inability of CEA to indicate the presence of cancer at day 10, ensemble peptide cleavage resulted in a strong 2-fold elevation in

urine fluorescence (FIG. 23d,e), allowing the detection of tumors ~2.5-times smaller than CEA (130 vs. 330 mm³) by urinary fluorescence alone. To determine the full diagnostic potential of multiplexed analysis, individual biomarkers were quantified by iCORE MS/MS. Similar to the liver studies described herein, the classifying power of the highest performing single biomarker (G1, AUC=0.78, FIG. 25) steadily improved in double (G1+G2, AUC=0.88) and triple biomarker combinations (G1+G2+G3, AUC=0.89) (n=16, FIG. 230. This triple combination significantly outperformed serum CEA, which detected disease poorly with an AUC of 0.61 at day 10 (FIG. 23g).

Discussion

[0184] Despite the fact that many new features of disease are rapidly being discovered with global profiling approaches, the development of this knowledge into next-generation biomarkers remains fundamentally limited by many technical and biological challenges intrinsic to endogenous targets. Ideally, a candidate biomarker would be secreted at high levels relative to native background, remain stable or persistent in circulation until detection, be readily accessible from compositionally simple host fluids, and be able to discriminate disease with high sensitivity and specificity. In practice, such parameters are difficult to improve or control, and many promising biomarkers fail during rigorous evaluation for clinical translation.

[0185] Here, a system of synthetic biomarkers was devised with the capacity to (i) amplify biomarker levels through substrate turnover by targeting aberrant protease activities, (ii) release stable, d-isomer enriched mass reporters designed to be detected within a narrow mass window free of host molecules, (iii) trigger reporter clearance from blood into urine to reduce matrix complexity and to facilitate facile extraction, and (iv) simultaneously profile libraries of candidate synthetic biomarkers in vivo to identify multiplexed combinations for highly sensitive and specific diagnosis. The work described herein shows that by engineering exogenous agents to interrogate diseased tissues, key biological and transport challenges can be separately addressed prior to integration, resulting in synthetic biomarkers that can be rapidly designed, tested and identified for distinct diseases.

[0186] The liver studies described herein demonstrate the potential of this technology for monitoring both liver fibrosis and resolution. Currently, the needle biopsy remains the gold standard despite the fact that biopsy results are often variable because only a small part of the liver is sampled and can lead to inaccurate diagnosis or repeat biopsies. Furthermore, gross architectural changes frequently occur at a longer time scale compared with alterations in protease activities as occurs during resolution, making it challenging to predict patient trajectory based on histological analysis. Notably, this study showed that nanoscale agents penetrate throughout the liver, delineated a panel of fibrosing and resolving biomarkers (0-3 and 7-11 weeks respectively), and even revealed biomarkers between week 3 and 7 (e.g. G3, G7) that could potentially represent an anticipatory signature in advance of resolution since liver sections within this time window were indistinguishable by clinically-utilized histological or matrix quantification assays. Early predictive biomarkers of fibrotic resolution would accelerate the rate in which new anti-fibrotic drugs are identified in preclinical

and clinical studies by quickly differentiating potential responders from non-responders with identical histology.

[0187] Currently, little clinical data exists correlating the size of tumors to circulating biomarker levels and it remains unclear whether small tumors ($<1\text{ cm}^3$) can be reliably detected. Recently, Gambhir and colleagues¹⁹ have estimated via mathematical modeling that solid tumors could potentially remain undetectable for 10-12 years and reach a spherical diameter $>2.5\text{ cm}$ before blood biomarkers could indicate disease. In order to be useful for early detection, current approaches dependent on endogenous species will need to be greatly improved. In comparison, this study showed that multiplexed synthetic biomarker panels can detect CRC tumors earlier and more precisely than serum CEA, enabling the detection of tumor burdens ~ 2.5 -fold smaller. It further illustrated that relatively small diseased sites ($\sim 130\text{ mm}^3$) outside of the liver are accessible for interrogation, opening the possibility of detecting disease at sites currently challenging to probe with near-infrared imaging (e.g. deep seeded tumors) or with other imaging modalities (e.g. diseased tissues with poor MRI or CT contrast such as liver metastases).³⁹

[0188] The technology described herein can be readily extended to encompass additional diseases and alternative applications. Given the cumulative wealth of approved and experimental nanoparticle formulations, the work described herein is transferrable to other nanomaterials and scaffolds that can actively target or transport peptide cargo to different organs, types of vasculature, and tissue depths.^{40, 41} Further, the iCORE mass-encoding scheme disclosed herein can be extended to create hundreds of orthogonal codes by incorporating additional isotope-enriched amino acids and by making use of distinct parent peptides. A larger encoding library would not only enable the simultaneous monitoring of hundreds of synthetic biomarkers, but could additionally allow the identities of dysregulated proteases to be revealed by mathematical deconvolution⁴² to overcome the challenges of designing specific substrates for promiscuous proteases. A collection of diverse delivery strategies, broader multiplexing capabilities, and precise mathematical algorithms would provide rich opportunities for systems-level monitoring of disease and clearly elucidate the roles that complex protease networks play in health and disease.

REFERENCES

- [0189] [1] Sawyers, C. L. The cancer biomarker problem. *Nature* 452, 548-552 (2008).
- [0190] [2] Hanash, S. M., Pitteri, S. J., and Faca, V. M. Mining the plasma proteome for cancer biomarkers. *Nature* 452, 571-579 (2008).
- [0191] [3] Sreekumar, A. et al. Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. *Nature* 457, 910-914 (2009).
- [0192] [4] Villanueva, J. et al. Differential exoprotease activities confer tumor-specific serum peptidome patterns. *J. Clin. Invest.* 116, 271-284 (2006).
- [0193] [5] Surinova, S. et al. On the development of plasma protein biomarkers. *J. Proteome Res.* 10, 5-16 (2011).
- [0194] [6] Schwarzenbach, H., Hoon, D. S. B., and Pantel, K. Cell-free nucleic acids as biomarkers in cancer patients. *Nat. Rev. Cancer* 11, 426-437 (2011).
- [0195] [7] Moon, P.-G., You, S., Lee, J.-E., Hwang, D., and Baek, M.-C. Urinary exosomes and proteomics. *Mass Spectrom. Rev.* 30, 1185-1202 (2011).
- [0196] [8] Nagrath, S. et al. Isolation of rare circulating tumour cells in cancer patients by microchip technology. *Nature* 450, 1235-1239 (2007).
- [0197] [9] Lutz, A. M., Willmann, J. K., Cochran, F. V., Ray, P., and Gambhir, S. S. Cancer screening: a mathematical model relating secreted blood biomarker levels to tumor sizes. *PloS Med.* 5, e170 (2008).
- [0198] [10] Anderson, N. L. and Anderson, N. G. The human plasma proteome: history, character, and diagnostic prospects. *Mol. Cell. Proteomics* 1, 845-867 (2002).
- [0199] [11] Haun, J. B. et al. Micro-nmr for rapid molecular analysis of human tumor samples. *Sci. Transl. Med.* 3, 71ra16 (2011).
- [0200] [12] Fonovic, M. and Bogoy, M. Activity-based probes as a tool for functional proteomic analysis of proteases. *Expert Rev. Proteomics* 5, 721-730 (2008).
- [0201] [13] Baruch, A., Jeffery, D. A., and Bogoy, M. Enzyme activity—it's all about image. *Trends Cell. Biol.* 14, 29-35 (2004).
- [0202] [14] Hilderbrand, S. A. and Weissleder, R. Near-infrared fluorescence: application to in vivo molecular imaging. *Curr. Opin. Chem. Biol.* 14, 71-79 (2010).
- [0203] [15] Braet, F. and Wisse, E. Structural and functional aspects of liver sinusoidal endothelial cell fenestrae: a review. *Comp. Hepatol.* 1, 1 (2002).
- [0204] [16] Jain, R. K. and Stylianopoulos, T. Delivering nanomedicine to solid tumors. *Nat. Rev. Clin. Oncol.* 7, 653-664 (2010).
- [0205] [17] López-Otín, C. and Bond, J. S. Proteases: multifunctional enzymes in life and disease. *J. Biol. Chem.* 283, 30433-30437 (2008).
- [0206] [18] Schuppan, D. and Afdhal, N. H. Liver cirrhosis. *Lancet* 371, 838-851 (2008).
- [0207] [19] Hori, S. S. and Gambhir, S. S. Mathematical model identifies blood biomarker-based early cancer detection strategies and limitations. *Sci. Transl. Med.* 3, 109ra116 (2011).
- [0208] [20] Bremer, C., Tung, C. H., and Weissleder, R. In vivo molecular target assessment of matrix metalloproteinase inhibition. *Nat. Med.* 7, 743-748 (2001).
- [0209] [21] Kridel, S. J. et al. A unique substrate binding mode discriminates membrane type-1 matrix metalloproteinase from other matrix metalloproteinases. *J. Biol. Chem.* 277, 23788-23793 (2002).
- [0210] [22] Lutolf, M. P. et al. Repair of bone defects using synthetic mimetics of collagenous extracellular matrices. *Nat. Biotechnol.* 21, 513-518 (2003).
- [0211] [23] Mahmood, U. and Weissleder, R. Near-infrared optical imaging of proteases in cancer. *Mol. Cancer Ther.* 2, 489-496 (2003).
- [0212] [24] Turk, B. E., Huang, L. L., Piro, E. T., and Cantley, L. C. Determination of protease cleavage site motifs using mixture-based oriented peptide libraries. *Nat. Biotechnol.* 19, 661-667 (2001).
- [0213] [25] Park, J.-H. et al. Systematic surface engineering of magnetic nanoworms for in vivo tumor targeting. *Small* 5, 694-700 (2009).
- [0214] [26] Fickert, P. et al. A new xenobiotic-induced mouse model of sclerosing cholangitis and biliary fibrosis. *Am. J. Pathol.* 171, 525-536 (2007).

- [0215] [27] Morris, T. A. et al. Urine and plasma levels of fibrinopeptide b in patients with deep vein thrombosis and pulmonary embolism. *Thromb. Res.* 110, 159-165 (2003).
- [0216] [28] Choi, H. S. et al. Renal clearance of quantum dots. *Nat. Biotechnol.* 25, 1165-1170 (2007).
- [0217] [29] Villanueva, J., Nazarian, A., Lawlor, K., Yi, S. S., Robbins, R. J., and Tempst, P. A sequence-specific exopeptidase activity test (sseat) for “functional” biomarker discovery. *Mol. Cell. Proteomics* 7, 509-518 (2008).
- [0218] [30] Brown, B. B., Wagner, D. S., and Geysen, H. M. A single-bead decode strategy using electrospray ionization mass spectrometry and a new photolabile linker: 3-amino-3-(2-nitrophenyl)propionic acid. *Mol. Divers.* 1, 4-12 (1995).
- [0219] [31] Ross, P. L. et al. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* 3, 1154-1169 (2004).
- [0220] [32] Thompson, A. et al. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal. Chem.* 75, 1895-1904 (2003).
- [0221] [33] Rockey, D. C., Caldwell, S. H., Goodman, Z. D., Nelson, R. C., Smith, A. D., and for the Study of Liver Diseases, A. A. Liver biopsy. *Hepatology* 49, 1017-1044 (2009).
- [0222] [34] Popov, Y. and Schuppan, D. Targeting liver fibrosis: strategies for development and validation of antifibrotic therapies. *Hepatology* 50, 1294-1306 (2009).
- [0223] [35] Bedossa, P., Dargère, D., and Paradis, V. Sampling variability of liver fibrosis in chronic hepatitis C. *Hepatology* 38, 1449-1457 (2003).
- [0224] [36] Etzioni, R. et al. The case for early detection. *Nat. Rev. Cancer* 3, 243-252 (2003).
- [0225] [37] Kulasingam, V., Pavlou, M. P., and Diamandis, E. P. Integrating high-throughput technologies in the quest for effective biomarkers for ovarian cancer. *Nat. Rev. Cancer* 10, 371-378 (2010).
- [0226] [38] D’Souza, A. L. et al. A strategy for blood biomarker amplification and localization using ultrasound. *Proc. Natl. Acad. Sci.* 106, 17152-17157 (2009).
- [0227] [39] Schima, W., Kulinna, C., Langenberger, H., and Ba-Ssalamah, A. Liver metastases of colorectal cancer: US, CT or MR? *Cancer Imaging* 5 Spec No A, S149-S156 (2005).
- [0228] [40] Ruoslahti, E., Bhatia, S. N., and Sailor, M. J. Targeting of drugs and nanoparticles to tumors. *J. Cell Biol.* 188, 759-768 (2010).
- [0229] [41] Sugahara, K. N. et al. *Science* 328, 1031-1035 (2010).
- [0230] [42] Miller, M. A. et al. Proteolytic activity matrix analysis (prama) for simultaneous determination of multiple protease activities. *Integr. Biol.* 3, 422-438 (2011).
- [0231] [43] Popov, Y., Patsenker, E., Fickert, P., Trauner, M., and Schuppan, D. Mdr2 (abcb4)^{-/-} mice spontaneously develop severe biliary fibrosis via massive dysregulation of pro- and antifibrogenic genes. *J. Hepatol.* 43, 1045-1054 (2005).
- [0232] The entire contents of all references described or listed herein, including references [1]-[43] listed above, any references described in the specification above, and any reference described below, are incorporated herein by reference, as if each and every individual reference was incorporated herein by reference. In case of a conflict between the teachings of any reference incorporated herein and the specification, the specification shall control.
- [0233] The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. The present invention is not limited in scope by the examples provided, since the examples are intended as illustrations of various aspect of the invention and other functionally equivalent embodiments are within the scope of the invention. Various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and fall within the scope of the appended claims. The advantages and objects of the invention are not necessarily encompassed by each embodiment of the invention.
- [0234] As used herein and in the appended claims, the singular forms “a”, “an”, and “the” include the plural reference unless the context clearly indicates otherwise. Thus, for example, a reference to “a cell” includes a plurality of such cells, and a reference to “an antibody” is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.
- [0235] Where ranges are given, endpoints are included. For each range an embodiment in which the described parameter assumes a single value within the range or a sub-range within the range is included.
- [0236] Not all embodiments have been described and combination of a feature or of multiple features of any embodiment(s) described herein with any feature or multiple features of any other embodiment(s) described herein is contemplated. Similarly, any limitations in any of the appended claims may be combined with any limitation in any of the other claims, or a combination of such limitations. Such combined embodiments have not been explicitly set forth to save space, but are expressly contemplated to be within the scope of the present invention. Similarly, any feature(s) of any embodiment described herein may be excluded from the scope of the appended claims.
- [0237] All references, patents and patent publications that are recited in this application are incorporated in their entirety herein by reference.

SEQUENCE LISTING

Sequence total quantity: 37
 SEQ ID NO: 1 moltype = AA length = 14
 FEATURE Location/Qualifiers
 REGION 1..14
 note = Synthetic polypeptide
 SITE 1
 note = May be isotope E+3, E+2 or E+1
 SITE 2

-continued

SITE	note = May be isotope G+6 8	
source	note = May be isotope E+1 or E+2 1..14 mol_type = protein organism = synthetic construct	
SEQUENCE: 1 EGVNDNEEGF FSAR		14
SEQ ID NO: 2 FEATURE REGION	moltype = AA length = 10 Location/Qualifiers 1..10 note = Synthetic polypeptide	
source	1..10 mol_type = protein organism = synthetic construct	
SEQUENCE: 2 KGGPVGLIGC		10
SEQ ID NO: 3 FEATURE REGION	moltype = AA length = 13 Location/Qualifiers 1..13 note = Synthetic polypeptide	
source	1..13 mol_type = protein organism = synthetic construct	
SEQUENCE: 3 KGGPWGIWQ GGC		13
SEQ ID NO: 4 FEATURE REGION	moltype = AA length = 6 Location/Qualifiers 1..6 note = Synthetic polypeptide	
source	1..6 mol_type = protein organism = synthetic construct	
SEQUENCE: 4 GFFSAR		6
SEQ ID NO: 5 FEATURE REGION	moltype = AA length = 8 Location/Qualifiers 1..8 note = Synthetic polypeptide	
source	1..8 mol_type = protein organism = synthetic construct	
SEQUENCE: 5 EGVNDNEE		8
SEQ ID NO: 6 FEATURE REGION	moltype = AA length = 7 Location/Qualifiers 1..7 note = Synthetic polypeptide	
source	1..7 mol_type = protein organism = synthetic construct	
SEQUENCE: 6 EGFFSAR		7
SEQ ID NO: 7 FEATURE REGION	moltype = AA length = 7 Location/Qualifiers 1..7 note = Synthetic polypeptide	
source	1..7 mol_type = protein organism = synthetic construct	
SEQUENCE: 7 EGVNDNE		7
SEQ ID NO: 8 FEATURE REGION	moltype = AA length = 24 Location/Qualifiers 1..24 note = Synthetic polypeptide	
SITE	1 note = D-Glutamic Acid	
SITE	1	

-continued

	note = Isotope E+3	
SITE	2	
	note = Isotope G+6	
SITE	14	
	note = D-Arginine	
MOD_RES	14..15	
	note = Modified by 3-amino-3-(2-nitrophenyl)propionic acid	
MOD_RES	15	
	note = Modified by carboxyfluorescein	
source	1..24	
	mol_type = protein	
	organism = synthetic construct	
SITE	4	
	note = D-Asparagine	
SITE	5	
	note = D-Aspartic acid	
SITE	6	
	note = D-Asparagine	
SITE	7	
	note = D-Glutamic Acid	
SITE	8	
	note = D-Glutamic Acid	
SITE	11	
	note = D-Phenylalanine	
SITE	12	
	note = D-Serine	
SEQUENCE: 8		
EGVNDNEEGF FSARKGGPQG IWGQ		24
SEQ ID NO: 9	moltype = AA length = 24	
FEATURE	Location/Qualifiers	
REGION	1..24	
	note = Synthetic polypeptide	
SITE	1	
	note = D-Glutamic Acid	
SITE	1	
	note = Isotope E+2	
SITE	2	
	note = Isotope G+6	
SITE	8	
	note = Isotope E+1	
SITE	14	
	note = D-Arginine	
MOD_RES	14..15	
	note = Modified by 3-amino-3-(2-nitrophenyl)propionic acid	
MOD_RES	15	
	note = Modified by carboxyfluorescein	
source	1..24	
	mol_type = protein	
	organism = synthetic construct	
SITE	4	
	note = D-Asparagine	
SITE	5	
	note = D-Aspartic Acid	
SITE	6	
	note = D-Asparagine	
SITE	7	
	note = D-Glutamic Acid	
SITE	8	
	note = D-Glutamic Acid	
SITE	11	
	note = D-Phenylalanine	
SITE	12	
	note = D-Serine	
SEQUENCE: 9		
EGVNDNEEGF FSARKGGLVP RGSG		24
SEQ ID NO: 10	moltype = AA length = 23	
FEATURE	Location/Qualifiers	
REGION	1..23	
	note = Synthetic polypeptide	
SITE	1	
	note = D-Glutamic Acid	
SITE	1	
	note = Isotope E+1	
SITE	2	

-continued

SITE	note = Isotope G+6 8	
SITE	note = Isotope E+2 14	
MOD_RES	note = D-Arginine 14..15	
MOD_RES	note = Modified by 3-amino-3-(2-nitrophenyl)propionic acid 15	
source	note = Modified by carboxyfluorescein 1..23 mol_type = protein organism = synthetic construct	
SITE	4 note = D-Asparagine	
SITE	5 note = D-Aspartic Acid	
SITE	6 note = D-Asparagine	
SITE	7 note = D-Glutamic Acid	
SITE	8 note = D-Glutamic Acid	
SITE	11 note = D-Phenylalanine	
SITE	12 note = D-Serine	
SEQUENCE: 10		
EGVNDNEEGF FSARKGGPVG LIG		23
SEQ ID NO: 11	moltype = AA length = 25	
FEATURE	Location/Qualifiers	
REGION	1..25 note = Synthetic polypeptide	
SITE	1 note = D-Glutamic Acid	
SITE	2 note = Isotope G+6	
SITE	8 note = Isotope E+2	
SITE	12 note = Isotope S+1	
SITE	14 note = D-Arginine	
MOD_RES	14..15 note = Modified by 3-amino-3-(2-nitrophenyl)propionic acid	
MOD_RES	15 note = Modified by carboxyfluorescein	
source	1..25 mol_type = protein organism = synthetic construct	
SITE	4 note = D-Asparagine	
SITE	5 note = D-Aspartic Acid	
SITE	6 note = D-Asparagine	
SITE	7 note = D-Glutamic Acid	
SITE	8 note = D-Glutamic Acid	
SITE	11 note = D-Phenylalanine	
SITE	12 note = D-Serine	
SEQUENCE: 11		
EGVNDNEEGF FSARKGGPWG IWGQG		25
SEQ ID NO: 12	moltype = AA length = 25	
FEATURE	Location/Qualifiers	
REGION	1..25 note = Synthetic polypeptide	
SITE	1 note = D-Glutamic Acid	
SITE	2 note = Isotope G+5	
SITE	12	

-continued

	note = Isotope S+4	
SITE	14	
	note = D-Arginine	
MOD_RES	14..15	
	note = Modified by 3-amino-3-(2-nitrophenyl)propionic acid	
MOD_RES	15	
	note = Modified by carboxyfluorescein	
source	1..25	
	mol_type = protein	
	organism = synthetic construct	
SITE	4	
	note = D-Asparagine	
SITE	5	
	note = D-Aspartic Acid	
SITE	6	
	note = D-Asparagine	
SITE	7	
	note = D-Glutamic Acid	
SITE	8	
	note = D-Glutamic Acid	
SITE	11	
	note = D-Phenylalanine	
SITE	12	
	note = D-Serine	
SEQUENCE: 12		
EGVNDNEEGF FSARKGGPVP LSLVM		25
SEQ ID NO: 13	moltype = AA length = 24	
FEATURE	Location/Qualifiers	
REGION	1..24	
	note = Synthetic polypeptide	
SITE	1	
	note = D-Glutamic Acid	
SITE	1	
	note = Isotope E+3	
SITE	2	
	note = Isotope G+1	
SITE	8	
	note = Isotope E+1	
SITE	12	
	note = Isotope S+4	
SITE	14	
	note = D-Arginine	
MOD_RES	14..15	
	note = Modified by 3-amino-3-(2-nitrophenyl)propionic acid	
MOD_RES	15	
	note = Modified by carboxyfluorescein	
source	1..24	
	mol_type = protein	
	organism = synthetic construct	
SITE	4	
	note = D-Asparagine	
SITE	5	
	note = D-Aspartic Acid	
SITE	6	
	note = D-Asparagine	
SITE	7	
	note = D-Glutamic Acid	
SITE	8	
	note = D-Glutamic Acid	
SITE	11	
	note = D-Phenylalanine	
SITE	12	
	note = D-Serine	
SEQUENCE: 13		
EGVNDNEEGF FSARKGGPLG LRSW		24
SEQ ID NO: 14	moltype = AA length = 24	
FEATURE	Location/Qualifiers	
REGION	1..24	
	note = Synthetic polypeptide	
SITE	1	
	note = D-Glutamic Acid	
SITE	1	
	note = Isotope E+3	
SITE	9	

-continued

SITE	note = Isotope G+6 14	
MOD_RES	note = D-Arginine 14..15	
MOD_RES	note = Modified by 3-amino-3-(2-nitrophenyl)propionic acid 15	
source	note = Modified by carboxyfluorescein 1..24 mol_type = protein organism = synthetic construct	
SITE	4 note = D-Asparagine	
SITE	5 note = D-Aspartic Acid	
SITE	6 note = D-Asparagine	
SITE	7 note = D-Glutamic Acid	
SITE	8 note = D-Glutamic Acid	
SITE	11 note = D-Phenylalanine	
SITE	12 note = D-Serine	
SEQUENCE: 14		
EGVNDNEEGF FSARKGGPLG VRGK		24
SEQ ID NO: 15	moltype = AA length = 23	
FEATURE	Location/Qualifiers	
REGION	1..23 note = Synthetic polypeptide	
SITE	1 note = D-Glutamic Acid	
SITE	1 note = Isotope E+2	
SITE	9 note = Isotope G+6	
SITE	12 note = Isotope S+1	
SITE	14 note = D-Arginine	
MOD_RES	14..15 note = Modified by 3-amino-3-(2-nitrophenyl)propionic acid	
MOD_RES	15 note = Modified by carboxyfluorescein	
SITE	18 note = D-Phenylalanine	
MOD_RES	18 note = Modified by pipecolic acid	
source	1..23 mol_type = protein organism = synthetic construct	
SITE	4 note = D-Asparagine	
SITE	5 note = D-Aspartic Acid	
SITE	6 note = D-Asparagine	
SITE	7 note = D-Glutamic Acid	
SITE	8 note = D-Glutamic Acid	
SITE	11 note = D-Phenylalanine	
SITE	12 note = D-Serine	
SEQUENCE: 15		
EGVNDNEEGF FSARKGGFRS GGG		23
SEQ ID NO: 16	moltype = AA length = 24	
FEATURE	Location/Qualifiers	
REGION	1..24 note = Synthetic polypeptide	
SITE	1 note = D-Glutamic Acid	
SITE	1	

-continued

SITE	note = Isotope E+1	
	8	
SITE	note = Isotope E+2	
	9	
SITE	note = Isotope G+6	
	14	
MOD_RES	note = D-Arginine	
	14..15	
MOD_RES	note = Modified by 3-amino-3-(2-nitrophenyl)propionic acid	
	15	
SITE	note = Modified by carboxyfluorescein	
	18	
source	note = D-Phenylalanine	
	1..24	
	mol_type = protein	
	organism = synthetic construct	
SITE	4	
SITE	note = D-Asparagine	
	6	
SITE	note = D-Asparagine	
	5	
SITE	note = D-Aspartic Acid	
	7	
SITE	note = D-Glutamic Acid	
	8	
SITE	note = D-Glutamic Acid	
	11	
SITE	note = D-Phenylalanine	
	12	
	note = D-Serine	
SEQUENCE: 16		
EGVNDNEEGF FSARKGGFPR	SGGG	24
SEQ ID NO: 17	moltype = AA length = 23	
FEATURE	Location/Qualifiers	
REGION	1..23	
	note = Synthetic polypeptide	
SITE	1	
	note = D-Glutamic Acid	
SITE	8	
	note = Isotope E+3	
SITE	9	
	note = Isotope G+6	
SITE	14	
	note = D-Arginine	
MOD_RES	14..15	
	note = Modified by 3-amino-3-(2-nitrophenyl)propionic acid	
MOD_RES	15	
	note = Modified by carboxyfluorescein	
SITE	18	
	note = D-Phenylalanine	
MOD_RES	18	
	note = Modified by pipecolic acid	
source	1..23	
	mol_type = protein	
	organism = synthetic construct	
SITE	4	
	note = D-Asparagine	
SITE	6	
	note = D-Asparagine	
SITE	5	
	note = D-Aspartic Acid	
SITE	7	
	note = D-Glutamic Acid	
SITE	8	
	note = D-Glutamic Acid	
SITE	11	
	note = D-Phenylalanine	
SITE	12	
	note = D-Serine	
SEQUENCE: 17		
EGVNDNEEGF FSARKGGFKS	GGG	23
SEQ ID NO: 18	moltype = AA length = 7	
FEATURE	Location/Qualifiers	
REGION	1..7	

-continued

source	note = Synthetic polypeptide 1..7 mol_type = protein organism = synthetic construct	
SEQUENCE: 18 PQGIWGQ		7
SEQ ID NO: 19 FEATURE REGION	moltype = AA length = 7 Location/Qualifiers 1..7 note = Synthetic polypeptide	
source	1..7 mol_type = protein organism = synthetic construct	
SEQUENCE: 19 LVPRGSG		7
SEQ ID NO: 20 FEATURE REGION	moltype = AA length = 6 Location/Qualifiers 1..6 note = Synthetic polypeptide	
source	1..6 mol_type = protein organism = synthetic construct	
SEQUENCE: 20 PVGLIG		6
SEQ ID NO: 21 FEATURE REGION	moltype = AA length = 8 Location/Qualifiers 1..8 note = Synthetic polypeptide	
source	1..8 mol_type = protein organism = synthetic construct	
SEQUENCE: 21 PWGIWGQG		8
SEQ ID NO: 22 FEATURE REGION	moltype = AA length = 8 Location/Qualifiers 1..8 note = Synthetic polypeptide	
source	1..8 mol_type = protein organism = synthetic construct	
SEQUENCE: 22 PVPLSLVM		8
SEQ ID NO: 23 FEATURE REGION	moltype = AA length = 7 Location/Qualifiers 1..7 note = Synthetic polypeptide	
source	1..7 mol_type = protein organism = synthetic construct	
SEQUENCE: 23 PLGLRSW		7
SEQ ID NO: 24 FEATURE REGION	moltype = AA length = 7 Location/Qualifiers 1..7 note = Synthetic polypeptide	
source	1..7 mol_type = protein organism = synthetic construct	
SEQUENCE: 24 PLGVRGK		7
SEQ ID NO: 25 FEATURE REGION	moltype = AA length = 6 Location/Qualifiers 1..6 note = Synthetic polypeptide	
SITE	1 note = D-Phenylalanine	
MOD_RES	1 note = Modified by pipecolic acid	
source	1..6	

-continued

	mol_type = protein organism = synthetic construct	
SEQUENCE: 25 FRSGGG		6
SEQ ID NO: 26 FEATURE REGION SITE source	moltype = AA length = 7 Location/Qualifiers 1..7 note = Synthetic polypeptide 1 note = D-Phenylalanine 1..7 mol_type = protein organism = synthetic construct	
SEQUENCE: 26 FPRSOGG		7
SEQ ID NO: 27 FEATURE REGION SITE MOD_RES source	moltype = AA length = 6 Location/Qualifiers 1..6 note = Synthetic polypeptide 1 note = D-Phenylalanine 1 note = Modified by pipecolic acid 1..6 mol_type = protein organism = synthetic construct	
SEQUENCE: 27 FKSGGG		6
SEQ ID NO: 28 FEATURE REGION source	moltype = AA length = 13 Location/Qualifiers 1..13 note = Synthetic polypeptide 1..13 mol_type = protein organism = synthetic construct	
SEQUENCE: 28 EGVNDNEEGF SAR		13
SEQ ID NO: 29 FEATURE REGION source	moltype = AA length = 5 Location/Qualifiers 1..5 note = Synthetic polypeptide 1..5 mol_type = protein organism = synthetic construct	
SEQUENCE: 29 GFSAR		5
SEQ ID NO: 30 FEATURE REGION source	moltype = AA length = 8 Location/Qualifiers 1..8 note = Synthetic polypeptide 1..8 mol_type = protein organism = synthetic construct	
SEQUENCE: 30 PPGFSPFR		8
SEQ ID NO: 31 FEATURE REGION source	moltype = AA length = 10 Location/Qualifiers 1..10 note = Synthetic polypeptide 1..10 mol_type = protein organism = synthetic construct	
SEQUENCE: 31 DRVYIHPFHL		10
SEQ ID NO: 32 FEATURE REGION source	moltype = AA length = 17 Location/Qualifiers 1..17 note = Synthetic polypeptide 1..17	

-continued

	mol_type = protein organism = synthetic construct	
SEQUENCE: 32 SYSMEHFRWG KPVGKKR		17
SEQ ID NO: 33 FEATURE REGION	moltype = AA length = 22 Location/Qualifiers 1..22 note = Synthetic polypeptide	
source	1..22 mol_type = protein organism = synthetic construct	
SEQUENCE: 33 RPVKVYPNGA EDESAEAFPL EF		22
SEQ ID NO: 34 FEATURE REGION	moltype = AA length = 32 Location/Qualifiers 1..32 note = Synthetic polypeptide	
source	1..32 mol_type = protein organism = synthetic construct	
SEQUENCE: 34 FRWGKPVGKK RRPVKVYPNG AEDESAEAFP LE		32
SEQ ID NO: 35 FEATURE REGION	moltype = AA length = 11 Location/Qualifiers 1..11 note = Synthetic polypeptide	
SITE	6 note = X is Norvaline (Nva)	
SITE	7 note = X is N - beta - (2,4 - dinitrophenyl) - L -alpha,beta-diaminopropionic acid (Dap)	
source	1..11 mol_type = protein organism = synthetic construct	
SEQUENCE: 35 GGPLAXXARG C		11
SEQ ID NO: 36 FEATURE REGION	moltype = AA length = 14 Location/Qualifiers 1..14 note = Synthetic polypeptide	
SITE	2 note = May be isotope G+1, G+2 or G+3	
SITE	3 note = May be isotope V+1, V+5 or V+6	
SITE	4 note = May be isotope N+6	
SITE	6 note = May be isotope N+2 or N+6	
REGION	10..11 note = May be isotope F+6 or F+10	
SITE	13 note = May be isotope A+1 or A+4	
source	1..14 mol_type = protein organism = synthetic construct	
SEQUENCE: 36 EGVNDNEEGF FSAR		14
SEQ ID NO: 37 FEATURE REGION	moltype = AA length = 18 Location/Qualifiers 1..18 note = Synthetic polypeptide	
SITE	1 note = D-Glutamic Acid	
SITE	2 note = Isotope G+2 or G+3	
SITE	3 note = Isotope V+6	
SITE	12 note = D-Serine	
SITE	13 note = May be isotope A+1	

-continued

SITE	14
	note = D-Arginine
MOD_RES	14..15
	note = Modified by 3-amino-3-(2-nitrophenyl)propionic acid
MOD_RES	15
	note = Modified by carboxyfluorescein
source	1..18
	mol_type = protein
	organism = synthetic construct
SITE	4
	note = D-Asparagine
SITE	6
	note = D-Asparagine
SITE	5
	note = D-Aspartic Acid
SITE	7
	note = D-Glutamic Acid
SITE	8
	note = D-Glutamic Acid
SEQUENCE: 37	
EGVNDNEEGF FSARKGGC	18

- 1-103.** (canceled)
- 104.** A composition comprising a plurality of isotope-labeled molecules wherein each isotope-labeled molecule of said plurality of isotope-labeled molecules comprises:
- a) a carrier, and
 - b) a mass tag conjugated to said carrier via a cleavable linker, wherein said cleavable linker is cleaved by a disease-associated protease.
- 105.** The composition of claim **104**, wherein said mass tag comprises an isobaric mass tag.
- 106.** The composition of claim **104**, wherein said mass tag comprises a peptide mass tag.
- 107.** The composition of claim **104**, wherein said isotope-labeled molecules are configured for use in mass spectrometry (MS).
- 108.** The composition of claim **107**, wherein said isotope-labeled molecules are configured for use in liquid-chromatography mass spectrometry (LC-MS).
- 109.** The composition of claim **107**, wherein said isotope-labeled molecules are configured for use in tandem mass spectrometry (MS/MS).
- 110.** The composition of claim **104**, wherein said disease-associated protease is selected from the group consisting of MMP2, MMP7, MMP9, MMP12, MMP14, tissue factor, factor Xa, cathepsin, and thrombin.
- 111.** The composition of claim **104**, wherein each isotope-labeled molecule of said plurality of isotope-labeled molecules is the same.

- 112.** The composition of claim **104**, wherein said plurality of isotope-labeled molecules comprises at least 6 different mass tag-conjugated cleavable linkers.
- 113.** The composition of claim **104**, wherein said plurality of isotope-labeled molecules comprises at least 10 different mass tag-conjugated cleavable linkers.
- 114.** The composition of claim **104**, wherein said plurality of isotope-labeled molecules comprises at least 20 different mass tag-conjugated cleavable linkers.
- 115.** The composition of claim **104**, wherein said cleavage occurs in vivo.
- 116.** The composition of claim **104**, wherein said cleavage occurs in vitro.
- 117.** The composition of claim **104**, wherein said cleavage occurs ex vivo.
- 118.** The composition of claim **104**, wherein said carrier comprises a particle.
- 119.** The composition of claim **118**, wherein said particle comprises a microparticle or a nanoparticle.
- 120.** The composition of claim **104**, wherein said carrier further comprises a binding agent.
- 121.** The composition of claim **120**, wherein said binding agent is selected from a group consisting of an antibody, an antibody fragment, an aptamer, and an adnectin.
- 122.** The composition of claim **104**, wherein said cleavable linker comprises a peptide.
- 123.** The composition of claim **104**, wherein each isotope-labeled molecule of said plurality of isotope-labeled molecules further comprises a reactive chemical moiety.

* * * * *