



US 20230194544A1

(19) **United States**

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

(10) **Pub. No.: US 2023/0194544 A1**

(43) **Pub. Date: Jun. 22, 2023**

(54) **METHODS AND PRODUCTS FOR IN VIVO
ENZYME PROFILING**

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

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

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

(21) Appl. No.: **18/056,752**

(22) Filed: **Nov. 18, 2022**

Related U.S. Application Data

(60) Continuation of application No. 17/570,847, filed on
Jan. 7, 2022, now Pat. No. 11,549,951, which is a
continuation of application No. 17/091,075, filed on
Nov. 6, 2020, which is a continuation of application
No. 15/966,385, filed on Apr. 30, 2018, now Pat. No.
10,883,998, which is a division of application No.
14/166,481, filed on Jan. 28, 2014, now Pat. No.

9,970,941, which is a continuation of application No.
12/715,965, filed on Mar. 2, 2010, now Pat. No.
8,673,267.

(60) Provisional application No. 61/156,660, filed on Mar.
2, 2009.

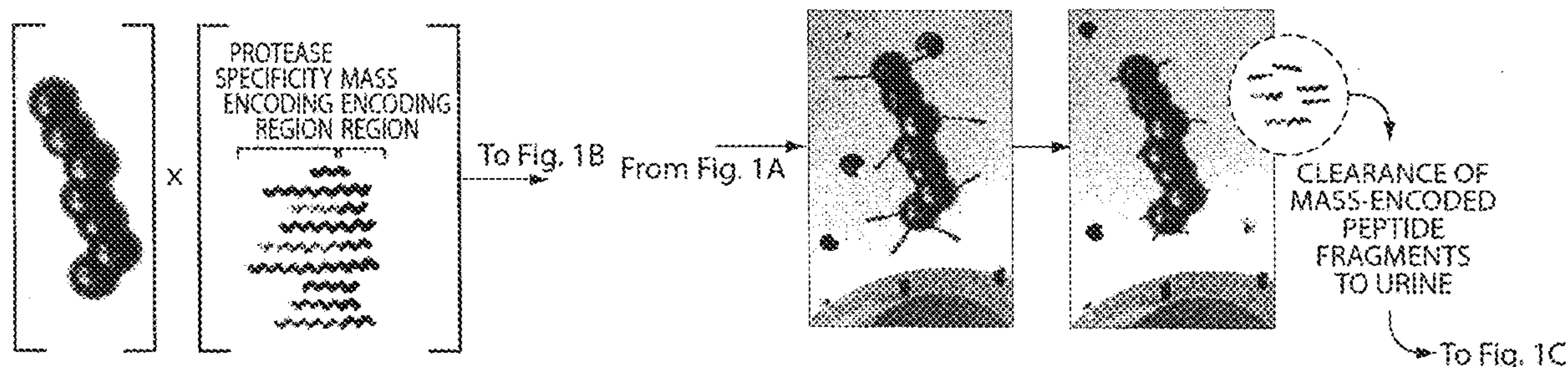
Publication Classification

(51) **Int. Cl.**
G01N 33/68 (2006.01)
C07K 7/06 (2006.01)
C12Q 1/37 (2006.01)
C12Q 1/56 (2006.01)
(52) **U.S. Cl.**
CPC **G01N 33/6848** (2013.01); **C07K 7/06**
(2013.01); **G01N 33/6842** (2013.01); **C12Q**
1/37 (2013.01); **C12Q 1/56** (2013.01)

(57) **ABSTRACT**

The present invention relates to methods and products
associated with in vivo enzyme profiling. In particular, the
invention relates to methods of in vivo processing of exog-
enous molecules followed by detection of signature mol-
ecules as representative of the presence of active enzymes
associated with diseases or conditions. The invention also
relates to products, kits, and databases for use in the methods
of the invention.

Specification includes a Sequence Listing.



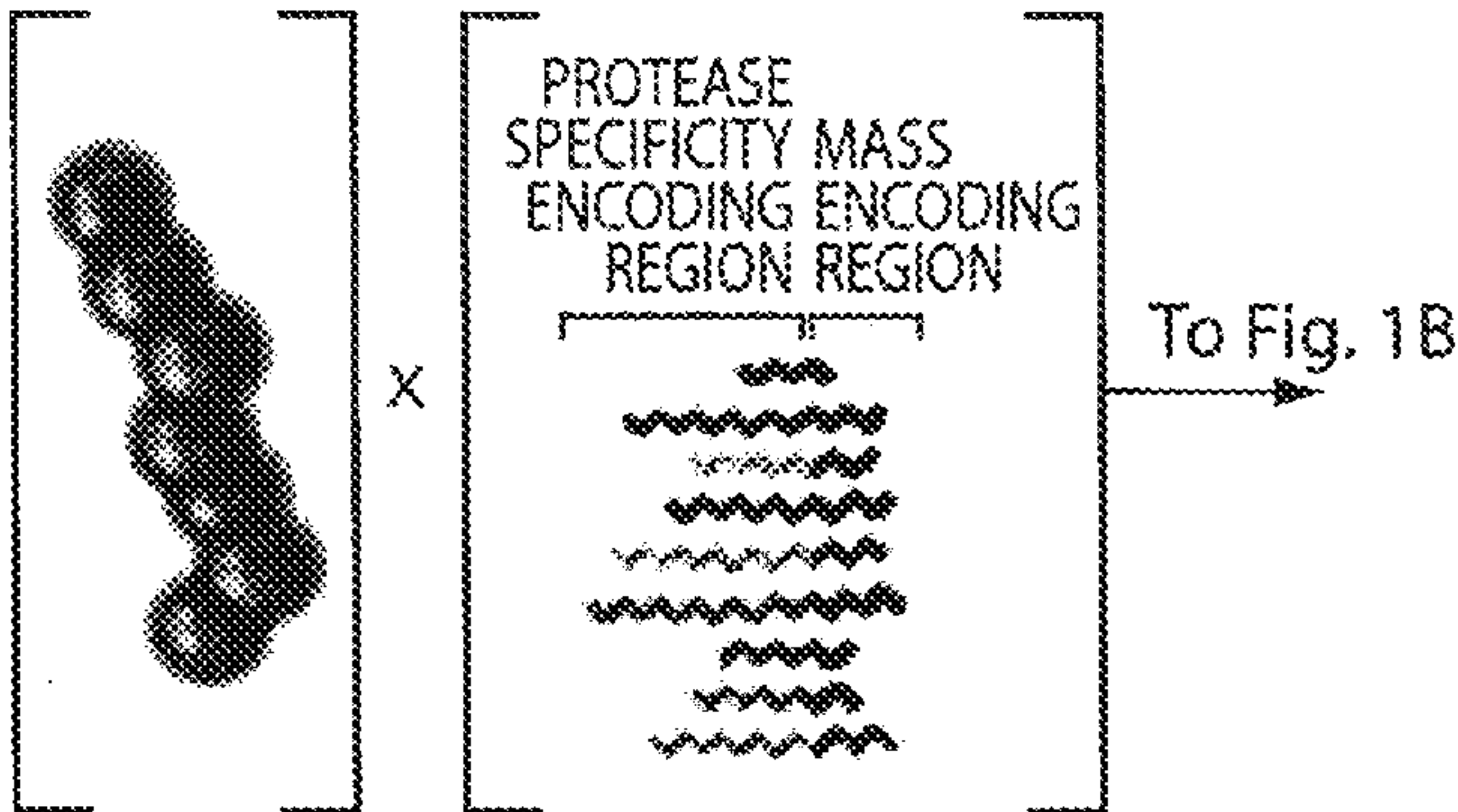


Fig. 1A

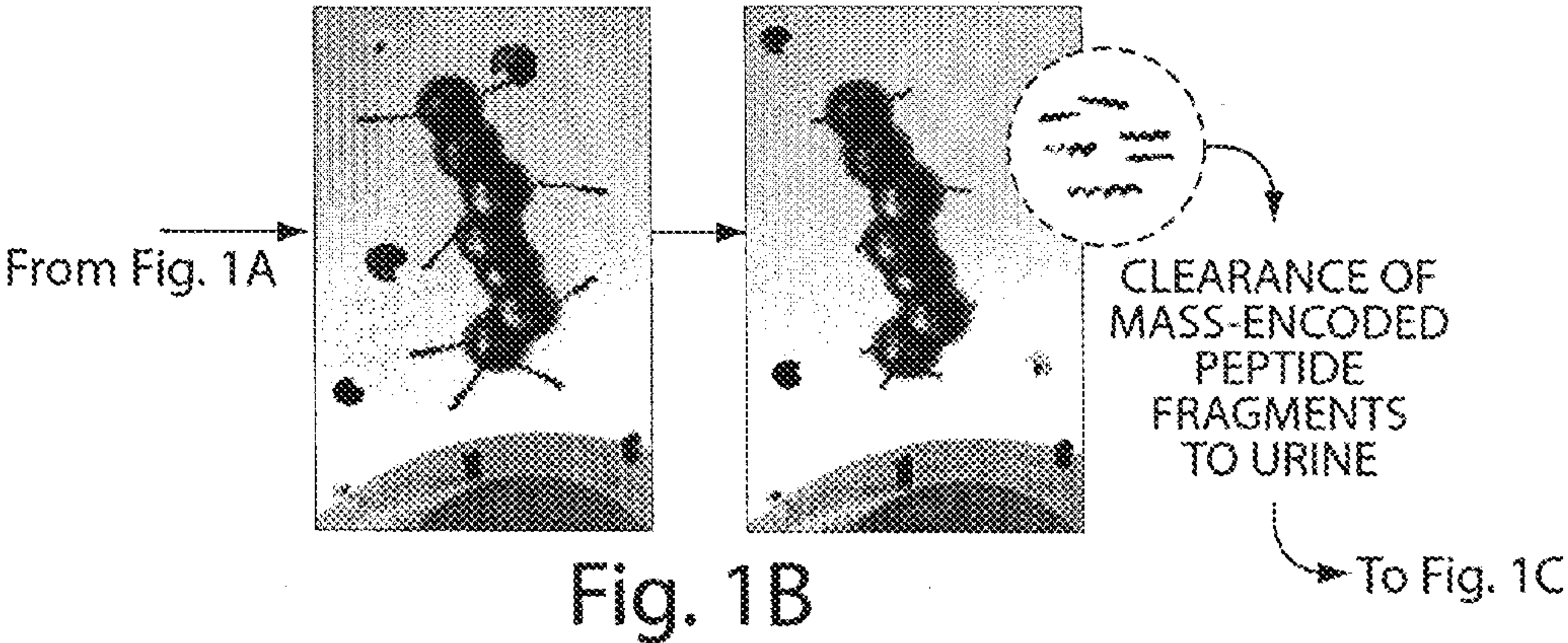
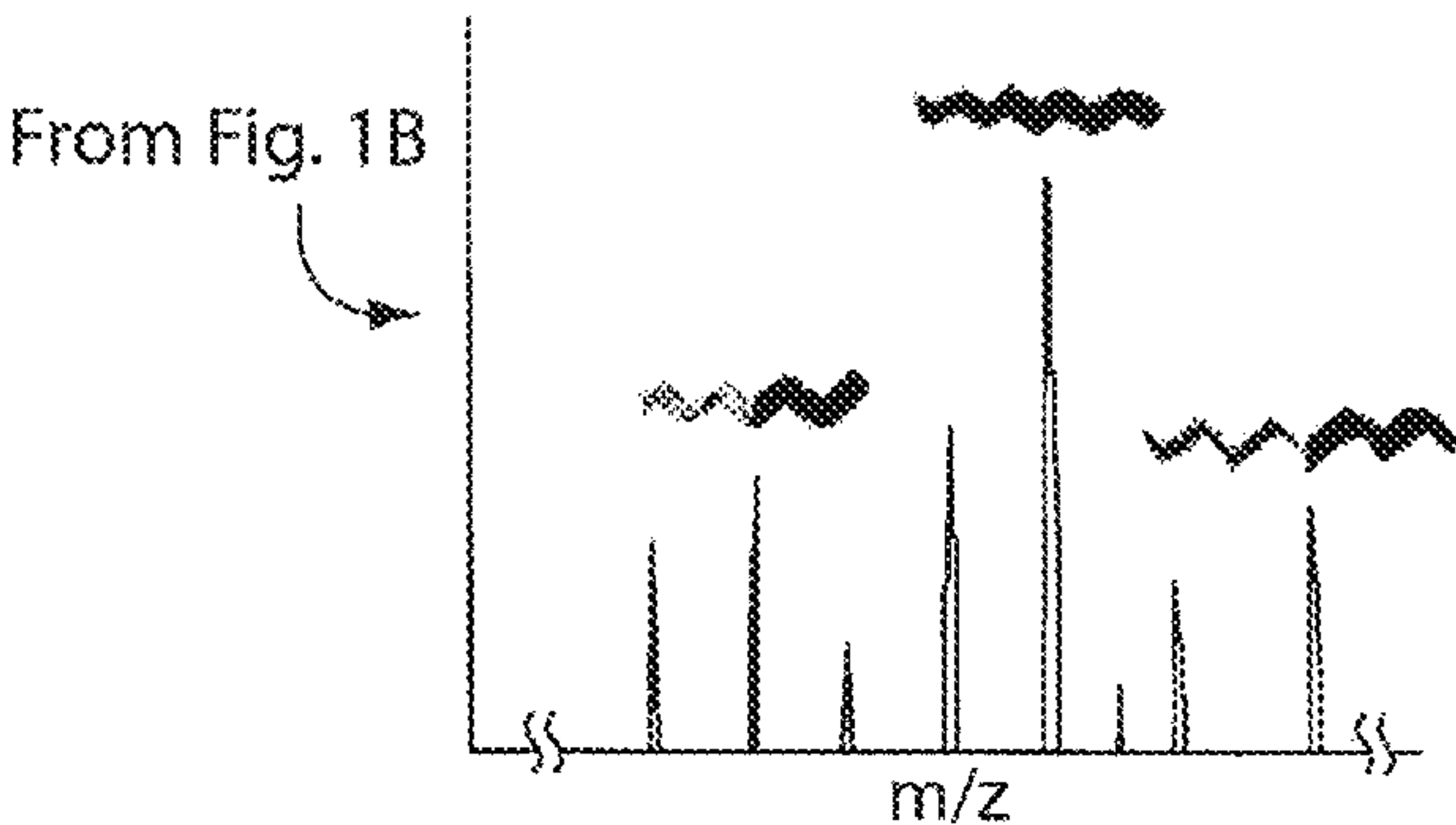
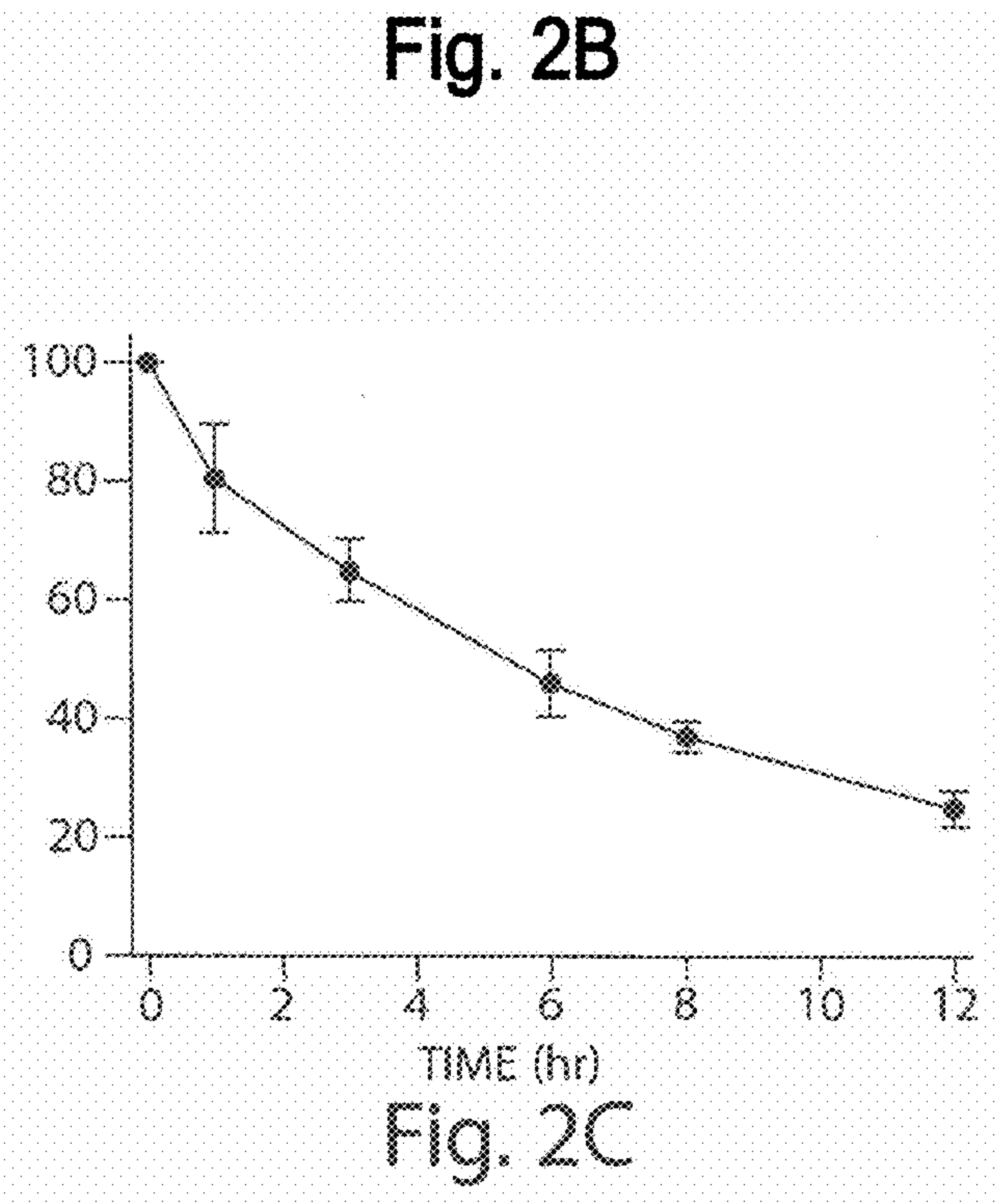
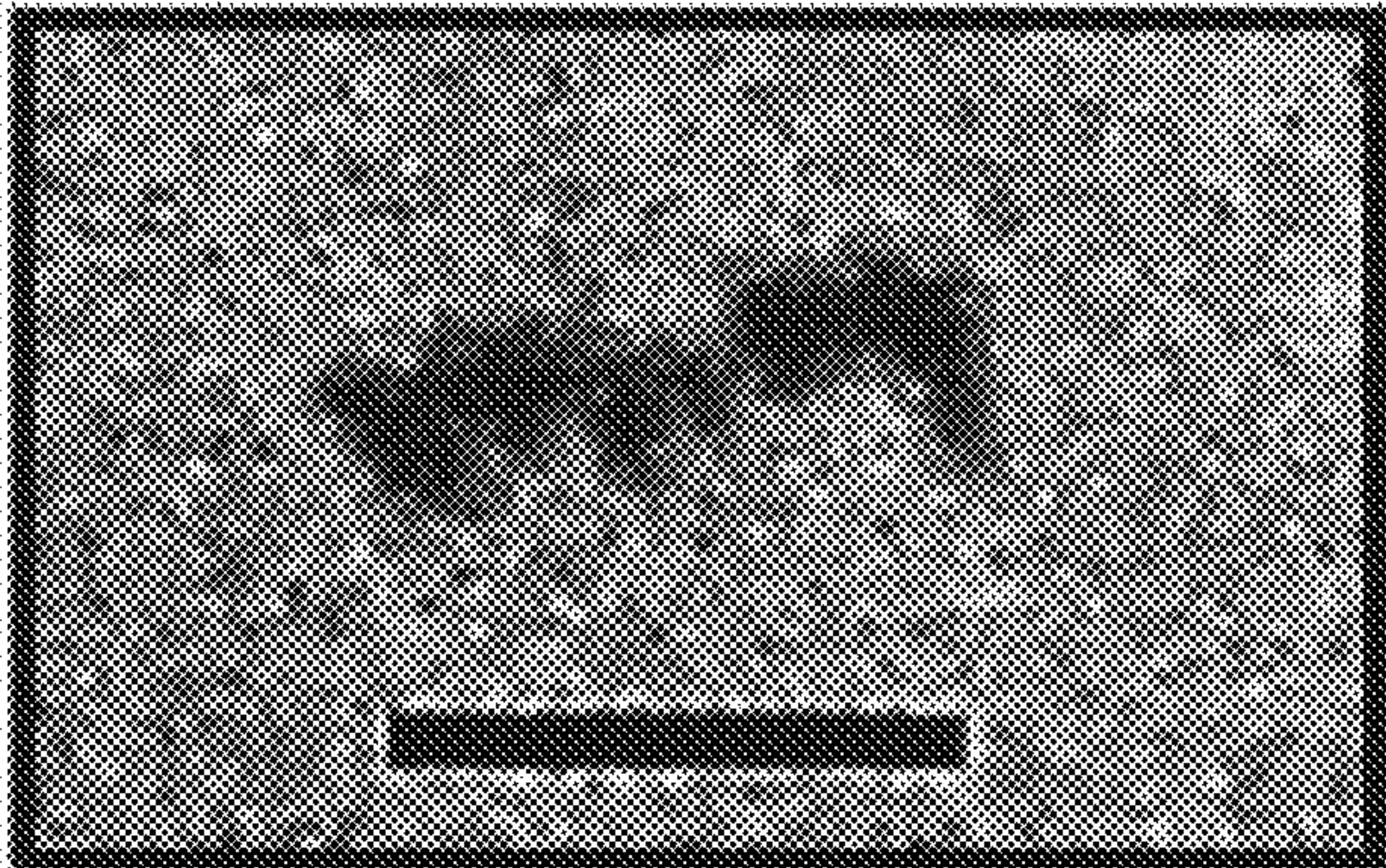
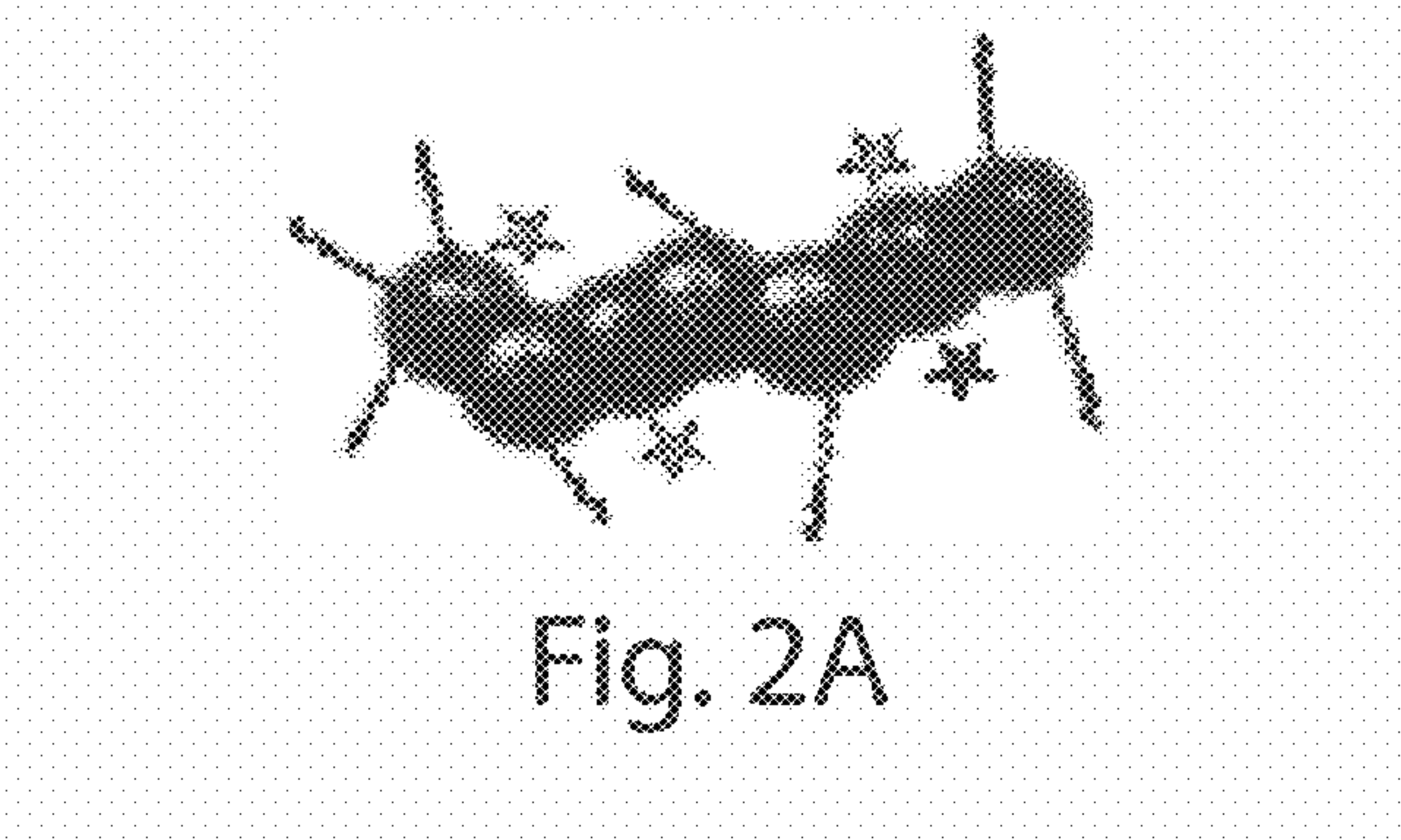


Fig. 1B



MULTIPLEXED LC-MS-MS
DETECTION OF
MASS-ENCODED FRAGMENTS

Fig. 1C



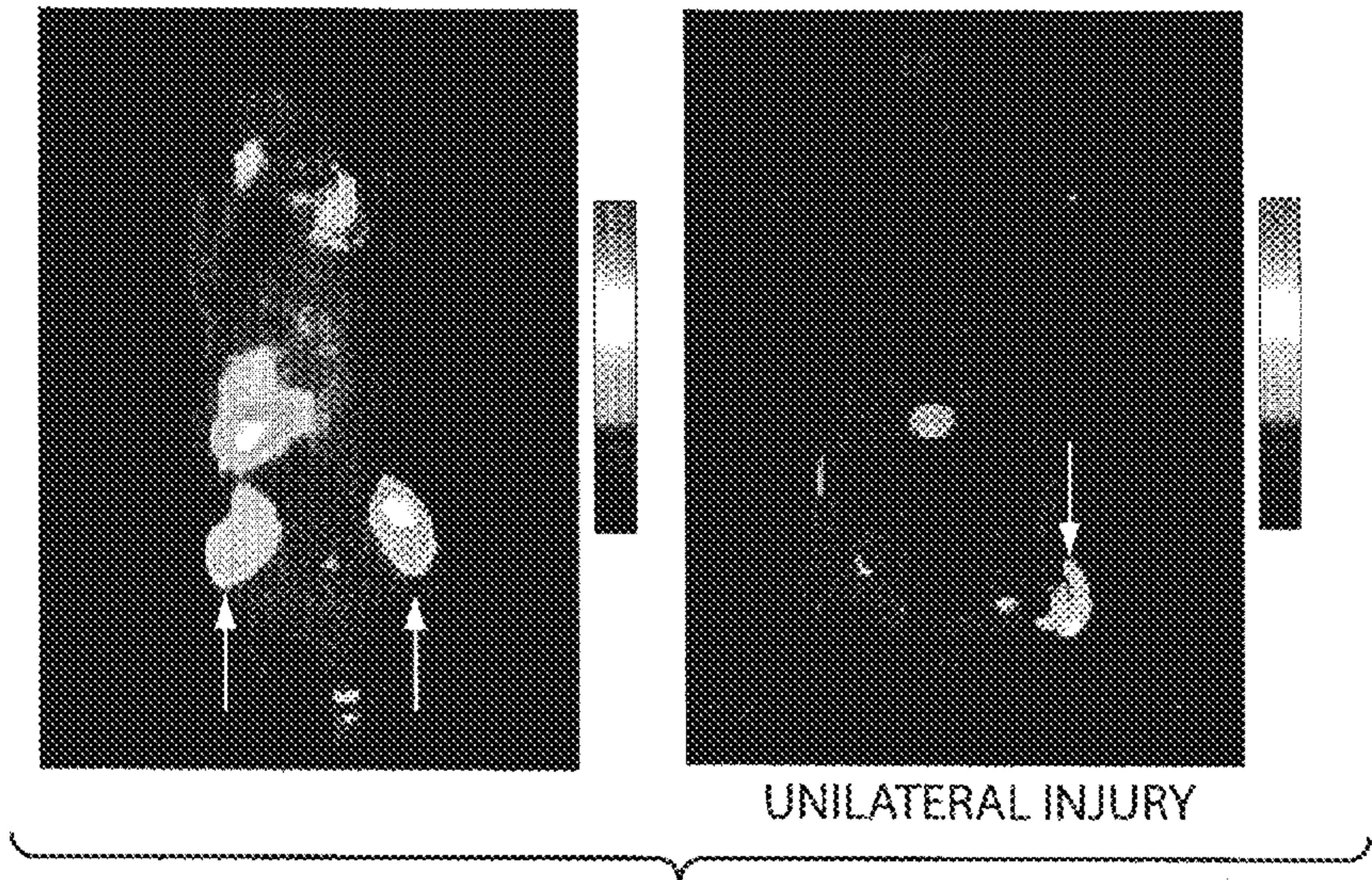


Fig. 2D

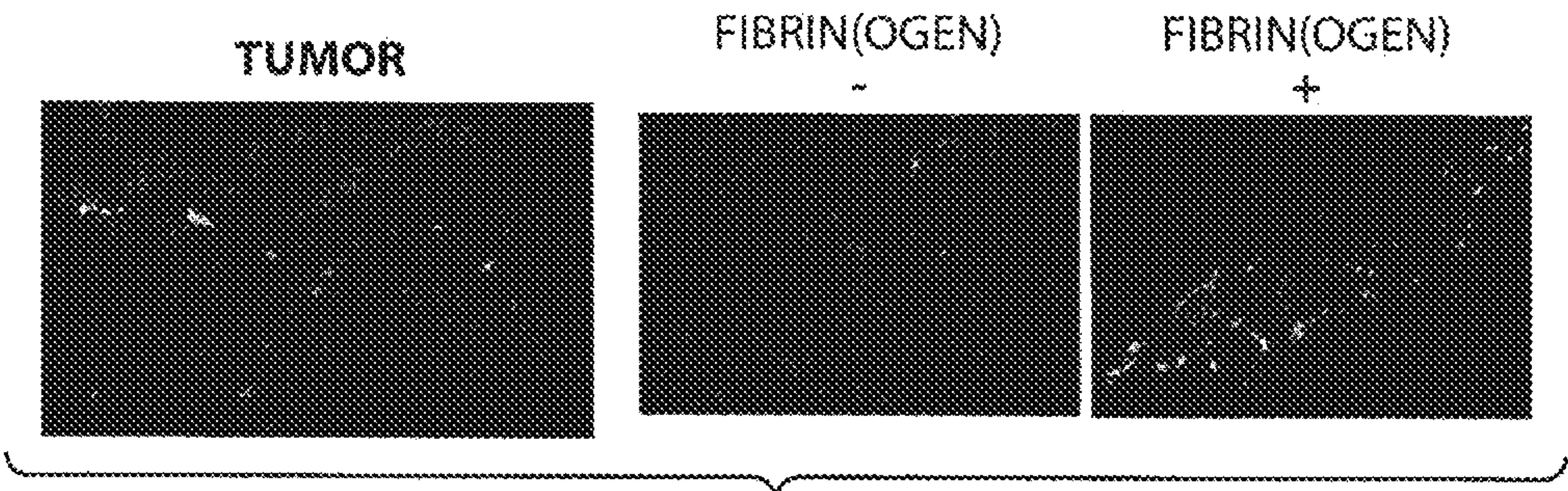


Fig. 2E

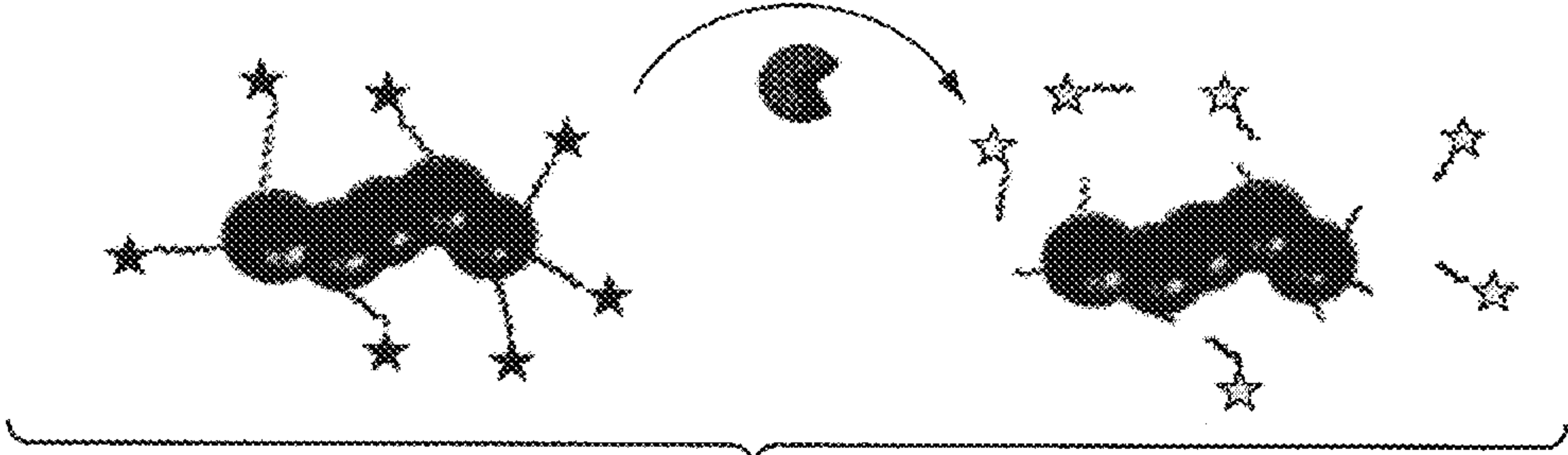


Fig. 3A

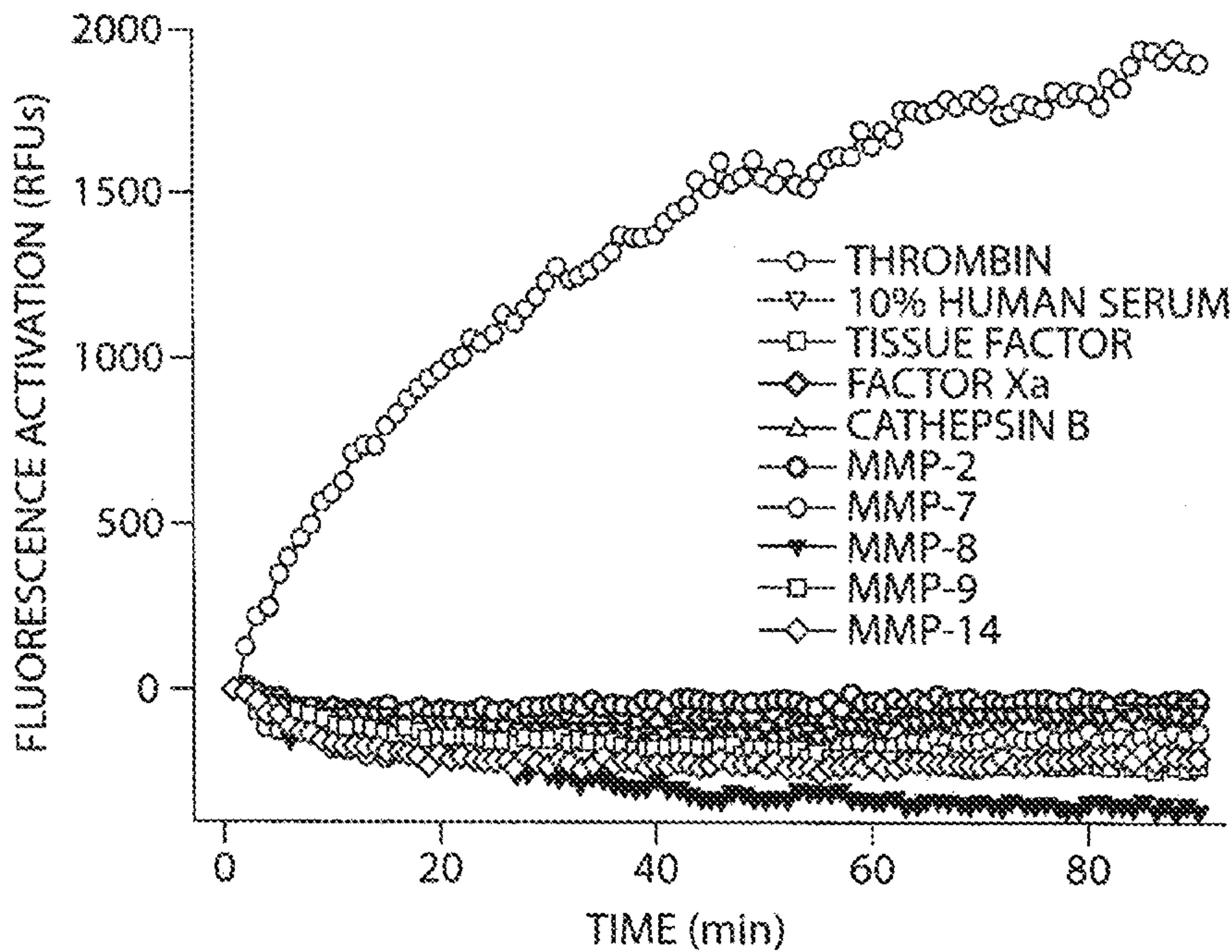
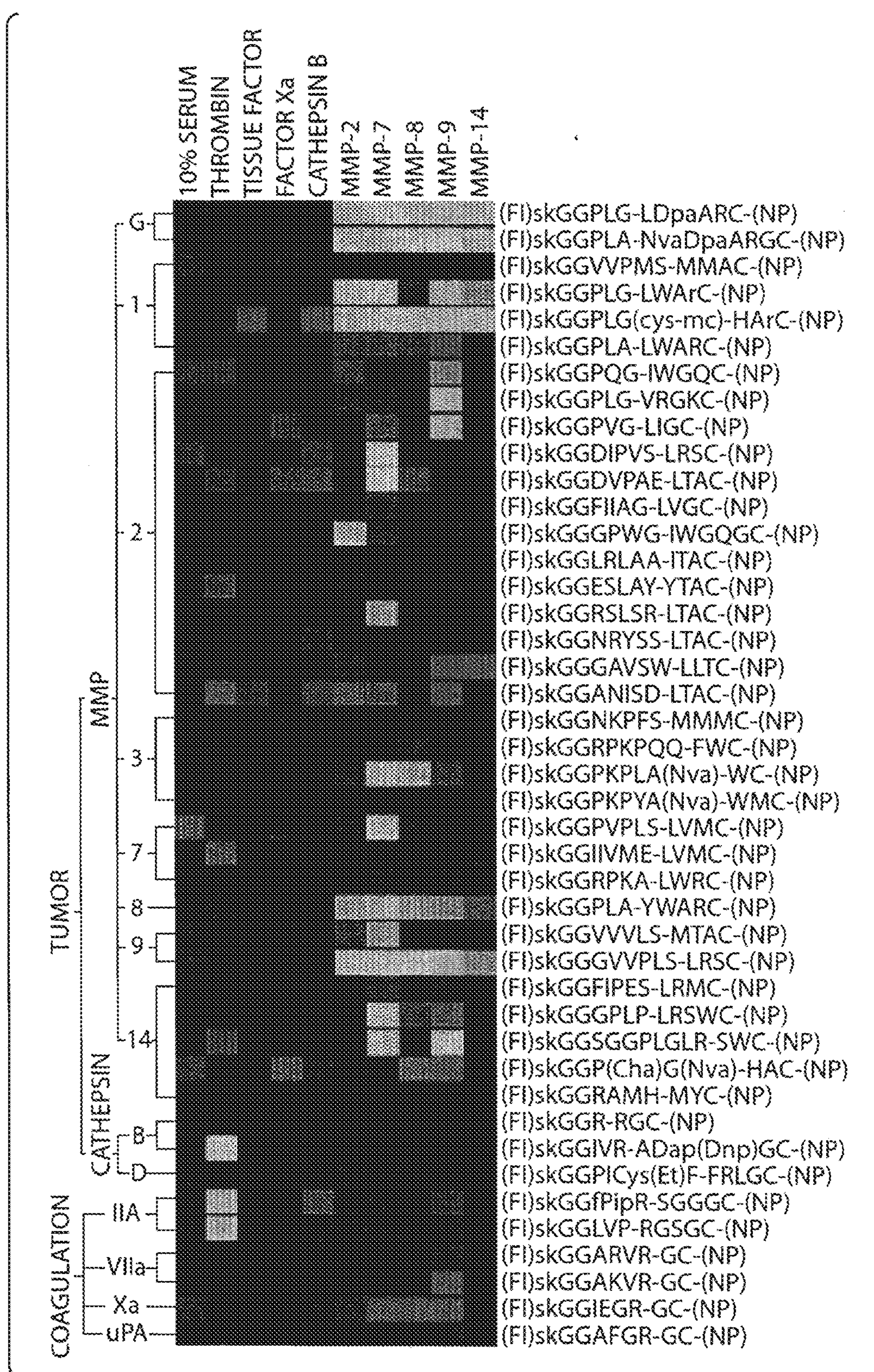


Fig. 3B



LC/MS/MS DETECTION OF PROTEOLYTIC REPORTERS

FLUOR-(PEPTIDE SEQUENCE)-NP	EXPECTED CLEAVAGE PRODUCT*	OBSERVED MAJOR CLEAVAGE PRODUCT*	EXPECTED MASS	OBSERVED MASS	PROTEASE ACTIVATION (MS)**	PROTEASE ACTIVATION (FLUORESCENCE)**
Fluor-dS-dK-G-G-P-Q-G-I-W-G-Q-C-(NP)	GGPQG	GGPQG	987.4	987.4	2, 7, 8, 9	2, 7, 8, 9
Fluor-dS-dK-G-G-P-L-G-V-R-G-K-C-(NP)	GGPLG	GGPLG	972.4	972.4	2, 7, 8, 9, th	2, 7, 9
Fluor-dS-dK-G-G-P-L-A-Nva-Dpa-A-R-G-C-(NP)	GGPLA	GGPLA	986.4	986.4	2, 7, 8, 9	2, 7, 8, 9, 14, th
Fluor-dS-dK-G-G-P-V-G-L-I-G-C-(NP)	GGPLG	GGPLG	958.4	958.4	2, 7, 8, 9, 14	7, 9
Fluor-dS-dK-G-G-P-V-P-L-S-L-V-M-C-(NP)	GGVPLS	GGVPLS	1198.7	998.4	7, 8, 9, th	7, 8
Fluor-dS-dK-G-G-S-G-P-L-G-L-R-S-W-C-(NP)	GGSGPLG	GGSGPLG	1173.5	1173.5	7, 9, 14, th	7, 8, 9, th
Fluor-dS-dK-G-G-P-W-G-I-W-G-Q-G-C-(NP)	GGGPWG	GGGPWG	1102.4	1102.4	2, 7, 8, 9	2, 7, 8, 9
Fluor-dS-dK-G-G-dF-Pip-R-S-G-G-C-(NP)	GGdFripR	GGdFripR	1119.5	1119.5	th	th
Fluor-dS-dK-G-G-L-V-P-R-G-S-G-C-(NP)	GGLVP	GGLVP	1170.5	1170.5	th	th
* ALL PEPTIDES HAVE FLUOR-dS-dK-(VARIABLE GROUP), ONLY THE [VARIABLE GROUPS] ARE SHOWN IN THESE COLUMNS						
**2 = MMP-2, 7 = MMP-7, 8 = MMP-8, 9 = MMP-9, 14 = MMP-14, th = THROMBIN						

Fig. 4

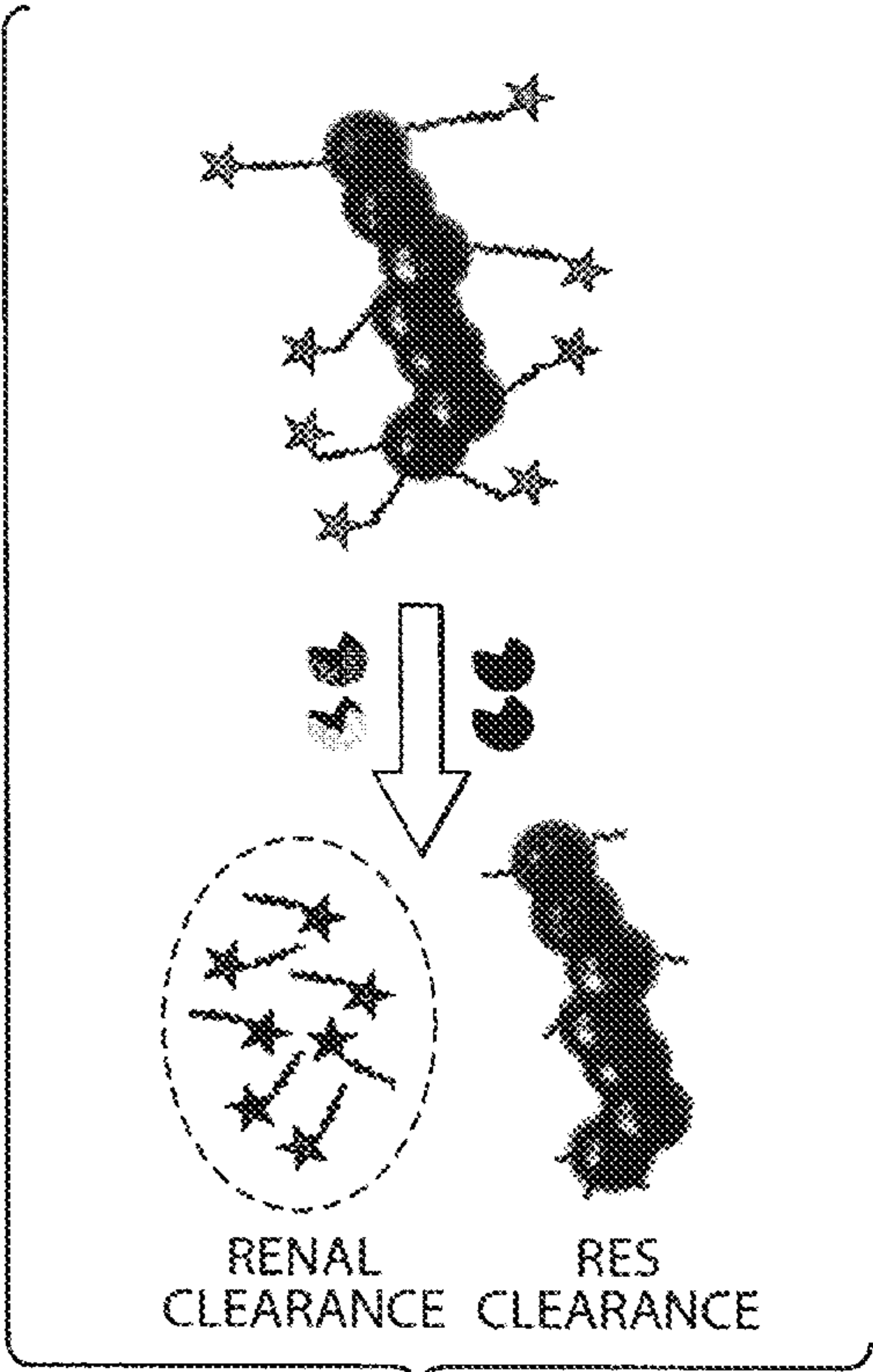


Fig. 5A

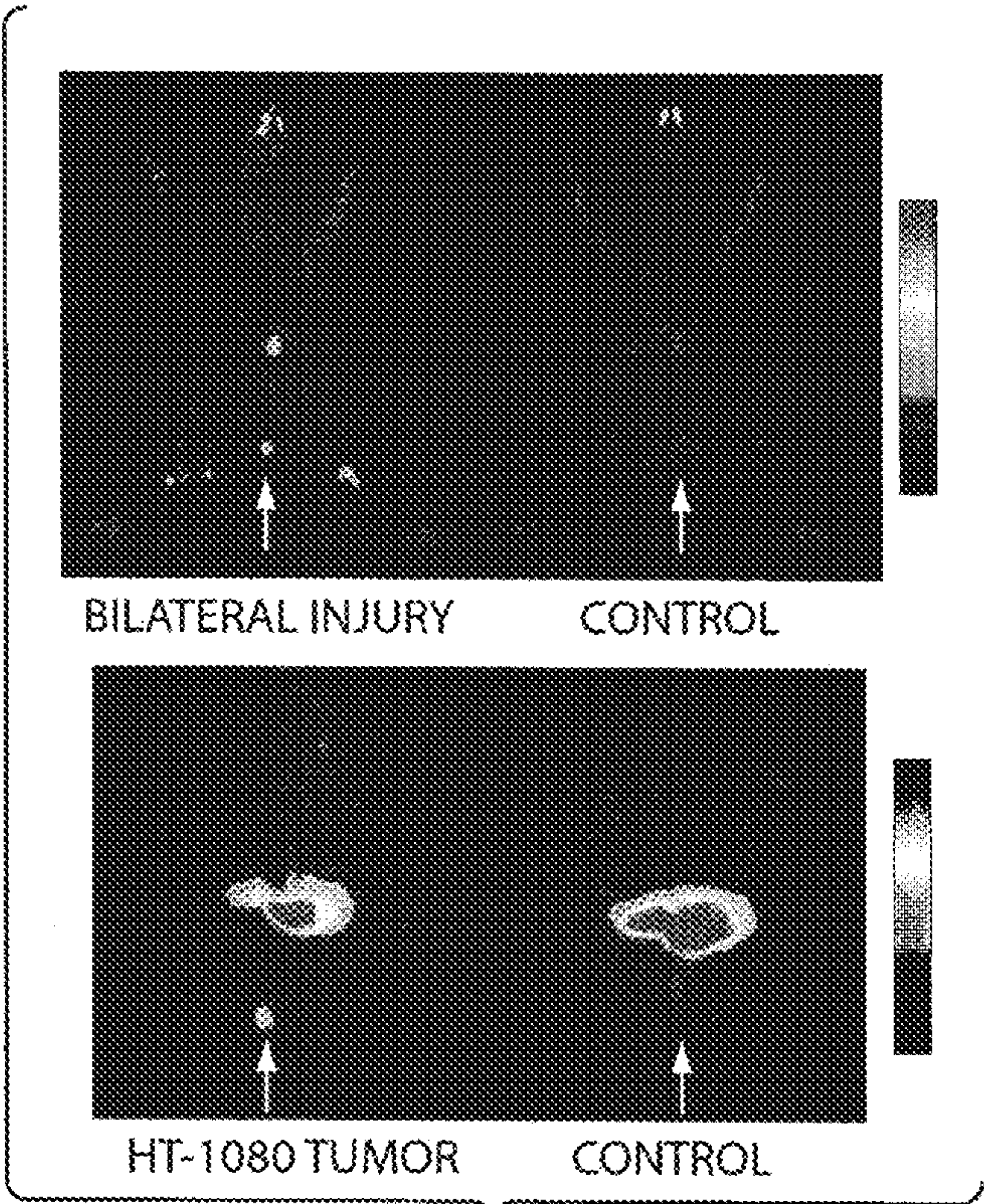


Fig. 5B

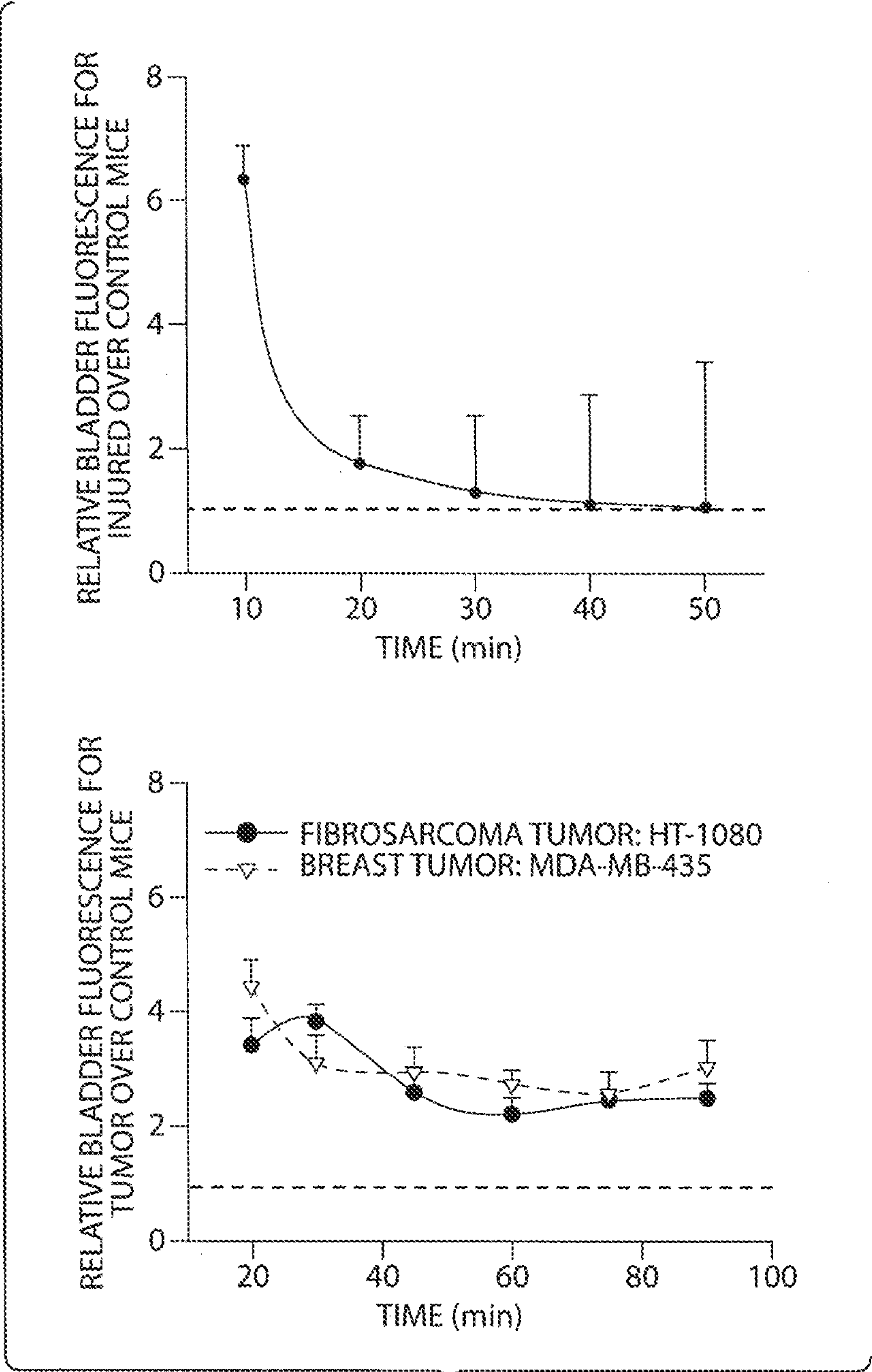


Fig. 5C

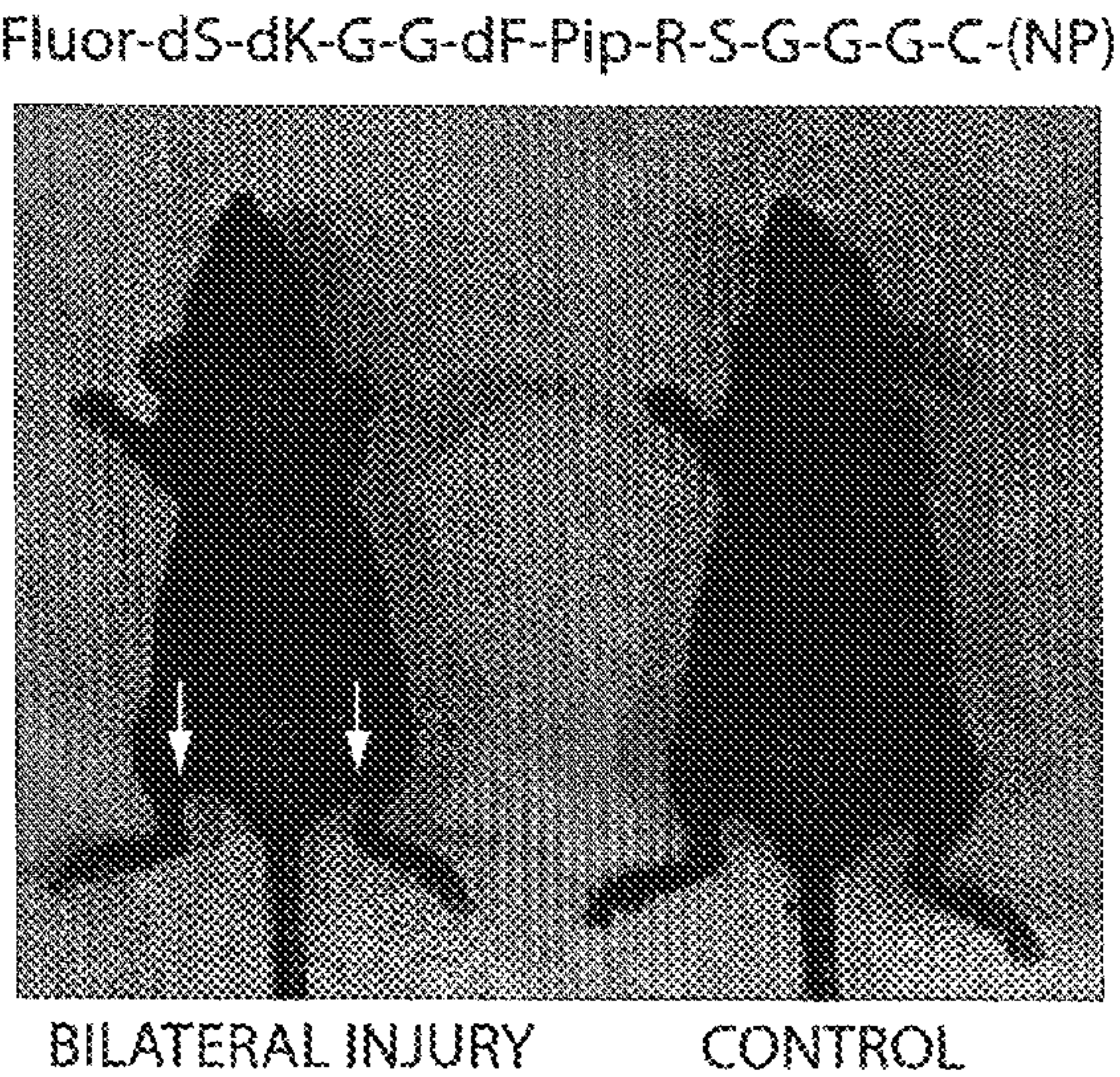


Fig. 6A

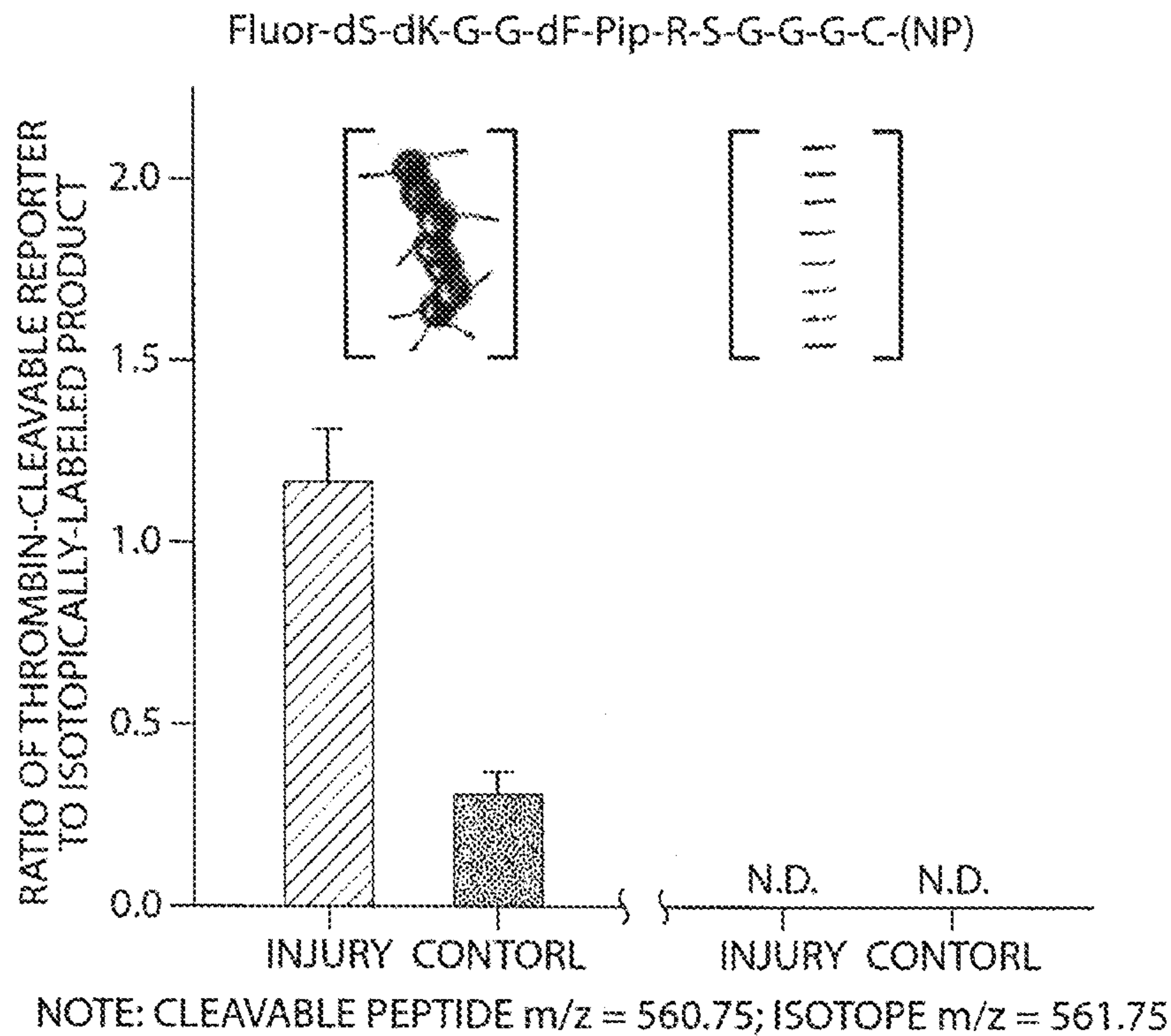


Fig. 6B

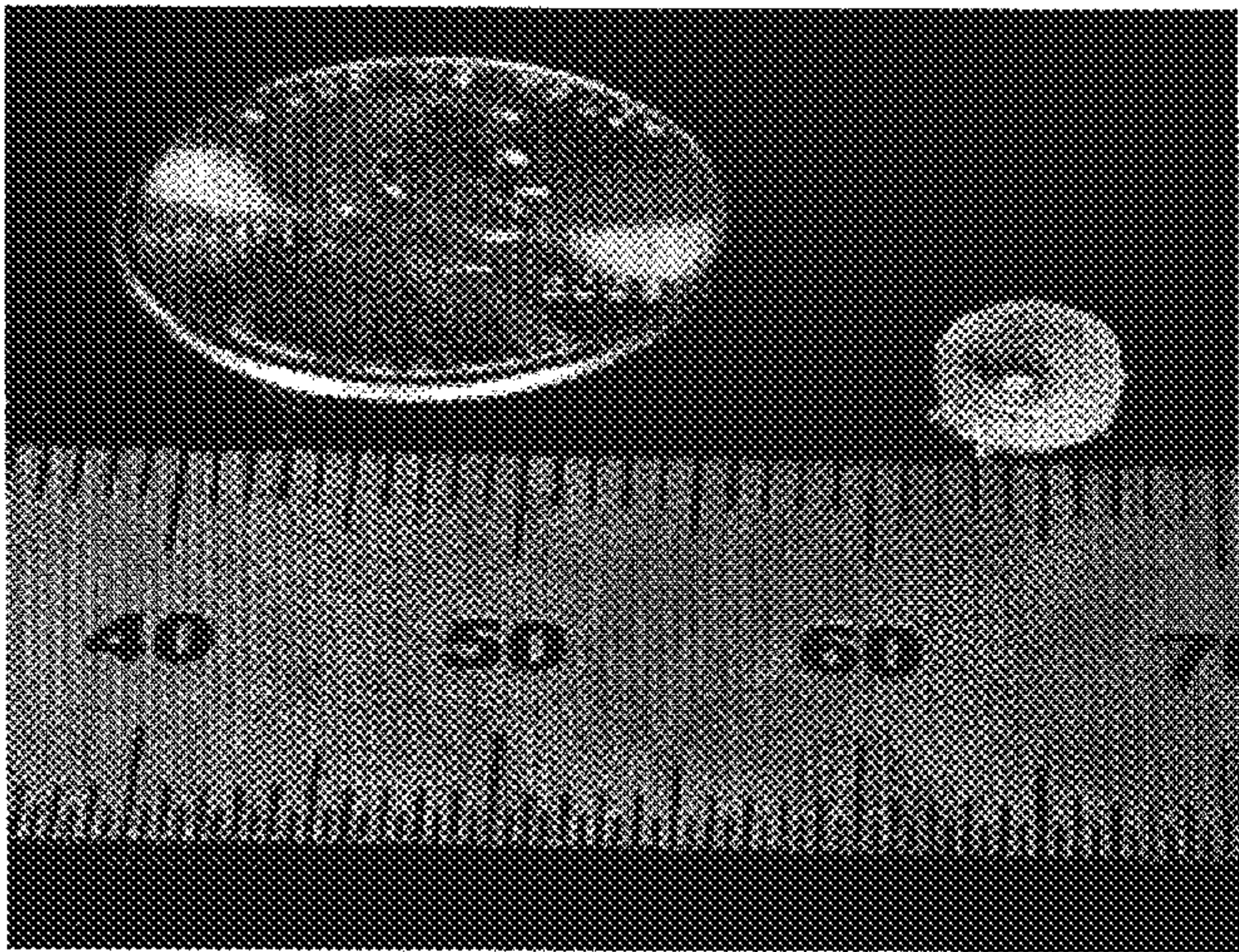


Fig. 7A

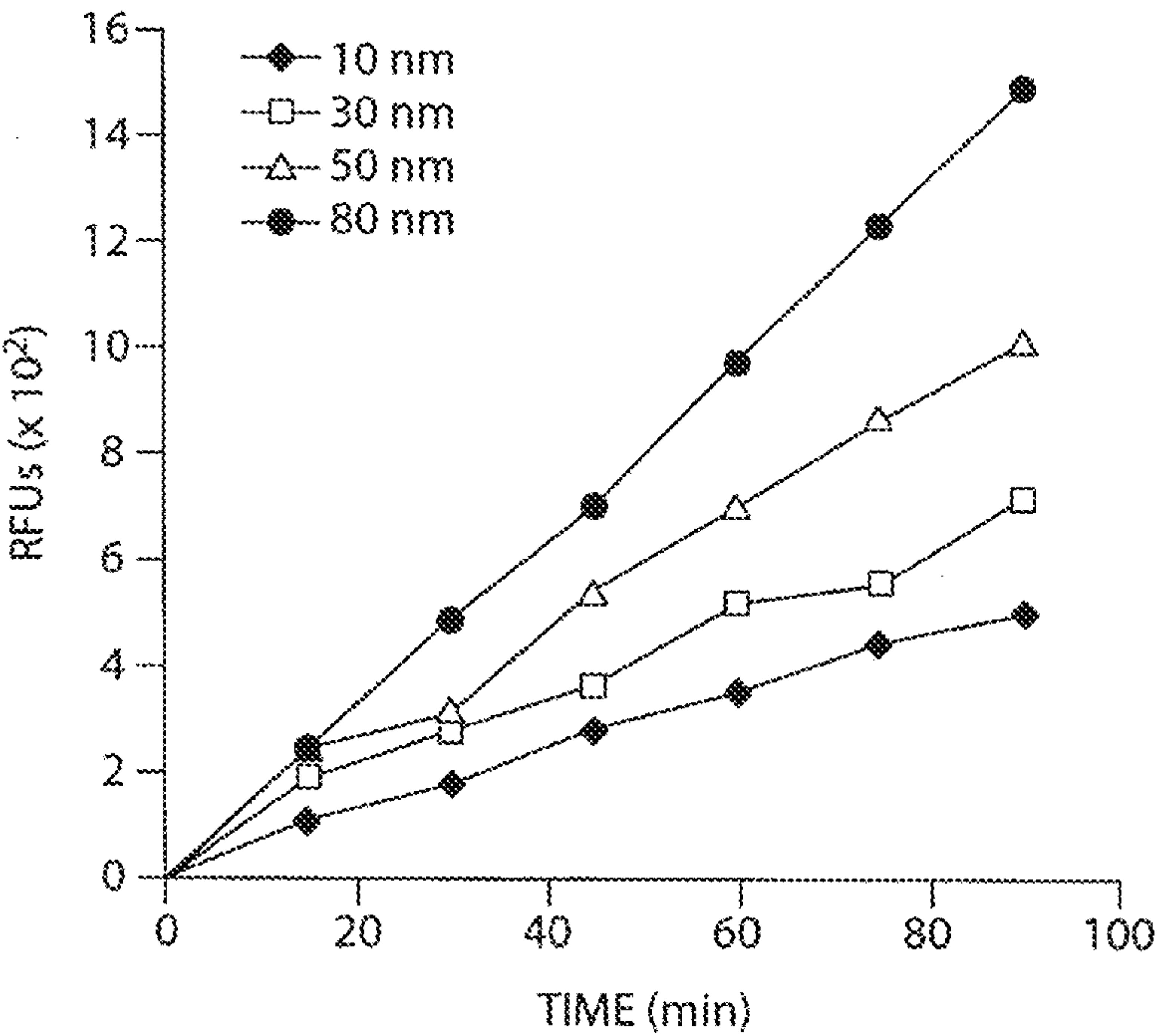


Fig. 7B

PEPTIDE	SEQUENCE
A ₁	FI-dR-G-G-dS-G-G-dS-G-G-dS-dR-V-V-V-L-S
A ₂	FI-dR-G-G-dS-G-G-dS-dR-G-G-V-V-V-L-S
A ₃	FI-dR-G-G-dS-dR-G-G-dS-G-G-V-V-V-L-S
A ₄	FI-dR-G-G-dS-dR-G-G-V-V-V-L-S
A ₅	FI-dR-G-dS-dR-G-G-V-V-V-L-S
A ₆	FI-dR-dS-dR-G-G-V-V-V-L-S
A ₇	FI-dR-G-G-dS-G-G-dS-G-G-dS-dR-P-V-G
A ₈	FI-dR-G-G-dS-G-G-dS-dR-G-G-P-V-G
A ₉	FI-dR-G-G-dS-dR-G-G-dS-G-G-P-V-G
A ₁₀	FI-dR-G-G-dS-dR-G-G-P-V-G
A ₁₁	FI-dR-G-dS-dR-G-G-P-V-G
A ₁₂	FI-dR-dS-dR-G-G-P-V-G
A ₁₃	FI-dR-G-G-dS-G-G-dS-G-G-dF-Pip-R
A ₁₄	FI-dR-G-G-dS-G-G-dF-Pip-R

Fig. 8A

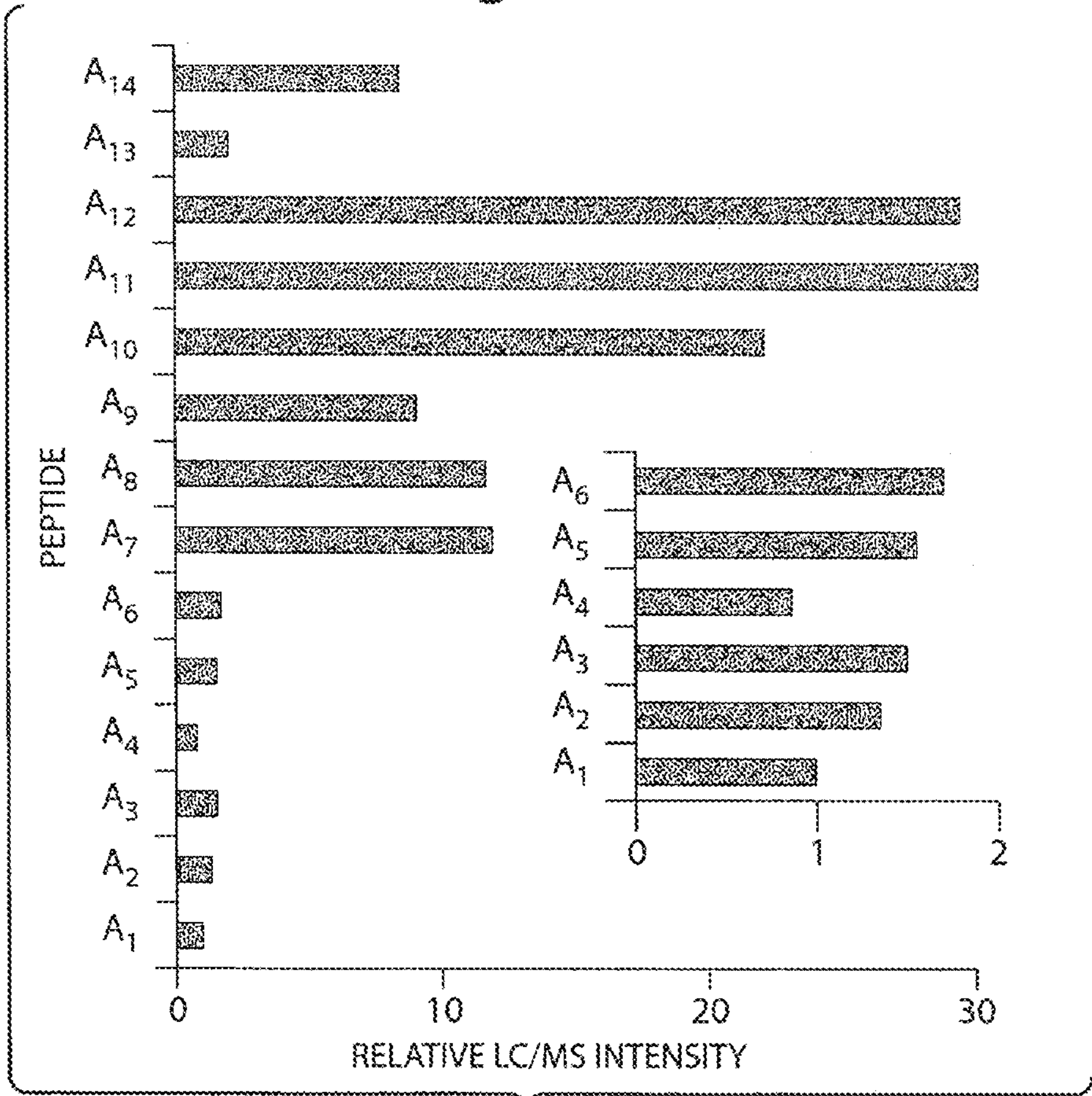


Fig. 8B

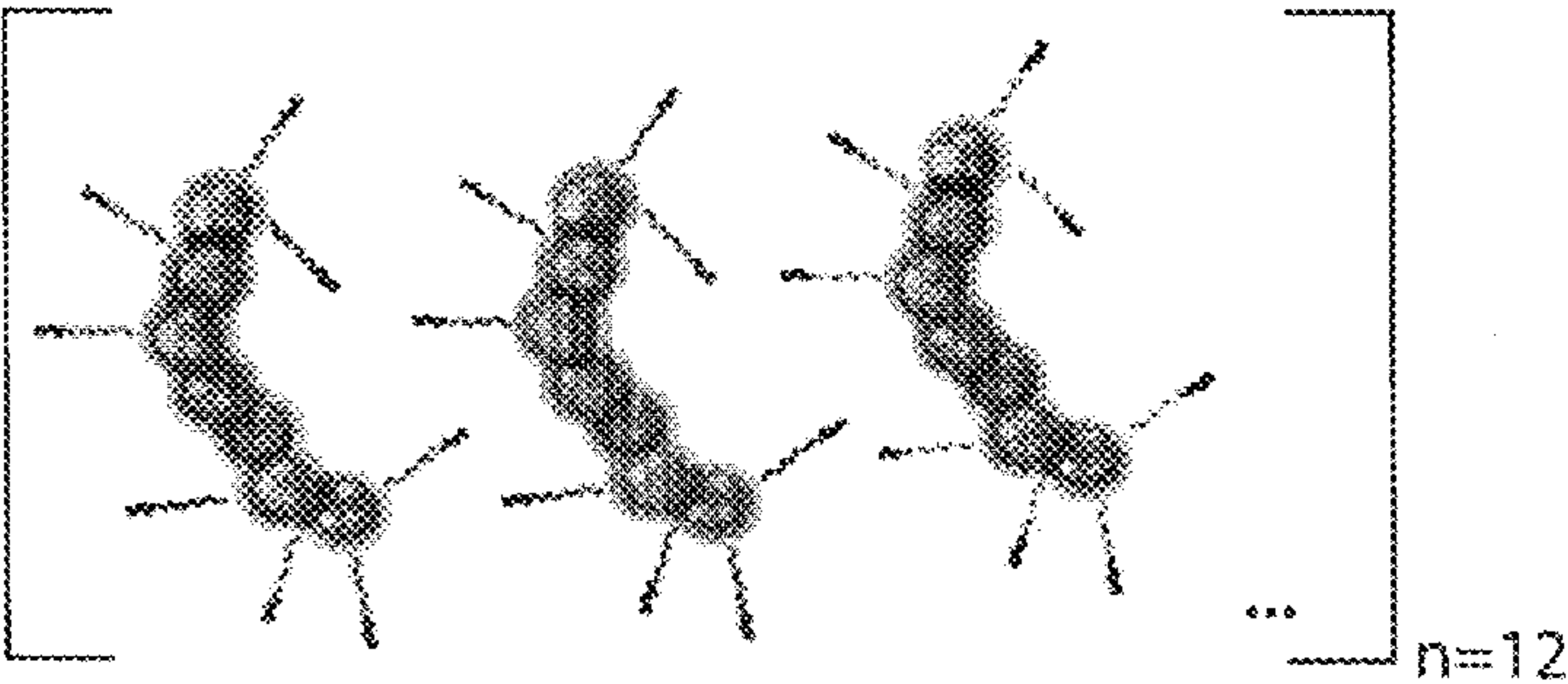


Fig. 9A

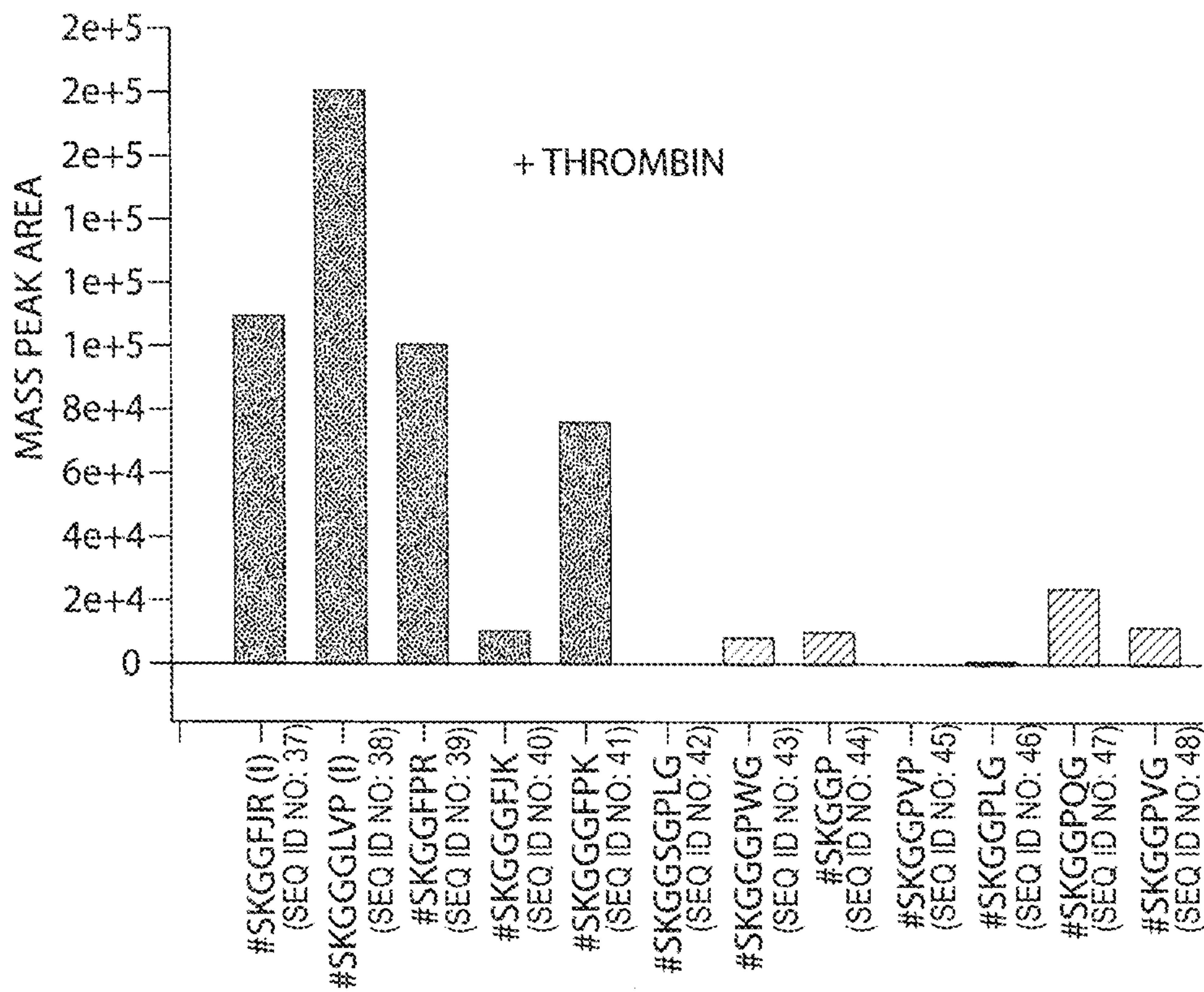


Fig. 9B

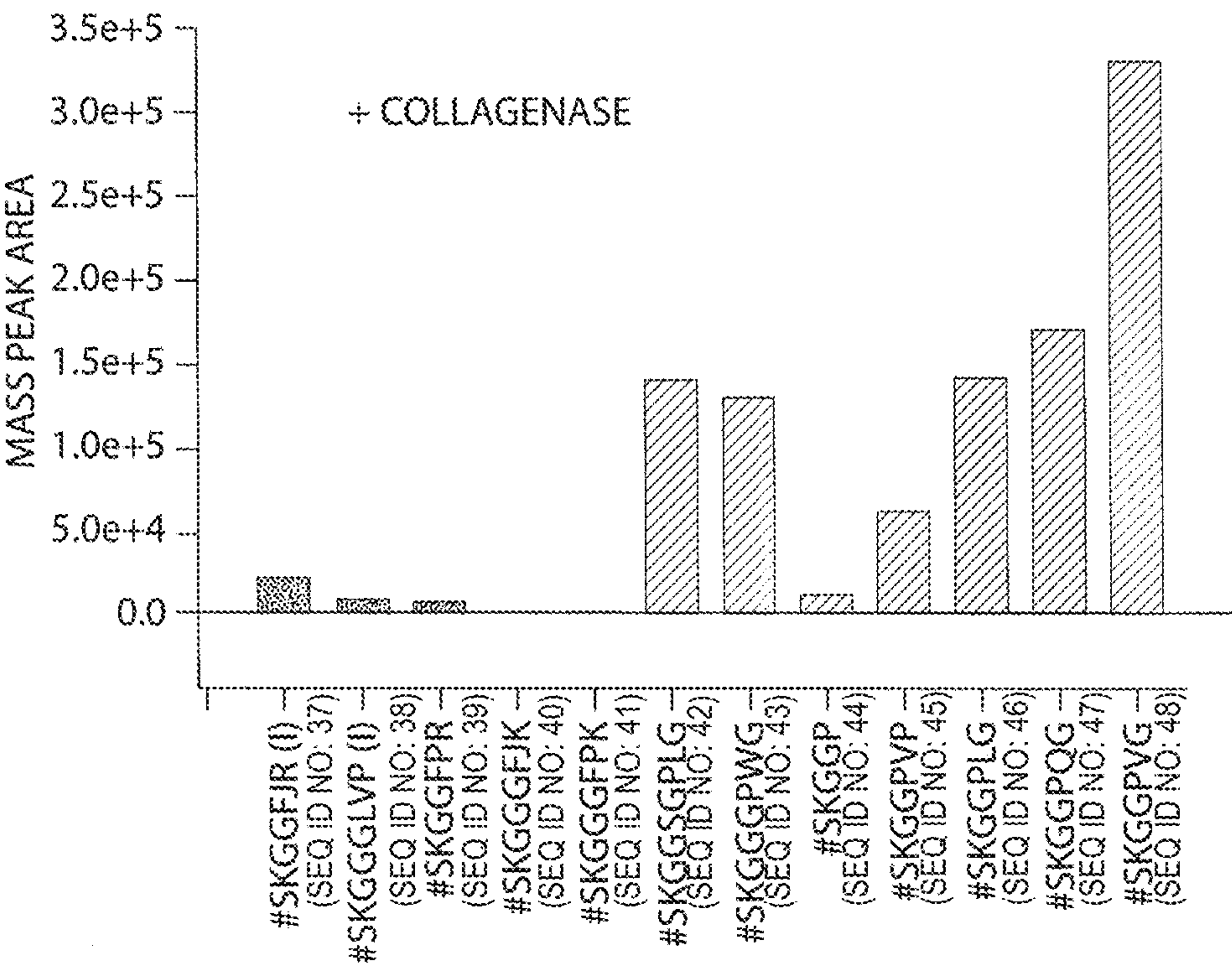


Fig. 9C

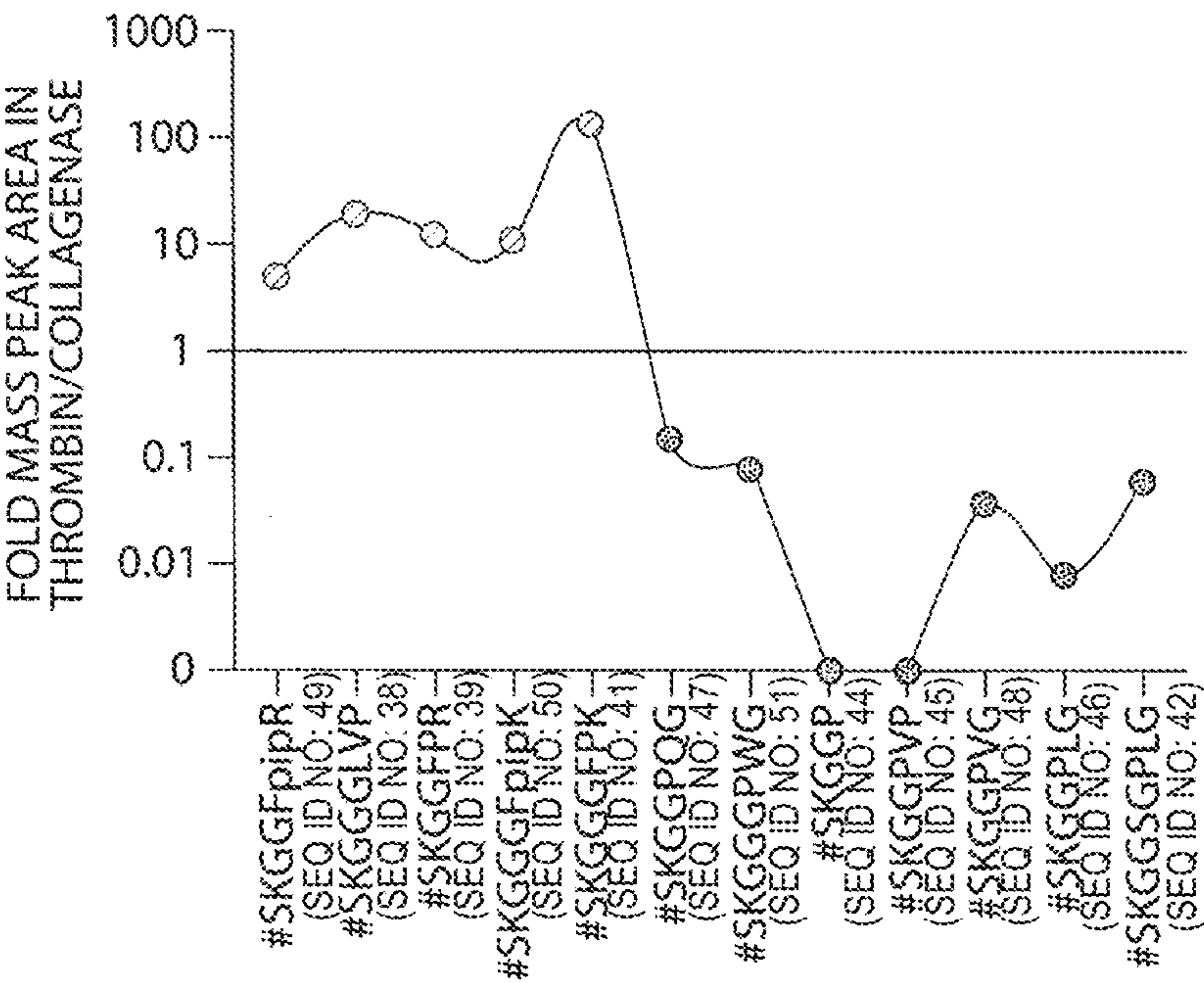


Fig. 9D

METHODS AND PRODUCTS FOR IN VIVO ENZYME PROFILING

RELATED APPLICATIONS

[0001] This application is a continuation of U.S. application Ser. No. 17/570,847, filed Jan. 7, 2022, entitled “METHODS AND PRODUCTS FOR IN VIVO ENZYME PROFILING”, which is a continuation of U.S. application Ser. No. 17/091,075, filed Nov. 6, 2020, entitled “METHODS AND PRODUCTS FOR IN VIVO ENZYME PROFILING”, which is a continuation of U.S. application Ser. No. 15/966,385, filed Apr. 30, 2018, entitled “METHODS AND PRODUCTS FOR IN VIVO ENZYME PROFILING”, now issued as U.S. Pat. No. 10,883,998 on Jan. 5, 2021, which is a divisional of U.S. application Ser. No. 14/166,481, filed Jan. 28, 2014, entitled “METHODS AND PRODUCTS FOR IN VIVO ENZYME PROFILING”, now issued as U.S. Pat. No. 9,970,941 on May 15, 2018, which is a continuation of U.S. application Ser. No. 12/715,965, filed on Mar. 2, 2010, entitled “METHODS AND PRODUCTS FOR IN VIVO ENZYME PROFILING,” now issued as U.S. Pat. No. 8,673,267 on Mar. 18, 2014, which claims the benefit under 35 U.S.C. § 119(e) of U.S. Provisional Application No. 61/156,660, filed on Mar. 2, 2009, entitled “METHODS AND PRODUCTS FOR IN VIVO ENZYME PROFILING,” the entire contents of each of which are incorporated herein by reference.

GOVERNMENT SUPPORT

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

REFERENCE TO AN ELECTRONIC SEQUENCE LISTING

[0003] The contents of the electronic sequence listing (M065670181US06-SEQ-FL.xml; Size: 84,091 bytes; and Date of Creation: Nov. 15, 2022) is herein incorporated by reference in its entirety.

FIELD OF THE INVENTION

[0004] The present invention relates to methods and products associated with in vivo enzyme profiling. Some aspects of the present invention relate to profiling enzymatic reaction products. In particular, the invention relates to methods of in vivo processing of exogenous molecules followed by detection of signature molecules as representative of the presence or absence of active enzymes associated with disease or conditions. The invention also relates to products, kits, and databases for use in the methods of the invention.

BACKGROUND OF THE INVENTION

[0005] Dysregulation of proteases in cancer has important consequences in cell signaling and helps drive cancer cell proliferation, invasion, angiogenesis, avoidance of apoptosis, and metastasis. Currently, in vivo analysis of proteases (and other enzymes such as glycosidase, esterase, etc.) activity is limited to biopsy or local fluorescent techniques, which are hindered by their invasiveness or low multiplexing potential, respectively.

SUMMARY OF THE INVENTION

[0006] The invention in some aspects is a method involving administering to a subject a pro-diagnostic reagent, wherein the pro-diagnostic reagent comprises a modular structure having a carrier domain linked to a signature producing domain, wherein the signature producing domain is capable of producing a signature molecule in the subject; identifying a biological sample for detection of the signature molecule, wherein the biological sample is at a site remote from the production of the signature molecule; and, subjecting the biological sample to an analysis method in order to detect the presence of the signature molecule, wherein the presence of the signature molecule in the biological sample is indicative of a biological predictor molecule within the subject.

[0007] In one embodiment the signature producing domain comprises an enzyme susceptible domain linked to a signature molecule, wherein the biological predictor molecule is an enzyme, wherein the enzyme susceptible domain is susceptible to modification by the enzyme in the subject, and wherein the presence of the signature molecule in the biological sample is indicative of an active enzyme within the subject.

[0008] In another embodiment the signature producing domain comprises an active signature producing agent, wherein the active signature producing agent is capable of modifying the biological predictor molecule to produce the signature molecule in the subject. The active signature producing agent may be an enzyme, such as a protease or a glycosidase.

[0009] In other embodiments the pro-diagnostic reagent further comprises an implantable microdelivery device that houses the modular structure. The implantable microdelivery device in some embodiments is an implantable capsule with a semi-permeable membrane that encapsulates the modular structure. In other embodiments the implantable microdelivery device is a chip having the modular structure attached thereto. In yet other embodiments the implantable microdelivery device is a sustained-release formulation.

[0010] In some aspects the invention is a method involving administering to a subject a pro-diagnostic reagent, wherein the pro-diagnostic reagent comprises a carrier domain linked to an enzyme susceptible domain which is linked to a signature molecule, wherein the enzyme susceptible domain is susceptible to modification by an enzyme in the subject; identifying a biological sample for detection of the signature molecule, wherein the biological sample is at a site remote from the enzyme; and, subjecting the biological sample to an analysis method in order to detect the presence of one or more signature molecules, wherein the presence of the signature molecule in the biological sample is indicative of an active enzyme within the subject.

[0011] In other aspects of the invention a method of administering to a subject a pro-diagnostic reagent, wherein the pro-diagnostic reagent comprises a carrier domain linked to an enzyme susceptible domain which is linked to a signature molecule; collecting a urine sample from the subject; and, subjecting the urine sample to an analysis method in order to detect the presence of the signature molecule, wherein the presence of the signature molecule in the biological sample is indicative of an active enzyme within the subject is provided.

[0012] In yet other aspects a method for diagnosing a disease is provided. The method involves administering to a

subject a pro-diagnostic reagent, wherein the pro-diagnostic reagent comprises a carrier domain linked to an enzyme susceptible domain which is linked to a signature molecule, and wherein the enzyme susceptible domain is susceptible to cleavage by an enzyme associated with a disease; collecting a urine sample from the subject; and, subjecting the urine sample to an analysis method in order to detect the presence of the signature molecule, wherein the presence of the signature molecule in the biological sample is indicative of the subject having the disease.

[0013] In another aspect of the invention a method of collecting a urine sample from a subject suspected of having a disorder or condition, wherein the subject has been administered a pro-diagnostic reagent, the pro-diagnostic reagent comprising a carrier domain linked to an enzyme susceptible domain which is linked to a signature molecule; and, subjecting the urine sample to a multiplex analysis method in order to detect the presence of the signature molecule, wherein the presence of the signature molecule in the biological sample is indicative of the disorder or condition within the subject. is provided In some embodiments, the subject is a healthy subject. In some embodiments, the subject is a subject at risk of developing a disease or condition. In some embodiments, the subject is suspected of having a disease or condition or a subject diagnosed with having a disease or condition.

[0014] A method of collecting a urine sample from a subject suspected of having a disorder or condition, wherein the subject has been administered a pro-diagnostic reagent, the pro-diagnostic reagent comprising a carrier domain linked to an enzyme susceptible domain which is linked to a signature molecule and, subjecting the urine sample to a multiplex analysis method in order to detect the presence or absence of the signature molecule, wherein the absence of the signature molecule in the biological sample is indicative of the disorder or condition within the subject is provided according to other aspects of the invention.

[0015] A method of treating a subject is provided according to an aspect of the invention. The method involves collecting a urine sample from a subject suspected of having a disorder or condition or diagnosed with a disorder or condition, wherein the subject has been administered a pro-diagnostic reagent, the pro-diagnostic reagent comprising a carrier domain linked to an enzyme susceptible domain which is linked to a signature molecule; subjecting the urine sample to a multiplex analysis method in order to detect the presence of the signature molecule, wherein the presence of the signature molecule in the biological sample is indicative of the disorder or condition within the subject; and, administering a therapeutic agent to the subject to treat the disorder.

[0016] In some embodiments a further step of collecting a biological sample from the subject is provided. In other embodiments the signature molecule is detected in the biological sample in the subject. The biological sample may be urine, blood, saliva, or mucous secretion.

[0017] A plurality of pro-diagnostic reagents having a plurality of signature molecules may be administered to the subject in some embodiments. The plurality of pro-diagnostic reagents may have a plurality of signature molecules. In other embodiments the pro-diagnostic reagent includes a plurality of signature molecules.

[0018] In some embodiments the enzyme susceptible domain is susceptible to modification, i.e. cleavage, addi-

tion, conformational or charge change, by an enzyme associated with a disease or condition. In some embodiments the enzyme susceptible domain is susceptible to cleavage by a protease associated with a disease or condition. The enzyme susceptible domain in other embodiments is susceptible to modification by an enzyme not associated with a disease or condition, but associated with a normal condition.

[0019] In some embodiments the enzyme susceptible domain is a peptide, such as, for instance, a MMP sensitive site, a kallikrein sensitive site, a cathepsin sensitive site, a plasminogen activator sensitive site and/or an ADAM sensitive site.

[0020] In some embodiments the disease or condition is cancer, cardiovascular disease, arthritis, viral, bacterial, parasitic or fungal infection, Alzheimer's disease emphysema, thrombosis, hemophilia, stroke, organ dysfunction, any inflammatory condition, vascular disease, parenchymal disease, or a pharmacologically-induced state.

[0021] In some embodiments, the carrier domain comprises a particle, for example, a microparticle or a nanoparticle. The carrier domain is greater than 5 nm in size in some embodiments and in other embodiments is smaller than 5 nm in size. In some embodiments, the carrier domain comprises a targeting domain and/or a therapeutic agent. In some embodiments, the carrier domain selectively interacts with a molecular target, for example, a protein or peptide, a nucleic acid, or a carbohydrate. In some embodiments, the carrier domain selectively binds a molecular target. In some embodiments, the carrier domain selectively interacts with a target molecule as part of an enzymatic reaction, for example, an enzymatic reaction carried out by the carrier domain or by the target molecule. In some embodiments, the carrier domain comprises a peptide, a protein, a nucleic acid or a small molecule, for example, a peptide, protein, nucleic acid or small molecule selectively binding a molecular target, for example, a target molecule (e.g., a peptide, protein, nucleic acid, or carbohydrate) expressed in a target cell or cell type, after administration to a subject. In some embodiments the molecular target is specifically expressed in a target cell or target cell type, for example, a cancer cell or a cell of a certain differentiation state or of a certain tissue. In some embodiments, the carrier domain comprises a therapeutic agent. In some embodiments, the carrier domain comprises a therapeutic agent selectively interacting with a molecular target, for example, a molecular target expressed in a target cell or target cell type. In some embodiments the carrier domain is a nanoparticle, a peptide, for example, an RGD peptide, an aptamer, an antibody, or a fragment thereof, an adnectin, or a targeting molecule.

[0022] The signature molecule in some embodiments is a peptide, nucleic acid, small molecule, fluorophore/quencher, carbohydrate, particle, radiolabel, MRI-active compound, inorganic material, and/or organic material, with encoded characteristics to facilitate optimal detection.

[0023] The analysis step used in the methods may be a multiplex analysis method or a singular analysis method. The analysis methods include but are not limited to mass spectrometry, liquid chromatography-mass spectrometry, PCR analysis, DNA microarray, and fluorescence analysis.

[0024] In some embodiments the method also includes a purification step, wherein the signature molecule is isolated from other components in the biological sample. The purification step may be, for instance, affinity chromatography.

[0025] In other aspects of the invention a reagent is provided. The reagent includes a carrier domain, wherein the carrier domain is a particle and is greater than 5 nm in size; an enzyme susceptible domain linked to the carrier domain; and, a signature molecule linked to the enzyme susceptible domain, wherein the signature molecule is a peptide or nucleic acid.

[0026] In other aspects, the invention is a reagent having an implantable microdelivery device housing a modular structure having a carrier domain linked to a signature producing domain.

[0027] In some embodiments the signature producing domain comprises an active signature producing agent, wherein the active signature producing agent is capable of modifying a biological predictor molecule to produce a signature molecule. The active signature producing agent may be an enzyme such as a protease or a glycosidase. In other embodiments the implantable microdelivery device is an implantable capsule with a semi-permeable membrane that encapsulates the modular structure, a chip having the modular structure attached thereto, or a sustained-release formulation.

[0028] In other aspects the invention is a reagent including a carrier domain having a plurality of enzyme susceptible domains linked to the carrier domain wherein each enzyme susceptible domain is linked to a non-fluorescent signature molecule.

[0029] In yet other aspects the invention is a composition having a plurality of pro-diagnostic reagents comprising a carrier domain, an enzyme susceptible domain linked to the carrier domain; and, a non-fluorescent signature molecule linked to the enzyme susceptible domain.

[0030] In some embodiments the carrier domain is polymer based microparticle, an iron oxide microparticle, or nanoparticle, an inorganic carrier, or an organic carrier. The carrier domain optionally includes a targeting domain and/or a therapeutic agent. The targeting domain may be, for instance, an antibody.

[0031] In some embodiments the enzyme susceptible domain is a peptide, such as for instance, GGPQGIWGQC (SEQ ID NO: 1), GGPLGVRGKC (SEQ ID NO: 2), GGPLANvaDpaARGC (SEQ ID NO: 3), GGPVGLIGL (SEQ ID NO: 4), GGPVPLSLVMC (SEQ ID NO: 5), GGSGGPLGLRSWC (SEQ ID NO: 6), GGGPWGIWGQC (SEQ ID NO: 7), GGdFPipRSGGQC (SEQ ID NO: 8), or GGLVPRGSGC (SEQ ID NO: 9).

[0032] In other embodiments the signature molecule is a peptide, nucleic acid, small molecule, fluorophore/quencher, carbohydrate, or particle. The signature molecule in some embodiments is a peptide of GGPQG (SEQ ID NO: 10), GGPLG (SEQ ID NO: 11), GGPLA (SEQ ID NO: 12), GGPVG (SEQ ID NO: 13), GGPVPLS (SEQ ID NO: 14), GGSGGPLG (SEQ ID NO: 15), GGGPWG (SEQ ID NO: 16), GGdFPipR (SEQ ID NO: 17), or GGLVP (SEQ ID NO: 18).

[0033] A kit is provided according to other aspects of the invention. The kit has a container housing a pro-diagnostic reagent, wherein the pro-diagnostic reagent comprises a carrier domain linked to an enzyme susceptible domain which is linked to a signature molecule; and, instructions for administering the pro-diagnostic reagent to a subject and for analyzing the signature molecule of the pro-diagnostic reagent in a biological sample of the subject.

[0034] In some embodiments the kit also includes a second container housing an analytical reagent. In other embodiments the kit also includes a box housing the containers. In yet other embodiments the kit includes a specimen collection device. Other embodiments of this invention would use a diversity of carriers, cleavage domains, and signature molecules to enable detection via modalities such as radiation, fluorescence, color, elemental detection, light scattering, magnetic techniques, MRI, electrical measurements, biochemical measurements, biological assays (including ELISA assays and others), among others.

[0035] 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.

BRIEF DESCRIPTION OF THE DRAWINGS

[0036] FIGS. 1A-1C are a schematic depicting a method according to the invention for multiplexed in vivo enzyme profiling of mass-coded nanoparticle based pro-diagnostic reagents.

[0037] FIGS. 2A-2E show data depicting the process of tumor and wound targeting with pro-diagnostic reagents. FIG. 2A is a schematic of the pro-diagnostic reagent, with the circles referring to the carrier, the star is a fluorescent molecule, and the zigzag line refers to the enzyme susceptible domain and the signature molecule (darker end region). FIG. 2B is an electron micrograph of the pro-diagnostic reagent. FIG. 2C is a graph depicting the circulation time of the pro-diagnostic reagent, by plotting detection of the carrier in the blood with respect to time after intravenous injection. FIG. 2D is photographs of mice having either tumors or injuries (left and right panels, respectively) administered the pro-diagnostic reagent. FIG. 2E is histopathological analysis of carrier homing to tumors or regions of injury.

[0038] FIG. 3A is a schematic of the pro-diagnostic reagent, with the circles referring to the carrier, the star is a signature molecule, and the zigzag line refers to the enzyme susceptible domain. FIG. 3B is a graph depicting fluorescence activation versus time. FIG. 3C depicts data on 43 pro-diagnostic reagents (with enzyme susceptible domains listed to the right for detection of tumor and injury enzymes).

[0039] FIG. 4 is a Table depicting the mass detection of ejected fragments in vitro. The results confirmed that the fluorescent results from the screen could also be detected by analyzing the mass of ejected fragments in vitro.

[0040] FIGS. 5A-5C depict the results of fluorescent detection of urinary reporter activation by tumors and injuries in vivo. FIG. 5A is a schematic of the pro-diagnostic reagent as shown in FIG. 3A, further depicting the portion of the molecule that undergoes renal clearance and the portion that undergoes RES clearance. FIG. 5B is a set of photographs of that were intravenously administered the optimized pro-diagnostic reagent for injury detection (top) or tumor detection (bottom). Half the mice that were administered the optimized pro-diagnostic agents for injury-detection suffered bilateral hind limb injuries (left side of the photograph) while the control mice had no injuries (right side of the photograph). Half the mice administered the optimized pro-diagnostic agents for tumor-detection harbored human fibrosarcoma tumors (HT-1080) (left side of photograph), while the other mice contained no tumors

(right side of photograph). FIG. 5C is a set of graphs depicting relative bladder fluorescence for tumor (bottom panel) or injured (top panel) versus control mice in order to track the entrance of cleaved signature peptide into the urine after injection.

[0041] FIGS. 6A-6B show LC/MS quantitation of signature molecules in urine. FIG. 6A is a photograph of an experimental mouse, having bilateral injury and a control uninjured mouse. FIG. 6B is a graph depicting the ratio of signature molecule (from thrombin cleavable proteolytic susceptible domain) to isotopically labeled product in injured versus control mice.

[0042] FIG. 7A shows a photograph of a typical implantable capsule in comparison to a penny and a ruler. FIG. 7B is a graph showing measurements of thrombin-cleaved peptide efflux from implantable diagnostic capsules sealed with semi-permeable membranes of different pore sizes. Capsules made with membranes of pore size 10, 30, 50 or 80 nm were loaded with nanoparticles functionalized with GGdFPipRSGGGC (SEQ ID NO: 8) and exposed to solutions of thrombin or factor Xa (a cognate and a non-cognate protease, respectively). Thrombin-specific cleavage was monitored over time and is shown in terms of kinetics of reporter release.

[0043] FIG. 8A shows representative proteolytic products appended with peptide caps of differing length and charge density. FIG. 8B is a graph showing normalized relative intensities of the peptide reporters. The inset of FIG. 8B shows a magnification of the normalized relative intensities of the peptide sequences A1-A6 as measured via LC/MS.

[0044] FIG. 9A is a schematic of the pro-diagnostic reagent, showing two identical cocktails of 12 pro-diagnostic nanoparticles, each functionalized with a different peptide, with the circles referring to the carrier, the star is a signature molecule, and the zigzag line refers to the enzyme susceptible domain. FIG. 9B is a graph depicting LC/MS peak area measurements of all twelve peptides after exposure of the first multiplex cocktail to thrombin. Left to right the sequences on the X-axis correspond to SEQ ID NOs. 37-48. FIG. 9C is a graph depicting LC/MS peak area measurements of all twelve peptides after exposure of the second multiplex cocktail to collagenase. Left to right the sequences on the X-axis correspond to SEQ ID NOs. 37-48. FIG. 9D is a graph showing the ratio of the LC/MS peak area for each peptide reporter after exposure to thrombin over the peak area measured after exposure to collagenase. Left to right the sequences on the X-axis correspond to SEQ ID NOs. 49, 38, 39, 50, 41, 47, 51, 44, 45, 48, 46, and 42.

BRIEF DESCRIPTION OF THE SEQUENCES

- [0045] SEQ ID NO: 1 is GGPQGIWQGC
- [0046] SEQ ID NO: 2 is GGPLGVRGKC
- [0047] SEQ ID NO: 3 is GGPLANvaDapARGC
- [0048] SEQ ID NO: 4 is GGPVGLIGL
- [0049] SEQ ID NO: 5 is GGPVPLSLVMC
- [0050] SEQ ID NO: 6 is GGS GGPLGLRSWC
- [0051] SEQ ID NO: 7 is GGGPWGIWQGC
- [0052] SEQ ID NO: 8 is GGdFPipRSGGGC
- [0053] SEQ ID NO: 9 is GGLVPRGSGC
- [0054] SEQ ID NO: 10 is GGPQG
- [0055] SEQ ID NO: 11 is GGPLG
- [0056] SEQ ID NO: 12 is GGPLA
- [0057] SEQ ID NO: 13 is GGPVG
- [0058] SEQ ID NO: 14 is GGPVPLS

- [0059] SEQ ID NO: 15 is GGS GGPLG
- [0060] SEQ ID NO: 16 is GGGPWG
- [0061] SEQ ID NO: 17 is GGdFPipR
- [0062] SEQ ID NO: 18 is GGLVP
- [0063] SEQ ID NO: 19 is GGVVVLS
- [0064] SEQ ID NO: 20 is Fl-dR-dS-dR
- [0065] SEQ ID NO: 21 is Fl-dR-G-dS-dR
- [0066] SEQ ID NO: 22 is Fl-dR-dS-dR-G-G-P-Q-G-I-W-G-Q-C
- [0067] SEQ ID NO: 23 is Fl-dR-G-dS-dR-G-G-P-L-G-V-R-G-K-C
- [0068] SEQ ID NO: 24 is Fl-dR-G-dS-dR-G-G-P-L-A-Nva-Dpa-A-R-G-C
- [0069] SEQ ID NO: 25 is Fl-dR-G-dS-dR-G-G-P-V-G-L-I-G-C
- [0070] SEQ ID NO: 26 is Fl-dR-dS-dR-G-G-P-V-P-L-S-L-V-M-C
- [0071] SEQ ID NO: 27 is Fl-dR-G-dS-dR-G-G-V-V-V-L-S-M-T-A-C
- [0072] SEQ ID NO: 28 is Fl-dR-G-dS-dR-G-G-S-G-G-P-L-G-L-R-S-W-C
- [0073] SEQ ID NO: 29 is Fl-dR-G-dS-dR-G-G-G-P-W-G-I-W-G-Q-G-C
- [0074] SEQ ID NO: 30 is Fl-dR-G-G-dS-G-G-dF-Pip-R-S-G-G-G-C
- [0075] SEQ ID NO: 31 is Fl-dR-dS-dR-G-G-L-V-P-R-G-S-G-C
- [0076] SEQ ID NO: 32 is Fl-dR-G-G-dS-F-P-R-S-G-G-G-C
- [0077] SEQ ID NO: 33 is Fl-dR-G-G-dS-G-G-G-dF-Pip-K-S-G-G-G-C
- [0078] SEQ ID NO: 34 is Fl-dR-G-G-dS-G-G-G-dF-P-K-S-G-G-G-C
- [0079] SEQ ID NO: 35 is dR-dS-dR
- [0080] SEQ ID NO: 36 is dR-G-dS-dR

[0081] (Fl: Fluorescein; Nva: Norvaline; Dap=(N-β-(2,4-dinitrophenyl))-L-α,β-diaminopropionic acid); Pip: pipecolic acid; d: D-isomer.)

DETAILED DESCRIPTION OF THE INVENTION

[0082] The status of physiological conditions of a subject can be assessed using the methods of the invention by identifying molecular properties also referred to as “molecular signatures”. Such molecular signatures are useful for diagnosing diseases such as cancer, rheumatoid arthritis and arteriosclerosis, as well as for prognostic indicators. The response of most cancers to medical intervention is currently monitored by physical exams and various clinical imaging modalities. A few cancers such as prostate and ovarian cancer are monitored by use of single biomarkers in the blood. Such diagnostic techniques are achieved, for instance using fluorescence detection of molecular markers which are activated in a particular disease state.

[0083] The invention relates to a platform for functional characterization of disease or condition specific enzymatic repertoire as a method to monitor both disease progression and regression as well as response to therapeutics. The methods provide orders of magnitude more in vivo enzyme-substrate information than current fluorescent detection technologies. The platform provides a unique opportunity to functionally monitor cancer and other disease progression and response to therapy. It is particularly useful for prolonged therapeutic regimens, where the discovery of prog-

nostic functional signatures would greatly assist intervention and where enzymatic signatures directly correlate to therapeutic efficacy.

[0084] By administering a pro-diagnostic reagent, such as an exogenous detectable substrate library into animal models of disease it is possible to gain information into substrate specific enzymatic activities associated with diseases, such as cancer, cardiovascular disease, arthritis, and others. The technology allows for the potential simultaneous profiling of hundreds of enzyme-substrate activities in vivo using, for instance, -chaperoned, enzyme sensitive detectable compounds, an example of a compound referred to as pro-diagnostic reagents. The method leverages the distinct pharmacokinetics of modular structures and small, optionally hydrophilic, marker peptides (RES and renal clearance, respectively). The pro-diagnostic reagents have long circulation times and thus remain in circulation or permeate into tumors via porous angiogenic vascular networks, where upon local molecules, such as enzymes (MMPs, kallikreins, cathepsins, plasminogen activators, ADAMs) gain access to the enzyme susceptible regions of the pro-diagnostic reagents or substrates gain access to the enzymes of the pro-diagnostic reagents.

[0085] When the pro-diagnostic reagents, for example, the reagents including an enzyme susceptible domain are exposed to enzymes, for instance, proteases, the reagent is cleaved, such that a marker, referred to herein as a signature molecule, is released. The marker is renally-cleared and thus functions as a “messenger” of enzyme activity. For instance, a marker may include a self-quenched dye, such as Cy5.5 which is bound to a larger molecule. When the peptide containing the self-quenched dye is cleaved or modified by specific enzymes at the disease site the fluorophores are no longer self-quenched but instead developed fluorescent properties which can be detected at remote sites. Alternatively, using mass-encoded substrate libraries, the mass of enzyme substrates are designed such that upon cleavage, a distinct mass-specific messenger of cleavage will enter the urine of a patient or animal for detection using LC-MS technology. LC-MS urine analysis can generate data that is organized into a barcode of, for instance, cancer enzyme activity. In the absence of enzyme activity the pro-diagnostic reagents remain uncleaved and the whole reagent including the signature molecule is cleared through RES organs (liver, spleen, and lymph nodes) without producing urine markers. The use of mass to identify substrates allows unprecedented multiplexing capability with the potential to assay greater than 1,000 substrates.

[0086] When the pro-diagnostic reagents includes an enzyme, such as a protease, the enzyme is exposed to endogenous substrates and the substrate is cleaved, such that a marker, referred to herein as a signature molecule, is released from the endogenous substrate. The marker is renally-cleared and thus functions as a “messenger” of enzyme activity. For instance, a marker may include a peptide, carbohydrate or nucleic acid fragment which has been cleaved from the substrate. Using the detection techniques, for instance, LC-MS technology, the mass of the signature can be detected. LC-MS urine analysis can generate data that is organized into a barcode of, for instance, cancer enzyme/substrate activity. In the absence of enzyme activity the signature molecule is not cleared through RES organs (liver, spleen, and lymph nodes) and does not produce urine markers.

[0087] Thus, the invention in some aspects involves administering to a subject a pro-diagnostic reagent, identifying a biological sample from the subject in which to detect the signature molecule and optionally collecting the sample; and, subjecting the biological sample to an analysis method in order to detect the presence of one or more signature molecules. The presence of the signature molecule in the biological sample is indicative of an active enzyme or a substrate within the subject.

[0088] For example the invention in some aspects involves methods for administering to a subject a pro-diagnostic reagent, such that the pro-diagnostic reagent has a modular structure having a carrier domain linked to a signature producing domain, wherein the signature producing domain is capable of producing a signature molecule in the subject; identifying a biological sample for detection of the signature molecule, wherein the biological sample is at a site remote from the production of the signature molecule; and, subjecting the biological sample to an analysis method in order to detect the presence of the signature molecule, wherein the presence of the signature molecule in the biological sample is indicative of a biological predictor molecule within the subject.

[0089] The pro-diagnostic reagent comprises a modular structure having a carrier domain linked to a signature producing domain. A modular structure, as used herein, refers to a molecule having multiple domains.

[0090] The signature producing domain may be, for instance, an enzyme that can react with an endogenous substrate in a subject to produce a signature molecule or it may be an enzyme susceptible domain which is linked to a signature molecule. The carrier domain may include a single type of signature producing domain, such as, a single enzyme susceptible domain and or signature molecule or it may include multiple signature producing domains, such as, different enzyme susceptible domains and signature molecules. For instance each carrier may include 1 type of signature producing domain or it may include 2-1,000 different signature producing domains or any integer therebetween. Alternatively each carrier may include greater than 1,000 signature producing domains. Multiple copies of the pro-diagnostic reagent are administered to the subject. Some mixtures of pro-diagnostic reagents may include signature producing domains that are enzymes, others may be enzymatic susceptible domains, and other may be mixtures of the two. Additionally a plurality of different pro-diagnostic reagents may be administered to the subject to determine whether multiple enzymes and/or substrates are present. In that instance, the plurality of different pro-diagnostic reagents includes a plurality of signature molecules, such that each enzyme susceptible domain is associated with a particular signature molecule or molecules.

[0091] The carrier domain may serve as the core of the pro-diagnostic agent. A purpose of the carrier domain is to serve as a platform for the signature producing domain. As such, the carrier can be any material or size as long as it can serve as a carrier or platform. Preferably the material is non-immunogenic, i.e. does not provoke an immune response in the body of the subject to which it will be administered. Another purpose is that it may function as a targeting means to target the modular structure to a tissue, cell or molecule. In some embodiments the carrier domain is a particle. A particle, for example, a nanoparticle, may, for instance, result in passive targeting to tumors by circulation.

Other types of carriers, include, for instance, compounds that cause active targeting to tissue, cells or molecules. Examples of carriers include, but are not limited to, microparticles, nanoparticles, aptamers, peptides (RGD, iRGD, LyP-1, CREKA, etc.) antibodies or antibody fragments (e.g. herceptin, cetuximab, panitumumab, etc.) and small molecules (e.g. erlotinib, gefitinib, sorafenib, etc.).

[0092] As used herein the term “particle” includes nanoparticles as well as microparticles. Nanoparticles are defined as particles of less than 1.0 μm in diameter. A preparation of nanoparticles includes particles having an average particle size of less than 1.0 μm in diameter. Microparticles are particles of greater than 1.0 μm in diameter but less than 1 mm. A preparation of microparticles includes particles having an average particle size of greater than 1.0 μm in diameter. The microparticles may therefore have a diameter of at least 5, at least 10, at least 25, at least 50, or at least 75 microns, including sizes in ranges of 5-10 microns, 5-15 microns, 5-20 microns, 5-30 microns, 5-40 microns, or 5-50 microns. A composition of particles may have heterogeneous size distributions ranging from 10 nm to mm sizes. In some embodiments the diameter is about 5 nm to about 500 nm. In other embodiments, the diameter is about 100 nm to about 200 nm. In other embodiment, the diameter is about 10 nm to about 100 nm.

[0093] The particles may be composed of a variety of materials including ceramic, metallic, natural polymer materials (including lipids, sugars, chitosan, hyaluronic acid etc), synthetic polymer materials (including poly-lactide-coglycolide, poly-glycerol sebacate, etc), and non-polymer materials, or combinations thereof.

[0094] The particles may be composed in whole or in part of polymers or non-polymer materials. Non-polymer materials, for example, may be employed in the preparation of the particles. Exemplary materials include alumina, calcium carbonate, calcium sulfate, calcium phosphosilicate, sodium phosphate, calcium aluminate, calcium phosphate, hydroxyapatite, tricalcium phosphate, dicalcium phosphate, tricalcium phosphate, tetracalcium phosphate, amorphous calcium phosphate, octacalcium phosphate, and silicates. In certain embodiments the particles may comprise a calcium salt such as calcium carbonate, a zirconium salt such as zirconium dioxide, a zinc salt such as zinc oxide, a magnesium salt such as magnesium silicate, a silicon salt such as silicon dioxide or a titanium salt such as titanium oxide or titanium dioxide.

[0095] A number of biodegradable and non-biodegradable biocompatible polymers are known in the field of polymeric biomaterials, controlled drug release and tissue engineering (see, for example, U.S. Pat. Nos. 6,123,727; 5,804,178; 5,770,417; 5,736,372; 5,716,404 to Vacanti; U.S. Pat. Nos. 6,095,148; 5,837,752 to Shastri; U.S. Pat. No. 5,902,599 to Anseth; U.S. Pat. Nos. 5,696,175; 5,514,378; 5,512,600 to Mikos; U.S. Pat. No. 5,399,665 to Barrera; U.S. Pat. No. 5,019,379 to Domb; U.S. Pat. No. 5,010,167 to Ron; U.S. Pat. No. 4,946,929 to d'Amore; and U.S. Pat. Nos. 4,806,621; 4,638,045 to Kohn; see also Langer, *Acc. Chem. Res.* 33:94, 2000; Langer, *J. Control Release* 62:7, 1999; and Uhrich et al., *Chem. Rev.* 99:3181, 1999; all of which are incorporated herein by reference).

[0096] Polymers include, but are not limited to: polyamides, polycarbonates, polyalkylenes, polyalkylene glycols, polyalkylene oxides, polyalkylene terephthalates, polyvinyl alcohols, polyvinyl ethers, polyvinyl esters, polyvinyl

halides, polyglycolides, polysiloxanes, polyurethanes and copolymers thereof, alkyl cellulose, hydroxyalkyl celluloses, cellulose ethers, cellulose esters, nitro celluloses, polymers of acrylic and methacrylic esters, methyl cellulose, ethyl cellulose, hydroxypropyl cellulose, hydroxy-propyl methyl cellulose, hydroxybutyl methyl cellulose, cellulose acetate, cellulose propionate, cellulose acetate butyrate, cellulose acetate phthalate, carboxylethyl cellulose, cellulose triacetate, cellulose sulphate sodium salt, poly(methyl methacrylate), poly(ethylmethacrylate), poly(butylmethacrylate), poly(isobutylmethacrylate), poly(hexylmethacrylate), poly(isodecylmethacrylate), poly(lauryl methacrylate), poly(phenyl methacrylate), poly(methyl acrylate), poly(isopropyl acrylate), poly(isobutyl acrylate), poly(octadecyl acrylate), polyethylene, polypropylene poly(ethylene glycol), poly(ethylene oxide), poly(ethylene terephthalate), poly(vinyl alcohols), poly(vinyl acetate, poly vinyl chloride and polystyrene.

[0097] Examples of non-biodegradable polymers include ethylene vinyl acetate, poly(meth) acrylic acid, polyamides, copolymers and mixtures thereof.

[0098] Examples of biodegradable polymers include synthetic polymers such as polymers of lactic acid and glycolic acid, polyanhydrides, poly(ortho)esters, polyurethanes, poly(butic acid), poly(valeric acid), poly(caprolactone), poly(hydroxybutyrate), poly(lactide-co-glycolide) and poly(lactide-co-caprolactone), and natural polymers such as alginate and other polysaccharides including dextran and cellulose, collagen, chemical derivatives thereof (substitutions, additions of chemical groups, for example, alkyl, alkylene, hydroxylations, oxidations, and other modifications routinely made by those skilled in the art), albumin and other hydrophilic proteins, zein and other prolamines and hydrophobic proteins, copolymers and mixtures thereof. In general, these materials degrade either by enzymatic hydrolysis or exposure to water in vivo, by surface or bulk erosion. The foregoing materials may be used alone, as physical mixtures (blends), or as co-polymers. In some embodiments the polymers are polyesters, polyanhydrides, polystyrenes, polylactic acid, polyglycolic acid, and copolymers of lactic and glycolic acid and blends thereof.

[0099] PVP is a non-ionogenic, hydrophilic polymer having a mean molecular weight ranging from approximately 10,000 to 700,000 and the chemical formula $(\text{C}_6\text{H}_9\text{NO})[\text{n}]$. PVP is also known as poly[1-(2-oxo-1-pyrrolidinyl)ethylene], PovidoneTM, PolyvidoneTM, RP 143TM, KollidonTM, Pregel STTM, PeristonTM, PlasdoneTM, PlasmosanTM, ProtagentTM SubtosanTM, and VinisilTM. PVP is non-toxic, highly hygroscopic and readily dissolves in water or organic solvents.

[0100] Polyethylene glycol (PEG), also known as poly(oxyethylene) glycol, is a condensation polymer of ethylene oxide and water having the general chemical formula $\text{HO}(\text{CH}_2\text{CH}_2\text{O})[\text{n}]\text{H}$.

[0101] Polyvinyl alcohol (PVA) is a polymer prepared from polyvinyl acetates by replacement of the acetate groups with hydroxyl groups and has the formula $(\text{CH}_2\text{CHOH})[\text{n}]$. Most polyvinyl alcohols are soluble in water.

[0102] PEG, PVA and PVP are commercially available from chemical suppliers such as the Sigma Chemical Company (St. Louis, Mo.).

[0103] In certain embodiments the particles may comprise poly(lactic-co-glycolic acid) (PLGA).

[0104] The carrier may be composed of inorganic materials. Inorganic materials include, for instance, magnetic materials, conductive materials, and semiconductor materials.

[0105] In addition to particles the carrier may be composed of any organic carrier, including biological and living carriers such as cells, viruses, bacteria, as well as any non-living organic carriers, or any composition enabling exposure of enzyme substrates to enzymes in disease (including extracellular, membrane-bound, and intracellular enzymes).

[0106] In some embodiments, the particles are porous. A porous particle can be a particle having one or more channels that extend from its outer surface into the core of the particle. In some embodiments, the channel may extend through the particle such that its ends are both located at the surface of the particle. These channels are typically formed during synthesis of the particle by inclusion followed by removal of a channel forming reagent in the particle.

[0107] The size of the pores may depend upon the size of the particle. In certain embodiments, the pores have a diameter of less than 15 microns, less than 10 microns, less than 7.5 microns, less than 5 microns, less than 2.5 microns, less than 1 micron, less than 0.5 microns, or less than 0.1 microns. The degree of porosity in porous particles may range from greater than 0 to less than 100% of the particle volume. The degree of porosity may be less than 1%, less than 5%, less than 10%, less than 15%, less than 20%, less than 25%, less than 30%, less than 35%, less than 40%, less than 45%, or less than 50%. The degree of porosity can be determined in a number of ways. For example, the degree of porosity can be determined based on the synthesis protocol of the carriers (e.g., based on the volume of the aqueous solution or other channel-forming reagent) or by microscopic inspection of the carriers post-synthesis.

[0108] The plurality of particles may be homogeneous for one or more parameters or characteristics. A plurality that is homogeneous for a given parameter, in some instances, means that particles within the plurality deviate from each other no more than about $\pm 10\%$, preferably no more than about $\pm 5\%$, and most preferably no more than about $\pm 1\%$ of a given quantitative measure of the parameter. As an example, the particles may be homogeneously porous. This means that the degree of porosity within the particles of the plurality differs by not more than $\pm 10\%$ of the average porosity. In other instances, a plurality that is homogeneous means that all the particles in the plurality were treated or processed in the same manner, including for example exposure to the same agent regardless of whether every particle ultimately has all the same properties. In still other embodiments, a plurality that is homogeneous means that at least 80%, preferably at least 90%, and more preferably at least 95% of particles are identical for a given parameter.

[0109] The plurality of particles may be heterogeneous for one or more parameters or characteristics. A plurality that is heterogeneous for a given parameter, in some instances, means that particles within the plurality deviate from the average by more than about $\pm 10\%$, including more than about $\pm 20\%$. Heterogeneous particles may differ with respect to a number of parameters including their size or diameter, their shape, their composition, their surface charge, their degradation profile, whether and what type of

agent is comprised by the particle, the location of such agent (e.g., on the surface or internally), the number of agents comprised by the particle, etc. The invention contemplates separate synthesis of various types of particles which are then combined in any one of a number of pre-determined ratios prior to contact with the sample. As an example, in one embodiment, the particles may be homogeneous with respect to shape (e.g., at least 95% are spherical in shape) but may be heterogeneous with respect to size, degradation profile and/or agent comprised therein.

[0110] Particle size, shape and release kinetics can also be controlled by adjusting the particle formation conditions. For example, particle formation conditions can be optimized to produce smaller or larger particles, or the overall incubation time or incubation temperature can be increased, resulting in particles which have prolonged release kinetics.

[0111] The particles may also be coated with one or more stabilizing substances, which may be particularly useful for long term depoting with parenteral administration or for oral delivery by allowing passage of the particles through the stomach or gut without dissolution. For example, particles intended for oral delivery may be stabilized with a coating of a substance such as mucin, a secretion containing mucopolysaccharides produced by the goblet cells of the intestine, the submaxillary glands, and other mucous glandular cells.

[0112] To enhance delivery the particles may be incorporated, for instance, into liposomes, virosomes, cationic lipids or other lipid based structures. The term "cationic lipid" refers to lipids which carry a net positive charge at physiological pH. Such lipids include, but are not limited to, DODAC, DOTMA, DDAB, DOTAP, DC-Chol and DMRIE. Additionally, a number of commercial preparations of cationic lipids are available. These include, for example, LIPO-FECTIN® (commercially available cationic liposomes comprising DOTMA and DOPE, from GIBCO/BRL, Grand Island, N.Y., USA); LIPOFECTAMINE® (commercially available cationic liposomes comprising DOSPA and DOPE, from GIBCO/BRL); and TRANSFECTAM® (commercially available cationic lipids comprising DOGS in ethanol from Promega Corp., Madison, Wis., USA). A variety of methods are available for preparing liposomes e.g., U.S. Pat. Nos. 4,186,183, 4,217,344, 4,235,871, 4,261,975, 4,485,054, 4,501,728, 4,774,085, 4,837,028, 4,946,787; and PCT Publication No. WO 91/17424. The particles may also be composed in whole or in part of GRAS components. i.e., ingredients are those that are Generally Regarded As Safe (GRAS) by the US FDA. GRAS components useful as particle material include non-degradable food based particles such as cellulose.

[0113] The carrier domain can serve several functions. As discussed above, it may be useful for targeting the product to a specific region, such as a tissue. In that instance it could include a targeting agent such as a glycoprotein, an antibody, or a binding protein.

[0114] Further, the size of the carrier domain may be adjusted based on the particular use of the pro-diagnostic reagent. For instance, the carrier domain may be designed to have a size greater than 5 nm. Particles, for instance, of greater than 5 nm are not capable of entering the urine, but rather, are cleared through the reticuloendothelial system (RES; liver, spleen, and lymph nodes). By being excluded from the removal through the kidneys any uncleaved pro-diagnostic reagent will not be detected in the urine during the analysis step. Additionally, larger particles can be useful

for maintaining the particle in the blood or in a tumor site where large particles are more easily shuttled through the vasculature. In some embodiments the carrier domain is 500 microns-5 nm, 250 microns-5 nm, 100 microns-5 nm, 10 microns-5 nm, 1 micron-5 nm, 100 nm-5 nm, 100 nm-10 nm, 50 nm-10 nm or any integer size range therebetween. In other instances the carrier domain is smaller than 5 nm in size. In such instance the pro-diagnostic reagent will be cleared into the urine. However, the presence of free signature molecule can still be detected for instance using mass spectrometry. In some embodiments the carrier domain is 1-5 nm, 2-5 nm, 3-5 nm, or 4-5 nm.

[0115] Optionally the carrier domain may include a biological agent. In one embodiment a biological agent could be incorporated in the carrier domain or it may make up the carrier domain. For instance, it may form the scaffold or platform that the proteolytic domain is attached to. Thus the compositions of the invention can achieve two purposes at the same time, the diagnostic methods and delivery of a therapeutic agent. In some embodiments the biological agent may be a enzyme inhibitor. In that instance the biological agent can inhibit proteolytic activity at a local site and the signature molecule can be used to test the activity of that particular therapeutic at the site of action. HIV is an example of the disease in which active proteases can be monitored. In this embodiment the composition may include a micro-particle or other delivery device carrying a protease inhibitor. The protease susceptible site may be sensitive to the HIV proteases such that feedback can be provided regarding the activity of the particular protease inhibitor.

[0116] Biological agents include diagnostic, cosmetic, and therapeutic agents, such as releasable drugs. Thus, any biological agent can be incorporated within the particles, which can locally or systemically deliver or maintain the incorporated agents following administration or application to a subject. Any biocompatible or pharmacologically acceptable material can be incorporated into the particles or trapped in the pores of the particles using technology known to those skilled in the art. Biological agents include but are not limited to synthetic inorganic and organic compounds, proteins and peptides, polysaccharides and other sugars, lipids, and DNA and RNA nucleic acid sequences having therapeutic, prophylactic, cosmetic or diagnostic activities. Nucleic acid sequences include genes, plasmids, vectors, antisense molecules that bind to complementary DNA to inhibit transcription, siRNA, shRNA, and ribozymes.

[0117] In certain instances, the biological agent is an anti-microbial agent. An anti-microbial agent, as used herein, refers to a naturally-occurring or synthetic compound which is capable of killing or inhibiting infectious microorganisms. The type of anti-microbial agent useful according to the invention will depend upon the type of microorganism with which the subject is infected or at risk of becoming infected. Anti-microbial agents include but are not limited to anti-bacterial agents, anti-viral agents, anti-fungal agents and anti-parasitic agents. Phrases such as “anti-infective agent”, “anti-bacterial agent”, “anti-viral agent”, “anti-fungal agent”, “anti-parasitic agent” and “parasiticide” have well-established meanings to those of ordinary skill in the art and are defined in standard medical texts.

[0118] Growth factors may also be incorporated into the carrier. As used herein, the term growth factor refers to any agent that stimulates cellular proliferation and/or differentiation. Growth factors include but are not limited to fibro-

blast growth factor (FGF), platelet-derived growth factor (PDGF), insulin-like growth factors (IGF) I and II, TGF- β , TGF- α , bone morphogenetic protein (BMP) (e.g., BMP-2, BMP-3, BMP-4, BMP-6, or BMP-7), hedgehog proteins, growth differentiation factors, hematopoietic colony-stimulating factors (CSF), vascular endothelium growth factor (VEGF), osteoid-inducing factor (OIF), angiogenins, endothelins, hepatocyte growth factor, keratinocyte growth factor, ADMP-1, interleukins (IL) (e.g., IL-3 and IL-6), epithelial growth factors, dexamethasone, leptin, sortilin, transglutaminase, prostaglandin E, 1,25-dihydroxyvitamin D3, ascorbic acid, pro-collagen, glycerol phosphate, TAK-778, statins, growth hormone, steel factor (SF), activin A (ACT), retinoic acid (RA), epidermal growth factor (EGF), hematopoietic growth factors, peptide growth factors, erythropoietin, tumor necrosis factors (TNF), interferons (IFN), heparin binding growth factor (HBGF), nerve growth factor (NGF) and muscle morphogenic factor (MMP).

[0119] The biological agent may also be an anti-cancer therapy. Anti-cancer therapies include for instance, radiotherapy, chemotherapy, adjuvant therapy, or any combination of the aforementioned.

[0120] The carrier domain may also be configured such that it can be detected in the body during the analysis. For instance, iron oxide can be incorporated into the particles so that the pro-diagnostic reagent can be tracked using MRI to provide non-invasive imaging data.

[0121] The carrier is linked to the signature producing domain. A signature producing domain, as used herein, is the portion of the modular structure that promotes the enzymatic reaction in the subject, causing the release of a signature molecule. The signature producing domain is either an active signature producing agent or an enzyme susceptible domain linked to a signature molecule. An active signature producing agent is a molecule that causes the formation of a signature molecule by modifying an endogenous molecule, referred to herein as a biological predictor molecule. For example the active signature producing agent is an enzyme that causes production of a signature molecule when it acts on an endogenous substrate. Enzymes include, for instance, proteases and glycosidases. The biological predictor molecule is the endogenous molecule acted upon by the active signature producing agent. Biological predictor molecules include for instance, substrates.

[0122] The enzyme susceptible site is dependent on enzymes that are active in a specific disease state. Alternatively, the enzyme specific site may be associated with enzymes that are ordinarily present but are absent in a particular disease state. For instance, tumors are associated with a specific set of enzymes. If the disease state being analyzed is a tumor then the product is designed with an enzyme susceptible site that matches that of the enzyme expressed by the tumor or other diseased tissue.

[0123] An enzyme, as used herein refers to any of numerous proteins produced in living cells that accelerate or catalyze the metabolic processes of an organism. Enzymes act on substrates. The substrate binds to the enzyme at a location called the active site just before the reaction catalyzed by the enzyme takes place. Enzymes include but are not limited to proteases, glycosidases, lipases, heparinases, phosphatases.

[0124] The enzyme susceptible site may be optimized to provide both high catalytic activity (or other enzymatic activity) for specified target enzymes but to also release

optimized signature molecules for detection. For the specific detection modality of mass spectrometry, inclusion of tags for rapid affinity purification, specific charges in the sequence to increase ionization efficiency, mass defects to eliminate endogenous organic background, and unique mass signatures can improve detection. In addition to improving detection sensitivity, by programming the mass of these molecules, mass selection techniques can be harnessed to remove background during detection and focus on expected signature masses.

[0125] Patient outcome depends on the phenotype of individual diseases at the molecular level, and this is often reflected in expression of enzymes. The recent explosion of bioinformatics has facilitated exploration of complex patterns of gene expression in human tissues (Fodor, S. A. Massively parallel genomics. Science 277, 393-395 (1997)). Sophisticated computer algorithms have been recently

developed capable of molecular diagnosis of tumors using the immense data sets generated by expression profiling (Khan J, Wei J S, Ringner M, Saal L H, Ladanyi M, Westermann F, et al. Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. Nat Med 2001; 7:673-679.). This information can be accessed in order to identify enzymes and substrates associated with specific diseases. Based on this information the skilled artisan can identify appropriate enzyme or substrates to incorporate into the pro-diagnostic reagent. Table 1 provides a non-limiting list of enzymes associated with (either increased or decreased with respect to normal) disease and in some instances, the specific substrate. Table 2 provides a non-limiting list of substrates associated with disease or other conditions. Numerous other enzyme/substrate combinations associated with specific diseases or conditions are known to the skilled artisan and are useful according to the invention.

TABLE 1

DISEASE	ENZYME	SUBSTRATE
Cancer	MMP	collagens, gelatin, various ECM proteins
Cancer	MMP-2	type IV collagen and gelatin
Cancer	MMP-9	type IV and V collagens and gelatin
Cancer	kallikreins	kininogens, plasminogen
Cancer	cathepsins	broad spectrum of substrates
Cancer	plasminogen activator, tPA	Plasminogen
Cancer	ADAM (A Diseintegrin And Metalloprotease, also MDC, Adamalysin)	various extracellular domains of transmembrane proteins
Pancreatic carcinoma	MMP-7	various, e.g. collagen 18, FasL, HLE, DCN, IGFBP-3, MAG, plasminogen, other MMPs
Pancreatic Cancer	ADAM9, ADAM 15	various extracellular domains of transmembrane proteins
Prostate adenocarcinoma	Matriptase, a type 11 transmembrane serine protease	unspecific, cleaves after Lys or Arg residues
Prostate cancer	Kallikrein 3	kininogens, plasminogen
Prostate cancer	ADAM 15	various extracellular domains of transmembrane proteins
Ovarian carcinoma	Kallikrein 6	kininogens, plasminogen
Epithelial-derived tumors (breast, prostate, ovarian, colon, oral)	Matriptase, a type 11 transmembrane serine protease	unspecific, cleaves after Lys or Arg residues
Ovarian Cancer	MMP-2, MMP-9, kallikrein-10 (hk-10)	type IV and V collagens and gelatin, kininogens, plasminogen
Breast, gastric, prostate cancer	cathepsins B, L and D	broad spectrum of substrates
Endometrial cancer	cathepsin B	unspecific cleavage of a broad spectrum of substrates without clear sequence specificity
esophageal adenocarcinoma	cathepsin B	unspecific cleavage of a broad spectrum of substrates without clear sequence specificity
Invasive cancers, metastases	type 11 integral serine proteases (dipeptidyl peptidase IV (DPP4/CD26),seprase/fibroblast activation protein alpha (FAPalpha) and related type II transmembrane prolyl serine peptidases))	
Invasive cancers, metastases	Seprase	various ECM proteins
viral infections		
All Retroviruses	viral protease	precursor GagPol fusion
HIV	HIV protease (HIV PR, an aspartic protease)	precursor Gag and GagPol proteins
Hepatitis C	NS3 serine protease	viral precursor polyprotein
Dengue	Dengue protease	auocleavage (NS2B/NS3), NS3/NS4A and NS4B/NS5 cleavage

TABLE 1-continued		
DISEASE	ENZYME	SUBSTRATE
West Nile bacterial infections	NS2B/NS3pro	viral precursor polypeptide
Legionella spp.	zinc metalloprotease	Me-Arg-Pro-Tyr
Meningococcal meningitis	histolytic cysteine protease	
Streptococcus pyogenes (Group A Streptococcus)	streptococcal pyrogenic exo toxin B (SpeB)	extracellular matrix, immunoglobulins, complement components
Chlostridium difficile	Cwp84	fibronectin, laminin, vitronectin and other ECM proteins
DISEASE	ENZYME	SUBSTRATE
Alzheimer's disease	BACE-1,2 (Alzheimer secretase)	-amyloid precursor protein
Stroke and recovery	MMP, tPA	
cardiovascular disease	Angiotensin Converting Enzyme (ACE)	angiotensin 1, bradykinin
Atherosclerosis	cathepsin K, L, S	broad spectrum of substrates
arthritis	MMP-1	triple-helical fibrillar collagens
rheumatoid arthritis	thrombin	Osteopontin
osteoarthritis	thrombin	Osteopontin
osteoporosis/osteoarthritis	cathepsin K, S	broad spectrum of substrates
Arthritis, inflammatory joint disease	Aggrecanase (ADAMTS4, ADAMTS 11)	aggrecans (proteoglycans)
thrombosis	factor Xa (thrombokinase)	Prothrombin
thrombosis	ADAMTS13	von Willebrand factor (vWF)
thrombosis	plasminogen activator, tPA	Plasminogen
Stress-induced Renal pressure natriuresis	Prostasin	epithelial Na channel subunits

TABLE 2		
DISEASE	TARGET SUBSTRATE	ENZYME
Inflammation	Interleukin 1 beta	MMP-2, MMP-3, MMP-9, Trypsin, chymotrypsin, pepsin, Lys-C, Glu-C, Asp-N, Arg-C
Pituitary gland dysfunction, abnormal bone density, growth disorders	IGFBP-3	MMP-1, MMP-3, MMP-9, Trypsin, chymotrypsin, pepsin, Lys-C, Glu-C, Asp-N, Arg-C
Cancer	TGF-beta	MMP-9, Trypsin, chymotrypsin, pepsin, Lys-C, Glu-C, Asp-N, Arg-C
Cancer, autoimmune disease	TNF	MMP-7, Trypsin, chymotrypsin, pepsin, Lys-C, Glu-C, Asp-N, Arg-C
Cancer, autoimmune disease	FASL	MMP-7, Trypsin, chymotrypsin, pepsin, Lys-C, Glu-C, Asp-N, Arg-C
Wound healing, cardiac disease	HB-EGF	MMP-3, Trypsin, chymotrypsin, pepsin, Lys-C, Glu-C, Asp-N, Arg-C
Pfeiffer syndrome	FGFR1	MMP-2, Trypsin, chymotrypsin, pepsin, Lys-C, Glu-C, Asp-N, Arg-C
Cancer	Decorin	MMP-2, MMP-3, MMP-7, Trypsin, chymotrypsin, pepsin, Lys-C, Glu-C, Asp-N, Arg-C
Cancer	Tumor associated carbohydrate antigens	Endoglycosidases
Cancer	Sialyl Lewis ^x	O-glycanase
Cancer	Sialyl Lewis ^x	O-glycanase
Cancer/ Rheumatoid Arthritis, pulmonary hypertension	VEGF	Trypsin, chymotrypsin, pepsin, Lys-C, Glu-C, Asp-N, Arg-C
Cancer	EGF	Trypsin, chymotrypsin, pepsin, Lys-C, Glu-C, Asp-N, Arg-C
Cancer	IL2	Trypsin, chymotrypsin, pepsin, Lys-C, Glu-C, Asp-N, Arg-C
Cancer	IL6	Trypsin, chymotrypsin, pepsin, Lys-C, Glu-C, Asp-N, Arg-C
inflammation/angiogenesis		
Cancer	IFN-γ	Trypsin, chymotrypsin, pepsin, Lys-C, Glu-C, Asp-N, Arg-C

TABLE 2-continued

DISEASE	TARGET SUBSTRATE	ENZYME
Cancer inflammation/angiogenesis, Rheumatoid Arthritis	TNF- α	Trypsin, chymotrypsin, pepsin, Lys-C, Glu-C, Asp-N, Arg-C
Cancer, Pulmonary fibrosis, Asthma	TGF- β	Trypsin, chymotrypsin, pepsin, Lys-C, Glu-C, Asp-N, Arg-C
Cancer, Pulmonary hypertension	PDGF	Trypsin, chymotrypsin, pepsin, Lys-C, Glu-C, Asp-N, Arg-C
Cancer, pulmonary cystadenoma	Fibroblast growth factor (FGF)	Trypsin, chymotrypsin, pepsin, Lys-C, Glu-C, Asp-N, Arg-C
Cancer	Brain-derived neurotrophic factor (BDNF)	Trypsin, chymotrypsin, pepsin, Lys-C, Glu-C, Asp-N, Arg-C
Cancer	Interferon regulatory factors (IRF-1, IRF-2)	Trypsin, chymotrypsin, pepsin, Lys-C, Glu-C, Asp-N, Arg-C
Inhibitor of tumor suppressors	MIF	Trypsin, chymotrypsin, pepsin, Lys-C, Glu-C, Asp-N, Arg-C
Lymphomas/carcinomas, alveolar proteinosis	GM-CSF	Trypsin, chymotrypsin, pepsin, Lys-C, Glu-C, Asp-N, Arg-C
Cancer invasion	M-CSF	Trypsin, chymotrypsin, pepsin, Lys-C, Glu-C, Asp-N, Arg-C
Chemical carcinogenesis, multiple sclerosis, rheumatoid arthritis, Crohn's disease	IL-12	Trypsin, chymotrypsin, pepsin, Lys-C, Glu-C, Asp-N, Arg-C
Natural Killer T cell leukemias, inflammatory bowel disease, rheumatoid arthritis	IL-15	Trypsin, chymotrypsin, pepsin, Lys-C, Glu-C, Asp-N, Arg-C
Cirrhosis	Tissue inhibitor of MMPs (TIMPs)	Trypsin, chymotrypsin, pepsin, Lys-C, Glu-C, Asp-N, Arg-C
Cirrhosis	Collagen 1, III	MMP-1, MMP-8, Trypsin, chymotrypsin, pepsin, Lys-C, Glu-C, Asp-N, Arg-C
Cirrhosis	Collagen IV, V	MMP-2, Trypsin, chymotrypsin, pepsin, Lys-C, Glu-C, Asp-N, Arg-C

[0126] Several of the enzyme/substrates described above are described in the following publications, all of which are incorporated herein in their entirety by reference: Parks, W. C. and R. P. Mecham (Eds): Matrix metalloproteinases. San Diego: Academic Press; 1998; Nagase, H. and J. F. Woessner, Jr. (1999) J. Biol. Chem. 274:21491; Ito, A. et al. (1996) J. Biol. Chem. 271:14657; Schonbeck, U. et al. (1998) J. Immunol. 161: 3340; *Rajah*, R. et al. (1999) Am. J. Cell Mol. Biol. 20:199; Fowlkes, J. L. et al. (1994) Endocrinology 135:2810; Manes, S. et al. (1999) J. Biol. Chem. 274:6935; Mira, E. et al. (1999) Endocrinology 140:1657; Yu, Q. and I. Stamenkovic (2000) Genes Dev. 14:163; Haro, H. et al. (2000) J. Clin. Invest. 105:143; Powell, C. P. et al. (1999) Curr. Biol. 9:1441; Suzuki, M. et al. (1997) J. Biol. Chem. 272:31730; Levi, E. et al. (1996) Proc. Natl. Acad. Sci. USA 93:7069; Imai, K. et al. (1997) Biochem. J. 322:809; Smith, M. M. et al. (1995) J. Biol. Chem. 270:6440; and Dranoff, G. (2004) Nat. Rev. Cancer 4: 11-22.

[0127] The signature producing domain may be attached directly to the carrier. For instance it may be coated directly on the surface of microparticles using known techniques. Alternatively if the carrier is a protein material it may be directly connected through a peptide bond. Additionally, the signature producing domain may be connected to the carrier domain through the use of a linker. As used herein “linked” or “linkage” means two entities are bound to one another by any physicochemical means. Any linkage known to those of ordinary skill in the art, covalent or non-covalent, is embraced. Thus, in some embodiments the carrier has a linker attached to an external surface, which can be used to

link the signature producing domain. Another molecule can also be attached to the linker.

[0128] The signature producing domain is preferably a polymer made up of a plurality of chemical units. A “chemical unit” as used herein is a building block or monomer which may be linked directly or indirectly to other building blocks or monomers to form a polymer. In some embodiments the signature producing domain is a peptide that is susceptible to cleavage by an enzyme or causes cleavage of a substrate associated with a disease or condition. A number of examples of when the proteolytic cleavage site is a peptide are presented in the table above.

[0129] The enzyme susceptible domain may also be a polysaccharide. Some polysaccharide specific degrading enzymes are associated with tumors, angiogenesis and other conditions. A “polysaccharide” is a biopolymer comprised of linked saccharide or sugar units. The polysaccharides used as proteolytic susceptible domains may be isolated or synthesized de novo. For example, the polysaccharides may be isolated from natural sources e.g. purified, as by cleavage and gel separation or may be synthesized e.g., by chemical synthesis and incorporated into the pro-diagnostic reagent.

[0130] For instance, HSGAG degrading enzymes are enzymes that can be analyzed according to the methods of the invention. HSGAG degrading enzymes include heparinase-I, heparinase-II, heparinase-III, D-glucuronidase and L-iduronidase. The heparinases cleave at the glycosidic linkage before a uronic acid. Heparinase I clips at a glycosidic linkage before a 2-O sulfated iduronic acid. Heparinase-III cleaves at a glycosidic linkage before an unsulfated

glucuronic acid. Heparinase-II cleaves at both Hep-I and Hep-III cleavable sites. Glucuronidase and iduronidase, as their name suggests cleave at the glycosidic linkage after a glucuronic acid and iduronic acid respectively. Nitrous acid clips randomly at glycosidic linkages after a N-sulfated hexosamine and converts the six membered hexosamine ring to a 5 membered anhydromannitol ring. Appropriate enzyme susceptible domains may be designed based on the known substrates and cleavage sites of these enzymes.

[0131] The pro-diagnostic reagent may also include an implantable microdelivery device that houses the modular structure. An implantable microdelivery device is any type of device, that is sized for implantation into a body and can retain the modular structure. For instance the device may be an implantable capsule that contains the modular structure housed there in. The capsule may have a semi-permeable membrane, such that the modular structure cannot pass through the membrane, but which is permeable to endogenous molecules such as enzymes and substrates as well as signature molecules. Alternatively the implantable microdelivery device may be a chip having the modular structure attached thereto. Examples of implantable microdelivery devices include but are not limited to implantable capsules, chips, sustained-release formulations, multi-pulse drug delivery resorbable polymeric microchip device (Grayson et al. *Nature Materials*, VOL 2, November 2003, p. 767), and controlled release microchips (Santini et al *Nature*, vol 397, 1999, p. 335). These devices may be made from many materials include many of the polymeric materials described herein. Preferably the implantable devices are biocompatible and non-toxic.

[0132] Modification of the enzyme susceptible domain by an enzyme in vivo, results in the production of a signature molecule. Alternatively, when the signature producing domain is an enzyme the enzyme cleaves an endogenous substrate producing a signature molecule from the endogenous substrate. The signature molecule is a detectable molecule. It can be part of the enzyme susceptible domain, e.g. the piece that is released or added upon cleavage or it can be a separate entity. The signature molecule may be, for instance, a peptide, nucleic acid, small molecule, fluorophore/quencher, carbohydrate, particle, radiolabel, MRI-active compound, inorganic material, organic material, with encoded characteristics to facilitate optimal detection.

[0133] The signature molecule may be detected by any known detection methods. A variety of methods may be used, depending on the nature of the signature molecule/label. Labels on signature molecules may be directly detected through optical or electron density, radioactive emissions, nonradiative energy transfers, or signature molecules may be indirectly detected with antibody conjugates, streptavidin-biotin conjugates, mass spectrometry, liquid chromatography-mass spectrometry, PCR analysis, DNA microarray, and fluorescence analysis.

[0134] The analysis step may be performed directly on the biological sample or the signature component may be purified to some degree first. For instance, a purification step may involve isolating the signature molecule from other components in the biological sample. Purification steps include methods such as affinity chromatography. As used herein an "isolated molecule" or "purified molecule" is a signature molecule that is isolated to some extent from its

natural environment. The isolated or purified molecule need not be 100% pure or even substantially pure prior to analysis.

[0135] The methods for analysing signature molecules by identifying the presence of a signature molecule may be used to provide a qualitative assessment of the molecule (e.g., whether the signature molecule is present or absent) or a quantitative assessment (e.g., the amount of signature molecule present to indicate a comparative activity level of the enzymes. The quantitative value may be calculated by any means, such as, by determining the percent relative amount of each fraction present in the sample. Methods for making these types of calculations are known in the art.

[0136] A signature molecule can be determined using mass spectrometry. The mass spectrometry data may be a valuable tool to ascertain information about the signature molecule in the biological sample. After a molecular weight of a signature molecule is identified, it may be compared to molecular weights of other known signature molecules.

[0137] Molecular weight may be determined by several methods including mass spectrometry. The use of mass spectrometry for determining the molecular weight of molecules is well known in the art. Liquid chromatography-mass spectrometry (LC-MS, or alternatively HPLC-MS) is an analytical chemistry technique that combines the physical separation capabilities of liquid chromatography (or HPLC) with the mass analysis capabilities of mass spectrometry. LC-MS is a powerful technique which has very high sensitivity and specificity. It can be used to detect signature molecules in a complex mixture. Other types of mass spectrometry known in the art, such as, matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS), electron spray-MS, fast atom bombardment mass spectrometry (FAB-MS) and collision-activated dissociation mass spectrometry (CAD) can also be used to identify the molecular weight of the signature molecule.

[0138] Methods for performing mass spectrometry using nucleic acid samples have been described. See e.g., U.S. Pat. No. 5,885,775. U.S. Pat. Nos. 7,412,332 and 6,597,996 describe methods for detecting polysaccharide molecules using mass spectrometry. As shown in these patent applications, one technique for comparing molecular weights is to generate a mass line and compare the molecular weight of the unknown polysaccharide to the mass line to determine a subpopulation of polysaccharides which have the same molecular weight. A "mass line" is an information database, preferably in the form of a graph or chart which stores information for each possible type of polysaccharide having a unique sequence based on the molecular weight of the polysaccharide. Because mass spectrometry data indicates the mass of a fragment to 1 Da accuracy, a length may be assigned uniquely to a fragment by looking up a mass on the mass line.

[0139] NMR spectroscopy is an analytical tool that allows for the determination of molecular structure. Utilizing the magnetic properties of some nuclei, the nuclear spins of the nuclei can be oriented randomly with an external magnetic field. Oriented nuclei that are subsequently irradiated at the correct frequency will absorb energy and transition to a higher energy state. Upon relaxation this energy is emitted and detected in various NMR systems. This irradiation of the nuclei occur in pulses. In basic one dimensional (1D) NMR the excitation is produced from a single pulse and emitted radiation is detected as free induction decay (FID). In two

dimensional (2D) NMR spectroscopy the nuclei is irradiated with two pulses, and acquisition of the FID occurs at many time points with a delay between the pulses.

[0140] When the signature molecule is a nucleic acid, it can also be analyzed using PCR and microarrays. PCR methods are well-known in the art. For instance, U.S. Pat. No. 5,333,675, issued to Mullis et al. describes an apparatus and method for performing automated PCR. In general, performance of a PCR method results in amplification of a selected region of DNA by providing two DNA primers, each of which is complementary to a portion of one strand within the selected region of DNA. The primer is hybridized to a template strand of nucleic acid in the presence of deoxyribonucleotide triphosphates (dATP, dCTP, dGTP, and dTTP) and a chain extender enzyme, such as DNA polymerase. The primers are hybridized with the separated strands, forming DNA molecules that are single stranded except for the region hybridized with the primer, where they are double stranded. The double stranded regions are extended by the action of the chain extender enzyme (e.g. DNA polymerase) to form an extended double stranded molecule between the original two primers. The double stranded DNA molecules are separated to produce single strands which can then be re-hybridized with the primers. The process is repeated for a number of cycles to generate a series of DNA strands having the same nucleotide sequence between and including the primers.

[0141] Chain extender enzymes are well known in the art and include, for example, *E. coli* DNA polymerase I, klenow fragment of *E. coli* DNA polymerase I, T4 DNA polymerase, T7 DNA polymerase, recombinant modified T7 DNA polymerase, reverse transcriptase, and other enzymes. Heat stable enzymes are particularly preferred as they are useful in automated thermal cycle equipment. Heat stable polymerases include, for example, DNA polymerases isolated from *Bacillus stearothermophilus* (Bio-Rad), *Thermus thermophilus* (finzyme, ATCC number 27634), *thermus* species (ATCC number 31674), *Thermus aquaticus* strain TV11518 (ATCC number 25105), *Sulfolobus acidocaldarius*, described by Bukhrashuili et al., *Biochem. Biophys. Acta.*, 1008:102-07 (1909), *Thermus filiformis* (ATCC number 43280), Taq DNA polymerase, commercially available from Perkin-Elmer-Cetus (Norwalk, Conn.), Promega (Madison, Wis.) and Stratagene (La Jolla, Calif.), and AmpliTaq™ DNA polymerase, a recombinant *Thermus equitus* Taq DNA polymerase, available from Perkin-Elmer-Cetus and described in U.S. Pat. No. 4,889,818.

[0142] Preferably, the PCR-based methods performed according to the invention are automated and performed using thermal cyclers. Many types of thermal cyclers are well-known in the art. For instance, M. J. Research (Watertown, Mass.) provides a thermal cycler having a peltier heat pump to provide precise uniform temperature control in the thermal cyclers; DeltaCycler thermal cyclers from Ericomp (San Diego, Calif.) also are peltier-based and include automatic ramping control, time/temperature extension programming and a choice of tube or microplate configurations. The RoboCycler™ by Stratagene (La Jolla, Calif.) incorporates robotics to produce rapid temperature transitions during cycling and well-to-well uniformity between samples; and a particularly preferred cycler, is the Perkin-Elmer Applied Biosystems (Foster City, Calif.) ABI Prism™ 877 Integrated Thermal cycler, which is operated through a

programmable interface that automates liquid handling and thermocycling processes for fluorescent DNA sequencing and PCR reactions.

[0143] The presence or absence of enzymes in the subject may also be determined using hybridization techniques. Standard hybridization techniques of microarray technology are utilized to assess the presence of nucleic acids in the biological sample. Microarray technology, which is also known by other names including: DNA chip technology, gene chip technology, and solid-phase nucleic acid array technology, is well known to those of ordinary skill in the art and is based on, but not limited to, obtaining an array of identified nucleic acid probes on a fixed substrate, labeling target molecules with reporter molecules (e.g., radioactive, chemiluminescent, or fluorescent tags such as fluorescein, Cy3-dUTP, or Cy5-dUTP), hybridizing target nucleic acids to the probes, and evaluating target-probe hybridization. A probe with a nucleic acid sequence that perfectly matches the target sequence will, in general, result in detection of a stronger reporter-molecule signal than will probes with less perfect matches. Many components and techniques utilized in nucleic acid microarray technology are presented in *The Chipping Forecast*, Nature Genetics, Vol. 21, January 1999, the entire contents of which is incorporated by reference herein.

[0144] According to the present invention, microarray substrates may include but are not limited to glass, silica, aluminosilicates, borosilicates, metal oxides such as alumina and nickel oxide, various clays, nitrocellulose, or nylon. In all embodiments a glass substrate is preferred. An “array” as used herein is a set of molecules arranged in a specific order with respect to a surface. Preferably the array is composed of polynucleotides attached to the surface. Oligonucleotide arrays can be used to screen nucleic acid samples for a target nucleic acid, which can be labeled with a detectable marker. A fluorescent signal resulting from hybridization between a target nucleic acid and a substrate-bound oligonucleotide provides information relating to the identity of the target nucleic acid by reference to the location of the oligonucleotide in the array on the substrate. Such a hybridization assay can generate thousands of signals which exhibit different signal strengths. These signals correspond to particular oligonucleotides of the array. Different signal strengths will arise based on the amount of labeled target nucleic acid hybridized with an oligonucleotide of the array.

[0145] Conditions for optimal hybridization are known. The hybridization conditions in general are those used commonly in the art, such as those described in Sambrook et al., “Molecular Cloning: A Laboratory Manual”, (1989), 2nd Ed., Cold Spring Harbor, N.Y.; Berger and Kimmel, “Guide to Molecular Cloning Techniques”, *Methods in Enzymology*, (1987), Volume 152, Academic Press, Inc., San Diego, Calif.; and Young and Davis, (1983), *PNAS* (USA) 80:1194. In general, incubation temperatures for hybridization of nucleic acids range from about 20° C. to 75° C. For probes 17 nucleotides residues and longer, a preferred temperature range for hybridization is from about 50° C. to 54° C. The hybridization temperature for longer probes is preferably from about 55° C. to 65° C. and for shorter probes is less than 52° C. Rehybridization may be performed in a variety of time frames. Preferably, hybridization of SNP and RCGs performed for at least 30 minutes.

[0146] The signature molecule may be labeled. For example, a label may be added directly to a nucleic acid

when the isolated signature molecule is subjected to PCR. For instance, a PCR reaction performed using labeled primers or labeled nucleotides will produce a labeled product. Labeled nucleotides (e.g., fluorescein-labeled CTP) are commercially available. Methods for attaching labels to nucleic acids are well known to those of ordinary skill in the art and, in addition to the PCR method, include, for example, nick translation and end-labeling.

[0147] Labels suitable for use in the methods of the present invention include any type of label detectable by standard means, including spectroscopic, photochemical, biochemical, electrical, optical, or chemical methods. Preferred types of labels include fluorescent labels such as fluorescein. A fluorescent label is a compound comprising at least one fluorophore. Commercially available fluorescent labels include, for example, fluorescein phosphoramidides such as fluoreprime (Pharmacia, Piscataway, N.J.), fluoredate (Millipore, Bedford, Mass.), FAM (ABI, Foster City, Calif.), rhodamine, polymethadine dye derivative, phosphores, Texas red, green fluorescent protein, CY3, and CYS. Polynucleotides can be labeled with one or more spectrally distinct fluorescent labels. "Spectrally distinct" fluorescent labels are labels which can be distinguished from one another based on one or more of their characteristic absorption spectra, emission spectra, fluorescent lifetimes, or the like. Spectrally distinct fluorescent labels have the advantage that they may be used in, combination ("multiplexed"). Radionuclides such as ^3H , ^{125}I , ^{35}S , ^{14}C , or ^{32}P are also useful labels according to the methods of the invention. A plurality of radioactively distinguishable radionuclides can be used. Such radionuclides can be distinguished, for example, based on the type of radiation (e.g., α , β , or δ radiation) emitted by the radionuclides. The ^{32}P signal can be detected using a phosphoimager, which currently has a resolution of approximately 50 microns. Other known techniques, such as chemiluminescence or colorimetric (enzymatic color reaction), can also be used.

[0148] Quencher compositions in which a "donor" fluorophore is joined to an "acceptor" chromophore by a short bridge that is the binding site for the enzyme may also be used. The signal of the donor fluorophore is quenched by the acceptor chromophore through a process believed to involve resonance energy transfer (RET). Cleavage of the peptide results in separation of the chromophore and fluorophore, removal of the quench, and generation of a subsequent signal measured from the donor fluorophore.

[0149] Once the data is obtained, e.g. as a two-dimensional image, a computer can be used to transform the data into a displayed image which varies in color depending on the intensity of light emission at a particular location. Any type of commercial software which can perform this type of data analysis can be used. In general, the data analysis involves the steps of determining the intensity of the fluorescence emitted as a function of the position on the substrate, removing the outliers, and calculating the relative binding affinity. One or more of the presence, absence, and intensity of signal corresponding to a label is used to assess the presence or absence of an signature molecule. The presence and absence of one or more signature molecules can be used to determine the disease status of an individual based on the presence or absence of an enzyme.

[0150] The data may also be observed and analyzed manually. For instance, the presence or absence of a fluorescent

label may be observed in order to provide the diagnostic or prognostic information from the data.

[0151] The disease or condition assessed according to the methods of the invention is any disease or condition that is associated with an enzyme. For instance, cancer, cardiovascular disease, arthritis, viral, bacterial, parasitic or fungal infection, Alzheimer's disease emphysema, thrombosis, hemophilia, stroke, organ disfunction, any inflammatory condition, vascular disease, parenchymal disease, or a pharmacologically-induced state are all known to be associated with enzymes. A pharmacologically induced state is a condition in which enzyme inhibitors and other agents directly or indirectly affect enzyme activities. Thus each of the these can be assessed or monitored or studied according to methods of the invention.

[0152] It is useful to be able to differentiate non-metastatic primary tumors from metastatic tumors, because metastasis is a major cause of treatment failure in cancer patients. If metastasis can be detected early, it can be treated aggressively in order to slow the progression of the disease. Metastasis is a complex process involving detachment of cells from a primary tumor, movement of the cells through the circulation, and eventual colonization of tumor cells at local or distant tissue sites. Additionally, it is desirable to be able to detect a predisposition for development of a particular cancer such that monitoring and early treatment may be initiated. For instance, an extensive cytogenetic analysis of hematologic malignancies such as lymphomas and leukemias have been described, see e.g., Solomon et al., Science 254, 1153-1160, 1991. Early detection or monitoring using the non-invasive methods of the invention may be useful.

[0153] Solid tumors progress from tumorigenesis through a metastatic stage and into a stage at which several different active proteases can be involved. Some protease are believed to alter the tumor such that it can progress to the next stage, i.e., by conferring proliferative advantages, the ability to develop drug resistance or enhanced angiogenesis, proteolysis, or metastatic capacity.

[0154] Alzheimer's disease causes progressive dementia with consequent formation of amyloid plaques, neurofibrillary tangles, gliosis and neuronal loss. The disease occurs in both genetic and sporadic forms whose clinical course and pathological features are quite similar. Three genes have been discovered to date which, when mutated, cause an autosomal dominant form of Alzheimer's disease. These encode the amyloid protein precursor (APP) and two related proteins, presenilin-1 (PS1) and presenilin-2 (PS2). Mutations in any of the three proteins have been observed to enhance proteolytic processing of APP via an intracellular pathway that produces amyloid beta peptide ($\text{A}\beta$ peptide), a 40-42 amino acid long peptide that is the primary component of amyloid plaque in Alzheimer's disease. Pathological processing of APP at the β - and γ -secretase sites, which are located N-terminal and C-terminal to the α -secretase site, respectively, produces a very different result than processing at the α site. Sequential processing at the β - and γ -secretase sites releases the $\text{A}\beta$ peptide, a peptide possibly very important in Alzheimer's disease pathogenesis. The β secretase enzyme, termed Aspartyl Protease 2 (Asp2) is thought to mediate this processing. The presence of Asp2 activity is important for the diagnosis and prognosis of Alzheimer's disease. This enzyme and its substrate can also be used in

the methods of the invention to monitor the ability of a therapeutic to function in slowing the progression of Alzheimer's disease.

[0155] As used herein, a subject is a human, non-human primate, cow, horse, pig, sheep, goat, dog, cat, or rodent. In all embodiments human subjects are preferred. In aspects of the invention pertaining to cancer diagnosis in general the subject preferably is a human suspected of having cancer, or a human having been previously diagnosed as having cancer. Methods for identifying subjects suspected of having cancer may include physical examination, subject's family medical history, subject's medical history, biopsy, or a number of imaging technologies such as ultrasonography, computed tomography, magnetic resonance imaging, magnetic resonance spectroscopy, or positron emission tomography.

[0156] As used herein, a biological sample is a tissue sample. The biological sample may be examined in the body, for instance, by detecting a label at the site of the tissue, i.e. urine. Alternatively the biological sample may be collected from the subject and examined in vitro. Biological samples include but are not limited to urine, blood, saliva, or mucous secretion. In preferred embodiments the tissue sample is obtained non-invasively, such as the urine.

[0157] A "plurality" of elements, as used throughout the application refers to 2 or more of the elements.

[0158] The pro-diagnostic reagents of the invention are administered to the subject in an effective amount for detecting enzyme activity. An "effective amount", for instance, is an amount necessary or sufficient to cause release of a detectable level of signature molecule in the presence of an enzyme. In some instances when a therapeutic is administered in the pro-diagnostic reagent, the effective amount is that amount necessary to realize a desired biologic effect. An "effective amount for treating Alzheimer's disease", for instance, could be that amount necessary to (i) prevent further memory loss and/or (ii) arresting or slowing memory loss with respect to memory loss in the absence of the therapy. According to some aspects of the invention, an effective amount is that amount of a compound of the invention alone or in combination with another medicament, which when combined or co-administered or administered alone, results in a therapeutic response to the disease, either in the prevention or the treatment of the disease. The biological effect may be the amelioration and or absolute elimination of symptoms resulting from the disease. In another embodiment, the biological effect is the complete abrogation of the disease, as evidenced for example, by the absence of a symptom of the disease.

[0159] The effective amount of a compound of the invention described herein may vary depending upon the specific compound used, the mode of delivery of the compound, and whether it is used alone or in combination. The effective amount for any particular application can also vary depending on such factors as the disease being assessed or treated, the particular compound being administered, the size of the subject, or the severity of the disease or condition as well as the detection method. One of ordinary skill in the art can empirically determine the effective amount of a particular molecule of the invention without necessitating undue experimentation. Combined with the teachings provided herein, by choosing among the various active compounds and weighing factors such as potency, relative bioavailabil-

ity, patient body weight, severity of adverse side-effects and preferred mode of administration, an effective regimen can be planned.

[0160] Pharmaceutical compositions of the present invention comprise an effective amount of one or more agents, dissolved or dispersed in a pharmaceutically acceptable carrier. The phrases "pharmaceutical or pharmacologically acceptable" refers to molecular entities and compositions that do not produce an adverse, allergic or other untoward reaction when administered to an animal, such as, for example, a human, as appropriate. Moreover, for animal (e.g., human) administration, it will be understood that preparations should meet sterility, pyrogenicity, general safety and purity standards as required by FDA Office of Biological Standards.

[0161] As used herein, "pharmaceutically acceptable carrier" includes any and all solvents, dispersion media, coatings, surfactants, antioxidants, preservatives (e.g., antibacterial agents, antifungal agents), isotonic agents, absorption delaying agents, salts, preservatives, drugs, drug stabilizers, gels, binders, excipients, disintegration agents, lubricants, sweetening agents, flavoring agents, dyes, such like materials and combinations thereof, as would be known to one of ordinary skill in the art (see, for example, Remington's Pharmaceutical Sciences (1990), incorporated herein by reference). Except insofar as any conventional carrier is incompatible with the active ingredient, its use in the therapeutic or pharmaceutical compositions is contemplated. The agent may comprise different types of carriers depending on whether it is to be administered in solid, liquid or aerosol form, and whether it need to be sterile for such routes of administration as injection.

[0162] Preferably the material is injected into the body but could also be administered by other routes. For instance, the compounds of the present invention can be administered intravenously, intradermally, intraarterially, intralesionally, intratumorally, intracranially, intraarticularly, intraprostatically, intrapleurally, intratracheally, intranasally, intravitreally, intravaginally, intrarectally, topically, intratumorally, intramuscularly, intraperitoneally, subcutaneously, subconjunctival, intravesicularly, mucosally, intrapericardially, intraumbilically, intraocularly, orally, topically, locally, inhalation (e.g., aerosol inhalation), injection, infusion, continuous infusion, localized perfusion bathing target cells directly, via a catheter, via a lavage, in creams, in lipid compositions (e.g., liposomes), or by other method or any combination of the foregoing as would be known to one of ordinary skill in the art (see, for example, Remington's Pharmaceutical Sciences (1990), incorporated herein by reference).

[0163] In certain embodiments, pharmaceutical compositions may comprise, for example, at least about 0.1% of an active compound. In other embodiments, the an active compound may comprise between about 2% to about 75% of the weight of the unit, or between about 25% to about 60%, for example, and any range derivable therein.

[0164] The agent may be formulated into a composition in a free base, neutral or salt form. Pharmaceutically acceptable salts, include the acid addition salts, e.g., those formed with the free amino groups of a proteinaceous composition, or which are formed with inorganic acids such as for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric or mandelic acid. Salts formed with the free carboxyl groups also can be derived from inorganic

bases such as for example, sodium, potassium, ammonium, calcium or ferric hydroxides; or such organic bases as isopropylamine, trimethylamine, histidine or procaine.

[0165] In embodiments where the composition is in a liquid form, a carrier can be a solvent or dispersion medium comprising but not limited to, water, ethanol, polyol (e.g., glycerol, propylene glycol, liquid polyethylene glycol, etc.), lipids (e.g., triglycerides, vegetable oils, liposomes) and combinations thereof. The proper fluidity can be maintained, for example, by the use of a coating, such as lecithin; by the maintenance of the required particle size by dispersion in carriers such as, for example liquid polyol or lipids; by the use of surfactants such as, for example hydroxypropylcellulose; or combinations thereof such methods. In many cases, it will be preferable to include isotonic agents, such as, for example, sugars, sodium chloride or combinations thereof.

[0166] According to the methods of the invention, the compound may be administered in a pharmaceutical composition. In general, a pharmaceutical composition comprises the compound of the invention and a pharmaceutically-acceptable carrier. Pharmaceutically-acceptable carriers for peptides, monoclonal antibodies, and antibody fragments are well-known to those of ordinary skill in the art. As used herein, a pharmaceutically-acceptable carrier means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredients.

[0167] Pharmaceutically acceptable carriers include diluents, fillers, salts, buffers, stabilizers, solubilizers and other materials which are well-known in the art. Exemplary pharmaceutically acceptable carriers for peptides in particular are described in U.S. Pat. No. 5,211,657. Such preparations may routinely contain salt, buffering agents, preservatives, compatible carriers, and optionally other therapeutic agents. When used in medicine, the salts should be pharmaceutically acceptable, but non-pharmaceutically acceptable salts may conveniently be used to prepare pharmaceutically-acceptable salts thereof and are not excluded from the scope of the invention. Such pharmacologically and pharmaceutically-acceptable salts include, but are not limited to, those prepared from the following acids: hydrochloric, hydrobromic, sulfuric, nitric, phosphoric, maleic, acetic, salicylic, citric, formic, malonic, succinic, and the like. Also, pharmaceutically-acceptable salts can be prepared as alkaline metal or alkaline earth salts, such as sodium, potassium or calcium salts.

[0168] The compounds of the invention may be formulated into preparations in solid, semi-solid, liquid or gaseous forms such as tablets, capsules, powders, granules, ointments, solutions, depositories, inhalants and injections, and usual ways for oral, parenteral or surgical administration. The invention also embraces pharmaceutical compositions which are formulated for local administration, such as by implants.

[0169] Compositions suitable for oral administration may be presented as discrete units, such as capsules, tablets, lozenges, each containing a predetermined amount of the active agent. Other compositions include suspensions in aqueous liquids or non-aqueous liquids such as a syrup, elixir or an emulsion.

[0170] The compounds of the invention may be administered directly to a tissue. Direct tissue administration may be achieved by direct injection. The compounds may be admin-

istered once, or alternatively they may be administered in a plurality of administrations. If administered multiple times, the compounds may be administered via different routes. For example, the first (or the first few) administrations may be made directly into the affected tissue while later administrations may be systemic.

[0171] For oral administration, the compounds can be formulated readily by combining the active compounds with pharmaceutically acceptable carriers well known in the art. Such carriers enable the compounds of the invention to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a subject to be treated. Pharmaceutical preparations for oral use can be obtained as solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate. Optionally the oral formulations may also be formulated in saline or buffers for neutralizing internal acid conditions or may be administered without any carriers.

[0172] Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, talc, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

[0173] Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added. Microspheres formulated for oral administration may also be used. Such microspheres have been well defined in the art. All formulations for oral administration should be in dosages suitable for such administration.

[0174] For buccal administration, the compositions may take the form of tablets or lozenges formulated in conventional manner.

[0175] For administration by inhalation, the compounds for use according to the present invention may be conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebulizer, with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of e.g. gelatin for use in an inhaler or insufflator may be formulated

containing a powder mix of the compound and a suitable powder base such as lactose or starch. Techniques for preparing aerosol delivery systems are well known to those of skill in the art. Generally, such systems should utilize components which will not significantly impair the biological properties of the active agent (see, for example, Sciarra and Cutie, "Aerosols," in Remington's Pharmaceutical Sciences, 18th edition, 1990, pp 1694-1712; incorporated by reference). Those of skill in the art can readily determine the various parameters and conditions for producing aerosols without resort to undue experimentation.

[0176] The compounds, when it is desirable to deliver them systemically, may be formulated for parenteral administration by injection, e.g., by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, e.g., in ampoules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

[0177] Preparations for parenteral administration include sterile aqueous or non-aqueous solutions, suspensions, and emulsions. Examples of non-aqueous solvents are propylene glycol, polyethylene glycol, vegetable oils such as olive oil, and injectable organic esters such as ethyl oleate. Aqueous carriers include water, alcoholic/aqueous solutions, emulsions or suspensions, including saline and buffered media. Parenteral vehicles include sodium chloride solution, Ringer's dextrose, dextrose and sodium chloride, lactated Ringer's, or fixed oils. Intravenous vehicles include fluid and nutrient replenishers, electrolyte replenishers (such as those based on Ringer's dextrose), and the like. Preservatives and other additives may also be present such as, for example, antimicrobials, anti-oxidants, chelating agents, and inert gases and the like. Lower doses will result from other forms of administration, such as intravenous administration. In the event that a response in a subject is insufficient at the initial doses applied, higher doses (or effectively higher doses by a different, more localized delivery route) may be employed to the extent that patient tolerance permits. Multiple doses per day are contemplated to achieve appropriate systemic levels of compounds.

[0178] The data generated according to the invention may optionally be converted into a bar code or other human- or machine-readable form. For example, each line of a bar code may indicate the presence or absence of a specific enzyme or groups of specific enzymes for a particular subject. The bar code data can be compared with a database of information on other subjects or information on the disease to aid in the diagnosis, prognosis or other analysis of the test subject.

[0179] In one embodiment of the invention, the data generated herein is used to select clinical treatment paradigms for cancers. Treatment options, as described herein, may include but are not limited to: radiotherapy, chemotherapy, adjuvant therapy, or any combination of the aforementioned methods. Aspects of treatment that may vary include, but are not limited to: dosages, timing of administration, or duration of therapy; and may or may not be combined with other treatments, which may also vary in dosage, timing, or duration. Another treatment for cancer is surgery, which can be utilized either alone or in combination with any of the aforementioned treatment methods. One of ordinary skill in the medical arts may determine an appro-

priate treatment paradigm based on evaluation of data generated by the methods described herein.

[0180] In certain embodiments, software for calculating and processing the data as described herein can be provided on a computer connected by data link to a data generating device, such as a mass spectrometer, microarray reader or PCR machine. Any standard data link can be used, including serial or parallel cables, radio frequency or infrared telemetry links, LAN connections, WAN connections, etc. Alternatively, data can be transferred by computer-readable medium (e.g., magnetic or optical medium) and read by the software. The data also can be entered directly by the user via user interface, such as a keyboard, monitor, mouse, graphical user interface such as touch screen, etc. The computer may be contained within the data generating device, providing an integrated system for generating raw data, calculating ratios, and displaying such ratios. One or more computers also may be linked to one or more data generating devices and one or more display devices, such as in a local area network or wide area network.

[0181] In one embodiment of the invention, a visual display is used to display the data for the classification, diagnosis, prediction of prognosis and/or therapeutic monitoring. The visual display can be a graphical user interface, such as a monitor, or a printer.

[0182] The data can be processed individually or by a computer. For instance, a computer-implemented method for generating a data structure, tangibly embodied in a computer-readable medium, representing a quantitative value of a set of signature molecules may be performed according to the invention. The quantitative values may be compared with a reference database. Alternatively a qualitative pattern may be compared with a reference database.

[0183] A computer system that may implement the above as a computer program typically may include a main unit connected to both an output device which displays information to a user and an input device which receives input from a user. The main unit generally includes a processor connected to a memory system via an interconnection mechanism. The input device and output device also may be connected to the processor and memory system via the interconnection mechanism.

[0184] One or more output devices may be connected to the computer system. Example output devices include a cathode ray tube (CRT) display, liquid crystal displays (LCD), printers, communication devices such as a modem, and audio output. One or more input devices also may be connected to the computer system. Example input devices include a keyboard, keypad, track ball, mouse, pen and tablet, communication device, and data input devices such as sensors. The subject matter disclosed herein is not limited to the particular input or output devices used in combination with the computer system or to those described herein.

[0185] The computer system may be a general purpose computer system which is programmable using a computer programming language, such as C++, Java, or other language, such as a scripting language or assembly language. The computer system also may include specially-programmed, special purpose hardware such as, for example, an Application-Specific Integrated Circuit (ASIC). In a general purpose computer system, the processor typically is a commercially-available processor, of which the series x86, Celeron, and Pentium processors, available from Intel, and similar devices from AMD and Cyrix, the 680X0 series

microprocessors available from Motorola, the PowerPC microprocessor from IBM and the Alpha-series processors from Digital Equipment Corporation, are examples. Many other processors are available. Such a microprocessor executes a program called an operating system, of which Windows NT, Linux, UNIX, DOS, VMS and OS8 are examples, which controls the execution of other computer programs and provides scheduling, debugging, input/output control, accounting, compilation, storage assignment, data management and memory management, and communication control and related services. The processor and operating system define a computer platform for which application programs in high-level programming languages may be written.

[0186] A memory system typically includes a computer readable and writeable nonvolatile recording medium, of which a magnetic disk, a flash memory and tape are examples. The disk may be removable, such as a “floppy disk,” or permanent, known as a hard drive. A disk has a number of tracks in which signals are stored, typically in binary form, i.e., a form interpreted as a sequence of one and zeros. Such signals may define an application program to be executed by the microprocessor, or information stored on the disk to be processed by the application program. Typically, in operation, the processor causes data to be read from the nonvolatile recording medium into an integrated circuit memory element, which is typically a volatile, random access memory such as a dynamic random access memory (DRAM) or static memory (SRAM). The integrated circuit memory element typically allows for faster access to the information by the processor than does the disk. The processor generally manipulates the data within the integrated circuit memory and then copies the data to the disk after processing is completed. A variety of mechanisms are known for managing data movement between the disk and the integrated circuit memory element, and the subject matter disclosed herein is not limited to such mechanisms. Further, the subject matter disclosed herein is not limited to a particular memory system.

[0187] The subject matter disclosed herein is not limited to a particular computer platform, particular processor, or particular high-level programming language. Additionally, the computer system may be a multiprocessor computer system or may include multiple computers connected over a computer network. It should be understood that each module may be separate modules of a computer program, or may be separate computer programs. Such modules may be operable on separate computers. Data may be stored in a memory system or transmitted between computer systems. The subject matter disclosed herein is not limited to any particular implementation using software or hardware or firmware, or any combination thereof. The various elements of the system, either individually or in combination, may be implemented as a computer program product tangibly embodied in a machine-readable storage device for execution by a computer processor. Various steps of the process may be performed by a computer processor executing a program tangibly embodied on a computer-readable medium to perform functions by operating on input and generating output. Computer programming languages suitable for implementing such a system include procedural programming languages, object-oriented programming languages, and combinations of the two.

[0188] The invention further provides efficient methods of identifying pharmacological agents or lead compounds for agents active in vivo. Generally, the screening methods involve assaying for compounds that beneficially alter enzyme activity in vivo. Such methods according to the invention are adaptable to automated, high-throughput screening of compounds.

[0189] The methods may be used in any subject. For instance animal models of disease may be used to screen multiple putative therapeutic agents in order to assess the activity level of the putative therapeutic agents on particular enzymes associated with disease. For instance, a library of pro-diagnostic reagents having different putative therapeutic agents associated with the carrier can be administered to the animal model. If each therapeutic agent is associated with a unique signature molecule, then the activity of the putative therapeutic agent could be assessed by analyzing the level of signature molecule in the urine as described herein.

[0190] Additionally, the methods may be used for the advancement of personalized medicine. For instance, a set of pro-diagnostic reagents having multiple therapeutic agents, each therapeutic agent associated with a discreet signature molecule could be administered to a subject having a disease to assess which therapeutic agent is most effective in that individual subject. Based on the data, an appropriate therapeutic strategy could be designed. An example of this may be seen in HIV. Protease inhibitors are used therapeutically to inhibit the activity of critical proteases associated with HIV survival and activity. A set of pro-diagnostic reagents could be generated having different enzyme inhibitors as the carrier or part of the carrier. Each enzyme inhibitor is associated with a particular signature molecule, such that the activity of the particular inhibitor on the enzyme can be assessed by monitoring the level of signature in the urine. For instance a particularly active inhibitor would cause a reduced level of signature molecule being transported to the urine.

[0191] Typically, a known active therapeutic agent may serve as a negative control, i.e., the known therapeutic agent is incorporated into a pro-diagnostic reagent. Putative therapeutic agents, also referred to herein as candidate agents encompass numerous chemical classes, although typically they are organic compounds. Preferably, the candidate pharmacological agents are small organic compounds, i.e., those having a molecular weight of more than 50 yet less than about 2500, preferably less than about 1000 and, more preferably, less than about 500. Candidate agents comprise functional chemical groups necessary for structural interactions with polypeptides and/or nucleic acids, and typically include at least an amine, carbonyl, hydroxyl, or carboxyl group, preferably at least two of the functional chemical groups and more preferably at least three of the functional chemical groups. The candidate agents can comprise cyclic carbon or heterocyclic structure and/or aromatic or polyaromatic structures substituted with one or more of the above-identified functional groups. Candidate agents also can be biomolecules such as peptides, saccharides, fatty acids, sterols, isoprenoids, purines, pyrimidines, derivatives or structural analogs of the above, or combinations thereof and the like. Where the agent is a nucleic acid, the agent typically is a DNA or RNA molecule, although modified nucleic acids as defined herein are also contemplated.

[0192] Candidate agents are obtained from a wide variety of sources including libraries of synthetic or natural com-

pounds. For example, numerous means are available for random and directed synthesis of a wide variety of organic compounds and biomolecules, including expression of randomized oligonucleotides, synthetic organic combinatorial libraries, phage display libraries of random peptides, and the like. Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant, and animal extracts are available or readily produced. Additionally, natural and synthetically produced libraries and compounds can be readily be modified through conventional chemical, physical, and biochemical means. Further, known pharmacological agents may be subjected to directed or random chemical modifications such as acylation, alkylation, esterification, amidification, etc. to produce structural analogs of the agents.

[0193] The invention also includes kits having a container housing a pro-diagnostic reagent, wherein the pro-diagnostic reagent comprises a carrier domain linked to a signature producing domain. The kit may also include a second container housing an analytical reagent.

[0194] The kits, also referred to as articles, include pharmaceutical or diagnostic grade compounds of the invention in one or more containers. The article may include instructions or labels promoting or describing the use of the compounds of the invention. For instance, the kit may include instructions for administering the pro-diagnostic reagent to a subject and for analyzing the signature molecule of the pro-diagnostic reagent in a biological sample of the subject.

[0195] As used herein, “promoted” includes all methods of doing business including methods of education, hospital and other clinical instruction, pharmaceutical industry activity including pharmaceutical sales, and any advertising or other promotional activity including written, oral and electronic communication of any form, associated with compositions of the invention.

[0196] “Instructions” can define a component of promotion, and typically involve written instructions on or associated with packaging of compositions of the invention. Instructions also can include any oral or electronic instructions provided in any manner.

[0197] Thus the agents described herein may, in some embodiments, be assembled into pharmaceutical or diagnostic or research kits to facilitate their use in therapeutic, diagnostic or research applications. A kit may include one or more containers housing the components of the invention and instructions for use. Specifically, such kits may include one or more agents described herein, along with instructions describing the intended therapeutic application and the proper administration of these agents. In certain embodiments agents in a kit may be in a pharmaceutical formulation and dosage suitable for a particular application and for a method of administration of the agents.

[0198] The kit may be designed to facilitate use of the methods described herein by physicians and can take many forms. Each of the compositions of the kit, where applicable, may be provided in liquid form (e.g., in solution), or in solid form, (e.g., a dry powder). In certain cases, some of the compositions may be constitutable or otherwise processable (e.g., to an active form), for example, by the addition of a suitable solvent or other species (for example, water or a cell culture medium), which may or may not be provided with the kit. As used herein, “instructions” can define a component of instruction and/or promotion, and typically involve

written instructions on or associated with packaging of the invention. Instructions also can include any oral or electronic instructions provided in any manner such that a user will clearly recognize that the instructions are to be associated with the kit, for example, audiovisual (e.g., videotape, DVD, etc.), Internet, and/or web-based communications, etc. The written instructions may be in a form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which instructions can also reflect approval by the agency of manufacture, use or sale for human administration.

[0199] The kit may contain any one or more of the components described herein in one or more containers. As an example, in one embodiment, the kit may include instructions for mixing one or more components of the kit and/or isolating and mixing a sample and applying to a subject. The kit may include a container housing agents described herein. The agents may be prepared sterilely, packaged in syringe and shipped refrigerated. Alternatively it may be housed in a vial or other container for storage. A second container may have other agents prepared sterilely. Alternatively the kit may include the active agents premixed and shipped in a syringe, vial, tube, or other container.

[0200] The 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. The kit may be sterilized after the accessories are added, thereby allowing the individual accessories in the container to be otherwise unwrapped. The kits can be sterilized using any appropriate sterilization techniques, such as radiation sterilization, heat sterilization, or other sterilization methods known in the art. The kit may also include other components, depending on the specific application, for example, containers, cell media, salts, buffers, reagents, syringes, needles, a fabric, such as gauze, for applying or removing a disinfecting agent, disposable gloves, a support for the agents prior to administration etc.

[0201] The compositions of the kit may be provided as any suitable form, for example, as liquid solutions or as dried powders. When the composition provided is a dry powder, the powder may be reconstituted by the addition of a suitable solvent, which may also be provided. In embodiments where liquid forms of the composition are used, the liquid form may be concentrated or ready to use. The solvent will depend on the compound and the mode of use or administration. Suitable solvents for drug compositions are well known and are available in the literature. The solvent will depend on the compound and the mode of use or administration.

[0202] The kits, in one set of embodiments, may comprise a carrier means being compartmentalized to receive in close confinement one or more container means such as vials, tubes, and the like, each of the container means comprising one of the separate elements to be used in the method. For example, one of the containers may comprise a positive control for an assay. Additionally, the kit may include containers for other components, for example, buffers useful in the assay.

[0203] The present invention also encompasses a finished packaged and labeled pharmaceutical product. This article of manufacture includes the appropriate unit dosage form in an appropriate vessel or container such as a glass vial or

other container that is hermetically sealed. In the case of dosage forms suitable for parenteral administration the active ingredient is sterile and suitable for administration as a particulate free solution. In other words, the invention encompasses both parenteral solutions and lyophilized powders, each being sterile, and the latter being suitable for reconstitution prior to injection. Alternatively, the unit dosage form may be a solid suitable for oral, transdermal, topical or mucosal delivery.

[0204] In a preferred embodiment, the unit dosage form is suitable for intravenous, intramuscular or subcutaneous delivery. Thus, the invention encompasses solutions, preferably sterile, suitable for each delivery route.

[0205] In another preferred embodiment, compositions of the invention are stored in containers with biocompatible detergents, including but not limited to, lecithin, taurocholic acid, and cholesterol; or with other proteins, including but not limited to, gamma globulins and serum albumins.

[0206] As with any pharmaceutical product, the packaging material and container are designed to protect the stability of the product during storage and shipment. Further, the products of the invention include instructions for use or other informational material that advise the physician, technician or patient on how to appropriately prevent or treat the disease or disorder in question. In other words, the article of manufacture includes instruction means indicating or suggesting a dosing regimen including, but not limited to, actual doses, monitoring procedures.

[0207] More specifically, the invention provides an article of manufacture comprising packaging material, such as a box, bottle, tube, vial, container, sprayer, insufflator, intravenous (i.v.) bag, envelope and the like; and at least one unit dosage form of a pharmaceutical agent contained within said packaging material. The invention further provides an article of manufacture comprising a needle or syringe, preferably packaged in sterile form, for injection of the formulation, and/or a packaged alcohol pad.

EXAMPLES

Example 1: Methods

Syntheses of Nanoparticle Pro-Diagnostic Reagents

[0208] 40 nm dextran-coated iron oxide nanoparticles (NP; 115,000 g/mole per iron core) were dissolved in 1×phosphate buffered saline (PBS; 137 mM NaCl, 10 mM Phosphate, 2.7 mM KCl, pH 7.4) at a concentration of 2 mg/mL. Vivotag-750 fluorophore was labeled on the NPs such that each NP has around 2 VT-750 fluorophores. The linker N-succinimidyl iodoacetate (SIA) was dissolved in DMSO at 20 mg/mL. The two solutions were mixed to obtain a 1-to-7 mass ratio between iron oxide NPs and SIA for 2 hr at room temperature with shaking. Size exclusion chromatography (column diameter×height=1 cm×30 cm; media: Sephadex G-50-coarse) was used to separate out the excess SIA and to exchange NPs into borate buffer (50 mM sodium borate, 5 mM EDTA, pH 8.3).

[0209] Each of the 43 fluorescein-labeled peptides; MIT Biopolymers Laboratory) was dissolved in DMSO at 25 mg/mL. Thiol polyethylene glycol (SH-PEG; MW=20 k) was dissolved in borate buffer at 1 mg/mL. Each of the 43 peptides, SH-PEGs, and the activated NPs were left to for >12 hr at molar ratio of 1:20:95=NP:SH-PEG:peptide, mak-

ing 43 different peptide-PEG-NPs. Additional borate buffer was added to the reactants to bring the DMSO to <10% of the total reaction volume.

[0210] After the linkers on the NP surface had reacted with 20 k SH-PEGs and fluorophore-peptides, the five final-product solutions were filtered on centrifugal filter columns (Amicon, Millipore; MW=100 k) at 4,200 rcf to remove the un-conjugated SH-PEGs and peptides to <0.1% of the original conjugated quantity. 1×HEPES salt buffer (100 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid, 150 mM NaCl, pH 7.5) was used to replace the borate buffer and DMSO during centrifugation until <0.1% volume of borate and <0.02% volume of the DMSO were left in the samples.

[0211] The 43 newly made peptide-PEG-NPs (pP-NPs) were analyzed with a spectrometer to assess the number of peptides bound to each nanoparticle. Nanoparticle spectra were normalized to particles that had not been reacted with peptides to allow quantification of the attached fluorophore-peptide absorbance (fluorescein peak absorbance=495 nm; extinction coefficient=72,000×10⁶ cm⁻¹ M⁻¹). NP concentration was assessed by recording its absorbance at 400 nm with an extinction coefficient of 2.07×10⁶ cm⁻¹ M⁻¹. Comparison of these concentrations allowed quantitation of the average peptide-to-NP ratios. All samples were normalized to 5 μM based on peptide concentration and stored in 4° C.

In Vitro Screen of Extracellular Protease Activations on Nanoparticle Pro-Diagnostic Reagents

[0212] The 43 normalized (based on 5 μM peptide concentrations) pP-NPs were aliquoted on a clear half-96-well plate. From the aliquoted “stock” plate, 20 μL of each sample was taken out and put into a black half-area 96-well plate. MMP-2 was dissolved in 1×HEPES salt buffer (with 5 mM CaCl₂) at 64 nM, and the solution was heated at 37° C. for 15 minutes to activate the enzyme. 20 μL of the MMP-2 solution was put into each of all 43 wells (which already contained the pP-NP samples) on the black half-96-well plate such that each well had 32 nM MMP-2 and 2.5 μM peptide concentrations. The plate was read immediately and every 1-min following for 90 min by using a microplate fluorimeter (Molecular Devices Corporation; Gemini EM; excitation: 485 nm, emission: 538 nm, cutoff: 530 nm) to sense the cleaved fluorophore signals. The protease activities for particular peptide substrate were measured by the fluorophore intensity levels over time. The basic physical concept is that when the peptide-fluorophore was attached to the NP iron core, the fluorescence of the fluorophore is quenched by the neighboring absorption iron cores. However, when the peptide were cleaved (now de-quench from iron core), an increase in fluorescence can be detected as it diffuses away from the particle core. Same steps were repeated for MMP-7, MMP-8, MMP-9, and MMP-14.

[0213] Other proteases (Thrombin, factor Xa, tissue factor, and Cathepsin B) dissolved in 1×HEPES (but no CaCl₂) were used to test the enzyme activities at a final concentration of 32 nM. 20 μL Dulbecco’s Modified Eagle Medium (DMEM, GibcoBRL, Rockville, Md.) with 10% fetal bovine serum was put into each well of the prepared black half-96-well plate as well. The DMEM with 10% serum served as the control for the experiment. The initial slope of time over fluorophore intensity level (V₀=milli-unit per minute) was plotted for each trials from each samples. There were total of 43 samples and 10 proteases (MMP-2, MMP-7, MMP-8, MMP-9, MMP-14, thrombin, factor Xa, tissue

factor, Cathepsin B, and DMEM with 10% serum). The data were then collected and analyzed, and the around 10 most effective samples were selected for future in vivo experiment.

HT-1080/MDA-MB-435 Cancer Cell Culture and Implantation

[0214] HT-1080 human fibrosarcoma cells (American Type Culture Collection; ATCC) were cultured in tissue culture flasks (150 cm²) by using cell media solution consist of DMEM (89%), fetal bovine serum (10%; Invitrogen), and penicillin/streptomycin (1%). The cells were grown over several generation (n=4-5) and were then concentrated to around 2×10⁶ million cells/mL in a serum-free solution (66% DMEM, 33% matrix gel, and 1% penicillin/streptomycin). 200 μL of cell solution were injected subcutaneously into each of the nude mice (Nu/Nu; strain code: 088; Charles River Laboratories, Inc., Wilmington, Mass.). For each mouse, 100 μL of the cell solution were injected into each of the nude mouse on the regions of upper right/left thigh (thus, 200 μL total per mouse) below the skin but not into the tissues. The cells would gradually settled and started growing into a sizable tumor (diameter=1 cm) in around 14 days. MDA-MB-435 was processed in the exact same way as HT-1080 with one exception: the injection solution contained 50% 2×10⁶ million cells/mL and 50% DMEM.

Circulation Assessment of Nanoparticle Pro-Diagnostic Reagents In Vivo

[0215] Thrombin-specific pP-NPs samples (2 n-mol based on peptide concentration) were injected into each of four 1-month old adult female white mice (Swiss Webster; strain code: 024; Charles River Laboratories, Inc., Wilmington, Mass.). The reagents (also referred to herein as NP-chaperones) were allowed to circulate in the bloodstream for 5 min (the endpoint of which was defined as 0 hr). 15 μL of blood samples were taken retro-orbitally with heparinized capillary tubes at the following time points: 0 hr, 1 hr, 3 hr, 6 hr, 8 hr, and 12 hr. 10 μL of the blood sample was mixed gently with 40 μL 1×PBS containing 10 mM EDTA to chelate blood calcium and prevent coagulation; the mixture was vortexed and centrifuged to pellet the red blood cells which were then discarded. 30 μL of the supernatant (blood plasma) was put into a well on a 24-by-36-microarray plate.

[0216] The intensity of the NP-linked Vivotag-750 fluorescence in the blood plasma sample was detected with the Odyssey imaging system (Westburg, Leusden, Netherlands). All fluorescence intensity levels were normalized to the 0 hr time point (defined as 100%). Bi-exponential lines of best fit were drawn for all samples to calculate the half-life points.

In Vivo Assessment of Tumor and Clotting-Factor Protease Activity

[0217] Out of the nine selected peptide-PEG-NP (pP-NPs), the first seven peptide were targeted for tumor monitoring while last two peptide were chosen for internal, vascular injury assessment. These nine pP-NPs were re-synthesized in larger quantity (at least 100 n-moles based on peptide concentrations). Each of the nine pP-NPs (around 5 n-moles) was reacted with VT-750 fluorophore (at five-fold molar excess to the peptide concentrations on the NPs) for 2 hr, and the excess un-reactive VT-750 were filtered out to <0.01% by spinning down on 100 k filter column. The

VT-750 was now covalently connected to the end of the peptides on pP-NPs, and we abbreviate these products as 750-pP-NPs.

[0218] The two thrombin-specific-cleavage 750-pP-NP chaperones were each taken out based on 1-nmol VT-750 concentrations, combined together, and re-suspended to 200 uL of 1×PBS. After fed on non-fluorophore diet for >1 wk, 9 nude mice were each injected intravenously with the 200 uL of the 750-pP-NPs; immediately, 4 of the 9 mice were slightly injured on both thigh muscles while 1 of the 9 was injured on only one thigh muscles. All mice were imaged for the bio-distribution of NP chaperones in vivo (Odyssey imaging systems; Westburg, Leusden, Netherlands) by tracing VT-750 for 2-hr period with 10-min intervals. The bio-distributions of 750-pP-NPs over time in each mouse were quantified (ImageJ; NIH), and the organs of interest included bladder, kidney, liver, and spleen. The seven MMP-specific cleavage 750-pP-NP chaperones were each taken out based on 0.3-nmol VT-750 concentrations, combined together, and re-suspended to 200 uL of 1×PBS. After fed on non-fluorophore diet for >1 wk, 9 nude mice (3 with HT-1080 tumors, 3 with MDA-MB-435 tumors, and 3 with no tumor) were each injected intravenously with the 200 uL of the 750-pP-NPs. The bio-distribution of 750-pP-NPs over time in each mouse were traced and quantified for 2-hr period with 10-min intervals. If the NP chaperones have been stored in the fridge for more than 6 hr, it is suggested to spin them down in 100 k filter column in a volume 20-fold of the original samples to get rid of the free peptides and/or VT-750.

[0219] The remaining pP-NPs from the nine selected samples were grouped into seven MMP-specific pP-NPs (MMP:pP-NPs) and two thrombin-specific pP-NPs (thrombin:pP-NPs); the pP-NPs in each group were present in the same amount based on peptide concentrations. 2 n-mol of MMP:pP-NPs (based on peptide concentrations) was injected into each of the 16 nude mice (4 with HT-1080 tumors, 4 with MDA-MB-435 tumors, 4 with no tumor but injury, and 4 with no tumors and no injury); the urine of all 12 mice were collected after 2 hr of injections. 2 n-mol of thrombin:pP-NPs were injected into 8 Swiss Webster mice (immediately after injection, 4 were slightly injured on both thigh muscles) and 4 nude mice with MDA-MB-435 tumors; the urine of all 12 mice were collected after 1 hr of injections. In addition, 2 n-mol of a free thrombin peptide was injected into each of the 8 mice, and 4 were immediately injured afterward; the urine samples of all 8 mice were collected after 1 hr of injections. In the waiting periods between the injections of the NP chaperones and the excretions of the urine, all mice were anesthetized with isoflurane (2-chloro-2-(difluoromethoxy)-1,1,1-trifluoro-ethane). After urine excretions for each mouse, each urine sample volume was diluted to 500 uL by adding ddH₂O. 100 μL of each 500 uL urine sample was put into a well on a black half-96-well plate, and a microplate fluorimeter (Molecular Devices Corporation; Gemini EM; excitation: 485 nm, emission: 538 nm, cutoff: 530 nm) was use to measure the relative fluorophore units (RFU) of each sample. Furthermore, 100 μL of each 250 uL urine sample was analyzed on HPLC-MS-MS for specific peptide sequences injected into the mice.

Statistical Analysis

[0220] Since that the sample sizes for both the nude mice and the white mice were very small and that a normal

distribution could not be assumed, a Student two-tailed t-test was performed for all statistical tests in this study. The null hypothesis was that the two groups (e.g. tumor vs. non-tumor; injured vs. non-injured) do not differ in the amount of peptides excreted, and $p\text{-value} < 0.05$ would reject the null hypothesis.

Example 2: Preparation of Pro-Diagnostic Reagents

[0221] 43 pro-diagnostic reagents composed of peptide-PEG-nanoparticles (pP-NPs) were prepared and analyzed with a spectrometer to assess the number of peptides bound to each nanoparticle. The reagents were designed to enable rapid high throughput screening methods. FIG. 1 is a schematic depicting a method according to the invention for multiplexed in vivo enzyme profiling of mass-coded nanoparticle based pro-diagnostic reagents.

Example 3: In Vitro Screen of Extracellular Protease Activations on Pro-Diagnostic Reagents

[0222] Initially, we established that the pro-diagnostic reagents could target tumors or injuries. We then conducted a screen of 43 pro-diagnostic reagents to fluorescently find optimal sequences for detection of tumor and injury proteases. The protease activities for a particular peptide substrate were measured by the fluorophore intensity levels over time. When the peptide-fluorophore (signature molecule) was attached to the nanoparticle iron core (carrier), the fluorescence of the fluorophore was quenched by the neighboring absorption iron cores. However, when the peptides were cleaved, the fluorophores were no longer quenched and, an increase in fluorescence was detected.

[0223] FIG. 3A is a schematic of the pro-diagnostic reagent, with the circles referring to the carrier, the star is a signature molecule, and the zigzag line refers to the enzyme susceptible domain. FIG. 3B is a graph depicting fluorescence activation versus time. FIG. 3C depicts data on 43 pro-diagnostic reagents (with enzyme susceptible domains listed to the right for detection of tumor and injury enzymes).

[0224] FIG. 4 is a Table depicting the mass detection of ejected fragments in vitro. The results confirmed that the fluorescent results from the screen could also be detected by analyzing the mass of ejected fragments in vitro.

Example 4: Tumor and Wound Targeting with Pro-Diagnostic Reagents

[0225] Tumor and wound targeting with pro-diagnostic reagents shows how the carriers enhance the circulation time of the peptides (which would otherwise clear within minutes) and enable targeting to either tumors or injuries. FIG. 2 shows data depicting the process of tumor and wound targeting with pro-diagnostic reagents. FIG. 2A is a schematic of the pro-diagnostic reagent, with the circles referring to the carrier, the star is a fluorescent molecule, and the zigzag line refers to the enzyme susceptible domain and the signature molecule (darker end region). FIG. 2B is an electron micrograph of the pro-diagnostic reagent. FIG. 2C is a graph depicting the circulation time of the pro-diagnostic reagent, by plotting detection of the carrier in the blood with respect to time after intravenous injection. FIG. 2D is photographs of mice having either tumors or injuries (left and right panels, respectively) administered the pro-diagnostic reagent. FIG. 2E is histopathological analysis of carrier homing to tumors or regions of injury.

Example 5: In Vivo Assessment of Tumor and Clotting-Factor Protease Activity

[0226] The circulation time of pro-diagnostic reagents in vivo was assessed. Optimized pro-diagnostic reagents were injected intravenously into mice and the cleaved peptide (signature component) was fluorescently tracked into the urine after injection. The intensity of the fluorescence in the blood plasma sample was detected with the Odyssey imaging system. FIG. 5 depicts the results of fluorescent detection of urinary reporter activation by tumors and injuries in vivo. FIG. 5A is a schematic of the pro-diagnostic reagent as shown in FIG. 3A, further depicting the portion of the molecule that undergoes renal clearance and the portion that undergoes RES clearance. FIG. 5B is a set of photographs of that were intravenously administered the optimized pro-diagnostic reagent for injury detection (top) or tumor detection (bottom). Half the mice that were administered the optimized pro-diagnostic agents for injury-detection suffered bilateral hind limb injuries (left side of the photograph) while the control mice had no injuries (right side of the photograph). Half the mice administered the optimized pro-diagnostic agents for tumor-detection harbored human fibrosarcoma tumors (HT-1080) (left side of photograph), while the other mice contained no tumors (right side of photograph). FIG. 5C is a set of graphs depicting relative bladder fluorescence for tumor (bottom panel) or injured (top panel) versus control mice in order to track the entrance of cleaved signature peptide into the urine after injection. Bi-exponential lines of best fit were drawn for all samples to calculate the half-life points.

[0227] FIG. 6 shows LC/MS quantitation of signature molecules in urine. FIG. 6A is a photograph of an experimental mouse, having bilateral injury and a control uninjured mouse. FIG. 6B is a graph depicting the ratio of signature molecule (from thrombin cleavable proteolytic susceptible domain) to isotopically labeled product in injured versus control mice.

Example 6: In Vitro Assessment of an Implantable Diagnostic Capsule

[0228] FIG. 7 shows an exemplary embodiment of an implantable diagnostic capsule. FIG. 7A shows a photograph of a typical implantable capsule as described, for example, in Daniel, K. D., et al., *Implantable diagnostic device for cancer monitoring*. Biosensors and Bioelectronics 24 (2009) 3252-3257 which can be used according to the invention for loading nanoparticles. Nanoparticles can be loaded into the well-shaped reservoir and sealed by a semi-permeable polycarbonate membrane enabling free traffic of enzymes and other molecules but limiting the extracapsule diffusion of nanoparticles. FIG. 7B shows measurements of thrombin-cleaved peptide efflux from implantable diagnostic capsules sealed with semi-permeable membranes of different pore sizes. Capsules made with membranes of pore size 10, 30, 50 or 80 nm were loaded with nanoparticles functionalized with GGdFPipRSGGGC (SEQ ID NO: 8) and exposed to solutions of thrombin or factor Xa (a cognate and a non-cognate protease, respectively). Thrombin-specific cleavage was monitored over time by measuring extracapsule fluorescence every 30 minutes, normalizing over factor Xa. The kinetics of reporter release is shown.

Example 7: Optimization of Peptide Reporters for LC/MS Detection

[0229] A series of peptides (A1-A14) of different sequence were investigated to determine optimal sequence length and charge density that would enable facile detection via LC/MS. Three representative proteolytic products, sequences GGVVLS (SEQ ID NO: 19), GGPVG (SEQ ID NO: 13), and GGdFPipR (SEQ ID NO: 17) were selected and appended with peptide caps of differing length and charge density (FIG. 8A). The sequences were then detected via LC/MS. FIG. 8B shows normalized relative intensities of the peptide reporters. The inset of FIG. 8B shows a magnification of the normalized relative intensities of the peptide sequences A1-A6 as measured via LC/MS. In general, hydrophilic sequences appended with shorter caps were more readily detected. For example, peptides A5 and A6, containing the shortest caps of the group of peptides containing the proteolytic product GGVVLS (SEQ ID NO: 19), were more readily detected than other peptides of that group, which contained longer, or less hydrophilic caps. A hydrophilic molecule is a molecule that can transiently hydrogen-bond with water and is, thus, soluble in water and, in some embodiments, in other polar solvents. Hydrophilicity of peptides can be modulated according to methods well known in the art. In some embodiments, removal of the fluorescein tag will increase the hydrophilicity of the peptides reported in this disclosure. In some embodiments, modification of the sequences to contain positively charged residues (Histidine, Lysine, and Arginine) and/or negatively charged residues (glutamic acid and aspartic acid) will increase the hydrophilicity. Similarly, peptides A11 and A12, containing the shortest caps of the group of peptides containing the proteolytic product GGPVG (SEQ ID NO: 13), were more readily detected than other peptides of that group, which contained longer, or less hydrophilic caps. Of the tested cap sequences F1-dR-dS-dR (SEQ ID NO: 20) and F1-dR-G-dS-dR (SEQ ID NO: 21) were determined to be advantageous. Accordingly, for non-fluorescent detection, the cap sequences dR-dS-dR (SEQ ID NO: 35) and dR-G-dS-dR (SEQ ID NO: 36) were determined to be advantageous. Using the results from the optimization experiments as guidelines, a revised list of pro-diagnostic peptides were designed for optimal LC/MS detection. The following pro-diagnostic peptides optimized for LC/MS detection were designed:

Peptide	Sequence
A	F1-dR-dS-dR-G-G-P-Q-G-I-W-G-Q-C (SEQ ID NO: 22)
B	F1-dR-G-dS-dR-G-G-P-L-G-V-R-G-K-C (SEQ ID NO: 23)
C	F1-dR-G-dS-dR-G-G-P-L-A-Nva-Dpa-A-R-G-C (SEQ ID NO: 24)
D	F1-dR-G-dS-dR-G-G-P-V-G-L-I-G-C (SEQ ID NO: 25)
E	F1-dR-dS-dR-G-G-P-V-P-L-S-L-V-M-C (SEQ ID NO: 26)
F	F1-dR-G-dS-dR-G-G-V-V-V-L-S-M-T-A-C (SEQ ID NO: 27)

-continued

Peptide	Sequence
G	F1-dR-G-dS-dR-G-G-S-G-G-P-L-G-L-R-S-W-C (SEQ ID NO: 28)
H	F1-dR-G-dS-dR-G-G-G-P-W-G-I-W-G-Q-G-C (SEQ ID NO: 29)
I	F1-dR-G-G-dS-G-G-dF-Pip-R-S-G-G-G-C (SEQ ID NO: 30)
J	F1-dR-dS-dR-G-G-L-V-P-R-G-S-G-C (SEQ ID NO: 31)
Ia	F1-dR-G-G-dS-G-G-F-P-R-S-G-G-G-C (SEQ ID NO: 32)
Ib	F1-dR-G-G-dS-G-G-G-dF-Pip-K-S-G-G-G-C (SEQ ID NO: 33)
Ic	F1-dR-G-G-dS-G-G-G-dF-P-K-S-G-G-G-C (SEQ ID NO: 34)

Example 8: In Vitro Multiplexed Analysis of Protease Activity by LC/MS

[0230] Two identical cocktails of 12 pro-diagnostic nanoparticles, each functionalized with a different peptide (FIG. 9A), were exposed to thrombin or collagenase and the proteolytically released reporters were analyzed by LC/MS. Six of the peptides contained target peptide sequences for thrombin, whereas the other six peptides contained target peptide sequences for collagenase. LC/MS peak area measurements of all twelve peptides after exposure of the first multiplex cocktail to thrombin and after exposure of the second multiplex cocktail to collagenase are shown in FIG. 9B and FIG. 9C, respectively. In the cocktail exposed to thrombin, peptide reporters cleaved from peptides containing a thrombin target site (left six bars) were predominantly detected over peptide reporters cleaved from peptides containing a collagenase target site (right six bars), as shown in FIG. 9B. In the cocktail exposed to collagenase, peptide reporters cleaved from peptides containing a collagenase target site (right six bars) were predominantly detected over peptide reporters cleaved from peptides containing a thrombin target site (left six bars), as shown in FIG. 9C. FIG. 9D shows the ratio of the LC/MS peak area for each peptide reporter after exposure to thrombin over the peak area measured after exposure to collagenase.

[0231] 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.

[0232] 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: 109

SEQ ID NO: 1 moltype = AA length = 10

FEATURE Location/Qualifiers

REGION 1..10

note = Synthetic polypeptide

source 1..10

mol_type = protein

organism = synthetic construct

SEQUENCE: 1

GGPQGIWGQC

10

SEQ ID NO: 2 moltype = AA length = 10

FEATURE Location/Qualifiers

REGION 1..10

note = Synthetic polypeptide

source 1..10

mol_type = protein

organism = synthetic construct

SEQUENCE: 2

GGPLGVRGKC

10

SEQ ID NO: 3 moltype = AA length = 11

FEATURE Location/Qualifiers

REGION 1..11

note = Synthetic polypeptide

source 1..11

mol_type = protein

organism = synthetic construct

MOD_RES

6

note = Norvaline (Nva)

SITE

7

note = N - beta - (2,4 - dinitrophenyl) - L -alpha,beta-diaminopropionic acid (Dap)

SEQUENCE: 3

GGPLAVXARG C

11

SEQ ID NO: 4 moltype = AA length = 9

FEATURE Location/Qualifiers

REGION 1..9

note = Synthetic polypeptide

source 1..9

mol_type = protein

organism = synthetic construct

SEQUENCE: 4

GGPVGLIGL

9

SEQ ID NO: 5 moltype = AA length = 11

FEATURE Location/Qualifiers

REGION 1..11

note = Synthetic polypeptide

source 1..11

mol_type = protein

organism = synthetic construct

SEQUENCE: 5

GGPVPLSLVM C

11

SEQ ID NO: 6 moltype = AA length = 13

FEATURE Location/Qualifiers

REGION 1..13

note = Synthetic polypeptide

source 1..13

mol_type = protein

organism = synthetic construct

SEQUENCE: 6

GGSGGPLGLR SWC

13

SEQ ID NO: 7 moltype = AA length = 12

FEATURE Location/Qualifiers

REGION 1..12

note = Synthetic polypeptide

source 1..12

mol_type = protein

organism = synthetic construct

SEQUENCE: 7

GGGPWGIWGQ GC

12

-continued

SEQ ID NO: 8	moltype = AA length = 10	
FEATURE	Location/Qualifiers	
REGION	1..10	
	note = Synthetic polypeptide	
source	1..10	
	mol_type = protein	
	organism = synthetic construct	
SITE	4	
	note = Pipecolic acid (Pip)	
SITE	3	
	note = D-glycine	
SEQUENCE: 8		
GGFXRSGGGC		10
SEQ ID NO: 9	moltype = AA length = 10	
FEATURE	Location/Qualifiers	
REGION	1..10	
	note = Synthetic polypeptide	
source	1..10	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 9		
GGLVPRGSGC		10
SEQ ID NO: 10	moltype = AA length = 5	
FEATURE	Location/Qualifiers	
REGION	1..5	
	note = Synthetic polypeptide	
source	1..5	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 10		
GGPQG		5
SEQ ID NO: 11	moltype = AA length = 5	
FEATURE	Location/Qualifiers	
REGION	1..5	
	note = Synthetic polypeptide	
source	1..5	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 11		
GGPLG		5
SEQ ID NO: 12	moltype = AA length = 5	
FEATURE	Location/Qualifiers	
REGION	1..5	
	note = Synthetic polypeptide	
source	1..5	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 12		
GGPLA		5
SEQ ID NO: 13	moltype = AA length = 5	
FEATURE	Location/Qualifiers	
REGION	1..5	
	note = Synthetic polypeptide	
source	1..5	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 13		
GGPVG		5
SEQ ID NO: 14	moltype = AA length = 7	
FEATURE	Location/Qualifiers	
REGION	1..7	
	note = Synthetic polypeptide	
source	1..7	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 14		
GGPVPLS		7
SEQ ID NO: 15	moltype = AA length = 8	

-continued

FEATURE	Location/Qualifiers	
REGION	1..8	
source	note = Synthetic polypeptide 1..8 mol_type = protein organism = synthetic construct	
SEQUENCE: 15		
GGSGGPLG		8
SEQ ID NO: 16	moltype = AA length = 6	
FEATURE	Location/Qualifiers	
REGION	1..6	
source	note = Synthetic polypeptide 1..6 mol_type = protein organism = synthetic construct	
SEQUENCE: 16		
GGGPWG		6
SEQ ID NO: 17	moltype = AA length = 5	
FEATURE	Location/Qualifiers	
REGION	1..5	
source	note = Synthetic polypeptide 1..5 mol_type = protein organism = synthetic construct	
SITE	3	
SITE	note = D-glycine 4	
SEQUENCE: 17	note = Pipecolic acid (Pip)	
GGFXR		5
SEQ ID NO: 18	moltype = AA length = 5	
FEATURE	Location/Qualifiers	
REGION	1..5	
source	note = Synthetic polypeptide 1..5 mol_type = protein organism = synthetic construct	
SEQUENCE: 18		
GGLVP		5
SEQ ID NO: 19	moltype = AA length = 7	
FEATURE	Location/Qualifiers	
REGION	1..7	
source	note = Synthetic polypeptide 1..7 mol_type = protein organism = synthetic construct	
SEQUENCE: 19		
GGVVLS		7
SEQ ID NO: 20	moltype = length =	
SEQUENCE: 20		
000		
SEQ ID NO: 21	moltype = AA length = 4	
FEATURE	Location/Qualifiers	
REGION	1..4	
source	note = Synthetic polypeptide 1..4 mol_type = protein organism = synthetic construct	
SITE	1	
SITE	note = N-terminal Fluorescein 1	
SITE	note = D-isomer 3..4	
SEQUENCE: 21	note = D-isomer	
RGSR		4
SEQ ID NO: 22	moltype = AA length = 13	
FEATURE	Location/Qualifiers	
REGION	1..13	

-continued

source	note = Synthetic polypeptide 1..13 mol_type = protein organism = synthetic construct	
SITE	1 note = N-terminal Fluorescein	
SITE	1..3 note = D-isomer	
SEQUENCE: 22 RSRGGPQGIW GQC		13
SEQ ID NO: 23 FEATURE REGION	moltype = AA length = 14 Location/Qualifiers 1..14 note = Synthetic polypeptide	
SITE	1 note = N-terminal Fluorescein	
SITE	3..4 note = D-isomer	
source	1..14 mol_type = protein organism = synthetic construct	
SITE	1 note = D-isomer	
SEQUENCE: 23 RGSRGGPLGV RGKC		14
SEQ ID NO: 24 FEATURE REGION	moltype = AA length = 15 Location/Qualifiers 1..15 note = Synthetic polypeptide	
SITE	1 note = N-terminal Fluorescein	
SITE	1 note = D-arginine	
SITE	4 note = D-arginine	
MOD_RES	10 note = Norvaline (Nva)	
SITE	11 note = N - beta - (2,4 - dinitrophenyl) - L -alpha,beta- diaminopropionic acid (Dap)	
source	1..15 mol_type = protein organism = synthetic construct	
SITE	3 note = D-serine	
SEQUENCE: 24 RGSRGGPLAV XARGC		15
SEQ ID NO: 25 FEATURE REGION	moltype = AA length = 13 Location/Qualifiers 1..13 note = Synthetic polypeptide	
SITE	1 note = N-terminal Fluorescein	
SITE	1 note = D-isomer	
SITE	3..4 note = D-isomer	
source	1..13 mol_type = protein organism = synthetic construct	
SEQUENCE: 25 RGSRGGPVGL IGC		13
SEQ ID NO: 26 FEATURE REGION	moltype = AA length = 14 Location/Qualifiers 1..14 note = Synthetic polypeptide	
SITE	1 note = N-terminal Fluorescein	
SITE	1..3 note = D-isomer	
source	1..14 mol_type = protein	

-continued

SEQUENCE: 26	organism = synthetic construct	
RSRGGPVPLS LVMC		14
SEQ ID NO: 27	moltype = AA length = 15	
FEATURE	Location/Qualifiers	
REGION	1..15	
	note = Synthetic polypeptide	
SITE	1	
	note = N-terminal Fluorescein	
SITE	1	
	note = D-isomer	
SITE	3..4	
	note = D-isomer	
source	1..15	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 27		
RGSRRGGVVVL SMTAC		15
SEQ ID NO: 28	moltype = AA length = 17	
FEATURE	Location/Qualifiers	
REGION	1..17	
	note = Synthetic polypeptide	
SITE	1	
	note = N-terminal Fluorescein	
SITE	1	
	note = D-isomer	
SITE	3..4	
	note = D-isomer	
source	1..17	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 28		
RGSRRGGSGGP LGLRSWC		17
SEQ ID NO: 29	moltype = AA length = 16	
FEATURE	Location/Qualifiers	
REGION	1..16	
	note = Synthetic polypeptide	
SITE	1	
	note = N-terminal Fluorescein	
SITE	1	
	note = D-isomer	
source	1..16	
	mol_type = protein	
	organism = synthetic construct	
SITE	3..4	
	note = D-isomer	
SEQUENCE: 29		
RGSRRGGGPWG IWGQGC		16
SEQ ID NO: 30	moltype = AA length = 14	
FEATURE	Location/Qualifiers	
REGION	1..14	
	note = Synthetic polypeptide	
SITE	1	
	note = N-terminal Fluorescein	
SITE	4	
	note = D-serine	
SITE	7	
	note = D-phenylalanine	
SITE	8	
	note = Pipecolic acid (Pip)	
source	1..14	
	mol_type = protein	
	organism = synthetic construct	
SITE	1	
	note = D-arginine	
SEQUENCE: 30		
RGGSGGFXRS GGGC		14
SEQ ID NO: 31	moltype = AA length = 13	
FEATURE	Location/Qualifiers	
REGION	1..13	
	note = Synthetic polypeptide	

-continued

SITE	1	
	note = N-terminal Fluorescein	
SITE	1..3	
	note = D-isomer	
source	1..13	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 31		
RSRGGLVPRG SGC		13
SEQ ID NO: 32	moltype = AA length = 14	
FEATURE	Location/Qualifiers	
REGION	1..14	
	note = Synthetic polypeptide	
SITE	1	
	note = N-terminal Fluorescein	
SITE	1	
	note = D-isomer	
SITE	4	
	note = D-isomer	
source	1..14	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 32		
RGGSGGFPRS GGGC		14
SEQ ID NO: 33	moltype = AA length = 15	
FEATURE	Location/Qualifiers	
REGION	1..15	
	note = Synthetic polypeptide	
SITE	1	
	note = N-terminal Fluorescein	
SITE	1	
	note = D-arginine	
SITE	8	
	note = D-phenylalanine	
SITE	9	
	note = Pipecolic acid (Pip)	
source	1..15	
	mol_type = protein	
	organism = synthetic construct	
SITE	4	
	note = D-serine	
SEQUENCE: 33		
RGGSGGGFXXK SGGGC		15
SEQ ID NO: 34	moltype = AA length = 15	
FEATURE	Location/Qualifiers	
REGION	1..15	
	note = Synthetic polypeptide	
SITE	1	
	note = N-terminal Fluorescein	
SITE	1	
	note = D-isomer	
SITE	8	
	note = D-isomer	
source	1..15	
	mol_type = protein	
	organism = synthetic construct	
SITE	4	
	note = D-isomer	
SEQUENCE: 34		
RGGSGGGFPPK SGGGC		15
SEQ ID NO: 35	moltype = length =	
SEQUENCE: 35		
000		
SEQ ID NO: 36	moltype = AA length = 4	
FEATURE	Location/Qualifiers	
REGION	1..4	
	note = Synthetic polypeptide	
SITE	1	
	note = D-isomer	
SITE	3..4	
	note = D-isomer	

-continued

source	1..4 mol_type = protein organism = synthetic construct	
SEQUENCE: 36 RGSR		4
SEQ ID NO: 37 FEATURE REGION	moltype = AA length = 7 Location/Qualifiers 1..7 note = Synthetic polypeptide	
VARIANT	6 note = X may be Leucine or Isoleucine	
source	1..7 mol_type = protein organism = synthetic construct	
SEQUENCE: 37 SKGGFXX		7
SEQ ID NO: 38 FEATURE REGION	moltype = AA length = 8 Location/Qualifiers 1..8 note = Synthetic polypeptide	
source	1..8 mol_type = protein organism = synthetic construct	
SEQUENCE: 38 SKGGGLVP		8
SEQ ID NO: 39 FEATURE REGION	moltype = AA length = 7 Location/Qualifiers 1..7 note = Synthetic Polypeptide	
source	1..7 mol_type = protein organism = synthetic construct	
SEQUENCE: 39 SKGGFPR		7
SEQ ID NO: 40 FEATURE REGION	moltype = AA length = 8 Location/Qualifiers 1..8 note = Synthetic Polypeptide	
VARIANT	7 note = X may be Leucine or Isoleucine	
source	1..8 mol_type = protein organism = synthetic construct	
SEQUENCE: 40 SKGGGFXX		8
SEQ ID NO: 41 FEATURE REGION	moltype = AA length = 8 Location/Qualifiers 1..8 note = Synthetic Polypeptide	
source	1..8 mol_type = protein organism = synthetic construct	
SEQUENCE: 41 SKGGGFPK		8
SEQ ID NO: 42 FEATURE REGION	moltype = AA length = 9 Location/Qualifiers 1..9 note = Synthetic Polypeptide	
source	1..9 mol_type = protein organism = synthetic construct	
SEQUENCE: 42 SKGSGPLG		9
SEQ ID NO: 43 FEATURE REGION	moltype = AA length = 8 Location/Qualifiers 1..8 note = Synthetic Polypeptide	
source	1..8 mol_type = protein	

-continued

SEQUENCE: 43	organism = synthetic construct	
SKGGGPWG		8
SEQ ID NO: 44	moltype = AA length = 5	
FEATURE	Location/Qualifiers	
REGION	1..5	
	note = Synthetic Polypeptide	
source	1..5	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 44		
SKGGP		5
SEQ ID NO: 45	moltype = AA length = 7	
FEATURE	Location/Qualifiers	
REGION	1..7	
	note = Synthetic Polypeptide	
source	1..7	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 45		
SKGGPVP		7
SEQ ID NO: 46	moltype = AA length = 7	
FEATURE	Location/Qualifiers	
REGION	1..7	
	note = Synthetic Polypeptide	
source	1..7	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 46		
SKGGPLG		7
SEQ ID NO: 47	moltype = AA length = 7	
FEATURE	Location/Qualifiers	
REGION	1..7	
	note = Synthetic Polypeptide	
source	1..7	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 47		
SKGGPQG		7
SEQ ID NO: 48	moltype = AA length = 7	
FEATURE	Location/Qualifiers	
REGION	1..7	
	note = Synthetic Polypeptide	
source	1..7	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 48		
SKGGPVG		7
SEQ ID NO: 49	moltype = AA length = 7	
FEATURE	Location/Qualifiers	
REGION	1..7	
	note = Synthetic Polypeptide	
source	1..7	
	mol_type = protein	
	organism = synthetic construct	
SITE	6	
	note = Pipecolic acid	
SEQUENCE: 49		
SKGGFXR		7
SEQ ID NO: 50	moltype = AA length = 8	
FEATURE	Location/Qualifiers	
REGION	1..8	
	note = Synthetic Polypeptide	
SITE	7	
	note = Pipecolic acid	
source	1..8	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 50		

-continued

SKGGGFXXK		8
SEQ ID NO: 51	moltype = AA length = 8	
FEATURE	Location/Qualifiers	
REGION	1..8	
	note = Synthetic Polypeptide	
source	1..8	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 51		
SKGGGPWG		8
SEQ ID NO: 52	moltype = AA length = 12	
FEATURE	Location/Qualifiers	
REGION	1..12	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
SITE	9	
	note = N-3-(2,4-dinitrophenyl)- L-2,3-diaminopropionyl (Dpa)	
source	1..12	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 52		
SKGGPLGLXA RC		12
SEQ ID NO: 53	moltype = AA length = 13	
FEATURE	Location/Qualifiers	
REGION	1..13	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
MOD_RES	8	
	note = Norvaline (Nva)	
SITE	9	
	note = N-3-(2,4-dinitrophenyl)- L-2,3-diaminopropionyl (Dpa)	
source	1..13	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 53		
SKGGPLAVXA RGC		13
SEQ ID NO: 54	moltype = AA length = 13	
FEATURE	Location/Qualifiers	
REGION	1..13	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..13	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 54		
SKGGVVPMSM MAC		13
SEQ ID NO: 55	moltype = AA length = 12	
FEATURE	Location/Qualifiers	
REGION	1..12	
	note = Synthetic	
SITE	11	
	note = D-isomer	
source	1..12	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 55		
SKGGPLGLWA RC		12
SEQ ID NO: 56	moltype = AA length = 12	
FEATURE	Location/Qualifiers	
REGION	1..12	
	note = Synthetic	

-continued

MOD_RES	8	
	note = modified by maleimidocaproyl	
SITE	11	
	note = D-isomer	
source	1..12	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 56		
SKGGPLGCHA RC		12
SEQ ID NO: 57	moltype = AA length = 12	
FEATURE	Location/Qualifiers	
REGION	1..12	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..12	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 57		
SKGGPLALWA RC		12
SEQ ID NO: 58	moltype = AA length = 12	
FEATURE	Location/Qualifiers	
REGION	1..12	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..12	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 58		
SKGGPQGIWG QC		12
SEQ ID NO: 59	moltype = AA length = 12	
FEATURE	Location/Qualifiers	
REGION	1..12	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..12	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 59		
SKGGPLGVRG KC		12
SEQ ID NO: 60	moltype = AA length = 11	
FEATURE	Location/Qualifiers	
REGION	1..11	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..11	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 60		
SKGGPVGLIG C		11
SEQ ID NO: 61	moltype = AA length = 13	
FEATURE	Location/Qualifiers	
REGION	1..13	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..13	
	mol_type = protein	

-continued

SEQUENCE: 61	organism = synthetic construct	
SKGGDIPVSL RSC		13
SEQ ID NO: 62	moltype = AA length = 13	
FEATURE	Location/Qualifiers	
REGION	1..13	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..13	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 62		
SKGGDVPAEL TAC		13
SEQ ID NO: 63	moltype = AA length = 13	
FEATURE	Location/Qualifiers	
REGION	1..13	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..13	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 63		
SKGGFIIAGL VGC		13
SEQ ID NO: 64	moltype = AA length = 14	
FEATURE	Location/Qualifiers	
REGION	1..14	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..14	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 64		
SKGGGPWGIW GQGC		14
SEQ ID NO: 65	moltype = AA length = 13	
FEATURE	Location/Qualifiers	
REGION	1..13	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..13	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 65		
SKGGLRLAAI TAC		13
SEQ ID NO: 66	moltype = AA length = 13	
FEATURE	Location/Qualifiers	
REGION	1..13	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..13	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 66		
SKGGESLAYY TAC		13
SEQ ID NO: 67	moltype = AA length = 13	
FEATURE	Location/Qualifiers	

-continued

REGION	1..13	
	note = Synthetic	
SITE	1	
	note = D-isomer	
source	1..13	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 67		
SKGGRSL SRL TAC		13
SEQ ID NO: 68	moltype = AA length = 13	
FEATURE	Location/Qualifiers	
REGION	1..13	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..13	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 68		
SKGGNRYSSL TAC		13
SEQ ID NO: 69	moltype = AA length = 13	
FEATURE	Location/Qualifiers	
REGION	1..13	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..13	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 69		
SKGGGAVSWL LTC		13
SEQ ID NO: 70	moltype = AA length = 13	
FEATURE	Location/Qualifiers	
REGION	1..13	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..13	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 70		
SKGGANISDL TAC		13
SEQ ID NO: 71	moltype = AA length = 13	
FEATURE	Location/Qualifiers	
REGION	1..13	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..13	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 71		
SKGGNKPFSM MMC		13
SEQ ID NO: 72	moltype = AA length = 13	
FEATURE	Location/Qualifiers	
REGION	1..13	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..13	
	mol_type = protein	

-continued

SEQUENCE: 72	organism = synthetic construct	
SKGGRPKPQQ FWC		13
SEQ ID NO: 73	moltype = AA length = 12	
FEATURE	Location/Qualifiers	
REGION	1..12	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
MOD_RES	10	
	note = Norvaline (Nva)	
source	1..12	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 73		
SKGGPKPLAV WC		12
SEQ ID NO: 74	moltype = AA length = 13	
FEATURE	Location/Qualifiers	
REGION	1..13	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
MOD_RES	10	
	note = Norvaline (Nva)	
source	1..13	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 74		
SKGGPKPYAV WMC		13
SEQ ID NO: 75	moltype = AA length = 13	
FEATURE	Location/Qualifiers	
REGION	1..13	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..13	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 75		
SKGGPVPLSL VMC		13
SEQ ID NO: 76	moltype = AA length = 13	
FEATURE	Location/Qualifiers	
REGION	1..13	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..13	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 76		
SKGGIIVMEL VMC		13
SEQ ID NO: 77	moltype = AA length = 12	
FEATURE	Location/Qualifiers	
REGION	1..12	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..12	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 77		

-continued

SKGGRPKALW RC		12
SEQ ID NO: 78	moltype = AA length = 12	
FEATURE	Location/Qualifiers	
REGION	1..12	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..12	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 78		
SKGGPLAYWA RC		12
SEQ ID NO: 79	moltype = AA length = 13	
FEATURE	Location/Qualifiers	
REGION	1..13	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..13	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 79		
SKGGVVVLSM TAC		13
SEQ ID NO: 80	moltype = AA length = 14	
FEATURE	Location/Qualifiers	
REGION	1..14	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..14	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 80		
SKGGGVVPLS LRSC		14
SEQ ID NO: 81	moltype = AA length = 13	
FEATURE	Location/Qualifiers	
REGION	1..13	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..13	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 81		
SKGGFIPESL RMC		13
SEQ ID NO: 82	moltype = AA length = 13	
FEATURE	Location/Qualifiers	
REGION	1..13	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..13	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 82		
SKGGGPLPLR SWC		13
SEQ ID NO: 83	moltype = AA length = 15	
FEATURE	Location/Qualifiers	
REGION	1..15	
	note = Synthetic	

-continued

SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..15	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 83		
SKGGSGGPLG LRSWC		15
SEQ ID NO: 84	moltype = AA length = 11	
FEATURE	Location/Qualifiers	
REGION	1..11	
	note = Synthetic	
SITE	6	
	note = beta-cyclohexyl-L-alanine (Cha)	
MOD_RES	8	
	note = Norvaline (Nva)	
source	1..11	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 84		
SKGGPXGVHA C		11
SEQ ID NO: 85	moltype = AA length = 11	
FEATURE	Location/Qualifiers	
REGION	1..11	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..11	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 85		
SKGGRAMHMY C		11
SEQ ID NO: 86	moltype = AA length = 8	
FEATURE	Location/Qualifiers	
REGION	1..8	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..8	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 86		
SKGRRGC		8
SEQ ID NO: 87	moltype = AA length = 12	
FEATURE	Location/Qualifiers	
REGION	1..12	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
SITE	9	
	note = X is Dap	
SITE	10	
	note = X is Dnp	
source	1..12	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 87		
SKGGIVRAXX GC		12
SEQ ID NO: 88	moltype = AA length = 13	
FEATURE	Location/Qualifiers	
REGION	1..13	
	note = Synthetic	
SITE	1	
	note = D-isomer	

-continued

SITE	2	
	note = D-isomer	
SITE	7	
	note = modified by ethyl group	
source	1..13	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 88		
SKGGPICFFR LGC		13
SEQ ID NO: 89	moltype = AA length = 12	
FEATURE	Location/Qualifiers	
REGION	1..12	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
SITE	6	
	note = X is pipecolic acid	
source	1..12	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 89		
SKGGFXRSGG GC		12
SEQ ID NO: 90	moltype = AA length = 12	
FEATURE	Location/Qualifiers	
REGION	1..12	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..12	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 90		
SKGGLVPRGS GC		12
SEQ ID NO: 91	moltype = AA length = 10	
FEATURE	Location/Qualifiers	
REGION	1..10	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..10	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 91		
SKGGARVRGC		10
SEQ ID NO: 92	moltype = AA length = 10	
FEATURE	Location/Qualifiers	
REGION	1..10	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..10	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 92		
SKGGAKVRGC		10
SEQ ID NO: 93	moltype = AA length = 10	
FEATURE	Location/Qualifiers	
REGION	1..10	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	

-continued

source	1..10	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 93		
SKGGIEGRGC		10
SEQ ID NO: 94	moltype = AA length = 10	
FEATURE	Location/Qualifiers	
REGION	1..10	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..10	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 94		
SKGGAFGRGC		10
SEQ ID NO: 95	moltype = AA length = 16	
FEATURE	Location/Qualifiers	
REGION	1..16	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	4	
	note = D-isomer	
SITE	7	
	note = D-isomer	
SITE	10	
	note = D-isomer	
SITE	11	
	note = D-isomer	
source	1..16	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 95		
RGGSGGSGGS RVVLS		16
SEQ ID NO: 96	moltype = AA length = 15	
FEATURE	Location/Qualifiers	
REGION	1..15	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	4	
	note = D-isomer	
SITE	7	
	note = D-isomer	
SITE	8	
	note = D-isomer	
source	1..15	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 96		
RGGSGGSRGG VVLS		15
SEQ ID NO: 97	moltype = AA length = 15	
FEATURE	Location/Qualifiers	
REGION	1..15	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	4	
	note = D-isomer	
SITE	5	
	note = D-isomer	
SITE	8	
	note = D-isomer	
source	1..15	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 97		
RGGSRGGSGG VVLS		15

-continued

SEQ ID NO: 98	moltype = AA length = 12	
FEATURE	Location/Qualifiers	
REGION	1..12	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	4	
	note = D-isomer	
SITE	5	
	note = D-isomer	
source	1..12	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 98		
RGGSRRGGVVV LS		12
SEQ ID NO: 99	moltype = AA length = 11	
FEATURE	Location/Qualifiers	
REGION	1..11	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	3	
	note = D-isomer	
SITE	4	
	note = D-isomer	
source	1..11	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 99		
RGSRRGGVVVL S		11
SEQ ID NO: 100	moltype = AA length = 10	
FEATURE	Location/Qualifiers	
REGION	1..10	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
SITE	3	
	note = D-isomer	
source	1..10	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 100		
RSRRGGVVVLS		10
SEQ ID NO: 101	moltype = AA length = 14	
FEATURE	Location/Qualifiers	
REGION	1..14	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	4	
	note = D-isomer	
SITE	7	
	note = D-isomer	
SITE	10	
	note = D-isomer	
SITE	11	
	note = D-isomer	
source	1..14	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 101		
RGGSGGSGGS RPVG		14
SEQ ID NO: 102	moltype = AA length = 13	
FEATURE	Location/Qualifiers	
REGION	1..13	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	4	
	note = D-isomer	

-continued

SITE	7	
	note = D-isomer	
SITE	8	
	note = D-isomer	
source	1..13	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 102		
RGGSGGSRGG PVG		13
SEQ ID NO: 103	moltype = AA length = 13	
FEATURE	Location/Qualifiers	
REGION	1..13	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	4	
	note = D-isomer	
SITE	5	
	note = D-isomer	
SITE	8	
	note = D-isomer	
source	1..13	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 103		
RGGSRGGSGG PVG		13
SEQ ID NO: 104	moltype = AA length = 10	
FEATURE	Location/Qualifiers	
REGION	1..10	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	4	
	note = D-isomer	
SITE	5	
	note = D-isomer	
source	1..10	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 104		
RGGSRGGPVG		10
SEQ ID NO: 105	moltype = AA length = 9	
FEATURE	Location/Qualifiers	
REGION	1..9	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	3	
	note = D-isomer	
SITE	4	
	note = D-isomer	
source	1..9	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 105		
RGSRGGPVG		9
SEQ ID NO: 106	moltype = AA length = 8	
FEATURE	Location/Qualifiers	
REGION	1..8	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
SITE	3	
	note = D-isomer	
source	1..8	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 106		
RSRGGPVG		8

-continued

SEQ ID NO: 107	moltype = AA length = 12	
FEATURE	Location/Qualifiers	
REGION	1..12	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	4	
	note = D-isomer	
SITE	7	
	note = D-isomer	
SITE	10	
	note = D-isomer	
SITE	11	
	note = X is pipecolic acid (Pip)	
source	1..12	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 107		
RGGSGGSGGF XR		12
SEQ ID NO: 108	moltype = AA length = 9	
FEATURE	Location/Qualifiers	
REGION	1..9	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	4	
	note = D-isomer	
SITE	7	
	note = D-isomer	
SITE	8	
	note = X is pipecolic acid (Pip)	
source	1..9	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 108		
RGGSGGFXR		9
SEQ ID NO: 109	moltype = AA length = 5	
FEATURE	Location/Qualifiers	
REGION	1..5	
	note = Synthetic	
SITE	1	
	note = D-isomer	
SITE	2	
	note = D-isomer	
source	1..5	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 109		
GGPVP		5

1-88. (canceled)

89. A method comprising:

introducing a diagnostic reagent, wherein said diagnostic reagent comprises a carrier linked to a signature molecule by a peptide, wherein said peptide is cleaved by a protease wherein said cleavage releases said signature molecule, and wherein said diagnostic reagent is not coupled to a chip;

performing an assay to detect said signature molecule cleaved from said carrier; and

evaluating a disease based on a quantifiable amount of said signature molecule,

wherein evaluating said disease is based on an amount of said released signature molecule, and

wherein said disease is a blood pathology.

90. The method of claim 89, wherein said detection is conducted in a solution.

91. The method of claim 89, wherein said signature molecule comprises a peptide sequence, a nucleic acid, a

small molecule, a fluorophore, a carbohydrate, a particle, a radiolabel, a MRI-active compound, an inorganic material, and/or an organic material.

92. The method of claim 91, wherein said signature molecule comprises encoded characteristics to facilitate optimal detection.

93. The method of claim 92, wherein said encoded characteristics are specific to detection using mass spectrometry.

94. The method of claim 91, wherein said diagnostic reagent further comprises a quencher configured to quench the fluorophore.

95. The method of claim 89, wherein said cleavage of said peptide generates a detectable signal.

96. The method of claim 89, wherein said protease cleaves said peptide in vivo.

97. The method of claim 89, wherein said protease cleaves said peptide in vitro.

98. The method of claim **89**, wherein said protease cleaves said peptide ex vivo.

99. The method of claim **89**, wherein evaluating said disease comprises analyzing a disease metric of said disease.

100. The method of claim **99**, wherein said disease metric includes a genetic predisposition for development of a particular disease.

101. The method of claim **99**, wherein said evaluating of said disease is non-invasive to a subject.

102. The method of claim **89**, wherein said protease comprises MMP-2, MMP-7, MMP-8, MMP-9, MMP-14, thrombin, factor Xa, tissue factor, cathepsin B, or a combination thereof.

103. The method of claim **89**, wherein said signature molecule is detected at a level above a normal value.

104. The method of claim **89**, wherein said carrier comprises a particle, and wherein said particle is a microparticle or a nanoparticle.

105. A method comprising:

introducing a diagnostic reagent, wherein said diagnostic reagent comprises a carrier linked to a signature mol-

ecule by a peptide, wherein said peptide is cleaved by a protease, wherein said cleavage releases said signature molecule, and wherein said carrier comprises a targeting agent;

performing an assay to detect said signature molecule cleaved from said carrier; and
evaluating a disease based on a quantifiable amount of said signature molecule.

106. The method of claim **105**, wherein said targeting agent is selected from the group consisting of a glycoprotein, an antibody, and a binding protein.

107. The method of claim **105**, wherein said signature molecule comprises a peptide sequence, a nucleic acid, a small molecule, a fluorophore, a carbohydrate, a particle, a radiolabel, a MRI-active compound, an inorganic material, and/or an organic material.

108. The method of claim **105**, wherein said protease comprises MMP-2, MMP-7, MMP-8, MMP-9, MMP-14, thrombin, factor Xa, tissue factor, cathepsin B, or a combination thereof.

* * * * *