



- (51) **International Patent Classification:**
A61K 38/16 (2006.01) *A61K 38/08* (2019.01)
- (21) **International Application Number:**
PCT/US2024/015002
- (22) **International Filing Date:**
08 February 2024 (08.02.2024)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
63/444,409 09 February 2023 (09.02.2023) US
- (71) **Applicant: THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY** [US/US];
Office of the General Counsel, Building 170, 3rd Floor, Main Quad, P.O. Box 20386, Stanford, California 94305-2038 (US).
- (72) **Inventors: LONG, Jonathan Z.;** Stanford Office of Technology Licensing, 415 Broadway Street, 2nd Floor, MC8854, Redwood City, California 94063 (US).
WIGGENHORN, Amanda L.; Stanford Office of Tech-

nology Licensing, 415 Broadway Street, 2nd Floor, MC8854, Redwood City, California 94063 (US).

(74) **Agent: FIELD, Bret E.;** 201 Redwood Shores Parkway, Suite 200, Redwood City, California 94065 (US).

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

(84) **Designated States** (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ,

(54) **Title:** CAPPED PEPTIDES AND METHODS FOR USING THE SAME

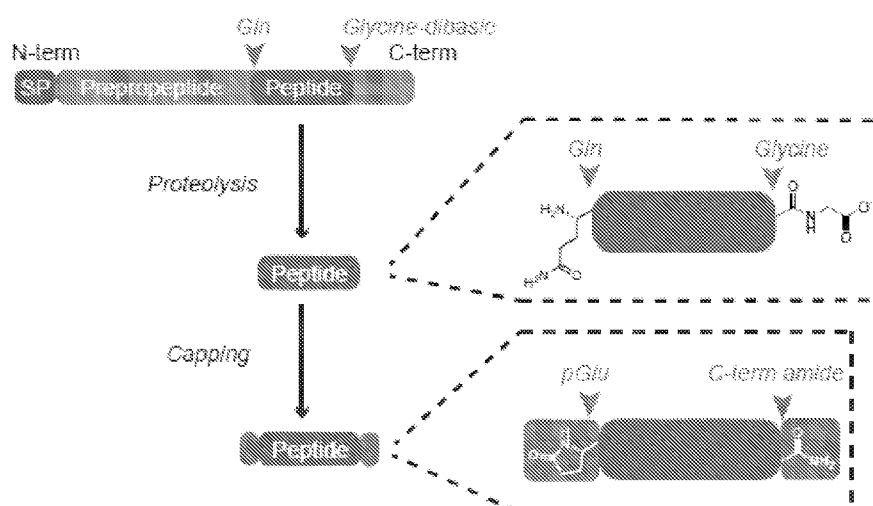


FIG. 1A

(57) **Abstract:** Capped peptides are provided. Aspects of the capped peptides include an N-terminal pyroglutamylation modification and a C-terminal amidation modification. In some instances, the capped peptides are not Thyrotropin-releasing hormone (TRH) or (Gonadotropin-releasing hormone (GnRH)/luteinizing hormone-releasing hormone (LHRH). Also provided are methods of administering capped peptides to a subject, e.g., to treat the subject for a condition, such as a disorder or disease condition, and pharmaceutical compositions that include the capped peptides.

DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT,
LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE,
SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN,
GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

- *without international search report and to be republished
upon receipt of that report (Rule 48.2(g))*
- *with sequence listing part of description (Rule 5.2(a))*

CAPPED PEPTIDES AND METHODS FOR USING THE SAME

CROSS-REFERENCE

Pursuant to 35 U.S.C. § 119 (e), this application claims priority to the filing date of U.S. Provisional Patent Application No. 63/444,409, filed February 9, 2023, which
5 application is incorporated herein by reference in its entirety.

INCORPORATION BY REFERENCE OF SEQUENCE LISTING XML FILE

A Sequence Listing is provided herewith as a Sequence Listing XML, "STAN-1907WO_S21-391_SEQ_LIST", created on February 8, 2024 and having a size of 389,698 bytes. The contents of the Sequence Listing XML are incorporated herein by
10 reference in their entirety.

INTRODUCTION

Peptide hormones and neuropeptides are fundamental signaling molecules that mediate cell-cell communication (Hook et al., J. Am. Soc. Mass Spectrom. 2018 29:807-816). These signaling molecules are produced by proteolytic cleavage of secreted
15 preproteins and released extracellularly via the classical secretory pathway. Once secreted, peptide hormones and neuropeptides act on cognate receptors to regulate nearly all aspects of homeostasis and physiology. Because of their potent and powerful physiologic actions, peptide hormones and neuropeptides have attracted considerable pharmaceutical interest as starting points for the development of therapeutics across for
20 multiple human diseases (Sanyal et al., Lancet. 2018 392: 2705-2717; Sikich et al., N. Engl. J. Med. 2021 385:1462-1473; Wilding et al., N. Engl. J. Med. 2021 384:989-1002).

The biosynthesis of signaling peptides often involves post-translational enzymatic modification of C- and/or N-termini. Approximately half of all peptide hormones and neuropeptides possess a C-terminal amide instead of a free carboxylate (e.g., GLP-1)
25 (Eipper et al., Annu. Rev. Neurosci. 1992 15:57-85). This modification is installed by the action of PAM (peptidyl-glycine alpha-amidating monooxygenase), an enzyme that

cleaves off a C-terminal glycine to leave an amide terminus (Czyzyk et al., Dev. Biol. 2005 287:301-313; Eipper et al., Annu. Rev. Neurosci. 1983 15:57-85). In addition, chemical modifications of the N-terminus include N-acetylation (Guo et al., Proc. Natl. Acad. Sci. 2004 101:11797-11802), N-formylation (Chen et al., Nat. Commun. 2022 13:1-13), and
5 N-pyroglutamylation (Lindemans et al., Front. Endocrinol. 2010 2:1-6). Combinations of C- and N-terminal modifications can also be found within the same peptide (Lindemans et al., Front. Endocrinol. 2010 2:1-6). In the cases where they have been examined, C- and N-termini modifications can both increase peptide stability and also enhance receptor affinity (Van Coillie et al., Nucleic Acids Res. 1998 51:D523-D531; Goren and Baucé, Mol
10 Pharmacol. 1977 13:606-614).

Historically, peptide hormones and neuropeptides-like molecules have been discovered via laborious and time-consuming activity-guided fractionation strategies (Harris and Lerner, Nature 1957 179:1346-1347; Schally et al., Science. 1971 173:1036-1038). More recently, shotgun mass spectrometry-based approaches using available
15 mammalian genome information have been developed for the detection of peptide fragments from tissues and fluids (Fälth et al., Mol. Cell. Proteomics 2007 6:1188-1197; Gupta et al., J. Proteome Res. 2010 9:5065-5075; Tagore et al., Nat. Chem. Biol. 2009 5:23-25). However, the existing mass spectrometry methods typically identify few peptides (e.g., low coverage). In addition, the potential bioactivity and/or signaling of the
20 peptides detected remains undefined.

SUMMARY

The present inventors reasoned that knowledge of primary sequences indicative of post-translational processing of peptide termini, combined with a targeted mass
25 spectrometry approach using authentic peptide standards, might enable de novo prediction and high-sensitivity detection of previously unknown peptide cleavage products. Importantly, these predicted cleavage products would harbor C- and N-terminal modifications that resemble that of other peptide hormones and neuropeptides, thereby suggesting also potential signaling activity. Using this strategy, the inventors report here
30 the discovery of a large and sequence-diverse class of orphan blood-borne mammalian

peptides called “capped peptides.” Capped peptides are defined by the co-incident presence of two post-translational modifications, N-terminal pyroglutamylation and C-terminal amidation, that function as terminal “caps” of the intervening peptide sequence. Capped peptides also exhibit regulatory characteristics similar to other signaling peptides, including dynamic changes to their circulating levels in response to physiologic and environmental state. In vitro and in vivo functional assays establish signaling and bioactivity for three orphan capped peptides: CAP-TAC1 is a tachykinin neuropeptide-like molecule that exhibits nanomolar agonist activity at multiple mammalian tachykinin receptors, CAP-GDF15, derived from the prepropeptide region of the anorexigenic hormone GDF15, is itself a novel 12-mer hypophagic peptide, and CAP-WNT9A is a growth hormone secretagogue peptide. The studies reported herein demonstrate that N- and C-terminal capping is present in a much larger number of endogenous secreted peptides than previously anticipated. Capped peptides therefore constitute a class of secreted signaling peptides with the potential to broadly regulate cell-cell communication in mammalian physiology.

Capped peptides are provided. Aspects of the capped peptides include an N-terminal pyroglutamylation modification and a C-terminal amidation modification. In some instances, the capped peptides are not Thyrotropin-releasing hormone (TRH) or (Gonadotropin-releasing hormone (GnRH)/luteinizing hormone-releasing hormone (LHRH). Also provided are methods of administering capped peptides to a subject, e.g., to treat the subject for a condition, such as a disease or disorder condition, and pharmaceutical compositions that include the capped peptides.

BRIEF DESCRIPTION OF THE FIGURES

Fig. 1. Genomic prediction and mass spectrometry detection of capped peptides in mouse plasma. (A) Schematic representation of capped peptide production from secreted, full-length preproprecursor proteins. Abbreviations are defined as: “N-term” is N-terminal; “C-term” is C-terminal; “SP” is signal peptide; “Gln” is glutamine; “pGlu” is pyroglutamyl. (B) Number of predicted mouse capped peptides and (C) the targeted mass spectrometry strategy for their detection. (D) Chemical structure of CAP-TAC1 (pGlu-

FFGLM-NH₂). (E) Representative extracted ion chromatogram at $m/z = 724.3$ and (F) mirror fragmentation spectra of authentic CAP-TAC1 standard (pink) and the endogenous mouse plasma peak (blue). (G) Pie chart showing the number of detectable capped peptides in mouse plasma and (H) their quantitation. For (H), data are shown as means $\pm$ SEM (N=5 mice).

Fig. 2. Sequence and gene-level analysis of capped peptides. (A) Pie chart and (B) representative examples of the distribution of mouse capped peptides based on their full-length preproprotein precursors. "SP" is signal peptide. (C) H-clustered heat map of mRNA expression for capped peptide preproprecursor home genes across mouse tissues and cell types. Data was obtained from BioGPS and shown as Z-score of the log-transformed value. (D) Frequency of each amino acid in capped peptides versus a Uniprot reference set of known peptide hormones and neuropeptides. (E) Heat map of amino acid frequency within four residues upstream and downstream of the N-terminal pyroglutamylation or (F) C-terminal amidation.

Fig. 3. Dynamic regulation of circulating capped peptide levels by different environmental and physiologic perturbations. (A) Heat map of the relative fold change for the indicated capped peptide in the indicated condition. (B-F) Representative peptide quantification for the indicated capped peptide in the indicated comparison. For (A-F), "Fasting/Fed" indicates a comparison between plasma from 3-month old male mice on chow diet versus age- and sex-matched controls after fasting for 16 h (N=3/group); "HFD/Chow" indicates a comparison between plasma from 4-month old male mice on chow diet versus 4-month old male mice after being on high fat diet (HFD, 60% kcal from fat) for 9 weeks (N=3/group); "LPS" indicates a comparison between plasma from 2-month old male mice following treatment with vehicle versus LPS (0.5 mg/kg, IP, 5 h, N=5/group); "6AM/6PM" indicates a comparison between plasma from 2-month old male mice at the beginning of the night cycle (6AM) versus age- and sex-matched mice at the beginning of the day cycle (6PM) (N=5/group); "Old/Young" indicates a comparison between plasma from 4-month old male mice versus 23-month old male mice (N=2/group); "Run/Sed" indicates a comparison between plasma from sedentary 3-month old male mice versus age- and sex-matched controls after an acute exhaustive 60 min

run on a treadmill. For (B-F), data are shown as means $\pm$ SEM. * $P < 0.05$, ** $P < 0.01$ versus control by Student's t-test.

Fig. 4. CAP-TAC1 is a potent tachykinin receptor agonist. (A) Schematic and annotation of the primary amino acid sequence for full-length mouse TAC1 preproprotein and its cleavage products. (B) Chemiluminescent signal intensity in PathHunter beta-arrestin CHO-K1 cells transfected with human TACR1 following treatment with the indicated concentration of CAP-TAC1 or substance P. (C) Relative levels of CAP-TAC1 or substance P in mouse plasma following incubation at 37°C for the indicated time. For (B), N=2 per concentration; for (C), N=3 per condition. For (B,C), data are shown as means $\pm$ SEM. * $P < 0.05$, ** $P < 0.01$ versus control by Student's t-test.

Fig. 5. Functional studies of CAP-GDF15 in feeding and energy balance in mice. (A) Schematic of the primary amino acid sequence for full-length mouse GDF15 preproprotein and annotation of CAP-GDF15 sequence as well as the canonical GDF15 hormone sequence. (B) Representative extracted ion chromatogram and (C) mirror tandem fragmentation spectra of CAP-GDF15 (pGlu-LELRVAAGR-NH₂; +2 ion m/z = 682.4) and the endogenous plasma peak. (D) Food intake, (E) respiratory exchange ratio (RER), and (F) ambulatory activity of 12-16-week diet-induced obese male mice following a single treatment of CAP-GDF15 (50 mg/kg, intraperitoneal) or vehicle control. (G) Food intake, (H) RER, and (I) ambulatory activity of 12-16-week diet-induced obese male mice following a single treatment of scrambled CAP-GDF15 (50 mg/kg, intraperitoneal, scrambled sequence: pGlu-GLEALRRLRV-NH₂) or vehicle control. (J) 3-day food intake and (K) body weight of 12-16-week diet-induced obese male mice following treatment with CAP-GDF15 or scrambled CAP-GDF15 (50 mg/kg/day, IP), or vehicle control. Data are shown as means $\pm$ SEM. For (D-F), N=12/group; for (G-I), N=8/group; for (J,K), N=7/group. For (D-I), injection occurred at time T=0 (5:00pm) and data was collected for the following 16 hrs. For (D-K), ** $P < 0.01$, *** $P < 0.001$ by two-way ANOVA.

Fig. 6. Additional measurements from metabolic chambers of mice treated with CAP-GDF15. (A) VO₂ and (B) VCO₂ of 12-16-week diet-induced obese male mice following a single treatment of CAP-GDF15 (50 mg/kg, intraperitoneal) or vehicle control. (C) VO₂ and (D) VCO₂ of 12-16-week diet-induced obese male mice following a single treatment of scrambled CAP-GDF15 (50 mg/kg, intraperitoneal, scrambled sequence:

pQGLEALRRLRV-NH₂) or vehicle control. Data are shown as means $\pm$ SEM. For (A-B), N=12/group; for (C-D), N=8/group. For (A-D), injection occurred at time T=0 (5:00pm) and data was collected for the subsequent 16 hrs.

Fig. 7. Detection of human capped peptides and sequence alignment comparison to mouse. (A) Pie graph of detectable human capped peptides and diagram of the overlap between mouse-specific and human-specific capped peptides. (B) Number of human or mouse capped peptides detected in human or mouse plasma, stratified by those capped peptide with 100% conservation (top) or those that are not 100% conserved (human-specific, middle; mouse-specific, bottom). (C,D) Representative extracted ion chromatograms of the indicated capped peptides in plasma from the indicated species. (E) Phylogenetic alignment of all detectable capped peptides in human and mouse, with primary capped peptide sequences for select subclusters A-D shown on the right.

Fig. 8. Prediction, detection, and composition analysis of human capped peptides. (A) Schematic of numbers of predicted human capped peptides. (B) Quantification of detectable human capped peptide concentrations in human plasma (N=3). Data is shown as means $\pm$ SEM, N=3. (C) Comparison of frequency of each amino acid between human capped peptides and known peptide hormones.

Fig. 9. Tissue distribution of mRNAs for home genes corresponding to human capped peptides. H-clustered heat map of mRNA expression for capped peptide preproprecursor home genes across human tissues and cell types, using GTEx as the reference database.

Fig. 10. CAP-TAC1 induces body weight loss independent of food intake. (A) Body weight change and (B) food intake of diet induced obese mice injected with vehicle (18:1:1 Saline:DMSO:Kolliphor EL), 10 mg/kg CAP-TAC1 (pGlu-FFGLM-NH₂), and vehicle pair fed to CAP-TAC1 group. (N=6 for vehicle, N=7 for CAP-TAC1 and pair fed groups, error bars are SEM, and P values calculated with two-way ANOVA.)

Fig. 11. CAP-TAC1 acutely induces oxygen consumption over 24 hours. Oxygen consumption (VO₂) of diet induced obese mice injected with a single injection at T=0 vehicle (18:1:1 Saline:DMSO:Kolliphor EL) or 10 mg/kg CAP-TAC1 (pGlu-FFGLM-NH₂). (N=4 for vehicle and CAP-TAC1 groups, error bars are SEM, and P values calculated with two-way ANOVA.)

Fig. 12. Human CAP-GDF15 homologue inhibits food intake. Food intake of diet induced obese mice 2 hours post injection with vehicle or 50 mg/kg CAP-GDF15 human homologue (pGlu-LELHLRPQAAR-NH₂). (N=8 mice/group, error bars are SEM, and P values calculated with two-sided Student's T test.)

Fig. 13. Functional screen in GH3 cells identifies CAP-WNT9A as a growth hormone secretagogue in vitro. (A) Z-score of media growth hormone concentration with 1 hour treatment of 10 μ M of 100% conserved capped peptides growth hormone secretagogue activity in GH3 cells (N=1 for all peptides and N=3 for DMSO). (B) Media growth hormone concentration with 1 hour treatment of 10 μ M CAP-WNT9A or DMSO (N=3/group). (Error bars are SEM.)

Fig. 14. Mouse capped peptides. (A) Predicted mouse capped peptides and (B-E) detected mouse capped peptides. (B-C) shows peptide information, (D) shows ion count integrations, and (E) shows concentrations in nM for detected mouse capped peptides.

Fig. 15. Human capped peptides. (A) Predicted human capped peptides and (B-E) detected human capped peptides. (B-C) shows peptide information, (D) shows ion count integrations, and (E) shows concentrations in nM for detected human capped peptides.

The invention is best understood from the following detailed description when read in conjunction with the accompanying drawings.

DEFINITIONS

Before describing exemplary embodiments in greater detail, the following definitions are set forth to illustrate and define the meaning and scope of the terms used in the description.

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Hale & Markham, THE HARPER COLLINS DICTIONARY OF BIOLOGY, Harper Perennial, N.Y. (1991) provide one of skill with the general meaning of many of the terms used herein. Still, certain terms are defined below for the sake of clarity and ease of reference.

Certain ranges are presented herein with numerical values being preceded by the term "about." The term "about" is used herein to provide literal support for the exact number that it precedes, as well as a number that is near to or approximately the number that the term precedes. In determining whether a number is near to or approximately a specifically recited number, the near or approximating unrecited number may be a number which, in the context in which it is presented, provides the substantial equivalent of the specifically recited number.

It must be noted that as used herein and in the appended claims, the singular forms "a", "an", and "the" include plural referents unless the context clearly dictates otherwise. For example, the term "an agent" refers to one or more agents, i.e., a single agent and multiple agents. It is further noted that the claims can be drafted to exclude any optional element. As such, this statement is intended to serve as antecedent basis for use of such exclusive terminology as "solely," "only" and the like in connection with the recitation of claim elements, or use of a "negative" limitation.

A "pharmaceutical composition" refers to a formulation of a compound and a medium generally accepted in the art for the delivery of the biologically active compound to a mammal, e.g., humans. Such a medium can include a pharmaceutically acceptable delivery vehicle, carrier, diluent, or excipient.

While it may be possible for the compounds of the subject invention to be administered as the raw chemical, it is also possible to present them as a pharmaceutical composition. Accordingly, provided herein are pharmaceutical compositions which comprise one or more of certain compounds disclosed herein, or one or more pharmaceutically acceptable salts, esters, prodrugs, amides, or solvates thereof, together with one or more pharmaceutically acceptable carriers thereof and optionally one or more other therapeutic ingredients. Pharmaceutically acceptable delivery vehicles and other therapeutic ingredients may include one or more excipients and/or one or more vehicles. The carrier(s) must be "acceptable" in the sense of being compatible with the other ingredients of the formulation and not deleterious to the recipient thereof. Proper formulation is dependent upon the route of administration chosen. Any of the well-known techniques, carriers, and excipients may be used as suitable and as understood in the art; e.g., in Remington's Pharmaceutical Sciences. The pharmaceutical compositions

disclosed herein may be manufactured in any manner known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or compression processes.

5 The term “peptide” used herein, refers to a polymeric form of amino acids of any length, which can include coded and non-coded amino acids, chemically or biochemically modified or derivatized amino acids, and polypeptides having modified peptide backbones.

10 The terms “subject,” “individual,” and “patient” are used interchangeably herein to refer to a mammal being assessed for treatment and/or being treated. In some embodiments, the mammal is a human. The terms “subject,” “individual,” and “patient” encompass, without limitation, individuals having a disease. Subjects may be human, but also include other mammals, particularly those mammals useful as laboratory models for human disease, e.g., mice, rats, etc.

15 The term “sample” with reference to a patient encompasses blood and other liquid samples of biological origin, solid tissue samples such as a biopsy specimen or tissue cultures or cells derived therefrom and the progeny thereof. The term also encompasses samples that have been manipulated in any way after their procurement, such as by treatment with reagents; washed; or enrichment for certain cell populations, such as diseased cells. The definition also includes samples that have been enriched for particular
20 types of molecules, e.g., nucleic acids, polypeptides, etc. The term “biological sample” encompasses a clinical sample, and also includes tissue obtained by surgical resection, tissue obtained by biopsy, cells in culture, cell supernatants, cell lysates, tissue samples, organs, bone marrow, blood, plasma, serum, and the like. A “biological sample” includes a sample obtained from a patient’s diseased cell, e.g., a sample comprising
25 polynucleotides and/or polypeptides that is obtained from a patient’s diseased cell (e.g., a cell lysate or other cell extract comprising polynucleotides and/or polypeptides); and a sample comprising diseased cells from a patient. A biological sample comprising a diseased cell from a patient can also include non-diseased cells.

30 The term “diagnosis” is used herein to refer to the identification of a molecular or pathological state, disease or condition in a subject, individual, or patient.

The term "prognosis" is used herein to refer to the prediction of the likelihood of death or disease progression, including recurrence, spread, and drug resistance, in a subject, individual, or patient. The term "prediction" is used herein to refer to the act of foretelling or estimating, based on observation, experience, or scientific reasoning, the likelihood of a subject, individual, or patient experiencing a particular event or clinical outcome. In one example, a physician may attempt to predict the likelihood that a patient will survive.

As used herein, the terms "treatment," "treating," and the like, refer to administering an agent, or carrying out a procedure, for the purposes of obtaining an effect on or in a subject, individual, or patient. The effect may be prophylactic in terms of completely or partially preventing a disease or symptom thereof and/or may be therapeutic in terms of effecting a partial or complete cure for a disease and/or symptoms of the disease. "Treatment," as used herein, may include treatment of fatty liver disease in a mammal, particularly in a human, and includes: (a) inhibiting the disease, i.e., arresting its development; and (b) relieving the disease or its symptoms, i.e., causing regression of the disease or its symptoms.

Treating may refer to any indicia of success in the treatment or amelioration or prevention of a disease, including any objective or subjective parameter such as abatement; remission; diminishing of symptoms or making the disease condition more tolerable to the patient; slowing in the rate of degeneration or decline; or making the final point of degeneration less debilitating. The treatment or amelioration of symptoms can be based on objective or subjective parameters; including the results of an examination by a physician. Accordingly, the term "treating" includes the administration of engineered cells to prevent or delay, to alleviate, or to arrest or inhibit development of the symptoms or conditions associated with disease or other diseases. The term "therapeutic effect" refers to the reduction, elimination, or prevention of the disease, symptoms of the disease, or side effects of the disease in the subject.

As used herein, a "therapeutically effective amount" refers to that amount of the therapeutic agent sufficient to treat or manage a disease or disorder. A therapeutically effective amount may refer to the amount of therapeutic agent sufficient to delay or minimize the onset of disease. A therapeutically effective amount may also refer to the

amount of the therapeutic agent that provides a therapeutic benefit in the treatment or management of a disease. Further, a therapeutically effective amount with respect to a therapeutic agent of the invention means the amount of therapeutic agent alone, or in combination with other therapies, that provides a therapeutic benefit in the treatment or management of a disease.

As used herein, the term "dosing regimen" refers to a set of unit doses (typically more than one) that are administered individually to a subject, typically separated by periods of time. In some embodiments, a given therapeutic agent has a recommended dosing regimen, which may involve one or more doses. In some embodiments, a dosing regimen comprises a plurality of doses each of which are separated from one another by a time period of the same length; in some embodiments, a dosing regimen comprises a plurality of doses and at least two different time periods separating individual doses. In some embodiments, all doses within a dosing regimen are of the same unit dose amount. In some embodiments, different doses within a dosing regimen are of different amounts. In some embodiments, a dosing regimen comprises a first dose in a first dose amount, followed by one or more additional doses in a second dose amount different from the first dose amount. In some embodiments, a dosing regimen comprises a first dose in a first dose amount, followed by one or more additional doses in a second dose amount same as the first dose amount. In some embodiments, a dosing regimen is correlated with a desired or beneficial outcome when administered across a relevant population (i.e., is a therapeutic dosing regimen).

"In combination with", "combination therapy" and "combination products" refer, in certain embodiments, to the concurrent administration to a patient of the engineered proteins and cells described herein in combination with additional therapies, e.g., surgery, radiation, chemotherapy, and the like. When administered in combination, each component can be administered at the same time or sequentially in any order at different points in time. Thus, each component can be administered separately but sufficiently closely in time so as to provide the desired therapeutic effect.

"Concomitant administration" means administration of one or more components, such as engineered proteins and cells, known therapeutic agents, etc. at such time that the combination will have a therapeutic effect. Such concomitant administration may

involve concurrent (*i.e.*, at the same time), prior, or subsequent administration of components. A person of ordinary skill in the art would have no difficulty determining the appropriate timing, sequence and dosages of administration.

5 The term "sustained-release", as in a sustained-release form, sustained-release composition or sustained-release formulation, is intended to include a form of an active ingredient, or formulation for an active ingredient, which has an extended in vivo half-life or duration of action. A sustained-release form may result from modification of the active ingredient, such as modifications that extend circulation residence time, decrease rates of degradation, decrease rates of clearance or the like, or may result from formulations or
10 compositions which provide for extended release of the active ingredient, such as use of various liposomes, emulsions, micelles, matrices and the like. A controlled-release form or formulation is a type of sustained-release form or formulation.

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

Where a range of values is provided, it is understood that each intervening value,
20 to the tenth of the unit of the lower limit unless the context clearly dictates otherwise, between the upper and lower limit of that range and any other stated or intervening value in that stated range, is encompassed within the invention. The upper and lower limits of these smaller ranges may independently be included in the smaller ranges and are also encompassed within the invention, subject to any specifically excluded limit in the stated
25 range. Where the stated range includes one or both of the limits, ranges excluding either or both of those included limits are also included in the invention.

Certain ranges are presented herein with numerical values being preceded by the term "about." The term "about" is used herein to provide literal support for the exact number that it precedes, as well as a number that is near to or approximately the number
30 that the term precedes. In determining whether a number is near to or approximately a specifically recited number, the near or approximating unrecited number may be a

number which, in the context in which it is presented, provides the substantial equivalent of the specifically recited number.

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

All publications and patents cited in this specification are herein incorporated by reference as if each individual publication or patent were specifically and individually indicated to be incorporated by reference and are incorporated herein by reference to disclose and describe the methods and/or materials in connection with which the publications are cited. The citation of any publication is for its disclosure prior to the filing date and should not be construed as an admission that the present invention is not entitled to antedate such publication by virtue of prior invention. Further, the dates of publication provided may be different from the actual publication dates which may need to be independently confirmed.

It is noted that, as used herein and in the appended claims, the singular forms “a”, “an”, and “the” include plural referents unless the context clearly dictates otherwise. It is further noted that the claims may be drafted to exclude any optional element. As such, this statement is intended to serve as antecedent basis for use of such exclusive terminology as “solely,” “only” and the like in connection with the recitation of claim elements, or use of a “negative” limitation.

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

While the apparatus and method has or will be described for the sake of grammatical fluidity with functional explanations, it is to be expressly understood that the claims, unless expressly formulated under 35 U.S.C. §112, are not to be construed as

necessarily limited in any way by the construction of "means" or "steps" limitations, but are to be accorded the full scope of the meaning and equivalents of the definition provided by the claims under the judicial doctrine of equivalents, and in the case where the claims are expressly formulated under 35 U.S.C. §112 are to be accorded full statutory equivalents under 35 U.S.C. §112.

DETAILED DESCRIPTION

Capped peptides are provided. Aspects of the capped peptides include an N-terminal pyroglutamylation modification and a C-terminal amidation modification. In some instances, the capped peptides are not Thyrotropin-releasing hormone (TRH) or (Gonadotropin-releasing hormone (GnRH)/luteinizing hormone-releasing hormone (LHRH). In some instances, the capped peptides exhibit physiological activity, including, but not limited to, mammalian tachykinin receptor agonist activity, activating metabolic activity, hypophagic activity and growth hormone secretagogue activity. Also provided are methods of administering capped peptides to a subject in need thereof, e.g., to treat the subject for a condition, such as a disorder or disease condition, as well as pharmaceutical compositions that include capped peptides.

CAPPED PEPTIDES

As reviewed above, capped peptides are provided. Capped peptides of embodiments of the invention include an N-terminal pyroglutamylation modification and a C-terminal amidation modification, wherein the capped peptide is not Thyrotropin-releasing hormone (TRH) or (Gonadotropin-releasing hormone (GnRH)/luteinizing hormone-releasing hormone (LHRH).

Capped peptides are modified fragments of proteins. By "capped peptide" it is meant that the N-terminus and C-terminus of the peptide are modified. These modifications include chemical modifications. Types of modifications include N-terminal pyroglutamylation modification and C-terminal amidation modification. In an N-terminal pyroglutamylation modification, an N-terminal glutamine is modified to an N-terminal

pyroglutamate. The pyroglutamylation reaction may be carried out enzymatically or chemically. In a C-terminal amidation modification, a C-terminal amino acid is modified to a C-terminal amide. The amidation reaction may be carried out enzymatically or chemically. In some cases, the amidation reaction may be carried out by the enzyme
5 peptidyl-glycine alpha-amidating monooxygenase (PAM).

Capped peptides may be naturally occurring peptides or synthetic peptides. Peptides may be capped via post-translational modification or modified synthetically. Capped peptides may be produced from secreted proteins or synthesized proteins. Secreted proteins may be preproprecursor proteins. By “preproprecursor protein” it is
10 meant that the protein is a protein precursor tagged with an N-terminal signal peptide that targets the protein for secretion. To produce a capped peptide from a secreted protein, the secreted protein is proteolyzed to produce a peptide, and then the peptide is post-translationally modified to produce a capped peptide. For example, an N-terminal glutamine amino acid can be post-translationally modified to an N-terminal pyroglutamyl
15 amino acid and a C-terminal amino acid can be post translationally modified to a C-terminal amide (Fig. 1).

Capped peptides may be chemically synthesized and/or modified. For example, methods of chemically synthesizing capped peptides may include, but are not limited to, solid phase peptide synthesis and/or chemoselective ligation.

20 In some embodiments, the capped peptides may vary in size, ranging in some instances from 2 to 30 amino acid residues, such as 3 to 25 amino acid residues, and including 3 to 20 amino acid residues. Capped peptides may include coded and non-coded amino acids. Capped peptides may include, but are not limited to, chemically or biochemically modified amino acids, derivatized amino acids, and peptides having
25 modified peptide backbones.

In some embodiments, the capped peptide may be from a variety of different species, including mammalian and non-mammalian species. In some embodiments, the capped peptide is a murine peptide. Examples of murine capped peptides of the invention include, but are not limited to, those listed in Fig. 14. In some embodiments, the capped
30 peptide is a peptide listed in Fig. 14. In some embodiments, the capped peptide is a peptide listed in Fig. 14A. In some embodiments, the murine capped peptides may vary

in size, ranging from 2 to 30 amino acid residues, such as 3 to 25 amino acid residues, and including 3 to 20 amino acid residues.

In some embodiments, the murine capped peptide may be modified with an N-terminal pyroglutamylation and C-terminal amidation from a sequence as shown in Fig. 14, or a sequence that is substantially the same as the sequence as shown in Fig. 14. By substantially the same as is meant a peptide having a region with a sequence that is 60% or greater, such as 75% or greater, such as 80% or greater, such as 85% or greater, such as 90% or greater, such as 95% or greater and including 98% or greater sequence identity with a sequence as shown in Fig. 14, e.g., as determined by BLAST using default settings.

In some embodiments, the murine capped peptide may be modified with an N-terminal pyroglutamylation and C-terminal amidation from a sequence as shown in Fig. 14A, or a sequence that has a sequence that is substantially the same as the sequence as shown in Fig. 14A. By substantially the same as is meant a peptide having a region with a sequence that is 60% or greater, such as 75% or greater, such as 80% or greater, such as 85% or greater, such as 90% or greater, such as 95% or greater and including 98% or greater sequence identity with a sequence as shown in Fig. 14A, e.g., as determined by BLAST using default settings.

In some embodiments, the murine capped peptide comprises a lysine residue. In some embodiments, the murine capped peptide comprises a lysine residue four or fewer residues away, including four residues away, three residues away, two residues away, and one residue away from the N-terminal pyroglutamylation modification. In some embodiments, the murine capped peptide comprises a glycine residue. In some embodiments, the murine capped peptide comprises a glycine residue four or fewer residues away, including four residues away, three residues away, two residues away, and one residue away from the N-terminal pyroglutamylation modification.

In some embodiments, the capped peptide is a human peptide. Examples of human capped peptides of the invention include, but are not limited to, those listed in Fig. 15. In some embodiments, the capped peptide is a peptide listed in Fig. 15. In some embodiments, the capped peptide is a peptide listed in Fig. 15A. In some embodiments, the human capped peptides may vary in size, ranging from 2 to 30 amino acid residues, such as 3 to 25 amino acid residues, and including 3 to 20 amino acid residues.

In some embodiments, the human capped peptide may be modified with an N-terminal pyroglutamylation and C-terminal amidation from a sequence as shown in Fig. 15, or a sequence that has a sequence that is substantially the same as the sequence as shown in Fig. 15. By substantially the same as is meant a peptide having a region with a sequence that is 60% or greater, such as 75% or greater, such as 80% or greater, such as 85% or greater, such as 90% or greater, such as 95% or greater and including 98% or greater sequence identity with a sequence as shown in Fig. 15, e.g., as determined by BLAST using default settings.

In some embodiments, the human capped peptide may be modified with an N-terminal pyroglutamylation and C-terminal amidation from a sequence as shown in Fig. 15A, or a sequence that has a sequence that is substantially the same as the sequence as shown in Fig. 15A. By substantially the same as is meant a peptide having a region with a sequence that is 60% or greater, such as 75% or greater, such as 80% or greater, such as 85% or greater, such as 90% or greater, such as 95% or greater and including 98% or greater sequence identity with a sequence as shown in Fig. 15A, e.g., as determined by BLAST using default settings.

In some embodiments, the human capped peptide comprises a lysine residue. In some embodiments, the human capped peptide comprises a lysine residue four or fewer residues away, including four residues away, three residues away, two residues away, and one residue away from the N-terminal pyroglutamylation modification. In some embodiments, the human capped peptide comprises a proline residue. In some embodiments, the murine capped peptide comprises a proline residue four or fewer residues away, including four residues away, three residues away, two residues away, and one residue away from the N-terminal pyroglutamylation modification. In some embodiments, the human capped peptide comprises a valine residue. In some embodiments, the human capped peptide comprises a valine residue four or fewer residues away, including four residues away, three residues away, two residues away, and one residue away from the N-terminal pyroglutamylation modification.

In some embodiments, the capped peptide of the present invention exhibits physiological activity. By physiological activity it is meant that the capped peptides exhibit a measurable biological response and/or change. Biological responses include, but are

not limited to, molecular responses, cellular responses, tissue-specific responses, organ-specific responses, organism-specific responses and combinations thereof. Measurable biological responses are able to be detected and/or quantified as a change from baseline activity.

5 Physiological activity may include mammalian tachykinin receptor agonist activity. Tachykinins are a family of peptides that perturb many biological processes via interaction with tachykinin receptors. There are three known mammalian tachykinin receptors: NK₁, NK₂, and NK₃. Tachykinins have been reported to participate in various physiological processes including, but not limited to, processes in the nervous, immune,
10 gastrointestinal, respiratory, urogenital and dermal systems (Steinhoff, M. S., *Physiol. Rev.* 2014 94:265-301). Agonists of tachykinins can bind to tachykinin receptors and thereby elicit and/or modify tachykinin-related responses. In certain embodiments, the capped peptide exhibits mammalian tachykinin receptor agonist activity. In certain
15 embodiments, the capped peptide exhibits mammalian tachykinin receptor NK₁ agonist activity. In certain embodiments, the capped peptide exhibits mammalian tachykinin receptor NK₂ agonist activity. In certain embodiments, the capped peptide exhibits mammalian tachykinin receptor NK₃ agonist activity. In certain embodiments, the capped peptide exhibits murine tachykinin receptor agonist activity. In certain embodiments, the capped peptide exhibits human tachykinin receptor agonist activity.

20 Physiological activity may include activating metabolic activity, wherein activation of metabolic activity means there is an increase in metabolic activity. Activating metabolic activity includes all of the processes that are necessary for breaking down compounds (i.e., catabolic processes) to cause an increase in energy consumption. The increase in energy consumption may occur in a cell, organ, organism, or combinations thereof.
25 Increases in metabolic activity (i.e., activating metabolic activity) may result in increased caloric consumption. Activating metabolic activity may also be characterized as, for example, increased food intake that is not accompanied by an increase in body weight, decreased body weight that is not accompanied by an increase in food intake, increased in oxygen consumption that is not accompanied by an increase in food intake, and/or
30 combinations thereof. In some embodiments, the capped peptide exhibits activating metabolic activity. In some embodiments, the capped peptide exhibits activating

metabolic activity in humans. In some embodiments, the capped peptide exhibits activating metabolic activity in mice.

Physiological activity may include hypophagic activity. Hypophagia is defined as the reduced ingestion of food. Hypophagic activity is characterized by decreased ingestion of food, decreased consumption of food, loss of appetite, not feeling hungry, and combinations thereof. In some embodiments, the capped peptide exhibits hypophagic activity. In some embodiments, the capped peptide exhibits human hypophagic activity. In some embodiments, the capped peptide exhibits murine hypophagic activity.

Physiological activity may include growth hormone secretagogue activity. A secretagogue is an agent that promotes the secretion of a molecule. As such, a growth hormone secretagogue is an agent that promotes the secretion of growth hormone. Growth hormone is a peptide hormone that stimulates growth, cell reproduction, and cell regeneration in mammals, including humans. In some embodiments, the capped peptide exhibits growth hormone secretagogue activity. In some embodiments, the capped peptide exhibits human growth hormone secretagogue activity. In some embodiments, the capped peptide exhibits murine growth hormone secretagogue activity.

In certain embodiments, the capped peptide is CAP-TAC1 comprising the sequence pGlu-FFGLM-NH₂ or a peptide having a region with a sequence that is 60% or greater, such as 75% or greater, such as 80% or greater, such as 85% or greater, such as 90% or greater, such as 95% or greater and including 98% or greater sequence identity with CAP-TAC1. In certain embodiments, CAP-TAC1 or a peptide having a region with a sequence that is 60% or greater, such as 75% or greater, such as 80% or greater, such as 85% or greater, such as 90% or greater, such as 95% or greater and including 98% or greater sequence identity with CAP-TAC1 exhibits tachykinin receptor agonist activity. In certain embodiments, CAP-TAC1 or a peptide having a region with a sequence that is 60% or greater, such as 75% or greater, such as 80% or greater, such as 85% or greater, such as 90% or greater, such as 95% or greater and including 98% or greater sequence identity with CAP-TAC1 exhibits activating metabolic activity. In certain embodiments, CAP-TAC1 or a peptide having a region with a sequence that is 60% or greater, such as 75% or greater, such as 80% or greater, such as 85% or greater, such as 90% or greater,

such as 95% or greater and including 98% or greater sequence identity with CAP-TAC1 exhibits tachykinin receptor agonist activity and activating metabolic activity.

In certain embodiments, the capped peptide is CAP-GDF15 comprising the sequence pGlu-LELRLRVAAGR-NH₂ or a peptide having a region with a sequence that is 60% or greater, such as 75% or greater, such as 80% or greater, such as 85% or greater, such as 90% or greater, such as 95% or greater and including 98% or greater sequence identity with CAP-GDF15. In certain embodiments, CAP-GDF15 or a peptide having a region with a sequence that is 60% or greater, such as 75% or greater, such as 80% or greater, such as 85% or greater, such as 90% or greater, such as 95% or greater and including 98% or greater sequence identity with CAP-GDF15 exhibits hypophagic activity.

In certain embodiments, the capped peptide is CAP-WNT9A comprising the sequence pGlu-WGGCGDNLKYSSKFVEFL-NH₂ or a peptide having a region with a sequence that is 60% or greater, such as 75% or greater, such as 80% or greater, such as 85% or greater, such as 90% or greater, such as 95% or greater and including 98% or greater sequence identity with CAP-WNT9A. In certain embodiments, CAP-WNT9A or a peptide having a region with a sequence that is 60% or greater, such as 75% or greater, such as 80% or greater, such as 85% or greater, such as 90% or greater, such as 95% or greater and including 98% or greater sequence identity with CAP-WNT9A exhibits growth hormone secretagogue activity.

METHODS OF USE

The present disclosure includes methods of using capped peptides. Aspects of such methods include administering a capped peptide to a subject in need thereof, wherein the capped peptide comprises an N-terminal pyroglutamylation modification and a C-terminal amidation modification, and is not Thyrotropin-releasing hormone (TRH) or (Gonadotropin-releasing hormone (GnRH)/luteinizing hormone-releasing hormone (LHRH).

As reviewed above, in some embodiments, capped peptides of the present invention exhibit physiological activity. Physiological activity includes, but is not limited to, mammalian tachykinin receptor agonist activity, activating metabolic activity, hypophagic

activity and growth hormone secretagogue activity. Such physiological activities may be beneficial in the treatment of various disorders, diseases and conditions. Aspects of the disclosure include methods of using capped peptides to treat a subject in need thereof. In some embodiments, the method is a method of treating a subject for a disorder, disease or condition.

Methods of treating neurological disorders

In some embodiments, the method is a method of treating a subject for a neurological disorder. Neurological disorders include disorders of the central nervous system and peripheral nervous system. In other words, neurological disorders include disorders of the brain, spinal cord, cranial nerves, peripheral nerves, nerve roots, autonomic nervous system, neuromuscular junction and muscles. Disorders of the brain include brain dysfunction such as apraxia (i.e., brain dysfunction related to movement), agnosia (i.e., brain dysfunction related to processing sensory information), amnesia (i.e. brain dysfunction related to memory), aphasia (i.e. brain dysfunction related to language) and dysarthria (i.e. brain dysfunction related to speech). Neurological disorders may be pain-related neurological disorders such as migraines, headaches, back pain, complex regional pain syndrome, fibromyalgia and chronic pain. Neurological disorders may be psychiatric disorders (i.e., mental disorders). Psychiatric disorders include anxiety and anxiety disorders, depression and depressive disorders, mood disorders, addiction and addictive disorders, and psychosis and psychotic disorders. Psychotic disorders include Schizophrenia. Other examples of neurological disorders include, but are not limited to, dysautonomia, multiple system atrophy, epilepsy, Parkinson's disease, Tourette's syndrome, amyotrophic lateral sclerosis, multiple sclerosis, pruritus, Alzheimer's disease and stroke.

As reviewed above, tachykinins and tachykinin-receptors participate in processes within the nervous system. As such, antagonism of tachykinin receptors is a method to control neurological disorders or symptoms thereof. In some cases, neurological disorders or symptoms of neurological disorders may be improved upon treatment with a capped peptide that exhibits mammalian tachykinin receptor agonist activity. In some embodiments, the method is a method of using capped peptides, wherein the capped

peptide exhibits mammalian tachykinin receptor agonist activity, to treat a subject for a neurological disorder. In certain embodiments, the neurological disorder may include, but is not limited to, a depressive disorder, anxiety disorder, chronic pain, addiction, pruritus, psychosis or schizophrenia.

5 Methods of treating inflammatory disorders

In some embodiments, the method is a method of treating a subject for an inflammatory disorder. Inflammation is a biological response to harmful stimuli such as pathogens, damaged cells, or irritants. As such, a disorder of inflammation (i.e. an inflammatory disorder) is one wherein inflammation is dysregulated within the organism.

10 Inflammation can present as heat, pain, redness, swelling and/or loss of function. Inflammation leads to a shift in the type of cells and/or signaling molecules present at the inflammation site. For example, inflammation may result in an increase of leukocytes and cytokines at the inflammation site. Inflammatory disorders include disorders wherein too little inflammation occurs and disorders wherein too much inflammation occurs.

15 Inflammatory disorders include acute inflammatory disorders (i.e., the initial response to harmful stimuli) and chronic inflammatory disorders (i.e. prolonged inflammation). Acute inflammation may last for hours to days. Acute inflammatory diseases include, but are not limited to, ileus and appendicitis. Chronic inflammation may last for one month to one year, or even longer than a year. Chronic inflammatory disorders include, but are not

20 limited to, irritable bowel syndrome, inflammatory bowel disease, Crohn's disease, asthma, colitis and fibrosis.

Tachykinins have been shown to be potent mediators of vasodilation, plasma extravasation, inflammatory cell recruitment and pain (Canning, B. J., J. Allergy Clin. Immunol. 1997. 99:579-582). As such, antagonism of tachykinin receptors is a method to

25 control inflammatory responses, as well as inflammatory disorders and symptoms thereof. In some cases, inflammatory disorders or symptoms of inflammatory disorders may be improved upon treatment with a capped peptide that exhibits mammalian tachykinin receptor agonist activity. In some embodiments, the method is a method of using capped peptides, wherein the capped peptide exhibits mammalian tachykinin receptor agonist

30 activity, to treat a subject for an inflammatory disorder. In certain embodiments, the

inflammatory disorder may include, but is not limited to, irritable bowel syndrome, inflammatory bowel disease, asthma, colitis, ileus or fibrosis.

Methods of treating nausea and/or vomiting

In some embodiments, the method is a method of treating a subject for nausea and/or vomiting. Nausea is a sensation that often includes the urge to vomit. Nausea and vomiting are symptoms resulting from a myriad of causes. Causes of nausea and vomiting include, but are not limited to, gastrointestinal disorders, food poisoning, motion sickness, dizziness, psychiatric disorders, low blood sugar, dehydration, lack of sleep, chemotherapy, anesthesia, and pregnancy. Nausea and vomiting are common side effects of chemotherapy treatment. Chemotherapy-induced nausea and vomiting includes acute (i.e., within the first 24 hours after chemotherapy), delayed (i.e., after 24 hours of chemotherapy), anticipatory (i.e., before chemotherapy dose administration but after prior cycles of chemotherapy), breakthrough (i.e. after nausea and vomiting treatment has already been administered), and chronic nausea and vomiting. Post-operative nausea and vomiting is a common side effect resulting from administration of anesthesia. Post-operative nausea and vomiting can result from administration of a variety of anesthetic agents including, but not limited to, nitrous oxide, ether, cyclopropaneetomidate, ketamine, propofol, opioids, and combinations thereof.

Tachykinins play a role in regulating nausea and vomiting. Specifically, tachykinin receptor NK₁ has been implicated in controlling nausea and vomiting. Antagonism of tachykinin receptors may treat nausea and vomiting. In some embodiments, the method is a method of using capped peptides, wherein the capped peptide exhibits mammalian tachykinin receptor agonist activity, to treat a subject for nausea and vomiting. In some embodiments, the method is a method of using capped peptides, wherein the capped peptide exhibits mammalian tachykinin receptor NK₁ agonist activity, to treat a subject for nausea and vomiting. In certain embodiments, the method is used to treat chemotherapy-induced nausea and vomiting. In certain embodiments, the method is used to treat postoperative nausea and vomiting.

Methods of treating obesity

In some embodiments, the method is a method of treating a subject for obesity or an obesity-related disorder. Obesity is defined as an excess of body fat relative to lean body mass. A subject is generally defined as obese if the subject has a body mass index of 30 kg/m² or greater. Obesity is often caused by excessive food intake coupled with limited energy expenditure and/or lack of physical exercise. Obesity increases the likelihood of various disorders. Obesity-related disorders may include, but are not limited to, hypertension, dyslipidemia, mellitus, atherosclerosis, gout, rheumatism, arthritis, type 2 diabetes, coronary heart disease, stroke, gallbladder disease, liver disease, sleep apnea and pain.

As reviewed above, activating metabolic activity increases energy consumption, resulting in a decrease in body weight. In some embodiments, the method is a method of using capped peptides, wherein the capped peptide exhibits activating metabolic activity, to treat a subject for obesity or an obesity-related disorder.

Hypophagic activity reduces ingestion of food, resulting in a decrease in body weight. In some embodiments, the method is a method of using capped peptides, wherein the capped peptide exhibits hypophagic activity, to treat a subject for obesity or an obesity-related disorder. In certain embodiments, the subject has a body mass index greater than 30 kg/m².

Efficacy of treatment for obesity can be readily determined by weight loss, for example the reduced food intake observed with administration is associated with weight loss, e.g., loss of 1% body weight, 2.5%, 5%, 7.5%, 10%, 12.5%, 15%, 17.5%, 20%, 22.5%, 25%, 27.5%, 30% or more, depending on the initial weight of the subject. Methods for measuring weight loss can be determined by weighing the subject.

Methods of treating growth hormone deficiency and disorders thereof

In some embodiments, the method is a method of treating a subject for a growth hormone deficiency. Growth hormone is produced by the pituitary gland and stimulates growth, cell reproduction and cell regeneration throughout the body. As such, growth hormone is required for human development. Deficiency of growth hormone in adolescents results in slow and limited growth. In general, growth hormone has many

effects on the body including, but not limited to, increasing calcium retention, increasing muscle mass, promoting lipolysis, increasing protein synthesis, stimulating growth of organs, stimulating growth of the brain, regulating homeostasis, reducing liver uptake of glucose, stimulating the immune system, and inducing insulin resistance. Lack of growth hormone or growth hormone deficiency results in a variety of disorders including, but not limited to, depression, anxiety, obesity, cardiac dysfunction, reduced energy, low libido, impaired concentration, memory loss, loss of hair growth, decreased muscle mass, insulin resistance and loss of thermoregulation. Growth hormone deficiency related-disorders may include, but are not limited to dwarfism, pituitary dwarfism, osteoporosis and dyslipidemia.

As reviewed above, growth hormone secretagogue activity promotes the secretion of growth hormone. Increased levels of growth hormone can have beneficial effects in treating growth hormone deficiency or disorders of growth hormone deficiency. In some embodiments, the method is a method of using capped peptides, wherein the capped peptide exhibits growth hormone secretagogue activity, to treat a subject for growth hormone deficiency or a growth hormone deficiency-disorder.

Efficacy of treatment for growth hormone deficiency or growth hormone deficiency-related disorders can be readily determined by measuring the change in the levels of growth hormone in a subject, e.g., a 1%, 2.5%, 5%, 7.5%, 10%, 12.5%, 15%, 17.5%, 20%, 22.5%, 25%, 27.5%, 30% or more, increase in growth hormone levels. Changes in the levels of growth hormone in a subject can be measured in blood and/or urine samples. Growth hormone levels can be measured using a variety of different methods including, but not limited to, immunoassays such as ELISA, gas chromatography mass spectrometry (GCMS) and liquid chromatography mass spectrometry (LC-MS/MS).

Methods of administration

Methods of administration may be carried out by any suitable means, including topical, oral, parenteral, intrapulmonary, and intranasal. Parenteral infusions include intramuscular, intravenous (bolus or slow drip), intraarterial, intraperitoneal, intrathecal or subcutaneous administration. An agent can be administered in any manner which is medically acceptable. This may include injections, by parenteral routes such as

intravenous, intravascular, intraarterial, subcutaneous, intramuscular, intratumor, intraperitoneal, intraventricular, intraepidural, or others as well as oral, nasal, ophthalmic, rectal, or topical. Sustained release administration is also included in the disclosure, by such means as depot injections or erodible implants.

5 Dosage and frequency of dosing may vary depending on the half-life of the agent (e.g., a capped peptide) in the patient. It will be understood by one of skill in the art that such guidelines will be adjusted for the molecular weight of the active agent, the clearance from the blood, the mode of administration, and other pharmacokinetic parameters. The dosage may also be varied for localized administration, e.g., intranasal, inhalation, etc.,
10 or for systemic administration, e.g., i.m., i.p., i.v., oral, and the like.

PHARMACEUTICAL COMPOSITIONS

Aspects of the present disclosure include pharmaceutical compositions comprising a capped peptide and a pharmaceutically acceptable delivery vehicle. Pharmaceutically
15 acceptable delivery vehicles may include one or more vehicles and/or carriers. The vehicle(s) or carrier(s) must be "acceptable" in the sense of being compatible with the other ingredients of the formulation and not deleterious to the recipient thereof. A pharmaceutically acceptable delivery vehicle is selected on the basis of the chosen route of administration and standard pharmaceutical practice. Proper formulation is dependent
20 upon the route of administration chosen. In instances where more than one vehicle and/or carrier is used, the vehicles and/or carriers must be "pharmaceutically acceptable" in the sense of being compatible with the other ingredients of the formulation and not injurious to the subject. Any of the well-known techniques, carriers, and excipients may be used as suitable and as understood in the art; e.g., in Remington's Pharmaceutical Sciences.
25 The pharmaceutical compositions disclosed herein may be manufactured in any manner known in the art.

As noted above, an agent, such as a capped peptide, can be formulated with an a pharmaceutically acceptable delivery vehicle (one or more organic or inorganic ingredients, natural or synthetic, with which a subject agent is combined to facilitate its
30 application). A suitable delivery vehicle includes sterile saline although other aqueous

and non-aqueous isotonic sterile solutions and sterile suspensions known to be pharmaceutically acceptable are known to those of ordinary skill in the art. An "effective amount" refers to that amount which is capable of ameliorating or delaying progression of the diseased, degenerative or damaged condition. An effective amount can be determined on an individual basis and will be based, in part, on consideration of the symptoms to be treated and results sought. An effective amount can be determined by one of ordinary skill in the art employing such factors and using no more than routine experimentation.

An agent, such as a capped peptide, can be administered as a pharmaceutical composition comprising a pharmaceutically acceptable excipient. The preferred form depends on the intended mode of administration and therapeutic application. The compositions can also include, depending on the formulation desired, pharmaceutically-acceptable, non-toxic carriers or diluents, which are defined as vehicles commonly used to formulate pharmaceutical compositions for animal or human administration. The diluent is selected so as not to affect the biological activity of the combination. Examples of such diluents are distilled water, physiological phosphate-buffered saline, Ringer's solutions, dextrose solution, and Hank's solution. In addition, the pharmaceutical composition or formulation may also include other carriers, adjuvants, or nontoxic, nontherapeutic, nonimmunogenic stabilizers and the like.

Compounds useful for co-administration with the active agents, such as capped peptides, of the invention can be made by methods known to one of ordinary skill in the art. As used herein, "methods known to one of ordinary skill in the art" may be identified through various reference books and databases. Suitable reference books and treatises that detail the synthesis of reactants useful in the preparation of compounds of the present invention, or provide references to articles that describe the preparation, include for example, "Synthetic Organic Chemistry", John Wiley & Sons, Inc., New York; S. R. Sandler et al., "Organic Functional Group Preparations," 2nd Ed., Academic Press, New York, 1983; H. O. House, "Modern Synthetic Reactions", 2nd Ed., W. A. Benjamin, Inc. Menlo Park, Calif. 1972; T. L. Gilchrist, "Heterocyclic Chemistry", 2nd Ed., John Wiley & Sons, New York, 1992; J. March, "Advanced Organic Chemistry: Reactions, Mechanisms and Structure", 4th Ed., Wiley-Interscience, New York, 1992. Specific and analogous

reactants may also be identified through the indices of known chemicals prepared by the Chemical Abstract Service of the American Chemical Society, which are available in most public and university libraries, as well as through on-line databases (the American Chemical Society, Washington, D.C., may be contacted for more details). Chemicals that are known but not commercially available in catalogs may be prepared by custom chemical synthesis houses, where many of the standard chemical supply houses (e.g., those listed above) provide custom synthesis services.

Compounds may also be purchased. Such "commercially available" compounds may be obtained from commercial sources including but not limited to Acros Organics (Pittsburgh PA), Aldrich Chemical (Milwaukee WI, including Sigma Chemical and Fluka), Apin Chemicals Ltd. (Milton Park UK), Avocado Research (Lancashire U.K.), BDH Inc. (Toronto, Canada), Bionet (Cornwall, U.K.), Chemservice Inc. (West Chester PA), Crescent Chemical Co. (Hauppauge NY), Eastman Organic Chemicals, Eastman Kodak Company (Rochester NY), Fisher Scientific Co. (Pittsburgh PA), Fisons Chemicals (Leicestershire UK), Frontier Scientific (Logan UT), ICN Biomedicals, Inc. (Costa Mesa CA), Key Organics (Cornwall U.K.), Lancaster Synthesis (Windham NH), Maybridge Chemical Co. Ltd. (Cornwall U.K.), Parish Chemical Co. (Orem UT), Pfaltz & Bauer, Inc. (Waterbury CN), Polyorganix (Houston TX), Pierce Chemical Co. (Rockford IL), Riedel de Haen AG (Hannover, Germany), Spectrum Quality Product, Inc. (New Brunswick, NJ), TCI America (Portland OR), Trans World Chemicals, Inc. (Rockville MD), Wako Chemicals USA, Inc. (Richmond VA), Novabiochem and Argonaut Technology.

In pharmaceutical dosage forms, the active agents and/or other compounds may be administered in the form of their pharmaceutically acceptable salts, or they may also be used alone or in appropriate association, as well as in combination with other pharmaceutically active compounds. The agents may be combined, as previously described, to provide a cocktail of activities. The following methods and excipients are exemplary and are not to be construed as limiting the invention.

Formulations are typically provided in a unit dosage form, where the term "unit dosage form," refers to physically discrete units suitable as unitary dosages for human subjects, each unit containing a predetermined quantity of active agent in an amount calculated sufficient to produce the desired effect in association with a pharmaceutically

acceptable diluent, carrier or vehicle. The specifications for the unit dosage forms of the present invention depend on the particular complex employed and the effect to be achieved, and the pharmacodynamics associated with each complex in the host.

In some embodiments a unit dose is at least about 0.1 mg/kg, at least about 0.5
5 mg/kg, at least about 1 mg/kg, at least about 5 mg/kg, at least about 10 mg/kg, at least about 20 mg/kg, at least about 50 mg/kg, at least about 100 mg/kg, in some embodiments the effective dose is from about 1 to 50 mg/kg. Dosing may be daily, every 2 days, every 3 or more days, e.g., weekly, semi-weekly, bi-weekly, monthly, etc. Dosing may be parenteral, including sustained release formulations. Dosing may be maintained for long
10 periods of time, e.g., months, or years, to maintain desirable glucose and fatty acid levels.

The pharmaceutically acceptable excipients, such as vehicles, adjuvants, carriers or diluents, are commercially available. Moreover, pharmaceutically acceptable auxiliary substances, such as pH adjusting and buffering agents, tonicity adjusting agents, stabilizers, wetting agents and the like, are commercially available. Any compound useful
15 in the methods and compositions of the invention can be provided as a pharmaceutically acceptable base addition salt. "Pharmaceutically acceptable base addition salt" refers to those salts which retain the biological effectiveness and properties of the free acids, which are not biologically or otherwise undesirable. These salts are prepared from addition of an inorganic base or an organic base to the free acid. Salts derived from inorganic bases
20 include, but are not limited to, the sodium, potassium, lithium, ammonium, calcium, magnesium, iron, zinc, copper, manganese, aluminum salts and the like. Preferred inorganic salts are the ammonium, sodium, potassium, calcium, and magnesium salts. Salts derived from organic bases include, but are not limited to, salts of primary, secondary, and tertiary amines, substituted amines including naturally occurring
25 substituted amines, cyclic amines and basic ion exchange resins, such as isopropylamine, trimethylamine, diethylamine, triethylamine, tripropylamine, ethanolamine, 2-dimethylaminoethanol, 2-diethylaminoethanol, dicyclohexylamine, lysine, arginine, histidine, caffeine, procaine, hydrabamine, choline, betaine, ethylenediamine, glucosamine, methylglucamine, theobromine, purines, piperazine,
30 piperidine, N-ethylpiperidine, polyamine resins and the like. Particularly preferred organic

bases are isopropylamine, diethylamine, ethanolamine, trimethylamine, dicyclohexylamine, choline and caffeine.

In some embodiments, pharmaceutical compositions can also include large, slowly metabolized macromolecules such as proteins, polysaccharides such as chitosan, polylactic acids, polyglycolic acids and copolymers (such as latex functionalized Sepharose™, agarose, cellulose, and the like), polymeric amino acids, amino acid copolymers, and lipid aggregates (such as oil droplets or liposomes).

A carrier may bear the agents in a variety of ways, including covalent bonding either directly or via a linker group, and non-covalent associations. Suitable covalent-bond carriers include proteins such as albumins, peptides, and polysaccharides such as aminodextran, each of which have multiple sites for the attachment of moieties. The nature of the carrier can be either soluble or insoluble for purposes of the invention.

Acceptable carriers, excipients, or stabilizers are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, and other organic acids; antioxidants including ascorbic acid and methionine; preservatives (such as octadecyldimethylbenzyl ammonium chloride; hexamethonium chloride; benzalkonium chloride, benzethonium chloride; phenol, butyl or benzyl alcohol; alkyl parabens such as methyl or propyl paraben; catechol; resorcinol; cyclohexanol; 3-pentanol; and m-cresol); low molecular weight (less than about 10 residues) polypeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids such as glycine, glutamine, asparagine, histidine, arginine, or lysine; monosaccharides, disaccharides, and other carbohydrates including glucose, mannose, or dextrans; chelating agents such as EDTA; sugars such as sucrose, mannitol, trehalose or sorbitol; salt-forming counter-ions such as sodium; metal complexes (e.g., Zn-protein complexes); and/or non-ionic surfactants such as TWEEN™, PLURONICS™ or polyethylene glycol (PEG). Formulations to be used for in vivo administration must be sterile. This is readily accomplished by filtration through sterile filtration membranes.

The active ingredients may also be entrapped in microcapsule prepared, for example, by coacervation techniques or by interfacial polymerization, for example, hydroxymethylcellulose or gelatin-microcapsule and poly-(methylmethacrylate)

microcapsule, respectively, in colloidal drug delivery systems (for example, liposomes, albumin microspheres, microemulsions, nano-particles and nanocapsules) or in macroemulsions. Such techniques are disclosed in Remington's Pharmaceutical Sciences 16th edition, Osol, A. Ed. (1980).

5 Compositions can be prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection can also be prepared. The preparation also can be emulsified or encapsulated in liposomes or micro particles such as polylactide, polyglycolide, or copolymer for enhanced adjuvant effect, as discussed above. Langer, Science 249: 1527, 1990 and
10 Hanes, Advanced Drug Delivery Reviews 28: 97-119, 1997. The agents of this invention can be administered in the form of a depot injection or implant preparation which can be formulated in such a manner as to permit a sustained or pulsatile release of the active ingredient. The pharmaceutical compositions are generally formulated as sterile, substantially isotonic and in full compliance with all Good Manufacturing Practice (GMP)
15 regulations of the U.S. Food and Drug Administration.

The following example(s) is/are offered by way of illustration and not by way of limitation.

EXAMPLES

20 The following examples are put forth so as to provide those of ordinary skill in the art with a complete disclosure and description of how to make and use the present invention, and are not intended to limit the scope of what the inventors regard as their invention nor are they intended to represent that the experiments below are all or the only experiments performed. Efforts have been made to ensure accuracy with respect to
25 numbers used (e.g. amounts, temperature, etc.) but some experimental errors and deviations should be accounted for. Unless indicated otherwise, parts are parts by weight, molecular weight is weight average molecular weight, temperature is in degrees Centigrade, and pressure is at or near atmospheric.

30 General methods in molecular and cellular biochemistry can be found in such standard textbooks as Molecular Cloning: A Laboratory Manual, 3rd Ed. (Sambrook et

al., HaRBor Laboratory Press 2001); Short Protocols in Molecular Biology, 4th Ed. (Ausubel et al. eds., John Wiley & Sons 1999); Protein Methods (Bollag et al., John Wiley & Sons 1996); Nonviral Vectors for Gene Therapy (Wagner et al. eds., Academic Press 1999); Viral Vectors (Kapliff & Loewy eds., Academic Press 1995); Immunology Methods
5 Manual (I. Lefkovits ed., Academic Press 1997); and Cell and Tissue Culture: Laboratory Procedures in Biotechnology (Doyle & Griffiths, John Wiley & Sons 1998), the disclosures of which are incorporated herein by reference. Reagents, cloning vectors, cells, and kits for methods referred to in, or related to, this disclosure are available from commercial vendors such as BioRad, Agilent Technologies, Thermo Fisher Scientific, Sigma-Aldrich,
10 New England Biolabs (NEB), Takara Bio USA, Inc., and the like, as well as repositories such as e.g., Addgene, Inc., American Type Culture Collection (ATCC), and the like.

Example 1 – Prediction and detection of capped peptides in mouse plasma

Here we provide evidence to show that capped peptides constitute a large class
15 of previously unstudied mammalian signaling peptides. Capped peptides are endogenously present in mouse and human plasma, where their levels are dynamically regulated by physiologic perturbations. Capped peptides exhibit post-translational N- and C-terminal modifications (pyroglutmylation and amidation, respectively) that resemble that of other peptide hormones and neuropeptides. Functional studies for previously
20 orphan capped peptides uncover functional bioactivity for previously orphan members of this class. The vast majority of the precise capped peptide sequences reported here have not been previously described as chemically defined, endogenous substances in mammals. In fact, our detection of >100 capped peptides in mouse and human plasma suggests that the N- and C-terminal “capping,” rather than being unusual and rare post-
25 translational modifications, instead defines a distinct chemical motif that is present in a large class of peptides with potential to mediate diverse axes of cell-cell communication.

Our ability to endogenously detect capped peptides was enabled by a custom mass spectrometry pipeline that uses a targeted mass spectrometry approach with authentic peptide standards. This method was inspired by classical approaches in
30 targeted metabolomics, where small molecule mass-to-charge ratios and retention times

are routinely compared against synthetic standards. The generality and simplicity of this approach was demonstrated by profiling the capped peptides present in blood plasma from two different species. Importantly, such a targeted approach obviates the need for large-scale database searching, thereby overcoming the problem of potential false discovery rate over-correction which may exclude true positive detection events. Because of potential limits of detection, it is not unreasonable to imagine that additional capped peptides that were undetectable here might indeed be endogenously present using more sensitive mass spectrometry methods. Projecting forward, these data also suggest that more targeted mass spectrometry approaches using synthetic standards, particularly on subsets of privileged peptide modifications and/or sequences, might be a powerful approach to identify peptides of specific structure and/or sequence within large, complex mass spectrometry datasets.

Peptide hormones and neuropeptides are fundamental signaling molecules that control diverse aspects of mammalian homeostasis and physiology. Using a hybrid computational-biochemical approach, we demonstrate the endogenous presence of a large and sequence diverse family of orphan blood-borne peptides that we call “capped peptides.” Capped peptides are modified fragments of secreted proteins and present in human and mouse blood. They exhibit structural and regulatory characteristics similar to other signaling peptides, including N-terminal pyroglutamylation, C-terminal amidation, cell type-specific expression, and dynamic environmental and physiologic regulation.

To define potential modified fragments from existing classically secreted proteins, we first used Uniprot (Consortium, 2023) to curate a collection of protein sequences corresponding to classically secreted mouse proteins. Our initial collection of N=2,835 sequences contained many known classically secreted proteins, including apolipoproteins, secreted enzymes, and preprohormones. Next, we sought to define potential fragments of these sequences that might also harbor additional post-translational modifications of the N- and C-termini. For these modifications, we focused on N-terminal pyroglutamylation and C-terminal amidation (**Fig. 1A**). C-terminal amidation sequences were identified by the presence of a glycine-dibasic GKK/GKR tripeptide (**Fig. 1A**). N-terminal pyroglutamylation sequences were identified by the

presence of a glutamine (Q) 3 to 20 amino acids upstream of the glycine-dibasic motif (**Fig. 1A** and **Fig. 14A**).

Our rationale for this set of prediction criteria, and for the focus on peptide fragments with coincident N-pyroglutamyl and C-amidation modifications, was based on the conceptual basis that: (1) these modifications have previously been shown to enhance receptor binding and plasma stability for other peptide hormones, therefore providing higher likelihood that our predicted sequences would be detectable by mass spectrometry and also exhibit bona fide signaling and bioactivity; (2) the glycine dibasic sequence is a well-established C-terminal cleavage motif that, after proteolysis, results a C-terminal glycine which can subsequently be enzymatically modified via PAM to generate the C-terminal amide; (3) biosynthesis of other N-terminal pyroglutamylated signaling peptides often occurs via an N-terminal glutamine, rather than a glutamate. Using this computational framework, we predicted a total of 216 potential cleaved and modified mouse peptides from 186 classically secreted proteins encoded in the mouse genome (**Fig. 1B**). We call this class of modified peptides “capped peptides” owing to their terminal modifications which function as “caps” of the intervening peptide sequence (**Figs. 1A-1B**). Our prediction of potentially hundreds of such modified sequences suggests that chemical capping, rather than being unusual and rare modifications that have been previously reported in only three mammalian peptide hormones (GnRH, TRH, and gastrin), might in fact define a distinct chemical class and potentially functional motif present in a large number of signaling peptides.

To determine whether capped peptides are produced endogenously, we used a targeted mass spectrometry approach to directly measure the levels of all predicted capped peptides in mouse plasma. Our targeted mass spectrometry strategy was inspired by standard approaches in small molecule metabolomics, where analytes are confidently identified by comparison of retention times and mass-to-charge (m/z) ratios to authentic standards that produced via chemical synthesis. Importantly, this mass spectrometry-based workflow obviates the need for identification via database searching that is classically associated with untargeted peptidomics or shotgun proteomics. Towards this end, we first isolated total plasma peptidome from mouse plasma. Briefly, mouse plasma was boiled to inactivate proteases, reduced with DTT, alkylated with iodoacetamide, and

concentrated using a C8 column. In parallel, we used solid phase peptide synthesis to generate authentic peptide standards for all 216 predicted capped peptides (**Fig. 1C**). A mixture of the 216 peptide standards was also reduced, alkylated, and concentrated. Total mouse plasma peptides were then compared to the synthetic peptide mixture by high resolution liquid chromatography-mass spectrometry (LC-MS) on a quadrupole-time-of-flight (qTOF) mass spectrometer.

A representative example of a positive detection event for CAP-TAC1 (pGlu-FFGLM-NH₂, **Fig. 1D**), a capped peptide derived from amino acids 63-68 of full-length TAC1 (protachykinin-1), is shown in **Fig. 1E**. Here, the authentic CAP-TAC1 standard exhibits identical m/z ratio and retention time to an endogenous plasma peak (m/z = 724.3, retention time = 28.5 min, **Fig. 1E**). As further validation of this structural assignment and detection event, MS/MS studies of the authentic CAP-TAC1 standard and the endogenous peak revealed nearly identical fragmentation patterns (**Fig. 1F**). Manual inspection of the daughter ions revealed that many could be directly assigned to the CAP-TAC1 sequence, including b₃(+1, m/z = 406.2), b₄(+1, m/z = 463.2), and a₅(+1, m/z = 548.3) (**Fig. 1F**).

Across all predicted capped peptides, 64 of the 216 (30%) exhibited high confidence detection in blood plasma by mass spectrometry (**Fig. 1G** and **Figs. 14B-14E**). The remaining 152 capped peptides (70%) did not exhibit a peak at the retention time of the authentic standard. By comparison to an external standard curve, the detected mouse capped peptides exhibited circulating concentrations in the range of ~0.1-100 nM (**Fig. 1H**). We conclude that capped peptides are a large family of endogenously circulating molecules. Our inability to detect 152 of the predicted capped peptides may either reflect true absence of post-translational processing to generate those fragments, or alternatively, the endogenous presence of the capped peptide is at a level below our detection limit. Regardless, that many predicted capped peptides could in fact be endogenously detected by a targeted mass spectrometry pipeline also suggest that specific proteolytic processing and capping to produce protected peptide fragments is much more prevalent than previously anticipated.

Example 2 – Sequence and gene-level analysis of mouse capped peptides

We next examined the sequences and set of full-length preproprecursor proteins of the detected capped peptides. Two of the detected capped peptides, CAP-GNRH1 (pGlu-HWSYGLRPG-NH₂) and CAP-GAST (pGlu-RPRMEEEEAYGWMDf-NH₂) directly corresponded with the known hormone sequences for GnRH and gastrin (**Figs. 2A-2B**). These data demonstrate that our hybrid computational-analytical approach can “re-discover” two of the known signaling peptides that harbor both N-terminal pyroglutamylation and C-terminal amidation. An additional 23 detected capped peptides mapped to preproprecursor proteins and corresponding genes that had previously been annotated to generate polypeptides with signaling bioactivity, but for which a shorter cleavage fragment had not been previously identified. For instance, we observed a capped tripeptide, CAP-FGF18 (pGlu-VL-NH₂), derived from amino acids 66-69 of the prepro-FGF18 (fibroblast growth factor 18) (**Fig. 2B**). A second capped peptide, CAP-GDNF (pGlu-AAAASPENSRGK-NH₂), mapped to amino acids 94-106 of GDNF (glial cell line-derived neurotrophic factor) (**Fig. 2B**). Classically, FGF18 is a so-called “paracrine” member of the fibroblast growth factor (FGF) family and has diverse roles in the development of several tissues, including bone, lung, and the nervous system (Itoh et al., 2016). GDNF is a major growth factor that promotes the survival of dopaminergic and motor neurons; outside of the nervous system, GDNF is also a morphogen in the kidney and a spermatogonia differentiation factor (Airaksinen and Saarma, 2002). Our detection of these two capped peptides in blood plasma suggests that FGF18 and GDNF might also exhibit endocrine functions via cleavage fragments generated from their canonical polypeptide sequences. Lastly, the remaining capped peptides (39/64, ~61%) mapped to preproprecursors for proteins and genes that had not been previously suggested to have any roles in signaling. These include CAP-COL27A1 (pGlu-LGPP-NH₂), which is cleaved from a collagen protein; CAP-PLA2G2A (pGlu-FGEMIRLKT-NH₂) which is cleaved from a phospholipase sequence; and CAP-SNED1 (pGlu-STEVDRSVDRLTFGDLLP-NH₂) which is derived from the poorly studied protein SNED1 (**Fig. 2B**).

To further understand the cellular and tissue origin of capped peptides, we next examined the mRNAs of the home genes encoding the capped peptide preproprecursors. We used BioGPS (Wu et al., 2016) as a reference mouse tissue gene expression dataset.

As shown in **Fig. 2C**, this set of mRNAs exhibited both cell type-specific as well as widespread tissue expression. For instance, a strong enrichment of home gene mRNAs for certain capped peptides was found in the brain (e.g., CAP-CBLN2, CAP-TENM1, CAP-TAC3, CAP-ADCYAP1), in bone (e.g., CAP-MATN4, CAP-CEMIP), and in
5 macrophages (e.g., CAP-GDF15, CAP-TREM2). Conversely, mRNAs corresponding to other capped peptides exhibited more diffuse tissue expression across multiple cell types and organs (e.g., CAP-VIP enrichment in both brain and gut and CAP-COL5A2 expression in > 10 tissues).

Lastly, we performed more detailed amino acid composition and sequence
10 analysis of the detectable capped peptides from mouse plasma. As a reference and comparison set, we once again used Uniprot to manually curate a set of known mouse peptide hormones and neuropeptides. Glutamine was enriched in capped peptides compared to the reference set of known peptide hormones and neuropeptides, which was expected based on our original computational search criteria. In addition, two hydrophobic
15 amino acids, valine and leucine, were also more prevalent in capped peptides, whereas two polar amino acids (arginine, serine) were less represented (**Fig. 2D**). To understand whether there might be additional sequence-specific determinants of capping beyond our original N-terminal Q and C-terminal GKK/GKR motifs, we examined the amino acid sequences centered around the N- and C-termini. A modest enrichment of glycine and
20 leucine were observed to flank both the N-terminal pyroglutamylation motif (**Fig. 2E**). In addition to glycine and leucine enriched around the C-terminal amidation motif, we also observed a strong enrichment for alanine at the +2 position (**Fig. 2F**). Together, these data demonstrate that capped peptides are produced from diverse tissues and exhibit specific patterns of amino acid composition and sequence.

Example 3 – Dynamic regulation of capped plasma levels in mice

Many signaling peptides exhibit dynamic regulation in a manner dependent on internal physiologic state or external environmental conditions. We therefore measured the circulating levels of capped peptides after six distinct perturbations that spanned a
30 wide range of physiologic processes, environmental stimuli, organ systems, and time

scales: 16 h fasting vs. fed, 8-weeks high fat diet feeding vs. chow feeding, lipopolysaccharide (LPS, 0.5 mg/kg, intraperitoneal) vs. vehicle, 6AM vs. 6PM, acute treadmill running (1 h) vs. sedentary, and 3 months vs. 24 months old. For each comparison, mouse plasma was collected and processed as described previously and capped peptides were quantified by LC-MS (**Fig. 1**).

As shown in **Fig. 3A**, each physiologic comparison resulted in bidirectional regulation of a unique subset of capped peptides. Across all measurements (N=384, corresponding to N=64 capped peptides in each of N=6 conditions), we observed a total of N=93 (24%) instances for a particular capped peptide being elevated (N=39) or suppressed (N=54) by more than 2-fold in any condition. 73% (47/64) of the capped peptides exhibited at least one instance of 2-fold change across any condition; similarly, 100% of the conditions (6/6) yielded at least one capped peptide that was changed by more than 2-fold.

The capped peptide/perturbation pair resulting in the most dramatic regulation was CAP-CSF1 (pGlu-LLLPSKSHSWGIVLPLGELE-NH₂), derived from amino acids 419-438 of full-length prepro-CSF1. Plasma CAP-CSF1 levels were induced by ~84-fold after LPS treatment ($P < 0.01$, **Figs. 3A-3B**). Importantly, CAP-CSF1 levels were unchanged in any of the other comparisons (**Fig. 3A**), establishing that induction of CAP-CSF1 in plasma is a specific response to an inflammatory stimulus. Previously, the most well-known polypeptide product derived from full-length prepro-CSF1 is m-CSF1 (macrophage colony-stimulating factor 1), which is itself an LPS-inducible cytokine (Benmerzoug et al., 2018). The co-induction of CAP-CSF1 may therefore represent additional, LPS-inducible proteolytic processing of m-CSF1. In addition, we could also identify several other interesting examples of individually regulated dynamic peptides in each of the conditions. For instance, CAP-GDNF was selectively downregulated in plasma collected at 6PM versus 6AM (pGlu-AAAASPENSRGK-NH₂, 58% reduction, $P < 0.05$, **Fig. 3C**) and CAP-FGF5 was selectively induced by a single bout of treadmill running (1 h) versus sedentary mice (pGlu-WSPS-NH₂, 2.6-fold increase, $P < 0.05$, **Fig. 3D**).

Beyond high magnitude changes in individual capped peptides in each condition, we also identified examples of capped peptides that exhibited coordinate regulation across multiple physiologic states. For instance, we observed a cluster capped peptides

that were coordinately regulated in two distinct nutritional stressors, fasting and high fat diet feeding. Individual examples of dynamic capped peptides within this nutrition-regulated cluster included CAP-COL27A1 (pGlu-LGPP-NH₂, ~75% reduction, $P < 0.05$, **Fig. 3E**) and CAP-VGF (pGlu-VEA-NH₂, >3-fold induction, $P < 0.05$, **Fig. 3F**). The nutritional regulation of this subset of capped peptides, and of CAP-COL27A1 and CAP-VGF in particular, might point to specific functions in nutrient harvesting, fuel metabolism, or energy homeostasis. Together, we conclude that capped peptide levels in the circulation are dynamically regulated in a manner dependent on the specific capped peptide and specific physiologic perturbation.

Example 4 – CAP-TAC1 is a novel tachykinin with homology to Substance P

Our data so far suggests that capped peptides exhibit many structural and regulatory features of other well-established peptide hormones and neuropeptides. We next sought to determine whether any of the detectable capped peptides exhibited signaling and/or functional bioactivity. We first focused on CAP-TAC1 (pGlu-FFGLM-NH₂), an orphan peptide of previously unknown function. As we already showed in **Fig. 1**, CAP-TAC1 is robustly detected in blood plasma. The full-length TAC1 preproprotein encodes multiple members of the tachykinin neuropeptides, including Neurokinin A/Substance K, Neuropeptide K/Neurokinin K, Neuropeptide gamma, and Substance P (**Fig. 4A**) (Steinhoff et al., 2014). Of these known tachykinin neuropeptides, CAP-TAC1 exhibits most homology to substance P.

We noted that the sequence of CAP-TAC1 contains the key consensus C-terminal FXGLM-NH₂ motif which is characteristic of all known tachykinin neuropeptides (**Fig. 4A**). In addition, the C-terminal methionyl amide has been shown to be critical for agonist activity (Escher et al., 1982; Patacchini et al., 1993). The presence of the tachykinin motif and an intact methionyl amide in CAP-TAC1 directly suggested that this capped peptide might also be a tachykinin neuropeptide-like molecule. We therefore used a cellular human TACR1-beta arrestin recruitment assay with a fluorescence readout to directly determine the ability of CAP-TAC1 to agonize the TACR1, a high affinity receptor for substance P (Bhatia et al., 1998). As shown in **Fig. 4B**, CAP-TAC1 exhibited dose-

dependent and high potency agonism of TACR1 ($ED_{50} = 0.7$ nM). As a positive control, substance P exhibited a similar dose-dependent activation ($ED_{50} = 1.7$ nM). Both CAP-TAC1 and substance P exhibited similar levels of maximal activation (CAP-TAC1, 96.6% of maximal response; substance P, 99.8% of maximal response). We conclude that CAP-TAC1 is a full agonist of the TACR1 receptor with potency similar to the previously known TACR1 ligand substance P.

Despite similar agonist activity, CAP-TAC1 and substance P exhibit differences in chemical structure: CAP-TAC1 contains an additional N-terminal pyroglutamylation and is also shorter in length compared to substance P. We reasoned these two chemical differences might result in important functional differences in terms of stability and resistance to proteolytic degradation. To directly test this possibility, CAP-TAC1 (10 μ M) and substance P (10 μ M) were individually incubated with mouse plasma and incubated at 37°C. and their levels over time were measured by LC-MS. Substance P exhibited time-dependent degradation with a $t_{1/2} = 8.4$ min. By contrast, the rate of CAP-TAC1 degradation was substantially slower ($t_{1/2} = 14.4$ min) (**Fig. 4C**). In fact, levels of CAP-TAC1 were still detectable after 90 min, a time point when substance P was undetectable (**Fig. 4C**). Together, these data demonstrate that CAP-TAC1 exhibits similarities (e.g., TACR1 agonism) as well as important differences (e.g., plasma stability) compared to previously described tachykinin neuropeptides.

Our identification of CAP-TAC1, and then demonstration of this molecule as a high affinity agonist of mammalian tachykinin receptors, shows that additional fragments of full-length tachykinin preproteins may be important endogenous mediators tachykinin signaling.

Example 5 – A 12-mer hypophagic capped peptide from the precursor region of GDF15

We next sought to understand whether signaling and bioactivity might be indeed a general feature of many capped peptides beyond CAP-TAC1 alone. We next focused on functional studies of a second orphan capped peptide, CAP-GDF15 (pGlu-LELRLRVAAGR-NH₂, **Fig. 5A**). Full-length GDF15 is a secreted, 303 amino acid preproprecursor that, upon cleavage at R188, produces a C-terminal 114-amino acid

anorexigenic protein hormone which is also called GDF15 (Chrysovergis et al., 2014; Johnen et al., 2007; Macia et al., 2012). Interestingly, CAP-GDF15 mapped to amino acids 174-185, a region just upstream of the canonical GDF15 hormone and localized in the GDF15 prepropeptide region (**Fig. 5A**). CAP-GDF15 showed identical co-elution with an authentic standard (**Fig. 5B**). In addition, tandem mass spectrometry fragmentation also revealed a similar fragmentation pattern between the endogenous peak and the authentic standard (**Fig. 5C**). These data demonstrate that full-length GDF15 precursor encodes at least two polypeptide products.

Unlike CAP-TAC1, the amino acid sequence of CAP-GDF15 did not immediately provide insights into its potential functions. However, we reasoned that the hypophagic effects previously demonstrated by overexpression of full-length GDF15 might extend beyond the classical GDF15 hormone alone to also include CAP-GDF15. To test this possibility, we administered a single dose of CAP-GDF15 (50 mg/kg, intraperitoneally) to diet-induced obese mice and measured whole body parameters of energy balance in metabolic chambers. CAP-GDF15 strongly suppressed food intake by ~60% compared to vehicle-treated mice (**Fig. 5D**). A corresponding and expected suppression of respiratory exchange ratio (RER) was also observed (**Fig. 5E**). Notably, CAP-GDF15 did not alter movement (**Fig. 5F**), oxygen consumption (**Fig. 6A**), or carbon dioxide production (**Fig. 6B**), demonstrating that the pharmacological effects of this peptide are specific to feeding control rather than other pathways of energy expenditure.

We next synthesized a control CAP-GDF15 peptide that preserved amino acid composition but scrambled the intervening amino acid sequence (scrambled CAP-GDF15, pGlu-GLEALRRLRV-NH₂). This scrambled peptide control was completely ineffective in suppressing food intake and RER in metabolic chambers under identical experimental conditions (**Figs. 5G-5I** and **Figs. 6C-6D**). We conclude that the hypophagic effects of CAP-GDF15 are specific to this amino acid sequence.

Lastly, to determine whether the acute hypophagic effects of CAP-GDF15 would lead to long-term suppression of feeding and obesity, we administered CAP-GDF15 or scrambled CAP-GDF15 (50 mg/kg/day, IP), or vehicle control to diet-induced obese mice. Food intake and body weight were monitored over a three-day period. A durable suppression of food intake in CAP-GDF15-treated mice was observed over the three-day

experiment (**Fig. 5J**). Consequently, and as expected, an increasing reduction in body weight was also detected (**Fig. 5K**). Importantly, mice treated with the scrambled CAP-GDF15 peptide were indistinguishable in body weight or food intake from control mice (**Figs. 5J-5K**). These data show that chronic CAP-GDF15 administration suppresses food intake and reduces body weight in a sequence-dependent manner.

Together with CAP-TAC1, these data on CAP-GDF15 provide functional evidence for the signaling and bioactivity of orphan capped peptides in both cell and animal models. Like CAP-TAC1, the detection of CAP-GDF15 also demonstrates that a single full-length preproprecursor (in this case, full-length GDF15) can generate more than a single bioactive polypeptide product. The observation that CAP-GDF15 is an anorexigenic peptide like the canonical GDF15 hormone raises new questions about the relative physiologic contribution of each CAP-GDF15 and canonical GDF15. In addition, because the sequences are largely distinct, we suspect that the downstream receptor(s) of CAP-GDF15 are likely to be distinct from that of the canonical GDF15 hormone.

Example 6 – Detection of human capped peptides and sequence comparison to mice

The capped peptide discovery pipeline described here only requires a full genome sequence and authentic peptide standards. Therefore, such an approach should also be readily amenable for discovering capped peptides in other species. Towards this end, we used the same hybrid computational-biochemical workflow as shown in **Fig. 1**, but now applied to protein sequences corresponding to classically secreted human proteins. Starting from N=3,791 secreted proteins, we predicted a total of 260 potential human capped peptides from 231 proteins (**Fig. 15** and **Fig. 8A**). We synthesized authentic peptide standards by solid phase peptide synthesis corresponding to all 260 possible human capped peptides. Once again, the retention time and parent masses of all authentic capped peptides were compared to endogenous peaks from commercially available pooled human plasma. In total, we robustly detected N=85 human capped peptides by mass spectrometry (**Fig. 7A**), a number that, by percentage, is similar to that previously observed with mice (30% detected/predicted for mouse, and 33% detected/predicted for human). Human capped peptides exhibited a similar distribution in plasma abundance (as quantitated using an external standard curve, **Fig. 8B**) and similar

sequence characteristics in terms of amino acid composition (**Fig. 8C**) as mouse capped peptides. Using GTEx as a reference gene expression dataset, human capped peptides were also derived from preproprecursors whose mRNA levels also exhibited tissue-restricted, as well as more broad expression (**Fig. 9**).

5 Because of sequence differences between the mouse and human proteome, we predicted that the set of humans capped peptides should be overlapping, but still distinct compared to those present in mouse plasma. Indeed, we detected many capped peptides in both human and mouse plasma that were 100% sequence conserved. In addition, none of the 61 human-specific peptides were detected in mouse plasma, and none of the 41
10 mouse-specific peptides were found in human plasma (**Fig. 7B**). Representative extracted ion chromatograms corresponding to conserved and species-specific capped peptides are shown in **Figs. 7C-7D**.

Next, we performed a multiple sequence alignment to globally understand the sequence relationship and homology across all capped peptide sequences from both
15 mice and humans. We also performed this analysis to understand whether the human- and mouse-specific capped peptide constituted entirely distinct sequences, high homologous sequences, or some combination of these two possibilities. The resulting dendrogram is shown in **Fig. 7E**. We selected several sub-clusters as illustrative examples here. Clusters "A" consists of a pair of capped peptides, mouse and human
20 CAP-CSF1. These two are highly homologous sequences derived from amino acids 419-438 and 424-441 of full-length mouse and human CSF1, respectively. However, because the sequences are not 100% identical, mouse and human CAP-CSF1 are considered species-specific in our analysis. A similar example is shown in Cluster "D" where mouse and human CAP-EDN3 differ only by two amino acid residues and are once again
25 considered species-specific, homologous sequences. Clusters "A" and "D" demonstrate that at least a subset of the species-specific sequences is due to differences in amino acid sequences of the corresponding full-length preproproteins from which the capped peptides are derived.

In another example, cluster "B" contained four short 3- and 5-mer capped peptides,
30 which were amongst the shortest sequences in the entire dataset. The 3-mer capped peptides (pGlu-VL-NH₂) were derived from the full-length mouse and human FGF18

sequences and exhibited identity between the two species. The other two capped peptides, derived from CNPY4 and again identical between mouse and human, constitute CAP-FGF18 homologs with an aspartyl-threonyl C-terminal extension (pGlu-VLDT-NH₂). This cluster demonstrates that highly homologous capped peptides can also be produced from distinct full-length preproprotein precursors.

Lastly, the cluster labeled “C” contained four peptides, three of which were derived from the full-length mouse or human VIP preproprecursor. The three VIP-derived capped peptides correspond with C-terminal fragments of the known PHI-27 and VIP peptide hormones. We also identified a non-VIP-derived peptide, CAP-ADCYAP1, which also exhibited high sequence alignment within this CAP-VIP-enriched cluster. CAP-ADCYAP1 is a C-terminal fragment of the neuropeptide PACAP (pituitary adenylate cyclase-activating polypeptide). These data suggest that the similar signal transduction pathways of VIP and PACAP peptides might also extend to additional fragments of these canonical sequences. We conclude that at least a subset of the human- and mouse-specific capped peptides represent highly homologous sequences. These data also globally identify similarities as well as important differences in the sequences of capped peptides between two species.

Example 7 – CAP-TAC1 induces body weight loss independent of food intake

CAP-TAC1 induces body weight loss in diet induced obese mice (**Fig. 10A**), independent of any changes of food intake (**Fig. 10B**), suggesting CAP-TAC1 chronically activates energy expenditure. CAP-TAC1 induces an increase in oxygen consumption with a single injection in diet induced obese mice (**Fig. 11**), showing CAP-TAC1 acutely activates energy expenditure.

Example 8 – Human CAP-GDF15 inhibits food intake

The human homologue of CAP-GDF15 acutely inhibits food intake in diet induced obese mice (**Fig. 12**), showing the human homologue has a similar hypophagic effect in mice as the endogenous mouse CAP-GDF15, despite some sequence differences.

Example 9 – CAP-WNT9A shows growth hormone secretagogue activity

In a functional screen in GH3 cells, a rat pituitary cell line, various capped peptides are shown to activate growth hormone secretion (**Fig. 13A**). Of the capped peptides that were screened, CAP-WNT9A shows the highest growth hormone secretagogue activity.

- 5 After a one hour treatment of CAP-WNT9A, growth hormone concentration in media is increased by four times (**Fig. 13B**).

Materials and methods

Mice and treatments. Animal experiments were performed according to a
10 procedure approved by the Stanford University Administrative Panel on Laboratory Animal Care. Mice were maintained in 12-h light-dark cycles at 22°C and about 50% relative humidity and fed a standard irradiated rodent diet. Where indicated, a high-fat diet (D12492, Research Diets 60% kcal from fat) was used. Male C57BL/6J (stock number 000664) and male C57BL/6J DIO mice were purchased from the Jackson
15 Laboratory (stock number 380050). For studies in high fat diet-fed mice, peptides were dissolved in 18:1:1 (by volume) of saline:Kolliphor EL (Sigma Aldrich):DMSO and administered to mice by intraperitoneal injections at a volume of 10 µl/g at the indicated doses for the indicated times. For lipopolysaccharide injection, LPS was dissolved in saline and administered to mice at a volume of 5 µl/g at indicated dose. For fasting, food
20 was removed from mice for 16 h. For running, a six-lane Columbus Instruments animal treadmill (product 1055-SRM-D65) was used with following 1 h protocol: 10 min at 6 m/min, 50 min at 18 m/min, and increase every 2 min by 2 m/min for the last 10 minutes, all at 12° incline. For all treatment experiments, mice were mock injected with the vehicle for 3-5 days until body weights were stabilized. Heparin plasma was harvested by
25 submandibular bleed. For all experiments, mice were randomly assigned to treatment groups. Experimenters were not blinded to groups.

CAP-TAC1 treatments in mice. Injected diet induced obese (on high fat diet, 4 months old, male, C57BL/6J, Jackson cat # 380050) mice daily with vehicle (18:1:1
30 Saline:DMSO:Kolliphor EL) or 10 mg/kg CAP-TAC1 (pGlu-FFGLM-NH₂). The pair fed

group was started one day later and fed the same food the CAP-TAC1 mice ate daily. Mice were singly housed and injected daily ~12 pm. Mouse body weights and food weights were measured at the same time. (N=6 for vehicle, N=7 for CAP-TAC1 and pair fed groups.)

5

Human CAP-GDF15 treatments in mice. Singly housed diet induced obese (on high fat diet, 4 months old, male, C57BL/6J, Jackson cat # 380050) mice were injected with vehicle (18:1:1 Saline:DMSO:Kolliphor EL) or 50 mg/kg CAP-GDF15 human homologue (pGlu-LELHLRPQAAR-NH₂) once at ~ 5 pm. Food weight was measured before injection and 2 hours post to calculated food intake. (N=8 mice/group)

10

In vitro studies. GH3 cells (rat pituitary cell line, ATCC cat. #CCL-82.1) were cultured in F-12K medium with 2.5% fetal bovine serum and 15% horse serum. GH3 cells were washed, resuspended in Krebs Ringer Hepes Buffer, diluted to 0.5 million cells/mL, and aliquoted into tubes (300 uL/sample). The GH3 cells were then starved for 60 minutes in Krebs Ringer Hepes Buffer. The capped peptides (N=1 in screen and N=3 in repeat with CAP-WNT9A) or DMSO (N=3) control were added at a 10 uM final concentration. GH3 cells were incubated for 1 hour. Cells were pelleted with 5 min centrifuge spin at 500 xg at 4C. Growth hormone was measured in the supernatant with an ELISA kit (Thermo Cat # KRC5311) according to product protocol. Z-score of growth hormone concentration is shown. CAP-WNT9A (pGlu-WGGCGDNLKYSSKFVEFL-NH₂) (the top hit from the screen) was validated with the same protocol in triplicate (N=3).

15

20

Uniprot dataset curation. Lists of classically secreted proteins were obtained from Uniprot using the keyword "secreted" and filtering for either human or mouse species. Known peptide hormone sequences were obtained from Uniprot by first filtering for proteins annotated with keyword "hormone" for function and subsequently extracting out the specific hormone sequences from the peptides listed under the PTM annotations.

25

Computational prediction of capped peptides. Capped peptide prediction was accomplished using an in-house custom algorithm written in python. First, a list of

30

classically secreted proteins was obtained from Uniprot using the keyword "secreted." Next, C-terminal amidation motifs were identified based on a GKR or GKK sequence indicative of dibasic cleavage and then amidation. N-terminal pyroglutamylation was identified by searching for Q residues within 20 amino acids upstream of the amidation motif, and capped peptides were predicted to be the inclusive sequence between the N-terminal (pyro)glutamine and the C-terminal amidation.

Solid-phase synthesis of capped peptides. Capped peptides were synthesized using standard solid phase synthesis protocols.

Plasma and authentic peptide standards preparation for peptidomics. To carry out plasma peptidomics, 1 µl of protease inhibitor (HALT) was added to 100 µl of plasma. Plasma was diluted to 1:6 plasma:Tris-HCl buffer (100 mM Tris-HCl, pH 8.2) and boiled at 95°C for 10 minutes. In total, 1 ml of pooled plasma was used per replicate. 1 mM dithiothreitol (DTT) was added, samples were vortexed and incubated for 50 minutes at 60°C. Iodoacetamide was added to obtain a final concentration of 5 mM and incubated at room temperature for 1 hour in the dark. Formic acid was added to 0.2% final concentration. Samples were centrifuged at 15,000 rpm for 20 min. Supernatants were concentrated with C8 columns (Waters, WAT054965), washed/desalted with water, and eluted in 100 µl of 80% acetonitrile. Samples were centrifuged at 15,000 rpm for 10 min. Supernatant was collected for liquid chromatography-mass spectrometry (LC-MS) analysis. All authentic peptide standards were pooled into 1 ml of Tris-HCl buffer (100 mM Tris-HCl, pH 8.2) and prepared in the same way as described above.

LC-MS detection of capped peptides. LC-MS was performed on an Agilent 6520 Quadrupole time-of-flight LC-MS instrument. MS analysis was performed using electrospray ionization (ESI) in positive mode. The dual ESI source parameters were set as follows: the gas temperature at 325°C, the drying gas flow rate at 13 l/min, the nebulizer pressure at 30 psig, the capillary voltage at 4,000 V, and the fragmentor voltage at 175 V. Separation of peptides was conducted using a C18 column with reverse phase chromatography. Mobile phases were as follows: buffer A, 100% water with 0.1% formic

acid; buffer B, 90:10 acetonitrile:water with 0.1% formic acid. The LC gradient started at 95% A with a flow rate of 0.7 ml/min from 0 to 3 minutes. The gradient was then linearly increased to 40%A/60%B from 3 to 28 minutes and subsequently flushed at 5%A/95%B for 4 minutes and equilibrated back to 95%A/5%B for 6 minutes all at a flow rate of 0.7 ml/min. A positive capped peptide detection required a peak of exact mass (± 50 ppm) with total area >1000 ion count and co-elution with the corresponding authentic synthetic standard. Exact masses and retention times of all detected peptides were recorded.

Targeted LC-MS/MS. Targeted LC-MS/MS spectra were obtained using Agilent 6545 Quadrupole time-of-flight LC-MS instrument. The dual ESI source parameters were set as follows: the gas temperature at 325°C, the drying gas flow rate at 13 l/min, the nebulizer pressure at 30 psig, the capillary voltage at 4,000 V, the fragmentor voltage at 185 V, the sheath gas temperature at 350°C, and the sheath gas flow rate at 11 l/min. The LC separation was done as described above. For CAP-GDF15, targeted LC-MS/MS method fragmented the +2 ion ($m/z = 682.4$) over the entire run with a collision energy of 40 and a narrow isolation width (~ 1.3 m/z). For CAP-TAC1, targeted LC-MS/MS method fragmented the +1 ion ($m/z = 724.3$) over the entire run with a collision energy of 25 and a narrow isolation width (~ 1.3 m/z).

TACR agonist assay. Dose-response curves for CAP-TAC1 and positive control Substance P on the agonism of TACR1 was measured by a Eurofins Discovery using human TACR1-transfected PathHunter beta-arrestin CHO-K1 cells.

Half-life calculations of CAP-TAC1 in plasma. 10 μ M of either synthetic CAP-TAC1 or SP was incubated with 100 μ l of mouse plasma. Samples were incubated at 37°C for 0, 10, 30, 60, or 90 minutes (N=3 for each time point). At the indicated time point, 1 μ l of HALT protease inhibitor was immediately added, and plasma was boiled and prepared using the peptidomics workflow described above. LC-MS spectra were obtained using Agilent 6545 Quadrupole time-of-flight LC-MS instrument as described above. Relative peptide levels were determined by total ion count area of exact mass (± 50 ppm, $m/z = 724.3$, +1 for CAP-TAC1 and $m/z = 674.4$, +1 for Substance P) peak that coeluted with

synthetic standards. Basal plasma concentrations of Substance P and CAP-TAC1 were determined in samples with no synthetic peptide spiked in (N=3) with a standard curve comparison. Half-lives were calculated using an exponential decay fit.

5 Metabolic chamber studies. For acute energy expenditure studies, food intake, RER, movement, VO₂, and VCO₂ were collected with CLAMS Oxymax Metabolic Cages. Mice were placed individual metabolic cages for 24 hours prior to experiment. For CAP-GDF15 experiments, mice were injected with either vehicle, 50 mg/kg CAP-GDF15, or 50 mg/kg scrambled CAP-GDF15 at 5 pm at T=0. For CAP-TAC1 experiments, mice were
10 injected with either vehicle (18:1:1 Saline:DMSO:Kolliphor EL) or 15 mg/kg CAP-TAC1 (pGlu-FFGLM-NH₂) at ~12 pm (T=0). Food intake, RER, movement, VO₂, and VCO₂ were collected every 7 minutes for 16 hours immediately after injection. All experiments were done with 3-5 month old DIO mice (Jackson, stock 380050), fed high fat diet ad libitum.

15

In at least some of the previously described embodiments, one or more elements used in an embodiment can interchangeably be used in another embodiment unless such a replacement is not technically feasible. It will be appreciated by those skilled in the art that various other omissions, additions and modifications may be made to the methods
20 and structures described above without departing from the scope of the claimed subject matter. All such modifications and changes are intended to fall within the scope of the subject matter, as defined by the appended claims.

It will be understood by those within the art that, in general, terms used herein, and especially in the appended claims (*e.g.*, bodies of the appended claims) are generally
25 intended as “open” terms (*e.g.*, the term “including” should be interpreted as “including but not limited to,” the term “having” should be interpreted as “having at least,” the term “includes” should be interpreted as “includes but is not limited to,” etc.). It will be further understood by those within the art that if a specific number of an introduced claim recitation is intended, such an intent will be explicitly recited in the claim, and in the
30 absence of such recitation no such intent is present. For example, as an aid to understanding, the following appended claims may contain usage of the introductory

phrases “at least one” and “one or more” to introduce claim recitations. However, the use of such phrases should not be construed to imply that the introduction of a claim recitation by the indefinite articles “a” or “an” limits any particular claim containing such introduced claim recitation to embodiments containing only one such recitation, even when the same claim includes the introductory phrases “one or more” or “at least one” and indefinite articles such as “a” or “an” (*e.g.*, “a” and/or “an” should be interpreted to mean “at least one” or “one or more”); the same holds true for the use of definite articles used to introduce claim recitations. In addition, even if a specific number of an introduced claim recitation is explicitly recited, those skilled in the art will recognize that such recitation should be interpreted to mean at least the recited number (*e.g.*, the bare recitation of “two recitations,” without other modifiers, means at least two recitations, or two or more recitations). Furthermore, in those instances where a convention analogous to “at least one of A, B, and C, etc.” is used, in general such a construction is intended in the sense one having skill in the art would understand the convention (*e.g.*, “a system having at least one of A, B, and C” would include but not be limited to systems that have A alone, B alone, C alone, A and B together, A and C together, B and C together, and/or A, B, and C together, etc.). In those instances where a convention analogous to “at least one of A, B, or C, etc.” is used, in general such a construction is intended in the sense one having skill in the art would understand the convention (*e.g.*, “a system having at least one of A, B, or C” would include but not be limited to systems that have A alone, B alone, C alone, A and B together, A and C together, B and C together, and/or A, B, and C together, etc.). It will be further understood by those within the art that virtually any disjunctive word and/or phrase presenting two or more alternative terms, whether in the description, claims, or drawings, should be understood to contemplate the possibilities of including one of the terms, either of the terms, or both terms. For example, the phrase “A or B” will be understood to include the possibilities of “A” or “B” or “A and B.”

In addition, where features or aspects of the disclosure are described in terms of Markush groups, those skilled in the art will recognize that the disclosure is also thereby described in terms of any individual member or subgroup of members of the Markush group.

As will be understood by one skilled in the art, for any and all purposes, such as in

terms of providing a written description, all ranges disclosed herein also encompass any and all possible sub-ranges and combinations of sub-ranges thereof. Any listed range can be easily recognized as sufficiently describing and enabling the same range being broken down into at least equal halves, thirds, quarters, fifths, tenths, etc. As a non-limiting example, each range discussed herein can be readily broken down into a lower third, middle third and upper third, etc. As will also be understood by one skilled in the art all language such as “up to,” “at least,” “greater than,” “less than,” and the like include the number recited and refer to ranges which can be subsequently broken down into sub-ranges as discussed above. Finally, as will be understood by one skilled in the art, a range includes each individual member. Thus, for example, a group having 1-3 articles refers to groups having 1, 2, or 3 articles. Similarly, a group having 1-5 articles refers to groups having 1, 2, 3, 4, or 5 articles, and so forth.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it is readily apparent to those of ordinary skill in the art in light of the teachings of this invention that certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims.

Accordingly, the preceding merely illustrates the principles of the invention. It will be appreciated that those skilled in the art will be able to devise various arrangements which, although not explicitly described or shown herein, embody the principles of the invention and are included within its spirit and scope. Furthermore, all examples and conditional language recited herein are principally intended to aid the reader in understanding the principles of the invention and the concepts contributed by the inventors to furthering the art, and are to be construed as being without limitation to such specifically recited examples and conditions. Moreover, all statements herein reciting principles, aspects, and embodiments of the invention as well as specific examples thereof, are intended to encompass both structural and functional equivalents thereof. Additionally, it is intended that such equivalents include both currently known equivalents and equivalents developed in the future, i.e., any elements developed that perform the same function, regardless of structure. Moreover, nothing disclosed herein is intended to

be dedicated to the public regardless of whether such disclosure is explicitly recited in the claims.

5 The scope of the present invention, therefore, is not intended to be limited to the exemplary embodiments shown and described herein. Rather, the scope and spirit of present invention is embodied by the appended claims. In the claims, 35 U.S.C. §112(f) or 35 U.S.C. §112(6) is expressly defined as being invoked for a limitation in the claim only when the exact phrase "means for" or the exact phrase "step for" is recited at the beginning of such limitation in the claim; if such exact phrase is not used in a limitation in the claim, then 35 U.S.C. § 112 (f) or 35 U.S.C. §112(6) is not invoked.

WHAT IS CLAIMED IS:

1. A pharmaceutical composition comprising:
a capped peptide comprising an N-terminal pyroglutamylation modification and a C-terminal amidation modification, wherein the capped peptide is not Thyrotropin-releasing hormone (TRH) or (Gonadotropin-releasing hormone (GnRH)/luteinizing hormone-releasing hormone (LHRH); and
a pharmaceutically acceptable delivery vehicle.
2. The pharmaceutical composition according to Claim 1, wherein the capped peptide comprises from 2 to 30 amino acid residues.
3. The pharmaceutical composition according to Claim 2, wherein the capped peptide comprises from 3 to 25 amino acid residues.
4. The pharmaceutical composition according to any of the preceding claims, wherein the capped peptide is a murine peptide.
5. The pharmaceutical composition according to Claim 4, wherein the capped peptide is a peptide listed in Fig. 14.
6. The pharmaceutical composition according to any of Claims 1 to 3, wherein the capped peptide is a human peptide.
7. The pharmaceutical composition according to Claim 6, wherein the capped peptide is a peptide listed in Fig. 15.
8. The pharmaceutical composition according to any of the preceding claims, wherein the capped peptide exhibits physiological activity.

9. The pharmaceutical composition according to Claim 8, wherein the physiological activity is mammalian tachykinin receptor agonist activity.
10. The pharmaceutical composition according to Claim 9, wherein capped peptide is CAP-TAC1 or a peptide having 80% or more identity with CAP-TAC1.
11. The pharmaceutical composition according to Claim 8, wherein the physiological activity is activating metabolic activity.
12. The pharmaceutical composition according to Claim 11, wherein capped peptide is CAP-TAC1 or a peptide having 80% or more identity with CAP-TAC1.
13. The pharmaceutical composition according to Claim 8, wherein the physiological activity is hypophagic activity.
14. The pharmaceutical composition according to Claim 13, wherein capped peptide is CAP-GDF15 or a peptide having 80% or more identity with CAP-GDF15.
15. The pharmaceutical composition according to Claim 8, wherein the physiological activity is growth hormone secretagogue activity.
16. The pharmaceutical composition according to Claim 15, wherein capped peptide is CAP-WNT9A or a peptide having 80% or more identity with CAP-WNT9A.
17. A method comprising administering a capped peptide to a subject in need thereof, wherein the capped peptide comprises an N-terminal pyroglutamylation modification and a C-terminal amidation modification, and is not Thyrotropin-releasing hormone (TRH) or (Gonadotropin-releasing hormone (GnRH)/luteinizing hormone-releasing hormone (LHRH).

18. The method according to Claim 17, wherein the capped peptide comprises from 2 to 30 amino acid residues.
19. The method according to Claim 18, wherein the capped peptide comprises from 3 to 25 amino acid residues.
20. The method according to any of Claims 17 to 19 wherein the capped peptide is a murine peptide.
21. The method according to Claim 20, wherein the capped peptide is a peptide listed in Fig. 14.
22. The method according to any of Claims 19 to 21, wherein the capped peptide is a human peptide.
23. The method according to Claim 22, wherein the capped peptide is a peptide listed in Fig. 15.
24. The method according to any of Claims 17 to 23, wherein the capped peptide exhibits physiological activity.
25. The method according to Claim 24, wherein the physiological activity is mammalian tachykinin receptor agonist activity.
26. The method according to Claim 25, wherein capped peptide is CAP-TAC1 or a peptide having 80% or more identity with CAP-TAC1.
27. The method according to Claim 24, wherein the physiological activity is activating metabolic activity.

28. The method according to Claim 27, wherein capped peptide is CAP-TAC1 or a peptide having 80% or more identity with CAP-TAC1.
29. The method according to Claim 24, wherein the physiological activity is hypophagic activity.
30. The method according to Claim 29, wherein capped peptide is CAP-GDF15 or a peptide having 80% or more identity with CAP-GDF15.
31. The method according to Claim 24, wherein the physiological activity is growth hormone secretagogue activity.
32. The method according to Claim 31, wherein capped peptide is CAP-WNT9A or a peptide having 80% or more identity with CAP-WNT9A.
33. The method according to Claims 25 or 26, wherein the method is a method of treating a subject for a neurological disorder.
34. The method according to Claim 33, wherein the neurological disorder is depressive disorder, anxiety disorder, chronic pain, addiction, pruritus, psychosis or schizophrenia.
35. The method according to Claims 25 or 26, wherein the method is a method of treating a subject for an inflammatory disorder.
36. The method according to Claim 35, wherein the inflammatory disorder is irritable bowel syndrome, inflammatory bowel disease, asthma, colitis, ileus or fibrosis.
37. The method according to Claims 25 or 26, wherein the method is a method of treating a subject for nausea and vomiting.

38. The method according to any of Claims 27 to 30, wherein the method is a method of treating a subject for obesity or an obesity-related disorder.

39. The method according to Claims 31 or 32, wherein the method is a method of treating a subject for a growth hormone deficiency or a growth hormone deficiency-related disorder.

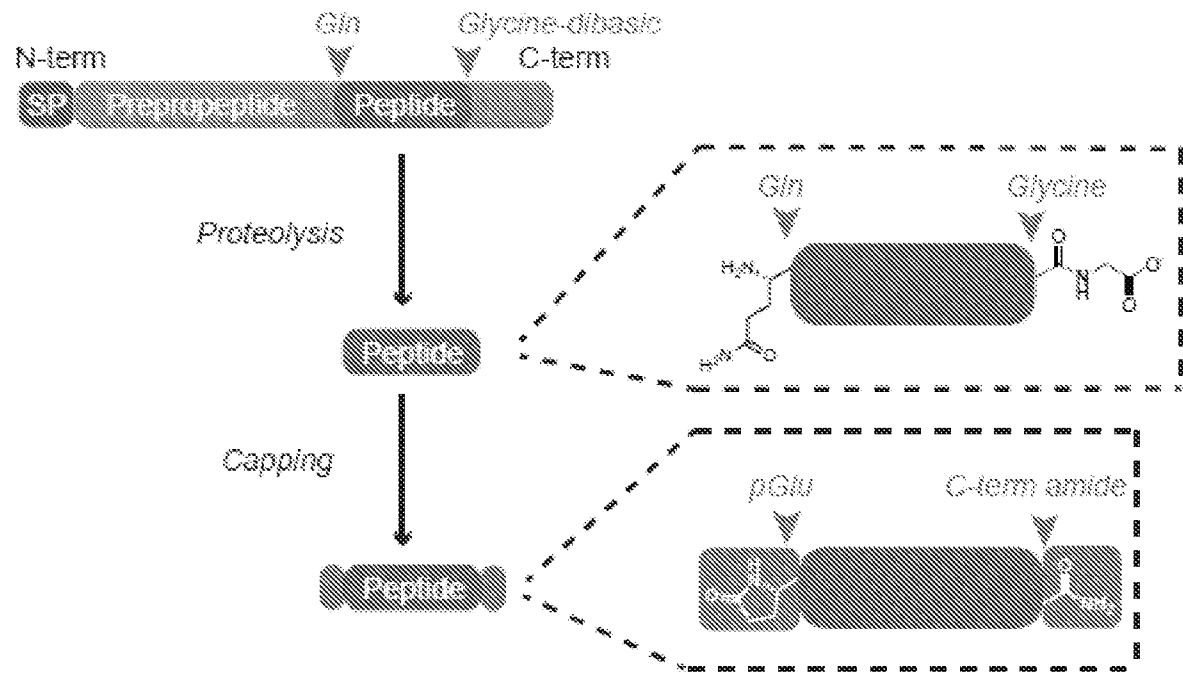


FIG. 1A

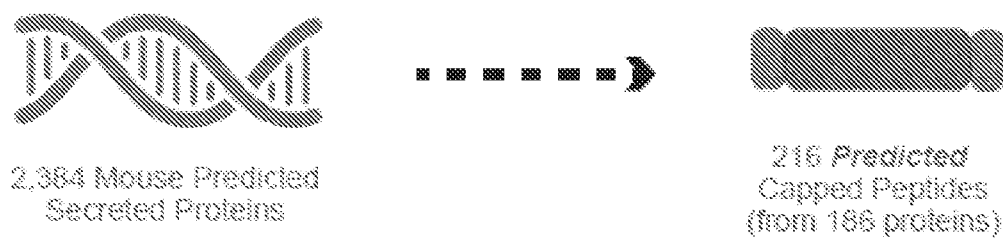


FIG. 1B

2/88

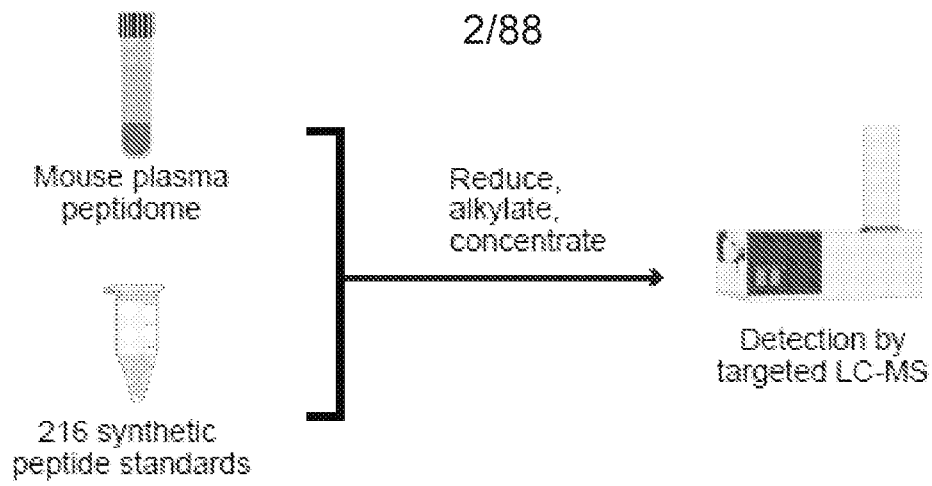
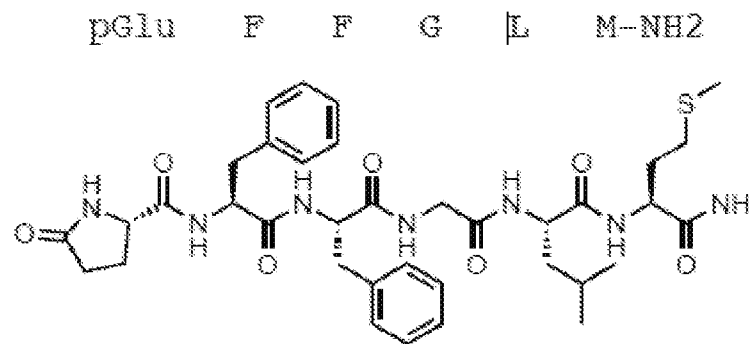


FIG. 1C



CAP-TAC1

FIG. 1D

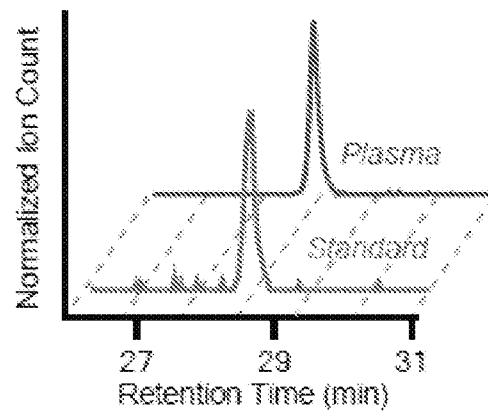


FIG. 1E

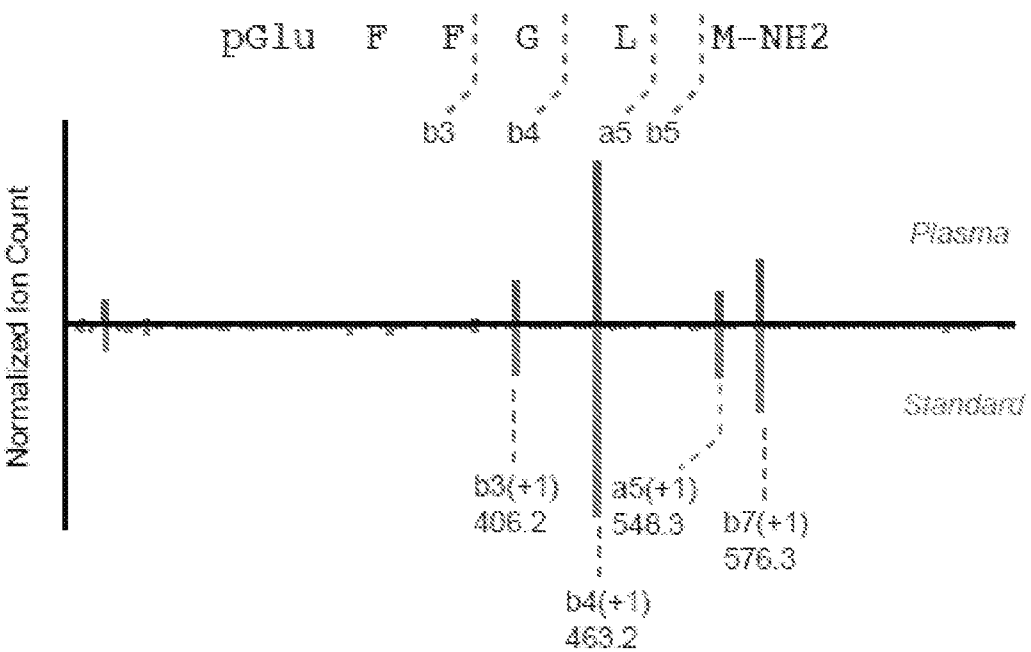


FIG. 1F

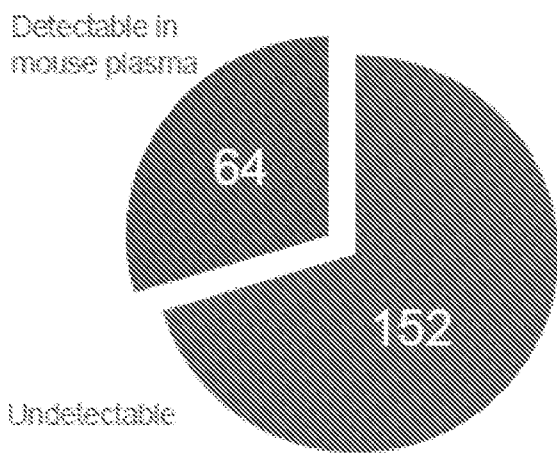


FIG. 1G

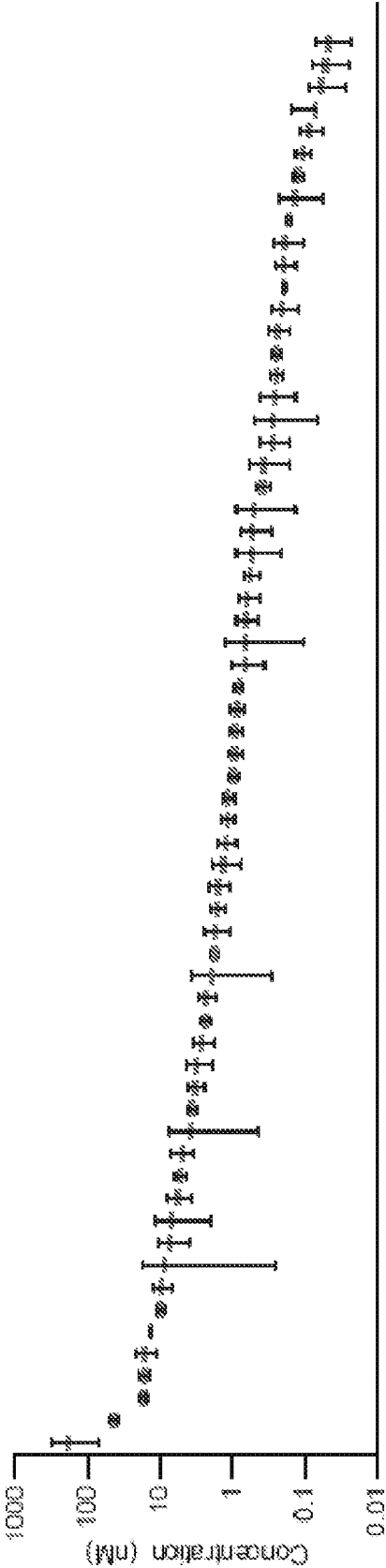


FIG. 1H

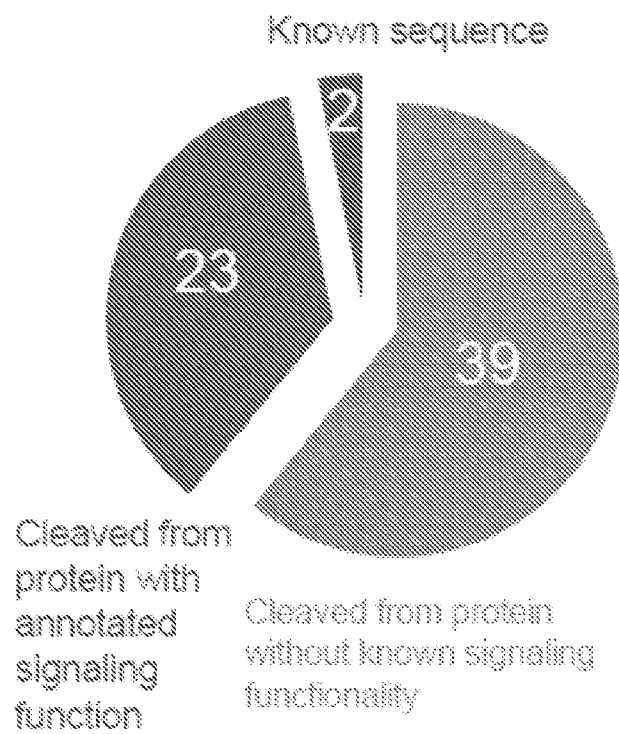


FIG. 2A

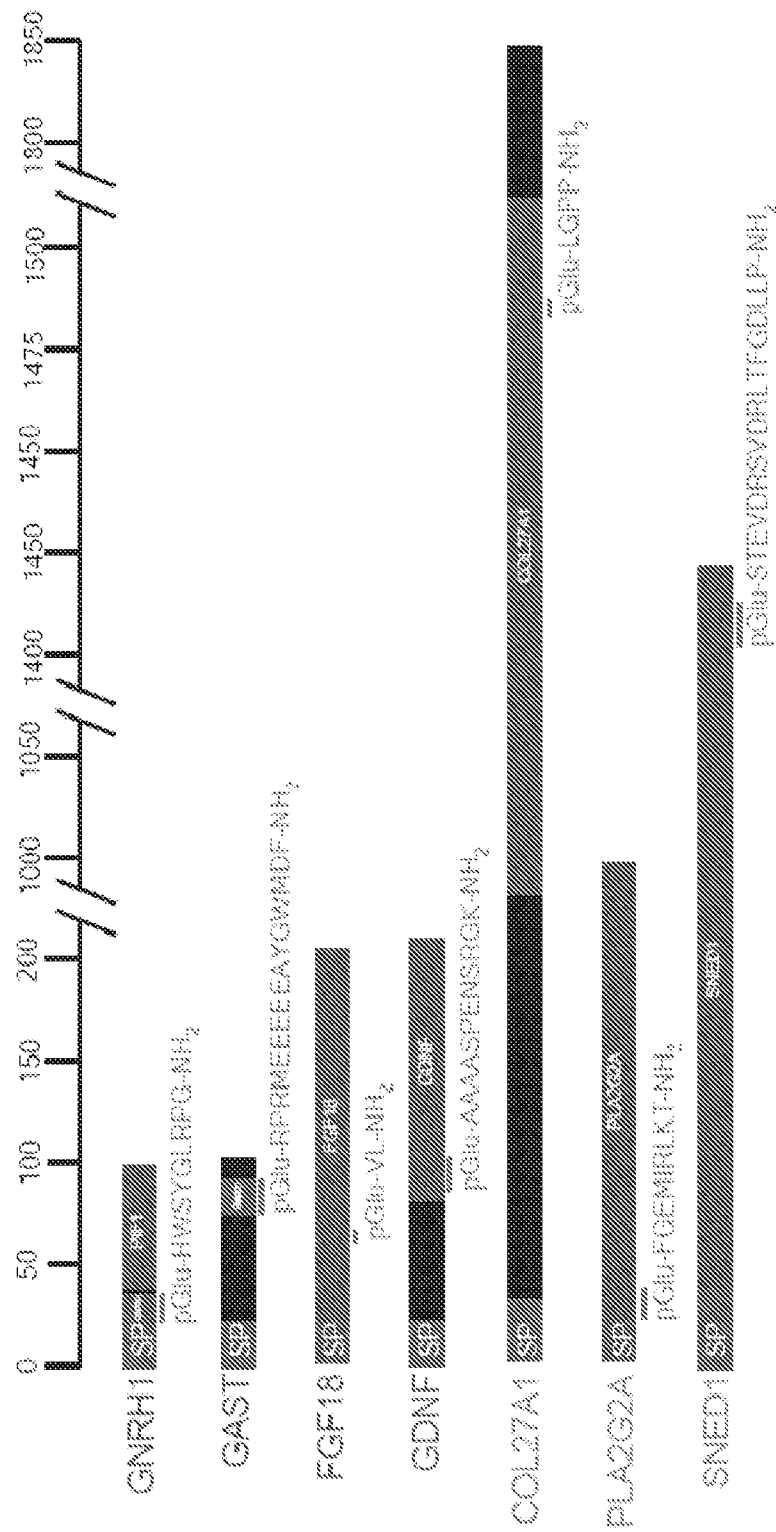


FIG. 2B

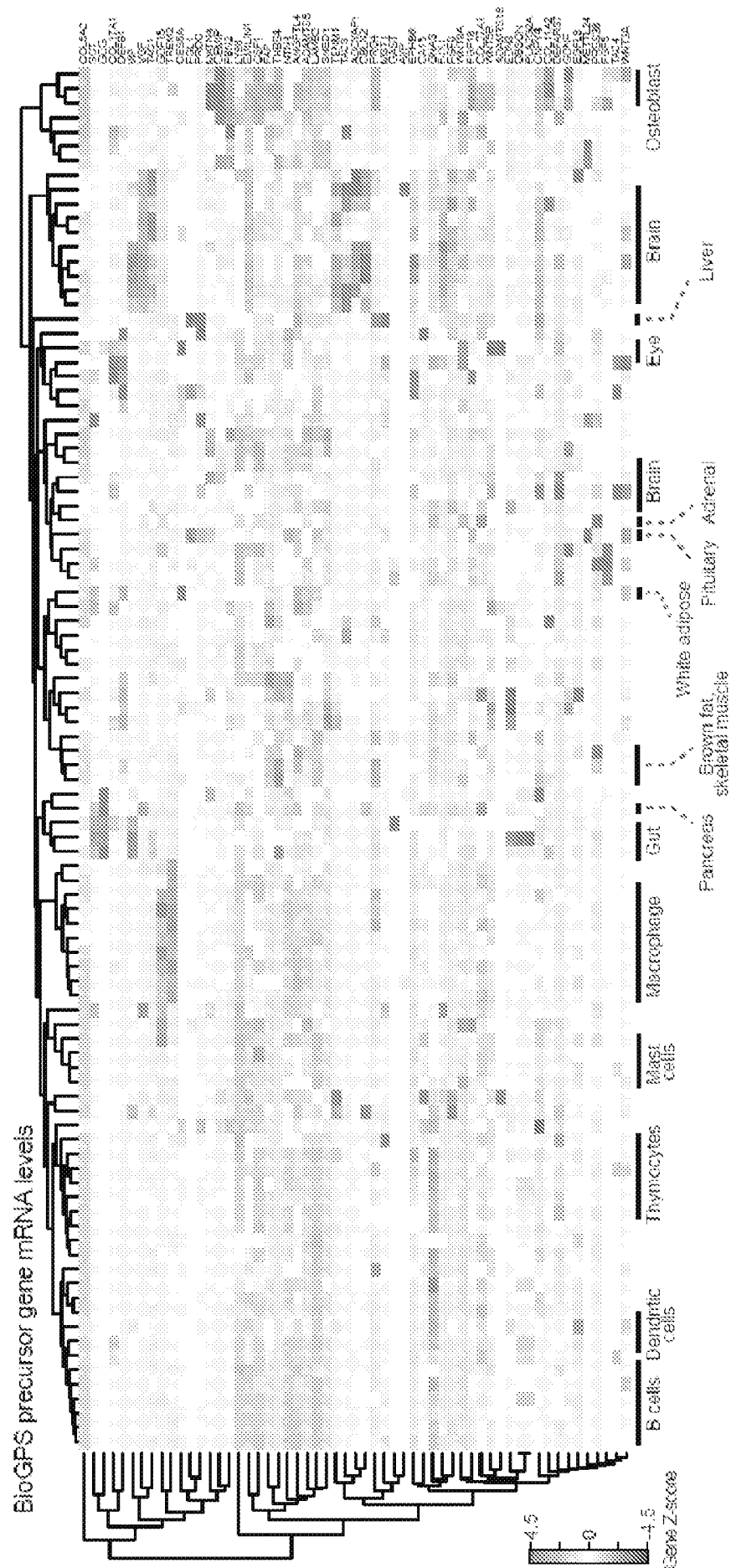


FIG. 2C

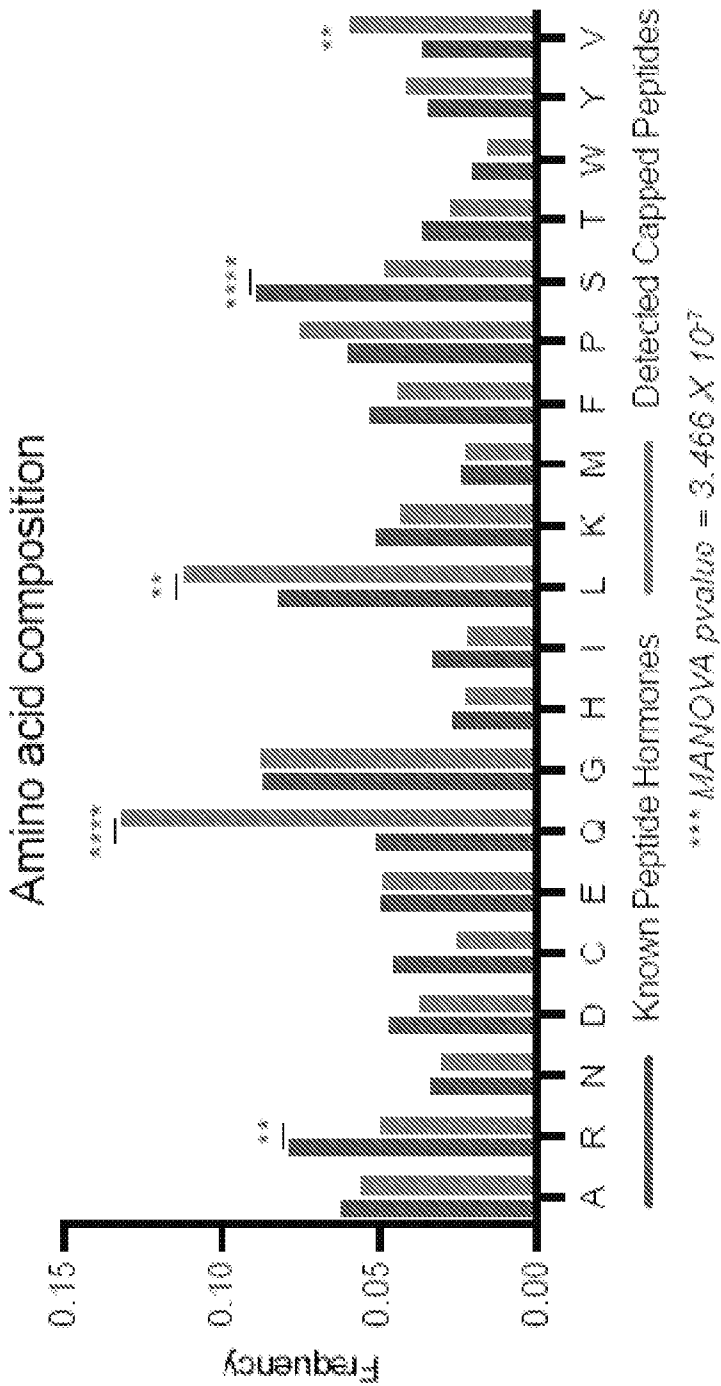


FIG. 2D

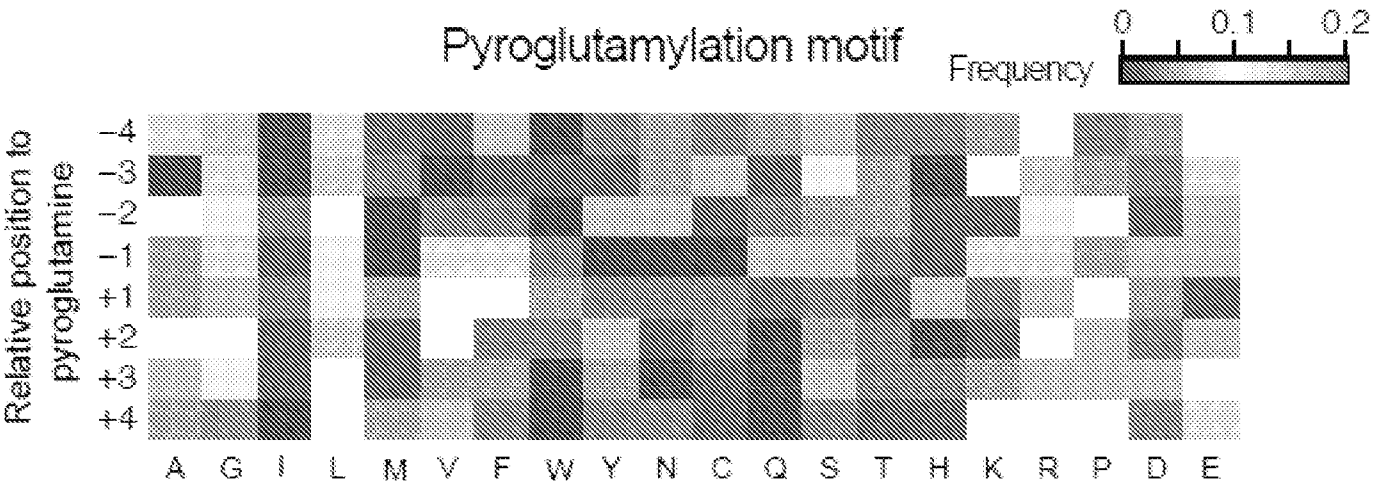


FIG. 2E

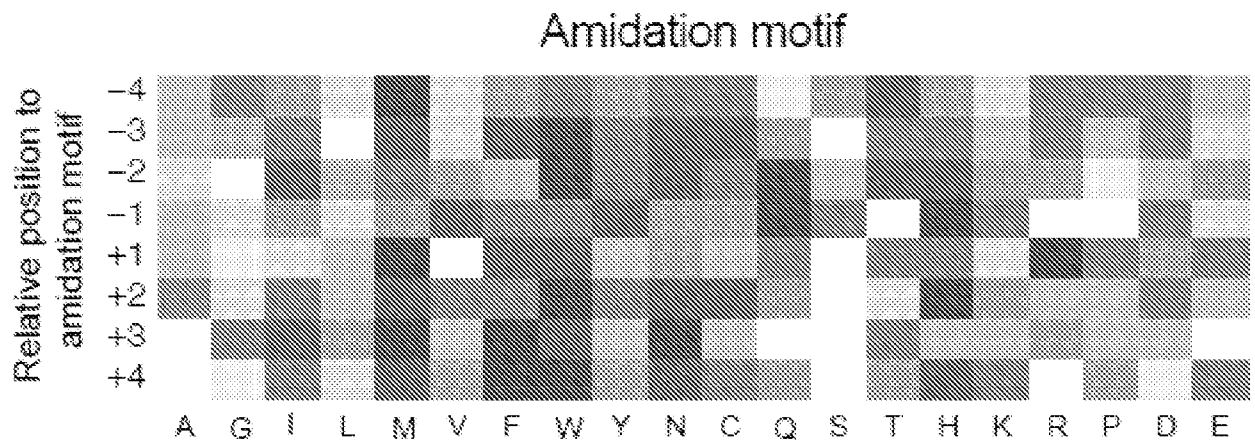


FIG. 2F

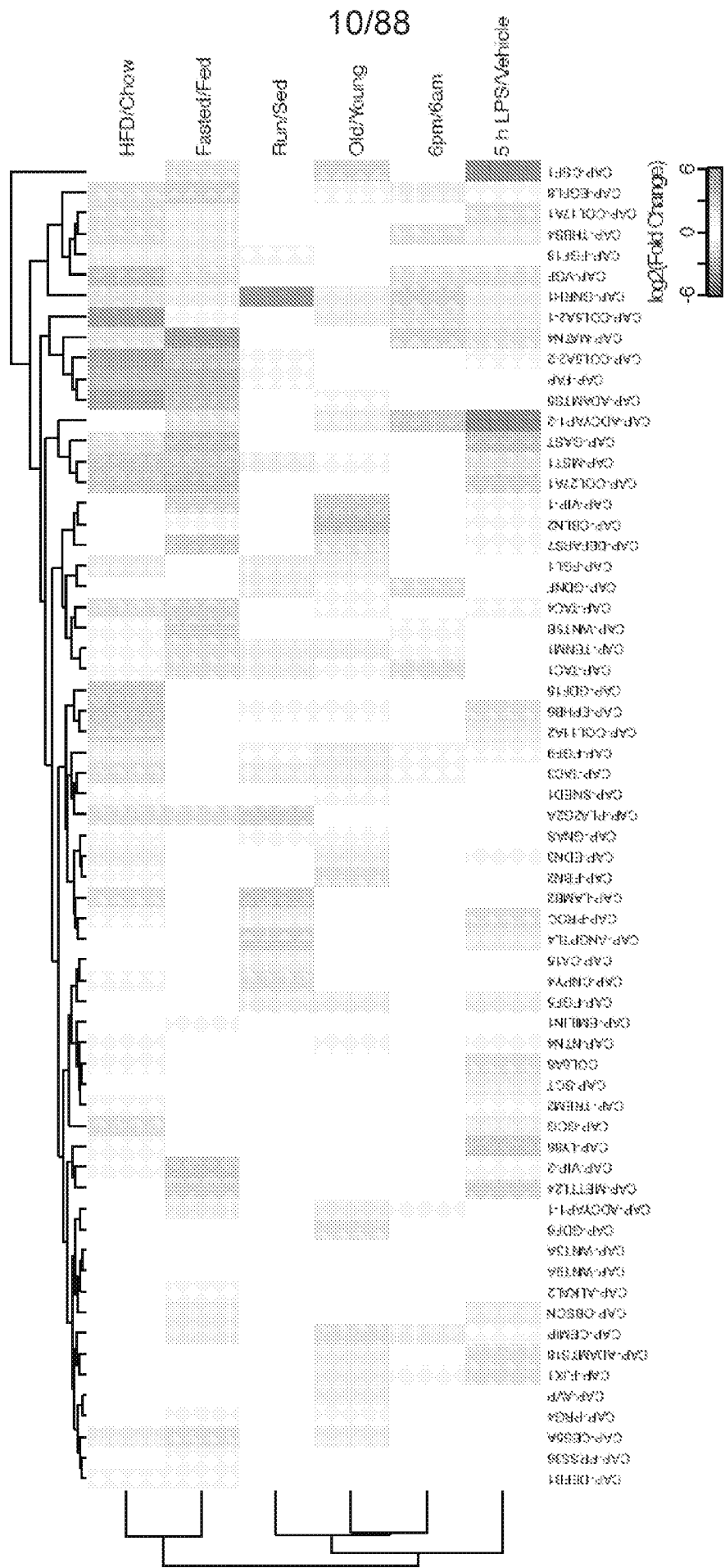


FIG. 3A

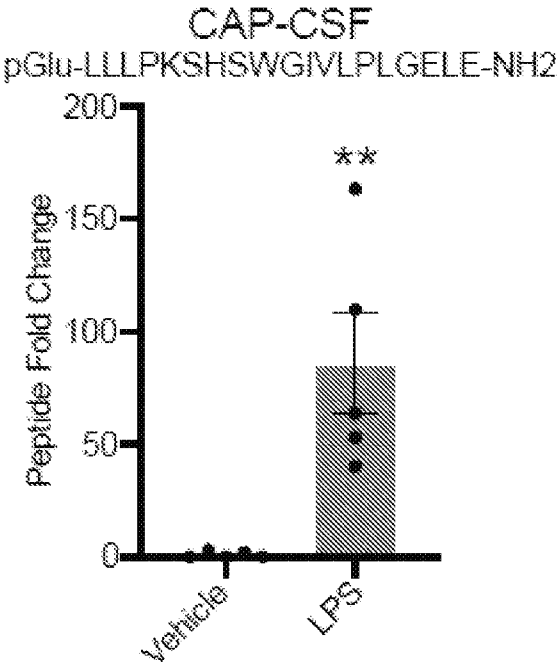


FIG. 3B

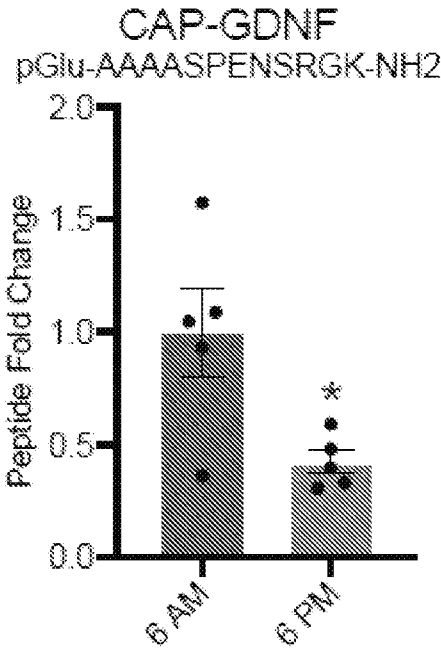


FIG. 3C

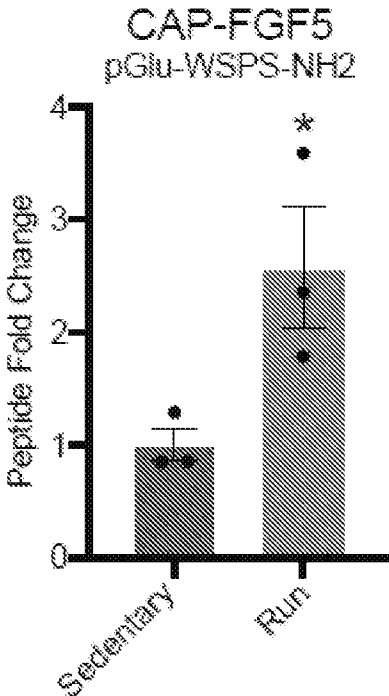


FIG. 3D

12/88

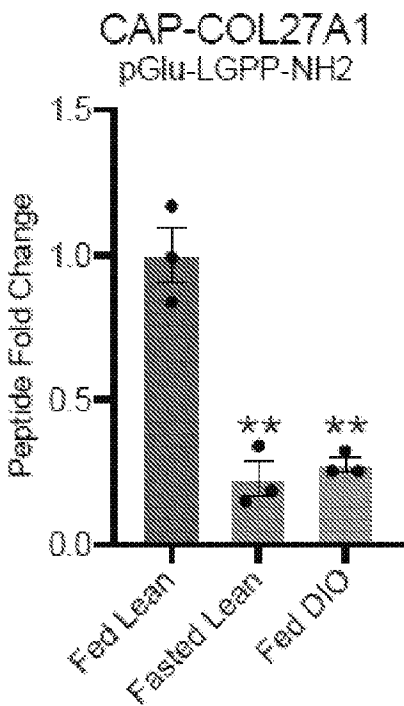


FIG. 3E

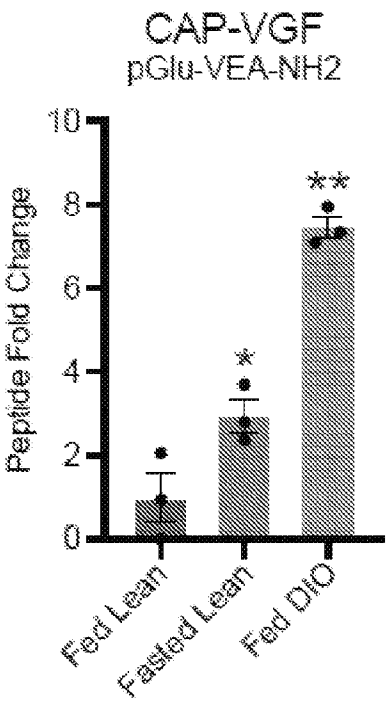


FIG. 3F

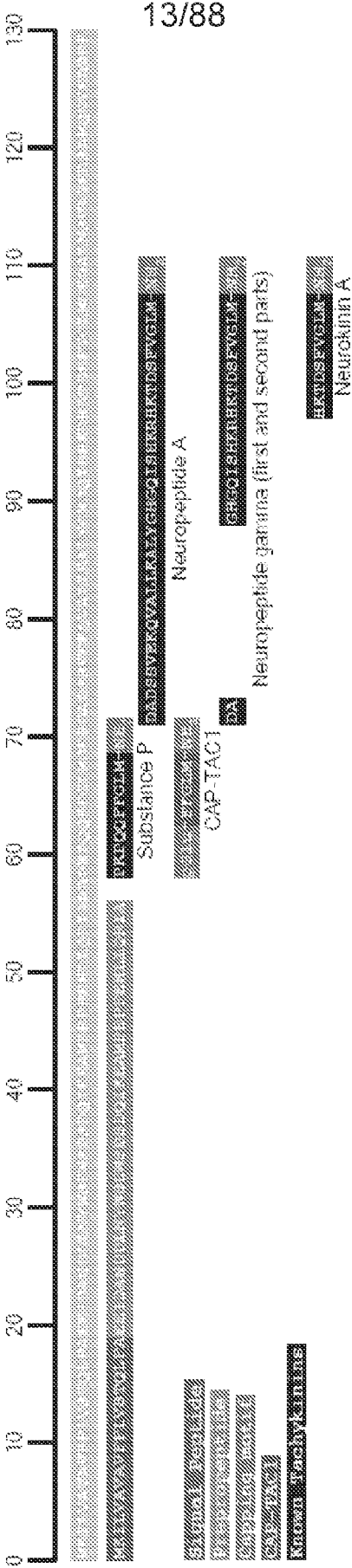


FIG. 4A

14/88

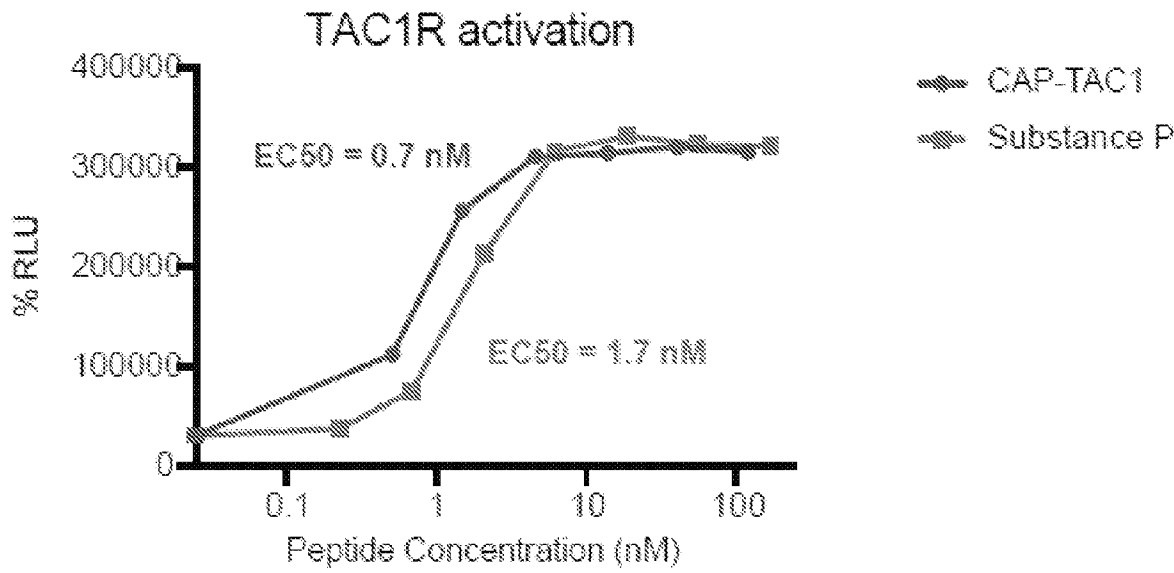


FIG. 4B

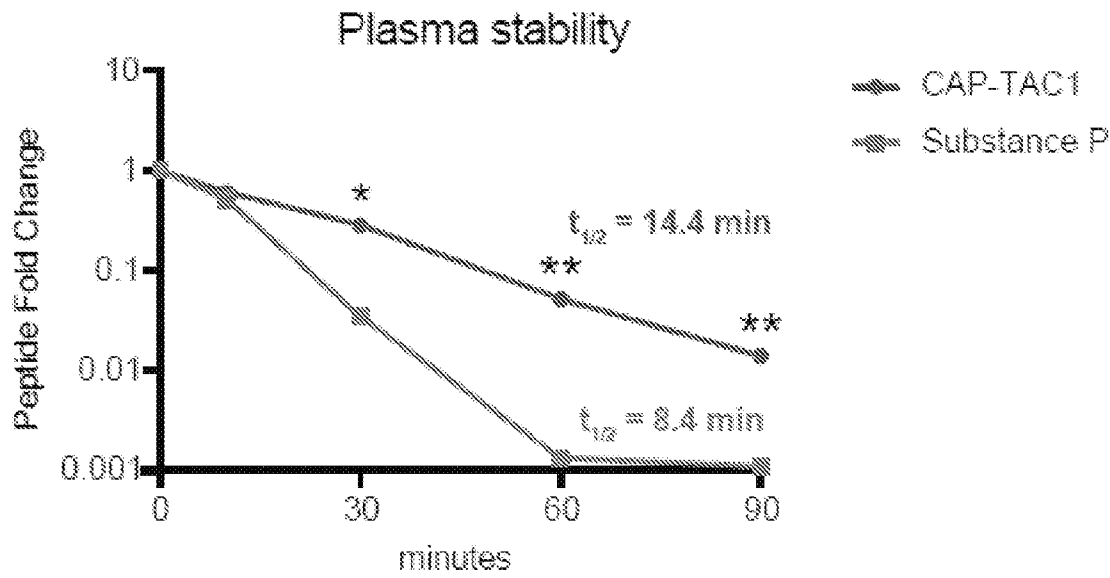


FIG. 4C

15/88

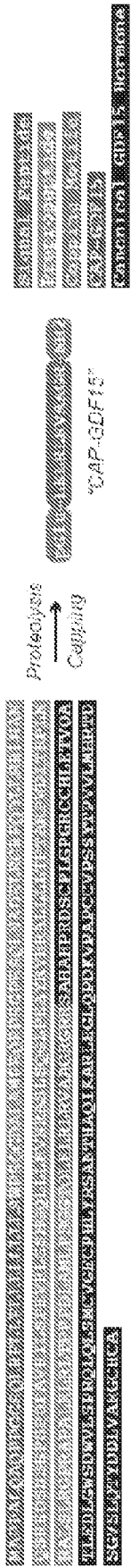


FIG. 5A

16/88

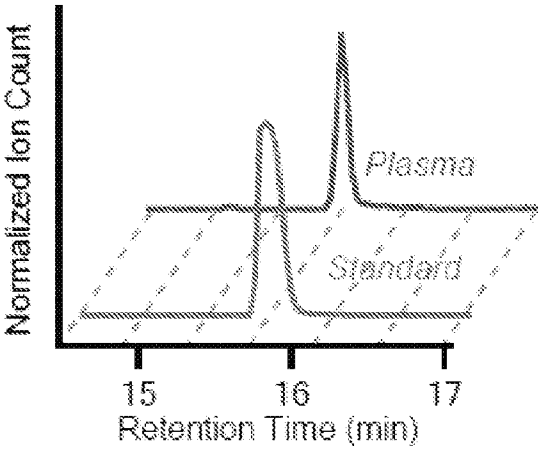


FIG. 5B

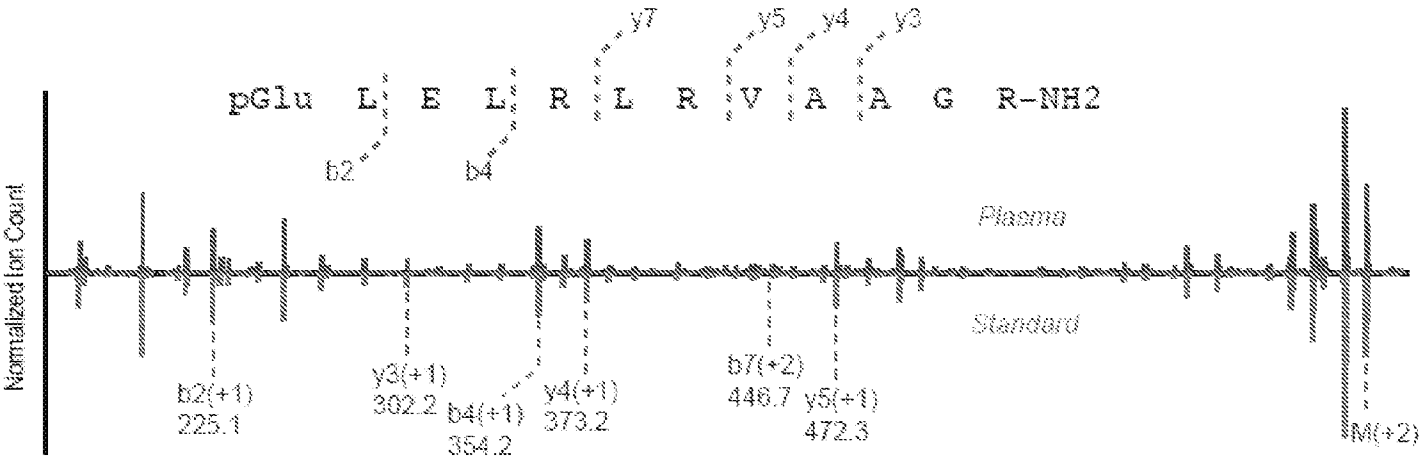


FIG. 5C

17/88

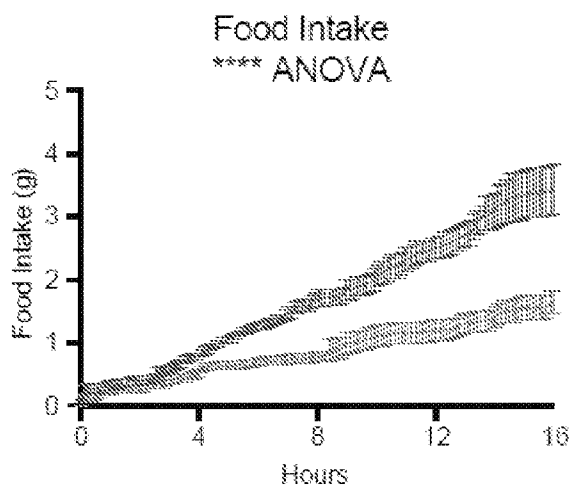


FIG. 5D

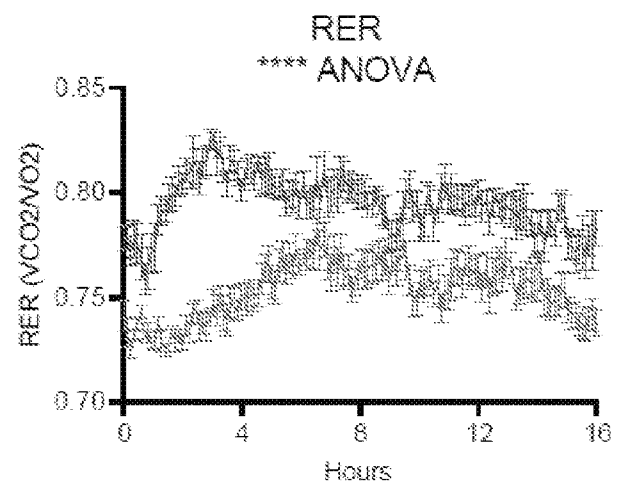


FIG. 5E

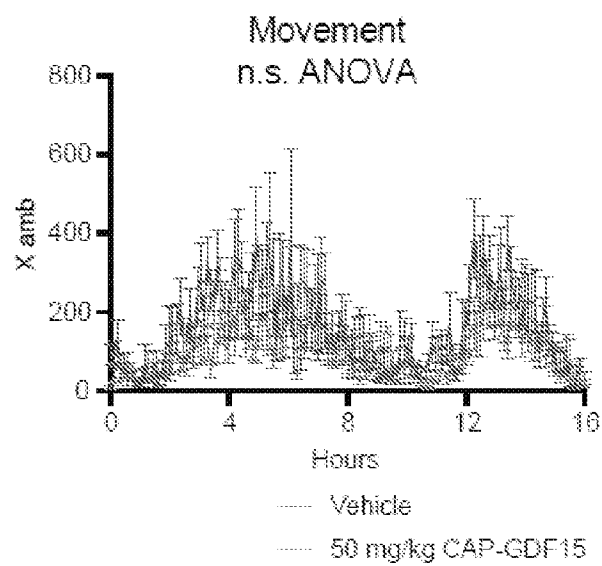


FIG. 5F

18/88

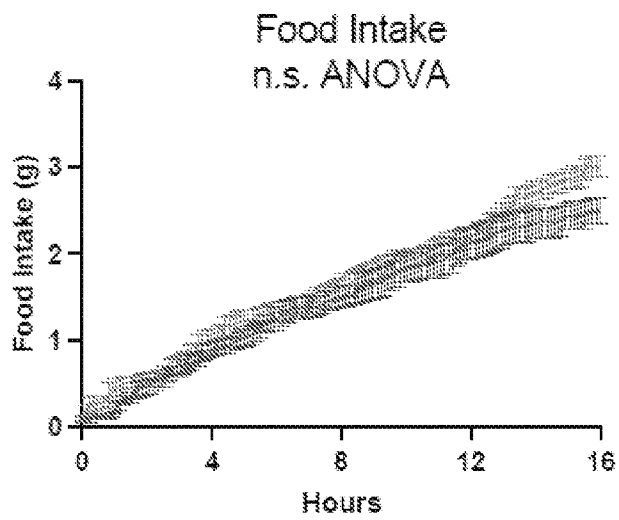


FIG. 5G

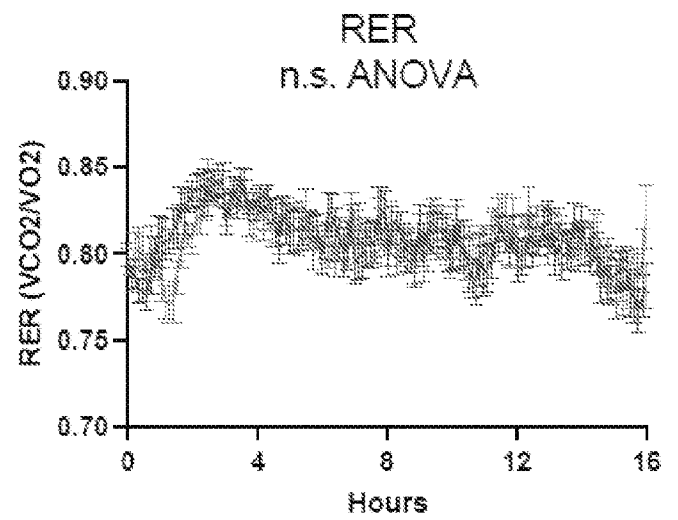


FIG. 5H

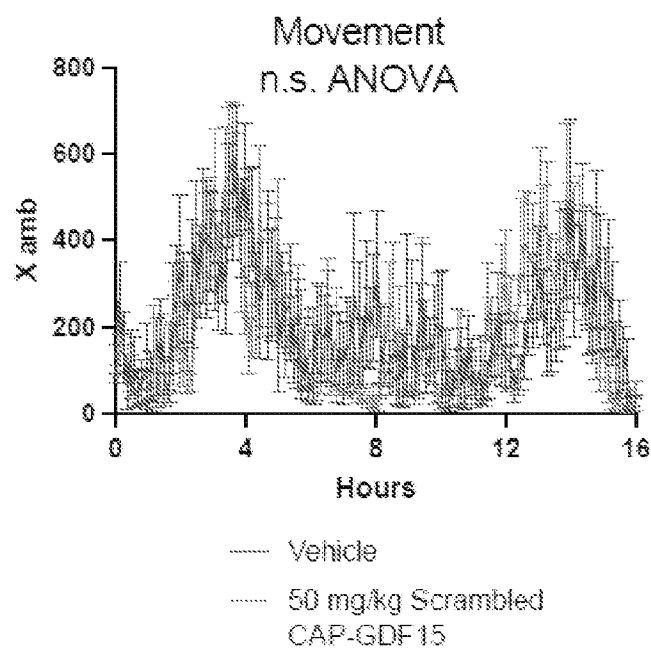


FIG. 5I

19/88

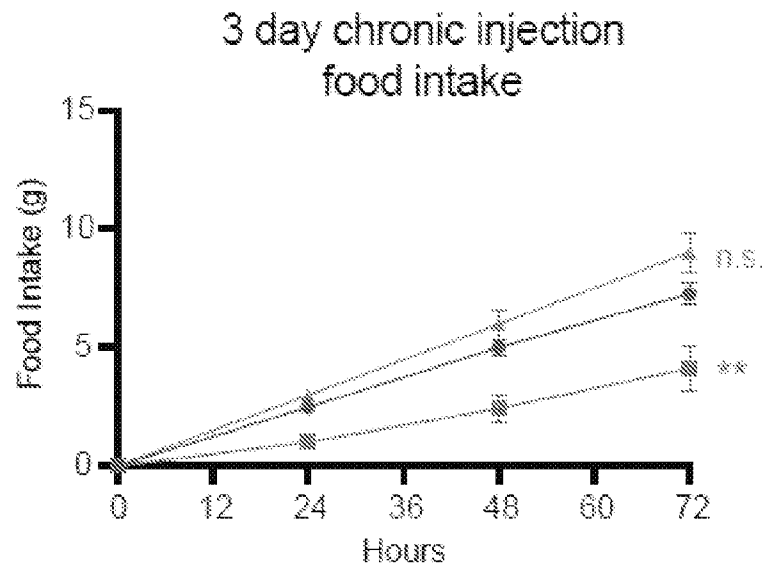


FIG. 5J

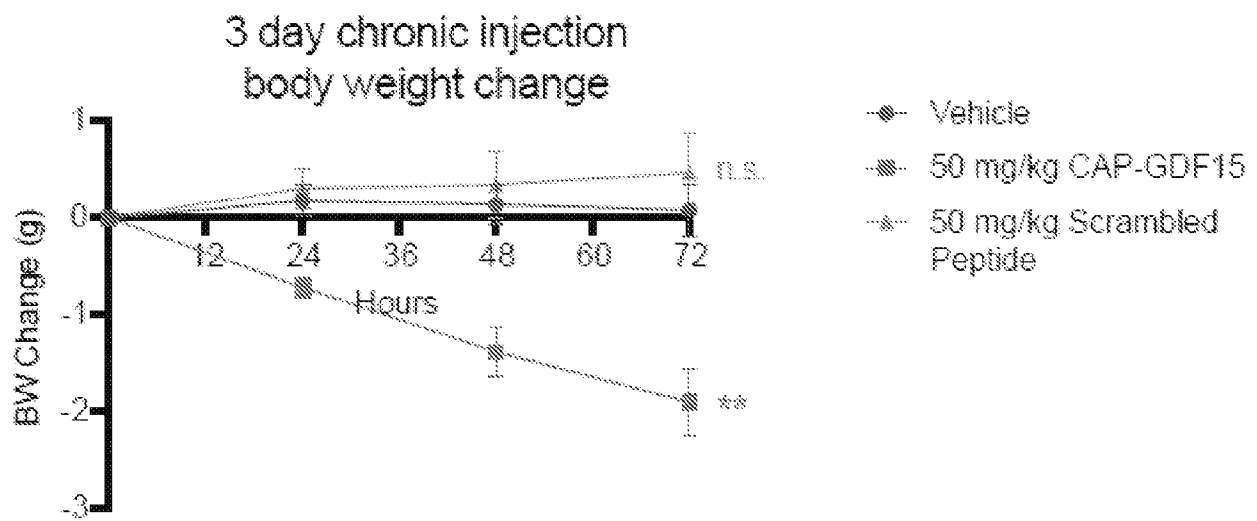


FIG. 5K

20/88

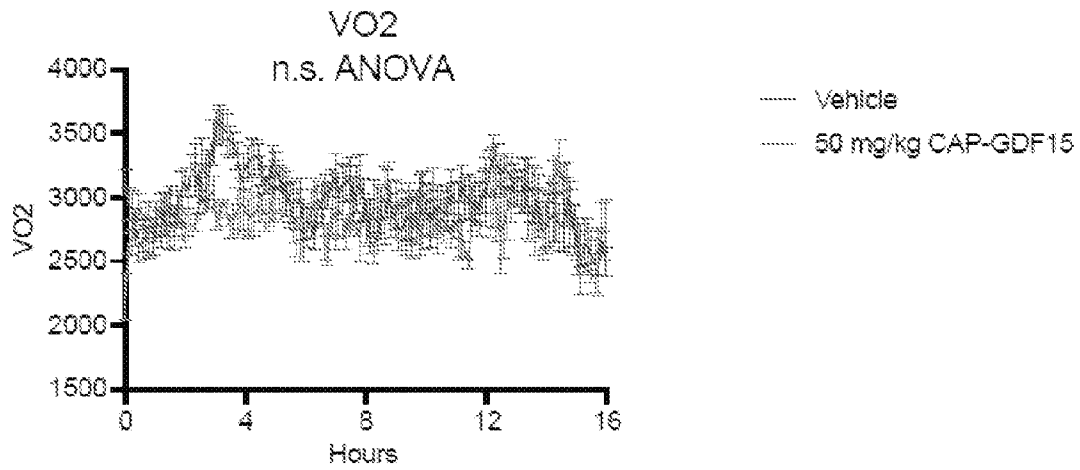


FIG. 6A

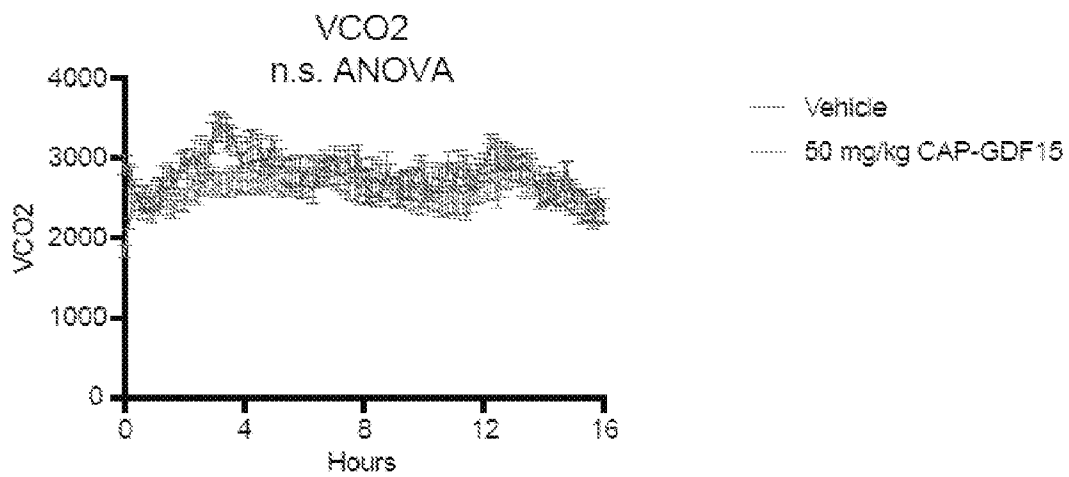


FIG. 6B

21/88

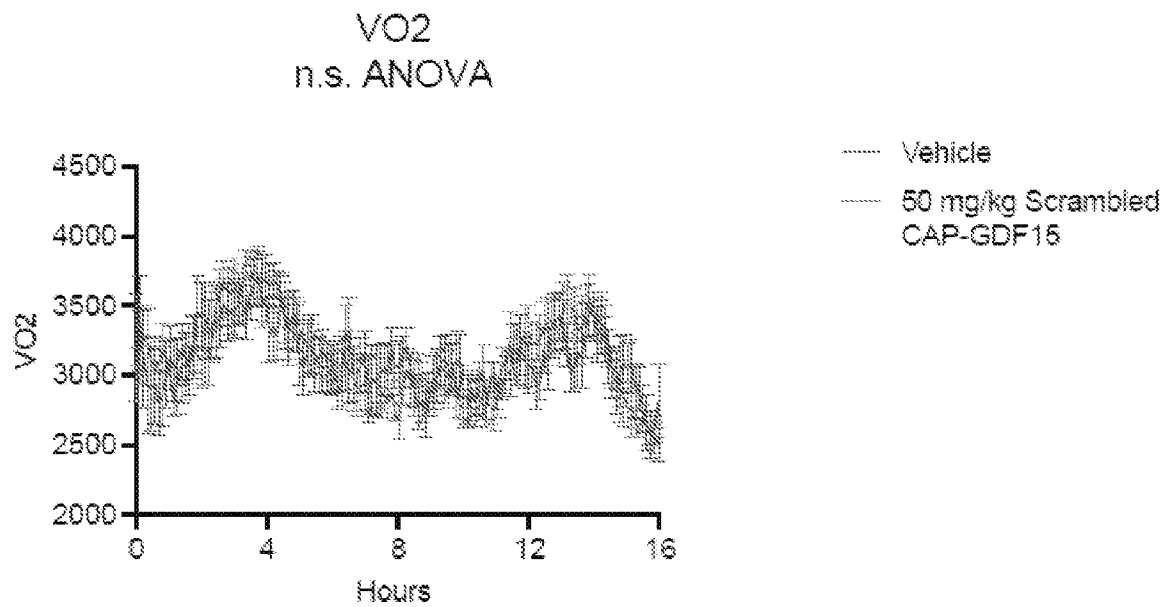


FIG. 6C

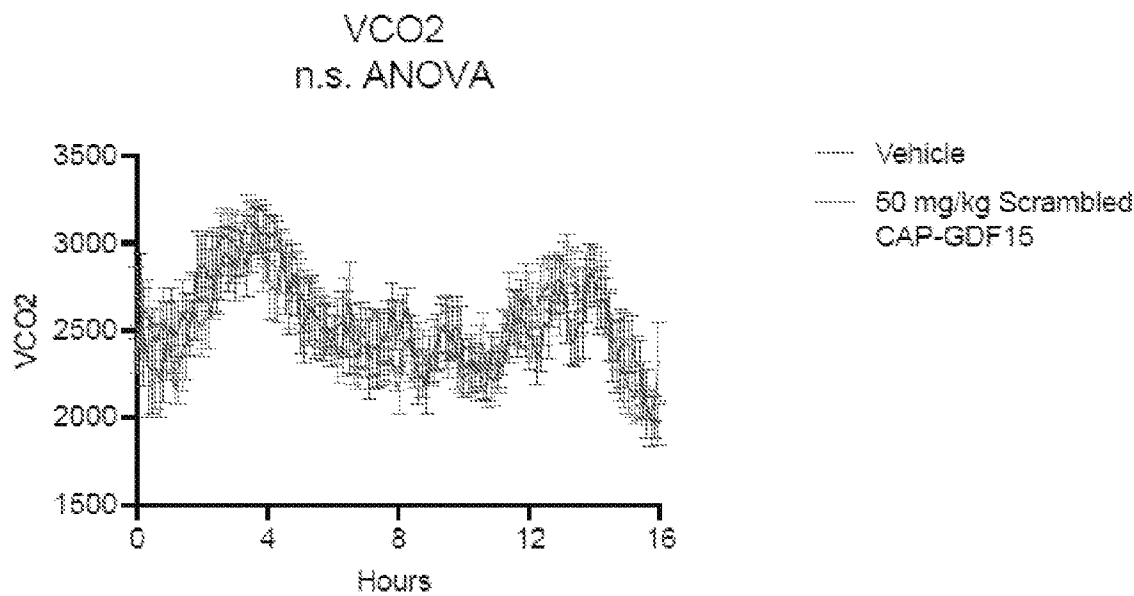


FIG. 6D

22/88

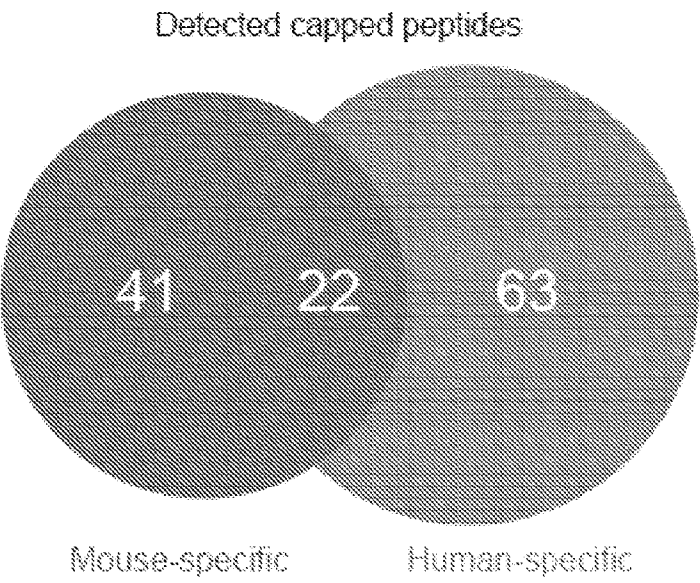
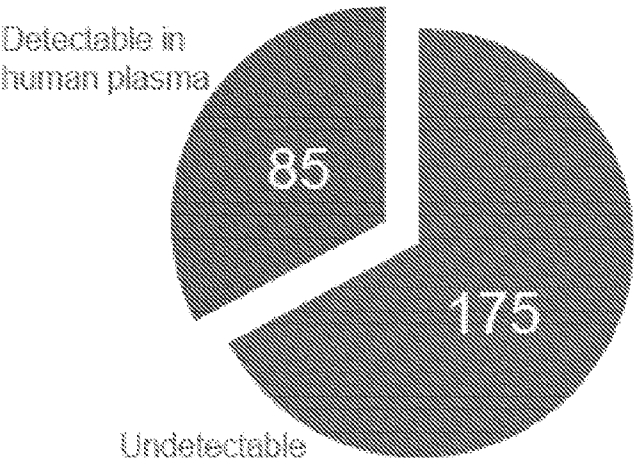


FIG. 7A

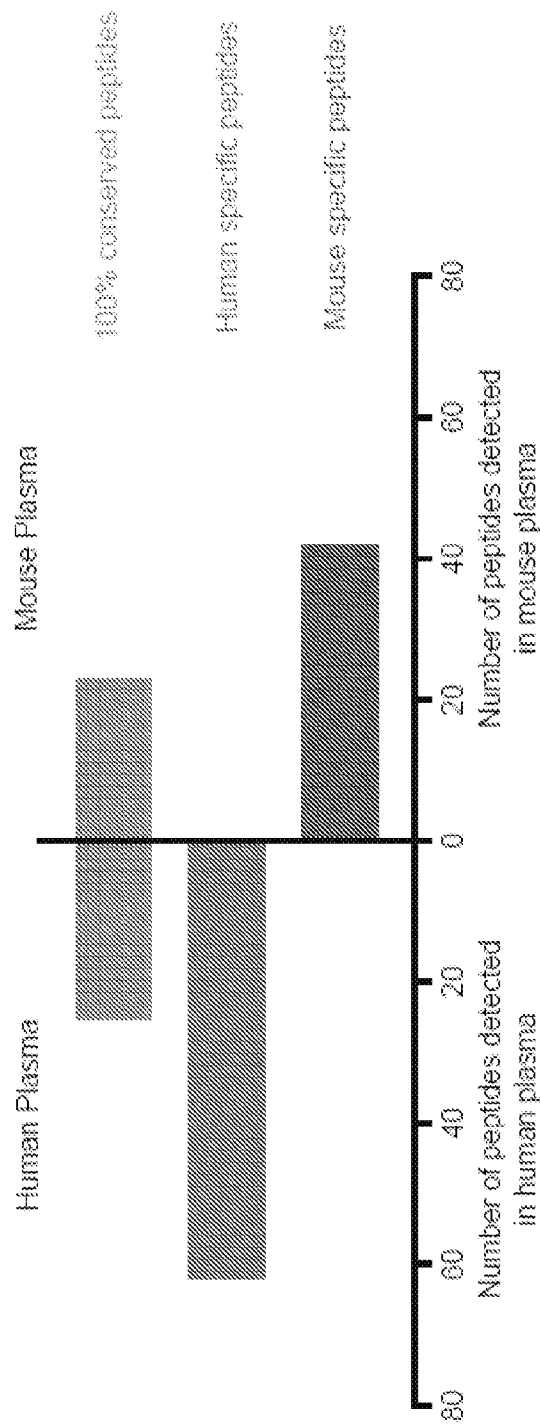


FIG. 7B

24/88

mouse CAP-NTN4
pGlu-HWKPAL-NH₂

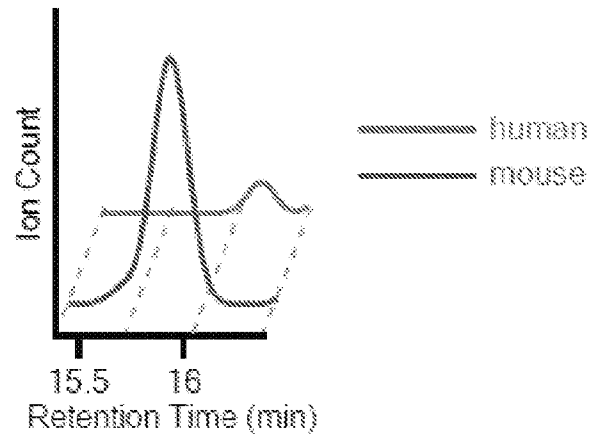


FIG. 7C

human CAP-LEFTY2
pGlu-YVLLRRSHGDRSR-NH₂

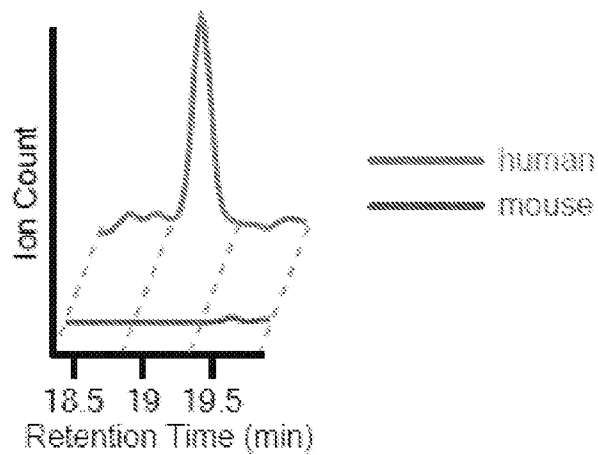


FIG. 7D

Alignment of all detectable mouse and human capped peptides

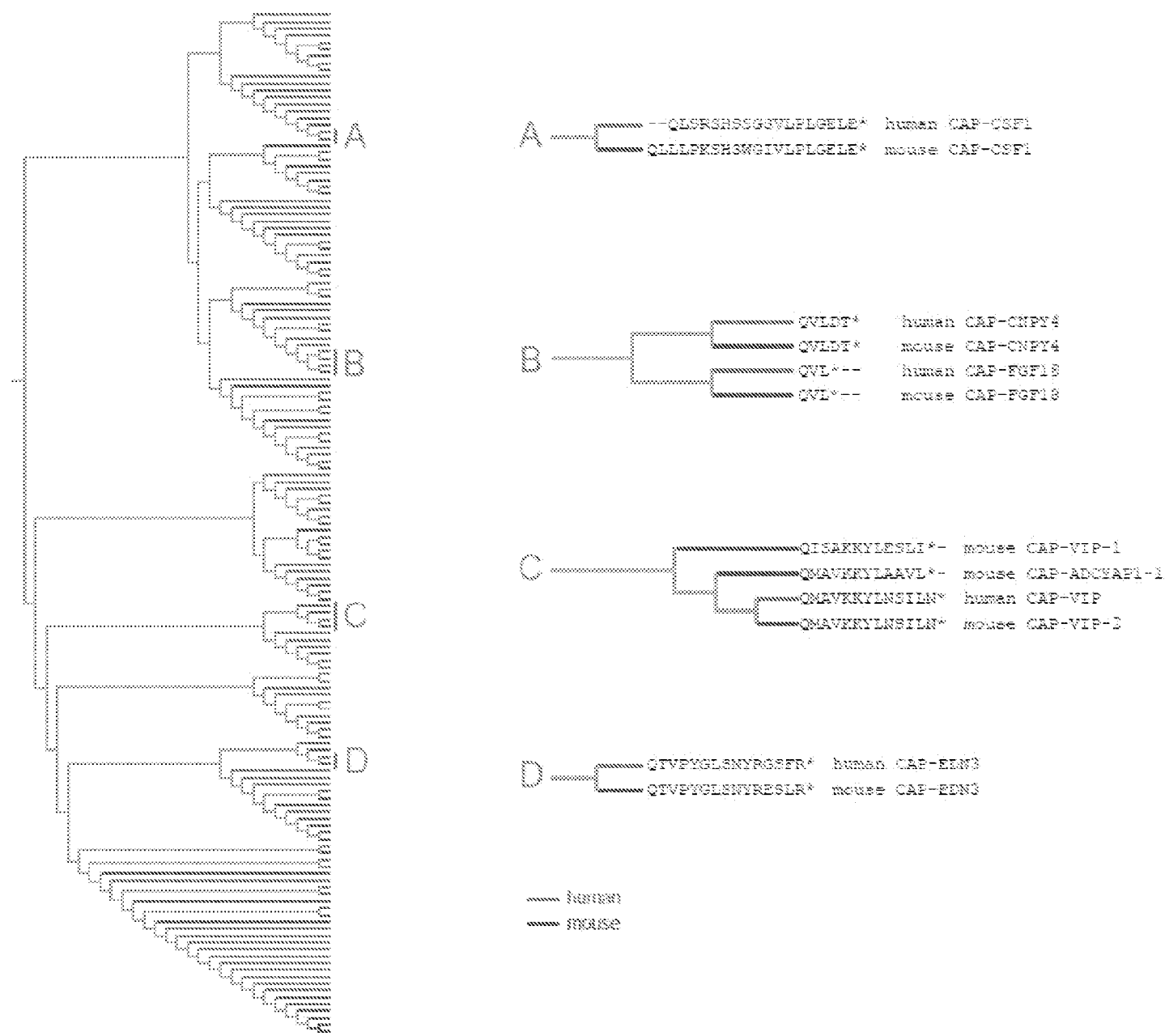


FIG. 7E

26/88

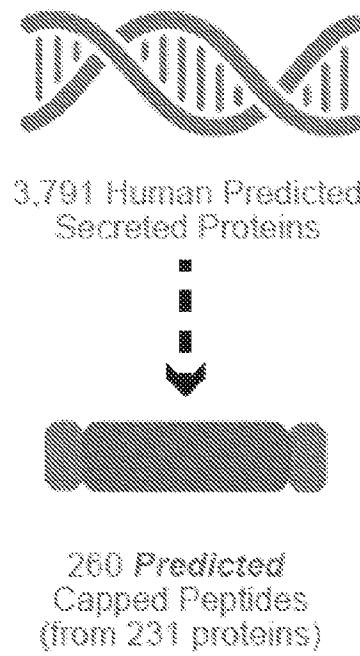


FIG. 8A

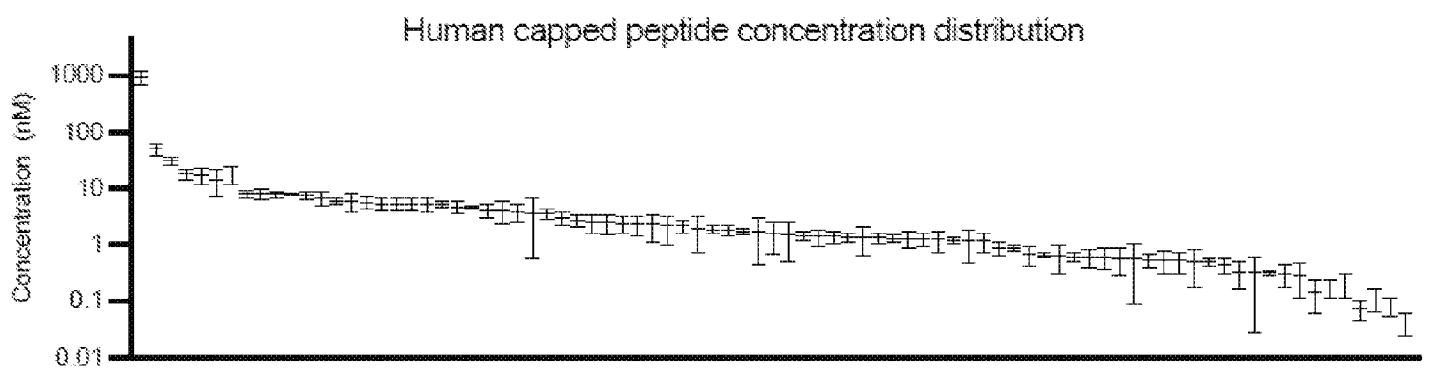


FIG. 8B

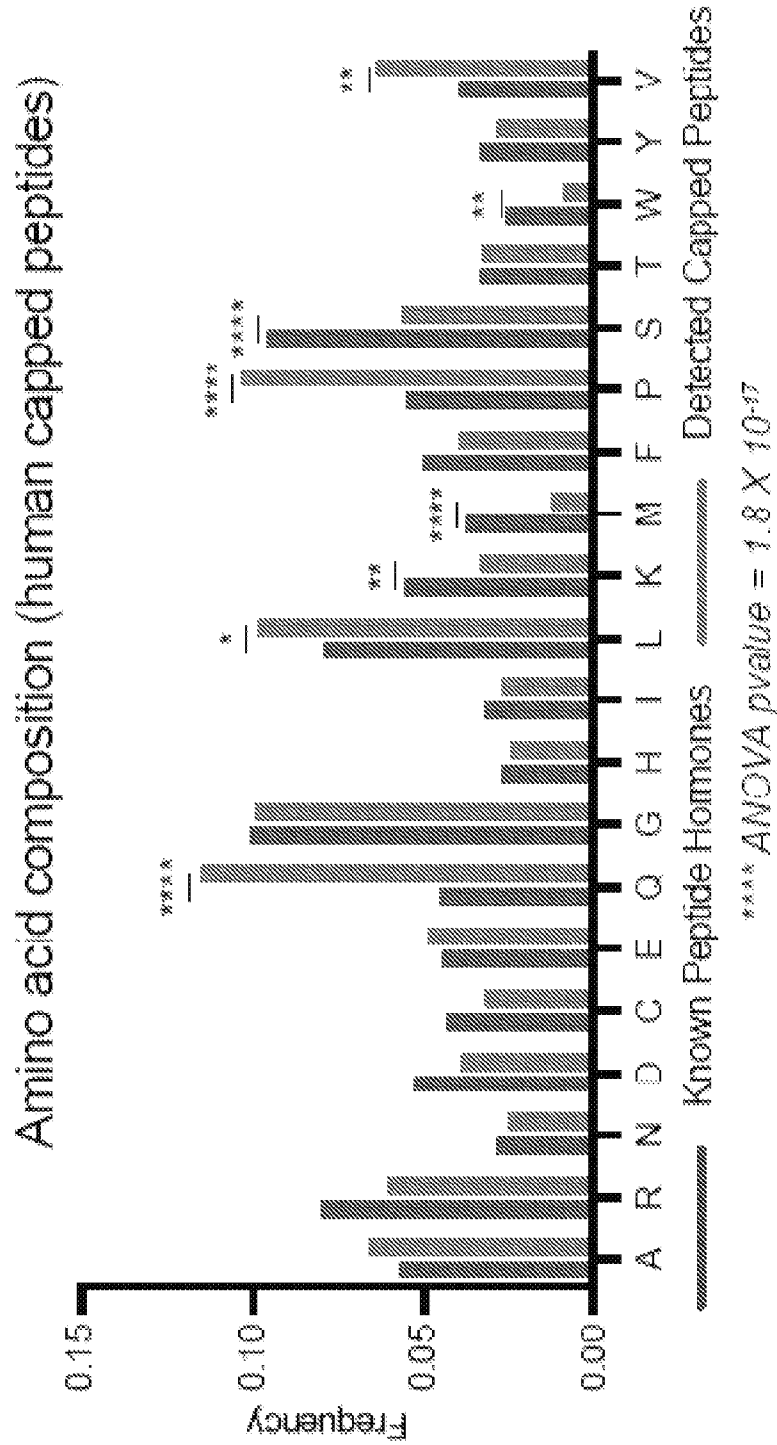


FIG. 8C

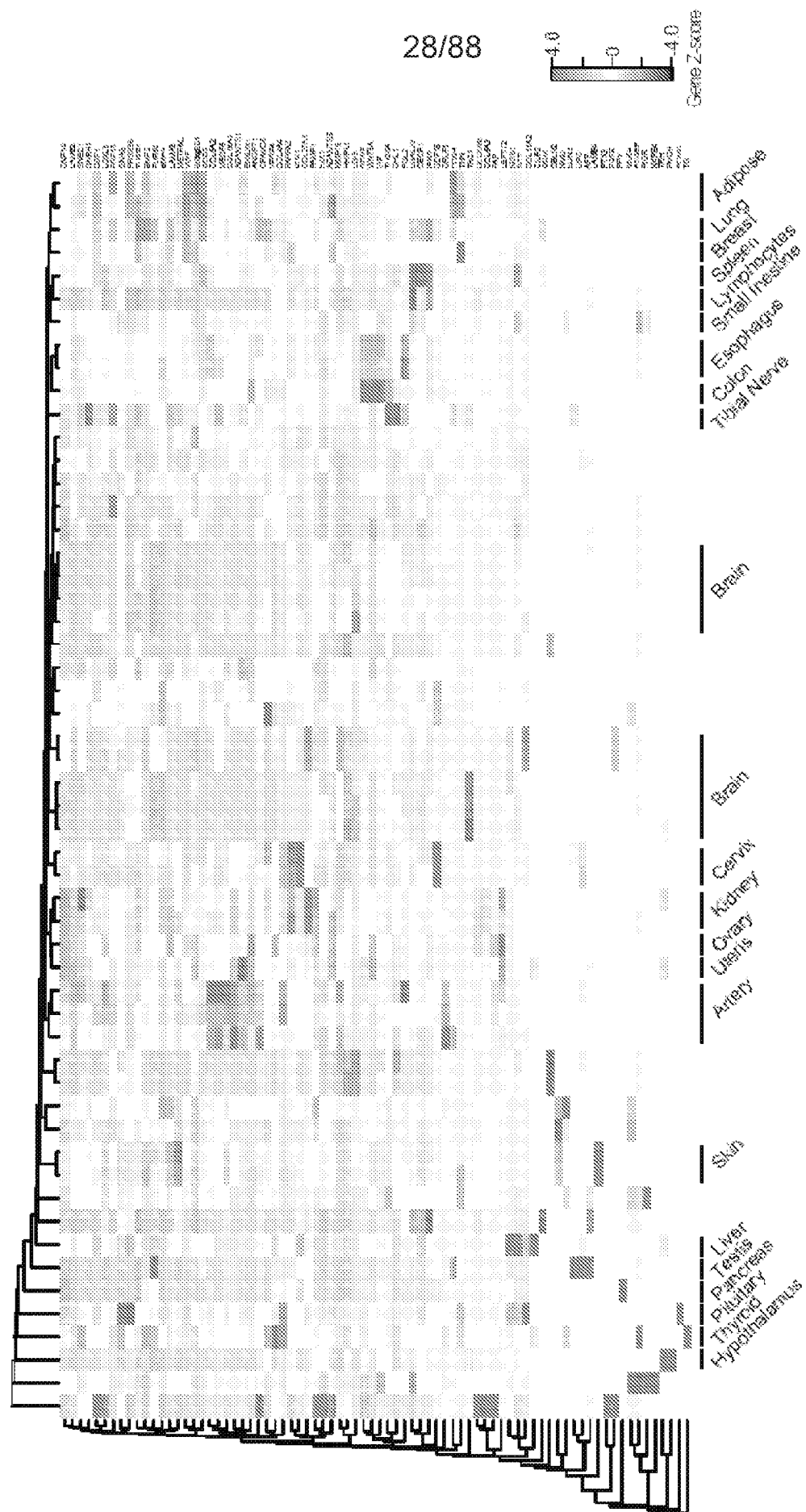


FIG. 9

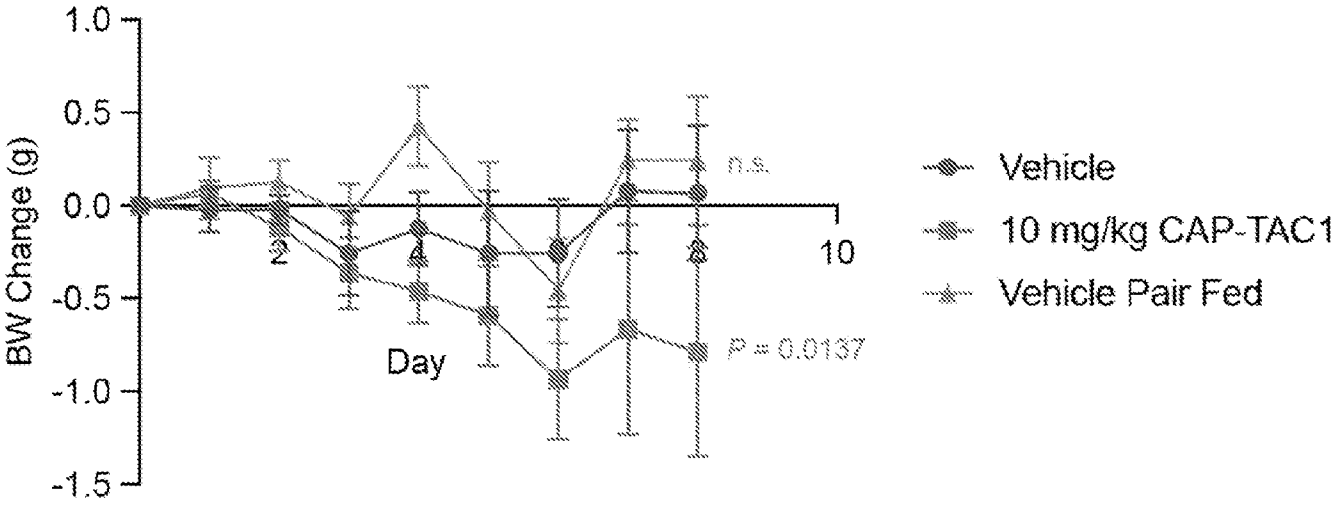


FIG. 10A

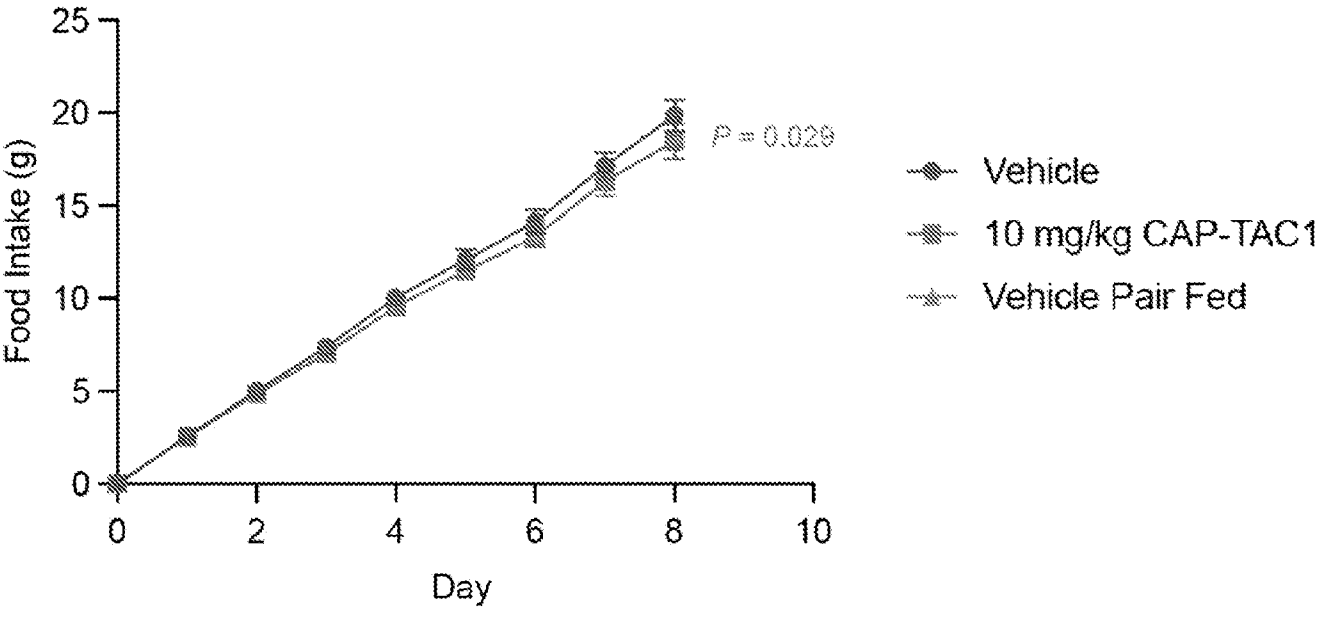


FIG. 10B

30/88

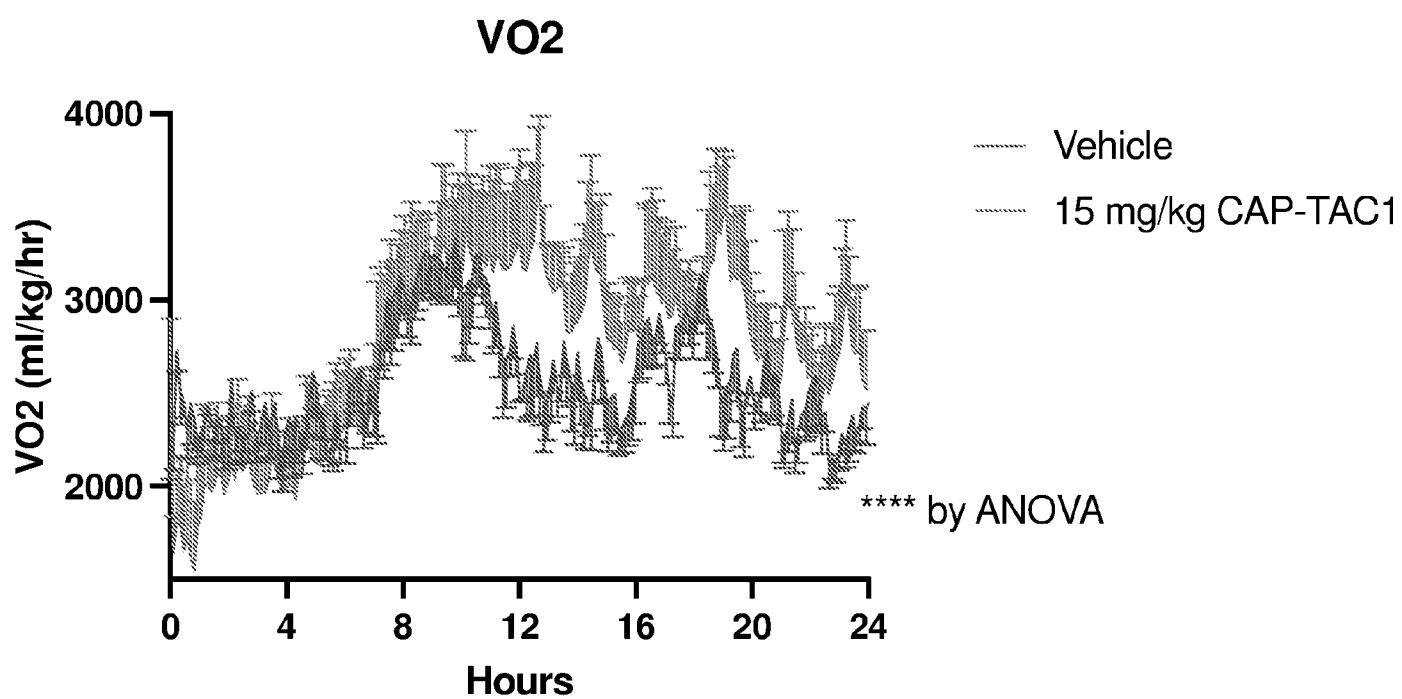


FIG. 11

31/88

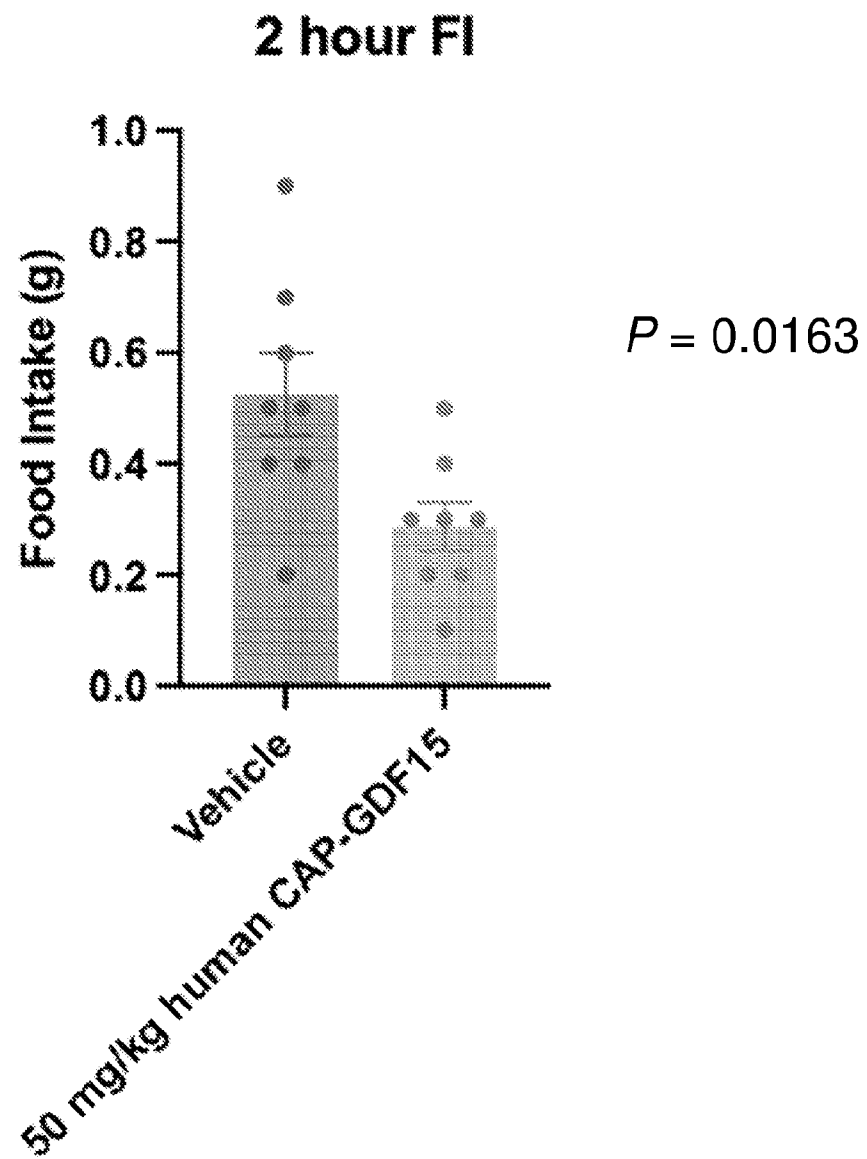


FIG. 12

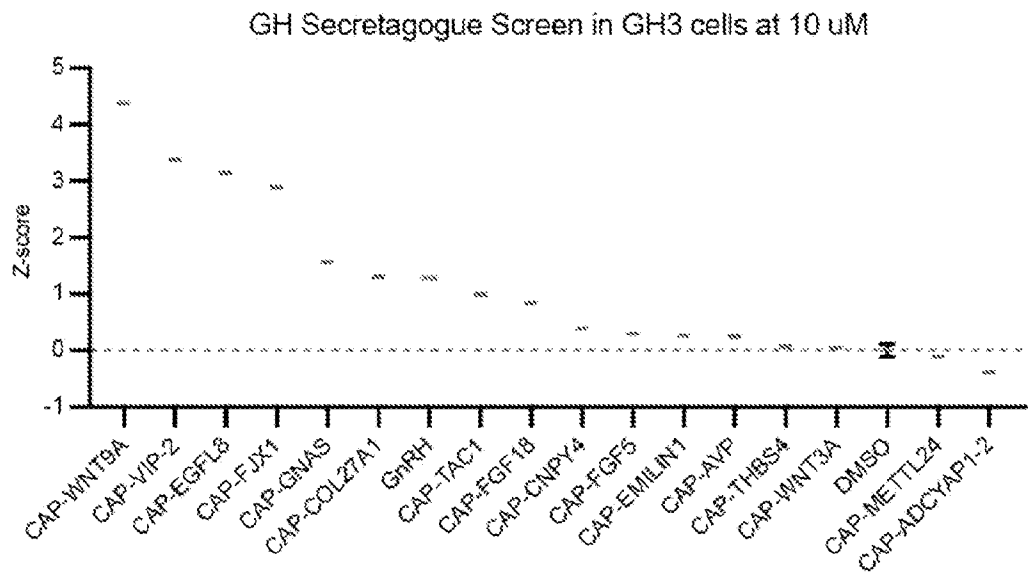


FIG. 13A

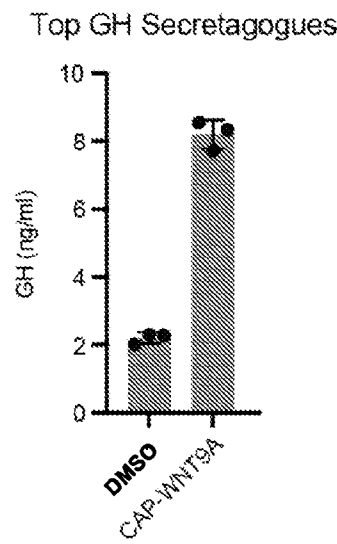


FIG. 13B

33/88

Fig. 14A

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
QATCDACAPGYT	Q05793	Basement membrane-specific heparan sulfate proteoglycan core protein OS=Mus musculus OX=10090 GN=Hspg2 PE=1 SV=1	835	HSPG2	1
QAFENGVRM	Q04592	Proprotein convertase subtilisin/kexin type 5 OS=Mus musculus OX=10090 GN=Pcsk5 PE=1 SV=3	291	PCSK5	2
QKCVWGHNKER	Q9ET66	Peptidase inhibitor 16 OS=Mus musculus OX=10090 GN=Pi16 PE=1 SV=2	73	PI16	3
QSTEVDRSVDRLTFGDLLP	Q70E20	Sushi nidogen and EGF-like domain-containing protein 1 OS=Mus musculus OX=10090 GN=Sned1 PE=2 SV=2	1050	SNED1	4
QNEASPTPRCMMHSRAMGR	Q99JR5	Tubulointerstitial nephritis antigen-like OS=Mus musculus OX=10090 GN=Tinagl1 PE=1 SV=1	298	TINAGL1	5
QNASR	Q2TJ95	R-spondin-3 OS=Mus musculus OX=10090 GN=Rspo3 PE=1 SV=2	21	RSPO3	6
QDCRVNNGGCLHYCLEESN	P33587	Vitamin K-dependent protein C OS=Mus musculus OX=10090 GN=Proc PE=1 SV=2	136	PROC	7
QSDRIKD	P33587	Vitamin K-dependent protein C OS=Mus musculus OX=10090 GN=Proc PE=1 SV=2	343	PROC	8
QGCVNTKGSYECVCP	Q6NZL8	Signal peptide CUB and EGF-like domain-containing protein 1 OS=Mus musculus OX=10090 GN=Scube1 PE=2 SV=2	404	SCUBE1	9
QFFGLM	P41539	Protachykinin-1 OS=Mus musculus OX=10090 GN=Tac1 PE=2 SV=1	62	TAC1	10
QISHKRHKTDSEFVGLM	P41539	Protachykinin-1 OS=Mus musculus OX=10090 GN=Tac1 PE=2 SV=1	91	TAC1	11
QFYGLM	Q99N14	Tachykinin-4 OS=Mus musculus OX=10090 GN=Tac4 PE=1 SV=1	60	TAC4	12
QGLV	Q08535	Secretin OS=Mus musculus OX=10090 GN=Sct PE=2 SV=1	54	SCT	13
QGPPGPAGEE	P11087	Collagen alpha-1(I) chain OS=Mus musculus OX=10090 GN=Col1a1 PE=1 SV=4	445	COL1A1	14
QVTVA	Q8K1S1	Leucine-rich repeat LGI family member 4 OS=Mus musculus OX=10090 GN=Lgi4 PE=1 SV=1	506	LGI4	15

34/88

Fig. 14A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
QHSHASRSR	Q64280	Left-right determination factor 1 OS=Mus musculus OX=10090 GN=Lefty1 PE=2 SV=1	65	LEFTY1	16
QGSHADRSR	P57785	Left-right determination factor 2 OS=Mus musculus OX=10090 GN=Lefty2 PE=2 SV=2	65	LEFTY2	17
QLYSGGKVGVLVYAG	Q9R001	A disintegrin and metalloproteinase with thrombospondin motifs 5 OS=Mus musculus OX=10090 GN=Adamts5 PE=2 SV=2	79	ADAMTS5	18
QFAYRHCNNPAPRNSGRYCT	Q9R001	A disintegrin and metalloproteinase with thrombospondin motifs 5 OS=Mus musculus OX=10090 GN=Adamts5 PE=2 SV=2	587	ADAMTS5	19
QDC	O35453	Serine protease hepsin OS=Mus musculus OX=10090 GN=Hpn PE=1 SV=3	169	HPN	20
QGNPSRSHGRG	Q9CXD9	Leucine-rich repeat-containing protein 17 OS=Mus musculus OX=10090 GN=Lrrc17 PE=2 SV=1	28	LRRC17	21
QCVSEGLSFHCLCPE	O89029	Matrilin-4 OS=Mus musculus OX=10090 GN=Matn4 PE=1 SV=1	353	MATN4	22
QKRHFC	Q4VC17	A disintegrin and metalloproteinase with thrombospondin motifs 18 OS=Mus musculus OX=10090 GN=Adamts18 PE=2 SV=2	248	ADAMTS18	23
QSALVGKN	Q4VC17	A disintegrin and metalloproteinase with thrombospondin motifs 18 OS=Mus musculus OX=10090 GN=Adamts18 PE=2 SV=2	371	ADAMTS18	24
QDPSG	Q62059	Versican core protein OS=Mus musculus OX=10090 GN=Vcan PE=1 SV=2	2335	VCAN	25
QASHH	Q6PAL1	Draxin OS=Mus musculus OX=10090 GN=Draxin PE=1 SV=2	52	DRAXIN	26
QFEENWYNTYSSNIYKHGDT	Q9ESL9	Fibroblast growth factor 20 OS=Mus musculus OX=10090 GN=Fgf20 PE=2 SV=2	141	FGF20	27
QRKCIVEILVEPKDI	P23188	Furin OS=Mus musculus OX=10090 GN=Furin PE=1 SV=2	446	FURIN	28

35/88

Fig. 14A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
QVEEVRLRPAVEW	Q9Z175	Lysyl oxidase homolog 3 OS=Mus musculus OX=10090 GN=Loxl3 PE=1 SV=2	165	LOXL3	29
QVYT	Q07563	Collagen alpha-1(XVII) chain OS=Mus musculus OX=10090 GN=Col17a1 PE=1 SV=3	1455	COL17A1	30
QGNDDPGPVGPA	Q5QNNQ9	Collagen alpha-1(XXVII) chain OS=Mus musculus OX=10090 GN=Col27a1 PE=2 SV=1	1443	COL27A1	31
QLGPP	Q5QNNQ9	Collagen alpha-1(XXVII) chain OS=Mus musculus OX=10090 GN=Col27a1 PE=2 SV=1	1484	COL27A1	32
QLLLPKSHSWGIVLPLGELE	P07141	Macrophage colony-stimulating factor 1 OS=Mus musculus OX=10090 GN=Csf1 PE=1 SV=2	418	CSF1	33
QRRDLFLAKHILHCKAKY	Q8BQB4	Four-jointed box protein 1 OS=Mus musculus OX=10090 GN=Fjx1 PE=1 SV=1	413	FJX1	34
QVVEK	P26928	Hepatocyte growth factor-like protein OS=Mus musculus OX=10090 GN=Mst1 PE=2 SV=2	470	MST1	35
QIDLLAPTHLDNGVDKTSR	Q9Z1P8	Angiopoietin-related protein 4 OS=Mus musculus OX=10090 GN=Angptl4 PE=1 SV=1	146	ANGPTL4	36
QED	Q80V70	Multiple epidermal growth factor-like domains protein 6 OS=Mus musculus OX=10090 GN=Megf6 PE=2 SV=3	237	MEGF6	37
QKWYNR	Q8BLU0	Leucine-rich repeat transmembrane protein FLRT2 OS=Mus musculus OX=10090 GN=Flrt2 PE=1 SV=1	574	FLRT2	38
QWSPS	P15656	Fibroblast growth factor 5 OS=Mus musculus OX=10090 GN=Fgf5 PE=2 SV=1	77	FGF5	39
QFLDKGSENSFIDL	Q71KU9	Fibrinogen-like protein 1 OS=Mus musculus OX=10090 GN=Fgl1 PE=1 SV=2	62	FGL1	40
QVTYVRS	Q8CCB5	Probable methyltransferase-like protein 24 OS=Mus musculus OX=10090 GN=Mettl24 PE=2 SV=1	67	METTL24	41
QHWSYGLRPG	P13562	Progonadoliberin-1 OS=Mus musculus OX=10090 GN=Gnrh1 PE=3 SV=1	21	GNRH1	42

36/88

Fig. 14A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
QHWKPAL	Q9JI33	Netrin-4 OS=Mus musculus OX=10090 GN=Ntn4 PE=1 SV=2	608	NTN4	43
QCRCRPNVA	Q61001	Laminin subunit alpha-5 OS=Mus musculus OX=10090 GN=Lama5 PE=1 SV=4	1553	LAMA5	44
QFEENWYNTYSSNLYKHVDT	P54130	Fibroblast growth factor 9 OS=Mus musculus OX=10090 GN=Fgf9 PE=2 SV=2	138	FGF9	45
QREKDLSTYNWNSFGLRY	Q6Y4S4	Metastasis-suppressor KiSS-1 OS=Mus musculus OX=10090 GN=Kiss1 PE=1 SV=2	101	KISS1	46
QTPEILAALPESAPAVRPTC	Q9R098	Hepatocyte growth factor activator OS=Mus musculus OX=10090 GN=Hgfac PE=1 SV=1	372	HGFAC	47
QGSLSFPLLER	O55233	Cerberus OS=Mus musculus OX=10090 GN=Cer1 PE=1 SV=1	26	CER1	48
QGPIGPPGPSGV	Q05306	Collagen alpha-1(X) chain OS=Mus musculus OX=10090 GN=Col10a1 PE=2 SV=1	207	COL10A1	49
QSLDPSHRISDRDYMGMWMD	P09240	Cholecystokinin OS=Mus musculus OX=10090 GN=Cck PE=1 SV=3	83	CCK	50
QDAALGW	P50715	Alpha-defensin-related sequence 7 OS=Mus musculus OX=10090 GN=Defa-rs7 PE=2 SV=2	54	DEFA-RS7	51
QPR	O08644	Ephrin type-B receptor 6 OS=Mus musculus OX=10090 GN=Ephb6 PE=2 SV=4	682	EPHB6	52
QDGRAECKCPAGEDWWYM	Q61847	Meprin A subunit beta OS=Mus musculus OX=10090 GN=Mep1b PE=1 SV=2	624	MEP1B	53
QRCLPEFENAAF	Q9R0B6	Laminin subunit gamma-3 OS=Mus musculus OX=10090 GN=Lamc3 PE=1 SV=2	41	LAMC3	54
QRCDDADP	Q9R0B6	Laminin subunit gamma-3 OS=Mus musculus OX=10090 GN=Lamc3 PE=1 SV=2	83	LAMC3	55
QPGKAS	Q9CWP4	Lymphocyte antigen 6K OS=Mus musculus OX=10090 GN=Ly6k PE=1 SV=1	120	LY6K	56
QGP PGPTGKP	Q61245	Collagen alpha-1(XI) chain OS=Mus musculus OX=10090 GN=Col11a1 PE=1 SV=2	564	COL11A1	57

37/88

Fig. 14A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
QPGPPGPSGEAGPPGPP	Q61245	Collagen alpha-1(XI) chain OS=Mus musculus OX=10090 GN=Col11a1 PE=1 SV=2	1340	COL11A1	58
QTR	Q8K209	Adhesion G-protein coupled receptor G1 OS=Mus musculus OX=10090 GN=Adgrg1 PE=1 SV=1	275	ADGRG1	59
QGPRGPLLSMP	Q8BGU2	Cerebellin-2 OS=Mus musculus OX=10090 GN=Cbln2 PE=1 SV=1	5	CBLN2	60
QMEPGVGILTSL	P56386	Beta-defensin 1 OS=Mus musculus OX=10090 GN=Defb1 PE=2 SV=1	16	DEFB1	61
QRPRMEEEEEAYGWMDF	P48757	Gastrin OS=Mus musculus OX=10090 GN=Gast PE=2 SV=2	75	GAST	62
QHEEF	P59509	A disintegrin and metalloproteinase with thrombospondin motifs 19 OS=Mus musculus OX=10090 GN=Adamts19 PE=2 SV=1	406	ADAMTS19	63
QVL	O89101	Fibroblast growth factor 18 OS=Mus musculus OX=10090 GN=Fgf18 PE=2 SV=1	66	FGF18	64
QSGVYDLRL	Q8R0Z6	Angiopoietin-related protein 6 OS=Mus musculus OX=10090 GN=Angptl6 PE=2 SV=1	256	ANGPTL6	65
QPLG	Q99N23	Carbonic anhydrase 15 OS=Mus musculus OX=10090 GN=Ca15 PE=1 SV=1	281	CA15	66
QFSSRDCTRPVPRNGGKYCE	Q8BNJ2	A disintegrin and metalloproteinase with thrombospondin motifs 4 OS=Mus musculus OX=10090 GN=Adamts4 PE=2 SV=2	536	ADAMTS4	67
QALY	Q9JHI0	Matrix metalloproteinase-19 OS=Mus musculus OX=10090 GN=Mmp19 PE=2 SV=1	252	MMP19	68
QKLY	Q10738	Matrilysin OS=Mus musculus OX=10090 GN=Mmp7 PE=2 SV=1	254	MMP7	69
QYSSCTSD	P41245	Matrix metalloproteinase-9 OS=Mus musculus OX=10090 GN=Mmp9 PE=1 SV=2	356	MMP9	70
QLLDATET	Q80XH2	Interphotoreceptor matrix proteoglycan 2 OS=Mus musculus OX=10090 GN=Impg2 PE=1 SV=1	56	IMPG2	71

38/88

Fig. 14A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
QLL	Q80XH2	Interphotoreceptor matrix proteoglycan 2 OS=Mus musculus OX=10090 GN=Impg2 PE=1 SV=1	243	IMPG2	72
QCSRCATGYYGFPYCKPCNC	Q61789	Laminin subunit alpha-3 OS=Mus musculus OX=10090 GN=Lama3 PE=1 SV=4	1291	LAMA3	73
QDEGILPS	Q9JM58	Cytokine receptor-like factor 1 OS=Mus musculus OX=10090 GN=Crlf1 PE=1 SV=1	406	CRLF1	74
QAAKEFIAWLVKGR	P55095	Pro-glucagon OS=Mus musculus OX=10090 GN=Gcg PE=1 SV=1	113	GCG	75
QPKFSFC	O88188	Lymphocyte antigen 86 OS=Mus musculus OX=10090 GN=Ly86 PE=1 SV=1	105	LY86	76
QAAAASPENSRGK	P48540	Glial cell line-derived neurotrophic factor OS=Mus musculus OX=10090 GN=Gdnf PE=1 SV=1	93	GDNF	77
QFGEMIRLKT	P31482	Phospholipase A2 membrane associated OS=Mus musculus OX=10090 GN=Pla2g2a PE=1 SV=3	24	PLA2G2A	78
QGPVSS	Q8K4G1	Latent-transforming growth factor beta-binding protein 4 OS=Mus musculus OX=10090 GN=Ltbp4 PE=1 SV=2	1214	LTBP4	79
QEKREIGDEENSAKFPI	P56942	Pro-MCH OS=Mus musculus OX=10090 GN=Pmch PE=1 SV=2	126	PMCH	80
QKGEKGMEGEKGSRGDVGSE	Q4ZJM7	Otolin-1 OS=Mus musculus OX=10090 GN=Otol1 PE=1 SV=2	259	OTOL1	81
QKSVWHGSDPS	P39061	Collagen alpha-1(XVIII) chain OS=Mus musculus OX=10090 GN=Col18a1 PE=1 SV=4	1705	COL18A1	82
QPAVSWSKD	A2AAJ9	Obscurin OS=Mus musculus OX=10090 GN=Obscn PE=1 SV=3	136	OBSCN	83
QLVTE	A2AAJ9	Obscurin OS=Mus musculus OX=10090 GN=Obscn PE=1 SV=3	273	OBSCN	84
QVSA	A2AAJ9	Obscurin OS=Mus musculus OX=10090 GN=Obscn PE=1 SV=3	752	OBSCN	85
QLRP	A2AAJ9	Obscurin OS=Mus musculus OX=10090 GN=Obscn PE=1 SV=3	2279	OBSCN	86
QYPEVGM	Q9R0E1	Multifunctional procollagen lysine hydroxylase and glycosyltransferase LH3 OS=Mus musculus OX=10090 GN=Plod3 PE=1 SV=1	155	PLOD3	87

39/88

Fig. 14A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
QSRPGCYGDRSSLPGVRSY	P55066	Neurocan core protein OS=Mus musculus OX=10090 GN=Ncan PE=1 SV=1	220	NCAN	88
QSPGSSAAPGP	Q5RJI4	Extracellular tyrosine-protein kinase PKDCC OS=Mus musculus OX=10090 GN=Pkdcc PE=1 SV=2	36	PKDCC	89
QMAVKKYLA AVL	O70176	Pituitary adenylate cyclase-activating polypeptide OS=Mus musculus OX=10090 GN=Adcyap1 PE=2 SV=1	145	ADCYA P1	90
QRVKNK	O70176	Pituitary adenylate cyclase-activating polypeptide OS=Mus musculus OX=10090 GN=Adcyap1 PE=2 SV=1	162	ADCYA P1	91
QGPRDTP	Q8K262	Placenta-specific protein 9 OS=Mus musculus OX=10090 GN=Plac9 PE=2 SV=1	22	PLAC9	92
QFAEMKLRPSNFRNLE	Q80YX1	Tenascin OS=Mus musculus OX=10090 GN=Tnc PE=1 SV=1	2088	TNC	93
QERRRAKRRKRKKEVPFW	Q9QXU7	Prokineticin-2 OS=Mus musculus OX=10090 GN=Prok2 PE=2 SV=1	80	PROK2	94
QRRR	O54907	Tumor necrosis factor ligand superfamily member 12 OS=Mus musculus OX=10090 GN=Tnfsf12 PE=2 SV=2	6	TNFSF 12	95
QLLPGMILCAGYPE	Q5K2P8	Polyserase-2 OS=Mus musculus OX=10090 GN=Prss36 PE=2 SV=1	220	PRSS3 6	96
QASPPRRL	Q8C5G2	TubulinyI-Tyr carboxypeptidase 2 OS=Mus musculus OX=10090 GN=Vash2 PE=1 SV=2	318	VASH2	97
QVEA	Q0VGU4	Neurosecretory protein VGF OS=Mus musculus OX=10090 GN=Vgf PE=1 SV=1	305	VGF	98
QGCGVVSSGSSLYFSKA	Q60841	Reelin OS=Mus musculus OX=10090 GN=Reln PE=1 SV=3	1096	RELN	99
QVGTGCGTLNDGRSLYFNGL	Q60841	Reelin OS=Mus musculus OX=10090 GN=Reln PE=1 SV=3	1469	RELN	100
QNGNA	O09164	Extracellular superoxide dismutase [Cu-Zn] OS=Mus musculus OX=10090 GN=Sod3 PE=1 SV=1	210	SOD3	101
QCPDGFV	P11214	Tissue-type plasminogen activator OS=Mus musculus OX=10090 GN=Plat PE=1 SV=3	105	PLAT	102

40/88

Fig. 14A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
QNCPRG	P35455	Vasopressin-neurophysin 2-copeptin OS=Mus musculus OX=10090 GN=Avp PE=2 SV=1	26	AVP	103
QEPMKKCT	Q9JM99	Proteoglycan 4 OS=Mus musculus OX=10090 GN=Prg4 PE=1 SV=2	923	PRG4	104
QDSLIFYSSNEFEGYP	P97321	Prolyl endopeptidase FAP OS=Mus musculus OX=10090 GN=Fap PE=1 SV=1	404	FAP	105
QSLRVSCYDALKHW	Q99NH8	Triggering receptor expressed on myeloid cells 2 OS=Mus musculus OX=10090 GN=Trem2 PE=1 SV=1	29	TREM2	106
QFR	P27467	Protein Wnt-3a OS=Mus musculus OX=10090 GN=Wnt3a PE=1 SV=1	79	WNT3A	107
QLRRYINTLTRPY	P10601	Pancreatic prohormone OS=Mus musculus OX=10090 GN=Ppy PE=3 SV=1	51	PPY	108
QVEVL	Q62522	Zona pellucida-binding protein 1 OS=Mus musculus OX=10090 GN=Zbp1 PE=1 SV=2	271	ZPBP	109
QESDASNFLKRR	Q14BU0	Unique cartilage matrix-associated protein OS=Mus musculus OX=10090 GN=Ucma PE=1 SV=1	49	UCMA	110
QSSDCPCVHA	Q8CIZ8	von Willebrand factor OS=Mus musculus OX=10090 GN=Vwf PE=1 SV=2	343	VWF	111
QDCPVCEVA	Q8CIZ8	von Willebrand factor OS=Mus musculus OX=10090 GN=Vwf PE=1 SV=2	1193	VWF	112
QWGGCGDNLKYSSKFVKEFL	Q8R5M2	Protein Wnt-9a OS=Mus musculus OX=10090 GN=Wnt9a PE=2 SV=1	163	WNT9A	113
QGDPNCKIDLFLIDGSTSI	Q8VHI5	Vitrin OS=Mus musculus OX=10090 GN=Vit PE=1 SV=2	256	VIT	114
QRKPRPLDPGEGSKDTS	Q9EPL0	Xylosyltransferase 2 OS=Mus musculus OX=10090 GN=Xylt2 PE=2 SV=3	48	XYLT2	115
QSRSSYSGEYGGKG	Q8K0C5	Zymogen granule membrane protein 16 OS=Mus musculus OX=10090 GN=Zg16 PE=1 SV=1	19	ZG16	116
QPPTRFC	Q3UW41	Protein WFDC9 OS=Mus musculus OX=10090 GN=Wfdc9 PE=3 SV=1	37	WFDC9	117
QNEA	P22726	Protein Wnt-5b OS=Mus musculus OX=10090 GN=Wnt5b PE=1 SV=2	199	WNT5B	118

41/88

Fig. 14A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
QLSVERGGW	P43021	Nodal OS=Mus musculus OX=10090 GN=Nodal PE=2 SV=2	229	NODAL	119
QHMI	Q61292	Laminin subunit beta-2 OS=Mus musculus OX=10090 GN=Lamb2 PE=1 SV=2	547	LAMB2	120
QCRCKPGVV	Q61292	Laminin subunit beta-2 OS=Mus musculus OX=10090 GN=Lamb2 PE=1 SV=2	801	LAMB2	121
QVGCLH	Q9Z0L3	Otoconin-90 OS=Mus musculus OX=10090 GN=Oc90 PE=1 SV=2	367	OC90	122
QPLKTPGKKKKGKP	P22858	Parathyroid hormone-related protein OS=Mus musculus OX=10090 GN=Pthlh PE=2 SV=1	116	PTHLH	123
QKRDMHDFVGLM	P55099	Tachykinin-3 OS=Mus musculus OX=10090 GN=Tac3 PE=3 SV=2	78	TAC3	124
QFNAFSYHFR	Q3UQ28	Peroxidasin homolog OS=Mus musculus OX=10090 GN=Pxdn PE=1 SV=2	1320	PXDN	125
QISAKKYLESLI	P32648	VIP peptides OS=Mus musculus OX=10090 GN=Vip PE=1 SV=1	95	VIP	126
QMAVKKYLSILN	P32648	VIP peptides OS=Mus musculus OX=10090 GN=Vip PE=1 SV=1	139	VIP	127
QESC	O35188	Fractalkine OS=Mus musculus OX=10090 GN=Cx3cl1 PE=1 SV=3	54	CX3CL1	128
QHP	Q62361	Pro-thyrotropin-releasing hormone OS=Mus musculus OX=10090 GN=Trh PE=2 SV=2	76	TRH	129
QILAHV	Q99ML8	Urocortin-2 OS=Mus musculus OX=10090 GN=Ucn2 PE=1 SV=2	104	UCN2	130
QLAD	Q9Z1T2	Thrombospondin-4 OS=Mus musculus OX=10090 GN=Thbs4 PE=1 SV=1	119	THBS4	131
QTLHLGT	Q9WTS4	Teneurin-1 OS=Mus musculus OX=10090 GN=Tenm1 PE=1 SV=1	1963	TENM1	132
QYYDGL	Q9WTS4	Teneurin-1 OS=Mus musculus OX=10090 GN=Tenm1 PE=1 SV=1	2252	TENM1	133
QSEI	Q9WTS4	Teneurin-1 OS=Mus musculus OX=10090 GN=Tenm1 PE=1 SV=1	2724	TENM1	134
QYTMRECDNPVPGNGGKYCE	P97857	A disintegrin and metalloproteinase with thrombospondin motifs 1 OS=Mus musculus OX=10090 GN=Adamts1 PE=1 SV=4	580	ADAMTS1	135

42/88

Fig. 14A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
QSKGYFLFRPRN	Q9QXK8	Neuromedin-U OS=Mus musculus OX=10090 GN=Nmu PE=2 SV=1	154	NMU	136
QTNCPKKDWKKPECTIKPN	Q9DD06	Retinoic acid receptor responder protein 2 OS=Mus musculus OX=10090 GN=Rarres2 PE=1 SV=1	75	RARRE S2	137
QRY	Q9EPS2	Peptide YY OS=Mus musculus OX=10090 GN=Pyy PE=3 SV=2	61	PYY	138
QLVRPA	Q7TNK8	Protein ADM2 OS=Mus musculus OX=10090 GN=Adm2 PE=2 SV=1	127	ADM2	139
QMD	E9Q414	Apolipoprotein B-100 OS=Mus musculus OX=10090 GN=Apob PE=1 SV=1	3683	APOB	140
QHDGTGNDCESI	Q68SA9	A disintegrin and metalloproteinase with thrombospondin motifs 7 OS=Mus musculus OX=10090 GN=Adamts7 PE=1 SV=3	380	ADAM TS7	141
QYGEPEG	Q3USZ8	Divergent protein kinase domain 2A OS=Mus musculus OX=10090 GN=Dipk2a PE=1 SV=2	94	DIPK2 A	142
QSCGPGRGPV	Q61488	Desert hedgehog protein OS=Mus musculus OX=10090 GN=Dhh PE=1 SV=1	20	DHH	143
QLLGESDFAL	Q80UG6	ALK and LTK ligand 2 OS=Mus musculus OX=10090 GN=Alkal2 PE=2 SV=3	58	ALKAL 2	144
QRRNPENG	O08538	Angiopoietin-1 OS=Mus musculus OX=10090 GN=Angpt1 PE=1 SV=2	21	ANGPT 1	145
QVT	P63075	Fibroblast growth factor 17 OS=Mus musculus OX=10090 GN=Fgf17 PE=2 SV=1	66	FGF17	146
QCICPPGFVLSDN	Q61555	Fibrillin-2 OS=Mus musculus OX=10090 GN=Fbn2 PE=1 SV=2	2072	FBN2	147
QKKKEE	Q9JJZ5	Epidermal growth factor-like protein 6 OS=Mus musculus OX=10090 GN=Egfl6 PE=1 SV=1	328	EGFL6	148
QWRRARHNYNDLCPPI	Q9Z0F1	Neuroendocrine secretory protein 55 OS=Mus musculus OX=10090 GN=Gnas PE=2 SV=2	8	GNAS	149
QFPRAEGMPEAILKVK	Q8BJD1	Inter-alpha-trypsin inhibitor heavy chain H5 OS=Mus musculus OX=10090 GN=Itih5 PE=1 SV=1	865	ITI5	150
QSPG	Q9ESQ8	Pro-FMRamide-related neuropeptide VF OS=Mus	163	NPVF	151

43/88

Fig. 14A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
		musculus OX=10090 GN=Npvf PE=2 SV=1			
QDCDFKDAEEAATGECTATV	O08677	Kininogen-1 OS=Mus musculus OX=10090 GN=Kng1 PE=1 SV=1	91	KNG1	152
QCYSGYTLAED	O08746	Matrilin-2 OS=Mus musculus OX=10090 GN=Matn2 PE=2 SV=2	303	MATN2	153
QHFFHLRNDFAFIEVD	Q8BI06	Cell migration-inducing and hyaluronan-binding protein OS=Mus musculus OX=10090 GN=Cemip PE=1 SV=4	1219	CEMIP	154
QLRSYSETIEK	Q3V1M1	Immunoglobulin superfamily member 10 OS=Mus musculus OX=10090 GN=Igsf10 PE=2 SV=2	1547	IGSF10	155
QPVPKPYLTLC	Q6GUQ1	Epidermal growth factor-like protein 8 OS=Mus musculus OX=10090 GN=Egfl8 PE=2 SV=2	54	EGFL8	156
QHLYKDCSDHYVL	P12804	Fibroleukin OS=Mus musculus OX=10090 GN=Fgl2 PE=1 SV=1	198	FGL2	157
QLCNIPVT	Q8CG19	Latent-transforming growth factor beta-binding protein 1 OS=Mus musculus OX=10090 GN=Ltbp1 PE=1 SV=2	1565	LTBP1	158
QPLYH	Q5FW85	Extracellular matrix protein 2 OS=Mus musculus OX=10090 GN=Ecm2 PE=1 SV=1	229	ECM2	159
QLADDGLVALDWVIGPCAR	Q5F2F2	Protein ABHD15 OS=Mus musculus OX=10090 GN=Abhd15 PE=1 SV=1	118	ABHD15	160
QGPPGPSGEE	Q01149	Collagen alpha-2(I) chain OS=Mus musculus OX=10090 GN=Col1a2 PE=1 SV=2	374	COL1A2	161
QGAPGPAGEE	P28481	Collagen alpha-1(II) chain OS=Mus musculus OX=10090 GN=Col2a1 PE=1 SV=2	478	COL2A1	162
QLFE	P01029	Complement C4-B OS=Mus musculus OX=10090 GN=C4b PE=1 SV=3	1435	C4B	163
QTGSPGPTGEPGPSGPP	O88207	Collagen alpha-1(V) chain OS=Mus musculus OX=10090 GN=Col5a1 PE=1 SV=2	1372	COL5A1	164
QGAP	Q3U962	Collagen alpha-2(V) chain OS=Mus musculus OX=10090 GN=Col5a2 PE=1 SV=1	347	COL5A2	165

44/88

Fig. 14A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
QGPIGPPGEE	Q3U962	Collagen alpha-2(V) chain OS=Mus musculus OX=10090 GN=Col5a2 PE=1 SV=1	488	COL5A2	166
QRGEPGSRGPVGPGR	Q3U962	Collagen alpha-2(V) chain OS=Mus musculus OX=10090 GN=Col5a2 PE=1 SV=1	1096	COL5A2	167
QGAKGDSGRP	Q02788	Collagen alpha-2(VI) chain OS=Mus musculus OX=10090 GN=Col6a2 PE=1 SV=3	383	COL6A2	168
QGPKGEK	A2AX52	Collagen alpha-4(VI) chain OS=Mus musculus OX=10090 GN=Col6a4 PE=1 SV=2	1539	COL6A4	169
QGEDGNPGRGGEGKAGKIR	A2AX52	Collagen alpha-4(VI) chain OS=Mus musculus OX=10090 GN=Col6a4 PE=1 SV=2	1694	COL6A4	170
QGPP	A6H584	Collagen alpha-5(VI) chain OS=Mus musculus OX=10090 GN=Col6a5 PE=1 SV=4	1746	COL6A5	171
QGDHGIP	Q8C6K9	Collagen alpha-6(VI) chain OS=Mus musculus OX=10090 GN=Col6a6 PE=1 SV=2	1470	COL6A6	172
QGNP	Q8C6K9	Collagen alpha-6(VI) chain OS=Mus musculus OX=10090 GN=Col6a6 PE=1 SV=2	1488	COL6A6	173
QGKEAIKATSKL	Q8C6K9	Collagen alpha-6(VI) chain OS=Mus musculus OX=10090 GN=Col6a6 PE=1 SV=2	2248	COL6A6	174
QGLGGPPGLPGSAGLP	P25318	Collagen alpha-2(VIII) chain OS=Mus musculus OX=10090 GN=Col8a2 PE=2 SV=2	349	COL8A2	175
QASPYLFEAT	Q9D7Z6	Calcium-activated chloride channel regulator 1 OS=Mus musculus OX=10090 GN=Clca1 PE=1 SV=2	58	CLCA1	176
QPSG	Q9EPW4	C-type lectin domain family 3 member A OS=Mus musculus OX=10090 GN=Clec3a PE=3 SV=1	157	CLEC3A	177
QGVKGHP	Q07643	Collagen alpha-2(IX) chain OS=Mus musculus OX=10090 GN=Col9a2 PE=2 SV=1	187	COL9A2	178
QIGNTGPK	Q07643	Collagen alpha-2(IX) chain OS=Mus musculus OX=10090 GN=Col9a2 PE=2 SV=1	588	COL9A2	179

45/88

Fig. 14A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
QVLDT	Q8BQ47	Protein canopy homolog 4 OS=Mus musculus OX=10090 GN=Cnpy4 PE=1 SV=1	71	CNPY4	180
QILRLKPLTGEGTAGGGGG	Q9Z1W4	Growth/differentiation factor 11 OS=Mus musculus OX=10090 GN=Gdf11 PE=1 SV=1	193	GDF11	181
QLELRRLRVAAGR	Q9Z0J7	Growth/differentiation factor 15 OS=Mus musculus OX=10090 GN=Gdf15 PE=1 SV=2	173	GDF15	182
QPPPLDLRSLGF	P43028	Growth/differentiation factor 6 OS=Mus musculus OX=10090 GN=Gdf6 PE=2 SV=2	256	GDF6	183
QVK	P13020	Gelsolin OS=Mus musculus OX=10090 GN=Gsn PE=1 SV=3	188	GSN	184
QGPKGEKGAI	Q8BMF8	Gliomedin OS=Mus musculus OX=10090 GN=Gldn PE=1 SV=1	164	GLDN	185
QVYAPEVSPESIMECAT	Q9ESY9	Gamma-interferon-inducible lysosomal thiol reductase OS=Mus musculus OX=10090 GN=Ifi30 PE=1 SV=3	160	IFI30	186
QGHAG	P21275	Bone morphogenetic protein 4 OS=Mus musculus OX=10090 GN=Bmp4 PE=1 SV=2	38	BMP4	187
QHFC	Q69Z28	A disintegrin and metalloproteinase with thrombospondin motifs 16 OS=Mus musculus OX=10090 GN=Adamts16 PE=2 SV=2	243	ADAM TS16	188
QLGPSGHDA	Q9DB72	BTB/POZ domain-containing protein 17 OS=Mus musculus OX=10090 GN=Btbd17 PE=1 SV=1	341	BTBD1 7	189
QGLTGPPGKA	Q64739	Collagen alpha-2(XI) chain OS=Mus musculus OX=10090 GN=Col11a2 PE=2 SV=3	524	COL11 A2	190
QPGSPGPTGENGPPGPL	Q64739	Collagen alpha-2(XI) chain OS=Mus musculus OX=10090 GN=Col11a2 PE=2 SV=3	1300	COL11 A2	191
QVVFSETVITSVGDEE	Q9QUN9	Dickkopf-related protein 3 OS=Mus musculus OX=10090 GN=Dkk3 PE=2 SV=1	124	DKK3	192
QFRKKWNKWALSR	P97297	Pro-adrenomedullin OS=Mus musculus OX=10090 GN=Adm PE=2 SV=2	28	ADM	193

46/88

Fig. 14A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
QGY	P97297	Pro-adrenomedullin OS=Mus musculus OX=10090 GN=Adm PE=2 SV=2	141	ADM	194
QNTDYAEKALKNPYKDRA	Q7TN31	Angiogenic factor with G patch and FHA domains 1 OS=Mus musculus OX=10090 GN=Aggf1 PE=2 SV=1	568	AGGF1	195
QNPDAPE	Q8CIE6	Coatomer subunit alpha OS=Mus musculus OX=10090 GN=Copa PE=1 SV=2	402	COPA	196
QTPYPGLSNYRESLR	P48299	Endothelin-3 OS=Mus musculus OX=10090 GN=Edn3 PE=2 SV=1	122	EDN3	197
QYLELDVNIST	Q6AW46	Carboxylesterase 5A OS=Mus musculus OX=10090 GN=Ces5a PE=2 SV=1	516	CES5A	198
QEANP	O55189	Ameloblastin OS=Mus musculus OX=10090 GN=Ambn PE=2 SV=1	292	AMBN	199
QECCPPEL	Q99K41	EMILIN-1 OS=Mus musculus OX=10090 GN=Emilin1 PE=1 SV=1	351	EMILIN 1	200
QAGL	Q99K41	EMILIN-1 OS=Mus musculus OX=10090 GN=Emilin1 PE=1 SV=1	783	EMILIN 1	201
QCYNFGWYPESPICE	Q07968	Coagulation factor XIII B chain OS=Mus musculus OX=10090 GN=F13b PE=1 SV=2	254	F13B	202
QEATGRAVSTAHPPS	Q62507	Cochlin OS=Mus musculus OX=10090 GN=Coch PE=1 SV=2	132	COCH	203
QVVPVDGSHREVRVSGLDLP	O35452	Tenascin X OS=Mus musculus OX=10090 GN=Tnxb PE=1 SV=1	1892	TNXB	204
QPGRKYKFLLYGLTG	O35452	Tenascin X OS=Mus musculus OX=10090 GN=Tnxb PE=1 SV=1	3273	TNXB	205
QAEAACRPS	Q3TBV5	Interleukin-1 OS=Mus musculus OX=10090 GN=Il1rn PE=1 SV=1	4	IL1RN	206
QFGEMIWLKT	V5TCT6	Phospholipase A2 OS=Mus musculus OX=10090 GN=Pla2g2a PE=3 SV=1	24	PLA2G 2A	207
QMMAMLM	W8R375	Lipocalin 16 OS=Mus musculus OX=10090 GN=Gm39773 PE=2 SV=1	130	GM39 773	208
QPVTPLLEFE	V9GZG9	MHC class III H2-Slp (haplotype a) sex-limited protein mRNA (Fragment) OS=Mus musculus OX=10090 PE=2 SV=1	1407	none	209
QLPGMLCAGYPA	E9QJT0	Polyserase-2 OS=Mus musculus OX=10090 GN=Prss36 PE=4 SV=1	221	PRSS3 6	210

Fig. 14A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
QPGST	Q9ESQ1	Type IV collagen alpha 6 chain OS=Mus musculus OX=10090 GN=Col4a6 PE=2 SV=1	852	COL4A6	211
QTNTIKGLNGRPGSPGSM	Q9ESQ1	Type IV collagen alpha 6 chain OS=Mus musculus OX=10090 GN=Col4a6 PE=2 SV=1	1005	COL4A6	212
QLFGLKD	Q3TAQ2	Uncharacterized protein OS=Mus musculus OX=10090 GN=Tnxb PE=2 SV=1	420	TNXB	213
QKLTAKE	G3UXC7	ADAMTS-like 3 OS=Mus musculus OX=10090 GN=Adamtsl3 PE=1 SV=1	859	ADAMTSL3	214
QHHTACSHNSSSDCED	G3UXC7	ADAMTS-like 3 OS=Mus musculus OX=10090 GN=Adamtsl3 PE=1 SV=1	1580	ADAMTSL3	215
QAMVSDRNHTD	A2BHR0	Odorant-binding protein 2B OS=Mus musculus OX=10090 GN=Obp2b PE=3 SV=1	36	OBP2B	216

48/88

Fig. 14B

Capped peptide names	Sequence	Precursor Uniprot	Precursor Fasta header	Seq ID
CAP-SNED1	QSTEVDRSVDRLTFGDLLP	Q70E20	Sushi nidogen and EGF-like domain-containing protein 1 OS=Mus musculus OX=10090 GN=Sned1 PE=2 SV=2	4
CAP-PROC	QDCRVNNGGCLHYCLEESN	P33587	Vitamin K-dependent protein C OS=Mus musculus OX=10090 GN=Proc PE=1 SV=2	7
CAP-TAC1	QFFGLM	P41539	Protachykinin-1 OS=Mus musculus OX=10090 GN=Tac1 PE=2 SV=1	10
CAP-TAC4	QFYGLM	Q99N14	Tachykinin-4 OS=Mus musculus OX=10090 GN=Tac4 PE=1 SV=1	12
CAP-SCT	QGLV	Q08535	Secretin OS=Mus musculus OX=10090 GN=Sct PE=2 SV=1	13
CAP-ADAMTS5	QLYSGGGKVGILVYAG	Q9R001	A disintegrin and metalloproteinase with thrombospondin motifs 5 OS=Mus musculus OX=10090 GN=Adamts5 PE=2 SV=2	18
CAP-MATN4	QCVSEGLSFHCLCPE	O89029	Matrilin-4 OS=Mus musculus OX=10090 GN=Matn4 PE=1 SV=1	22
CAP-ADAMTS18	QKRHFC	Q4VC17	A disintegrin and metalloproteinase with thrombospondin motifs 18 OS=Mus musculus OX=10090 GN=Adamts18 PE=2 SV=2	23
CAP-COL17A1	QVYT	Q07563	Collagen alpha-1(XVII) chain OS=Mus musculus OX=10090 GN=Col17a1 PE=1 SV=3	30
CAP-COL27A1	QLGPP	Q5QNQ9	Collagen alpha-1(XXVII) chain OS=Mus musculus OX=10090 GN=Col27a1 PE=2 SV=1	32
CAP-CSF1	QLLLPKSHSWGIVLPLGELE	P07141	Macrophage colony-stimulating factor 1 OS=Mus musculus OX=10090 GN=Csf1 PE=1 SV=2	33
CAP-FJX1	QRRDLFLAKHILHCKAKY	Q8BQB4	Four-jointed box protein 1 OS=Mus musculus OX=10090 GN=Fjx1 PE=1 SV=1	34
CAP-MST1	QVVFEC	P26928	Hepatocyte growth factor-like protein OS=Mus musculus OX=10090 GN=Mst1 PE=2 SV=2	35
CAP-ANGPTL4	QIDLLAPTHLDNGVDKTSR	Q9Z1P8	Angiopoietin-related protein 4 OS=Mus musculus OX=10090 GN=Angptl4 PE=1 SV=1	36
CAP-FGF5	QWSPS	P15656	Fibroblast growth factor 5 OS=Mus musculus OX=10090 GN=Fgf5 PE=2 SV=1	39

49/88

Fig. 14B (cont.)

Capped peptide names	Sequence	Precursor Uniprot	Precursor Fasta header	Seq ID
CAP-FGL1	QFLDKGSENSFIDLG	Q71KU9	Fibrinogen-like protein 1 OS=Mus musculus OX=10090 GN=Fgl1 PE=1 SV=2	40
CAP-METTL24	QVTYVRS	Q8CCB5	Probable methyltransferase-like protein 24 OS=Mus musculus OX=10090 GN=Mettl24 PE=2 SV=1	41
CAP-GNRH1	QHWSYGLRPG	P13562	Progonadoliberin-1 OS=Mus musculus OX=10090 GN=Gnrh1 PE=3 SV=1	42
CAP-NTN4	QHWKPAL	Q9JI33	Netrin-4 OS=Mus musculus OX=10090 GN=Ntn4 PE=1 SV=2	43
CAP-FGF9	QFEENWYNTYSSNLYKHVDT	P54130	Fibroblast growth factor 9 OS=Mus musculus OX=10090 GN=Fgf9 PE=2 SV=2	45
CAP-DEFA-RS7	QDAALGW	P50715	Alpha-defensin-related sequence 7 OS=Mus musculus OX=10090 GN=Defa-rs7 PE=2 SV=2	51
CAP-EPHB6	QPR	O08644	Ephrin type-B receptor 6 OS=Mus musculus OX=10090 GN=Ephb6 PE=2 SV=4	52
CAP-CBLN2	QGPRGPILLSMP	Q8BGU2	Cerebellin-2 OS=Mus musculus OX=10090 GN=Cbln2 PE=1 SV=1	60
CAP-DEFB1	QMEPGVGILTSL	P56386	Beta-defensin 1 OS=Mus musculus OX=10090 GN=Defb1 PE=2 SV=1	61
CAP-GAST	QRPRMEEEEAYGWMDF	P48757	Gastrin OS=Mus musculus OX=10090 GN=Gast PE=2 SV=2	62
CAP-FGF18	QVL	O89101	Fibroblast growth factor 18 OS=Mus musculus OX=10090 GN=Fgf18 PE=2 SV=1	64
CAP-CA15	QPLG	Q99N23	Carbonic anhydrase 15 OS=Mus musculus OX=10090 GN=Ca15 PE=1 SV=1	66
CAP-GCG	QAAKEFIAWLVKGR	P55095	Pro-glucagon OS=Mus musculus OX=10090 GN=Gcg PE=1 SV=1	75
CAP-LY86	QPKFSFC	O88188	Lymphocyte antigen 86 OS=Mus musculus OX=10090 GN=Ly86 PE=1 SV=1	76
CAP-GDNF	QAAAASPENSRGK	P48540	Glial cell line-derived neurotrophic factor OS=Mus musculus OX=10090 GN=Gdnf PE=1 SV=1	77
CAP-PLA2G2A	QFGEMIRLKT	P31482	Phospholipase A2 membrane associated OS=Mus musculus OX=10090 GN=Pla2g2a PE=1 SV=3	78

50/88

Fig. 14B (cont.)

Capped peptide names	Sequence	Precursor Uniprot	Precursor Fasta header	Seq ID
CAP-OBSCN	QLVTE	A2AAJ9	Obscurin OS=Mus musculus OX=10090 GN=Obscn PE=1 SV=3	84
CAP-ADCYAP1	QMAVKKYLA AVL	O70176	Pituitary adenylate cyclase-activating polypeptide OS=Mus musculus OX=10090 GN=Adcyap1 PE=2 SV=1	90
CAP-ADCYAP1-2	QRVKNK	O70176	Pituitary adenylate cyclase-activating polypeptide OS=Mus musculus OX=10090 GN=Adcyap1 PE=2 SV=1	91
CAP-PRSS36	QLLPGMLCAGYPE	Q5K2P8	Polyserase-2 OS=Mus musculus OX=10090 GN=Prss36 PE=2 SV=1	96
CAP-VGF	QVEA	Q0VGU4	Neurosecretory protein VGF OS=Mus musculus OX=10090 GN=Vgf PE=1 SV=1	98
CAP-AVP	QNCPRG	P35455	Vasopressin-neurophysin 2-copeptin OS=Mus musculus OX=10090 GN=Avp PE=2 SV=1	103
CAP-PRG4	QEPMKKCT	Q9JM99	Proteoglycan 4 OS=Mus musculus OX=10090 GN=Prg4 PE=1 SV=2	104
CAP-FAP	QDSLFYSSNEFEGYP	P97321	Prolyl endopeptidase FAP OS=Mus musculus OX=10090 GN=Fap PE=1 SV=1	105
CAP-TREM2	QSLRVSTYDALKHW	Q99NH8	Triggering receptor expressed on myeloid cells 2 OS=Mus musculus OX=10090 GN=Trem2 PE=1 SV=1	106
CAP-WNT3A	QFR	P27467	Protein Wnt-3a OS=Mus musculus OX=10090 GN=Wnt3a PE=1 SV=1	107
CAP-WNT9A	QWGGCGDNLKYSSKFVKEFL	Q8R5M2	Protein Wnt-9a OS=Mus musculus OX=10090 GN=Wnt9a PE=2 SV=1	113
CAP-WNT5B	QNNEA	P22726	Protein Wnt-5b OS=Mus musculus OX=10090 GN=Wnt5b PE=1 SV=2	118
CAP-LAMB2	QHMI	Q61292	Laminin subunit beta-2 OS=Mus musculus OX=10090 GN=Lamb2 PE=1 SV=2	120
CAP-TAC3	QKRDMHDFVGLM	P55099	Tachykinin-3 OS=Mus musculus OX=10090 GN=Tac3 PE=3 SV=2	124
CAP-VIP	QISAKKYLESLI	P32648	VIP peptides OS=Mus musculus OX=10090 GN=Vip PE=1 SV=1	126
CAP-VIP-2	QMAVKKYLSILN	P32648	VIP peptides OS=Mus musculus OX=10090 GN=Vip PE=1 SV=1	127
CAP-THBS4	QLAD	Q9Z1T2	Thrombospondin-4 OS=Mus musculus OX=10090 GN=Thbs4 PE=1 SV=1	131
CAP-TENM1	QYYDGL	Q9WTS4	Teneurin-1 OS=Mus musculus OX=10090 GN=Tenm1 PE=1 SV=1	133

51/88

Fig. 14B (cont.)

Capped peptide names	Sequence	Precursor Uniprot	Precursor Fasta header	Seq ID
CAP-ALKAL2	QLLGESDFAL	Q80UG6	ALK and LTK ligand 2 OS=Mus musculus OX=10090 GN=Alkal2 PE=2 SV=3	144
CAP-FBN2	QCICPPGFVLSDN	Q61555	Fibrillin-2 OS=Mus musculus OX=10090 GN=Fbn2 PE=1 SV=2	147
CAP-GNAS	QWRRARHNYNDLCPPI	Q9Z0F1	Neuroendocrine secretory protein 55 OS=Mus musculus OX=10090 GN=Gnas PE=2 SV=2	149
CAP-CEMIP	QHFFHLRNDFAIEVD	Q8BI06	Cell migration-inducing and hyaluronan-binding protein OS=Mus musculus OX=10090 GN=Cemip PE=1 SV=4	154
CAP-EGFL8	QPVYKPYLTICA	Q6GUQ1	Epidermal growth factor-like protein 8 OS=Mus musculus OX=10090 GN=Egfl8 PE=2 SV=2	156
CAP-COL5A2	QGAP	Q3U962	Collagen alpha-2(V) chain OS=Mus musculus OX=10090 GN=Col5a2 PE=1 SV=1	165
CAP-COL5A2-2	QRGEPGSRGPVGPGR	Q3U962	Collagen alpha-2(V) chain OS=Mus musculus OX=10090 GN=Col5a2 PE=1 SV=1	167
CAP-COL6A6	QGDHGIP	Q8C6K9	Collagen alpha-6(VI) chain OS=Mus musculus OX=10090 GN=Col6a6 PE=1 SV=2	172
CAP-CNPY4	QVLDT	Q8BQ47	Protein canopy homolog 4 OS=Mus musculus OX=10090 GN=Cnp4 PE=1 SV=1	180
CAP-GDF15	QLELRVAAGR	Q9Z0J7	Growth/differentiation factor 15 OS=Mus musculus OX=10090 GN=Gdf15 PE=1 SV=2	182
CAP-GDF6	QPPPLDLRSLGF	P43028	Growth/differentiation factor 6 OS=Mus musculus OX=10090 GN=Gdf6 PE=2 SV=2	183
CAP-COL11A2	QPGSPGPTGENGPPL	Q64739	Collagen alpha-2(XI) chain OS=Mus musculus OX=10090 GN=Col11a2 PE=2 SV=3	191
CAP-EDN3	QTPYGLSNYRESLR	P48299	Endothelin-3 OS=Mus musculus OX=10090 GN=Edn3 PE=2 SV=1	197
CAP-CES5A	QYLELDVNIST	Q6AW46	Carboxylesterase 5A OS=Mus musculus OX=10090 GN=Ces5a PE=2 SV=1	198
CAP-EMILIN1	QAGL	Q99K41	EMILIN-1 OS=Mus musculus OX=10090 GN=Emilin1 PE=1 SV=1	201

52/88

Fig. 14C

Capped peptide names	Amino Acid Start	MW (g/mol)	m/z detected	Z	RT (min)	Seq ID
CAP-SNED1	1050	2129.07534	1065.5455	2	24	4
CAP-PROC	136	2305.92632	769.6499367	3	16	7
CAP-TAC1	62	723.34147	724.3493	1	30	10
CAP-TAC4	60	739.33638	740.34421	1	26	12
CAP-SCT	54	397.23257	398.2404	1	15.5	13
CAP-ADAMTS5	79	1612.82499	807.420325	2	22.3	18
CAP-MATN4	353	1803.7379	902.87678	2	21.3	22
CAP-ADAMTS18	248	856.41267	857.4205	1	3.9	23
CAP-COL17A1	1455	491.23805	492.24588	1	13.4	30
CAP-COL27A1	1484	492.26968	493.27751	1	15.5	32
CAP-CSF1	418	2210.24637	737.75662	3	27.8	33
CAP-FJX1	413	2278.26337	760.4289533	3	22.7	34
CAP-MST1	470	890.43207	446.223865	2	17	35
CAP-ANGPTL4	146	2074.08076	692.3680833	3	20	36
CAP-FGF5	77	585.25476	586.26259	1	16.5	39
CAP-FGL1	62	1651.79627	826.905965	2	24.4	40
CAP-METTL24	67	833.4396	834.44743	1	13.6	41
CAP-GNRH1	21	1181.57307	591.794365	2	17.5	42
CAP-NTN4	608	860.46575	861.47358	1	15.9	43
CAP-FGF9	138	2519.10301	840.7088333	3	23	45
CAP-DEFA-RS7	54	741.34464	742.35247	1	21.5	51
CAP-EPHB6	682	381.2125	382.22033	1	3.5	52
CAP-CBLN2	5	1133.6016	1134.60943	1	20.8	60
CAP-DEFB1	16	1225.63771	1226.64554	1	27.5	61
CAP-GAST	75	2183.90412	1092.95989	2	23.2	62
CAP-FGF18	66	340.2111	341.21893	1	15.5	64
CAP-CA15	281	395.21692	396.22475	1	13.3	66
CAP-GCG	113	1597.90933	799.962495	2	17	75
CAP-LY86	105	894.40585	895.41368	1	19.9	76
CAP-GDNF	93	1267.62696	634.82131	2	2	77
CAP-PLA2G2A	24	1203.64346	602.82956	2	22.5	78
CAP-OBSCN	273	570.30138	571.30921	1	16.1	84
CAP-ADCYAP1	145	1315.76866	658.89216	2	25	90
CAP-ADCYAP1-2	162	753.461	754.46883	2	3.5	91
CAP-PRSS36	220	1429.67344	715.84455	2	26.5	96
CAP-VGF	305	427.20675	428.21458	1	12.3	98
CAP-AVP	26	712.30754	713.31537	1	3.9	103
CAP-PRG4	923	1002.46272	502.23919	2	2	104
CAP-FAP	404	1763.73154	882.8736	2	24.8	105
CAP-TREM2	29	1844.89923	923.457445	2	20	106
CAP-WNT3A	79	431.22815	432.23598	1	11.24	107
CAP-WNT9A	163	2344.13108	782.3848567	3	23.8	113
CAP-WNT5B	199	556.22419	557.23202	1	3.5	118
CAP-LAMB2	547	509.24209	510.24992	1	13	120
CAP-TAC3	78	1604.75924	803.38745	2	23.6	124

53/88

Fig. 14C (cont.)

Capped peptide names	Amino Acid Start	MW (g/mol)	m/z detected	Z	RT (min)	Seq ID
CAP-VIP	95	1373.7919	687.90378	2	24	126
CAP-VIP-2	139	1502.82797	752.421815	2	24	127
CAP-THBS4	119	427.20675	428.21458	1	12.5	131
CAP-TENM1	2252	902.38108	903.38891	1	21.8	133
CAP-ALKAL2	58	1073.53937	537.777515	2	26.6	144
CAP-FBN2	2072	1487.65376	744.83471	2	23	147
CAP-GNAS	8	2077.01772	693.34707	3	16	149
CAP-CEMIP	1219	2031.95919	678.32756	3	22.2	154
CAP-EGFL8	54	1433.7377	717.87668	2	21.8	156
CAP-COL5A2	347	353.16997	354.1778	1	2	165
CAP-COL5A2-2	1096	1655.86048	552.9613233	3	10	167
CAP-COL6A6	1470	704.32424	705.33207	1	11.7	172
CAP-CNPY4	71	556.28573	557.29356	1	15.7	180
CAP-GDF15	173	1362.82085	682.418255	2	17.7	182
CAP-GDF6	256	1320.71907	661.367365	2	25.7	183
CAP-COL11A2	1300	1539.73182	770.87374	2	18.2	191
CAP-EDN3	122	1763.89553	882.955595	2	19	197
CAP-CES5A	516	1275.63473	1276.64256	1	31	198
CAP-EMILIN1	783	369.20127	370.2091	1	12.5	201

54/88

Fig. 14D

Capped peptide names	Plasma 1	Plasma 2	Plasma 3	100 nM Standard	10 nM Standard	Seq ID
CAP-SNED1	793.28	0	3037.28	230003.87	16451.28	4
CAP-PROC	1288.44	605.51	481.55	48113.96	3933.45	7
CAP-TAC1	4832.06	9149.99	8587.76	608761.66	35231.19	10
CAP-TAC4	7847.84	8618.17	17924.81	74559.29	5411.36	12
CAP-SCT	1133.75	981.97	1281.42	606347.3	55338.82	13
CAP-ADAMTS5	52566.53	3734.18	1067.82	510355.52	41507.08	18
CAP-MATN4	1670.1	812.89	4023.29	287702.92	17563.04	22
CAP-ADAMTS18	3247.78	2390.25	4045.57	21042.96	2150.9	23
CAP-COL17A1	1276.76	1201.05	2970.97	259478.12	23795.74	30
CAP-COL27A1	24128.24	20450.49	17291.22	1102893.94	85954.13	32
CAP-CSF1	564.4	2519.97	3471.65	115686.76	4766.96	33
CAP-FIX1	10947.61	4362.4	5863.79	71162.33	0	34
CAP-MST1	546.14	0	407.44	509545.78	45842.98	35
CAP-ANGPTL4	780.62	0	1127.12	935109.66	88606.07	36
CAP-FGF5	1401.56	1629.79	3756.16	339194.18	32259.94	39
CAP-FGL1	8904.89	3541.7	7126.9	380066.33	30398.8	40
CAP-METTL24	879.31	1790.45	2995.75	68046.87	3197.97	41
CAP-GNRH1	217.73	206.31	830.31	755953.48	74629.37	42
CAP-NTN4	6524.05	4453.27	4761.08	569460.77	23688.84	43
CAP-FGF9	3950.48	2696.49	4972.3	73261.61	5673.59	45
CAP-DEFA-RS7	473.9	0	2177.47	286691.35	25637.79	51
CAP-EPHB6	2121.92	3310.07	3488.45	31292.94	2476.56	52
CAP-CBLN2	684.53	2075.59	3809.82	703339.14	68035.89	60
CAP-DEFB1	2183.22	883.51	336.88	372696.42	22609.96	61
CAP-GAST	18429.96	235.51	202.46	80803.45	10583.21	62
CAP-FGF18	0	967.1	1588.96	215194.96	23194.6	64
CAP-CA15	13852.54	4196.09	7545.23	162306.47	7553.44	66
CAP-GCG	1399.44	1873.71	2943.32	202650.34	10728.64	75
CAP-LY86	598.35	249.11	1098.93	275493.14	21010.03	76
CAP-GDNF	1326.83	672	861.11	239455.41	35324.23	77
CAP-PLA2G2A	1088.93	848.8	1367.05	751036.06	50308.59	78
CAP-OBSCN	1598.66	254.99	2035.69	602908.79	47922.99	84
CAP-ADCYAP1	0	1061.91	2257.59	666433.75	27797.65	90
CAP-ADCYAP1-2	214.74	0	1591.43	79833.75	4367.92	91
CAP-PRSS36	11779.11	2637.56	6181.87	210510.68	10363.76	96
CAP-VGF	1102.89	0	506.73	87242.15	2401.11	98
CAP-AVP	6321.04	5753.2	7566.09	43606.99	6011.7	103
CAP-PRG4	3138.57	3419.89	5914.97	347703.16	38573.84	104
CAP-FAP	23211.72	1700.63	4488.57	154601.12	17847.61	105
CAP-TREM2	1809.46	1368.76	1079.29	38359.27	2760.02	106
CAP-WNT3A	12291.35	6459.14	10443.02	987669.88	67912.6	107
CAP-WNT9A	3350.07	4958.97	3723.58	163306.02	10935.31	113
CAP-WNT5B	5528.67	4597.62	6604.62	18264.97	5307.8	118
CAP-LAMB2	1070.8	473.74	1574.48	486598.01	52303.6	120

55/88

Fig. 14D (cont.)

Capped peptide names	Plasma 1	Plasma 2	Plasma 3	100 nM Standard	10 nM Standard	Seq ID
CAP-TAC3	569.31	1268	1103.95	334161.07	15848.97	124
CAP-VIP	2169.54	1805.25	4932.52	1060547.4	51908.72	126
CAP-VIP-2	4777.89	4731.26	4527.05	33345.35	0	127
CAP-THBS4	0	0	1110.53	439741.8	39678.46	131
CAP-TENM1	1981.36	1766.17	3606.97	397546.35	25404.63	133
CAP-ALKAL2	13399.51	2980.29	5283.27	100591.4	6769.05	144
CAP-FBN2	392.89	521.97	1466.5	135619.04	9159.9	147
CAP-GNAS	4051.82	6771.78	9912.89	206758.01	11735.85	149
CAP-CEMIP	3646.28	1030.42	1322.19	139021.31	8497.87	154
CAP-EGFL8	461.43	1139.04	910.69	34181.08	1547.48	156
CAP-COL5A2	591.18	1812.03	1292.72	100779.26	14028.02	165
CAP-COL5A2-2	251.97	1133.72	1049.34	15130.53	1690.7	167
CAP-COL6A6	5798.56	3829.7	6214.94	516388.88	50845.99	172
CAP-CNPY4	1926.18	1015.38	1126.32	1087903.55	100271.41	180
CAP-GDF15	954.32	252.74	1069.48	561542.33	88142.49	182
CAP-GDF6	4636.7	3782.82	4335.56	1874075.86	112765.88	183
CAP-COL11A2	1153.51	1722.95	1087.69	479449.15	48463.05	191
CAP-EDN3	939.54	286.57	1377.19	839225.16	68738.62	197
CAP-CES5A	25642.29	7779.22	2028.59	13046.27	6204.38	198
CAP-EMILIN1	1519.46	2847.6	2359.68	251910.38	30169.28	201

56/88

Fig. 14E

Capped peptide names	Plasma 1	Plasma 2	Plasma 3	Average concentration (nM)	Seq ID
CAP-SNED1	0.334321396	0	1.280036922	0.53811944	4
CAP-PROC	2.624677714	1.233482819	0.980964231	1.61304159	7
CAP-TAC1	0.758260324	1.435841935	1.347615237	1.1805725	10
CAP-TAC4	10.21441423	11.21704294	23.3301691	14.9205421	12
CAP-SCT	0.185183175	0.160391905	0.209303131	0.1849594	13
CAP-ADAMTS5	10.09065467	0.716812026	0.204978393	3.67081503	18
CAP-MATN4	0.556411738	0.270823027	1.340402239	0.72254567	22
CAP-ADAMTS18	15.47211897	11.38692657	19.27271563	15.3772537	23
CAP-COL17A1	0.487556176	0.458644808	1.13452393	0.69357497	30
CAP-COL27A1	2.135368857	1.809884992	1.530287028	1.82518029	32
CAP-CSF1	0.457952503	2.044696258	2.816886615	1.77317846	33
CAP-FIX1	13.8455964	5.517188659	7.416017716	8.92626759	34
CAP-MST1	0.106000223	0	0.079079962	0.0616934	35
CAP-ANGPTL4	0.082995277	0	0.11983505	0.06761011	36
CAP-FGF5	0.410968812	0.477890965	1.101390317	0.6634167	39
CAP-FGL1	2.292006066	0.911588788	1.834373927	1.67932293	40
CAP-METTL24	1.220342982	2.484860961	4.157626421	2.62094345	41
CAP-GNRH1	0.028761201	0.027252668	0.109680399	0.05523142	42
CAP-NTN4	1.075842248	0.734362245	0.78512136	0.86510862	43
CAP-FGF9	5.260447044	3.590637808	6.62109942	5.15739476	45
CAP-DEFA-RS7	0.163380266	0	0.750697673	0.30469265	51
CAP-EPHB6	6.627230763	10.33808896	10.8952096	9.28684311	52
CAP-CBLN2	0.096973689	0.29403769	0.539716742	0.31024271	60
CAP-DEFB1	0.561260781	0.227132178	0.086604892	0.29166595	61
CAP-GAST	23.62134336	0.301848869	0.259489287	8.06089384	62
CAP-FGF18	0	0.453327275	0.744823604	0.39938363	64
CAP-CA15	8.056246782	2.440327663	4.38809308	4.96155584	66
CAP-GCG	0.656255129	0.878659891	1.380244131	0.97171972	75
CAP-LY86	0.211611293	0.088099756	0.388645439	0.22945216	76
CAP-GDNF	0.584990005	0.296280069	0.379657336	0.42030914	77
CAP-PLA2G2A	0.139859937	0.109018132	0.1755811	0.14148639	78
CAP-OBSCN	0.259248795	0.041350788	0.330120338	0.21023997	84
CAP-ADCYAP1	0	0.149650012	0.318151605	0.15593387	90
CAP-ADCYAP1-2	0.256097362	0	1.897927844	0.7180084	91
CAP-PRSS36	5.296708538	1.186030742	2.779799459	3.08751291	96
CAP-VGF	1.169953834	0	0.537542916	0.56916558	98
CAP-AVP	15.13204447	13.7726827	18.1125907	15.6724393	103
CAP-PRG4	0.913764181	0.995667768	1.722086084	1.21050601	104
CAP-FAP	15.27605982	1.11921588	2.954010467	6.44976206	105
CAP-TREM2	4.574573903	3.460421217	2.728599619	3.58786491	106
CAP-WNT3A	1.202731986	0.63203914	1.021869378	0.9522135	107
CAP-WNT9A	1.978768098	2.929088537	2.199387271	2.3690813	113
CAP-WNT5B	38.40192727	31.9348901	45.87543422	38.7374172	118
CAP-LAMB2	0.221904767	0.098174416	0.32628373	0.2154543	120

57/88

Fig. 14E (cont.)

Capped peptide names	Plasma 1	Plasma 2	Plasma 3	Average concentration (nM)	Seq ID
CAP-TAC3	0.160967491	0.35851606	0.312132338	0.2772053	124
CAP-VIP	0.19358627	0.161080973	0.440124704	0.26493065	126
CAP-VIP-2	12.89565412	12.76979849	12.21863018	12.6280276	127
CAP-THBS4	0	0	0.24982969	0.08327656	131
CAP-TENM1	0.479178739	0.427136468	0.872321706	0.59287897	133
CAP-ALKAL2	12.85361004	2.858872113	5.068028034	6.92683673	144
CAP-FBN2	0.2796168	0.371482046	1.043696802	0.56493188	147
CAP-GNAS	1.869858277	3.125081786	4.574660131	3.18986673	149
CAP-CEMIP	2.514224265	0.710506864	0.911691417	1.37880752	154
CAP-EGFL8	1.272574892	3.141351245	2.511586218	2.30850412	156
CAP-COL5A2	0.613319187	1.879888979	1.341131262	1.27811314	165
CAP-COL5A2-2	1.68732045	7.591971029	7.026919239	5.43540357	167
CAP-COL6A6	1.12099317	0.740367875	1.201488868	1.02094997	172
CAP-CNPY4	0.175527095	0.092528581	0.102638215	0.12356463	180
CAP-GDF15	0.181429719	0.048049446	0.203323263	0.14426748	182
CAP-GDF6	0.23692763	0.193295788	0.221539879	0.21725443	183
CAP-COL11A2	0.240879926	0.359792346	0.227135168	0.27593581	191
CAP-EDN3	0.109747018	0.033474044	0.160868612	0.10136322	197
CAP-CES5A	337.3053498	102.3298825	26.68460031	155.439944	198
CAP-EMILIN1	0.616716522	1.155780322	0.957743964	0.91008027	201

58/88

Fig. 15A

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
QCTRCATGHYGFPRCKPCSC	Q16787	Laminin subunit alpha-3 OS=Homo sapiens OX=9606 GN=LAMA3 PE=1 SV=3	1295	LAMA3	217
QHMOV	P55268	Laminin subunit beta-2 OS=Homo sapiens OX=9606 GN=LAMB2 PE=1 SV=2	545	LAMB2	218
QCLCKPGVV	P55268	Laminin subunit beta-2 OS=Homo sapiens OX=9606 GN=LAMB2 PE=1 SV=2	801	LAMB2	219
QATCDACAPGYT	P98160	Basement membrane-specific heparan sulfate proteoglycan core protein OS=Homo sapiens OX=9606 GN=HSPG2 PE=1 SV=4	836	HSPG2	220
QAFENGVRM	Q92824	Proprotein convertase subtilisin/kexin type 5 OS=Homo sapiens OX=9606 GN=PCSK5 PE=1 SV=4	290	PCSK5	221
QGL	O43157	Plexin-B1 OS=Homo sapiens OX=9606 GN=PLXNB1 PE=1 SV=3	1298	PLXNB 1	222
QCVWGHNKER	Q6UXB8	Peptidase inhibitor 16 OS=Homo sapiens OX=9606 GN=PI16 PE=1 SV=1	73	PI16	223
QILPGMLCAGYPE	Q5K4E3	Polyserase-2 OS=Homo sapiens OX=9606 GN=PRSS36 PE=1 SV=2	221	PRSS3 6	224
QAGITSFGFGC	Q5K4E3	Polyserase-2 OS=Homo sapiens OX=9606 GN=PRSS36 PE=1 SV=2	257	PRSS3 6	225
QGHAG	A0A097 PIG0	Bone morphogenetic protein 4 (Fragment) OS=Homo sapiens OX=9606 GN=BMP4 PE=3 SV=1	39	BMP4	226
QSRGYFLFRPRN	D6RBC9	Neuromedin U isoform CRA_a OS=Homo sapiens OX=9606 GN=NMU PE=1 SV=1	128	NMU	227
QLLDEGLCVESTECPCVHS	L8E853	von Willebrand factor OS=Homo sapiens OX=9606 GN=VWF PE=1 SV=1	335	VWF	228
QTCVDPEDCPVCEVA	L8E853	von Willebrand factor OS=Homo sapiens OX=9606 GN=VWF PE=1 SV=1	1188	VWF	229
QPGPPGPSGEAGPPGPP	Q59HB5	Alpha 1 type XI collagen isoform C preproprotein variant (Fragment) OS=Homo sapiens OX=9606 PE=2 SV=1	554	none	230

59/88

Fig. 15A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
QASC	Q53FD8	Chemokine (C-X3-C motif) ligand 1 variant (Fragment) OS=Homo sapiens OX=9606 PE=2 SV=1	55	none	231
QVLLGGAAGLVPEL	C8C061	Bone morphogenetic protein 2 OS=Homo sapiens OX=9606 GN=BMP2 PE=2 SV=1	15	BMP2	232
QTDAPSPSYHLLP	J3KPA4	Vascular endothelial growth factor A OS=Homo sapiens OX=9606 GN=VEGFA PE=1 SV=1	5	VEGFA	233
QGPPGPSGEE	A0A384 MDU2	Collagen type I alpha 2 isoform CRA_c OS=Homo sapiens OX=9606 GN=COL1A2 PE=2 SV=1	369	COL1A2	234
QGPPGPAGKP	A0A024 R8E5	Collagen type V alpha 1 isoform CRA_a OS=Homo sapiens OX=9606 GN=COL5A1 PE=4 SV=1	597	COL5A1	235
QTGSPGPTGEPGPSGPP	A0A024 R8E5	Collagen type V alpha 1 isoform CRA_a OS=Homo sapiens OX=9606 GN=COL5A1 PE=4 SV=1	1373	COL5A1	236
QAGL	A0A0A0 MT20	EMILIN-1 (Fragment) OS=Homo sapiens OX=9606 GN=EMILIN1 PE=1 SV=1	109	EMILIN1	237
QSLKPEPF	A6H8L5	Hepatic triacylglycerol lipase OS=Homo sapiens OX=9606 GN=LIPC PE=2 SV=1	25	LIPC	238
QNGGRYCL	Q59EU4	ADAMTS-8 variant (Fragment) OS=Homo sapiens OX=9606 PE=2 SV=1	88	none	239
QGATYPVSLVAFKG	B3KXB6	cDNA FLJ45101 fis clone BRAWH3032298 highly similar to Tenascin-N OS=Homo sapiens OX=9606 PE=2 SV=1	1033	none	240
QPLSPGVG	B7ZW00	COL6A3 protein OS=Homo sapiens OX=9606 GN=COL6A3 PE=2 SV=1	614	COL6A3	241
QFVGHP	A0A494 C0S9	C3 and PZP-like alpha-2-macroglobulin domain-containing protein 8 OS=Homo sapiens OX=9606 GN=CPAMD8 PE=1 SV=1	895	CPAMD8	242
QLLLDKHMGMKRYEVA	A0A494 C0S9	C3 and PZP-like alpha-2-macroglobulin domain-containing protein 8 OS=Homo sapiens OX=9606 GN=CPAMD8 PE=1 SV=1	1638	CPAMD8	243
QFAEMKLRPSNFRNLE	D3YHM4	Tenascin-C isoform 14/AD1/16 OS=Homo sapiens OX=9606 PE=2 SV=1	1819	none	244

60/88

Fig. 15A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
QLCNIPVT	C9JD84	Latent-transforming growth factor beta-binding protein 1 OS=Homo sapiens OX=9606 GN=LTBP1 PE=1 SV=1	1194	LTBP1	245
QVTGALRSGRGKMLFS	B7Z7V0	92 kDa gelatinase OS=Homo sapiens OX=9606 PE=2 SV=1	181	none	246
QFEENWYNTYSSNLKHYVD	A0A7U3L6D0	Fibroblast growth factor OS=Homo sapiens OX=9606 GN=FGF4B PE=3 SV=1	139	FGF4B	247
QCETSM DGEAGLCWCVPWN	Q6PEY6	Insulin-like growth factor-binding protein 1 OS=Homo sapiens OX=9606 GN=IGFBP1 PE=2 SV=1	216	IGFBP1	248
QVEEVRI RPAVGW	B8ZZT6	Lysyl oxidase homolog OS=Homo sapiens OX=9606 GN=LOXL3 PE=1 SV=1	165	LOXL3	249
QHP	D6RFM1	Pro-thyrotropin-releasing hormone OS=Homo sapiens OX=9606 GN=TRH PE=1 SV=1	80	TRH	250
QPLTENPRKYVMGHFRWDRF	Q5TZZ7	Corticotropin-lipotropin OS=Homo sapiens OX=9606 PE=2 SV=1	68	none	251
QGLTGPPGKA	A0A1U9X7I1	COL11A2 OS=Homo sapiens OX=9606 PE=4 SV=1	525	none	252
QPGSPGPTGENGPPGTL	A0A1U9X7I1	COL11A2 OS=Homo sapiens OX=9606 PE=4 SV=1	1301	none	253
QFSSRDCTRPVTRNGGKYCE	A0A384ME91	Epididymis secretory sperm binding protein OS=Homo sapiens OX=9606 PE=2 SV=1	541	none	254
QILRLKPLTGEGTAGGGGG	A0A024RB20	Growth differentiation factor 11 isoform CRA_a OS=Homo sapiens OX=9606 GN=GDF11 PE=3 SV=1	196	GDF11	255
QKNMTALEKKKEAVES	A0A087WYF1	Laminin subunit alpha-2 OS=Homo sapiens OX=9606 GN=LAMA2 PE=1 SV=1	1808	LAMA2	256
QGPVGS	A0A0C4DH07	Latent-transforming growth factor beta-binding protein 4 OS=Homo sapiens OX=9606 GN=LTBP4 PE=1 SV=1	1136	LTBP4	257
QLFE	A0A3B3ISA6	Complement C4-B (Fragment) OS=Homo sapiens OX=9606 GN=C4B PE=1 SV=1	1011	C4B	258
QVMAELSPA AKT	A8YQE4	Interferon gamma OS=Homo sapiens OX=9606 GN=IFNG PE=2 SV=1	138	IFNG	259

61/88

Fig. 15A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
QMLFR	A8YQE4	Interferon gamma OS=Homo sapiens OX=9606 GN=IFNG PE=2 SV=1	156	IFNG	260
QGPTGPSGPPGV	A0A650 AXN9	Collagen type X alpha 1 chain OS=Homo sapiens OX=9606 GN=COL10A1 PE=2 SV=1	208	COL10A1	261
QLSWEW	H7C0E4	Nodal homolog (Fragment) OS=Homo sapiens OX=9606 GN=NODAL PE=1 SV=1	171	NODAL	262
QFSSRDCTRPVPRNGGKYCE	A8K6A8	cDNA FLJ76304 highly similar to Homo sapiens ADAM metalloproteinase with thrombospondin type 1 motif 4 (ADAMTS4) mRNA OS=Homo sapiens OX=9606 PE=2 SV=1	541	none	263
QSGDA	A0A2R8 Y214	Endothelin-3 OS=Homo sapiens OX=9606 GN=EDN3 PE=1 SV=1	24	EDN3	264
QTPYGLSNYRGSFR	A0A2R8 Y214	Endothelin-3 OS=Homo sapiens OX=9606 GN=EDN3 PE=1 SV=1	123	EDN3	265
QGLGGPPGLPGSAGLP	R9UH47	Collagen type VIII alpha 2 (Fragment) OS=Homo sapiens OX=9606 GN=COL8A2 PE=4 SV=1	289	COL8A2	266
QVA	B7ZMJ3	ADAMTSL4 protein OS=Homo sapiens OX=9606 GN=ADAMTSL4 PE=2 SV=1	293	ADAMTSL4	267
QRAVPEGF	Q8NE26	ADAM metalloproteinase with thrombospondin type 1 motif 1 OS=Homo sapiens OX=9606 GN=ADAMTS1 PE=2 SV=1	2	ADAMTS1	268
QYTMRECDNPVPKNGGKYCE	Q8NE26	ADAM metalloproteinase with thrombospondin type 1 motif 1 OS=Homo sapiens OX=9606 GN=ADAMTS1 PE=2 SV=1	580	ADAMTS1	269
QEVWLGSPAASDD	B7Z544	cDNA FLJ51742 highly similar to Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens OX=9606 PE=2 SV=1	838	none	270
QDVDECA	Q6ZNB8	cDNA FLJ16233 fis clone FEBRA2028256 moderately similar to FIBRILLIN 2 OS=Homo sapiens OX=9606 PE=2 SV=1	672	none	271
QGPPGPTGKP	D3DT72	Collagen type XI alpha 1 isoform CRA_c OS=Homo sapiens OX=9606 GN=COL11A1 PE=4 SV=1	579	COL11A1	272

62/88

Fig. 15A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
QSRPGCYGDRSSLPGVRSY	Q4LE67	CSPG3 variant protein (Fragment) OS=Homo sapiens OX=9606 GN=CSPG3 variant protein PE=2 SV=1	236	CSPG3	273
QFFGLM	Q9Y494	PPT (Fragment) OS=Homo sapiens OX=9606 PE=2 SV=1	28	none	274
QISHKRHKTDSEFVGLM	Q9Y494	PPT (Fragment) OS=Homo sapiens OX=9606 PE=2 SV=1	42	none	275
QVDINGFTT	A0A024 R0P8	Transforming growth factor beta OS=Homo sapiens OX=9606 GN=TGFB1 PE=3 SV=1	233	TGFB1	276
QPGSPGPTGENGPPGPL	A0A1U9 X7I9	COL11A2 OS=Homo sapiens OX=9606 PE=4 SV=1	1301	none	277
QEVCFAGE	J3KMX3	Alpha-fetoprotein OS=Homo sapiens OX=9606 GN=AFP PE=1 SV=1	588	AFP	278
QELGPPALVAPRAFAHITMA	Q658V8	Uncharacterized protein DKFZp666C182 (Fragment) OS=Homo sapiens OX=9606 GN=DKFZp666C182 PE=2 SV=1	404	DKFZP 666C1 82	279
QRGDPSRGPPIPPGRA	A0A087 WYX9	Collagen alpha-2(V) chain OS=Homo sapiens OX=9606 GN=COL5A2 PE=1 SV=1	712	COL5A 2	280
QWSPS	H0Y9E2	Fibroblast growth factor (Fragment) OS=Homo sapiens OX=9606 GN=FGF5 PE=3 SV=1	41	FGF5	281
QYVLLRRSHGDRSR	A1NY82	Left-right determination factor OS=Homo sapiens OX=9606 GN=LEFTY2 PE=3 SV=1	60	LEFTY2	282
QPLKTPGKKKKGKP	F5GZD9	Parathyroid hormone-related protein (Fragment) OS=Homo sapiens OX=9606 GN=PTH1H PE=1 SV=1	117	PTH1H	283
QGCGVISSGSSLYFSKA	A0A7P0 TA83	Reelin OS=Homo sapiens OX=9606 GN=RELN PE=1 SV=1	1096	RELN	284
QVGTGCGTLNDGKSLYFNGP	A0A7P0 TA83	Reelin OS=Homo sapiens OX=9606 GN=RELN PE=1 SV=1	1469	RELN	285
QLKEYVLTDG	A0A0K1 P7S2	Usher syndrome 2A OS=Homo sapiens OX=9606 PE=4 SV=1	4960	none	286
QMVSETVITSVGDEE	Q8N294	cDNA FLJ33633 fis clone BRAMY2022786 highly similar to Homo sapiens dickkopf-3 (DKK-3) mRNA OS=Homo sapiens OX=9606 PE=2 SV=1	170	none	287

63/88

Fig. 15A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
QRG	A8K8Z4	Complement component C6 OS=Homo sapiens OX=9606 PE=2 SV=1	595	none	288
QKPYS	B4DRD3	Plasminogen activator OS=Homo sapiens OX=9606 PE=2 SV=1	116	none	289
QESDASNFLKRR	R4GMV7	Unique cartilage matrix-associated protein OS=Homo sapiens OX=9606 GN=UCMA PE=1 SV=1	28	UCMA	290
QAAAAANPENSRGK	A0A0S2Z3T2	Glial cell line-derived neurotrophic factor (Fragment) OS=Homo sapiens OX=9606 GN=GDNF PE=2 SV=1	68	GDNF	291
QTSMGDGEAGLCWCVPWN	C9JXF9	Insulin-like growth factor-binding protein 1 OS=Homo sapiens OX=9606 GN=IGFBP1 PE=1 SV=1	216	IGFBP1	292
QNNEA	F5H7Q6	Protein Wnt (Fragment) OS=Homo sapiens OX=9606 GN=WNT5B PE=1 SV=8	200	WNT5B	293
QGEP	A0A0A0U994	Collagen type IV alpha 5 transcript variant 2 OS=Homo sapiens OX=9606 GN=COL4A5 PE=2 SV=1	291	COL4A5	294
QPLTENPRKYVMGHRWDRL	K7QX41	Pro-opiomelanocortin (Fragment) OS=Homo sapiens OX=9606 GN=POMC PE=3 SV=1	24	POMC	295
QHFC	B3KXQ2	cDNA FLJ45846 fis clone OCBBF2007039 highly similar to ADAMTS-16 OS=Homo sapiens OX=9606 PE=2 SV=1	246	none	296
QNLDPSHRISDRDYMGMWMD F	Q6FG82	CCK protein (Fragment) OS=Homo sapiens OX=9606 GN=CCK PE=2 SV=1	84	CCK	297
QKGIPGDPAFGHLGPP	J3KNM7	Collagen type IV alpha 4 isoform CRA_a OS=Homo sapiens OX=9606 GN=COL4A4 PE=4 SV=1	755	COL4A4	298
QSGGGGAGRGH	Q75RY1	Growth differentiation factor 7 OS=Homo sapiens OX=9606 GN=gdf7 PE=2 SV=1	334	GDF7	299
QAAKNGKGNVCHRKKHH	A0N0Q3	C-C motif chemokine OS=Homo sapiens OX=9606 GN=CCL28 PE=2 SV=1	88	CCL28	300
QEVDECA	A0A494C0D8	Fibrillin-3 OS=Homo sapiens OX=9606 GN=FBN3 PE=1 SV=1	2651	FBN3	301

64/88

Fig. 15A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
QPPPPDLRSLGF	A0A0S2A5D6	Growth differentiation factor 6 OS=Homo sapiens OX=9606 GN=GDF6 PE=3 SV=1	254	GDF6	302
QGPWLEEEEEAYGWMDF	A0A0E3VY36	Gastrin OS=Homo sapiens OX=9606 GN=GAS PE=3 SV=1	76	GAS	303
QFN	A0A024R035	Complement component C9 OS=Homo sapiens OX=9606 GN=C9 PE=3 SV=1	72	C9	304
QPR	K7EPJ4	Cartilage intermediate layer protein 2 OS=Homo sapiens OX=9606 GN=CILP2 PE=1 SV=1	134	CILP2	305
QVGPPGAACRRAYAAFSV	A0A3B0IWW5	Adiponectin O OS=Homo sapiens OX=9606 GN=ADIO PE=4 SV=1	104	ADIO	306
QGCLRF	Q59F13	Cat eye syndrome critical region protein 1 isoform a variant (Fragment) OS=Homo sapiens OX=9606 PE=3 SV=1	42	none	307
QVPSVEDLSEDCMFGN	A0A0F7G8J1	Plasminogen OS=Homo sapiens OX=9606 GN=PLG PE=2 SV=1	469	PLG	308
QWGGCGDNLKYSSKFVKEFL	D9ZGG3	Protein Wnt OS=Homo sapiens OX=9606 GN=WNT9A PE=3 SV=1	164	WNT9A	309
QRSHGDRSR	A0A0A0N0K8	Lefty1 isoform B OS=Homo sapiens OX=9606 GN=Lefty1 PE=2 SV=1	66	LEFTY1	310
QALRKGVRRTSPSRK	B5MCX1	Oncostatin-M OS=Homo sapiens OX=9606 GN=OSM PE=1 SV=1	205	OSM	311
QRRR	O43508	Tumor necrosis factor ligand superfamily member 12 OS=Homo sapiens OX=9606 GN=TNFSF12 PE=1 SV=1	7	TNFSF12	312
QELTLV	O75095	Multiple epidermal growth factor-like domains protein 6 OS=Homo sapiens OX=9606 GN=MEGF6 PE=1 SV=4	51	MEGF6	313
QACTCPVGVACDSVSGEC	O75095	Multiple epidermal growth factor-like domains protein 6 OS=Homo sapiens OX=9606 GN=MEGF6 PE=1 SV=4	697	MEGF6	314
QINGTGLL	Q99102	Mucin-4 OS=Homo sapiens OX=9606 GN=MUC4 PE=1 SV=5	4873	MUC4	315
QNCPRG	P01185	Vasopressin-neurophysin 2-copeptin OS=Homo sapiens OX=9606 GN=AVP PE=1 SV=2	23	AVP	316

65/88

Fig. 15A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
QVTFIHLKCDSSRKGK	Q8IX30	Signal peptide CUB and EGF-like domain-containing protein 3 OS=Homo sapiens OX=9606 GN=SCUBE3 PE=1 SV=1	520	SCUBE3	317
QNATNNGCAMASRSDGR	Q9UJW2	Tubulointerstitial nephritis antigen OS=Homo sapiens OX=9606 GN=TINAG PE=2 SV=3	313	TINAG	318
QGR	Q9GZM7	Tubulointerstitial nephritis antigen-like OS=Homo sapiens OX=9606 GN=TINAGL1 PE=1 SV=1	24	TINAGL1	319
QNASR	Q9BXY4	R-spondin-3 OS=Homo sapiens OX=9606 GN=RSPO3 PE=1 SV=1	22	RSPO3	320
QYAADLRRYINMLTRPRY	P01298	Pancreatic prohormone OS=Homo sapiens OX=9606 GN=PPY PE=1 SV=1	48	PPY	321
QYPEVGT	O60568	Multifunctional procollagen lysine hydroxylase and glycosyltransferase LH3 OS=Homo sapiens OX=9606 GN=PLOD3 PE=1 SV=1	153	PLOD3	322
QAGYM	Q8TER0	Sushi nidogen and EGF-like domain-containing protein 1 OS=Homo sapiens OX=9606 GN=SNED1 PE=2 SV=2	684	SNED1	323
QATDVDRSVDRFTFRALLP	Q8TER0	Sushi nidogen and EGF-like domain-containing protein 1 OS=Homo sapiens OX=9606 GN=SNED1 PE=2 SV=2	1051	SNED1	324
QGCVNTKGSYECVCP	Q8IWY4	Signal peptide CUB and EGF-like domain-containing protein 1 OS=Homo sapiens OX=9606 GN=SCUBE1 PE=1 SV=3	375	SCUBE1	325
QGLL	Q86UU9	Tachykinin-4 OS=Homo sapiens OX=9606 GN=TAC4 PE=1 SV=1	92	TAC4	326
QPFLVG	Q8N145	Leucine-rich repeat LGI family member 3 OS=Homo sapiens OX=9606 GN=LGI3 PE=2 SV=1	471	LGI3	327
QVLT	P04003	C4b-binding protein alpha chain OS=Homo sapiens OX=9606 GN=C4BPA PE=1 SV=2	548	C4BPA	328
QDCEYKDAKAATGECTATV	P01042	Kininogen-1 OS=Homo sapiens OX=9606 GN=KNG1 PE=1 SV=2	92	KNG1	329

66/88

Fig. 15A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
QREKDLPNYNWNSFGLRF	Q15726	Metastasis-suppressor KiSS-1 OS=Homo sapiens OX=9606 GN=KISS1 PE=1 SV=4	104	KISS1	330
QNEILTIHFTGHSFIY	P12259	Coagulation factor V OS=Homo sapiens OX=9606 GN=F5 PE=1 SV=4	618	F5	331
QCPCHGEVS	A4D0S4	Laminin subunit beta-4 OS=Homo sapiens OX=9606 GN=LAMB4 PE=2 SV=1	835	LAMB4	332
QCPCKLGYG	A4D0S4	Laminin subunit beta-4 OS=Homo sapiens OX=9606 GN=LAMB4 PE=2 SV=1	1098	LAMB4	333
QVSA	Q9NRM6	Interleukin-17 receptor B OS=Homo sapiens OX=9606 GN=IL17RB PE=1 SV=2	485	IL17RB	334
QALATPNKEEH	Q99075	Proheparin-binding EGF-like growth factor OS=Homo sapiens OX=9606 GN=HBEGF PE=1 SV=1	81	HBEGF	335
QAALPLLSNAECKKSW	Q6GPI1	Chymotrypsinogen B2 OS=Homo sapiens OX=9606 GN=CTRB2 PE=2 SV=2	175	CTRB2	336
QLMIIKNSTYREI	P24071	Immunoglobulin alpha Fc receptor OS=Homo sapiens OX=9606 GN=FCAR PE=1 SV=1	59	FCAR	337
QRRSPENS	Q15389	Angiopoietin-1 OS=Homo sapiens OX=9606 GN=ANGPT1 PE=1 SV=2	22	ANGPT1	338
QLS	P80217	Interferon-induced 35 kDa protein OS=Homo sapiens OX=9606 GN=IFI35 PE=1 SV=5	129	IFI35	339
QDI	Q641Q3	Meteorin-like protein OS=Homo sapiens OX=9606 GN=METRNL PE=2 SV=1	159	METRNL	340
QSAPPKENPSERKTNE	Q8TAT2	Fibroblast growth factor-binding protein 3 OS=Homo sapiens OX=9606 GN=FGFBP3 PE=1 SV=1	191	FGFBP3	341
QHLYYKDCSDYYAI	Q14314	Fibroleukin OS=Homo sapiens OX=9606 GN=FGL2 PE=1 SV=1	206	FGL2	342
QNCPLG	P01178	Oxytocin-neurophysin 1 OS=Homo sapiens OX=9606 GN=OXT PE=1 SV=1	23	OXT	343
QAPGRGPLGLRLMMP	Q8IU8	Cerebellin-2 OS=Homo sapiens OX=9606 GN=CBLN2 PE=2 SV=1	2	CBLN2	344

67/88

Fig. 15A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
QSPKKSVDLLGSFE	Q76M96	Coiled-coil domain-containing protein 80 OS=Homo sapiens OX=9606 GN=CCDC80 PE=1 SV=1	606	CCDC80	345
QCHCVPGVA	P25391	Laminin subunit alpha-1 OS=Homo sapiens OX=9606 GN=LAMA1 PE=1 SV=2	968	LAMA1	346
QNVTPLTFFEEVPYSELEN	P51512	Matrix metalloproteinase-16 OS=Homo sapiens OX=9606 GN=MMP16 PE=1 SV=2	160	MMP16	347
QRFS	P01732	T-cell surface glycoprotein CD8 alpha chain OS=Homo sapiens OX=9606 GN=CD8A PE=1 SV=1	87	CD8A	348
QDEELASTHP	Q6P0A1	Protein FAM180B OS=Homo sapiens OX=9606 GN=FAM180B PE=1 SV=3	69	FAM180B	349
QLYSGGGKVGYLVIYAG	Q9UNA0	A disintegrin and metalloproteinase with thrombospondin motifs 5 OS=Homo sapiens OX=9606 GN=ADAMTS5 PE=1 SV=2	79	ADAMTS5	350
QFAYRHCNNPAPRNNGRYCT	Q9UNA0	A disintegrin and metalloproteinase with thrombospondin motifs 5 OS=Homo sapiens OX=9606 GN=ADAMTS5 PE=1 SV=2	588	ADAMTS5	351
QSLAVN	Q63HQ2	Pikachurin OS=Homo sapiens OX=9606 GN=EGFLAM PE=1 SV=2	536	EGFLAM	352
QGDPHYTTFD	Q9Y6R7	IgGFc-binding protein OS=Homo sapiens OX=9606 GN=FCGBP PE=1 SV=3	475	FCGBP	353
QGSGBP HYVSFD	Q9Y6R7	IgGFc-binding protein OS=Homo sapiens OX=9606 GN=FCGBP PE=1 SV=3	865	FCGBP	354
QASGBP HYTTFD	Q9Y6R7	IgGFc-binding protein OS=Homo sapiens OX=9606 GN=FCGBP PE=1 SV=3	1253	FCGBP	355
QPGNGNEGSVTGSCYC	Q9H2A7	C-X-C motif chemokine 16 OS=Homo sapiens OX=9606 GN=CXCL16 PE=2 SV=4	25	CXCL16	356
QNISNCTYADRP	P12724	Eosinophil cationic protein OS=Homo sapiens OX=9606 GN=RNASE3 PE=1 SV=2	118	RNASE3	357

68/88

Fig. 15A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
QLGPSGHDA	A6NE02	BTB/POZ domain-containing protein 17 OS=Homo sapiens OX=9606 GN=BTBD17 PE=3 SV=1	342	BTBD17	358
QRKCIIDILTEPKDI	P09958	Furin OS=Homo sapiens OX=9606 GN=FURIN PE=1 SV=2	447	FURIN	359
QHWSHGWYPG	O43555	Progonadoliberin-2 OS=Homo sapiens OX=9606 GN=GNRH2 PE=2 SV=1	24	GNRH2	360
QCLCPPGFVLSDN	P35556	Fibrillin-2 OS=Homo sapiens OX=9606 GN=FBN2 PE=1 SV=3	2080	FBN2	361
QHWSYGLRPG	P01148	Progonadoliberin-1 OS=Homo sapiens OX=9606 GN=GNRH1 PE=1 SV=1	24	GNRH1	362
QWRRARHNYNDLCPI	O95467	Neuroendocrine secretory protein 55 OS=Homo sapiens OX=9606 GN=GNAS PE=2 SV=1	9	GNAS	363
QFEKC	P26927	Hepatocyte growth factor-like protein OS=Homo sapiens OX=9606 GN=MST1 PE=1 SV=2	464	MST1	364
QAC	Q04756	Hepatocyte growth factor activator OS=Homo sapiens OX=9606 GN=HGFA PE=1 SV=1	392	HGFA	365
QKSVWHGSDPN	P39060	Collagen alpha-1(XVIII) chain OS=Homo sapiens OX=9606 GN=COL18A1 PE=1 SV=5	1687	COL18A1	366
QVHAGHWEWIGGEPVAFTN	Q5H8C1	FRAS1-related extracellular matrix protein 1 OS=Homo sapiens OX=9606 GN=FREM1 PE=1 SV=3	2118	FREM1	367
QLSRSHSSGSVLPLGELE	P09603	Macrophage colony-stimulating factor 1 OS=Homo sapiens OX=9606 GN=CSF1 PE=1 SV=2	424	CSF1	368
QESSWGKKNVSSTESGRLFK	P00451	Coagulation factor VIII OS=Homo sapiens OX=9606 GN=F8 PE=1 SV=1	975	F8	369
QRPEDAEDGAEGG	Q08J23	RNA cytosine C(5)-methyltransferase NSUN2 OS=Homo sapiens OX=9606 GN=NSUN2 PE=1 SV=2	14	NSUN2	370
QHDGSGNDCEPV	Q9UKP4	A disintegrin and metalloproteinase with thrombospondin motifs 7 OS=Homo sapiens OX=9606 GN=ADAMTS7 PE=1 SV=2	397	ADAMTS7	371

69/88

Fig. 15A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
QIECVCP	Q6UXH9	Inactive serine protease PAMR1 OS=Homo sapiens OX=9606 GN=PAMR1 PE=1 SV=1	50	PAMR1	372
QSLDTPGASCGI	Q8N1E2	Lysozyme g-like protein 1 OS=Homo sapiens OX=9606 GN=LYG1 PE=1 SV=1	29	LYG1	373
QLHILLEVRDHL	Q6UWF7	NXPE family member 4 OS=Homo sapiens OX=9606 GN=NXPE4 PE=2 SV=1	116	NXPE4	374
QVT	O60258	Fibroblast growth factor 17 OS=Homo sapiens OX=9606 GN=FGF17 PE=1 SV=1	67	FGF17	375
QCKCRPNVT	O15230	Laminin subunit alpha-5 OS=Homo sapiens OX=9606 GN=LAMA5 PE=1 SV=8	1549	LAMA5	376
QVR	P48357	Leptin receptor OS=Homo sapiens OX=9606 GN=LEPR PE=1 SV=2	307	LEPR	377
QRRDLFLAKHILHCKAKY	Q86VR8	Four-jointed box protein 1 OS=Homo sapiens OX=9606 GN=FJX1 PE=2 SV=1	414	FJX1	378
QKVEEVKPLE	O00533	Neural cell adhesion molecule L1-like protein OS=Homo sapiens OX=9606 GN=CHL1 PE=1 SV=4	458	CHL1	379
QNPDAPE	P53621	Coatomer subunit alpha OS=Homo sapiens OX=9606 GN=COPA PE=1 SV=2	403	COPA	380
QFEENWYNTYSSNIYKHGDT	Q9NP95	Fibroblast growth factor 20 OS=Homo sapiens OX=9606 GN=FGF20 PE=1 SV=1	142	FGF20	381
QPVCFP RPR	O00182	Galectin-9 OS=Homo sapiens OX=9606 GN=LGALS9 PE=1 SV=2	166	LGALS9	382
QDGKIVDDVWVVRPLF	Q6ZNF0	Acid phosphatase type 7 OS=Homo sapiens OX=9606 GN=ACP7 PE=2 SV=2	417	ACP7	383
QCYSGYALAED	O00339	Matrilin-2 OS=Homo sapiens OX=9606 GN=MATN2 PE=1 SV=4	304	MATN2	384
QVEVNITEFPYTPSPPEH	Q15223	Nectin-1 OS=Homo sapiens OX=9606 GN=NECTIN1 PE=1 SV=3	328	NECTIN1	385
QPTFTTRLDLDAAGAGA	Q9BZZ2	Sialoadhesin OS=Homo sapiens OX=9606 GN=SIGLEC1 PE=1 SV=2	600	SIGLEC1	386
QKCP	Q92954	Proteoglycan 4 OS=Homo sapiens OX=9606 GN=PRG4 PE=1 SV=3	1279	PRG4	387

70/88

Fig. 15A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
QVDSEEGLLST	Q9NPG4	Protocadherin-12 OS=Homo sapiens OX=9606 GN=PCDH12 PE=1 SV=1	78	PCDH12	388
QSSGENCDVVVNTL	P01833	Polymeric immunoglobulin receptor OS=Homo sapiens OX=9606 GN=PIGR PE=1 SV=4	273	PIGR	389
QEESLFLISGKPE	P10163	Basic salivary proline-rich protein 4 OS=Homo sapiens OX=9606 GN=PRB4 PE=1 SV=4	25	PRB4	390
QASVENGNA	P08294	Extracellular superoxide dismutase [Cu-Zn] OS=Homo sapiens OX=9606 GN=SOD3 PE=1 SV=2	193	SOD3	391
QEESPSVISGKPE	Q04118	Basic salivary proline-rich protein 3 OS=Homo sapiens OX=9606 GN=PRB3 PE=1 SV=2	25	PRB3	392
QVSCPYDSMKHW	Q9NZC2	Triggering receptor expressed on myeloid cells 2 OS=Homo sapiens OX=9606 GN=TREM2 PE=1 SV=1	33	TREM2	393
QERRKRKRKRKKEVPFF	Q9HC23	Prokineticin-2 OS=Homo sapiens OX=9606 GN=PROK2 PE=1 SV=2	82	PROK2	394
QRGGPCTNHEPPALEEG	A0A3B3IS91	POLG alternative reading frame OS=Homo sapiens OX=9606 GN=POLG PE=1 SV=2	11	POLG	395
QREGTREHLRERAKG	Q8N3Z0	Inactive serine protease 35 OS=Homo sapiens OX=9606 GN=PRSS35 PE=1 SV=2	215	PRSS35	396
QGLV	P09683	Secretin OS=Homo sapiens OX=9606 GN=SCT PE=1 SV=2	51	SCT	397
QASPPRRL	Q86V25	Tubulinyl-Tyr carboxypeptidase 2 OS=Homo sapiens OX=9606 GN=VASH2 PE=1 SV=2	319	VASH2	398
QDEGILPS	O75462	Cytokine receptor-like factor 1 OS=Homo sapiens OX=9606 GN=CRLF1 PE=1 SV=1	404	CRLF1	399
QEVFLKTKDDAINGDKKAKK	Q9NR99	Matrix-remodeling-associated protein 5 OS=Homo sapiens OX=9606 GN=MXRA5 PE=1 SV=3	720	MXRA5	400
QKVAPSSTMSTHPSRRRPN	Q9NR99	Matrix-remodeling-associated protein 5 OS=Homo sapiens OX=9606 GN=MXRA5 PE=1 SV=3	1134	MXRA5	401
QLEMEKNAEPTSKGTPRRKH	Q9NR99	Matrix-remodeling-associated protein 5 OS=Homo sapiens OX=9606 GN=MXRA5 PE=1 SV=3	1215	MXRA5	402

71/88

Fig. 15A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
QRKPRPLDPGEGSKDTS	Q9H1B5	Xylosyltransferase 2 OS=Homo sapiens OX=9606 GN=XYLT2 PE=1 SV=2	49	XYLT2	403
QDSLFYSSNEFEY	Q12884	Prolyl endopeptidase FAP OS=Homo sapiens OX=9606 GN=FAP PE=1 SV=5	405	FAP	404
QEEKNKDNIGFHHL	Q9NQ76	Matrix extracellular phosphoglycoprotein OS=Homo sapiens OX=9606 GN=MEPE PE=1 SV=1	37	MEPE	405
QFR	P56704	Protein Wnt-3a OS=Homo sapiens OX=9606 GN=WNT3A PE=1 SV=2	80	WNT3A	406
QFNAFSYHFR	Q92626	Peroxidasin homolog OS=Homo sapiens OX=9606 GN=PXDN PE=1 SV=2	1324	PXDN	407
QMAVKKYLA	P18509	Pituitary adenylate cyclase-activating polypeptide OS=Homo sapiens OX=9606 GN=ADCYAP1 PE=1 SV=3	147	ADCYAP1	408
QRVKNK	P18509	Pituitary adenylate cyclase-activating polypeptide OS=Homo sapiens OX=9606 GN=ADCYAP1 PE=1 SV=3	164	ADCYAP1	409
QVLETVA	P17948	Vascular endothelial growth factor receptor 1 OS=Homo sapiens OX=9606 GN=FLT1 PE=1 SV=2	342	FLT1	410
QDAA	Q9Y334	von Willebrand factor A domain-containing protein 7 OS=Homo sapiens OX=9606 GN=VWA7 PE=2 SV=4	705	VWA7	411
QVEIL	Q9BS86	Zona pellucida-binding protein 1 OS=Homo sapiens OX=9606 GN=ZPBP PE=2 SV=1	273	ZPBP	412
QEKREIGDEENSAKFPI	P20382	Pro-MCH OS=Homo sapiens OX=9606 GN=PMCH PE=1 SV=3	127	PMCH	413
QATRDA	P26022	Pentraxin-related protein PTX3 OS=Homo sapiens OX=9606 GN=PTX3 PE=1 SV=3	122	PTX3	414
QRRPYRRRRFPYYMRRPY	P67809	Y-box-binding protein 1 OS=Homo sapiens OX=9606 GN=YBX1 PE=1 SV=3	184	YBX1	415
QPGFS	Q4LDE5	Sushi von Willebrand factor type A EGF and pentraxin domain-containing protein 1 OS=Homo	1410	SVEP1	416

72/88

Fig. 15A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
		sapiens OX=9606 GN=SVEP1 PE=1 SV=3			
QLELPCDPSTP	Q6W4X9	Mucin-6 OS=Homo sapiens OX=9606 GN=MUC6 PE=1 SV=3	2405	MUC6	417
QVGLPILYFS	Q9BXP8	Pappalysin-2 OS=Homo sapiens OX=9606 GN=PAPPA2 PE=1 SV=4	257	PAPPA2	418
QSIVCT	Q9BXP8	Pappalysin-2 OS=Homo sapiens OX=9606 GN=PAPPA2 PE=1 SV=4	1710	PAPPA2	419
QLAD	P35443	Thrombospondin-4 OS=Homo sapiens OX=9606 GN=THBS4 PE=1 SV=2	118	THBS4	420
QVEA	O15240	Neurosecretory protein VGF OS=Homo sapiens OX=9606 GN=VGF PE=1 SV=2	303	VGF	421
QPTRRPRPGTGP	P51888	Prolargin OS=Homo sapiens OX=9606 GN=PRELP PE=1 SV=1	21	PRELP	422
QPKK	Q8IW00	V-set and transmembrane domain-containing protein 4 OS=Homo sapiens OX=9606 GN=VSTM4 PE=1 SV=3	238	VSTM4	423
QLSAKKYLESIM	P01282	VIP peptides OS=Homo sapiens OX=9606 GN=VIP PE=1 SV=1	96	VIP	424
QMAVKKYLSILN	P01282	VIP peptides OS=Homo sapiens OX=9606 GN=VIP PE=1 SV=1	140	VIP	425
QKLPK	Q7Z7G0	Target of Nesh-SH3 OS=Homo sapiens OX=9606 GN=ABI3BP PE=1 SV=2	22	ABI3BP	426
QGVVTVFKEAPRYP	Q9NS62	Thrombospondin type-1 domain-containing protein 1 OS=Homo sapiens OX=9606 GN=THSD1 PE=1 SV=1	269	THSD1	427
QSPEPSPAPGP	Q504Y2	Extracellular tyrosine-protein kinase PKDCC OS=Homo sapiens OX=9606 GN=PKDCC PE=1 SV=2	37	PKDCC	428
QARSSSYSGEYGGGG	O60844	Zymogen granule membrane protein 16 OS=Homo sapiens OX=9606 GN=ZG16 PE=1 SV=3	20	ZG16	429
QDVPLAALE	P01266	Thyroglobulin OS=Homo sapiens OX=9606 GN=TG PE=1 SV=5	843	TG	430
QKDRGSGKAFCVDGE	P01266	Thyroglobulin OS=Homo sapiens OX=9606 GN=TG PE=1 SV=5	1533	TG	431
QETLANSSIGRLCLAFL	Q8N2E2	von Willebrand factor D and EGF domain-containing protein	811	VWDE	432

73/88

Fig. 15A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
		OS=Homo sapiens OX=9606 GN=VWDE PE=2 SV=4			
QRY	P10082	Peptide YY OS=Homo sapiens OX=9606 GN=PYY PE=1 SV=3	62	PYY	433
QAGGKP	Q8IUX7	Adipocyte enhancer-binding protein 1 OS=Homo sapiens OX=9606 GN=AEBP1 PE=1 SV=1	70	AEBP1	434
QYGEPREG	Q8NDZ4	Divergent protein kinase domain 2A OS=Homo sapiens OX=9606 GN=DIPK2A PE=1 SV=1	95	DIPK2A	435
QLADDGLVALDWVVGPCVR	Q6UXT9	Protein ABHD15 OS=Homo sapiens OX=9606 GN=ABHD15 PE=1 SV=2	128	ABHD15	436
QRLAAK	P82987	ADAMTS-like protein 3 OS=Homo sapiens OX=9606 GN=ADAMTSL3 PE=1 SV=4	847	ADAMTSL3	437
QSALIGKN	Q8TE60	A disintegrin and metalloproteinase with thrombospondin motifs 18 OS=Homo sapiens OX=9606 GN=ADAMTS18 PE=1 SV=3	372	ADAMTS18	438
QAVSTAHPPT	O43405	Cochlin OS=Homo sapiens OX=9606 GN=COCH PE=1 SV=1	136	COCH	439
QVL	O76093	Fibroblast growth factor 18 OS=Homo sapiens OX=9606 GN=FGF18 PE=1 SV=1	67	FGF18	440
QNLAKAETKLHTSNSG	Q8WU39	Marginal zone B- and B1-cell-specific protein OS=Homo sapiens OX=9606 GN=MZB1 PE=1 SV=1	62	MZB1	441
QDC	Q9Y5Q5	Atrial natriuretic peptide-converting enzyme OS=Homo sapiens OX=9606 GN=CORIN PE=1 SV=2	788	CORIN	442
QVYA	Q9UMD9	Collagen alpha-1(XVII) chain OS=Homo sapiens OX=9606 GN=COL17A1 PE=1 SV=3	1482	COL17A1	443
QGKTVI	Q8NFW1	Collagen alpha-1(XXII) chain OS=Homo sapiens OX=9606 GN=COL22A1 PE=2 SV=2	399	COL22A1	444
QVTYVRS	Q5JXM2	Probable methyltransferase-like protein 24 OS=Homo sapiens OX=9606 GN=METTL24 PE=2 SV=1	72	METTL24	445
QAAKEFIAWLVKGR	P01275	Pro-glucagon OS=Homo sapiens OX=9606 GN=GCG PE=1 SV=3	114	GCG	446

74/88

Fig. 15A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
QPGSRRGG	Q96JK4	HHIP-like protein 1 OS=Homo sapiens OX=9606 GN=HHIPL1 PE=1 SV=2	647	HHIPL1	447
QKLY	P09237	Matrilysin OS=Homo sapiens OX=9606 GN=MMP7 PE=1 SV=1	255	MMP7	448
QISRSRFLDSD	Q14667	Protein KIAA0100 OS=Homo sapiens OX=9606 GN=KIAA0100 PE=1 SV=3	158	KIAA0100	449
QGY	P35318	Pro-adrenomedullin OS=Homo sapiens OX=9606 GN=ADM PE=1 SV=1	144	ADM	450
QSCGPGRGPV	O43323	Desert hedgehog protein OS=Homo sapiens OX=9606 GN=DHH PE=1 SV=1	21	DHH	451
QGPPGPAGEE	P02452	Collagen alpha-1(I) chain OS=Homo sapiens OX=9606 GN=COL1A1 PE=1 SV=6	457	COL1A1	452
QGPPGSRGPPGMRGAK	Q8IZC6	Collagen alpha-1(XXVII) chain OS=Homo sapiens OX=9606 GN=COL27A1 PE=1 SV=1	1047	COL27A1	453
QGKAGAP	Q8IZC6	Collagen alpha-1(XXVII) chain OS=Homo sapiens OX=9606 GN=COL27A1 PE=1 SV=1	1408	COL27A1	454
QGDDGDPGPMGPA	Q8IZC6	Collagen alpha-1(XXVII) chain OS=Homo sapiens OX=9606 GN=COL27A1 PE=1 SV=1	1459	COL27A1	455
QLGPP	Q8IZC6	Collagen alpha-1(XXVII) chain OS=Homo sapiens OX=9606 GN=COL27A1 PE=1 SV=1	1500	COL27A1	456
QGAPGPAGEE	P02458	Collagen alpha-1(II) chain OS=Homo sapiens OX=9606 GN=COL2A1 PE=1 SV=3	479	COL2A1	457
QGPIGPPGEE	P05997	Collagen alpha-2(V) chain OS=Homo sapiens OX=9606 GN=COL5A2 PE=1 SV=3	491	COL5A2	458
QGGKGDPRP	P12110	Collagen alpha-2(VI) chain OS=Homo sapiens OX=9606 GN=COL6A2 PE=1 SV=4	369	COL6A2	459
QGSP	A6NMZ7	Collagen alpha-6(VI) chain OS=Homo sapiens OX=9606 GN=COL6A6 PE=1 SV=2	1490	COL6A6	460
QASLYLLEAT	A8K7I4	Calcium-activated chloride channel regulator 1 OS=Homo	59	CLCA1	461

75/88

Fig. 15A (cont.)

Sequence	Uniprot	Fasta Header	Start Amino Acid	Gene	Seq ID
		sapiens OX=9606 GN=CLCA1 PE=1 SV=3			
QGVKGHA	Q14055	Collagen alpha-2(IX) chain OS=Homo sapiens OX=9606 GN=COL9A2 PE=1 SV=2	189	COL9A2	462
QIGNTGPK	Q14055	Collagen alpha-2(IX) chain OS=Homo sapiens OX=9606 GN=COL9A2 PE=1 SV=2	590	COL9A2	463
QDY	O75596	C-type lectin domain family 3 member A OS=Homo sapiens OX=9606 GN=CLEC3A PE=1 SV=1	115	CLEC3A	464
QPNG	O75596	C-type lectin domain family 3 member A OS=Homo sapiens OX=9606 GN=CLEC3A PE=1 SV=1	158	CLEC3A	465
QVLDT	Q8N129	Protein canopy homolog 4 OS=Homo sapiens OX=9606 GN=CNPY4 PE=1 SV=1	66	CNPY4	466
QAAR	Q99988	Growth/differentiation factor 15 OS=Homo sapiens OX=9606 GN=GDF15 PE=1 SV=3	187	GDF15	467
QVK	P06396	Gelsolin OS=Homo sapiens OX=9606 GN=GSN PE=1 SV=1	191	GSN	468
QGPKGEKGAN	Q6ZMI3	Gliomedin OS=Homo sapiens OX=9606 GN=GLDN PE=1 SV=1	165	GLDN	469
QPVYKPYLTICA	Q99944	Epidermal growth factor-like protein 8 OS=Homo sapiens OX=9606 GN=EGFL8 PE=1 SV=1	55	EGFL8	470
QNTEYEDEKTLKNPKYKDRA	Q8N302	Angiogenic factor with G patch and FHA domains 1 OS=Homo sapiens OX=9606 GN=AGGF1 PE=1 SV=2	572	AGGF1	471
QRHLAGLAV	Q9Y6C2	EMILIN-1 OS=Homo sapiens OX=9606 GN=EMILIN1 PE=1 SV=3	337	EMILIN1	472
QECCSPEL	Q9Y6C2	EMILIN-1 OS=Homo sapiens OX=9606 GN=EMILIN1 PE=1 SV=3	351	EMILIN1	473
QPEPAPPRPSGPATAEDP	Q9BXX0	EMILIN-2 OS=Homo sapiens OX=9606 GN=EMILIN2 PE=1 SV=3	803	EMILIN2	474
QCYNFGWYPESPVCE	P05160	Coagulation factor XIII B chain OS=Homo sapiens OX=9606 GN=F13B PE=1 SV=3	254	F13B	475
QKWKYNR	O43155	Leucine-rich repeat transmembrane protein FLRT2 OS=Homo sapiens OX=9606 GN=FLRT2 PE=1 SV=1	575	FLRT2	476

76/88

Fig. 15B

Capped peptide names	Sequence	Precursor Uniprot	Precursor Fasta header	Seq ID
CAP-FJX1	QRRDLFLAKHILHCKAKY	Q86VR8	Four-jointed box protein 1 OS=Homo sapiens OX=9606 GN=FJX1 PE=2 SV=1	378
CAP-MMP16	QNVTPLTFFEEVPYSELEN	P51512	Matrix metalloproteinase-16 OS=Homo sapiens OX=9606 GN=MMP16 PE=1 SV=2	347
CAP-CPAMD8-1	QLLLDKHMGMKRYEVA	A0A494C0S9	C3 and PZP-like alpha-2-macroglobulin domain-containing protein 8 OS=Homo sapiens OX=9606 GN=CPAMD8 PE=1 SV=1	243
CAP-KISS1	QREKDLPNYNWNSFGLRF	Q15726	Metastasis-suppressor KiSS-1 OS=Homo sapiens OX=9606 GN=KISS1 PE=1 SV=4	330
CAP-TRH	QHP	D6RFM1	Pro-thyrotropin-releasing hormone OS=Homo sapiens OX=9606 GN=TRH PE=1 SV=1	250
CAP-WNT9A	QWGGCGDNLKYSSKFVKEFL	D9ZGG3	Protein Wnt OS=Homo sapiens OX=9606 GN=WNT9A PE=3 SV=1	309
CAP-OSM	QALRKGVRRTRPSRK	B5MCX1	Oncostatin-M OS=Homo sapiens OX=9606 GN=OSM PE=1 SV=1	311
CAP-PMCH	QEKREIGDEENSAKFPI	P20382	Pro-MCH OS=Homo sapiens OX=9606 GN=PMCH PE=1 SV=3	413
CAP-BMP4	QGHAG	A0A097PIG0	Bone morphogenetic protein 4 (Fragment) OS=Homo sapiens OX=9606 GN=BMP4 PE=3 SV=1	226
CAP-EDN3	QTVPYGLSNYRGSFR	A0A2R8Y214	Endothelin-3 OS=Homo sapiens OX=9606 GN=EDN3 PE=1 SV=1	265
CAP-FGL2	QHLYKDCSDYYAI	Q14314	Fibroleukin OS=Homo sapiens OX=9606 GN=FGL2 PE=1 SV=1	342
CAP-PRSS36	QILPGMLCAGYPE	Q5K4E3	Polyserase-2 OS=Homo sapiens OX=9606 GN=PRSS36 PE=1 SV=2	224
CAP-COL11A1-1	QPGPPGPSGEAGPPGPP	Q59HB5	Alpha 1 type XI collagen isoform C preproprotein variant (Fragment) OS=Homo sapiens OX=9606 PE=2 SV=1	230
CAP-SOD3	QASVENGNA	P08294	Extracellular superoxide dismutase [Cu-Zn] OS=Homo sapiens OX=9606 GN=SOD3 PE=1 SV=2	391

77/88

Fig. 15B (cont.)

Capped peptide names	Sequence	Precursor Uniprot	Precursor Fasta header	Seq ID
CAP-LAMA3	QCTRCATGHYGFPRCKPCSC	Q16787	Laminin subunit alpha-3 OS=Homo sapiens OX=9606 GN=LAMA3 PE=1 SV=3	217
CAP-FGF5	QWSPS	H0Y9E2	Fibroblast growth factor (Fragment) OS=Homo sapiens OX=9606 GN=FGF5 PE=3 SV=1	281
CAP-VIP	QMAVKKYLSILN	P01282	VIP peptides OS=Homo sapiens OX=9606 GN=VIP PE=1 SV=1	425
CAP-LEFTY2	QYVVLLRRSHGDRSR	A1NY82	Left-right determination factor OS=Homo sapiens OX=9606 GN=LEFTY2 PE=3 SV=1	282
CAP-FURIN	QRKCIIDILTEPKDI	P09958	Furin OS=Homo sapiens OX=9606 GN=FURIN PE=1 SV=2	359
CAP-COL6A3	QPLSPGVG	B7ZW00	COL6A3 protein OS=Homo sapiens OX=9606 GN=COL6A3 PE=2 SV=1	241
CAP-ADIO	QVGPPGAACRRAYAAFSV	A0A3B0I WW5	Adiponectin O OS=Homo sapiens OX=9606 GN=ADIO PE=4 SV=1	306
CAP-LGI3	QPFLVG	Q8N145	Leucine-rich repeat LGI family member 3 OS=Homo sapiens OX=9606 GN=LGI3 PE=2 SV=1	327
CAP-CRLF1	QDEGILPS	O75462	Cytokine receptor-like factor 1 OS=Homo sapiens OX=9606 GN=CRLF1 PE=1 SV=1	399
CAP-FGF18	QVL	O76093	Fibroblast growth factor 18 OS=Homo sapiens OX=9606 GN=FGF18 PE=1 SV=1	440
CAP-PPY	QYAADLRRYINMLTRPRY	P01298	Pancreatic prohormone OS=Homo sapiens OX=9606 GN=PPY PE=1 SV=1	321
CAP-CILP2	QPR	K7EPJ4	Cartilage intermediate layer protein 2 OS=Homo sapiens OX=9606 GN=CILP2 PE=1 SV=1	305
CAP-PRG4	QKCP	Q92954	Proteoglycan 4 OS=Homo sapiens OX=9606 GN=PRG4 PE=1 SV=3	387
CAP-LGALS9	QPVCFPPRR	O00182	Galectin-9 OS=Homo sapiens OX=9606 GN=LGALS9 PE=1 SV=2	382
CAP-IFNG-1	QVMAELSPAAKT	A8YQE4	Interferon gamma OS=Homo sapiens OX=9606 GN=IFNG PE=2 SV=1	259

78/88

Fig. 15B (cont.)

Capped peptide names	Sequence	Precursor Uniprot	Precursor Fasta header	Seq ID
CAP-PAPPA2	QVGLPILYFS	Q9BXP8	Pappalysin-2 OS=Homo sapiens OX=9606 GN=PAPPA2 PE=1 SV=4	418
CAP-F5	QNEILTIHFTGHSFIY	P12259	Coagulation factor V OS=Homo sapiens OX=9606 GN=F5 PE=1 SV=4	331
CAP-EMILIN1	QAGL	A0A0A0MT20	EMILIN-1 (Fragment) OS=Homo sapiens OX=9606 GN=EMILIN1 PE=1 SV=1	237
CAP-SCT	QGLV	P09683	Secretin OS=Homo sapiens OX=9606 GN=SCT PE=1 SV=2	397
CAP-GNAS	QWRRARHNYNDLCPPI	O95467	Neuroendocrine secretory protein 55 OS=Homo sapiens OX=9606 GN=GNAS PE=2 SV=1	363
CAP-METRNL	QDI	Q641Q3	Meteorin-like protein OS=Homo sapiens OX=9606 GN=METRNL PE=2 SV=1	340
CAP-BMP2	QVLLGGAAGLVPEL	C8C061	Bone morphogenetic protein 2 OS=Homo sapiens OX=9606 GN=BMP2 PE=2 SV=1	232
CAP-POMC	QPLTENPRKYVMGHFRWDRL	K7QX41	Pro-opiomelanocortin (Fragment) OS=Homo sapiens OX=9606 GN=POMC PE=3 SV=1	295
CAP-GDF15	QAAR	Q99988	Growth/differentiation factor 15 OS=Homo sapiens OX=9606 GN=GDF15 PE=1 SV=3	467
CAP-VEGFA	QTDTPSPSYHLLP	J3KPA4	Vascular endothelial growth factor A OS=Homo sapiens OX=9606 GN=VEGFA PE=1 SV=1	233
CAP-VWF	QTCVDPEDCPVCEVA	L8E853	von Willebrand factor OS=Homo sapiens OX=9606 GN=VWF PE=1 SV=1	229
CAP-MEGF6-1	QACTCPVGACDSVSGEC	O75095	Multiple epidermal growth factor-like domains protein 6 OS=Homo sapiens OX=9606 GN=MEGF6 PE=1 SV=4	314
CAP-GDNF	QAAAANPENSRGK	A0A0S2Z3T2	Glial cell line-derived neurotrophic factor (Fragment) OS=Homo sapiens OX=9606 GN=GDNF PE=2 SV=1	291

79/88

Fig. 15B (cont.)

Capped peptide names	Sequence	Precursor Uniprot	Precursor Fasta header	Seq ID
CAP-BRAWH3032298	QGATYPVSLVAFKG	B3KXB6	cDNA FLJ45101 fis clone BRAWH3032298 highly similar to Tenascin-N OS=Homo sapiens OX=9606 PE=2 SV=1	240
CAP-HHIPL1	QPGSRRGG	Q96JK4	HHIP-like protein 1 OS=Homo sapiens OX=9606 GN=HHIPL1 PE=1 SV=2	447
CAP-LAMB4	QCPCCKLGYG	A4D0S4	Laminin subunit beta-4 OS=Homo sapiens OX=9606 GN=LAMB4 PE=2 SV=1	333
CAP-MEGF6-2	QELTLV	O75095	Multiple epidermal growth factor-like domains protein 6 OS=Homo sapiens OX=9606 GN=MEGF6 PE=1 SV=4	313
CAP-COL4A4	QKGIPGDPAFGHLGPP	J3KNM7	Collagen type IV alpha 4 isoform CRA_a OS=Homo sapiens OX=9606 GN=COL4A4 PE=4 SV=1	298
CAP-LYG1	QSLDTPGASCGI	Q8N1E2	Lysozyme g-like protein 1 OS=Homo sapiens OX=9606 GN=LYG1 PE=1 SV=1	373
CAP-F8	QESSWGKNVSTESGRLFK	P00451	Coagulation factor VIII OS=Homo sapiens OX=9606 GN=F8 PE=1 SV=1	369
CAP-SIGLEC1	QPTFTTRLDLDAAGAGA	Q9BZZ2	Sialoadhesin OS=Homo sapiens OX=9606 GN=SIGLEC1 PE=1 SV=2	386
CAP-THBS4	QLAD	P35443	Thrombospondin-4 OS=Homo sapiens OX=9606 GN=THBS4 PE=1 SV=2	420
CAP-PTX3	QATRDA	P26022	Pentraxin-related protein PTX3 OS=Homo sapiens OX=9606 GN=PTX3 PE=1 SV=3	414
CAP-ADAMTS5	QLYSGGGKVGVLVYAG	Q9UNAO	A disintegrin and metalloproteinase with thrombospondin motifs 5 OS=Homo sapiens OX=9606 GN=ADAMTS5 PE=1 SV=2	350
CAP-COL8A2	QGLGGPPGLPGSAGLP	R9UH47	Collagen type VIII alpha 2 (Fragment) OS=Homo sapiens OX=9606 GN=COL8A2 PE=4 SV=1	266
CAP-PPT	QFFGLM	Q9Y494	PPT (Fragment) OS=Homo sapiens OX=9606 PE=2 SV=1	274

80/88

Fig. 15B (cont.)

Capped peptide names	Sequence	Precursor Uniprot	Precursor Fasta header	Seq ID
CAP-CSF1	QLSRSHSSGSVLPLGELE	P09603	Macrophage colony-stimulating factor 1 OS=Homo sapiens OX=9606 GN=CSF1 PE=1 SV=2	368
CAP-COL11A1-2	QPGSPGPTGENGPPGPL	A0A1U9X7I9	COL11A2 OS=Homo sapiens OX=9606 PE=4 SV=1	277
CAP-CCDC80	QSPKKSVDLLGSFE	Q76M96	Coiled-coil domain-containing protein 80 OS=Homo sapiens OX=9606 GN=CCDC80 PE=1 SV=1	345
CAP-MUC4	QINGTGLL	Q99102	Mucin-4 OS=Homo sapiens OX=9606 GN=MUC4 PE=1 SV=5	315
CAP-AFP	QEVCFAGE	J3KMX3	Alpha-fetoprotein OS=Homo sapiens OX=9606 GN=AFP PE=1 SV=1	278
CAP-TG	QDVPLAALE	P01266	Thyroglobulin OS=Homo sapiens OX=9606 GN=TG PE=1 SV=5	430
CAP-FCGBP	QGDPHYTTFD	Q9Y6R7	IgGFC-binding protein OS=Homo sapiens OX=9606 GN=FCGBP PE=1 SV=3	353
CAP-COL11A2	QPGSPGPTGENGPPGTL	A0A1U9X7I1	COL11A2 OS=Homo sapiens OX=9606 PE=4 SV=1	253
CAP-DKFZP666C182	QELGPPALVAPRAFAHITMA	Q658V8	Uncharacterized protein DKFZp666C182 (Fragment) OS=Homo sapiens OX=9606 GN=DKFZp666C182 PE=2 SV=1	279
CAP-ADAMTS1	QRAVPEGF	Q8NE26	ADAM metalloproteinase with thrombospondin type 1 motif 1 OS=Homo sapiens OX=9606 GN=ADAMTS1 PE=2 SV=1	268
CAP-NXPE4	QLHILLEVRDHL	Q6UWF7	NXPE family member 4 OS=Homo sapiens OX=9606 GN=NXPE4 PE=2 SV=1	374
CAP-COL27A1	QLGPP	Q8IZC6	Collagen alpha-1(XVII) chain OS=Homo sapiens OX=9606 GN=COL27A1 PE=1 SV=1	456
CAP-FAP	QDSLFISSNEFEEYP	Q12884	Prolyl endopeptidase FAP OS=Homo sapiens OX=9606 GN=FAP PE=1 SV=5	404
CAP-COL18A1	QKSVWHGSDPN	P39060	Collagen alpha-1(XVIII) chain OS=Homo sapiens OX=9606 GN=COL18A1 PE=1 SV=5	366

81/88

Fig. 15B (cont.)

Capped peptide names	Sequence	Precursor Uniprot	Precursor Fasta header	Seq ID
CAP-Q59F13	QGCLRF	Q59F13	Cat eye syndrome critical region protein 1 isoform a variant (Fragment) OS=Homo sapiens OX=9606 PE=3 SV=1	307
CAP-CNPY4	QVLDT	Q8N129	Protein canopy homolog 4 OS=Homo sapiens OX=9606 GN=CNPY4 PE=1 SV=1	466
CAP-PIGR	QSSGENCDVVNTL	P01833	Polymeric immunoglobulin receptor OS=Homo sapiens OX=9606 GN=PIGR PE=1 SV=4	389
CAP-LIPC	QSLKPEPF	A6H8L5	Hepatic triacylglycerol lipase OS=Homo sapiens OX=9606 GN=LIPC PE=2 SV=1	238
CAP-PAMR1	QIECVCP	Q6UXH9	Inactive serine protease PAMR1 OS=Homo sapiens OX=9606 GN=PAMR1 PE=1 SV=1	372
CAP-CBLN2	QAPGRGPLGLRLMMP	Q8IUK8	Cerebellin-2 OS=Homo sapiens OX=9606 GN=CBLN2 PE=2 SV=1	344
CAP-GNRH1	QHWSYGLRPG	P01148	Progonadoliberin-1 OS=Homo sapiens OX=9606 GN=GNRH1 PE=1 SV=1	362
CAP-NMU	QSRGYFLFRPN	D6RBC9	Neuromedin U isoform CRA_a OS=Homo sapiens OX=9606 GN=NMU PE=1 SV=1	227
CAP-SNED1	QATDVDRSVDRFTFRALLP	Q8TER0	Sushi nidogen and EGF-like domain-containing protein 1 OS=Homo sapiens OX=9606 GN=SNED1 PE=2 SV=2	324
CAP-IFNG-2	QMLFR	A8YQE4	Interferon gamma OS=Homo sapiens OX=9606 GN=IFNG PE=2 SV=1	260
CAP-LAMB2	QCLCKPGVV	P55268	Laminin subunit beta-2 OS=Homo sapiens OX=9606 GN=LAMB2 PE=1 SV=2	219
CAP-ANGPT1	QRRSPENS	Q15389	Angiopoietin-1 OS=Homo sapiens OX=9606 GN=ANGPT1 PE=1 SV=2	338
CAP-THSD1	QGVVTVFKEAPRYP	Q9NS62	Thrombospondin type-1 domain-containing protein 1 OS=Homo sapiens OX=9606 GN=THSD1 PE=1 SV=1	427

Fig. 15B (cont.)

Capped peptide names	Sequence	Precursor Uniprot	Precursor Fasta header	Seq ID
CAP-CPAMD8-2	QFVGHP	A0A494C0S9	C3 and PZP-like alpha-2-macroglobulin domain-containing protein 8 OS=Homo sapiens OX=9606 GN=CPAMD8 PE=1 SV=1	242
CAP-LOXL3	QVEEVRI RPAVGW	B8ZZT6	Lysyl oxidase homolog OS=Homo sapiens OX=9606 GN=LOXL3 PE=1 SV=1	249
CAP-CHL1	QKVEEVKPLE	O00533	Neural cell adhesion molecule L1-like protein OS=Homo sapiens OX=9606 GN=CHL1 PE=1 SV=4	379

83/88

Fig. 15C

Capped peptide names	Amino Acid Start	MW (g/mol)	m/z detected	Z	Seq ID
CAP-FJX1	414		760.42844	3	378
CAP-MMP16	160		1045.99953	2	347
CAP-CPAMD8-1	1638	1913.00159	957.508625	2	243
CAP-KISS1	104	2265.108	756.0438233	3	330
CAP-TRH	80		363.17758	1	250
CAP-WNT9A	164	2344.13108	782.3848567	3	309
CAP-OSM	205	1790.09763	597.70704	3	311
CAP-PMCH	127		986.4922	2	413
CAP-BMP4	39		451.2049	1	226
CAP-EDN3	123	1725.85875	863.937205	2	265
CAP-FGL2	206	1770.81618	885.912005	2	342
CAP-PRSS36	221	1429.673	1430.68127	1	224
CAP-COL11A1-1	554		739.3571	2	230
CAP-SOD3	193	870.38321	871.39104	1	391
CAP-LAMA3	1295	2484.01264	622.01099	4	217
CAP-FGF5	41	585.25476	586.26259	1	281
CAP-VIP	140	1502.82797	752.421815	2	425
CAP-LEFTY2	60	1823.003	912.509195	2	282
CAP-FURIN	447	1822.99755			359
CAP-COL6A3	614				241
CAP-ADIO	104				306
CAP-LGI3	471				327
CAP-CRLF1	404				399
CAP-FGF18	67	340.2111	341.21893	1	440
CAP-PPY	48				321
CAP-CILP2	134	381.2125	382.22033	1	305
CAP-PRG4	1279				387
CAP-LGALS9	166				382
CAP-IFNG-1	138				259
CAP-PAPPA2	257				418
CAP-F5	618				331
CAP-EMILIN1	109	369.20127	370.2091	1	237
CAP-SCT	51	397.23257	398.2404	1	397
CAP-GNAS	9	2077.01772	693.34707	3	363
CAP-METRNL	159				340
CAP-BMP2	15				232
CAP-POMC	24				295
CAP-GDF15	187				467
CAP-VEGFA	5				233
CAP-VWF	1188				229
CAP-MEGF6-1	697				314
CAP-GDNF	68				291
CAP-BRAWH3032298	1033				240
CAP-HHIPL1	647				447
CAP-LAMB4	1098				333

84/88

Fig. 15C (cont.)

Capped peptide names	Amino Acid Start	MW (g/mol)	m/z detected	Z	Seq ID
CAP-MEGF6-2	51				313
CAP-COL4A4	755				298
CAP-LYG1	29				373
CAP-F8	975				369
CAP-SIGLEC1	600				386
CAP-THBS4	118	427.20675	428.21458	1	420
CAP-PTX3	122				414
CAP-ADAMTS5	79	1612.82499	807.420325	2	350
CAP-COL8A2	289				266
CAP-PPT	28	723.34147	724.3493	1	274
CAP-CSF1	424				368
CAP-COL11A1-2	1301	1539.73182	770.87374	2	277
CAP-CCDC80	606				345
CAP-MUC4	4873				315
CAP-AFP	588				278
CAP-TG	843				430
CAP-FCGBP	475				353
CAP-COL11A2	1301				253
CAP-DKFZP666C182	404				279
CAP-ADAMTS1	2				268
CAP-NXPE4	116				374
CAP-COL27A1	1500	492.26968	493.27751	1	456
CAP-FAP	405				404
CAP-COL18A1	1687				366
CAP-Q59F13	42				307
CAP-CNPY4	66	556.28573	557.29356	1	466
CAP-PIGR	273				389
CAP-LIPC	25				238
CAP-PAMR1	50				372
CAP-CBLN2	2				344
CAP-GNRH1	24	1181.57307	591.794365	2	362
CAP-NMU	128				227
CAP-SNED1	1051				324
CAP-IFNG-2	156				260
CAP-LAMB2	801				219
CAP-ANGPT1	22				338
CAP-THSD1	269				427
CAP-CPAMD8-2	895				242
CAP-LOXL3	165				249
CAP-CHL1	458				379

85/88

Fig. 15D

Capped peptide names	Plasma 1	Plasma 2	Plasma 3	1 uM Standard	100 nM Standard	10 nM Standard	Seq ID
CAP-FJX1	216305.25	185008.91	447009.09	NA	32581.6	6160.66	378
CAP-MMP16	2517.62	1190.92	1498.01	NA	3084.72	0	347
CAP-CPAMD8-1	5907.35	4241.16	7290.07	NA	16917.76	218.98	243
CAP-KISS1	25796.91	22307.73	43459.78	NA	171227.14	22172.4	330
CAP-TRH	2755.39	917.38	3814.64	NA	15016.59	2073.85	250
CAP-WNT9A	10567.2	3093.5	24937.32	NA	84897.3	5317.42	309
CAP-OSM	0	0	4408.08	109390.02	0	0	311
CAP-PMCH	5865.94	3614.75	5220.94	NA	60160.33	5656.41	413
CAP-BMP4	1900.19	1064.19	2192.45	NA	21139.59	1967.11	226
CAP-EDN3	41677.57	28623.52	31575	NA	432864	46396.62	265
CAP-FGL2	3808.75	3730.81	3568.62	NA	47502.28	4565.35	342
CAP-PRSS36	3622.24	2888.18	4931.85	NA	50518.14	4472.56	224
CAP-COL11A1-1	9862.66	5363.91	14260.9	NA	144686.22	16124.96	230
CAP-SOD3	917.44	1124.38	1347.05	NA	18543.36	1749.19	391
CAP-LAMA3	7702.25	2591.05	3285.31	NA	123038.77	123038.77	217
CAP-FGF5	7089.97	7182.71	14083.71	NA	167198.47	18725.24	281
CAP-VIP	19100.19	10339.2	22792.78	NA	321821.72	30782.06	425
CAP-LEFTY2	30856.22	20592.03	46657.39	NA	604547.76	56964.49	282
CAP-FURIN	30369.4	20592.03	46657.39	NA	604547.76	56964.49	359
CAP-COL6A3	10573.42	7329.83	17722.68	NA	225480.1	24145.1	241
CAP-ADIO	16111.94	10608.57	15456.98	NA	266697.12	22385.12	306
CAP-LGI3	44240.78	34442.86	75244.89	NA	1099266.15	124574.93	327
CAP-CRLF1	8625.62	8541.58	7441.7	NA	178556.43	18239.41	399
CAP-FGF18	4741.06	6280.02	1983.05	NA	113198.13	19821.85	440
CAP-PPY	9922.9	3799.58	21040.7	NA	265808.9	14205.09	321
CAP-CILP2	1507.2	1295.13	3699.79	NA	53872.33	4498.02	305
CAP-PRG4	1005.94	0	8433.51	NA	92607.66	15199.57	387
CAP-LGALS9	31179.74	19715.02	45565.74	NA	897424.56	97939.77	382
CAP-IFNG-1	38332.14	18592.34	57081.71	NA	1296231.98	169733.26	259
CAP-PAPPA2	856.24	753.86	1534.24	NA	40385.89	6088.69	418
CAP-F5	5460.95	1515.09	2846.34	NA	134686.3	16275.46	331
CAP-EMILIN1	2408.75	1589.76	5339.38	NA	129685.09	16637.69	237
CAP-SCT	5595.29	2999.33	9487	NA	254241.34	29303.94	397
CAP-GNAS	2924.47	545.97	3063.83	NA	92430.04	8463.16	363
CAP-METRNL	1133.88	219.05	2455.35	NA	58072.38	9134.38	340
CAP-BMP2	2672.33	0	3898.55	NA	106363.03	15018.81	232
CAP-POMC	6104.95	2657.13	6177.4	NA	228937	20923	295
CAP-GDF15	473	199	2092.04	NA	46342.36	5058.26	467
CAP-VEGFA	10643.52	7147.37	12587.8	NA	555485.77	71748.82	233
CAP-VWF	8050.02	5507	11947.76	NA	467931.42	52116.07	229
CAP-MEGF6-1	1824.11	1839.02	1177.52	NA	92905.11	11092.64	314
CAP-GDNF	2818.77	701.14	16882.54	NA	428244.42	80209.2	291
CAP-BRAWH3032298	4239.49	1622.73	13333.32	NA	412461.85	48443.67	240

86/88

Fig. 15D (cont.)

Capped peptide names	Plasma 1	Plasma 2	Plasma 3	1 uM Standard	100 nM Standard	10 nM Standard	Seq ID
CAP-HHIPL1	940.33	721.93	5899.37	NA	163691.96	11617.22	447
CAP-LAMB4	3123.99	2516.61	4777.43	NA	238171.66	22865.58	333
CAP-MEGF6-2	5117.33	2347.52	8986.18	NA	405047.11	55157.42	313
CAP-COL4A4	16540.7	8794.14	23036.73	NA	1175608.68	142498.15	298
CAP-LYG1	1850.41	3056.19	1867.62	NA	175151.15	26560.06	373
CAP-F8	1748.99	1534.84	7154.46	NA	270044.9	37182.6	369
CAP-SIGLEC1	4756.54	2112.64	3475.53	NA	274596.9	42596.52	386
CAP-THBS4	3571.32	3426.41	2263.6	NA	239608.21	28439.64	420
CAP-PTX3	861.99	519.62	1574.77	NA	88215.77	19688.82	414
CAP-ADAMTS5	4695.14	1869.02	4803.12	NA	305328.04	34970.37	350
CAP-COL8A2	2527.26	1228.76	5690.5	NA	250308.45	25298.39	266
CAP-PPT	2514.8	2399.3	3581.01	NA	236442.87	23258.05	274
CAP-CSF1	1276.53	0	2674.91	NA	114710.45	14026	368
CAP-COL11A1-2	4446.27	858.84	2664.29	NA	227298.2	23939.2	277
CAP-CCDC80	4172.33	2434.31	6402.82	NA	501821.88	65445.94	345
CAP-MUC4	2353.88	1584.72	2623.14	NA	256579.31	36401.27	315
CAP-AFP	1110.48	638.86	2252.13	NA	206094.23	25951.54	278
CAP-TG	2053.86	2324.83	2885.56	NA	377350.44	47179.97	430
CAP-FCGBP	1977.37	0	2397.96	NA	241977.82	36034.37	353
CAP-COL11A2	1702.19	1687.23	968.35	NA	241484.79	31440.69	253
CAP-DKFZP666C182	1620.8	568.67	2665.78	NA	257545.87	22192.08	279
CAP-ADAMTS1	9232.49	5068.68	12066.1	NA	1455816.22	164281.88	268
CAP-NXPE4	1748.5	739.79	2359.79	NA	278276.26	30401.54	374
CAP-COL27A1	1937.07	649.18	4602.96	NA	420347.68	42628.4	456
CAP-FAP	6984.05	4083.13	6102.5	NA	1005269.62	57248.07	404
CAP-COL18A1	3872.46	1770.56	4773.09	NA	685462.05	105625.47	366
CAP-Q59F13	9578.4	5674.14	13524.73	NA	1790853.8	169215.62	307
CAP-CNPY4	2253.64	646.19	3222.99	NA	418659.47	50019.38	466
CAP-PIGR	4024.53	5475.31	5393.13	NA	925087.81	19295.92	389
CAP-LIPC	6499.3	5862.21	10048.76	NA	1656378.35	153113.71	238
CAP-PAMR1	2058.87	718.34	1254.72	NA	420504.77	50235.16	372
CAP-CBLN2	2013.93	2549.62	9942.27	NA	1460220.09	122731.65	344
CAP-GNRH1	1223.85	1375.39	1440.03	NA	430135.38	48464.46	362
CAP-NMU	3325.22	1695.97	4289.95	NA	1028542.47	119308.96	227
CAP-SNED1	3960.03	1777.86	6910.7	NA	1518269.46	211827.06	324
CAP-IFNG-2	3514.94	3134.29	8482.93	NA	3415075.07	325748.62	260
CAP-LAMB2	0	425.48	1102.5	NA	451689.76	64353.34	219
CAP-ANGPT1			1843.77	NA	583460.98	98187.62	338
CAP-THSD1	1936.36	1021.83	2392.42	NA	2439386.14	240421	427
CAP-CPAMD8-2	0	481.57	3898.08	NA	2188825.76	227119.24	242
CAP-LOXL3	0	481.21	1087.79	NA	973573.52	114038.82	249
CAP-CHL1	202.68	0	1863.64	NA	2936664.37	418844.44	379

87/88

Fig. 15E

Capped peptide names	Plasma 1	Plasma 2	Plasma 3	Seq ID
CAP-FJX1	736.8198293	630.2123202	1522.686857	378
CAP-MMP16	73.45425193	34.74636272	43.70604139	347
CAP-CPAMD8-1	31.83834388	22.85822078	39.29067273	243
CAP-KISS1	15.57630371	13.46951932	26.24123325	330
CAP-TRH	19.1601701	6.379190187	26.52588246	250
CAP-WNT9A	11.95085994	3.498560189	28.20259091	309
CAP-OSM	0	0	36.26722072	311
CAP-PMCH	9.686176701	5.968882605	8.621115692	413
CAP-BMP4	8.919925852	4.995550915	10.29186104	226
CAP-EDN3	9.705816051	6.665806568	7.353143233	265
CAP-FGL2	7.983512096	7.820142241	7.480176156	342
CAP-PRSS36	7.079975972	5.645193306	9.639720034	224
CAP-COL11A1-1	6.904408062	3.755033981	9.983419578	230
CAP-SOD3	4.916563307	6.025555297	7.218844397	391
CAP-LAMA3	10.14791647	3.413776362	4.328482129	217
CAP-FGF5	4.297726264	4.353942458	8.537120799	281
CAP-VIP	5.906470273	3.197254972	7.048352792	425
CAP-LEFTY2	5.071484014	3.384476483	7.668541627	282
CAP-FURIN	4.991470978	3.384476483	7.668541627	359
CAP-COL6A3	4.726489681	3.276552512	7.922324484	241
CAP-ADIO	5.935339238	3.908000016	5.694064147	306
CAP-LGI3	4.085058035	3.180348131	6.94788253	327
CAP-CRLF1	4.842316805	4.795137784	4.177678702	399
CAP-FGF18	4.569633744	6.05294835	1.911347293	440
CAP-PPY	3.549473277	1.359129657	7.526368539	321
CAP-CILP2	2.747339659	2.360776282	6.744015258	305
CAP-PRG4	1.169575428	0	9.805382099	387
CAP-LGALS9	3.509981222	2.219369051	5.129449179	382
CAP-IFNG-1	3.062491363	1.485408346	4.560461374	259
CAP-PAPPA2	2.246877296	1.978219796	4.02603128	418
CAP-F5	4.150679955	1.151567711	2.163404972	331
CAP-EMILIN1	1.917669049	1.26564963	4.250820452	237
CAP-SCT	2.238738867	1.200065885	3.795856092	397
CAP-GNAS	3.134596641	0.585198593	3.283969822	363
CAP-METRNL	2.085275246	0.402846459	4.515540071	340
CAP-BMP2	2.633004037	0	3.841179004	232
CAP-POMC	2.641387118	1.149642332	2.672733566	295
CAP-GDF15	1.031147585	0.433823191	4.560680746	467
CAP-VEGFA	1.980243188	1.329779129	2.341979458	233
CAP-VWF	1.742364249	1.191947339	2.585999771	229
CAP-MEGF6-1	2.006661087	2.023063232	1.295362431	314
CAP-GDNF	0.72891847	0.18131096	4.365732296	291
CAP-BRAWH3032298	1.048173198	0.401204412	3.296535354	240
CAP-HHIPL1	0.556500705	0.42724847	3.491331302	447
CAP-LAMB4	1.305857689	1.051967042	1.997011417	333

88/88

Fig. 15E (cont.)

Capped peptide names	Plasma 1	Plasma 2	Plasma 3	Seq ID
CAP-MEGF6-2	1.316299717	0.603838313	2.311460506	313
CAP-COL4A4	1.440952305	0.766106411	2.006857582	298
CAP-LYG1	1.120773123	1.851100897	1.131197032	373
CAP-F8	0.675975029	0.59320723	2.765159496	369
CAP-SIGLEC1	1.84520646	0.819557278	1.34826374	386
CAP-THBS4	1.522095831	1.460335219	0.964745843	420
CAP-PTX3	1.132096204	0.682443914	2.068227172	414
CAP-ADAMTS5	1.562976186	0.622182459	1.598921902	350
CAP-COL8A2	1.010858803	0.491482025	2.276098233	266
CAP-PPT	1.061670338	1.012909831	1.511791037	274
CAP-CSF1	1.141066967	0	2.391053435	368
CAP-COL11A1-2	1.967772757	0.380094316	1.179127061	277
CAP-CCDC80	0.8605188	0.502062281	1.320544391	345
CAP-MUC4	0.962172249	0.647770323	1.072234997	315
CAP-AFP	0.554800198	0.319176981	1.12517305	278
CAP-TG	0.559854429	0.633717183	0.786564589	430
CAP-FCGBP	0.864136733	0	1.047940102	353
CAP-COL11A2	0.729356835	0.722946753	0.41492001	253
CAP-DKFZP666C182	0.619798814	0.217461125	1.019402322	279
CAP-ADAMTS1	0.643361988	0.353208727	0.840820849	268
CAP-NXPE4	0.634856996	0.268607868	0.85680823	374
CAP-COL27A1	0.461549911	0.15468154	1.096757359	456
CAP-FAP	0.66302765	0.387630112	0.579338096	404
CAP-COL18A1	0.601068322	0.274819502	0.74086064	366
CAP-Q59F13	0.531595772	0.314911554	0.750614851	307
CAP-CNPY4	0.550204944	0.157761192	0.786862601	466
CAP-PIGR	0.399879602	0.544029932	0.53586448	389
CAP-LIPC	0.389111128	0.350968742	0.601616226	238
CAP-PAMR1	0.500441557	0.174604122	0.304979931	372
CAP-CBLN2	0.135517956	0.171564698	0.669018343	344
CAP-GNRH1	0.288590234	0.324324159	0.339566609	362
CAP-NMU	0.32914515	0.167874697	0.424638441	227
CAP-SNED1	0.27280399	0.122475664	0.476073802	324
CAP-IFNG-2	0.102399214	0.091309903	0.2471295	260
CAP-LAMB2	0	0.098862895	0.256172657	219
CAP-ANGPT1	0	0	0.341950154	338
CAP-THSD1	0.079252007	0.041821809	0.097917787	427
CAP-CPAMD8-2	0	0.022093672	0.17883776	242
CAP-LOXL3	0	0.050386447	0.113900114	249
CAP-CHL1	0.007244839	0	0.066616202	379