Multiplexed Temporal Quantification of the Exercise-regulated Plasma Peptidome*[S]

Benjamin L. Parker‡, James G. Burchfield‡, Daniel Clayton§, Thomas A. Geddes‡, Richard J. Payne§, Bente Kiens¶, Jørgen F. P. Wojtaszewski¶, Erik A. Richter¶, and David E. James‡‖**

Exercise is extremely beneficial to whole body health reducing the risk of a number of chronic human diseases. Some of these physiological benefits appear to be mediated via the secretion of peptide/protein hormones into the blood stream. The plasma peptidome contains the entire complement of low molecular weight endogenous peptides derived from secretion, protease activity and PTMs, and is a rich source of hormones. In the current study we have quantified the effects of intense exercise on the plasma peptidome to identify novel exercise regulated secretory factors in humans. We developed an optimized 2D-LC-MS/MS method and used multiple fragmentation methods including HCD and EThcD to analyze endogenous peptides. This resulted in quantification of 5,548 unique peptides during a time course of exercise and recovery. The plasma peptidome underwent dynamic and large changes during exercise on a time-scale of minutes with many rapidly reversible following exercise cessation. Among acutely regulated peptides, many were known hormones including insulin, glucagon, ghrelin, bradykinin, cholecystokinin and secretogranins validating the method. Prediction of bioactive peptides regulated with exercise identified *C*-terminal peptides from Transgelins, which were increased in plasma during exercise. *In vitro* experiments using synthetic peptides identified a role for transgelin peptides on the regulation of cell-cycle, extracellular matrix remodeling and cell migration. We investigated the effects of exercise on the regulation of PTMs and proteolytic processing by building a site-specific network of protease/substrate activity. Collectively, our deep peptidomic analysis of plasma revealed that exercise rapidly modulates the circulation of hundreds of bioactive peptides through a network of proteases and PTMs. These findings illustrate that peptidomics is an ideal method for quantifying changes in circulating factors on a global scale in response to physiological perturbations such as exercise. This will likely be a key method for pinpointing exercise regulated factors that generate health benefits. ***Molecular & Cellular Proteomics 16: 10.1074/mcp.RA117.000020, 2055–2068, 2017.***

Multicellular organisms have evolved sophisticated mechanisms to enable cell-cell communication. Such mechanisms are fundamental to homeostasis enabling the organism to respond appropriately to the environment. One of the most common methods of communication involves the secretion or release of proteins and peptides from one cell in response to an environmental perturbation. These signals travel via the blood to modulate physiological pathways in other cells and tissues. Thus, the comprehensive measurement of peptides in blood provides a systematic record of this complex interorgan communication system and how it changes under certain conditions. There has been an extensive effort to develop more sensitive and comprehensive methods for quantifying blood borne peptides. Traditional analysis relied on the use of antibodies to measure just one or a handful of factors. These assays have been applied to numerous hormones and are commonplace in clinical diagnostics. More recently, the ability to globally characterize the complete repertoire of endogenous peptides in blood, referred to as the peptidome, has been advanced by the field of mass spectrometry-based proteomics. This is primarily attributed to advances in isolation, separation, fragmentation, quantification and computational analysis of hundreds or thousands of peptides. A variety of methods have been used to isolate the peptidome including molecular weight separation techniques such as size-exclusion chromatography or filtration (1, 2) or depletion of larger proteins with acid precipitation or organic solvents (3, 4). The combination of these extraction techniques with multidimensional liquid chromatography has led to the identification of hundreds of endogenous low molecular weight peptides in plasma (5). This has been coupled to a variety of mass spectrometry platforms and fragmentation approaches (for an ex-

From the ‡Charles Perkins Centre, School of Life and Environmental Sciences, University of Sydney, Sydney, NSW 2006, Australia; §School of Chemistry, University of Sydney, Sydney, NSW 2006, Australia; ¶Department of Nutrition, Exercise and Sports, August Krogh Centre, University of Copenhagen, Copenhagen 2100, Denmark; ‖School of Medicine, University of Sydney, Sydney, NSW 2006, Australia

Received August 27, 2017
Published, MCP Papers in Press, October 5, 2017, DOI 10.1074/mcp.RA117.000020
Author contributions: B.L.P. and D.E.J. designed research; B.L.P., J.G.B., D.C., T.A.G., B.K., J.W., and E.A.R. performed research; B.L.P. and J.G.B. analyzed data; B.L.P., J.G.B., J.W., E.A.R., and D.E.J. wrote the paper; D.C. and R.J.P. contributed new reagents/analytic tools.

tensive review see (6)). To identify differentially regulated peptides between two or more states, quantitative peptidomic analysis has been performed using label-free approaches (7), stable isotope labeling with chemical derivatisation (8, 9), or stable isotope labeling with metabolic incorporation (10). These quantification strategies have been applied to identify biomarkers for prognosis or diagnosis of disease. This includes analysis of urine for the identification of biomarkers in chronic kidney disease (11), transplant rejection (12), and cardiovascular disease (13), and validated peptide biomarkers have also been identified for cancer (14, 15) and diabetes (16). Despite these advances in peptidomic technologies, further developments are required to increase throughput for clinical analysis to overcome the large dynamic range in plasma. Overcoming these hurdles will facilitate the identification of new peptides with bioactive properties and reveal the temporal regulation and stability of the peptidome.

A key feature of the peptidome is it will undergo dynamic and detectable change in response to physiological perturbations and/or disease. The goal of the present study was to determine the influence of exercise on the peptidome. This is important because exercise is a common physiological perturbation that is likely to have a profound impact on the peptidome. Exercise also has many health benefits including improvements in heart function, neurological function and insulin sensitivity and, these effects are thought to be mediated at least in part by the regulation of circulating plasma factors (17). These factors, including peptide hormones, interleukins and growth factors have profound effects throughout the body. For example, to meet the large amounts of energy essential for exercise, increased oxygen delivery is required. This is achieved by increasing blood flow through the action of a variety of vasodilators. Exercise increases the activity of the kinin-kallikrein system, a protease cascade producing the bradykinin peptide, a potent endogenous vasodilator (18). Therefore, exercise and the inhibition of enzymes that degrade vasodilators are important therapies for the treatment of hypertensive patients. We hypothesize that characterizing the plasma peptidome in response to exercise will allow an investigation of protease regulation and simultaneously identify new signaling factors that contribute to physiological adaptions. Moreover, such an analysis will provide another step forward in validating the use of peptidomics as a useful tool for discovery of both novel regulatory factors and clinical diagnostics. In the present study, we performed a temporal peptidomic analysis of human plasma in response to high-intensity exercise. Using isobaric tagging, multidimensional liquid chromatography and tandem mass spectrometry with complementary fragmentation techniques, we identified 6,652 unique endogenous plasma peptides. Our data reveal the peptidome is rapidly modulated by exercise and involves the coordinated regulation of a network of proteases. The unexplored complexity of the exercise peptidome may reveal important signaling molecules mediating the beneficial effects of exercise.

EXPERIMENTAL PROCEDURES

Human Subjects and Sample Collection—Four healthy male volunteers (age: 26–28; BMI: 23.3–25.8 kg/m2, VO2 max: 41.6–47.3 ml/kg/min, Wmax: 280–295 W) abstained from strenuous exercise for 2 days before the experiment. They reported to the laboratory in the overnight fasted state and rested in the supine position for 30 min. A venous catheter was inserted in a forearm vein and a blood sample was obtained. Venous blood was collected from the forearm and obtained in heparinized syringes and quickly transferred to Eppendorf tubes containing 30 μl 200 mM EDTA/1500 μl blood. Following warm up for 2 min, subjects underwent cycle exercise for 6 min at 77% of individual Wmax and then to exhaustion at 87–88% of Wmax, which occurred after 9–11 min total exercise time following warm up. Blood samples were collected in the last minute of exercise followed by three additional samples taken at 1, 2, and 5 h postexercise. Postexercise, subjects lay fasted in the supine position with access only to water. The study was approved by the regional ethics committee in Denmark (Journal number: H-1-2012-006) and carried out in accordance with the Declaration of Helsinki II. Written informed consent was obtained from each subject.

Peptidomic Sample Preparation—For plasma peptidome isolation comparisons, 100 μl aliquots of the identical plasma was mixed 1:1 with either PBS or Urea Buffer (8 M urea, 20 mM dithiothreitol (DTT) in 50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acids (HEPES), pH 8.0), and incubated at room temperature for 5 min. The diluted plasma aliquots were mixed with either; (1) 1 volume of 20% trichloroacetic (TCA), (2) 5 volumes of 100% acetonitrile (AcN), or (3) 5 volumes of 100% acetone and incubated for 60 min at 4 °C to precipitate proteins. The precipitate was centrifuged at 16,000 $\times$ *g* for 10 min and the supernatant containing peptides collected. Supernatants from the AcN and acetone precipitations were dried by vacuum centrifugation and resuspended in 5% AcN, 0.1% TFA. Additional 100 μl aliquots of the same plasma sample were diluted with either PBS or Urea Buffer and applied to 10 kDa molecular weight cut-off (MWCO)[1] filters and centrifuged at 16,000 $\times$ *g* for 30 min at 4 °C. The filters were washed once with either PBS or Urea Buffer by a second centrifugation at 16,000 $\times$ *g* for 30 min and filtrates adjusted to 5% AcN, 0.1% TFA. All peptide preparations were desalted with hydrophilic-lipophilic balance solid-phase extraction (HLB-SPE) columns (Waters, Milford, MA). Peptides were eluted with 50% acetonitrile, 0.1% trifluoroacetic acid (TFA) and dried by vacuum centrifugation. For quantification of exercise plasma samples, 200 μl of plasma was processed with the Urea Buffer and TCA precipitation approach described above. Desalted peptides were resuspended in 40 μl of 100 mM HEPES and adjusted to pH 8.0. Peptides were labeled with 10-plex tandem mass tags (TMT; Thermo Scentific, CA) for 90 min at room temperature, quenched with 2 μl of 5% hydroxylamine and acidified to 2% formic acid (FA). To cover the 20 samples in total (4 subjects x 5 time-points) two TMT 10-plex experiments were performed. Each 10-plex experiment contained two subjects with the following labeling of channels; 126/129N: preexercise, 127N/129C: exercise, 127C/130N: 1h postexercise, 128N/130C: 2h postexercise, 128C/131: 5h postexercise. The peptides from each TMT 10-plex experiment were combined and concentrated by HLB-SPE followed by removal of excess unreacted TMT reagent and fractionation with hydrophilic interaction liquid chromatography (HILIC), as previously described (19).

[1] The abbreviations used are: MWCO, molecular weight cut-off; TFA, trifluoracetic acid; FA, formic acid; TMT, tandem mass tags; HILIC, hydrophilic interaction liquid chromatography.

Proteomic Sample Preparation—L6 myoblasts were harvested by scraping in 8 M guanidine containing 10 mM (Tris(2-carboxyethyl) phosphine) and 40 mM chloroacetamide in 100 mM Tris, pH 7.5 and tip-probe sonicated for 30 s. Lysates were heated to 95 °C for 5 min followed by centrifugation and 20,000 × *g* for 15 min at 4 °C. Protein containing supernatant was precipitated with 4 volumes of acetone overnight at −30 °C and protein pellets washed with 80% acetone. Protein pellets were resuspension in 100 mM Tris containing 10% trifluoroethanol and digested with trypsin (1:50 enzyme/substrate) overnight at 37 °C with vortexing. Peptides were acidified to a final concentration of 1% TFA and desalted using SDB-RPS microcolumns (3 M Empore, Sigma). Peptides were eluted in 80% acetonitrile, 1% ammonium hydroxide and dried by vacuum centrifugation.

Liquid Chromatography - Tandem Mass Spectrometry—For peptidomics analysis, peptides were analyzed on a Dionex 3500RS nanoUHPLC coupled to an Orbitrap Fusion mass spectrometer in positive mode. Peptides were separated using an in-house packed 75 μm × 40 cm pulled column (1.9 μm particle size, C18AQ; Dr Maisch, Germany) with a gradient of 2–30% acetonitrile containing 0.1% FA over 120 min at 250 nl/min at 55 °C. An MS1 scan was acquired from 350–1550 (120,000 resolution, 5e5 AGC, 100 ms injection time) followed by MS/MS data-dependent acquisition with HCD and detection in the Orbitrap (60,000 resolution, 2e5 AGC, 120 ms injection time, 40 NCE, 2.0 m/z quadrupole isolation width) and, EThcD (20) and detection in the Orbitrap (60,000 resolution, 2e5 AGC, 120 ms injection time, calibrated charge-dependent ETD reaction times (2 + 121 ms; 3 + 54 ms; 4 + 30 ms; 5 + 20ms; 6 + 13 ms; 7+; 10 ms), 25 NCE for HCD supplemental activation, 2.0 *m/z* quadrupole isolation width).

For proteomic analysis, peptides were analyzed on a Easy-nLC 1200 nanoUHPLC coupled to a Q Exactive HF mass spectrometer in positive mode. Peptides were separated using an in-house packed 75 μm × 50 cm pulled column (1.9 μm particle size, C18AQ; Dr Maisch, Germany) with a gradient of 2–30% acetonitrile containing 0.1% FA over 120 min at 300 nl/min at 60 °C. An MS1 scan was acquired from 300–1650 (60,000 resolution, 3e6 AGC, 50 ms injection time) followed by MS/MS data-dependent acquisition with HCD and detection in the Orbitrap (15,000 resolution, 2e5 AGC, 25 ms injection time, 27 NCE, 1.4 *m/z* quadrupole isolation width).

Data Analysis—Data were processed with Proteome Discoverer (v2.1) using the Byonic node (v1.0.334) (21) or MaxQuant (v1.5.3.30) using Andromeda (22) against the UniProt human database containing only the primary accession of an open reading frame without isoforms (January 2016, 20,955 entries). For Byonic analysis, the precursor MS, HCD MS/MS and EThcD MS/MS tolerance were set to 20 ppm with nonspecific enzyme searching. The peptides were searched with oxidation of methionine, C terminus amidation and asparagine N-glycan modification in the NxS/T motif (48 N-glycan monosaccharide compositions) set as variable modification, and TMT tags on peptide N terminus/lysine set as a fixed modification. A precursor isotope off set was enabled (narrow) to account for incorrect precursor monoisotopic reporting (± 1.0 Da). All data were searched as a single batch with PSM and protein FDR set to 1% using a target decoy approach in Byonic. For MaxQuant analysis of the peptidomics data, all settings were default with precursor-ion and product-ion tolerance set to 20 ppm and 0.02 Da, respectively. No enzyme specificity was employed and peptides searched with oxidation of methionine and C terminus amidation set as variable modifications, and TMT tags on peptide N terminus/lysine set as a fixed modification. All data were searched as a single batch with PSM and protein FDR set to 1% using a target decoy approach. For MaxQuant analysis of proteomics data, all settings were default with precursor-ion and product-ion tolerance set to 20 ppm and 0.02 Da, respectively. Full trypsin specificity was employed with a maximum of 2-missed cleavages and peptides searched with oxidation of methionine and acetylation on protein N terminus set as variable modifications, and carbamidomethylation of cysteine set as fixed modification. All data was searched as a single batch with PSM and protein FDR set to 1% using a target decoy approach. The match between runs and label-free quantification (MaxLFQ) options were selected (23). Quantification of peptidomics data was performed with either the precursor area detector node for LFQ, or reporter ion quantification node for TMT quantification in Proteome Discoverer. For LFQ, extracted ion chromatograms were generated at 2 ppm, and for TMT precision was set to 10 ppm and corrected for isotopic impurities. Only spectra with <50% coisolation interference were used for quantification with an average signal-to-noise filter of >10. Statistical analysis including the determination of differentially regulated peptides and enrichment analysis, and visualization including heat maps were performed in Perseus (v1.5.3.0) (24). Fuzzy c-means clustering (five to nine clusters) was performed in GPROX (25) using 100 iterations and a fuzzification factor of 2.0.

Peptide Synthesis—Transgelin 1: MGSNRGASQAGMTGYGRPRQIIS (TAGLN), Transgelin 2: MGTNRGASQAGMTGYGMPRQIL (TAGLN2), and scramble Transgelin 1 sequence as a negative control: RGMINGRMIQSTGSYPSARGQAG (control) were prepared at 0.1 mmol scale using Fmoc solid-phase peptide synthesis (SPPS) on Wang resin using standard side chain protecting groups. The negative control for the cell migration experiments was based on the Transgelin 1 sequence and randomized using the ExPASy RandSeq tool (Random protein sequence generator, http://web.expasy.org/randseq/). The first residue was loaded to the Wang resin using a 3-fold molar excess of Fmoc-amino acid and HBTU (113.7 mg, 0.3 mmol) with a catalytic amount of DMAP (0.1 equiv, 1.2 mg, 0.01 mmol) in DMF (0.6 ml). After overnight loading the resin was capped with a 1vol.% acetic anhydride, 2vol.% N,N,-diisopropylethylamine solution in DMF (1 ml) for 10 min, then the extent of the first residue loading was determined from the combined pool (8 ml) of 2 × 15 min treatments with 20% piperidine in DMF; 50 μl of the pool was diluted to 1 ml and the UV absorbance of the piperidine-fulvine ($\lambda = 301$ nm, $\varepsilon = 7800$ M-1 cm-1) adduct determined. Elongation of the target peptides was performed on a CEM Liberty Blue automated microwave peptide synthesizer (USA, NC) using a 4 min coupling cycle (2 min coupling (90 °C), 1 min deprotection (90 °C), with the additional 1 min for associated washes and liquid handling), and a 5-fold excess of Fmoc amino acid, Oxyma and DIC as recommended by the manufacturer. Peptides were cleaved from the resin and deprotected with a TFA cleavage solution (TFA/TIPS/H_2O/DODT, 92.5:2.5:2.5:2.5, 5 ml per 100 mg of resin) for 2 h at room temperature.

Synthetic Peptide Purification—A mobile phase of 0.1% TFA in water (Solvent A) and 0.1% TFA in acetonitrile (Solvent B) was used in all cases. Preparative reverse-phase HPLC was performed using a Waters 2535 Quaternary gradient Module with a Waters 2489 UV/Vis detector operating at 230 and 254 nm. Transgelin peptides were purified on a 19 mm x 150 mm Waters Sunfire column (C18 OBD, 5 μm particle size) at a flow rate of 20 ml/min. Transgelin 1 was purified using a linear 0–30% buffer B gradient over 30 min whereas transgelin 2 was purified using a linear 3–35% buffer B gradient over 30 min, and for the negative control peptide a linear 10–40% buffer B gradient over 40 min. The identity of the peptides was confirmed by LC-MS on a Shimadzu UPLC-MS 2020 instrument consisting of LC-M20A pumps and a SPD-M30A diode array detector with a Shimadzu 2020 mass spectrometer (ESI) operating in positive mode (0.1% formic acid); Transgelin 1, Calculated Mass $(M+2H)^{2+}$: 1198.1 (100%), $(M+3H)^{3+}$: 799.1 (100%), $(M+4H)^{4+}$: 599.6; Observed Mass (ESI+) $(M+2H)^{2+}$: 1199.4 (100%), $(M+3H)^{3+}$: 799.8 (100%), $(M+4H)^{4+}$: 600.0 (100%). Transgelin 2, Calculated Mass $(M+2H)^{2+}$: 1149.1 (100%), $(M+3H)^{3+}$: 766.4 (100%), $(M+4H)^{4+}$: 575.0; Observed Mass (ESI+) $(M+2H)^{2+}$: 1150.3 (100%), $(M+3H)^{3+}$: 767.0

(100%), $(M+4H)^{4+}$: 575.4 (100%), negative control peptide, calculated Mass (M+2H)2+: 1198.1 (100%), (M+3H)3+: 799.1 (100%), (M+4H)4+: 599.6; Observed Mass (ESI+) (M+2H)2+: 1198.6 (100%), (M+3H)3+: 799.5 (100%), (M+4H)4+: 599.8 (100%). The purity of the purified peptides was assessed on a Waters Acquity UPLC system and on a Waters Acquity C18 BEH 1.7 μm 2.1 $\times$ 50 mm column (0–50% B, 5 min at 0.6 ml/min). Yields for the final peptides were: Transgelin 1, 4.6 mg, overall yield 11% (purity ≥98%); Transgelin 2, 1.6 mg, overall yield 4% (purity ≥98%); negative control peptide, 7.2 mg, overall yield 24% (purity ≥98%).

Cell Culture, Live Cell Microscopy and Wound Healing Assay—L6 myoblasts were cultured in alpha-MEM containing 10% fetal calf serum (FCS) to a maximum of 80% confluency before being split every 2–3 days. Cells were treated for 7 d with either a synthetic *C*-terminal peptide from TAGLN spanning amino acids 179–200, or a synthetic *C*-terminal peptide from TAGLN2 spanning amino acids 178–199. The peptides were prepared at 10 mg/ml in PBS and diluted 1:10,000 to a final concentration of 1 μg/ml in alpha-MEM containing 10% FCS. Peptides were present for the duration of the experiment with media replaced every 2–3 d. For proteomics experiments, additional control cells were cultured throughout the duration of the experiment and treated only with PBS at a 1:10,000 dilution. For microscopy experiments, additional control cells were treated with a synthetic scramble control peptide prepared at 10 mg/ml in PBS and diluted 1:10,000 to a final concentration of (1 μg/ml in alpha-MEM containing 10% FCS). Cells were trypsinized and seeded into each well of an ibidi Culture-Insert 2 Well silicone insert in a 35-mm tissue culture treated μ-dish. Following a 4 h recovery, the inserts were removed and the dishes placed into an incubation chamber (37 °C, 10% CO_2; Okolabs) on a Nikon TiE inverted microscope. The 500 μm wide wound created by removal of the insert was then imaged using differential interference contrast (DIC) every 5 min for 48 h. Peptide treatments including the negative control were present throughput the imaging and imaged simultaneously. Closure of the gap was measured using Fiji (26). Segmentation of images into the "cell" and "wound" areas was achieved using Trainable Weka Segmentation v3.1.2 (27). Briefly, images were randomly chosen from multiple conditions to train the model. The model was applied to all images to achieve segmentation and the respective areas measured.

Experimental Design and Statistical Rationale—All presented data shows a minimum of three independent replicates. For peptidomic isolation comparisons, three replicates were performed. For quantification of the exercise-regulated peptidome, four replicates were performed. For treatment of L6 myoblasts with transgelin peptides, four replicates were performed. ANOVA with adjustment for multiple testing with permutation-based correction was used to determine significantly regulated peptides and proteins. Fisher's Exact tests with Benjamini Hochberg correction was used for enrichment analysis. Significant regulation was determined using an adjusted p value <0.05 unless otherwise stated.

RESULTS

Isolation of the Plasma Peptidome—We further expanded on previous optimizations to isolate the plasma peptidome and compared various methods to maximize identifications with speed and ease (Fig. 1*A*). Previous studies have reported loss of polypeptides binding to high abundant proteins during depletion strategies (28). We hypothesized that rapid chaotropic denaturation of plasma with urea under reducing conditions would liberate noncovalently bound peptides to improve recovery during protein depletion. We also compared depletion strategies to isolate the peptidome including protein precipitation and removal with either TCA, acetone or acetonitrile (AcN). Following centrifugation of precipitated proteins, the supernatant containing peptides was collected. For acetone and AcN precipitations, an additional vacuum centrifugation step was required followed by resuspension of peptides in aqueous buffer. Our comparison of peptidome isolation also included removal of proteins with size-exclusion 10 kDa MWCO filters. All peptide isolations were acidified to 0.1% TFA, adjusted to 5% acetonitrile and desalted with HLB-SPE. Peptides were analyzed by single-shot nanoUHPLC-MS/MS employing both HCD and EThcD and quantified by LFQ (supplemental Table S1). Hierarchal clustering of the data separated the four peptidome isolation strategies (Fig. 1*B*). The greatest number of peptides identified was achieved using TCA precipitation, consistent with previous studies (4) (Fig. 1*C*), followed by MWCO isolation, acetone and then AcN precipitation. As hypothesized, depletion of proteins under denaturing and reducing conditions increased the number of peptides identified for all isolation strategies except AcN precipitation. The TCA precipitation method was also the fastest, simplest and cheapest procedure directly compatible with HLB-SPE (precipitation with organic solvents requires additional evaporation before HLB-SPE, and MWCO requires extended centrifugation and washing of the filters likely resulting in additional sample loss). We also assessed the reproducibility of TCA precipitation under denaturing conditions using peptide LFQ revealing an average Pearson correlation of 0.95 (Fig. 1*D*) with 67% of the peptides quantified resulting in <20% CV (Fig. 1*E*).

Characterizing the Exercise-regulated Peptidome of Human Plasma—Having established a suitable peptidomics pipeline we extended the analysis to the physiological setting of acute high-intensity exercise. Four healthy male volunteers abstained from strenuous exercise for 2 d and fasted overnight. Forearm venous blood was collected followed by a single bout of high-intensity cycle exercise increasing to 87–88% of Wmax, which required the subjects to exercise for ~9–11 min. In the last minute of exercise, a second venous blood sample was collected followed by three further blood samples at 1, 2, and 5 h postexercise (Fig. 2*A*). The peptidome was isolated using TCA precipitation under denaturing conditions followed by HLB-SPE (Fig. 2*B*). Peptides were labeled with 10-plex TMT and fractionated by HILIC into 13 fractions. Each fraction was analyzed by nanoUHPLC-MS/MS employing alternating fragmentation with both HCD and EThcD. The multidimensional LC-MS/MS analysis resulted in a total of 24,017 PSMs (11,612 HCD and 12,405 EThcD) (supplemental Table S2). This covered a total of 6,652 unique peptide sequences (4947 HCD and 4218 EThcD) with an overlap of 2513 unique peptides between the two fragmentation approaches (38%) (Fig. 2*C*). EThcD identified higher charge states than HCD (average charge for EThcD and HCD was 4.1 and 3.5, respectively) and produced higher peptide identification scores. Interestingly, EThcD trended to identify shorter length peptides

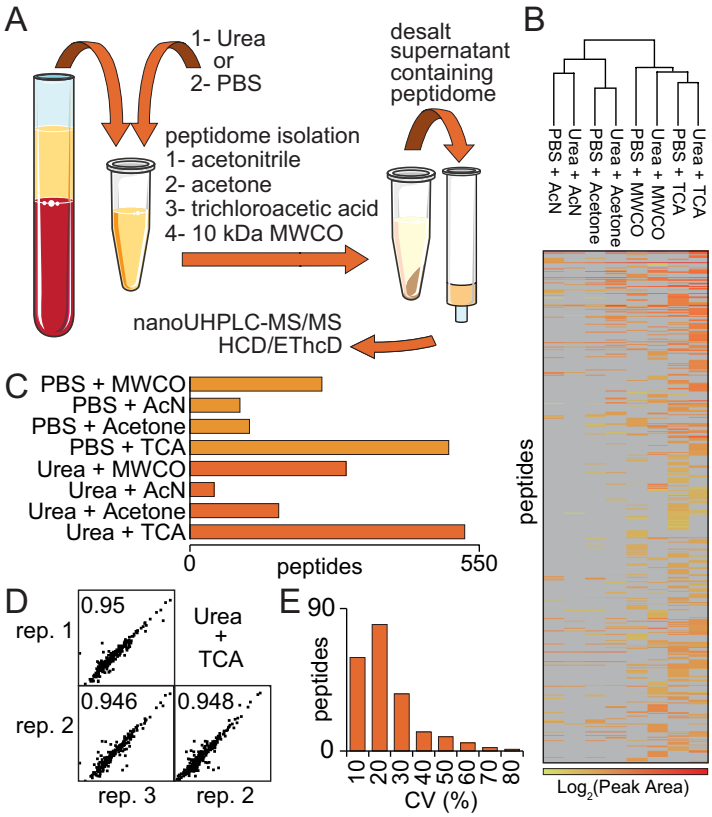


FIG. 1. **Comparison of strategies to isolate the plasma peptidome.** *A*, Overview of the methods compared with isolate the plasma peptidome. Plasma aliquots were mixed with either PBS or Urea Buffer (8 M urea, 20 mM DTT in 50 mM HEPES, pH 8.0). The diluted plasma was mixed with either acetonitrile (AcN), acetone or trichloroacetic acid (TCA), or filtered through 10 kDa molecular weight cut-off (MWCO) filters to remove plasma proteins. Peptides were desalted and analyzed by nanoUHPLC-MS/MS employing alternative fragmentation with HCD and EThcD. *B*, Hierarchical clustering of the quantified peptides using the various isolation strategies. *C*, Number of unique peptides identified using the various isolation strategies. *D*, Pearson correlation of the quantified peptides using the Urea denatured and TCA precipitation isolation approach ($n = 3$). *E*, Coefficient of variation (CV) of the quantified peptides using the Urea denatured and TCA precipitation isolation approach ($n = 3$).

whereas longer peptides were identified equally by HCD and EThcD (Fig. 2*E*). We identified 733 *N*-terminal and 942 *C*-terminal peptides, defined as starting or ending within 50 amino acids of the *N*- and *C*-terminus of the protein primary amino acid sequence, respectively (Fig. 2*F*). The 6,652 peptides mapped to 522 proteins with 65 peptides mapping to >1 protein. Interestingly, 32 and 18 of the peptides that mapped to >1 protein were defined as N-terminal and C-terminal, respectively. Gene Ontology cellular compartment analysis of these proteins revealed the majority were extracellular and membrane localized but also contained organelle annotated proteins from the mitochondria (56 proteins), Golgi apparatus (51 proteins), endoplasmic reticulum (60 proteins), lysosomes (14 proteins), exosomes (22 proteins) and other vesicle proteins (85 proteins), (Fig. 2*G*). Gene Ontology biological process enrichment analysis of the peptidome revealed the most significantly enriched processes were muscle filament sliding, platelet degranulation/activation, exocytosis, glucose metabolic process and secretion by the cell ($p < 0.05$; Fisher's exact test with Benjamini-Hochberg correction) (Fig. 2*H*).

Quantification of the Exercise-regulated Peptidome—Our multiplexed strategy quantified 5548 unique peptides in at least two out of the four subjects across all five time points (supplemental Table S3). All data were expressed as a Log_2(fold-change) relative to the individual's preexercise levels and normalized to a median of zero. The distribution of fold-changes showed the majority of peptides were not regulated but there was a rapid perturbation of the peptidome during exercise, which returned to a normal distribution at 1–5 h postexercise (Fig. 3*A*). A total of 425 unique peptides derived from 81 proteins were significantly regulated during or in the immediate hours post exercise ($p < 0.05$; ANOVA with permutation-based correction). We next used these proteins to perform a pathway enrichment analysis. We first mapped all the proteins identified to both Reactome and KEGG databases. A Fisher's exact test was used to investigate which pathways were overrepresented based on the 81 proteins containing a regulated peptide using the human UniProt database as a background. The most significantly enriched pathways included those associated with blood coagulation,

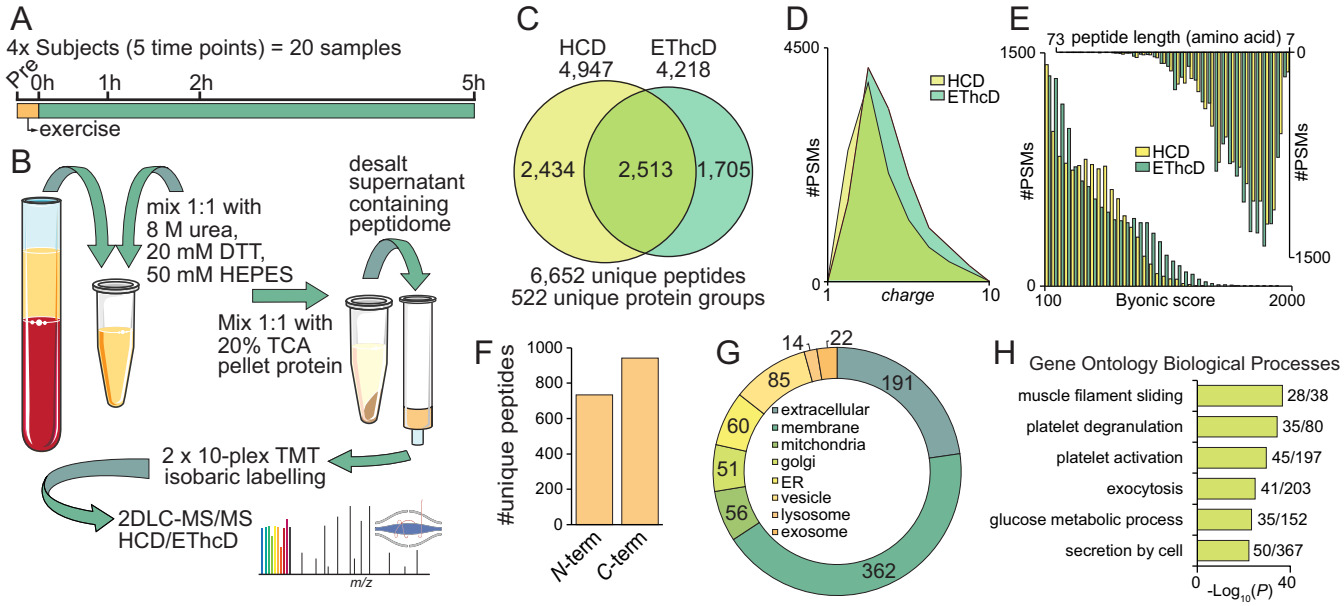


FIG. 2. **Multiplexed temporal quantification of the exercise-regulated plasma peptidome.** *A*, Exercise and sample collection time points of the four human subjects. *B*, Sample preparation strategy to isolate and quantify the plasma peptidome. Plasma proteins are precipitated with trichloroacetic acid (TCA) under denaturing conditions. The supernatant containing the peptidome is purified before isobaric labeling using 10-plex tandem mass tags (TMT) and analysis by multidimensional liquid chromatography - tandem mass spectrometry employing alternative fragmentation with HCD and EThcD. *C*, Overlap of unique peptides identified by HCD and EThcD. D, Charge distribution of peptide spectral matches (PSMs) between HCD and EThcD. *E*, Byonic score and peptide length distributions of PSMs between HCD and EThcD. *F*, Identified peptides starting or ending within 50 amino acids of the *N*- or *C*-terminus of the protein primary amino acid sequence. *G*, Gene Ontology cellular compartmentalization of peptidome derived proteins. *H*, Gene Ontology biological processes enrichment analysis of the peptidome derived proteins.

extracellular matrix interactions, muscle contraction, protein digestion and infection ($p < 0.05$; Fisher's exact test with Benjamini-Hochberg correction) (Fig. 3*B*). The proteins from these enriched pathways were analyzed by STRING (29) to retrieve high-confidence known protein-protein interactions. This revealed two interaction networks containing >3 protein members, including a network of collagens and apolipoproteins that were decreased postexercise (Fig. 3*C*), and a network consisting of fibrinogen and proteases that increased postexercise (Fig. 3*D*). This is consistent with previous reports of the promotion of thrombin and procoagulants by acute high-intensity exercise and the regulation of proteases (30). We next mapped the peptidome to a combined database of bioactive peptides retrieved from SATPdb (31), SwePep (32), and BIOPEP (33). Because exo-proteases generate a ladder of peptide variants with a single amino acid removed, we mapped each peptide identified to the full length amino sequence of the precursor protein present in the database. 1396 unique peptides could be mapped to the combined database, and they aligned to 48 proteins, simplifying the analysis of peptides with known bioactive properties (supplemental Table S3). The regulation of peptides derived from several hormones and growth factors was observed during or in the immediate hours post exercise (Fig. 3*E*). Three *C*-terminal insulin peptides (INS) were decreased following high-intensity exercise, and we observed a rise in a peptide fragment from the hormone glucagon (GCG) peaking at one-hour postexercise and progressively declined below preexercise levels by five hours. This provides validation of the method as exercise is well known to modulate the secretion of these pancreatic hormones in this way via enhanced circulating catecholamines (34, 35). Interestingly, the peptide spanning amino acids 57–85 of INS was quantified with and without *C*-terminal amidation. The amidated version returned to preexercise levels within 5 h of the cessation of exercise whereas the abundance of the nonamidated isoform remained lower than preexercise levels for the duration of the experiment. This suggests an increase in amidation activity. We also detected peptides derived from the appetite-regulating hormone, ghrelin (GHRL) in the peptidome. This was of interest as the effects of exercise on ghrelin levels is highly controversial (36). We observed five different ghrelin derived peptides with 1 peptide from the *N*-terminal region rapidly downregulated during exercise and then progressively returning to preexercise levels by 5 h, and 3 peptides from the *N*-terminal region progressively downregulated at 2 h postexercise and then returning to preexercise levels by 5 h (Fig. 3*E*). In addition, we found that a peptide spanning amino acids 35–46 of ghrelin progressively increased following exercise suggesting a further complex processing. A peptide spanning amino acids 143–183 of Proenkephalin-A (PENK) was rapidly downregulated with exercise. Endogenous opioid peptides includ-

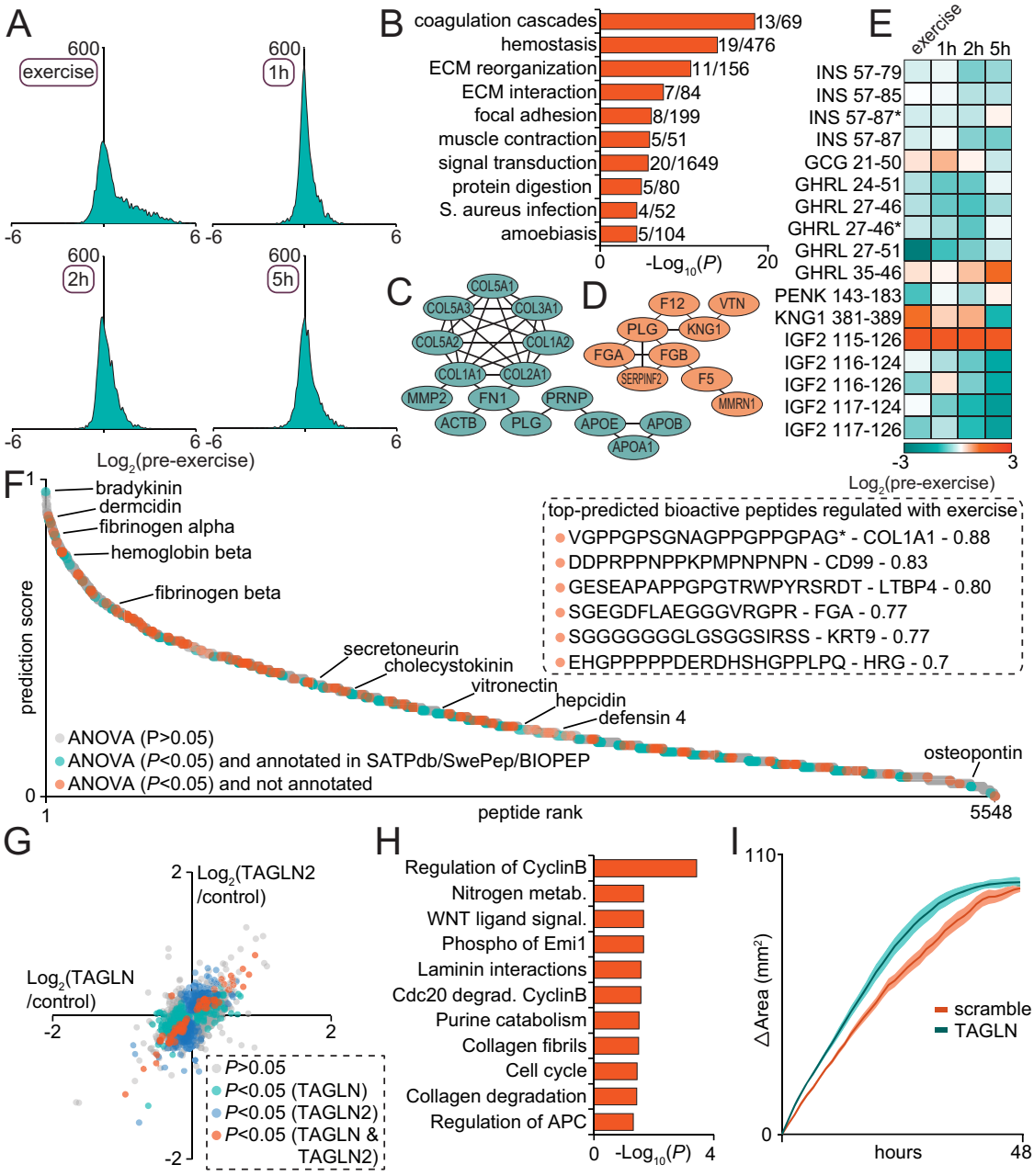


FIG. 3. **Quantification of the exercise-regulated plasma peptidome and analysis of bioactive peptides.** *A*, Distribution of quantified peptides expressed as a Log_2(fold-change) to preexercise levels at each of the four time points; during exercise, 1 h post exercise, 2 h post exercise, and 5 h post exercise ($n = 4$). *B*, Reactome and KEGG pathway enrichment analysis of the proteins contributing to exercise-increased or -decreased peptides at any of the time points ($p < 0.02$; Fischer exact test with Benjamini-Hochberg FDR). *C*, STRING protein:protein interaction analysis of proteins contributing to decreased peptides during or following exercise. *D*, STRING protein:protein interaction analysis of proteins contributing to increased peptides during or following exercise. *E*, Quantification of exercise-regulated peptides derived from growth factors and hormones. Asterisks indicate a *C*-terminal amidated peptide. *F*, Prediction of quantified peptides with bioactivity using PeptideRanker. *G*, Proteomic analysis of L6 myoblasts treated with either a synthetic *C*-terminal peptide from TAGLN spanning amino acids 179–200, a synthetic *C*-terminal peptide from TAGLN2 spanning amino acids 178–199, or a PBS control (1 μg/ml; 7 days; $n = 4$). Plotted is the Log_2(fold-change to control) of TAGLN compared with TAGLN2. Highlighted is significantly regulated proteins unique or common to the two peptide treatments (*t* test with Benjamini-Hochberg FDR). *H*, Reactome and KEGG pathway enrichment analysis of the proteins regulated following treatment of L6 myoblasts with either a synthetic *C*-terminal peptide from TAGLN spanning amino acids 179–200, or a synthetic *C*-terminal peptide from TAGLN2 spanning amino acids 178–199 (1 μg/ml; 7 days; $n = 4$) ($p < 0.05$; Fischer exact test with Benjamini-Hochberg FDR). *I*, Live cell migration assay of L6 myoblasts treated with a synthetic *C*-terminal peptide from TAGLN spanning amino acids 179–200, or a scramble control peptide (1 μg/ml; 7 days; $n = 4$). Plotted is the average difference in surface area migration over time. Shaded plots S.E.

ing enkephalins, endorphins and dynorphins play important roles in mood, pain perception, cardiac function and cellular growth, and have well established roles in the beneficial effects of exercise on cardiovascular function (37). Previous studies quantifying changes in opioid peptides primarily rely on the use of radioimmunoassays targeting a single epitope. However, these hormones display a complex pattern of protease processing producing a variety of peptides. Therefore, quantifying the suite of peptides generated is required to delineate specific opioid peptide regulation. More than 22 peptides derived from Kininogenin-1 (KNG1) were significantly increased during exercise including the peptide spanning amino acids 381–389 producing Bradykinin which has well established roles in blood vessel dilation and the regulation of blood pressure following exercise (38). A complex series of sequential proteolytic events generate this peptide which also resulted in several other peptides concomitantly decreasing. Exercise also resulted in a complex regulation of peptides derived from insulin-like growth factor 2 (IGF2). Plasma IGF1/2 levels have previously been shown to be regulated with exercise however the synthesis and degradation in response to different exercise intensities is incompletely understood (39). These data highlight the ability of peptidomics to quantify complex hormonal regulation following exercise.

Prediction of Peptides with Bioactivity—We next performed an analysis of peptides with novel bioactivity using the neural network machine learning predictor, PeptideRanker (40). Peptides were ranked and classified as either; (1) regulated with exercise ($p < 0.05$; ANOVA with permutation-based correction), and/or (2) previously annotated in the combined database of bioactive peptides described above (Fig. 3*F*). Of the 425 unique peptides regulated with exercise, 141 were not annotated in the combined bioactive peptide database. We were particularly interested in the prediction of peptides from the *C*-terminus of both transgelin (TAGLN) and transgelin-2 (TAGLN2) that were upregulated in plasma with exercise. TAGLN and TAGLN2 are evolutionary conserved cytoplasmic proteins associated with actin cytoskeleton remodeling, and smooth muscle/epithelial cell differentiation (41). More recently, TAGLN and TAGLN2 have been associated with altered expression in the development of several cancers. TAGLN and TAGLN2 share 65% sequence similarity and both contain actin binding calponin homology (CH) domains taking up the majority of the protein sequence, and a calponin-like repeat in the *C*-terminus. Interestingly, the peptides identified arose from exclusive cleavage of the calponin-like repeat. To our knowledge, this is the first identification and regulation of TAGLN and TAGLN2 peptides in plasma. To investigate bioactivity of these peptides, we treated L6 myoblasts *in vitro* with a synthetic *C*-terminal peptide from TAGLN spanning amino acids 179–200, or a synthetic *C*-terminal peptide from TAGLN2 spanning amino acids 178–199 (1 μg/ml; 7 days), and compared the treatments to control myoblasts cultured for the same amount of time with the addition of only PBS. Proteomic analysis by single-shot nanoUHPLC-MS/MS was performed on four biological replicates and quantification performed by LFQ. A total of 6095 proteins were identified with 4710 quantified in at least three biological replicates (supplemental Table S4). A total of 222 and 307 proteins were regulated by the TAGLN and TAGLN2 peptides relative to the control treatments, respectively, with 61 proteins coregulated by both TAGLN and TAGLN2 ($p < 0.05$; two-sample *t* test with Benjamini-Hochberg correction) (Fig. 3*G*). To gain an overview of the regulated pathways, all the proteins were mapped to Reactome and KEGG databases. A pathway enrichment analysis was performed with a Fisher's exact test using all the proteins significantly regulated by TAGLN and/or TAGLN2. The most significantly enriched pathways were associated with the regulation of Cyclin B, the anaphase promoting complex and cell cycle; regulation of nitrogen/purine metabolism; and regulation of extracellular matrix including collagen and laminin interactions ($p < 0.05$; Fisher's exact test with Benjamini-Hochberg correction) (Fig. 3*H*). These data suggest the *C*-terminal peptides of TAGLN and TAGLN2 regulate cell growth or cell movement. To functionally assess this regulation, we performed an *in vitro* cell migration assay on L6 myoblasts treated with the TAGLN synthetic peptide. Experiments were also performed with a control scramble peptide made up of the same amino acid composition of TAGLN *C*-terminal peptide using the ExPASy RandSeq tool. Live-cell microscopy revealed chronic treatment with the TAGLN C-terminal peptide significantly increased cell migration compared with scramble control using a wound healing assay over a period of 48 h (Fig. 3*I*). Taken together, these analyses reveal the ability of quantitative peptidomics to identify and predict peptides with bioactive properties.

Protease Mapping—The regulation of proteases and the peptidome are vital for many normal and disease processes (42). Therefore, understanding the activity of proteases and the substrate cleavage sites is a fundamental endeavor to unravel the peptidome. We used the Proteasix peptide-centric tool to perform large-scale automated mapping of cleavage sites to experimentally observed proteases (43). A total of 541 cleavage sites were mapped to 94 experimentally observed upstream proteases (supplemental Table S5). The 541 cleavage sites were mapped to 509 unique peptide sequences (some cleavage sites were mapped to both the *N*- and *C*-terminus of one peptide). Using the direction of fold-change, an enrichment analysis revealed that 9 proteases were associated with decreased activity (Fig. 4*A*) and 14 proteases were associated with increased activity during or post exercise (Fig. 4*B*) ($p < 0.05$; Fischer exact test with Benjamini-Hochberg correction). This identified a variety of matrix metalloproteases (MMPs) associated with decreased activity immediately following exercise, as well as several proteases associated with increased activity including kallikreins, plasminogen and calpain. A more detailed site-specific analysis of the cleavage sites was next performed. Fuzzy c-means clustering

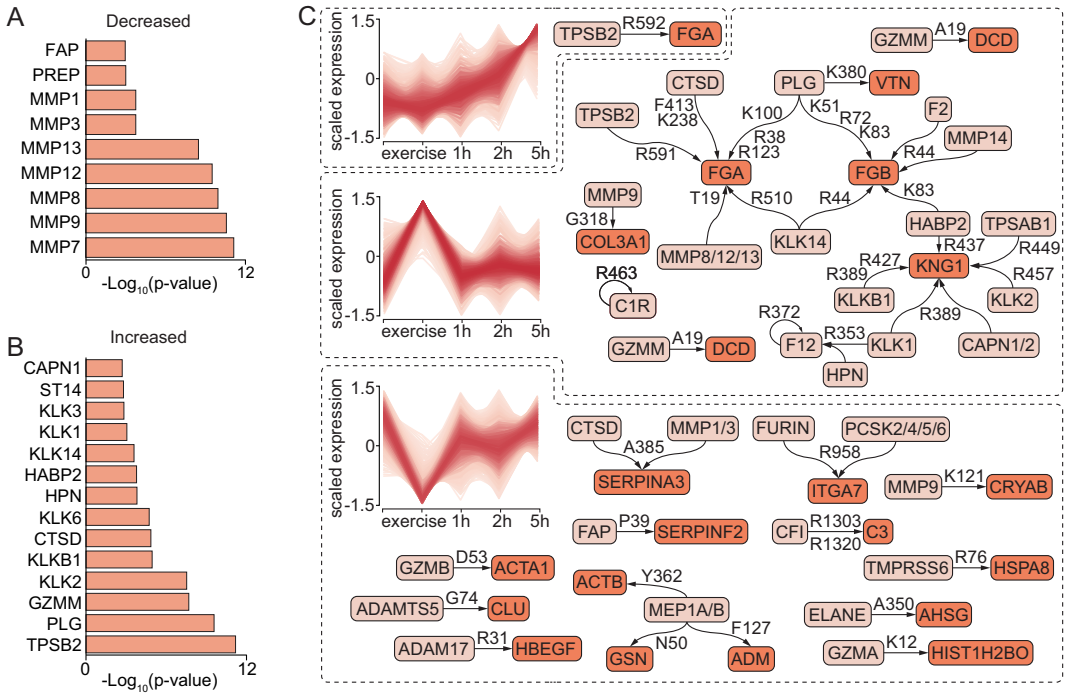
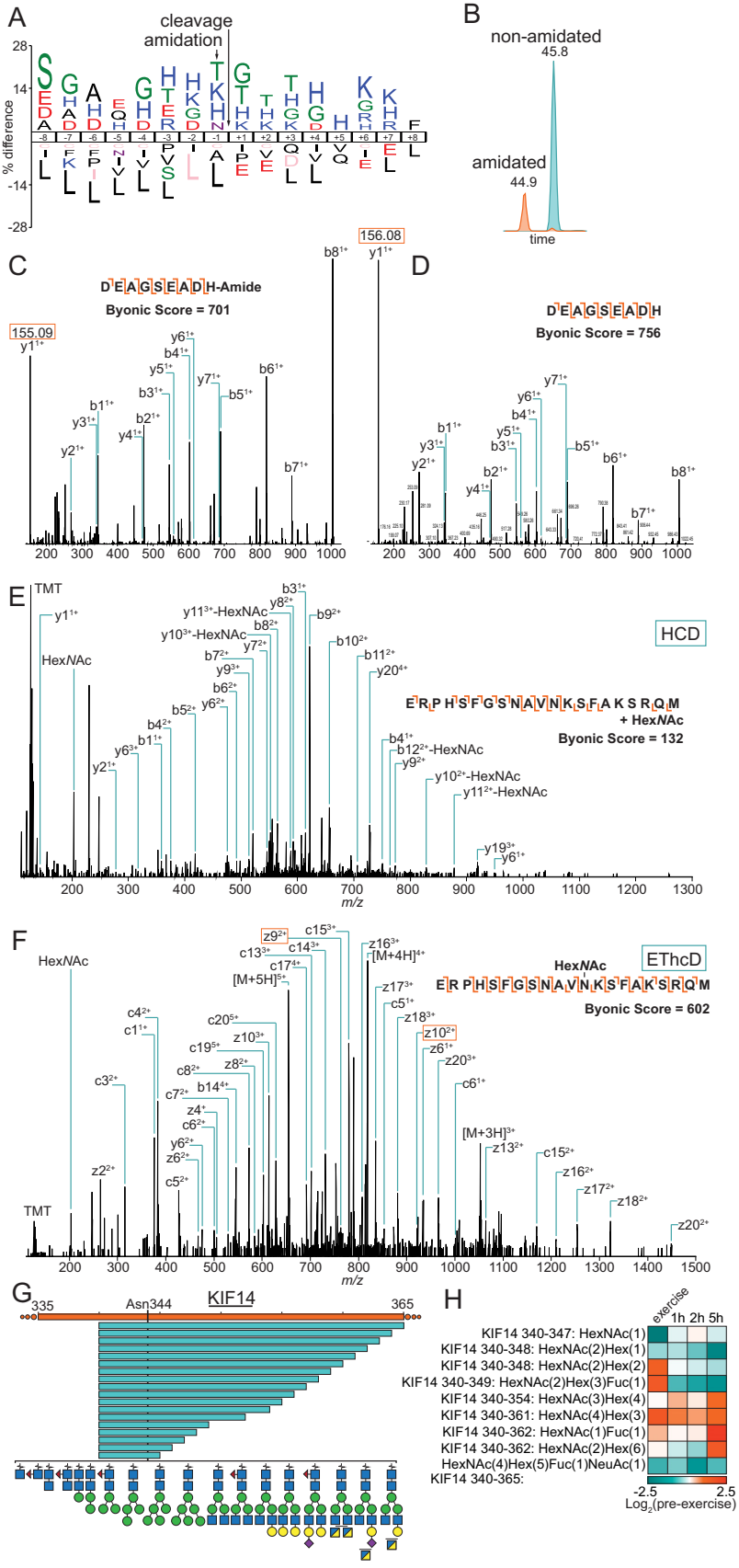


FIG. 4. **Plasma protease activity during and following exercise.** *A*, Enrichment of proteases with decreased activity based on the observed fold-changes of known peptide substrate cleavage events during or post exercise ($p < 0.02$; Fischer exact test; p). *B*, Enrichment of proteases with increased activity based on the observed fold-changes of known peptide substrate cleavage events during or post exercise ($p < 0.02$; Fischer exact test). *C*, Fuzzy c-means clustering and site-specific protease mapping of peptides during or post exercise. Dotted lines indicate the clusters containing the site-specific network.

of the temporal data revealed three distinct clusters with the strongest membership (Fig. 4*C*). This included an immediate up or downregulation of peptides during high-intensity exercise that returned to preexercise levels with 1 h, and a minor but later regulation of peptides >2 h following exercise cessation. This suggests there are multiple underlying temporal process potentially driven by a time-dependent regulation of proteases. The experimentally observed upstream proteases from each of the clusters were mapped onto a site-specific protease-substrate network. This clustering and network analysis provides a global view of plasma protease regulation with exercise combined with site-specific protease cleavage events.

Post-translational Modifications (PTMs) of the Exercise-regulated Peptidome—Many hormones are post-translationally modified. Two common modifications include *C*-terminal amidation and glycosylation which have been associated with regulating hormonal activity. Enzymatic *C*-terminal amidation is generated by the peptidylglycine alpha-amidating monooxygenase (PAM), which catalyzes the oxidative cleavage of a Gly in the +1 position resulting in amidation plus glyoxylate (44). Peptides containing Leu, Ala, Lys or Glu in the +1 position are not substrates of PAM suggesting that Gly is mandatory. Nonenzymatic *C*-terminal alpha-amidation through metal-catalyzed oxidative cleavage has previously been described (45). However, much less is known about the specific substrates or conditions and the alpha-amidation reactions occurring *in vivo* are incompletely understood. A total of 691 unique *C*-terminally amidated peptides were quantified at a 1% FDR (supplemental Table S3). Interestingly, only 117 peptides contained Gly in the +1 position. Motif analysis revealed overrepresentation of Gly, Thr, His and Lys in the +1 position, and an enrichment of basic amino acids surrounding the *C*-terminal cleavage and amidation site (Fig. 5*A*). A recent comprehensive neuropeptidomic analysis of rat brain identified 438 unique *C*-terminal amidated peptides with only 130 containing Gly in the +1 position (1). Consistent with our observations, basic amino acids were also enriched around the *C*-terminal cleavage and amidation site. The observation of a large percentage of amidated peptides which do not contain Gly in the +1 position suggest nonenzymatic *C*-terminal amidation is more prevalent than previously thought and/or these represent miss-identifications. The amidation reaction induces a mass shift of −0.984 Da and may result in miss annotation by incorrect mono-isotopic peak detection. To investigate this further we also analyzed the plasma peptidomics data with MaxQuant/Andromeda and manually inspected MS/MS spectra using the Expert Annotation Viewer (46). MaxQuant/Andromeda identified a total of 3,376 unique peptides at a 1% PSM and protein FDR (2416 HCD and 2543 EThcD), which included 139 unique *C*-terminally amidated peptides (supplemental Table S6). Of 139 *C*-



terminally amidated peptides identified with MaxQuant/Andromeda, 115 were also identified with Byonic. Amidated peptides identified by both Byonic and MaxQuant/Andromeda were selected for manual analysis of the precursor MS monoisotopic annotation, and MS/MS spectra to inspect the diagnostic y1- or z1-ion monoisotopic mass for *C*-terminal amidation or a series of >3 y-ions monoisotopic annotation. Of the 115 unique peptides investigated, a total of 99 unique peptides were manually verified to contain *C*-terminal amidation with 75 of these not containing Gly in the +1 position (supplemental Table S6 and supplemental Fig. S1). Interestingly, of the 75 *C*-terminally amidated peptides manually verified that did not contain Gly in the +1 position, 55 were also identified in their nonamidated form *i.e.* the same peptide sequence was identified with and without amidation. These peptides contained distinct chromatographic elution profiles further supporting the observation of *C*-terminal amidation without the presence of Gly in the +1 position. Fig. 5*B* shows the chromatographic elution profile of a peptide from Fibrinogen Alpha spanning amino acids 605–613 with and without *C*-terminal amidation which does not contain Gly in the +1 position. Fig. 5*C* and 5*D* show the annotated HCD MS/MS spectra highlighting the diagnostic y1-ion of the amidated and nonamidated peptides, respectively.

To our knowledge, the *N*-linked glycopeptidome of human plasma has not been investigated. Therefore, our analysis also included an investigation of this PTM using a modification database consisting of 48 *N*-glycan monosaccharide compositions most frequently observed in human plasma. To reduce the search space, we restricted our analysis of *N*-glycan modification to only modified asparagine residues in the *N*-linked glycosylation motif NxS/T (x≠P) using the Byonic algorithm. Out of the 6652 unique peptides identified by Byonic, a total of 733 unique *N*-glycopeptides were identified at a 1% PSM FDR. The remaining 5919 nonglycosylated peptides identified by Byonic were next compared with the 3376 nonglycosylated peptides identified by MaxQuant/Andromeda. Interestingly, 2945 unique peptides were identified by both algorithms whereas 2974 and 431 peptides were identified exclusively by Byonic and MaxQuant/Andromeda, respectively. Out of the 733 unique N-glycopeptides identified, 408 were quantified in at least two out of the four subjects across all five time points, and 92 regulated by ± 2-fold during or in the immediate hours postexercise (supplemental Table S2). The combined use of complementary fragmentation approaches with HCD and EThcD enabled inference of monosaccharide compositions, and localization of glycosylation sites. Fig. 5*E* and 5*F* shows the HCD and EThcD MS/MS spectra of the glycopeptide derived from Complement Component C8-Beta (C8B) spanning amino acids 33–53 with a single Hex*N*Ac modification of Asn44, respectively. The glycopeptide data also showed extensive glycan heterogeneity combined with peptide variants arising from exopeptidase activity. More than 77 unique glycopeptides were identified from Kinesin-like Protein 14 (KIF14). These were made up of 18 peptides containing glycosylation of Asn344 with one of 17 different glycan compositions (Fig. 5*G*). These glycopeptides displayed a complex pattern of regulation during or in the immediate hours following exercise with 10 *N*-linked glycopeptides regulated by ± 2-fold (Fig. 5F). Considering the regulatory features of these data, future investigations of the functional significance are warranted. Collectively, these data demonstrate the human plasma peptidome is subject to PTM regulation, and the combined use of complementary fragmentation techniques with TMT labeling is an effective strategy to characterize and quantify modifications on endogenous peptides.

DISCUSSION

This study presents a comprehensive analysis of the human plasma peptidome combined with multiplexed temporal quantification during and following exercise. This analysis, which combined alternative fragmentation approaches, and high resolution and accurate mass detection to significantly expand coverage (47, 48), is one of the deepest profiles of the plasma peptidome. This is attributed to; (1) the charge-density dependent fragmentation efficiency of the dissociation method in which collisional-induced approaches favor lower charge densities and electron-induced approaches favor higher charge densities (49) and, (2) the high-resolution and accurate mass measurements afford decreased fragment ion mass tolerance to increase identifications using an FDR-controlled target-decoy strategy with no enzyme searching (48).

Our multiplexed temporal analysis of the human plasma peptidome identified a complex regulation of endogenous peptides in response to exercise. Surprisingly, this regulation was extremely dynamic, occurring within minutes of exercise commencement suggesting higher resolution time-points would reveal finer details of temporal regulation. Importantly, the analysis quantified hormones known to be regulated with exercise that have a well described physiological role such as an increase in bradykinin which regulates vasodilation, a de-

FIG. 5. **Analysis of the post-translationally modified peptidome.** *A*, Logo analysis of *C*-terminally amidated peptides. *B*, Extracted ion chromatograms of the amidated and nonamidated peptides derived from Fibrinogen Alpha spanning amino acids 605–613. *C*, HCD MS/MS annotation of the amidated and, *D*, HCD MS/MS annotation of the nonamidated peptide derived from Fibrinogen Alpha spanning amino acids 605–613. All fragment ions are annotated within 20 ppm. Box indicates diagnostic ions for localization of the modification site. *E*, HCD MS/MS and, *F*, EThcD MS/MS annotation of the peptide glycopeptide derived from Complement Component C8-Beta (C8B) spanning amino acids 33–53 with a single HexNAc modification of Asn44. All fragment ions are annotated within 20 ppm. Box indicates diagnostic ions for localization of the modification site. *G*, Peptide variants and glycan heterogeneity at Asn344 on Kinesin-like Protein 14 (KIF14). *H*, Quantification of glycopeptide variants modified at Asn344 during and following exercise.

crease in insulin which regulates several outcomes including blood glucose levels, and a decrease in the level of several peptides derived from ghrelin, which regulates hunger. This emphasizes the use of peptidomics to simultaneously quantify multiple peptide hormones to obtain a systems-wide view of acute peptide regulation in the plasma of exercising humans. In addition, it also enables identification of novel peptides that may influence interorgan communication, cellular physiology and contribute to the beneficial effects of exercise. We show novel peptides generated from the *C*-terminus of transgelin and transgelin-2 containing calponin-like domains are increased in the plasma following exercise. *In vitro* experiments using cultured cells suggest these peptides have bioactivity and regulate cell migration via effects on cell cycle and extracellular matrix remodeling. Transgelins are primarily expressed in smooth muscle cells and are markers of differentiation (50). The calponin-like domains of transgelins are required for actin binding, and mice lacking trangelin ($TAGLN^{-/-}$) have reduced actin content in arteries and veins with altered basal and stimulated force contractions (51). Several studies have also observed altered expression of transgelin in a variety of cancers (for an extensive review see (41)), and associated with a tumor suppressor role via the regulation of MMP9 (52) and androgen stimulated cell growth (53). It is exciting to speculate if the regulation of transgelins may play a role in the anticancer effects of exercise however further studies are required to investigate this.

A further layer of information that we could obtain from our analysis concerned protein degradation. Proteolytic products of precursor proteins indicated a surprisingly complex pattern of regulated protein degradation in the plasma after exercise. As to whether any of these peptides have *in vivo* function remains to be determined. For instance, we observed multiple peptides derived from ghrelin, and it will be interesting to see if any of these might bind to the ghrelin receptor and regulate its activity. This is exciting because the role of ghrelin in the anorectic effects of exercise is highly controversial (36). Alternatively, the degradation of hormones may represent another layer of regulation to deactivate them at or near their putative site of action. Complementary quantification of the plasma proteome may provide additional information to investigate these differences.

Our deep peptidome analysis with complementary fragmentation techniques enabled us to investigate other PTMs commonly occurring on secreted hormones including *C*-terminal amidation and glycosylation that regulated peptide stability. We provide further evidence that *C*-terminal amidation is more abundant than previously thought and many of these peptides do not contain Gly in the +1 position. This suggests nonenzymatic amidation is prevalent in the plasma or other unidentified mechanisms are present that regulate peptide stability. It is well known that exercise regulates a complex series of reactive oxygen species (ROS) and it will be interesting to further investigate the oxidative cleavage of proteins in plasma to generate amidated peptides.

Exercise has beneficial effects on many organs of the body and it is well known that many of these effects are mediated by circulating factors. It is highly likely that several of these beneficial effects are mediated at least in part by the peptidome. The quantification of peptides regulated with exercise will be a valuable resource to further characterize bioactive endogenous peptides. Finally, peptidomics shows great promise as a diagnostic tool considering the myriad of peptide patterns that emerge from this relatively simple method.

Acknowledgments—We thank Ben Crossett, Stuart Cordwell and The Sydney University Mass Spectrometry Facility. We also thank Sean Humphrey for critical evaluation of the manuscript.

DATA AVAILABILITY

All RAW and processed data associated with the manuscript has been deposited in PRIDE proteomeXchange and can be accessed at http://www.proteomexchange.org/. Peptidomics data associated with optimisation of sample preparation is available in project accession PXD007191 (username: reviewer97192@ebi.ac.uk; password: zkVstwAw). Data associated with temporal quantification of the exercise-regulated peptidome is available in project accession PXD004781 (username: reviewer61660@ebi.ac.uk; password: NaHrEPqj). Data associated with the proteomic analysis of L6 myoblasts treated with transgelin C-terminal peptides is available in project accession PXD007176 (username: reviewer06819@ebi.ac.uk; password: 8pQVmNeV).

* This work was supported by a National Health and Medical Research Council (NHMRC) project grant (D.E.J and B.L.P) and by the Danish Research Council for Independent Research/Medicine (E.A.R.). B.L.P is a recipient of an NHMRC Early Career Fellowship. D.E.J. is a recipient of an NHRMC Senior Research Fellowship. The contents of the published material are solely the responsibility of the individual authors and do not reflect the view of NHMRC.

S This article contains supplemental material.

** To whom correspondence should be addressed: Charles Perkins Centre, School of Life and Environmental Sciences, University of Sydney, Sydney, NSW 2006, Australia. Tel.: +61-2-8627-5731. E-mail: david.james@sydney.edu.au.

REFERENCES

1. Secher, A., Kelstrup, C. D., Conde-Frieboes, K. W., Pyke, C., Raun, K., Wulff, B. S., and Olsen, J. V. (2016) Analytic framework for peptidomics applied to large-scale neuropeptide identification. *Nat. Commun.* **7,** 11436
2. Shen, Y., Liu, T., Tolic, N., Petritis, B. O., Zhao, R., Moore, R. J., Purvine, S. O., Camp, D. G., and Smith, R. D. (2010) Strategy for degradomic-peptidomic analysis of human blood plasma. *J. Proteome Res.* **9,** 2339–2346
3. Aristoteli, L. P., Molloy, M. P., and Baker, M. S. (2007) Evaluation of endogenous plasma peptide extraction methods for mass spectrometric biomarker discovery. *J. Proteome Res.* **6,** 571–581
4. Polson, C., Sarkar, P., Incledon, B., Raguvaran, V., and Grant, R. (2003) Optimization of protein precipitation based upon effectiveness of protein removal and ionization effect in liquid chromatography-tandem mass spectrometry. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **785,** 263–275
5. Hu, L., Boos, K. S., Ye, M., Wu, R., and Zou, H. (2009) Selective on-line serum peptide extraction and multidimensional separation by coupling a

restricted-access material-based capillary trap column with nanoliquid chromatography-tandem mass spectrometry. *J. Chromatogr. A* **1216,** 5377–5384
6. Boonen, K., Landuyt, B., Baggerman, G., Husson, S. J., Huybrechts, J., and Schoofs, L. (2008) Peptidomics: the integrated approach of MS, hyphenated techniques and bioinformatics for neuropeptide analysis. *J. Sep. Sci.* **31,** 427–445
7. Nanni, P., Levander, F., Roda, G., Caponi, A., James, P., and Roda, A. (2009) A label-free nano-liquid chromatography-mass spectrometry approach for quantitative serum peptidomics in Crohn's disease patients. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **877,** 3127–3136
8. Che, F. Y., and Fricker, L. D. (2005) Quantitative peptidomics of mouse pituitary: comparison of different stable isotopic tags. *J. Mass Spectrom.* **40,** 238–249
9. Hardt, M., Witkowska, H. E., Webb, S., Thomas, L. R., Dixon, S. E., Hall, S. C., and Fisher, S. J. (2005) Assessing the effects of diurnal variation on the composition of human parotid saliva: quantitative analysis of native peptides using iTRAQ reagents. *Anal. Chem.* **77,** 4947–4954
10. Bourdetsky, D., Schmelzer, C. E., and Admon, A. (2014) The nature and extent of contributions by defective ribosome products to the HLA peptidome. *Proc. Natl. Acad. Sci. U.S.A.* **111,** E1591–E1599
11. Good, D. M., Zurbig, P., Argiles, A., Bauer, H. W., Behrens, G., Coon, J. J., Dakna, M., Decramer, S., Delles, C., Dominiczak, A. F., Ehrich, J. H., Eitner, F., Fliser, D., Frommberger, M., Ganser, A., Girolami, M. A., Golovko, I., Gwinner, W., Haubitz, M., Herget-Rosenthal, S., Jankowski, J., Jahn, H., Jerums, G., Julian, B. A., Kellmann, M., Kliem, V., Kolch, W., Krolewski, A. S., Luppi, M., Massy, Z., Melter, M., Neususs, C., Novak, J., Peter, K., Rossing, K., Rupprecht, H., Schanstra, J. P., Schiffer, E., Stolzenburg, J. U., Tarnow, L., Theodorescu, D., Thongboonkerd, V., Vanholder, R., Weissinger, E. M., Mischak, H., and Schmitt-Kopplin, P. (2010) Naturally occurring human urinary peptides for use in diagnosis of chronic kidney disease. *Mol. Cell. Proteomics* **9,** 2424–2437
12. Ling, X. B., Sigdel, T. K., Lau, K., Ying, L., Lau, I., Schilling, J., and Sarwal, M. M. (2010) Integrative urinary peptidomics in renal transplantation identifies biomarkers for acute rejection. *J. Am. Soc. Nephrol.* **21,** 646–653
13. Zimmerli, L. U., Schiffer, E., Zurbig, P., Good, D. M., Kellmann, M., Mouls, L., Pitt, A. R., Coon, J. J., Schmieder, R. E., Peter, K. H., Mischak, H., Kolch, W., Delles, C., and Dominiczak, A. F. (2008) Urinary proteomic biomarkers in coronary artery disease. *Mol. Cell. Proteomics* **7,** 290–298
14. Antwi, K., Hostetter, G., Demeure, M. J., Katchman, B. A., Decker, G. A., Ruiz, Y., Sielaff, T. D., Koep, L. J., and Lake, D. F. (2009) Analysis of the plasma peptidome from pancreas cancer patients connects a peptide in plasma to overexpression of the parent protein in tumors. *J. Proteome Res.* **8,** 4722–4731
15. Bassani-Sternberg, M., Barnea, E., Beer, I., Avivi, I., Katz, T., and Admon, A. (2010) Soluble plasma HLA peptidome as a potential source for cancer biomarkers. *Proc. Natl. Acad. Sci. U.S.A.* **107,** 18769–18776
16. Liu, F., Zhao, C., Liu, L., Ding, H., Huo, R., and Shi, Z. (2016) Peptidome profiling of umbilical cord plasma associated with gestational diabetes-induced fetal macrosomia. *J. Proteomics* **139,** 38–44
17. Hawley, J. A., Hargreaves, M., Joyner, M. J., and Zierath, J. R. (2014) Integrative biology of exercise. *Cell* **159,** 738–749
18. Hellsten, Y., Nyberg, M., Jensen, L. G., and Mortensen, S. P. (2012) Vasodilator interactions in skeletal muscle blood flow regulation. *J. Physiol.* **590,** 6297–6305
19. Palmisano, G., Lendal, S. E., Engholm-Keller, K., Leth-Larsen, R., Parker, B. L., and Larsen, M. R. (2010) Selective enrichment of sialic acid-containing glycopeptides using titanium dioxide chromatography with analysis by HILIC and mass spectrometry. *Nat. Protoc.* **5,** 1974–1982
20. Frese, C. K., Altelaar, A. F., van den Toorn, H., Nolting, D., Griep-Raming, J., Heck, A. J., and Mohammed, S. (2012) Toward full peptide sequence coverage by dual fragmentation combining electron-transfer and higher-energy collision dissociation tandem mass spectrometry. *Anal. Chem.* **84,** 9668–9673
21. Bern, M., Kil, Y. J., and Becker, C. (2012) Byonic: advanced peptide and protein identification software. *Curr. Protoc. Bioinformatics* Chapter 13, Unit 13.20
22. Tyanova, S., Temu, T., and Cox, J. (2016) The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protocols* **11,** 2301–2319
23. Cox, J., Hein, M. Y., Luber, C. A., Paron, I., Nagaraj, N., and Mann, M. (2014) Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell. Proteomics* **13,** 2513–2526
24. Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M. Y., Geiger, T., Mann, M., and Cox, J. (2016) The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13,** 731–740
25. Rigbolt, K. T., Vanselow, J. T., and Blagoev, B. (2011) GProX, a user-friendly platform for bioinformatics analysis and visualization of quantitative proteomics data. *Mol. Cell. Proteomics* **10,** O110.007450
26. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J. Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., and Cardona, A. (2012) Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9,** 676–682
27. Arganda-Carreras, I., Kaynig, V., Rueden, C., Eliceiri, K. W., Schindelin, J., Cardona, A., and Seung, H. S. (2017) Trainable Weka Segmentation: a machine learning tool for microscopy pixel classification. *Bioinformatics* **33,** 2424–2426
28. Gundry, R. L., Fu, Q., Jelinek, C. A., Van Eyk, J. E., and Cotter, R. J. (2007) Investigation of an albumin-enriched fraction of human serum and its albuminome. *Proteomics. Clin. Appl.* **1,** 73–88
29. Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K. P., Kuhn, M., Bork, P., Jensen, L. J., and von Mering, C. (2015) STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **43,** D447–D452
30. Chen, Y. W., Chen, J. K., and Wang, J. S. (2010) Strenuous exercise promotes shear-induced thrombin generation by increasing the shedding of procoagulant microparticles from platelets. *Thromb. Haemost.* **104,** 293–301
31. Singh, S., Chaudhary, K., Dhanda, S. K., Bhalla, S., Usmani, S. S., Gautam, A., Tuknait, A., Agrawal, P., Mathur, D., and Raghava, G. P. (2016) SATPdb: a database of structurally annotated therapeutic peptides. *Nucleic Acids Res.* **44,** D1119–D1126
32. Falth, M., Skold, K., Norrman, M., Svensson, M., Fenyo, D., and Andren, P. E. (2006) SwePep, a database designed for endogenous peptides and mass spectrometry. *Mol. Cell. Proteomics* **5,** 998–1005
33. Minkiewicz, P., Dziuba, J., Iwaniak, A., Dziuba, M., and Darewicz, M. (2008) BIOPEP database and other programs for processing bioactive peptide sequences. *J AOAC Int* **91,** 965–980
34. Marliss, E. B., and Vranic, M. (2002) Intense exercise has unique effects on both insulin release and its roles in glucoregulation: implications for diabetes. *Diabetes* **51,** S271–S283
35. Galbo, H., Holst, J. J., and Christensen, N. J. (1975) Glucagon and plasma catecholamine responses to graded and prolonged exercise in man. *J. Appl. Physiol.* **38,** 70–76
36. King, J. A., Wasse, L. K., Stensel, D. J., and Nimmo, M. A. (2013) Exercise and ghrelin. A narrative overview of research. *Appetite* **68,** 83–91
37. Grossman, A., and Sutton, J. R. (1985) Endorphins: what are they? How are they measured? What is their role in exercise? *Med. Sci. Sports Exercise* **17,** 74–81
38. Duncker, D. J., and Bache, R. J. (2008) Regulation of coronary blood flow during exercise. *Physiol. Rev.* **88,** 1009–1086
39. Bang, P., Brandt, J., Degerblad, M., Enberg, G., Kaijser, L., Thoren, M., and Hall, K. (1990) Exercise-induced changes in insulin-like growth factors and their low molecular weight binding protein in healthy subjects and patients with growth hormone deficiency. *Eur. J. Clin. Invest.* **20,** 285–292
40. Mooney, C., Haslam, N. J., Pollastri, G., and Shields, D. C. (2012) Towards the improved discovery and design of functional peptides: common features of diverse classes permit generalized prediction of bioactivity. *PLoS ONE* **7,** e45012
41. Dvorakova, M., Nenutil, R., and Bouchal, P. (2014) Transgelins, cytoskeletal proteins implicated in different aspects of cancer development. *Expert Rev. Proteomics* **11,** 149–165
42. Turk, B. (2006) Targeting proteases: successes, failures and future prospects. *Nat. Rev. Drug Discov.* **5,** 785–799
43. Klein, J., Eales, J., Zurbig, P., Vlahou, A., Mischak, H., and Stevens, R. (2013) Proteasix: a tool for automated and large-scale prediction of proteases involved in naturally occurring peptide generation. *Proteomics* **13,** 1077–1082

44. Bradbury, A. F., Finnie, M. D., and Smyth, D. G. (1982) Mechanism of C-terminal amide formation by pituitary enzymes. *Nature* **298,** 686–688
45. Berlett, B. S., and Stadtman, E. R. (1997) Protein oxidation in aging, disease, and oxidative stress. *J. Biol. Chem.* **272,** 20313–20316
46. Neuhauser, N., Michalski, A., Cox, J., and Mann, M. (2012) Expert system for computer-assisted annotation of MS/MS spectra. *Mol. Cell. Proteomics* **11,** 1500–1509
47. Mommen, G. P., Frese, C. K., Meiring, H. D., van Gaans-van den Brink, J., de Jong, A. P., van Els, C. A., and Heck, A. J. (2014) Expanding the detectable HLA peptide repertoire using electron-transfer/higher-energy collision dissociation (EThcD). *Proc. Natl. Acad. Sci. U.S.A.* **111,** 4507–4512
48. Shen, Y., Tolic, N., Xie, F., Zhao, R., Purvine, S. O., Schepmoes, A. A., Moore, R. J., Anderson, G. A., and Smith, R. D. (2011) Effectiveness of CID, HCD, and ETD with FT MS/MS for degradomic-peptidomic analysis: comparison of peptide identification methods. *J. Proteome Res.* **10,** 3929–3943
49. Good, D. M., Wirtala, M., McAlister, G. C., and Coon, J. J. (2007) Performance characteristics of electron transfer dissociation mass spectrometry. *Mol. Cell. Proteomics* **6,** 1942–1951
50. Assinder, S. J., Stanton, J. A., and Prasad, P. D. (2009) Transgelin: an actin-binding protein and tumour suppressor. *Int. J. Biochem. Cell Biol.* **41,** 482–486
51. Zhang, J. C., Kim, S., Helmke, B. P., Yu, W. W., Du, K. L., Lu, M. M., Strobeck, M., Yu, Q., and Parmacek, M. S. (2001) Analysis of SM22alpha-deficient mice reveals unanticipated insights into smooth muscle cell differentiation and function. *Mol. Cell. Biol.* **21,** 1336–1344
52. Ai, T., Chen, F., Zhou, S., Zhang, J., Zheng, H., Zhou, Y., Hu, W., Liu, X., Li, L., and Lin, J. (2015) Magnetic bead-based serum peptidome profiling in patients with gestational diabetes mellitus. *Biomed. Res. Int.* **586309,** 2015
53. Yang, Z., Chang, Y. J., Miyamoto, H., Ni, J., Niu, Y., Chen, Z., Chen, Y. L., Yao, J. L., di Sant'Agnese, P. A., and Chang, C. (2007) Transgelin functions as a suppressor via inhibition of ARA54-enhanced androgen receptor transactivation and prostate cancer cell growth. *Mol. Endocrinol.* **21,** 343–358