ONLINE METHODS

Commercial instrumentation and materials are identified in this work to adequately describe the experimental procedure. Such identification does not imply recommendation or endorsement by the authors and the National Institute of Standards and Technology nor does it imply that the equipment, instruments or materials are necessarily the best available for the purpose.

Materials. The light ($^{12}C/^{14}N$) and heavy ($^{13}C/^{15}N$) forms of 11 unique signature peptides corresponding to the target proteins were synthesized and purified by Anaspec. Signature tryptic peptides containing C-terminal arginine and lysine residues were synthesized as $^{13}C_6$ and $^{13}C_6{}^{15}N_2$ analogs, respectively. Two tryptic peptides contained a $^{13}C_5$-valine residue and one N-terminal partial tryptic peptide was prepared as the $^{13}C_6$-leucine analog (**Table 1**). Target proteins were purchased either from Sigma (equine myoglobin, bovine myelin basic protein, bovine aprotinin, murine leptin and horseradish peroxidase) or from Scripps Laboratories (human C-reactive protein and human prostate-specific antigen). Pooled and filtered (0.2 μm) human K_2-EDTA plasma was purchased from Bioreclamation. ReproSil-Pur C18-AQ resin (3 μm particle size) was purchased from Dr. Maisch. Mass spectrometry grade Trypsin Gold was obtained from Promega. Iodoacetamide, dithiothreitol and urea were purchased from Sigma Chemical or from ThermoFisher Scientific.

Peptide purity of synthetic peptides and amino acid analysis. Peptide and isotopic purity of the synthetic peptides was estimated to be >98% as determined by LC-UV and matrix-assisted laser desorption ionization (MALDI)-MS at Anaspec. In addition, isotopic purity of heavy peptides was assessed at the National Institute of Standards and Technology (NIST) by MALDI-MS on a 4700 tandem time of flight (TOF)/TOF mass spectrometer (Applied Biosystems/MDS Analytical Technologies) in reflector mode. The laser intensity was adjusted to keep the maximum ion count below 25,000, preventing detector saturation and distortion of the observed peptide isotope distribution. For all heavy peptides, comparison between the observed isotope distributions obtained in the acquired spectra and simulated distributions, with varying percent incorporation of the heavy label, indicated that the isotopic peptide purity was ~99%. The exact concentrations of synthetic peptides and target proteins were determined by amino acid analysis after gas-phase acid hydrolysis followed by isotope dilution-LC-tandem MS (MS/MS)[28] at NIST.

ELISA. To determine the endogenous CRP and PSA concentrations of the human plasma used to prepare the study samples, the plasma was analyzed using a Quantikine Human C-Reactive Protein ELISA kit and a Quantikine Human Kallikrein 3/Prostrate Specific Antigen Protein ELISA kit (R&D Systems). The stock CRP and PSA solutions used to prepare the spiked plasma study samples were used to prepare calibrants for the respective ELISAs. For the CRP assay, the plasma sample was diluted 200-fold before analysis, whereas for the PSA assay, the plasma was assayed undiluted. Both assays were performed according to the manufacturer's instructions. The assay response was measured using a BioTek Synergy HT microplate reader. The plasma CRP concentration of the unspiked plasma was ~6.0 mg/l. The PSA level in the unspiked plasma was below the detection limit of the ELISA, which is about 1 ng/ml.

Sample preparation for study I: digested plasma spiked with signature peptides. The samples for studies I and II were prepared (including tryptic digestion) at NIST, and shipped to the eight participating laboratories. Stock solutions (100 pmol/μl) of the individual heavy and light peptides were dissolved in an aqueous solution of 30% acetonitrile (vol/vol) and 0.1% formic acid (vol/vol). Equimolar mixtures containing either the light or the heavy peptides were prepared at 1 pmol/μl.

As the background matrix, 1 ml of pooled human K_2EDTA plasma was diluted with 2 ml of 150 mM Tris, pH 8.0, containing 9 M urea and 30 mM dithiothreitol. The final protein concentration of human plasma before dilution was ~63.7 g/l by a bicinchoninic acid colorimetric assay (Pierce Biotechnology) and a 7% solution of BSA as a standard (NIST Standard Reference Material). Plasma proteins were reduced and denatured by heating for 30 min at 37 °C. The sample was cooled to ~20–23 °C before a 260 μl aliquot of 500 mM aqueous iodoacetamide was added to achieve a final concentration of 40 mM. The sample was incubated at ~20–23 °C for 30 min in the dark. Next, the plasma sample was diluted approximately tenfold with 100 mM Tris, pH 8.0, and digested with Promega Trypsin Gold (1 mg) at 37 °C. After 18 h, proteolysis was stopped by acidifying the solution to pH 2 with 1% (vol/vol) formic acid. The digested plasma was desalted using a 35 ml Oasis HLB LP solid-phase extraction cartridge (Waters) and the peptide eluate was lyophilized. Finally, the mixture of plasma peptides was reconstituted with 60 ml of an aqueous solution containing 0.6% (vol/vol) acetonitrile and 1% (vol/vol) formic acid, achieving a 60-fold dilution of the plasma that resulted in a final concentration of ~1 μg/μl (total protein).

A multistep process was used to prepare trypsin-digested plasma samples that contained varying amounts of the light peptide mixture (500, 275, 151, 83, 46, 25, 8.6, 2.9, 1.0 fmol/μl) and 50 fmol/μl of the heavy peptide mixture. First, aqueous 1,000 fmol/μl stock solutions of the light (solution A) and the heavy (solution B) peptides were prepared. Second, a 50 fmol/μl solution of the heavy signature peptides (solution C) was prepared by lyophilizing an aliquot (1 ml) of solution B and reconstituting with 20 ml of digested diluted plasma. Third, an aliquot (0.9 ml) of solution A was lyophilized and reconstituted with an appropriate volume (1.8 ml) of solution C to produce solution D (study I, sample J) that contained a mixture of light and heavy signature peptides at concentrations of 500 and 50 fmol/μl, respectively. Fourth, the remaining standards were prepared by serial dilution of solution D with solution C. Aliquots (25 μl) of the working standards were dispensed into polypropylene sample tubes, stored at −80 °C, and shipped frozen on dry ice to the participating laboratories.

Sample preparation for study II: digested plasma spiked with digested proteins. Individual solutions of the seven target proteins were prepared in water (ranging between 62 pmol/μl and 145 pmol/μl). Aliquots of these stock solutions were co-lyophilized and reconstituted in 100 mM Tris, pH 8.0, containing 6 M urea and 5 mM dithiothreitol to produce an equimolar mixture (100 pmol/μl). Reduction, denaturation and alkylation of the proteins were carried out as described for study I. Next, the protein mixture was diluted tenfold with 100 mM Tris, pH 8.0, and Promega Trypsin Gold was added at an enzyme/substrate ratio of 1:50 (wt/wt). Tryptic digestion was carried out as described above and the resulting peptide mixture was desalted using a 1 ml Oasis HLB solid phase extraction cartridge. The eluted peptides were lyophilized to dryness and reconstituted with an aqueous solution of 6% (vol/vol) acetonitrile and 1% (vol/vol) formic acid. Finally, study II samples containing 500, 275, 151, 83, 46, 25, 8.6, 2.9 or 1.0 fmol/μl of the trypsin-digested protein mixture and 50 fmol/μl of isotopically labeled signature peptides were prepared as described for study I above. Aliquots (25 μl) were dispensed into polypropylene sample tubes, stored at −80 °C, and shipped frozen on dry ice to the participating laboratories.

Sample preparation for study III: digestion of plasma spiked with target proteins. Stock solutions of human pooled K_2EDTA plasma, human pooled K_2EDTA plasma spiked with the seven target proteins spanning a concentration range of 0.06–30 pmol/μl, and a 500 fmol/μl mixture of the 11 isotopically labeled signature peptides were prepared and aliquoted at NIST (SOP, **Supplementary Methods**). A working solution was prepared by lyophilizing an aliquot of the 50 pmol/μl mixture of the seven target proteins and reconstituting the sample with whole plasma to a final concentration of 30 pmol/μl. Solutions with lower concentrations of spiked-in proteins were prepared by serial dilution of the 30 pmol/μl solution with plasma (SOP, **Supplementary Methods**). Study samples were aliquoted (35 μl), stored at −80 °C and shipped to the eight participating sites.

The remaining sample preparation steps were performed in triplicate (study IIIa, IIIb and IIIc) at each site. Plasma and spiked plasma samples were digested with trypsin using a scaled-down version of the protocol described for study I. Briefly, a 25 μl aliquot of each plasma sample was combined with 50 μl of buffer (300 mM Tris, pH 8/9M urea/20 mM DTT), reduced (30 min at 37 °C), and alkylated by adding 500 mM aqueous iodoacetamide (40 mM final concentration) and incubating at ~20–23 °C for 30 min in the dark. Samples were diluted tenfold with 100 mM Tris (pH 8) and digested with Promega Trypsin Gold (enzyme/substrate ratio of 1:50 (wt/wt), 37 °C for 18 h). The trypsin was provided to each participating site by NIST and was from the same lot as that used to prepare samples for studies I and II. Proteolysis was

doi:10.1038/nbt.1546

stopped by lowering the pH to 2 with 1% formic acid and the resulting peptide mixtures were desalted off-line by using Waters Oasis HLB 1 cc, 30 mg solid phase extraction cartridges (**Supplementary Methods**). Eluted tryptic peptides were lyophilized to dryness and resuspended in 25 µl of aqueous solution containing 3% acetonitrile and 5% formic acid. A mixture of the labeled signature peptides was added to aliquots of each reconstituted plasma digestion solution to yield standards that contained 50 fmol/µl $^{13}C/^{15}N$-signature peptides and tryptic ^{12}C-peptides (derived from the digested added-in target proteins) that spanned a range of concentrations (500, 275, 151, 83, 46, 25, 8.6, 2.9, 1.0 fmol/µl).

Reversed phase nanoflow liquid chromatography (nanoLC). Peptide mixtures were separated by online reversed phase nano high-performance liquid chromatography using dual pumping systems equipped with autosamplers: specifically six nanoLC-2D and one nanoLC-1D Plus System from Eksigent Technologies and one model 1100 Nanosystem from Agilent Technologies. PicoFrit (New Objective) columns, 75 µm internal diameter (i.d.) × 120 mm long, 10 µm i.d. tip, were self-packed with ReproSil-Pur C18-AQ (3 µm particle size and 120 Å pore size). Separations were performed at mobile phase flow rates of either 200 nl/min (Agilent) or 300 nl/min (Eksigent) on the binary pump systems using 0.1% (vol/vol) formic acid in water (mobile phase A) and 90% (vol/vol) acetonitrile with 0.1% (vol/vol) formic acid (mobile phase B). One microliter injections of the peptide digestion mixtures were separated using a binary gradient of 3–20% B in 3 min, 20–60% B in 35 min, 60–90% B in 2 min and at 90% B for 4 min (**Supplementary Methods** and **Supplementary Table 1a**).

4000 QTRAP instruments. Seven 4000 QTRAP hybrid triple quadrupole/linear ion trap mass spectrometers (Applied Biosystems/MDS Analytical Technologies) located at different sites were used to acquire MRM data for studies I, II and IIIa–c. General instrument operating parameters for the 4000 QTRAP instruments were kept uniform across the seven sites (see SOP, **Supplementary Methods** and **Supplementary Table 1a**). Typically, these mass spectrometers were required to operate with ion spray voltages of 2,200 ± 200 V, curtain gas 20, nebulizer gas (GS1) 5 ± 2, and interface heater temperature (IHT) 150 °C. MRM transitions were optimized for maximum transmission efficiency and sensitivity for individual instruments by infusion of unlabeled signature peptides. Optimized declustering potential, collision energy and collision cell exit potential are reported in **Supplementary Tables 1b–e** for each MRM transition along with the corresponding instrument used at each site. A total of 66 MRM transitions (3 per peptide) were monitored during an individual sample analysis. Identical instrument parameters were used for each unlabeled/labeled peptide pair. Due to the complexity of the matrix and to achieve maximum specificity, MRM transitions were acquired at unit resolution in the first and third quadrupoles (Q1 and Q3). Dwell times of 10 ms were used for all transitions and cycle times were set to 0.99 s.

TSQ Quantum Ultra instrument. A TSQ Quantum Ultra triple quadrupole mass spectrometer (ThermoFisher Scientific) was used to acquire MRM data for studies I, II and IIIa–c. Instrument operating parameters, based on precursor ion charge states and *m/z* values, were optimized for all MRM transitions by direct infusion of each unlabeled signature peptide (**Supplementary Table 1b**). Representative instrument operating parameters for the TSQ Quantum Ultra included a spray voltage of 1,200 ± 200 V, a capillary offset voltage of 35 V, a skimmer offset voltage of −5 V and a capillary temperature of 210 °C. Tube lens voltages used for all unlabeled and labeled peptides, which were based on values generated during the automatic tuning and calibration process, were not individually optimized. A single scan event was used to monitor a total of 66 MRM transitions, 3 MRM transitions per peptide, using the following parameters: Q1 and Q3 unit resolution of 0.7 FWHM, Q2 gas pressure of 1.5 mTorr, scan width of 0.004 *m/z* and a scan time of 15 ms (**Supplementary Table 1a**).

MRM data acquisition. Mass spectrometric data were acquired for the three separate studies that used the sample sets described above. Additional samples were analyzed for quality control (QC) purposes. These included equimolar mixtures of ^{12}C- and $^{13}C/^{15}N$-signature peptides (no plasma background matrix). Study samples were analyzed in a specified order, from lowest to highest concentration of added-in peptides, with four technical replicates for each sample as described in detail in the accompanying SOP (**Supplementary Methods**). A total of 57 LC-MRM runs were recorded per study.

Data analysis platforms. Instrument-specific data analysis software was employed for quantitative analyses: MultiQuant (Applied Biosystems/MDS Analytical Technologies) was used to process 4000 QTRAP data and SRM Workflow (prototype, ThermoFisher Scientific) was used to process TSQ Quantum Ultra data. Briefly, the MRM transitions for each peptide were individually integrated to generate ion current peak areas representing each of the ^{12}C and $^{13}C/^{15}N$ peptide fragment ion signals. A peak area ratio characteristic for each MRM transition was calculated by dividing the ^{12}C peak area by that of its corresponding $^{13}C/^{15}N$ counterpart. All data analysis and peak area integrations were initially performed at each of the eight sites, however, for the seven 4000 QTRAP instruments data sets were further 're-integrated' at one central site to guarantee uniform data analysis and uniform determination of outlier peaks. MultiQuant data were directly exported as a text file for further statistical analyses. SRM Workflow data were exported as .csv files and reformatted using an in-house Perl script, which was written to enable cross-site comparisons. Software versions and other details for data analysis for each site are listed in **Supplementary Table 1a**.

Statistical methods. For all statistical calculations, final concentrations of heavy and light peptides and added proteins were adjusted according to the gravimetric measurements described in **Supplementary Table 6a–f**.

Graphical methods. Data from MRM experiments were exported from MultiQuant (MQ) or SRM Workflow and imported into the R statistical computing environment (http://www.R-project.org/) for graphical review and statistical analyses[29]. Comprehensive plots were made of all experiments (studies I, II, IIIa–c) for all peptides (ten) and sites (eight) with estimated concentration on the vertical axis and theoretical concentrations on the horizontal axis. Estimated concentrations were based upon the following equation: [calculated concentration (fmol/µl) = peak area ratio of analyte to internal standard × 50 fmol/µl of internal standard. Plots were made on the linear scale (**Supplementary Fig. 5**), with additional plots used for data visualization shown in **Supplementary Fig. 6** and **Supplementary Appendix**. Plots were sent to the sites to identify and adjust integration errors by inspecting integrations in MQ or SRM Workflow from visually identified outliers in the plots.

Statistical models for linear calibration curves. When added into solution (buffer or plasma) the observed concentration y of a signature peptide should be identical to x the concentration at time of addition. However, full recovery of the peptide by the assay does not always occur, and percent recovery r is often $<100\%$. Thus, the relation between the observed and the expected concentration is $y = rx$. Variation between replicates as measured by the s.d. usually increases in proportion to the concentration x, although CV = s.d./mean generally decreases across the range of concentrations. Thus, the statistical model for the linear regression lines is:

$$E(y) = r \times x + c, \text{s.d.}(y) = k \times x$$

where E(y) is the average observed concentration (based on the 12 observations arising from the three transitions and four replicates) at the concentration x at time of addition. s.d.(y) is the standard deviation of the observations at x, and increases proportionally with x, k being the proportionality constant. The slope of the line is the percent recovery r and the line has an expected intercept c of 0. A statistically significant nonzero intercept can be interpreted as the endogenous level of the peptide existing in the solution with no added-in signature peptide. The s.d. increasing with the spiked-in concentration x requires the linear regression to be weighted proportionally to the inverse of the variance ((s.d.)^2), and so the weight is $1/x^2$ (refs. 30,31).

Robust linear regression. In some cases, data points were observed that did not fall near the linear trend line. These points, assigned as outliers, had plausible explanations in the majority of cases, including interference in the

heavy or light channels, saturation, presence of endogenous protein or were unexplained. A robust linear regression method was applied to down-weight the influence of outliers based on the estimation of the parameters in a linear regression: slope, intercept and s.d. about the regression line. In addition, this model also reports the standard error for the slope and the intercept so that 95% confidence intervals can be calculated as the estimate ± 1.96 × standard error (**Supplementary Table 4**). Data points from all three MRM transitions and the four replicate injections were included in estimating the linear trend. The robust linear model function in R was applied[32,33] using Tukey's biweight function to minimize the influence of outliers. For example, where there was an interference in one transition, the robust method was little influenced by the aberrant transition, and the parameter estimates resulted mainly from the data contributed by the other two transitions, which were usually coincident on the linear trend. This approach allowed all data points to be included in the estimation process and did not require subjective elimination of outliers.

Assessment metrics for quantitative MRM assays. The metrics used for assessing reproducibility of the MRM assays for the seven target proteins (ten peptides) were: (i) intralaboratory precision, represented by the median CV calculated from all concentration points for a particular peptide (based on quadruplicate measurements for a single transition used to calculate LOQ/LOD, see below) for each site, and for each study, and (ii) interlaboratory precision, represented by the median CV calculated at each concentration point for a particular peptide (based on quadruplicate measurements for a single transition used to calculate LOQ/LOD, see below) across all sites and for each study.

The CV is calculated as the ratio of the s.d. to the mean of a set of measurements. The CV calculations at each concentration point for a peptide at a given laboratory is based on four replicates for studies I and II and on 12 data points (four technical replicates for each of the three process replicates) for study III.

Determining LOD and LOQ. The following methods can be used to calculate the lower LOD for an analyte (defined as the concentration level at which the analyte can be reliably detected in the sample under consideration) and the lower LOQ, defined as the level at which the analyte can be detected and measured with sufficient precision. Methods range from straightforward modeling of blank sample variance using normal distributions[34], modeling variance as a function of concentration[35] fitting the relative s.d. along the concentration curve[36] and empirical methods[37]. For this study, a simple method was chosen for calculating LOD[34]. Once the LOD was determined separately for each transition of each peptide, the LOQ was calculated using the customary relation: LOQ = 3 × LOD[32]. The LOD was based on the variance of the blank sample (sample A1, with no analyte added in) and the variance of the lowest level added-in sample (sample B, with analyte at 1 fmol/µl). Assuming a type I error rate $\alpha = 0.05$ for deciding that the analyte is present when it is not, and a type II error rate $\beta = 0.05$ for not detecting the analyte when it is present, the LOD was derived as:

$$\mathrm{LOD} = \mathrm{LOB} + c_\beta \times \mathrm{s.d._S}$$

LOB (limit of blank) was defined as the 95th percentile of the blank A1 samples[38]. This was estimated as the mean plus $t_{1-\beta} \times \mathrm{s.d._b}$, where $\mathrm{s.d._b}$ was the standard deviation of the blank samples, and $\mathrm{s.d._S}$ was the standard deviation of the lowest analyte concentration point, sample B. For a relatively small number of repeated measurements for sample B, c_β was approximated as $t_{1-\beta}$, where $t_{1-\beta}$ is the $(1-\beta)$ percentile of the standard t distribution on f degrees of freedom. It is important to relate the LOD calculations to the measurement process. In this study, the final result of measuring the sample is obtained from the four replications as measured by the best transition. The LOD calculation when four values are averaged to obtain the final measurement requires the s.d. estimates to be halved, so the LOD equation becomes:

$$\mathrm{LOD} = \mathrm{mean_b} + t_{1-\beta} \times (\mathrm{s.d._b} + \mathrm{s.d._S})/2.$$

LOD values were initially calculated for all three transitions monitored for each peptide. The transition with the smallest root mean square deviation from the minimum LODs for both studies I and II was chosen as the best transition. This transition is used to report LOD and LOQ for both studies I and II, and for interlaboratory and intralaboratory CV calculations for all studies.

28. Burkitt, W. *et al.* Toward Système International d'Unité-traceable protein quantification: from amino acids to proteins. *Anal. Biochem.* **376**, 242–251 (2008).
29. R Development Core Team. *R: A language and environment for statistical computing* (R Foundation for Statistical Computing, Vienna, Austria, 2008).
30. Lavagnini, I. & Magno, F. A statistical overview on univariate calibration, inverse regression, and detection limits: Application to gas chromatography/mass spectrometry technique. *Mass Spectrom. Rev.* **26**, 1–18 (2007).
31. Xu, X., Keefer, L.K., Ziegler, R.G. & Veenstra, T.D. A liquid chromatography-mass spectrometry method for the quantitative analysis of urinary endogenous estrogen metabolites. *Nat. Protocols* **2**, 1350–1355 (2007).
32. Marazzi, A.. *Algorithms, Routines and S Functions for Robust Statistics* (CRC Press, Boca Raton, Florida, USA, 1993).
33. Venables, W.N. & Ripley, B.D. *Modern Applied Statistics with S* edn. 4 (Springer, New York, 2002).
34. Currie, L.A. Limits for qualitative detection and quantitative determination. *Anal. Chem.* **40**, 586–593 (1968).
35. Zorn, M.E., Gibbons, R.D. & Sonzogni, W.C. Weighted least-squares approach to calculating limits of detection and quantification by modeling variability as a function of concentration. *Anal. Chem.* **69**, 3069–3075 (1997).
36. Vial, J., Mapihan, K.L. & Jardy, A. What is the best means of estimating the detection and quantification limits of a chromatographic method? *Chromatographia* **57** Supplement, S303–S306 (2003).
37. Armbruster, D.A., Tillman, M.D. & Hubbs, L.M. Limit of detection (LOD)/limit of quantitation (LOQ): Comparison of the empirical and the statistical methods exemplified with GC-MS assays of abused Drugs. *Clin. Chem.* **40**, 1233–1238 (1994).
38. Linnet, K. & Kondratovich, M. Partly nonparametric approach for determining the limit of detection. *Clin. Chem.* **50**, 732–740 (2004).

doi:10.1038/nbt.1546