Journal of proteome research

Article

pubs.acs.org/jpr

Comprehensive Quantitative Analysis of Ovarian and Breast Cancer Tumor Peptidomes

Zhe Xu,[†,¶] Chaochao Wu,[†,¶] Fang Xie,[†] Gordon W. Slysz,[†] Nikola Tolic,[‡] Matthew E. Monroe,[†] Vladislav A. Petyuk,[†] Samuel H. Payne,[†] Grant M. Fujimoto,[†] Ronald J. Moore,[†] Thomas L. Fillmore,[‡] Athena A. Schepmoes,[†] Douglas A. Levine,[§] R. Reid Townsend,[∥] Sherri R. Davies,[∥] Shunqiang Li,[∥] Matthew Ellis,[∥] Emily Boja,[⊥] Robert Rivers,[⊥] Henry Rodriguez,[⊥] Karin D. Rodland,[†] Tao Liu,*[,†] and Richard D. Smith*[,†]

[†]Biological Sciences Division and [‡]Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, Washington 99354, United States

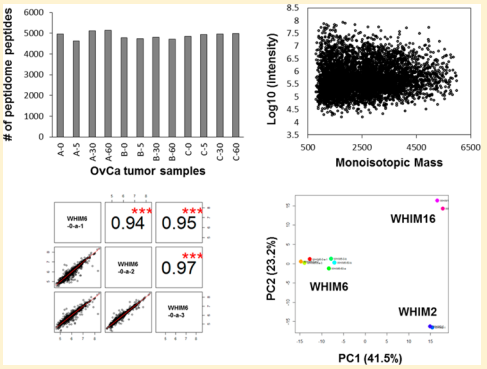
[§]Gynecology Service/Department of Surgery, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, United States

[∥]Department of Medicine, Washington University, St. Louis, Missouri 63130, United States

[⊥]Office of Cancer Clinical Proteomics Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, United States

Supporting Information

ABSTRACT: Aberrant degradation of proteins is associated with many pathological states, including cancers. Mass spectrometric analysis of tumor peptidomes, the intracellular and intercellular products of protein degradation, has the potential to provide biological insights on proteolytic processing in cancer. However, attempts to use the information on these smaller protein degradation products from tumors for biomarker discovery and cancer biology studies have been fairly limited to date, largely due to the lack of effective approaches for robust peptidomics identification and quantification and the prevalence of confounding factors and biases associated with sample handling and processing. Herein, we have developed an effective and robust analytical platform for comprehensive analyses of tissue peptidomes, which is suitable for high-throughput quantitative studies. The reproducibility and coverage of the platform, as well as the suitability of clinical ovarian tumor and patient-derived breast tumor xenograft samples with postexcision delay of up to 60 min before freezing for peptidomics analysis, have been demonstrated. Moreover, our data also show that the peptidomics profiles can effectively separate breast cancer subtypes, reflecting tumor-associated protease activities. Peptidomics complements results obtainable from conventional bottom-up proteomics and provides insights not readily obtainable from such approaches.



KEYWORDS: *protein degradation, peptidomics, proteases, tumor, ovarian cancer, breast cancer, ischemia*

INTRODUCTION

Proteolysis is tightly linked with cancer pathogenesis. For example, proteasomal activity in malignant tumors is significantly greater in breast cancer (BrCa) and endometrial cancer tissues compared to normal tissues;[1] the activity of extracellular proteases such as matrix metallopeptidase families is considerably elevated in pancreatic cancer, giving rise to much higher abundance of collagen fragments compared to healthy controls.[2] Therefore, studying the peptidome, that is, peptides derived from endogenous proteolysis events, can provide important insights into the activity of various endogenous peptidases under clinically relevant conditions, as well as potential information on biologically active peptides.[3,4]

Advanced liquid chromatography-tandem mass spectrometry (LC-MS/MS)-based peptidomics provides a powerful tool for identifying relatively large sets of endogenous peptides in a given biological sample, such as cell lines[5,6] and bodily fluids (e.g., blood serum/plasma[3,7−9] and urine[10,11]). Without converting proteins into peptides using trypsin or other proteases as applied in conventional bottom-up proteomic sample preparation, the natural forms of the peptidome peptides can be retained in the samples, including post-translational modifications and degradation products revealing the substrate cleavage specificities of their natural proteases. While a large number of peptidomics studies in mammals and a variety of invertebrates were focused on the identification of neuropeptides and peptide hormones,[12,13] many others have

Special Issue: Environmental Impact on Health

Received: August 11, 2014
Published: October 28, 2014

been carried out for discovering signature endogenous protein fragments in blood[3,7,8] or tissue[14−16] that are indicative of disease states such as cancer. More recent studies in cell lines have identified proteolytic products of intracellular and intercellular proteins that appear to be stable bioactive molecules with explicit roles in cellular signaling pathways, rather than simple transient protein degradation products.[5,17] Our recent study of blood plasma peptidome revealed significantly changed peptidomics profiles between samples from breast cancer patients and healthy controls but with very similar bottom-up proteomic profiles, showing the potential of peptidomics to reveal endogenous proteolytic processing information to which conventional bottom-up proteomics is effectively blind.[18] Thus, by globally assessing peptide concentrations and identifying relevant cleavage events, peptidomics could serve as a powerful tool in both characterizing the entire set of endogenous peptides and identifying novel physiologically active protease/substrates interactions in vivo for cancer studies.

The ability to identify a broader range of peptidome peptides in biological samples relies on not only advances in MS and related analytical tools (hence better sensitivity and reproducibility) but also improvements on other aspects of the analysis, such as sample preparation and data interpretation (e.g., informatics tools for analyzing the large peptidomics data sets[19]). Moreover, peptidome peptides may be processed differently in different cell types or even the same type of cells under different pathological conditions. Therefore, how to efficiently extract the full repertoire of peptides from complicated clinical tissue samples while maintaining the "original" status of the samples (e.g., avoiding potential confounding factors) and further detection of all peptides, given the limited dynamic range of measurement capabilities, are typical concerns for peptidomics. For example, post mortem stability is potentially problematic for peptidomics studies,[20] as are issues associated with the handling of clinical samples, including degradation associated with postexcision delay, making it difficult to characterize the "true" peptidome of clinical tumor samples. To the best of our knowledge, no detailed study assessing the effects of postexcision delay on tumor/tissue peptidomes has been performed to date. In addition, interpreting the significance of peptidome components identified from clinical tissue samples in a biological context is of great biomedical interest, specifically differentiating between specific regulated events that convey biological information and less specific general effects associated with protein degradation. Identifying the natural proteases and their associated pathways in clinical tissue samples and exploring the biological significance of the protease/substrate interaction has the potential to provide novel insights into the mechanisms of molecular tumor pathology.[21]

In the present study, we applied a quantitative peptidomics platform with in-house-developed data analysis tools to characterize the peptidomes in human ovarian cancer (OvCa) tumor and BrCa patient-derived xenograft (PDX) tumor samples. The ability of the platform to achieve highly reproducible measurements and comprehensive peptidomics coverage was demonstrated. In addition, analyses of tumor samples undergoing up to 60 min postexcision delay showed little or no effect of warm ischemia on peptidomes from both BrCa and OvCa tumors. Importantly, the peptidome profiles were found to effectively distinguish the different BrCa subtypes in the PDX tumor samples. In line with our previous BrCa cell line study and research from other groups, the proteasome was identified as the potential major contributor to human OvCa and BrCa tumor peptidomes. Taken together, our results demonstrate that robust quantitative peptidomics analysis can be effectively performed on clinical tumor samples using our peptidomics platform, which sets a stage for further determination of molecular details and functional significance of the peptidomic/degradomic activities in cancers.

MATERIALS AND METHODS

Tumor Samples and Postexcision Delay Experiments

After obtaining consent to IRB-approved protocols, tissue was collected from three patients with high-grade serous ovarian carcinoma. Each patient was under general anesthesia and had a large midline vertical incision that identified advanced disease (FIGO stage IIIC or IV). Prior to performing primary tumor resection and before any compromise to vascular supply, a portion of ovarian tumor attached to the omentum was rapidly resected using sharp or blunt dissection. The tumor specimen was immediately dissected into four contiguous and adjacent specimens strips, each no larger than 10 × 3 × 3 mm, and placed into cryovials and frozen in liquid nitrogen at four different time points (0, 5, 30, and 60 min, at room temperature). The first specimen (0 min) was processed as quickly as possible, with an elapsed time from resection to freezing of 1 min or less. All specimens were then stored in −80 °C freezers until cryopulverization (described below).

The PDX BrCa tumors, also referred to as "Washington University Human in Mouse" (or WHIM lines) from established basal (WHIM2 and WHIM6) and luminal (WHIM16) breast cancer subtypes, were raised subcutaneously in 8 week old NOD.Cg-*Prkdc*^{scid}*Il2rg*^{tm1Wjl}/SzJ mice (Jackson Laboratories, Bar Harbor, ME) as previously described (sample names and corresponding cancer subtypes are listed in Table 1).[22,23] For studying potential ischemic effect on tumor peptidomic analysis, WHIM6 basal tumors from each animal were harvested by surgical excision at approximately 1.5 cm^3, rapidly divided into four pieces, and snap-frozen by immersion in a liquid nitrogen bath at times 0 (~30 s) (WHIM6-0-a and WHIM6-0-b) and 60 min (WHIM6-60-a and WHIM6-60-b) postexcision. Another basal PDX tumor sample (WHIM2) and a luminal PDX tumor sample (WHIM16) were collected and processed without delay. The snap-frozen tumor tissues for individual time points were then placed in precooled tubes on dry ice and stored at −80 °C until cryopulverization.

Table 1. Summary of the PDX Breast Tumor Samples Analyzed in the Present Study

sample name	cancer subtype	WHIM line	delay
WHIM6-0-a	basal BrCa	WHIM6	0 min
WHIM6-60-a	basal BrCa	WHIM6	60 min
WHIM6-0-b	basal BrCa	WHIM6	0 min
WHIM6-60-b	basal BrCa	WHIM6	60 min
WHIM2	basal BrCa	WHIM2	N/A
WHIM16	luminal BrCa	WHIM16	N/A

The ovarian cancer tumor samples and breast cancer PDX tumor samples were further processed by cryopulverization. Briefly, tumor pieces were transferred into precooled Covaris Tissue-Tube 1 Extra (TT01xt) bags and processed in a Covaris CP02 Cryoprep device (Covaris, Woburn, MA) using different

dx.doi.org/10.1021/pr500840w | J. Proteome Res. 2015, 14, 422−433

impact settings according to the total tumor tissue weight: <250 mg = 3; 250−350 mg = 4; 350−440 mg = 5; 440−550 mg = 6. Tissue powder was transferred to an aluminum weighing dish (VWR, Radnor, PA), and the tissue was thoroughly mixed with a metal spatula precooled in liquid nitrogen. The tissue powder was then partitioned (~100 mg aliquots) into precooled Corning cryovials (Sigma-Aldrich, St. Louis, MO). All procedures were carried out on dry ice to maintain tissue in a powdered, frozen state.

Peptidomics Sample Preparation

For peptide extraction, extraction buffer containing 0.25% acetic acid and protease inhibitor cocktail (Sigma-Aldrich, St. Louis, MO) was added into approximately 20 mg (wet weight) of pulverized ovarian cancer tumor and breast cancer PDX samples at a ratio of 1:10 (w/v). Mixed samples were homogenized on ice for 1 min and then sonicated with three short bursts of 10 s followed by intervals of 5 min for cooling on the ice bed. After that, samples were centrifuged at 4 °C, 14 000*g*, for 30 min, and the supernatant was filtered in an Amicon Ultracel 30 kDa MWCO filter tube (Millipore, Billerica, MA) by centrifugation (4 °C, 8000*g*) to remove peptides larger than 30 kDa. The flow-through was concentrated in Speed-Vac (Thermo Fisher Scientific, Waltham, MA), and final yield of peptidome peptides was calculated by BCA protein assay kit (Pierce, Rockford, IL).

LC-MS/MS Analysis

All peptidomics samples were analyzed using nanoACQUITY UPLC system (Waters Corporation, Milford, MA) coupled online to a LTQ Orbitrap Velos mass spectrometer (Thermo Fisher Scientific, Waltham, MA). A 70 cm × 75 μm i.d. (flow rate 300 nL/min) and a 110 cm × 75 μm i.d. (flow rate 200 nL/min) fused-silica capillary column packed with 3 μm Jupiter C18-bonded particles (Phenomenex, Torrance, CA) was used for analysis of the PDX breast tumor and ovarian tumor samples, respectively. Mobile phases consisting of 0.1% formic acid in water (A) and 0.1% formic acid acetonitrile (B) were operated with effective gradient profiles as follows (min:%B): 0:1, 6:8, 60:12, 225:35, 291:45, 300:95 (110 cm × 75 μm i.d. column; flow rate 200 nL/min) and 0:1, 6:8, 36:12, 135:35, 175:45, 205:95 (70 cm × 75 μm i.d. column; flow rate 300 nL/min). The LTQ Orbitrap Velos mass spectrometer was operated in the data-dependent mode acquiring high-resolution CID scans (R = 15 000, 5 × 10^4 target ions) after each full MS scan (R = 60 000, 1 × 10^6 target ions) for the top six most abundant ions within the mass range of 400 to 2000 m/z. An isolation window of 2 Th and a normalized collision energy of 35 were used for CID. The dynamic exclusion time was 60 s.

Data Analysis

SEQUEST[24] with MS-GF rescoring,[25] MS-Align+,[26] and the Unique Sequence Tags (UStags)[27] algorithms were used to independently search all the MS/MS spectra against either the human NCBI database (Build37.3) for identification of the peptidome peptides in the ovarian tumor samples or a composite human−mouse NCBI database for the PDX tumor samples. For SEQUEST (no enzyme digestion, mass tolerance for parent ions is 5 Da, and 0.05 Da for fragment ions) and MS-Align+ methods (15 ppm for parent ions), target decoy database searching strategy was adopted for FDR estimation; UStags result was subjected to manual validation. MS/MS search results were then filtered using the following criteria: SEQUEST with MSGF SpecProb <1 × 10^{-9} (data were filtered so that each spectrum had at most one peptide identification event, and overall peptide level FDR was less than 1%), MS-Align+ with a p value <1 × 10^{-6} (less than 1% FDR), and further combined to create peptide databases (one for the ovarian tumor data and one for the PDX breast tumor data) for direct analysis of the LC-MS data in each individual analysis using the recently developed Informed Quantitation (IQ) approach (described below).

IQ first enumerates all confidently observed peptides. A precise empirical formula is used to generate the theoretical isotopic profile using a standard binomial expansion algorithm. For a given charge state for each peptide, IQ generates an extracted ion chromatogram (XIC) using a tight mass tolerance (e.g., 5 ppm) that is based on the most abundant peak of the theoretical isotopic profile. To determine which chromatographic peak represents the true LC-MS feature, IQ tests all detected chromatographic peaks that fall within an elution time window. Underlying MS1 mass spectra for each chromatographic peak are summed and mass spectral peaks detected using three-point peak detection, as described previously.[28] The m/z values of the theoretical isotopic profile are used to guide the extraction of the observed isotopic profile from the summed mass spectra. Least-squares fitting of the theoretical isotopic profile on the observed profile is then performed,[28] providing a measure of how well the observed isotopic profile matches the theoretical isotopic profile. This metric is called the "fit score" and is a key metric for resolving correct versus incorrect features. It is important to model the behavior of the score metric on false (random) data (the computational details are being prepared for a separate publication); therefore, we have modeled the score distribution of false hits and used that to guide the filtering. The fit scores used for each tested chromatographic peak are assembled, and IQ selects the "best" chromatographic peak as follows. The chromatographic peak candidates are filtered to remove all null results (when no isotopic profile was found). If there is only a single chromatographic peak candidate, it is selected. If the top two candidates have very similar fit scores (within 0.05), the most abundant LC-MS peak is selected. Otherwise, the LC-MS feature with the lowest (i.e., best) fit score is selected.

After a chromatographic peak is selected for a given peptide/charge state target, IQ then calculates the final abundance measurement. This comprises summing a total of five mass spectra, centered around the apex scan of the elution profile. A key step in IQ is the alignment of observed mass and elution times to database values in order to correct for variations in mass and elution time measurements taken across multiple data sets. Alignment of mass and the LC elution time make it possible to narrow the mass tolerance used in generating XICs and the elution time window for selecting the correct chromatographic peak. Currently, VIPER[29] is used in a first-pass analysis (using conventional accurate mass and time, or AMT tag approach) to output mass and normalized elution time (NET) alignment information, which is then loaded into IQ and used for mass and NET correction during subsequent processing. Data processed by the IQ approach was initially filtered by fit score (<0.2), NET tolerance (<0.025), mass accuracy (± 2.5 ppm), and followed by a manual validation using an in-house tool SIPPER[30] to eliminate false positives. The identified peptides were grouped to a nonredundant protein list using IDpicker3.[31]

The ProteinCoverageSummarizer tool developed in-house and corresponding protein databases were used to map the

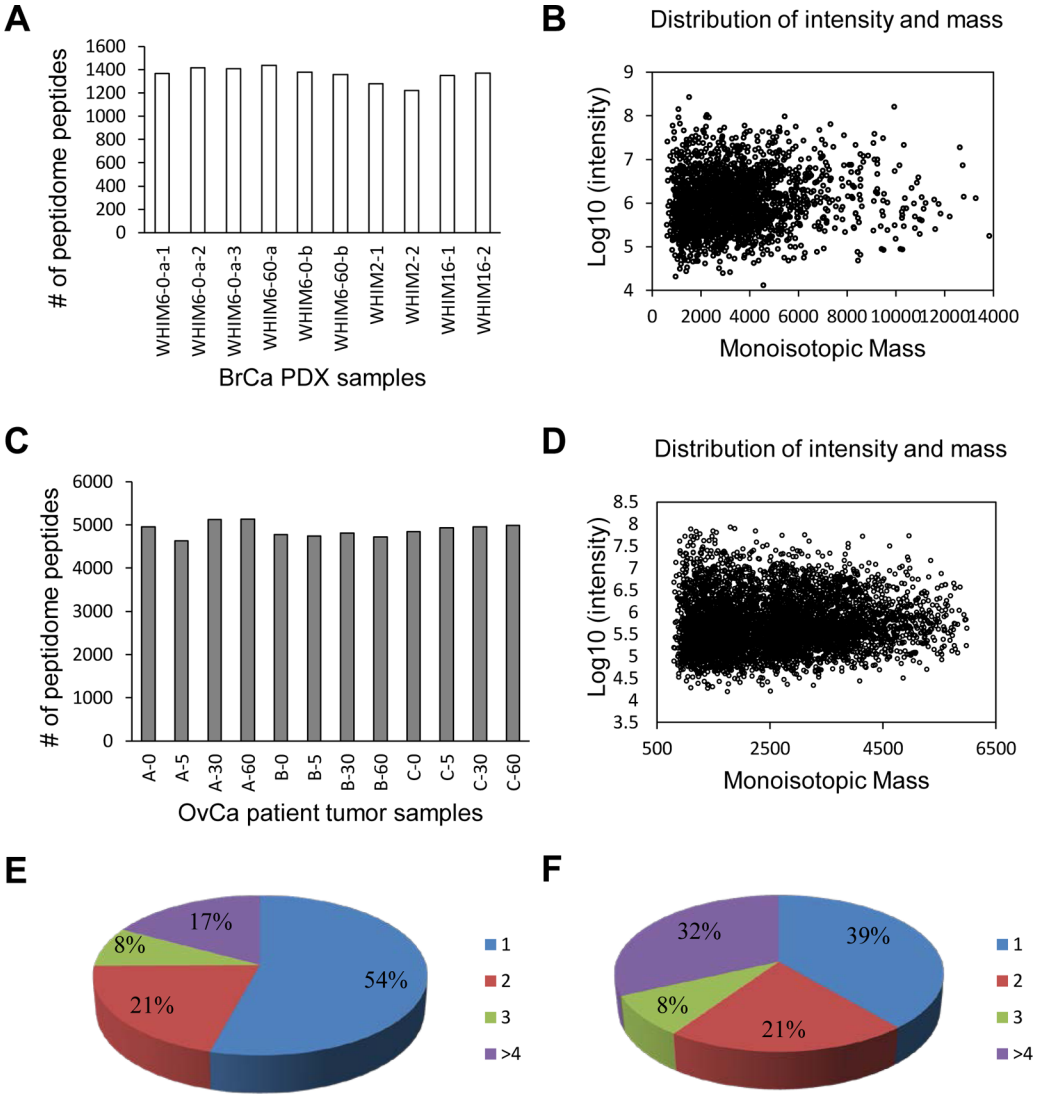


Figure 1. Summary of peptide identifications from OvCa tumor and PDX breast tumor samples. Bar graphs show the numbers of identified peptides from each PDX breast tumor (A) and OvCa tumor (C). The intensities of identified peptides ($\log_{10}$ values) from PDX (B) and OvCa (D) tumor samples are plotted against their monoisotopic masses (Da), showing a similar peptide intensity range but different mass ranges. The distributions of the precursor proteins in terms of number of unique peptidome peptide identifications are shown in pie graphs for the PDX samples (E) and OvCa samples (F).

unique peptide sequences into proteins. The intensities of peptides were first $\log_{10}$ transformed and then median normalized across all samples (most bioinformatics analysis was performed using R program unless specifically indicated). Unsupervised hierarchical clustering analysis was applied to peptide data sets using Euclidean distance as distance metrics and ward linkage for clustering. Only peptides detected in all samples were subjected to the clustering analysis. Datasheets containing normalized peptide intensities were imported into DanteR program[32] for principal component analysis (PCA). The same datasheets were also imported into R for Volcano plot and other statistical tests (two-factor ANOVA, *t* test), with an adjusted *p* value of 0.05 for significance. Venn Diagram Plotter (http://omics.pnl.gov/software/venn-diagram-plotter) was used for Venn diagrams. The frequency of various amino acids within the cleavage site (P6, P5, P4, P3, P2, P1, P1′, P2′, P3′, P4′, P5′, and P6′) from all identified peptides was analyzed.[33] For each peptide, N- and C-terminal cleavage sites were analyzed as PX_Up and PX_Down, respectively, the frequencies of which were later plotted together for individual P sites. In addition, amino acid frequencies at all P sites were adjusted to the relative abundance of each amino acid within all human proteins. Frequencies of amino acids at all 12 P sites were analyzed to identify potential proteases based on the protease cleavage rules which were adopted from the comprehensive publication by Keil et al.[38] Gene List Analysis was conducted using Panther (http://www.pantherdb.org/) to obtain biological information (e.g., function, pathway) of the precursor proteins.

In order to evaluate changes in the peptidome over time, a kinetics-based regression model was used. This model assumes that changes observed within a reasonably short period of time (1 h at most) following perturbation (cold ischemia) can be described by the law of unidirectional chemical reaction (eq 1).

$$\log(Y) = \log(B + (A - B)e^{-kt}) + \varepsilon \quad (1)$$

where Y is the measured relative protein concentration or abundance, A is the starting concentration, B is the final

dx.doi.org/10.1021/pr500840w | J. Proteome Res. 2015, 14, 422−433

concentration, k is the rate constant, t is time, and ε is the normally distributed measurement error. To infer the A, B, and k parameters, we applied constrained Nelder–Mead optimization that minimizes the sum of squared errors. The imposed constrains are along the lines of chemical kinetic principles: $A > 0$, $B > 0$ as the concentrations cannot be negative; $k > 0$ since no species can be exponentially growing. The statistical significance of the chemical kinetic model was tested against a null hypothesis model that assumes that there is not any change in the data associated with time points (eq 2).

$$\log(Y) = \log(C) + \varepsilon \quad (2)$$

where C is the constant protein abundance or concentration value. The fit of the null hypothesis model is represented by a flat line. The statistical significance of the alternative model can be assessed using F-statistic that looks at the ratio of the sums of squared error residuals taking into account the degrees of freedom of the alternative and null hypothesis models.

Datasheets containing peptides and their intensities from OvCa tumor time point samples, with $\log_2$ ratios normalized as specified above, were used for the statistical analysis. Each cold ischemia time point was treated as an independent group in an ANOVA analysis, with statistical significance being assessed using an F-test based on the ratio of between group to within group variability. This allowed capturing trends that are not unidirectional. To leverage information across observed peptides for more robust estimation of variance, we employed an empirical Bayes approach implemented as a moderated F-test in the Bioconductor "limma" package.[34] The resulting p values were adjusted for multiple testing using the Benjamini–Hochberg method.

RESULTS

Peptidomics Analysis of PDX Breast Tumor Samples

An average of 1358 ± 64 (CV = 4.7%) distinct peptides were identified by IQ from the analyses of each of the 10 PDX breast tumor samples (see Materials and Methods, Table 1, and Figure 1A); overall, a total of 2026 distinct peptides with molecular weight ranging from 600 to 14 000 Da (Figure 1B) were identified from all PDX tumor samples (the distribution of $\log_{10}$ transformed intensity of one sample is shown in Supporting Information Figure S3A). These peptides represented naturally occurring fragments of 824 different precursor proteins (peptidome peptides, their precursor proteins, and original peptide intensities from each PDX breast tumor sample were provided in Supporting Information Table S1), and the distribution of the numbers of peptides derived from the precursor proteins is shown in Figure 1E. These results demonstrate the good, consistent peptidome coverage of the analytical platform (Figure 1A).

The reproducibility of the peptidomics platform was evaluated through analysis of three full process replicates of the WHIM6 sample at the 0 min time point. The base peak chromatograms of the three WHIM600-a process replicates were well-aligned (Supporting Information Figure S1A); a considerable overlap (70.3%) of the identified peptides was also observed for the three WHIM6-0-a replicates (Supporting Information Figure S1B); furthermore, the peptide abundances in the three WHIM6-0-a replicates were well-correlated to each other (R^2 value 0.94–0.97; Supporting Information Figure S1C). These results demonstrated the good reproducibility of the integrated peptidomics platform, including sample preparation, LC-MS/MS instrument analysis, and IQ analysis.

We next explored whether peptidomics tumor sample analysis is capable of differentiating different BrCa subtypes. Two PDX tumor samples, WHIM2 and WHIM16, representing the basal and luminal BrCa subtypes, respectively, were analyzed in duplicate. The precursor proteins from the two WHIM2 and two WHIM16 analyses exhibited a considerable overlap (70.0%; Supporting Information Figure S2A). In the heatmap from unsupervised hierarchical clustering analysis, the two process replicates showed very similar relative peptide abundances; however, the WHIM2 samples showed substantially different peptidomics profiles from the WHIM16 samples (Figure 2A). Moderated t test analysis further showed that, when comparing WHIM16 to WHIM2 samples, out of a total of 815 common peptides, 450 were differentially abundant (p value <0.05) (Figure 2B and Supporting Information Figure S2B). This suggested a significant change in the peptidome between the two BrCa subtypes. To investigate if the difference in the peptidomic profile corresponds to that at the precursor protein level, we compared the changes in abundance of the differential peptidome peptides (through their precursor proteins) to that of the differential proteins (p value <0.05) from our previous bottom-up study of the same PDX samples (iTRAQ-labeled, with two WHIM2 replicates and two WHIM16 replicates; unpublished results). We found that peptides derived from 140 precursor proteins have the same propensity in abundance change as that of their precursor proteins, while 161 precursor proteins displayed opposite abundance changes compared to their corresponding peptidome peptides (Supporting Information Table S1). In addition, peptidome peptides derived from 62 precursor proteins showed inconsistent abundance changes (i.e., both upregulated and downregulated peptides from the same precursor protein), suggesting potential complicated proteolytic processes on the precursor protein from which the peptidome peptides were derived (Supporting Information Table S1). Taken together, these results suggest that peptidome could potentially provide insights on the BrCa subtypes beyond the protein expression level analysis.

Pearson correlation analysis of all the PDX breast tumor data showed that there was fairly good correlation within each WHIM line, again demonstrating the good reproducibility of the platform; however, the correlation became much less between different WHIM lines (R^2 value 0.50–0.66; Figure 3A). Notably, all the WHIM6 samples showed relatively good correlation to each other, independent of the postexcision delay of 60 min (R^2 value 0.74–0.96, Figure 3A). Unsupervised hierarchical clustering showed similar trends (Supporting Information Figure S2C). PCA analysis showed clear, complete separation of all three WHIM lines and again confirmed that the WHIM6 samples cannot be separated by any potential effects attributable to different time of warm ischemia (Figure 3B).

Peptidomics Analysis of Ovarian Cancer Tumor Samples

From analyses of all 12 ovarian cancer tumor tissue samples (from three patients, each with 4 time points), we identified a total of 5756 distinct peptides ranging from 500 to 6000 Da (Supporting Information Table S2 and Figure 1D); the number of identified peptides was highly consistent across samples (4952 ± 285; CV = 5.8%). These peptides were mapped into 974 distinct precursor proteins (peptidome peptides, their

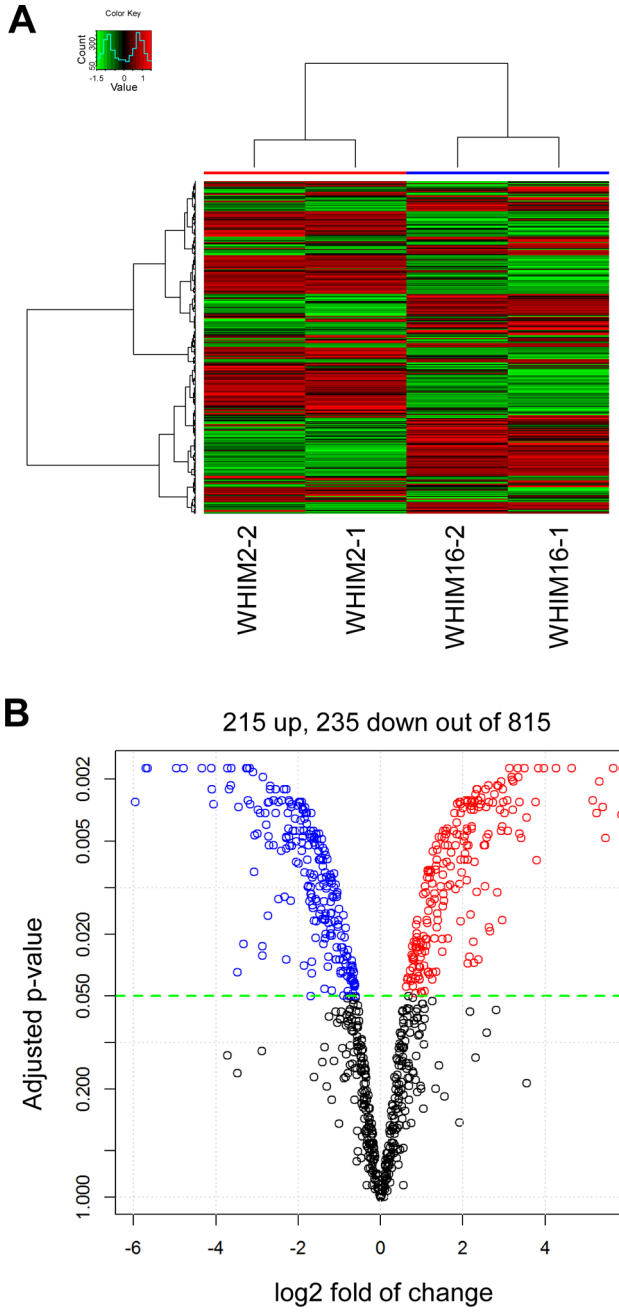


Figure 2. Drastic differences in peptidomics profiles of the PDX breast tumor WHIM2 and WHIM16 samples. (A) Unsupervised hierarchical clustering analysis of all peptides from WHIM2 and WHIM16 samples. (B) Volcano plot showing peptide abundance ratio (WHIM16/WHIM2) versus the adjusted p values (from the moderated t test) for 815 peptides that were detected in all four analyses. Blue (235 downregulated) and red (215 upregulated) hollow dots represent the peptides with p values <0.05; black hollow dots, not significantly changing peptides.

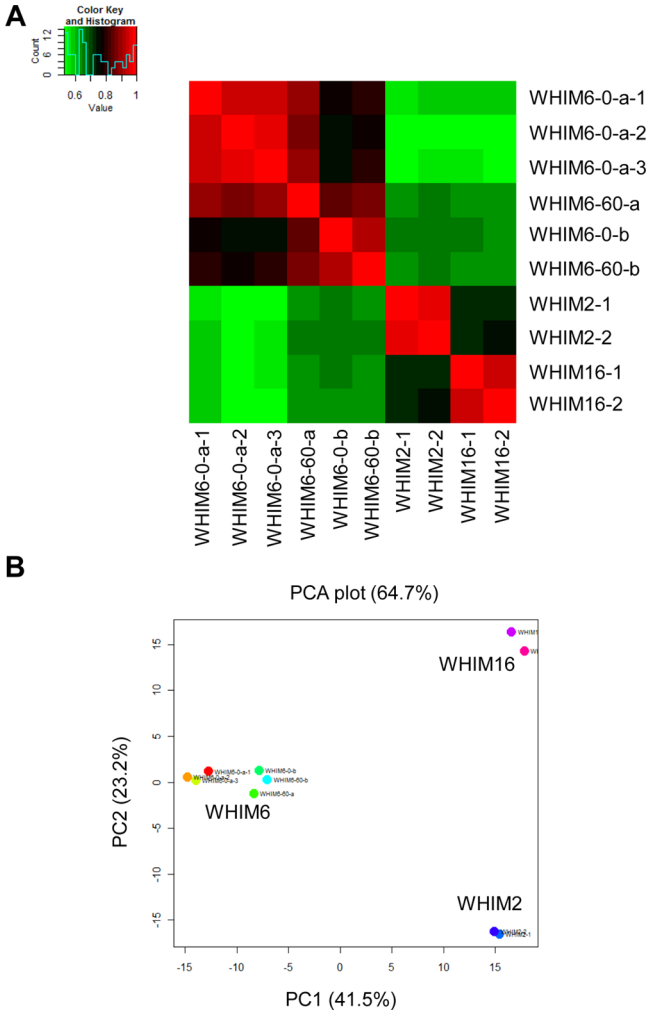


Figure 3. Peptidomics profiles clearly distinguish the different PDX breast tumor WHIM lines. Pearson correlation (A) and PCA analysis (B) of peptide intensities in all 10 PDX breast tumor samples.

precursor proteins, and original peptide intensities from each ovarian tumor sample are provided in Supporting Information Table S2), and the distribution of the numbers of peptides derived from the precursor proteins is shown in Figure 1F. Interestingly, although the same sample preparation protocols and analysis pipeline were applied for both the PDX and OvCa tumor peptidome analysis, for unknown reasons, the PDX samples yielded much smaller peptidomes (contain both human and mouse sequences) than the OvCa samples (~30%; Figure 1A vs 1C); however, the difference in number of precursor proteins was much smaller for the BrCa and OvCa peptidomes: 824 and 974, respectively. Compared to the PDX breast tumor samples, OvCa tumor peptidomes have similar peptide intensity range but a smaller mass range (Figure 1B vs 1D; the distribution of $\log_{10}$ transformed intensity of one OvCa sample is shown in Supporting Information Figure S3B).

Pearson correlation analysis, unsupervised hierarchical clustering analysis, and principal component analysis was performed to assess the ischemic effect on the OvCa tumor peptidome. Despite considerable overlap (95.4%) of the precursor proteins in all three patients (Supporting Information Figure S4A), the pairwise peptide intensities of all three patients were poorly correlated (Figure 4A). In contrast, peptides of the four time point samples from the same patient showed much better correlation, suggesting that any potential ischemic effect on the peptidome is much smaller than the patient-to-patient difference (Figure 4A). The clustering analysis (Figure 4B) and PCA analysis (Supporting Information Figure S4B) showed similar trends: the four time point samples for each patient had more similar peptidome profiles and were clustered together; the peptide intensity patterns changed significantly across the three patients and thus were clearly

dx.doi.org/10.1021/pr500840w | J. Proteome Res. 2015, 14, 422−433

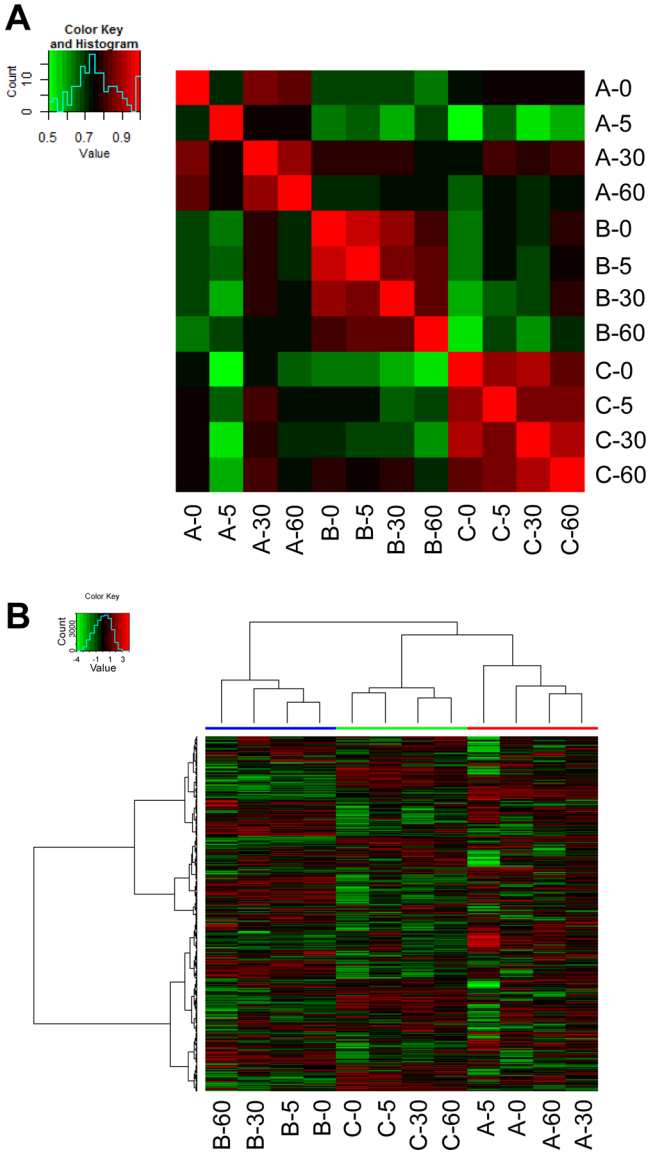


Figure 4. Analysis of the OvCa peptidome data demonstrated minimal ischemic effect within 60 min of postexcision delay. (A) Pearson correlation and (B) unsupervised hierarchical clustering analyses of all OvCa tumor samples (four time point samples from each of the three patients). It is evident that potential ischemic effect on tumor peptidome as a result of up to 60 min postexcision delay is much smaller than patient heterogeneity.

separated (grouped by patient) in PCA plots. These results indicate that variations in postexcision delay time provided a much smaller contribution to variations in peptidome peptide abundances than did patient heterogeneity.

A kinetics-based regression analysis was performed to further investigate whether there was any ischemic effect on the OvCa peptidome that is consistent across patients (the *p* values of the analysis are provided in Supporting Information Table S2). As shown by the volcano plot in Supporting Information Figure S4C, none of the peptides was significantly changing across the three patients as a result of 60 min ischemic stress. In addition, the two-factor ANOVA analysis showed a minimal contribution (<1%) of ischemic time to the significantly affected peptides, compared to other factors such as the patient heterogeneity (40%) (Supporting Information Figure S4D). Overall, results from all analyses on the OvCa tumor peptidome further indicated a negligible ischemic effect on the OvCa tumor peptidome within the first 60 min after excision, which is consistent with our PDX breast tumor peptidomics analysis. Previously, we have performed conventional bottom-up proteomics analysis for the same set of OvCa tumor time course samples, and the results showed that the overall protein abundance does not change significantly with up to 60 min postexcision delay,[35] which is also consistent with our tumor peptidome observations from the same samples.

IQ Analysis Improves Data Quality

The present peptidomics studies benefitted from the data analysis tool applied, IQ. As an example, we compared the IQ analysis results of the OvCa data to those obtained through conventional analyses such as MS-Align+ (MS/MS data) and the AMT tag strategy (LC-MS data). As expected, MS-Align+ suffers from the undersampling issue of typical data-dependent acquisition of MS/MS data and hence provided much lower peptidome coverage and less consistent results for each individual analysis (Supporting Information Figure S5A). Although, IQ and AMT tag analyses both use a database consisting of peptides identified from SEQUEST, MS-Align+, and UStags analysis of the entire peptidomics data sets (for each cancer type), IQ provided higher peptidome coverage and, more importantly, less missing data across the entire data sets (Supporting Information Figure S5B). This is because IQ uses all isotopic peaks as opposed to the AMT tag method[36,37] which uses individually deisotoped spectra, resulting in improved sensitivity, better distinguished overlapping features, more accurate quantification, and better reproducibility.

Proteolytic Cleavage Specificities in Ovarian and Breast Cancer Tumors

Peptidomics profiles are informative regarding cancer degradomic activities.[3] In order to investigate the degradomic activities of ovarian cancer and breast cancer tumors, precursor protein cleavage sites of all samples were examined for the frequency of signature amino acids. For peptides identified from OvCa samples, the most frequently observed amino acids at the P1 position were Leu, Lys, and Phe (Figure 5A). When normalized to the relative abundance of each amino acid within all of the human proteome, these amino acids were still the most common in the P1 position (Figure 5B). Thus, the preferred residues at the P1 position were basic and hydrophobic amino acids. In addition, predominant amino acids at the P1′ position of the cleavage sites were Leu, Val, Phe, Ala, and Ser, which were still among the most frequently seen amino acids after normalized to the amino acid abundance of the human proteome (Figure 5C,D). Chymotrypsin preferentially cleaves proteins at Phe, Tyr, and Trp with high specificity and at Leu, Met, and His with lower specificity.[38] Similar analyses of the residues were performed at other sites including P2, P3, P4, P5, P6, P2′, P3′, P4′, P5′, and P6′, showing the involvement of other proteases such as enterokinase and pepsin (Supporting Information Figure S6). Met also appears to be one of the most prominent residues at P1 and P1′ positions of the peptides after normalization (Figure 5B), but this can be attributed to N-terminal processing of the precursor proteins (discussed in the following paragraph). Collectively, the observed P1 amino acids indicate that chymotrypsin activity is most likely responsible for producing the peptides in the ovarian tumor peptidome. Previous studies indicated that the proteasome degradation pathway generates a large number of internal fragments, while

dx.doi.org/10.1021/pr500840w | J. Proteome Res. 2015, 14, 422−433

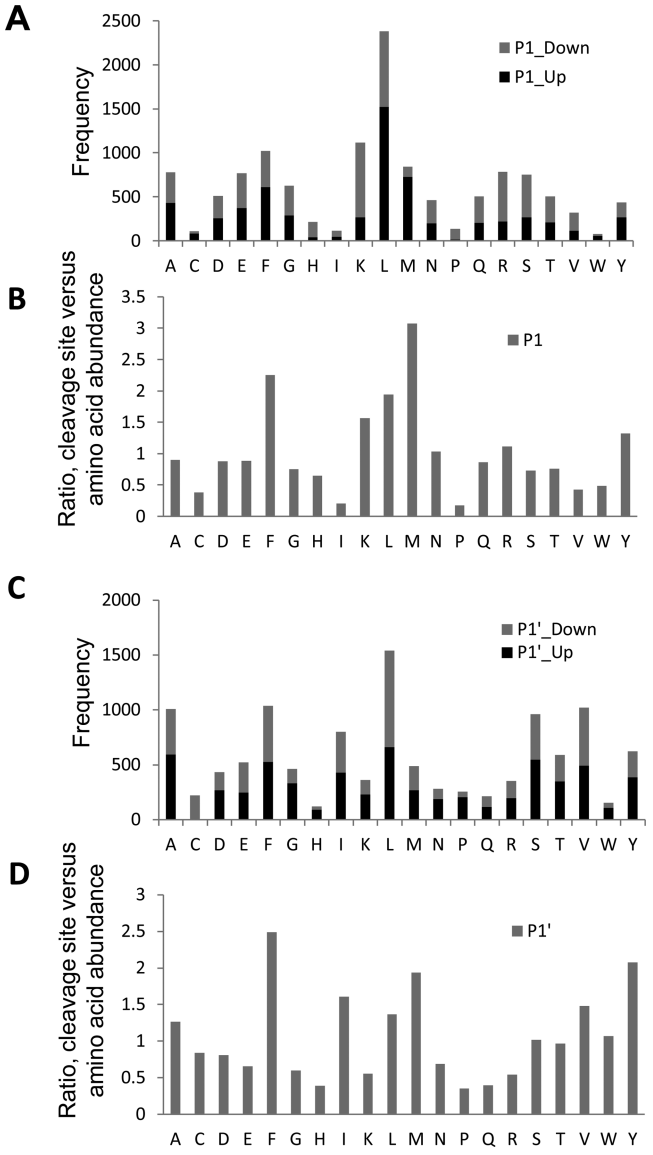


Figure 5. Cleavage specificity analysis at P1 and P1′ sites for OvCa peptidome peptides. (A) Number of peptides undergoing cleavages is shown on the *y*-axis for all P1 amino acids. P1_Down, downstream P1 site on the peptides; P1_Up, upstream P1 site. (B) P1 amino acid frequency in panel A was adjusted to the relative abundance of each amino acid in all human proteins. Ratio of 1.0 indicates that corresponding amino acid present in the cleavage site has the same frequency in overall human protein amino acid composition, while >1.0 suggests an amino acid found more frequently in the cleavage site than elsewhere in the peptides. (C) as in (A) except for the P1′ position. (D) as in (B), except for the P1′ position.

more terminal fragments are expected from the cytosolic proteolytic pathways.[39] Considering this, the percentages of terminal and internal peptides in OvCa peptidome were examined based on the sequence alignment of the peptides to their precursor proteins. The results showed that 82% of the 5756 peptides were internal peptides while 17% were terminal ones. In addition, 54% of the peptides were within the length range of proteasome products (3−23 amino acids).[40,41] Taken together, our data suggest that the chymotrypsin activity, which is possibly from the proteasome subunit, is the major, but not the only, contributor to the OvCa peptidome.

To investigate the other potential proteolytic mechanisms, we further examined the cleavage specificities of amino acid residues at the N-terminus and C-terminus of identified OvCa peptidome peptides. As shown by Supporting Information Figure S7A, Ala and Ser were the most frequently observed amino acids at the starting position of the peptides, which were derived from the N-terminus of precursor proteins with the first Met cleaved off. For the other N-terminal, all C-terminal, and overall peptidome peptides, no amino acid was observed with prominent frequency (Supporting Information Figure S7B−G). Based on the *MEROPS* database search (http://merops.sanger.ac.uk/index.shtml), an uncharacterized methionyl dipeptidase[42] (Met-Xaa dipeptidase) which most frequently cleaves Ala and Ser after Met matched the observed cleavage pattern, suggesting the involvement of N-terminal methionine excision (NME) in the generation of OvCa peptidome. In addition, ∼51% of the Met at P1 of all OvCa peptidome peptides was contributed by the NME process. Therefore, in addition to the contribution of proteasome subunit's low chymotrypsin specificity for Met, the NME process gives rise to the distinctive prevalence of Met at the P1 position of all OvCa peptidome peptides (Figure 5B).

Similar analyses were performed on the 2026 peptides identified from PDX breast tumor samples. The most frequently observed residues at the P1 position were Lys, Leu, Arg, and Phe, which were still among the most abundant amino acids when normalized to the amino acid abundance in the human proteome (data not shown), suggesting that trypsin and chymotrypsin activity were possible contributors to the peptides in the BrCa tumor peptidome. On the other hand, prevailing amino acids at the P1′ position were Met, Leu, Ala, Val, and Ser, including both polar and nonpolar amino acids. The cleavage specificities of other critical P and P′ sites were also analyzed (data not shown). In contrast to OvCa tumors, most of the peptides from PDX breast tumor were terminal peptides, 63%, and only 37% were internal ones. Taken together, these results suggest that cytosolic and extracellular proteolytic pathways appear to be the major contributor to the breast cancer peptidome rather than the proteasome-mediated proteolytic pathway.

However, our previous peptidomics studies of breast cancer cell lines suggested that chymotrypsin activity was the major contributor to the peptidome (unpublished results), which seemed to contradict the current data from the PDX tumor samples. To evaluate whether the xenograft system (i.e., "contamination" from mouse blood) contributed to the cleavage specificities of the BrCa tumor peptidome, we divided the peptides into human-only, mouse-only, and human−mouse-shared subgroups and performed similar cleavage specificity analyses for all groups. As shown in Supporting Information Figure S8, for human-only peptides, Lys, Leu, and Arg were the prevailing residues at the P1 site, while Ser, Val, Leu, Met, and Ala were mostly observed at the P1′ site. Residues at the other P and P′ sites excluded the possibility of proteases from the digestive system (Supporting Information Figure S9). Thus, the proteasome still appears to be the prominent contributor of the human-specific peptides, although possibly out of a combined contribution of the chymotrypsin (β5 subunit) and trypsin activities (β2 subunit).[43] For the other mouse and mixed mouse−human groups, possible cytosolic and extracellular proteases were contributing equally to proteasome subunits, which was also supported by the presence of about 50% internal peptides in the peptidome of these two

dx.doi.org/10.1021/pr500840w | J. Proteome Res. 2015, 14, 422−433

groups. This is also consistent with our previous peptidomics analysis of breast cancer patient blood plasma, where trypsin-like activity, instead of chymotrypsin-like activity, was prevalent.[18]

Are the Tumor Peptidome Precursor Proteins Highly Abundant or Unstable?

To examine whether the identified peptidome peptides are from known major cellular proteins, as an example, the OvCa peptidome precursor proteins were compared to the list of the most abundant proteins from our previous bottom-up study of exactly the same OvCa tumors.[35] Two hundred thirty-eight of the 974 nonredundant protein groups identified in the OvCa peptidome were among the 500 most abundant proteins in the OvCa tumors (Supporting Information Figure S10A), and 34 of these 238 proteins were among the 50 most abundant proteins (Supporting Information Figure S10B). We further compared the distributions of the spectral counts of the peptidome precursor proteins and the entire list of protein identifications in the bottom-up studies to examine the relative abundances of the peptidome precursor proteins. It appeared that more than 25% of the OvCa peptidome precursor proteins are relatively low abundance proteins in the proteome (with spectral count <5; Supporting Information Figure S10C); a similar trend was observed for the precursor proteins in the PDX BrCa peptidome (Supporting Information Figure S10D). Moreover, the relative abundances for the same precursor protein in the proteome and peptidome do not correlate well (R^2 < 0.2; results not shown). There are many potentially interesting low abundance precursor proteins that were identified with multiple peptidome peptide identifications, such as tumor protein D52 and protein FAM195B (transmembrane receptor protein tyrosine kinase activity) in the BrCa peptidome. The detailed information on these peptidome precursor proteins is provided in Supporting Information Table S3.

We also investigated whether the identified precursor proteins represent the most unstable proteins reported in studies of other cells. By comparing with the 600 proteins in A549 cells and 6000 proteins in HeLa cells with the highest turnover rates,[44,45] we found that 67 matched proteins in the A549 cell study and 684 matched proteins in the HeLa cell study. Only five precursor proteins were among the top 50 unstable proteins in A549 cells, and only 53 were among the top 500 category in HeLa cells (Supporting Information Figure S10E,F). Therefore, our data suggest that there is no correlation between the peptides identified in OvCa peptidome and the abundance/stability of the proteins from which these peptides are derived. Similar results were also obtained from the PDX breast tumor peptidome analysis (data not shown).

Degradation Substrates in BrCa and OvCa Tumors Are Involved in Broad Biological Processes

The biological processes in which the precursor proteins are involved were compared by gene ontology analyses between BrCa and OvCa data sets. As shown in Figure 5, approximately 25% of the 974 precursor proteins in ovarian cancer, representing the most prevalent proteolytic substrates, were participating in the metabolic processes. The second largest group of peptidome precursor proteins (~15%) was relevant to the cellular processes including cellular component morphogenesis, neural development, and immune responses. The rest of the precursor proteins were participating in processes such as transport, cell communication, reproduction, cellular component organization, and apoptosis, counting for two-thirds of total biological processes. However, none of these represents more than 10% of the total biological processes. Interestingly, the distribution pattern of the 824 precursor protein groups in the BrCa tumor peptidome was similar to that in the OvCa tumor (Supporting Information Figure S11). These results suggest that protein degradation is a global event in both OvCa and BrCa tumors.

DISCUSSION

We have developed an analytical platform for comprehensive and quantitative analyses of the tissue peptidomes. Replicate analysis of the WHIM6, WHIM2, and WHIM16 PDX breast tumors was first carried out to evaluate reproducibility and robustness of the platform. For studying potential ischemic effect on tumor peptidomic analysis, the WHIM6 tumors with ischemic times of 0 and 60 min (each with two replicates) and three clinical ovarian tumors each with ischemic times of 0, 5, 30, and 60 min were analyzed. The three different WHIM tumor samples were further analyzed to explore the potential of peptidomic profiles for distinguishing cancer subtypes. Both BrCa and OvCa tumor peptidomes were further characterized in terms of enzyme cleavage specificity and degradation substrates.

The study of the low molecular weight collection of bioreactive peptides, protein degradation products, and small intact proteins (i.e., peptidomics) is often criticized as being less sensitive, less reproducible, and as vulnerable to a lack of controls in sample collection and preparation.[46,47] In this study, we developed and evaluated a robust, efficient, and quantitative peptidomics platform and demonstrated its utility for in-depth characterization of human ovarian tumor and PDX breast tumor peptidomes. We also evaluated ischemic effects on peptidomics analysis and showed that the effects of up to 60 min of postexcision delay at room temperature are minimal in both tumor types. To our knowledge, our study represents the broadest characterization and first quantitative analysis of human cancer tumor peptidomes. This entire quantitative peptidomics platform encompassing sample preparation, LC-MS analysis, and data analysis provides a foundation for future peptidomics applications in biological and translational studies in cancer. Peptidomics patterns have been previously reported as capable of discriminating cancer types, such as breast, prostate, and bladder cancer, and healthy controls.[3] Our previous breast cancer plasma peptidomics study also showed that blood plasma peptidomic profiles are strikingly different between the early stage breast cancer patients and healthy controls, providing important, complementary information on top of conventional bottom-up proteomics.[18] In this study, substantially different peptidome profiles were observed for the three different WHIM lines (Figure 3A and Supporting Information S4C). Interestingly, the PCA plot showed a clear separation of the two WHIM lines that were both basal BrCa (i.e., WHIM2 and WHIM6), to the degree that was as strong as the separation between basal and luminal BrCa (e.g., WHIM2 and WHIM16; Figure 3B). Although the total number of PDX samples tested in this study was relatively small, the initial experimental results obtained herein demonstrate the potential of separating these BrCa subtypes using their peptidomics profiles. In the case of OvCa, samples from different patients were also clearly separated (Supporting Information Figure S5B). These results suggest that peptidomics profiles are sensitive to human pathophysical status and human hetero-

dx.doi.org/10.1021/pr500840w | J. Proteome Res. 2015, 14, 422−433

geneity and hence could provide highly valuable information on endogenous proteolytic events for cancer studies.

One of the major intracellular protein degradation pathways is mediated by the proteasome, which generally cleaves proteins after hydrophobic residues by its chymotrypsin-like β5 subunit and after basic residues by its trypsin-like β2 subunit.[48,49] The length of the proteasomal cleavage products ranges from 3 to 23 amino acid residues.[40,41] In addition, proteasomal degradation is expected to produce a larger number of internal fragments, whereas the proteolytic products of other cytoplasmic proteases are terminal ones. For OvCa, we found that the peptidome consisted of peptides rich in Leu, Lys, and Phe at the P1 position; more than half of the peptides fell within the range of proteasome products, the majority of which were internal peptides, suggesting major proteasome involvement in producing the OvCa tumor peptidome. The PDX breast tumor is harvested after the transplantation and growth of human BrCa tumor underneath the mouse skin[22] and hence exhibits a different tumor peptidome that appears to be the results of mixed proteasome and trypsin-like activities. This is presumably due to the vascularization that occurred during the tumor expansion phase and hence the shifted peptidomic–degradomic activities brought in by the mouse-origin proteins. Indeed, after removing the mouse-only and human–mouse-shared peptides from the PDX breast tumor peptidome, the proteasome still appears to be the major contributor for the human BrCa tumor peptidome, which is in agreement with the findings in our previous peptidomics studies in BrCa cell lines.

N-terminal methionine excision is an important proteolytic pathway regulating the diversity of N-terminal amino acids of proteins, conserved from prokaryotes to eukaryotes. Although Met is the first amino acid of newly synthesized proteins, it is often removed from the mature proteins for nonbulky N-terminal residue to avoid immediate degradation by the ubiquitin–proteasome system.[50] Besides being part of the protein-maturing mechanism, recent data suggest that NME also serves to regulate the protein's half-life based on the N-end rule at post-translational level throughout development, tumorigenesis, and in response to abiotic stress. For example, retained Met, Ser, Ala, Thr, Val, or Gly at N-terminus generally stabilize proteins for half-lives longer than 20 h, while proteins with N-terminal Phe, Leu, Asp, Lys, or Arg have half-lives of 3 min or less in mammals.[51] Therefore, NME is tightly regulated with other mechanisms including the ubiquitin–proteasome pathway to determine the fate of proteins under various circumstances.[52] Our data here revealed the prominent contribution of NME for the generation of OvCa tumor peptidome peptides. Based on the N-end rule, those proteins starting with nonbulky residues such as Ser, Ala, Thr, and Val (Supporting Information Figure S7A) are generally stabilized for longer half-lives instead of undergoing a proteolytic process, suggesting that NME contributes possibly at translational level rather than serves as a switch for protein degradation at post-translational levels in OvCa tumor cells.

In summary, in this study, we have evaluated a sensitive and robust tumor peptidomics platform that is amenable for high-throughput and reproducible quantitative cancer peptidomics studies. We have shown that protein degradation is a global event in both OvCa and BrCa tumors and that peptidomics profiling has the ability to differentiate cancer subtypes. Our study also indicates that warm ischemia effects are negligible up to 60 min of postexcision delay in both BrCa and OvCa tumor peptidomes, providing a basis for proper tumor specimen processing, handling, and selection for tumor peptidomics studies. Furthermore, we also show that chymotrypsin and trypsin activities of proteasome subunits are the major contributors to both the OvCa and BrCa tumor peptidome.

The observation that the cancer peptidome differs in a controlled, but unspecified, manner between the known subtypes of BrCa, as well as between BrCa and OvCa, suggests that there are cancer-specific changes in the regulation of normal proteolytic processes. The fact that varying times of warm ischemia, from 0 to 60 min, had little discernible effect on the peptidomic profile in either PDX xenografts or actual clinical samples of OvCa suggests that these cancer-specific proteolytic process are tightly regulated and not merely a nonspecific increase in protein degradation. These two observations suggest that a more extensive and tightly controlled analysis of tumor peptidomes, including the use of normal control tissues and clinical metadata, has the potential to provide new and potentially useful insights regarding cancer biology.

■ ASSOCIATED CONTENT

Supporting Information

Figure S1: Reproducibility of the tumor peptidomics platform. Figure S2: PDX breast tumor peptidome showed prominent peptidomics profile differences in different WHIM lines. Figure S3: Q–Q norm plot for normality test of WHIM-6-a-1 and A-0. Figure S4: Analysis of the OvCa peptidome data demonstrated minimal ischemic effect within 60 min of postexcision delay. Figure S5: Significantly improved peptide detection for OvCa peptidome using the IQ approach. Figure S6: Analysis of the cleavage specificities at P2–P6 and P2′–P6′ sites for OvCa tumor peptidome peptides. Figure S7: Analysis of the cleavage specificities at the N- and C-termini for peptidome peptides from OvCa tumor samples. Figure S8: Analysis of the cleavage specificities at P1 and P1′ sites for the human-only peptides from PDX breast tumor samples. Figure S9: Analysis of the cleavage specificities at P2–P6 and P2′–P6′ sites for human-only peptides from PDX breast tumor samples. Figure S10: The majority of OvCa peptidome peptides are not likely to derive from either the most abundant proteins or the least stable proteins. Figure S11: Functional classification of identified precursor proteins from OvCa and PDX breast tumor peptidomes. This material is available free of charge via the Internet at http://pubs.acs.org. The tumor peptidomics data sets are available at PeptideAtlas through accession PASS00572.

■ AUTHOR INFORMATION

Corresponding Authors

*Phone: 509-371-6346. Fax: 509-371-6564. E-mail: tao.liu@pnnl.gov.

*Phone: 509-371-6576. Fax: 509-371-6564. E-mail: dick.smith@pnnl.gov.

Author Contributions

¶Z.X. and C.W. contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Portions of this work were supported by Grant U24CA160019, from the National Cancer Institute Clinical Proteomic Tumor Analysis Consortium (CPTAC), National Institutes of Health

Grant P41GM103493, Department of Defense Interagency Agreement MIPR2DO89M2058, and Department of Energy (DOE) Early Career Award (to S.H.P.). The experimental work described herein was performed in the Environmental Molecular Sciences Laboratory, a national scientific user facility sponsored by the DOE and located at Pacific Northwest National Laboratory, which is operated by Battelle Memorial Institute for the DOE under Contract DE-AC05-76RL0 1830.

REFERENCES

(1) Kondakova, I. V.; Spirina, L. V.; Shashova, E. E.; Koval, V. D.; Kolomiets, L. A.; Chernysheva, A. L.; Slonimskaia, E. M. Proteasome activity in tumors of female reproductive system. *Bioorg. Khim.* **2012**, *38* (1), 106−110.

(2) Willumsen, N.; Bager, C. L.; Leeming, D. J.; Smith, V.; Karsdal, M. A.; Dornan, D.; Bay-Jensen, A. C. Extracellular matrix specific protein fingerprints measured in serum can separate pancreatic cancer patients from healthy controls. *BMC Cancer* **2013**, *13* (1), 554.

(3) Villanueva, J.; Shaffer, D. R.; Philip, J.; Chaparro, C. A.; Erdjument-Bromage, H.; Olshen, A. B.; Fleisher, M.; Lilja, H.; Brogi, E.; Boyd, J.; Sanchez-Carbayo, M.; Holland, E. C.; Cordon-Cardo, C.; Scher, H. I.; Tempst, P. Differential exoprotease activities confer tumor-specific serum peptidome patterns. *J. Clin. Invest.* **2006**, *116* (1), 271−284.

(4) Lopez-Otin, C.; Overall, C. M. Protease degradomics: a new challenge for proteomics. *Nat. Rev. Mol. Cell Biol.* **2002**, *3* (7), 509−519.

(5) Gelman, J. S.; Sironi, J.; Castro, L. M.; Ferro, E. S.; Fricker, L. D. Peptidomic analysis of human cell lines. *J. Proteome Res.* **2011**, *10* (4), 1583−1592.

(6) Fricker, L. D.; Gelman, J. S.; Castro, L. M.; Gozzo, F. C.; Ferro, E. S. Peptidomic analysis of HEK293T cells: effect of the proteasome inhibitor epoxomicin on intracellular peptides. *J. Proteome Res.* **2012**, *11* (3), 1981−1990.

(7) Marshall, J.; Kupchak, P.; Zhu, W.; Yantha, J.; Vrees, T.; Furesz, S.; Jacks, K.; Smith, C.; Kireeva, I.; Zhang, R.; Takahashi, M.; Stanton, E.; Jackowski, G. Processing of serum proteins underlies the mass spectral fingerprinting of myocardial infarction. *J. Proteome Res.* **2003**, *2* (4), 361−372.

(8) Zhang, Z.; Bast, R. C.; Yu, Y. H.; Li, J. N.; Sokoll, L. J.; Rai, A. J.; Rosenzweig, J. M.; Cameron, B.; Wang, Y. Y.; Meng, X. Y.; Berchuck, A.; van Haaften-Day, C.; Hacker, N. F.; de Bruijn, H. W. A.; van der Zee, A. G. J.; Jacobs, I. J.; Fung, E. T.; Chan, D. W. Three biomarkers identified from serum proteomic analysis for the detection of early stage ovarian cancer. *Cancer Res.* **2004**, *64* (16), 5882−5890.

(9) Koomen, J. M.; Li, D. H.; Xiao, L. C.; Liu, T. C.; Coombes, K. R.; Abbruzzese, J.; Kobayashi, R. Direct tandem mass spectrometry reveals limitations in protein profiling experiments for plasma biomarker discovery. *J. Proteome Res.* **2005**, *4* (3), 972−981.

(10) Jurgens, M.; Appel, A.; Heine, G.; Neitz, S.; Menzel, C.; Tammen, H.; Zucht, H. D. Towards characterization of the human urinary peptidome. *Comb. Chem. High Throughput Screening* **2005**, *8* (8), 757−765.

(11) Fiedler, G. M.; Baumann, S.; Leichtle, A.; Oltmann, A.; Kase, J.; Thiery, J.; Ceglarek, U. Standardized peptidome profiling of human urine by magnetic bead separation and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Clin. Chem.* **2007**, *53* (3), 421−428.

(12) Strand, F. L. Neuropeptides: general characteristics and neuropharmaceutical potential in treating CNS disorders. *Prog. Drug Res.* **2003**, *61*, 1−37.

(13) Hummon, A. B.; Amare, A.; Sweedler, J. V. Discovering new invertebrate neuropeptides using mass spectrometry. *Mass Spectrom Rev.* **2006**, *25* (1), 77−98.

(14) Yatskin, O. N.; Karelin, A. A.; Ivanov, V. T. Peptidomes of the brain, heart, lung, and spleen of a rat: similarity and differences. *Russ. J. Bioorg. Chem.* **2009**, *35* (4), 426−436.

(15) Lassout, O.; Pastor, C. M.; Fetaud-Lapierre, V.; Hochstrasser, D. F.; Frossard, J. L.; Lescuyer, P. Analysis of the pancreatic low molecular weight proteome in an animal model of acute pancreatitis. *J. Proteome Res.* **2010**, *9* (9), 4535−4544.

(16) Alvarez, M. T. B.; Shah, D. J.; Thulin, C. D.; Graves, S. W. Tissue proteomics of the low-molecular weight proteome using an integrated cLC-ESI-QTOFMS approach. *Proteomics* **2013**, *13* (9), 1400−1411.

(17) Cunha, F. M.; Berti, D. A.; Ferreira, Z. S.; Klitzke, C. F.; Markus, R. P.; Ferro, E. S. Intracellular peptides as natural regulators of cell signaling. *J. Biol. Chem.* **2008**, *283* (36), 24448−24459.

(18) Shen, Y.; Tolic, N.; Liu, T.; Zhao, R.; Petritis, B. O.; Gritsenko, M. A.; Camp, D. G.; Moore, R. J.; Purvine, S. O.; Esteva, F. J.; Smith, R. D. Blood peptidome−degradome profile of breast cancer. *PLoS One* **2010**, *5* (10), e13133.

(19) de Bruin, J. S.; Deelder, A. M.; Palmblad, M. Scientific workflow management in proteomics. *Mol. Cell. Proteomics* **2012**, *11* (7), M111.010595.

(20) Zhu, X.; Desiderio, D. M. Methionine enkephalin-like immunoreactivity, substance P-like immunoreactivity and beta-endorphin-like immunoreactivity post-mortem stability in rat pituitary. *J. Chromatogr.* **1993**, *616* (2), 175−187.

(21) Shahinian, H.; Loessner, D.; Biniossek, M. L.; Kizhakkedathu, J. N.; Clements, J. A.; Magdolen, V.; Schilling, O. Secretome and degradome profiling shows that Kallikrein-related peptidases 4, 5, 6, and 7 induce TGFβ-1 signaling in ovarian cancer cells. *Mol. Oncol.* **2014**, *8* (1), 68−82.

(22) Li, S.; Shen, D.; Shao, J.; Crowder, R.; Liu, W.; Prat, A.; He, X.; Liu, S.; Hoog, J.; Lu, C.; Ding, L.; Griffith, O. L.; Miller, C.; Larson, D.; Fulton, R. S.; Harrison, M.; Mooney, T.; McMichael, J. F.; Luo, J.; Tao, Y.; Goncalves, R.; Schlosberg, C.; Hiken, J. F.; Saied, L.; Sanchez, C.; Giuntoli, T.; Bumb, C.; Cooper, C.; Kitchens, R. T.; Lin, A.; Phommaly, C.; Davies, S. R.; Zhang, J.; Kavuri, M. S.; McEachern, D.; Dong, Y. Y.; Ma, C.; Pluard, T.; Naughton, M.; Bose, R.; Suresh, R.; McDowell, R.; Michel, L.; Aft, R.; Gillanders, W.; DeSchryver, K.; Wilson, R. K.; Wang, S.; Mills, G. B.; Gonzalez-Angulo, A.; Edwards, J. R.; Maher, C.; Perou, C. M.; Mardis, E. R.; Ellis, M. J. Endocrine-therapy-resistant ESR1 variants revealed by genomic characterization of breast-cancer-derived xenografts. *Cell Rep.* **2013**, *4* (6), 1116−1130.

(23) Ding, L.; Ellis, M. J.; Li, S.; Larson, D. E.; Chen, K.; Wallis, J. W.; Harris, C. C.; McLellan, M. D.; Fulton, R. S.; Fulton, L. L.; Abbott, R. M.; Hoog, J.; Dooling, D. J.; Koboldt, D. C.; Schmidt, H.; Kalicki, J.; Zhang, Q.; Chen, L.; Lin, L.; Wendl, M. C.; McMichael, J. F.; Magrini, V. J.; Cook, L.; McGrath, S. D.; Vickery, T. L.; Appelbaum, E.; Deschryver, K.; Davies, S.; Guintoli, T.; Lin, L.; Crowder, R.; Tao, Y.; Snider, J. E.; Smith, S. M.; Dukes, A. F.; Sanderson, G. E.; Pohl, C. S.; Delehaunty, K. D.; Fronick, C. C.; Pape, K. A.; Reed, J. S.; Robinson, J. S.; Hodges, J. S.; Schierding, W.; Dees, N. D.; Shen, D.; Locke, D. P.; Wiechert, M. E.; Eldred, J. M.; Peck, J. B.; Oberkfell, B. J.; Lolofie, J. T.; Du, F.; Hawkins, A. E.; O'Laughlin, M. D.; Bernard, K. E.; Cunningham, M.; Elliott, G.; Mason, M. D.; Thompson, D. M., Jr.; Ivanovich, J. L.; Goodfellow, P. J.; Perou, C. M.; Weinstock, G. M.; Aft, R.; Watson, M.; Ley, T. J.; Wilson, R. K.; Mardis, E. R. Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature* **2010**, *464* (7291), 999−1005.

(24) Eng, J. K.; McCormack, A. L.; Yates, J. R. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **1994**, *5* (11), 976−989.

(25) Kim, S.; Gupta, N.; Pevzner, P. A. Spectral probabilities and generating functions of tandem mass spectra: a strike against decoy databases. *J. Proteome Res.* **2008**, *7* (8), 3354−3363.

(26) Liu, X.; Sirotkin, Y.; Shen, Y.; Anderson, G.; Tsai, Y. S.; Ting, Y. S.; Goodlett, D. R.; Smith, R. D.; Bafna, V.; Pevzner, P. A. Protein identification using top-down. *Mol. Cell. Proteomics* **2012**, *11* (6), M111.008524.

(27) Shen, Y.; Liu, T.; Tolic, N.; Petritis, B. O.; Zhao, R.; Moore, R. J.; Purvine, S. O.; Camp, D. G.; Smith, R. D. Strategy for degradomic-

peptidomic analysis of human blood plasma. *J. Proteome Res.* **2010**, *9* (5), 2339−2346.
(28) Jaitly, N.; Mayampurath, A.; Littlefield, K.; Adkins, J. N.; Anderson, G. A.; Smith, R. D. Decon2LS: an open-source software package for automated processing and visualization of high resolution mass spectrometry data. *BMC Bioinf.* **2009**, *10*, 87.
(29) Monroe, M. E.; Tolic, N.; Jaitly, N.; Shaw, J. L.; Adkins, J. N.; Smith, R. D. VIPER: an advanced software package to support high-throughput LC-MS peptide identification. *Bioinformatics* **2007**, *23* (15), 2021−2023.
(30) Slysz, G. W.; Steinke, L.; Ward, D. M.; Klatt, C. G.; Clauss, T. R.; Purvine, S. O.; Payne, S. H.; Anderson, G. A.; Smith, R. D.; Lipton, M. S. Automated data extraction from in situ protein-stable isotope probing studies. *J. Proteome Res.* **2014**, *13* (3), 1200−1210.
(31) Holman, J. D.; Ma, Z. Q.; Tabb, D. L. Identifying proteomic LC-MS/MS data sets with Bumbershoot and IDPicker. *Current Protocols in Bioinformatics*; Wiley Online Library: New York, 2012; Chapter 13, Unit 13.17.
(32) Taverner, T.; Karpievitch, Y. V.; Polpitiya, A. D.; Brown, J. N.; Dabney, A. R.; Anderson, G. A.; Smith, R. D. DanteR: an extensible R-based tool for quantitative analysis of -omics data. *Bioinformatics* **2012**, *28* (18), 2404−2406.
(33) Schechter, I.; Berger, A. On the active site of proteases. 3. Mapping the active site of papain; specific peptide inhibitors of papain. *Biochem. Biophys. Res. Commun.* **1968**, *32* (5), 898−902.
(34) Smyth, G. K. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* **2004**, *3*, 3.
(35) Mertins, P.; Yang, F.; Liu, T.; Mani, D. R.; Petyuk, V. A.; Gillette, M. A.; Clauser, K. R.; Qiao, J. W.; Gritsenko, M. A.; Moore, R. J.; Levine, D. A.; Townsend, R.; Erdmann-Gilmore, P.; Snider, J. E.; Davies, S. A.; Ruggles, K. V.; Fenyo, D.; Kitchens, R. T.; Li, S.; Olvera, N.; Dao, F.; Rodriguez, H.; Chan, D. W.; Liebler, D.; White, F.; Rodland, K. D.; Mills, G. B.; Smith, R. D.; Paulovich, A. G.; Ellis, M.; Carr, S. A. Ischemia in tumors induces early and sustained phosphorylation changes in stress kinase pathways but does not affect global protein levels. *Mol. Cell. Proteomics* **2014**, *13* (7), 1690−1704.
(36) Smith, R. D.; Anderson, G. A.; Lipton, M. S.; Pasa-Tolic, L.; Shen, Y.; Conrads, T. P.; Veenstra, T. D.; Udseth, H. R. An accurate mass tag strategy for quantitative and high-throughput proteome measurements. *Proteomics* **2002**, *2* (5), 513−523.
(37) Zimmer, J. S.; Monroe, M. E.; Qian, W. J.; Smith, R. D. Advances in proteomics data analysis and display using an accurate mass and time tag approach. *Mass Spectrom. Rev.* **2006**, *25* (3), 450−482.
(38) Keil, B. *Specificity of Proteolysis*; Springer-Verlag: Berlin, 1992.
(39) Nixon, R. A.; Saito, K. I.; Grynspan, F.; Griffin, W. R.; Katayama, S.; Honda, T.; Mohan, P. S.; Shea, T. B.; Beermann, M. Calcium-activated neutral proteinase (calpain) system in aging and Alzheimer's disease. *Ann. N.Y. Acad. Sci.* **1994**, *747*, 77−91.
(40) Wenzel, T.; Eckerskorn, C.; Lottspeich, F.; Baumeister, W. Existence of a molecular ruler in proteasomes suggested by analysis of degradation products. *FEBS Lett.* **1994**, *349* (2), 205−209.
(41) Glickman, M. H.; Ciechanover, A. The ubiquitin−proteasome proteolytic pathway: destruction for the sake of construction. *Physiol. Rev.* **2002**, *82* (2), 373−428.
(42) Van Damme, P.; Maurer-Stroh, S.; Plasman, K.; Van Durme, J.; Colaert, N.; Timmerman, E.; De Bock, P. J.; Goethals, M.; Rousseau, F.; Schymkowitz, J.; Vandekerckhove, J.; Gevaert, K. Analysis of protein processing by N-terminal proteomics reveals novel species-specific substrate determinants of granzyme B orthologs. *Mol. Cell. Proteomics* **2009**, *8* (2), 258−272.
(43) Dick, T. P.; Nussbaum, A. K.; Deeg, M.; Heinemeyer, W.; Groll, M.; Schirle, M.; Keilholz, W.; Stevanovic, S.; Wolf, D. H.; Huber, R.; Rammensee, H. G.; Schild, H. Contribution of proteasomal beta-subunits to the cleavage of peptide substrates analyzed with yeast mutants. *J. Biol. Chem.* **1998**, *273* (40), 25637−25646.
(44) Doherty, M. K.; Hammond, D. E.; Clague, M. J.; Gaskell, S. J.; Beynon, R. J. Turnover of the human proteome: determination of protein intracellular stability by dynamic SILAC. *J. Proteome Res.* **2009**, *8* (1), 104−112.
(45) Boisvert, F. M.; Ahmad, Y.; Gierlinski, M.; Charriere, F.; Lamont, D.; Scott, M.; Barton, G.; Lamond, A. I. A quantitative spatial proteomics analysis of proteome turnover in human cells. *Mol. Cell. Proteomics* **2012**, *11* (3), M111.011429.
(46) Fricker, L. D.; Lim, J.; Pan, H.; Che, F. Y. Peptidomics: identification and quantification of endogenous peptides in neuroendocrine tissues. *Mass Spectrom. Rev.* **2006**, *25* (2), 327−344.
(47) Leichtle, A. B.; Dufour, J. F.; Fiedler, G. M. Potentials and pitfalls of clinical peptidomics and metabolomics. *Swiss Med. Wkly.* **2013**, *143*, w13801.
(48) Voges, D.; Zwickl, P.; Baumeister, W. The 26S proteasome: a molecular machine designed for controlled proteolysis. *Annu. Rev. Biochem.* **1999**, *68*, 1015−1068.
(49) Zwickl, P.; Voges, D.; Baumeister, W. The proteasome: a macromolecular assembly designed for controlled proteolysis. *Philos. Trans. R Soc. London, Ser. B* **1999**, *354* (1389), 1501−1511.
(50) Sherman, F.; Stewart, J. W.; Tsunasawa, S. Methionine or not methionine at the beginning of a protein. *Bioessays* **1985**, *3* (1), 27−31.
(51) Varshavsky, A. The N-end rule: functions, mysteries, uses. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93* (22), 12142−12149.
(52) Giglione, C.; Boularot, A.; Meinnel, T. Protein N-terminal methionine excision. *Cell. Mol. Life Sci.* **2004**, *61* (12), 1455−1474.