Off-Line Multidimensional Liquid Chromatography and Auto Sampling Result in Sample Loss in LC/LC−MS/MS

Sameh Magdeldin,*,†,‡,§ James J. Moresco,† Tadashi Yamamoto,‡ and John R. Yates, III*,†

†Department of Chemical Physiology, The Scripps Research Institute, 10550 North Torrey Pines Road, SR11, La Jolla, California 92037, United States

‡Department of Structural Pathology, Institute of Nephrology, Graduate School of Medical and Dental Sciences, Niigata University, 1-757 Asahimachi-dori, Niigata 951-8510, Japan

§Department of Physiology, Faculty of Veterinary Medicine, Suez Canal University, Ismailia 41522, Egypt

ⓢ *Supporting Information*

ABSTRACT: Large-scale proteomics often employs two orthogonal separation methods to fractionate complex peptide mixtures. Fractionation can involve ion exchange separation coupled to reversed-phase separation or, more recently, two reversed-phase separations performed at different pH values. When multidimensional separations are combined with tandem mass spectrometry for protein identification, the strategy is often referred to as multidimensional protein identification technology (MudPIT). MudPIT has been used in either an automated (online) or manual (offline) format. In this study, we evaluated the performance of different MudPIT strategies by both label-free and tandem mass tag (TMT) isobaric tagging. Our findings revealed that online MudPIT provided more peptide/protein identifications and higher sequence coverage than offline platforms. When employing an off-line fractionation method with direct loading of samples onto the column from an eppendorf tube via a high-pressure device, a 5.3% loss in protein identifications is observed. When off-line fractionated samples are loaded via an autosampler, a 44.5% loss in protein identifications is observed compared with direct loading of samples onto a triphasic capillary column. Moreover, peptide recovery was significantly lower after offline fractionation than in online fractionation. Signal-to-noise (S/N) ratio, however, was not significantly altered between experimental groups. It is likely that offline sample collection results in stochastic peptide loss due to noncovalent adsorption to solid surfaces. Therefore, the use of the offline approaches should be considered carefully when processing minute quantities of valuable samples.

KEYWORDS: *MudPIT, online, offline, label-free, TMT, isobaric, HEK293*
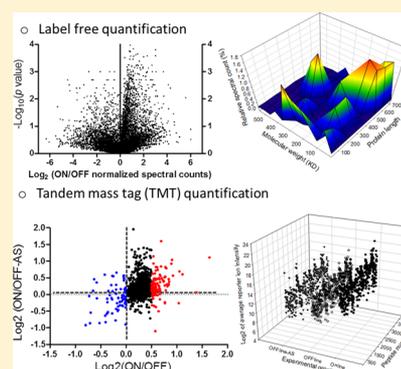
■ INTRODUCTION

The analysis of proteins is an integral component of biological studies. As genomes have been sequenced and sequence databases have been compiled, these databases have been used to identify proteins present in biological samples using tandem mass spectrometry data. In a process called "shotgun proteomics", mixtures of proteins from biological samples are digested to peptides using proteases, and liquid chromatography coupled to tandem mass spectrometry is used to directly identify the peptides present.[1−3] Identification of a peptide is used to infer the presence of the protein it is derived from. As more complex samples, such as cellular components and whole cells, have become targets of analysis, more comprehensive separations have become necessary to resolve complex peptide mixtures and increase dynamic range.[3] Giddings described multidimensional chromatography as a means to increase peak capacity by combining orthogonal separations.[4] For the analysis of complex peptide mixtures, several combinations of separation media have been used. The last phase in multidimensional liquid chromatography (MDLC) separations used for mass spectrometry is typically reversed-phase (RP), which separates peptides by hydrophobicity and is effective at removing salts or other small molecule contaminants prior to introduction of peptides into the mass spectrometer.[5] Many different forms of MDLC have appeared over the years, from combinations of ion exchange (IEX) with RP separations to combinations of capillary electrophoresis and liquid chromatography.[6−15]

In any multidimensional separation, how material is transferred from one separation stage to another is critical for maximizing peak capacity and optimizing sample recovery. In proteomics, several strategies have been used for multidimensional separations. Link et al. employed a biphasic column of strong cation exchange (SCX) and RP packing material in a single column. In this arrangement, a multidimensional separation is created by running buffer containing a set concentration of salt across the column to elute peptides from the SCX phase onto the RP column. Once the salt pulse is

completed, the RP buffer is applied to the column at 0% B to remove salt from the column prior to running the gradient. After a RP gradient is completed, a second salt pulse with a higher concentration of salt is applied to the column to move a new population of peptides to the RP material. The process of salt pulse/RP gradient is repeated until all peptides are removed from the SCX phase. There are a few unusual features to this strategy. A single column contains both the IEX phase and the RP phase, and all solvents flow over both phases. The sample is loaded directly onto the column from an eppendorf tube using a pressurized device, and the column is placed in line with the ion source with the voltage placed on a waste line at the backend of the column.[16,17] This method is commonly referred to as online MDLC and may employ a bi- or triphasic column.[2,3,5,18,19] A second strategy employs off-line fractionation.[20] This method usually employs SCX, but strong anion exchange (SAX) has also been used.[21] Advantages to off-line fractionation include the ability to add a high organic phase to the salt buffer (e.g., 25% organic in the IEX buffer) to minimize mixed-mode interactions, the capability to collect many fractions, and the capacity to load large amounts of material onto the column. In a clever use of off-line fractionation, Wang et al. showed improvement in peptide identifications by combining fractions from different parts of a high pH RP separation to produce collections of peptides with different physical characteristics like hydrophobicity for the second-dimension, low-pH RP separation.[22,23] After RP separation, excess solvent in the collected fractions is removed, and each fraction is loaded into an autosampler for introduction into the mass spectrometer. A third strategy for multidimensional separation is also an online method that employs a valving system to direct solvents to an IEX column, an enrichment column, a RP column, and waste. This system represents a compromise between the direct online and the off-line approaches. The valving system is used to redirect flow to shunt salt solutions used to elute peptides from an IEX column to waste rather than have it run through the RP analytical column, or in the case of a RP-RP LCLC system, the valves are used to direct peptides to the enrichment column to alter the pH of the buffer before the analytical separation.

Sample loading is a critical part of capillary chromatography, as these systems involve small diameter openings that must be aligned and low solvent flows, for which dead space can have a great impact. Kennedy and Jorgenson developed a pressurized device for both packing and loading capillary columns.[16] The end of the column was placed directly into a slurry of packing material and when the device was pressurized the packing material was driven into the column. This same strategy could be used as a means to load samples directly from eppendorf tubes into a column. This method has been adopted by others as a means to load small quantities of samples.[24]

An important consideration when analyzing peptides and proteins is sample loss associated with sample handling. Proteins and peptides can easily adhere to surfaces resulting in losses. A carrier protein is often used to minimize adherence to active surfaces during sample manipulations to protect low abundance proteins from losses.[25] An advantage of shotgun proteomics is the manipulation of complex protein mixtures where the more abundant proteins may presumably act as carrier proteins to protect lower abundance proteins from loss. Because losses can occur on active surfaces such a glass and metal, efforts have been directed to using biocompatible materials to reduce such sample losses. Two recent papers showed that peptides can be lost when analyzed from autosamplers, and peptide mixtures of inter-mediate complexity (in gel digestions) can be lost to the surface of autosampler vials made from a variety of materials, respectively.[26,27] Stejskal et al. tested a variety of carriers to determine which one improved recovery of peptides.[27] As these two papers have shown, and as experience has taught us, the more samples are handled and exposed to surfaces, the greater the loss. This study compares sample losses for a shotgun proteomics experiment using three different methods of sample introduction using two different quantitation methods.

■ MATERIALS AND METHODS

Cell Culture and Protein Extraction

HEK293 cells (CRL-1573) purchased from ATCC were seeded into a T25 flask in supplemented media (DMEM medium supplemented with 10% fetal bovine serum, 1% (v/v) penicillin/streptomycin, 2 mM L-glutamine, and 200 μg/mL G418) from (Gibco, Invitrogen)),and maintained with regular media changes for 3 weeks before they were considered to be stable cell lines.[28] Cultured cells were harvested at ∼80% confluency with 0.05% trypsin and EDTA, centrifuged for 5 min at 4000g at 4 °C, and washed twice with PBS. Cells were suspended in 8 M urea, 500 mM Tris-HCl pH 8.5 supplemented with complete ultra tablets, mini, EASYpack (Roche, Mannheim) for protein extraction.

In-Solution Digestion

Denaturated protein lysate was precipitated with acetone and assayed using modified bicinchoninic (BCA) method[29] (Pierce, Rockford IL). Resuspended protein was reduced with 5 mM tris(2-carboxyethyl)phosphine (TCEP) for 30 min. Cysteine residues were alkylated with 10 mM iodoacetamide for 20 min in the dark.[30,31] Samples were diluted to a final concentration 2 M urea with 100 mM Tris-HCl, pH 8.5 prior to digestion with trypsin. For endopeptidase digestion, modified trypsin (Promega, Madison, WI) was added at 50:1 (protein/protease mass ratio) along with 1 mM CaCl$_2$ and incubated overnight in a thermoshaker at 600 rpm at 37 °C. Digested peptide solution was acidified using 90% FA to a final pH of 3.0. The resulting peptide mixture was used for evaluating online and offline MudPIT techniques.

Online and Offline Multiprotein Identification Technology Methods

Capillary columns were prepared in-house using particle slurries in methanol. An analytical RPLC column was prepared by pulling a 100 μm ID/360 μm OD capillary (Polymicro Technologies, Phoenix, AZ) to 5 μm ID tip. Reversed-phase resin (Aqua C18, 3 μm dia., 90 Å pores, Phenomenex, Torrance, CA) was packed directly into the pulled column at 700 psi until 12 cm long. The column was washed and equilibrated at 100 bar with buffer B, followed by buffer A.[30] A multiprotein identification technology (MudPIT) trapping column was prepared by creating a Kasil frit at one end of an undeactivated 250 μm ID/360 μm OD capillary (Agilent Technologies, Santa Clara, CA). The frit was prepared by briefly dipping a 20 cm capillary in well-mixed 300 μL of Kasil 1624 (PQ Corporation, Malvern, PA) and 100 μL of formamide, curing at 100 °C overnight, and cutting the frit to ∼1.5 mm in length.[30] Triphasic[32] or biphasic[33] columns were successively packed with 2.5 cm SCX particles (Partisphere SCX, 5 μm dia., 100 Å pores, Phenomenex) and 2.5 cm RP resin (Aqua C18, 3 μm dia., 125 Å pores, Phenomenex), as shown in Figure 1. Peptide samples (∼100 μg) were loaded onto triphasic columns for online MudPIT. For offline MudPIT, samples were loaded onto a biphasic column and ten SCX offline fractions were collected in 1.5 mL eppendorf tubes. Fractions were then loaded into a 2.5 RP resin column (offline MudPIT) or purified by stage
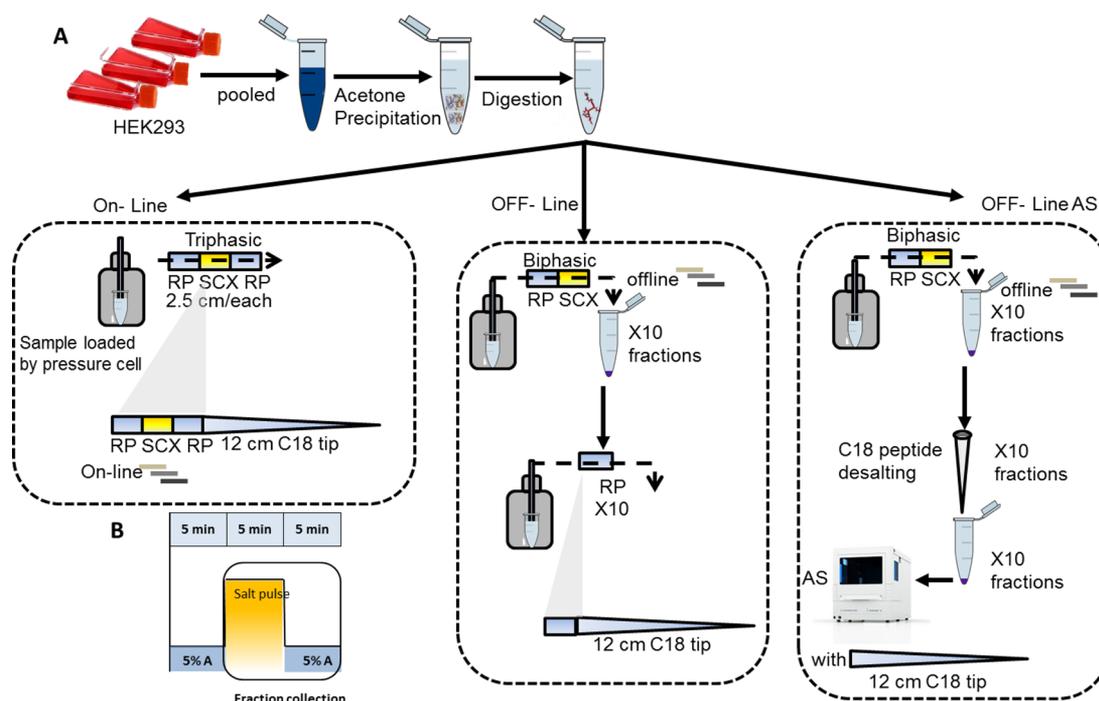
Figure 1. Schematic design workflow for the label-free MudPIT platforms quantification. (A) HEK293 protein lysate was digested and processed in three replicate runs with different MudPIT panels: online (automated panel), offline (manual collection of fractions), and offline-AS (offline with autosampler where fractions were collected manually then cleaned up with C18 stage tip columns before being placed into autosampler). (B) Offline fractions were collected for 5 min after the volatile salt pulse phase using 10−100% ammonium acetate.

tip and placed in autosampler vials (offline MudPIT with autosampler [EASY-nLC II, Thermo]). Both MudPIT and analytical columns were assembled using a zero-dead volume union (Upchurch Scientific, Oak Harbor, WA).

Tandem Mass Tag Isobaric Labeling

Sixplex tandem mass tag (TMT) labeling[34] was performed according to the manufacturer's instructions (Thermo Fisher Scientific, Rockford, IL). As illustrated in Figure 2, TMT reagents (0.8 mg) were dissolved in 40 μL of anhydrous ACN (Sigma, Milwaukee). Trypsin-digested HEK293 cells samples (25 μg/tag) were equilibrated to room temperature for 5 min with occasional vortexing. Samples were then resuspended in 100 mM TEAB and derivatized with sixplex chemical tags: 126, 127 for online MudPIT; 128, 129 for offline MudPIT; and 130 and 131 Th (Thomson) for offline MudPIT with autosampler. The reaction mixtures were incubated at room temperature for 1 h and quenched with 15 μL of 5% hydroxylamine solution in water. Equal ratios of TMT- tagged samples were mixed and analyzed prior to fractionation to ensure unbiased and impartial labeling. Each TMT-modified digest was fractionated either online or offline. Fractions were massed up to same volume with 100 mM TEAB and equally combined into one sample before vacuum drying. The lyophilized TMT-labeled peptides (25 μg) were reconstituted with 50 μL of buffer A (0.1% formic acid (FA), 5% acetonitrile (ACN) in water), centrifuged at 12 000g for 30 min prior to mass spectrometric analysis.

LC−MS/MS Analysis

Peptides were separated by an Eksigent NanoLC-2D system (Eksigent, Dublin) with or without autosampler unit (10 μL PEEK sample loop, six-port titanium injection valve, 50 mm SUS sample needle, 50 μm ID fused silica tubing). The HPLC system was either connected online or offline to Thermo LTQ XL (for label-free quantification) or LTQ-Orbitrap Velos (for TMT

quantification) using an in-house built nanoelectrospray stage. Electrospray was performed directly from the analytical column by applying the ESI voltage at a tee (150 μm ID, Upchurch Scientific) directly downstream of a 1:1000 split flow used to reduce the flow rate to 250 nL/min through the columns.[35] Ten-step MudPIT experiments were performed either online or offline, with steps corresponding to 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100% buffer C being run for 5 min at the beginning of a 120 min gradient. A three mobile phase system consisting of buffer A (5% ACN; 0.1% FA (Sigma-Aldrich, St. Louis, MO)), buffer B (80% ACN, 0.1% FA), and buffer C (500 mM ammonium acetate, 5% ACN, 0.1% FA) was used in the current experiment. The LC system was coupled to 224 nm laser-induced native fluorescence (LINF) detector with elliptical flow cell for real-time peptide detection.[36] Data-dependent acquisition of MS/MS spectra was performed by dynamically choosing up to 5 or 10 most intense precursor ions from the survey scan for LTQ XL or LTQ Orbitrap Velos, respectively. The following settings were applied: mass range 300−1600 Th., charge $\geq 2-5$, full-scan MS resolution of 30.000 (LTQ XL) and 60.000 (LTQ Orbitrap Velos) with a target value of 1×10^6, and the maximal injection time of 200 ms. The lower threshold for targeting a precursor ion in the MS scan was 5000 counts and 2.5 kV maximum injection time for higher-energy collisional dissociation (HCD)−MS/MS analysis in the Orbitrap. The HCD dissociation mode enables simultaneous production of TMT reporter ions and fragment ions of the peptides.[37] MS/MS scans were acquired in the Orbitrap with a mass resolution of 17 000. The target value was 30 000 ions with injection time of 150 ms. Once analyzed, the selected peptide ions were dynamically excluded from further analysis for 120 s to allow for the selection of lower-abundance ions for subsequent fragmentation and detection using the setting for repeat count = 1, repeat duration = 30 ms, and exclusion list size = 500. Ions with singly or
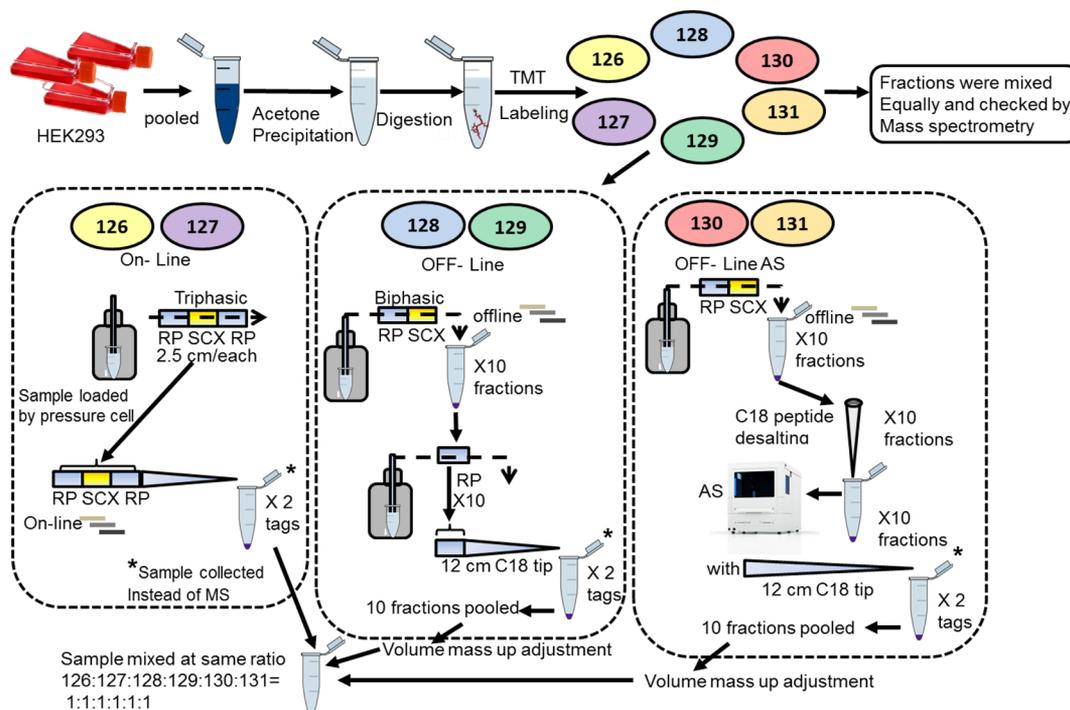
Figure 2. Schematic design workflow for the isobaric tandem mass tag (TMT[6]) quantification. HEK293 protein lysate was labeled in a sixplex TMT format: 126 and 127 for online MudPit, 128 and 129 for the offline MudPit, and 130 and 131 for the offline with autosampler MudPit (offline-AS). Directly after labeling, an equal volume of the TMT tags was mixed and analyzed to ensure successful labeling (labeling efficiency averaged 97.2%). Samples were then processed with different MudPit platforms, massed up to same volume with 100 mM TEAB, and equally combined into one sample before drying and mass analysis.

unassigned charge states were rejected. Activation time of 0.1 ms was used. The m/z isolation width for MS/MS fragmentation was set to 2 Th. For MS/MS, precursor ions were activated using 35% normalized collision energy.

Data Analysis

Tandem mass spectra were extracted from raw files using RawExtract 1.9.9[38] and searched with the ProLuCID algorithm[39] against *Homo sapiens* UniProt/Swiss-Prot database with reversed sequences (176 708 entries). The search space included all fully and semitryptic peptide candidates (at least six amino acids). Carbamidomethylation of cysteine (57.02146 amu) was considered as a static modification as well as static N-terminus and lysine modification (229.1629 amu) for sixplex TMT labels analysis.[37] The search parameters include 10 ppm precursor mass tolerance and 0.6 Da peptide mass tolerance. Exported ProLuCID files were assembled and filtered using the DTASelect2.0 which combines XCorr and DeltaCN values using a quadratic discriminate function to compute a confidence score.[40] The false-positive rate (FDR) was kept at 1% at the protein level. For quantitative analysis, Census was used to extract the relative intensities of reporter ions for each peptide from the identified tandem mass spectra for normalization.[41] The mass tolerance and intensity threshold for the reporter ions in Census were set at 0.05 Da and 5000, respectively.

Biostatistics

Statistical analysis was performed using a Kruskal−Wallis test with Dunns post hock test. $P \leq 0.05$ was considered to be statistically significant. For each result, ProLuCID XCorr, DeltaCN, and ZScore values were used to generate a Bayesian discriminator. Outlier points in the two distributions that had a Mahalanobis distance greater than four were discarded. For label-free quantification, normalized spectral abundance factor,

protein, peptide expression alteration (fold changes), log values, and confidence were calculated based on spectral peak intensities generated from the mass spectrometric analysis after extracting confident protein spectra with $P < 0.01$. For TMT analysis, the relative quantification between any experimental groups in the sixplex experiment was derived from the average ratio of the reporter ions of duplicate tags of one group over the average reporter ions of duplicate tags of the corresponding group. Statistical computing and graphics were performed in R software environment and Graphpad Prism 5.

■ RESULTS AND DISCUSSION

The analysis of samples with small amounts of material is challenging. Although the use of capillary chromatography has provided tremendous gains in sensitivity, the use of capillary chromatography also introduces sample-loading challenges. Kennedy and Jorgenson introduced a loading procedure that used a high-pressure device ("bomb") to transfer samples directly from an eppendorf tube into the capillary column.[16] When the bomb is pressurized, liquid is forced into the column from the eppendorf tube. The capability to analyze small amounts of materials allows access to samples such as biopsy samples, small sections of tissue such as brain sections, or even single-cell analysis such as neurons. As demonstrated by Masuda et al. and Thakur et al, as sample size decreases, sample handling and the nature of the chromatographic interface become important for good detection limits on peptides.[26,42] In this study, we tested three methods for introducing complex peptide mixtures into a tandem mass spectrometer. The three multidimensional LC methods consisted of a direct online method using an integrated triphasic capillary column for the introduction of samples and two off-line methods involving collection of samples in
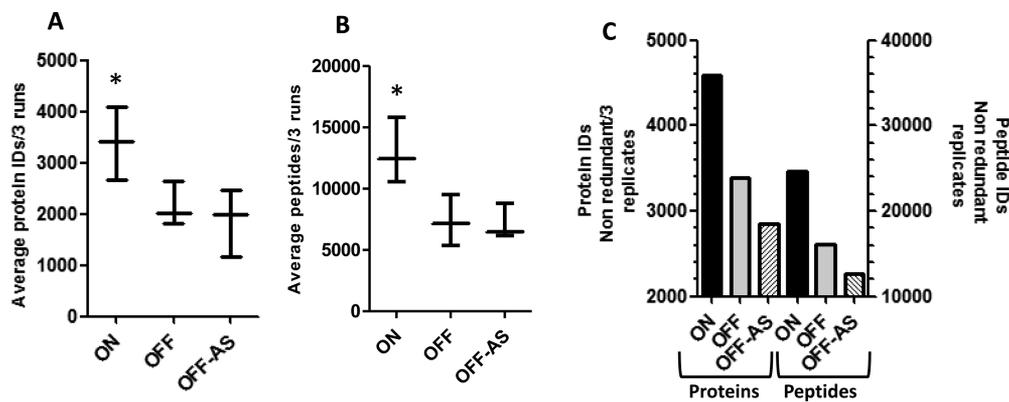
Figure 3. Box and whiskers plots of averaged identified proteins (A) and peptides (B) for different MudPit platforms. Whiskers represent minimum to maximum identification for three replicate runs. (C) Identified nonredundant proteins or peptides were merged. * represents significance at $p < 0.05$.

eppendorf tubes. One of the off-line methods used an autosampler to introduce the samples into the analytical RP column, and the other used the same direct pressure loading system used to load samples for the online MDLC analysis. These three approaches are shown in Figure 1. To be consistent, all approaches used the same type of SCX column to perform the IEX fractionation, which consisted of a biphasic column. In addition, 10 salt steps were used in the online system, and 10 fractions were collected in the two off-line methods using the same elution method. The same analytical RP method was employed. To perform the comparison, a trypsin-digested HEK293 sample was used that should be representative of the type of sample commonly analyzed in proteomics. For each method, the sample was run in triplicate to measure the variance in peptide identifications. Two different quantitation methods were used to measure differences in peptide recoveries between the different methods.

Even with the current advanced mass spectrometric capabilities, multiple dimensions of separation improve comprehensive analysis of peptide mixtures. Although a variety of multidimensional combinations have emerged in the past few years,[5,43−46] the SCX-RP combination is an efficient and highly resolving separation method for shotgun proteome analysis. The main impetus of this work is to provide a systematic comparative study of different LC/LC strategies in terms of performance, sensitivity, and recovery that should be applicable to all LC/LC methods regardless of the phases employed

Evaluation of Protein/Peptide Identification Efficiency

Proteins from a HEK293 cell line were digested with trypsin and aliquots were subjected to analysis by the three different platforms to measure peptide and protein identification efficiency for each method. For this measurement, the goal was simply to report how many peptides and proteins would be detected for the same sample using the different methods (Figure 1). Because of the complexity of the sample, 10 salt steps were used to ensure elution of most peptides from the column. Three replicate analyses of each platform showed an average of 187 366 ± 8545, 168 750 ± 3113, and 140 000 ± 7950 MS2 scans from online, offline, and offline-AS groups, respectively. The online approach generated the most MS2 scans. A total of 3383 ± 386, 2159 ± 243, and 1877 ± 413 proteins were identified from the 12937 ± 1533, 7351 ± 1201, and 7146 ± 829 peptides identified for the online, offline, and offline-AS groups, respectively (Figure 3A,B). The higher mean of the online triphasic column was statistically significant (confidence of 95%) compared with the other methods. Given the identical gradients and MS methods used for each approach, the peptide identification and protein identification

numbers were similar to an improvement by more than 1.7 to 1.8 fold in the online method compared with both offline methods (Figure 3B). Additionally, merging and removing redundant proteins/peptides increased protein/peptide identifications most in the online triphasic column method, less in the offline method, and least in the offline-AS method. (Figure 3C). This comparison showed that recovery of peptides was best in the online method by virtue of the most peptide and protein identifications and the poorest in the off-line fractionation method with the introduction of samples through an autosampler. This result makes sense, as the samples being fractionated offline and introduced through an autosampler are being exposed to more new surfaces and are being subjected to the most manipulations.

Reproducibility and Overlap between Different MudPIT Platforms

A key measure in proteomics is overlap in peptide and protein identifications as a function of technical replicates. Reproducibility between protein identifications is expected to be greater because a protein can be identified by different peptides. A high level of reproducibility in peptide identification is harder to achieve because it requires the system be near saturation. A comparison of reproducibility and overlap between the different systems is shown in Figure 4.
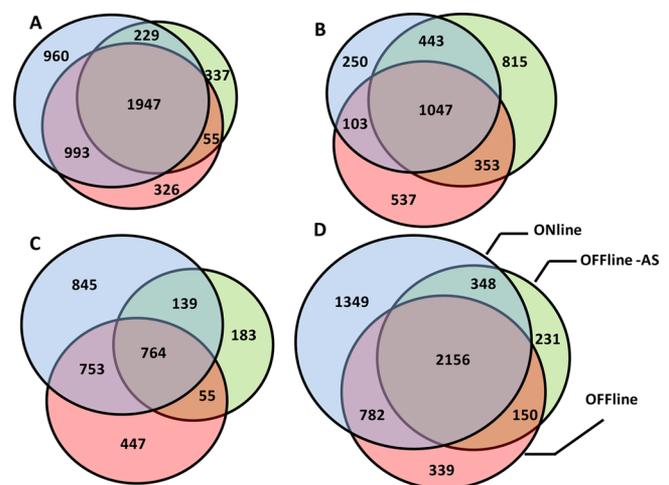
Figure 4. Proportional Venn diagram demonstrating the overlap between protein identification of different MudPit triplicate runs ((A) online, (B) offline; and (C) offline-AS) or between different MudPit platforms (D).
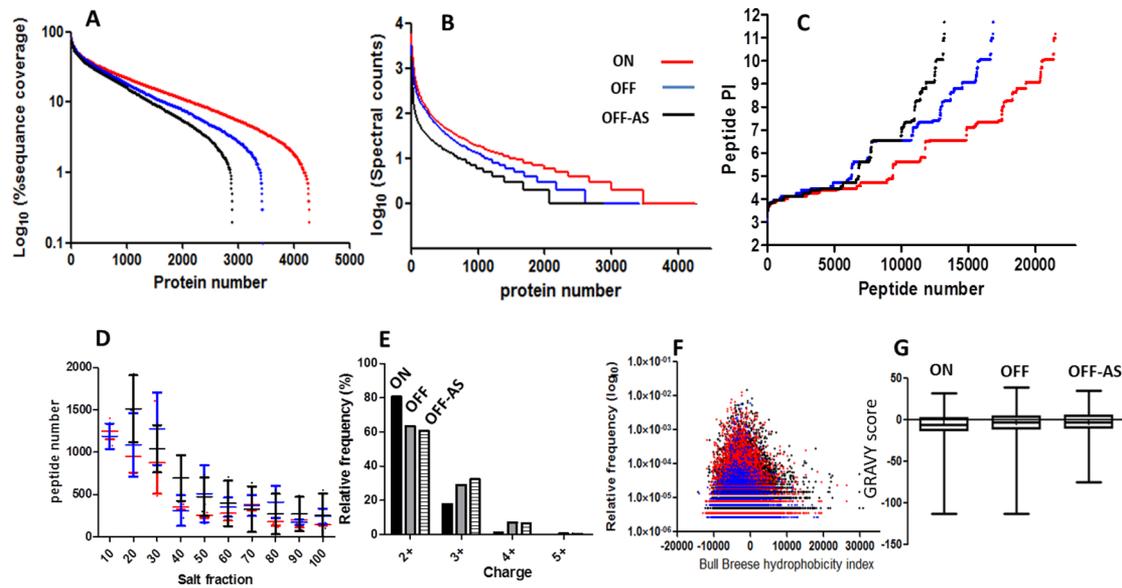
Figure 5. Proteome metrics and peptide physicochemical properties of the experimental MudPit platforms. (A) Percentage of protein sequence coverage. (B) Spectral count rank- abundance of proteins. (C) Peptide isoelectric point (pI). (D) Number of peptides eluted in each salt fraction of the MudPit Platforms. (E) Relative frequency of peptide charge. (F) Bull Breese hydrophobicity index was calculated based on the free energy of transfer to surface in KCl/mol. (G) Kyte−Dolittle hydrophathy scoring (GRAVY score) was calculated based on the average amino acids score for a given protein. Positive score is hydrophilic and negative score is hydrophobic in Bull Breese and vice versa for Kyte−Dolittle. Red is online MudPit, blue is offline MudPit, and black is offline-AS MudPit.
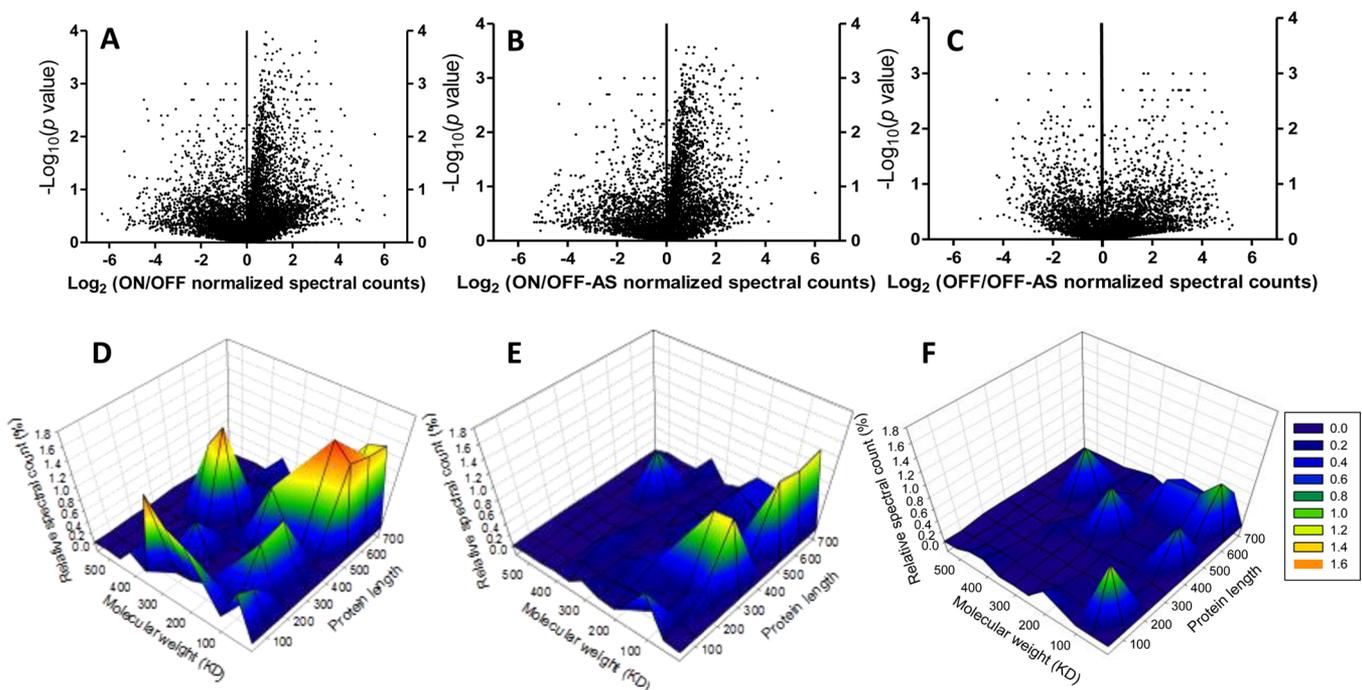
Figure 6. Label-free quantification based on normalized spectral abundance counts of different MudPit panels. (A−C) Volcano plots depict relationship between the P value and the magnitude of difference (Log_2) in expression value between average technical replicates of two compared groups. (D−F) Surface 3D plots of the identified proteins in each experimental group based on their relative spectral counts (%), molecular weight (KD), and protein length.

Reproducibility between runs was greatest among online replicates (60% overlap) and lowest in the offline analysis coupled to the autosampler (45% overlap). In addition, a comparison of the three experimental strategies shows online LC/LC identifies more distinct proteins at 1% FDR compared with the other platforms (Figure 4D). Differences in identification rates among the samples suggest that improved identification rates are a result of minimizing manual handling of

fractions. By using an online system sample, losses are decreased, and this leads to improved recovery of peptides through the system and acquisition of more MS/MS.

Proteome Metrics of MudPIT Formats

If there is an observed difference in peptide identification between the different methods, it begs the question of whether there are any differences in physicochemical characteristics of the
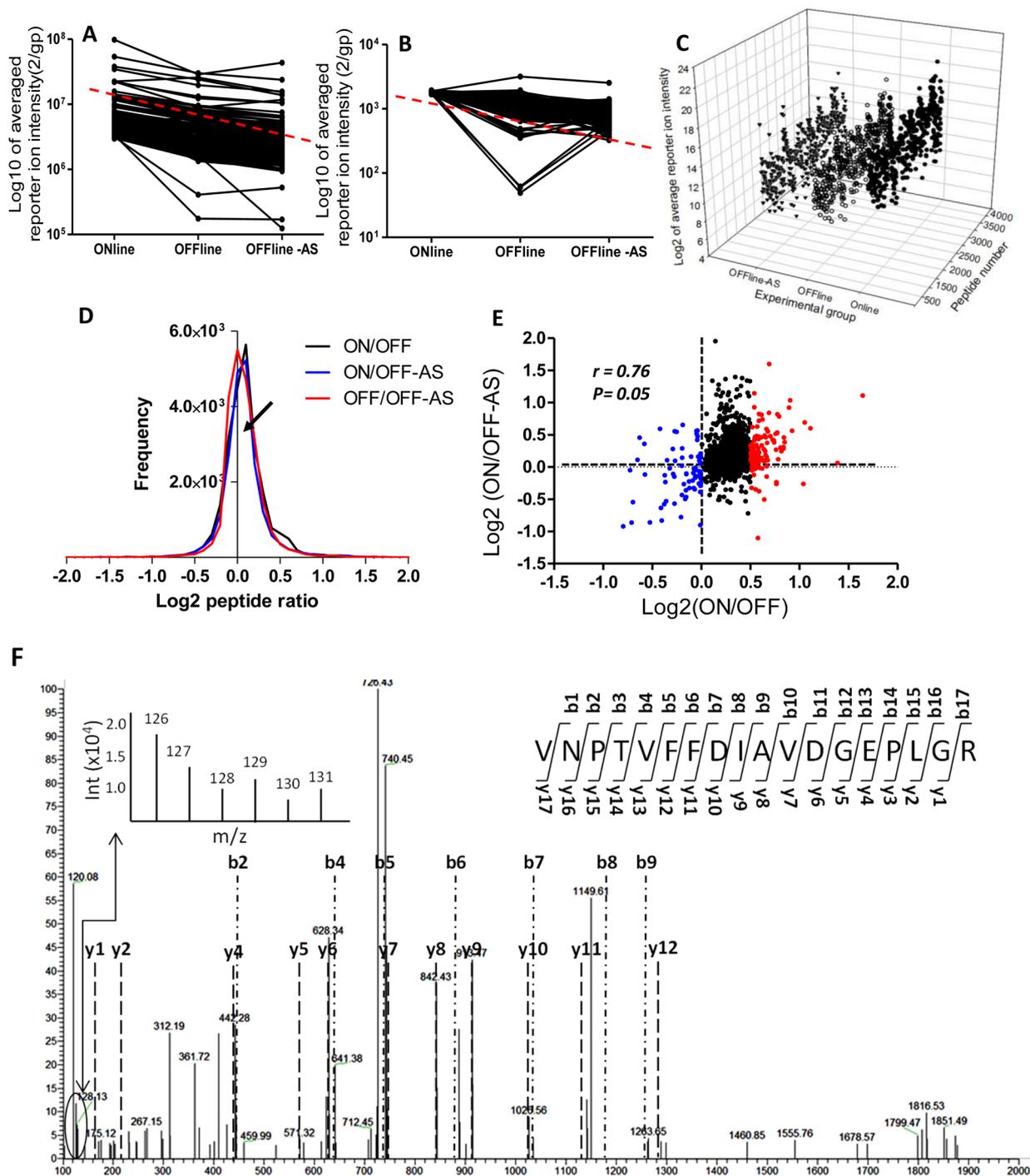
Figure 7. Tandem mass tag (TMT) isobaric quantification of different MudPit panels. Reporter ion intensities for highly abundant (A) and low abundant (B) peptides were plotted with trend line pattern (dashed red line). (C) Perpendicular 3D plot revealed that most of the identified reporter ions were relatively higher in online MudPit with respect to the other offline formats. (D) Frequency histogram of MudPit panels showing distribution of $\log_2$ peptide ratio observed between compared groups. (E) Log–Log correlation plot of protein expression ratio of the online panel over offline panels. Black and red dots represents proteins with higher intensities in online module with 1.4 to 2 fold, respectively. (F) Example spectrum for peptide labeled with TMT isobaric mass tag labeling reagent. The MS/MS fragmentations were used to sequence the peptide. On the basis of the amino acid ladder, the peptide was identified as VNPTVFFDIAVDGEPLGR with the N-terminus modified by TMT isobaric mass tag labeling reagent. This peptide belongs to peptidylprolyl isomerase (PPIA). Mass tags (126−131) observed in the lower m/z region (inserted figure) indicate the relative abundance of this peptide in each group. The samples were labeled in the following order: online MudPIT (126, 127), offline MudPIT (128, 129), and offline MudPIT with autosampler (130, 131).

peptides or proteins observed in the different methods. We analyzed the proteome metrics relevant to each experimental platform to determine if there were any peculiar physicochemical characteristics of peptides observed in one platform and not the other. In general, online MudPIT showed only a modest increase in protein sequence coverage (Figure 5A) over the offline methods. Spectral count rank, for example, abundance of proteins, however, was significantly lower in the offline-AS
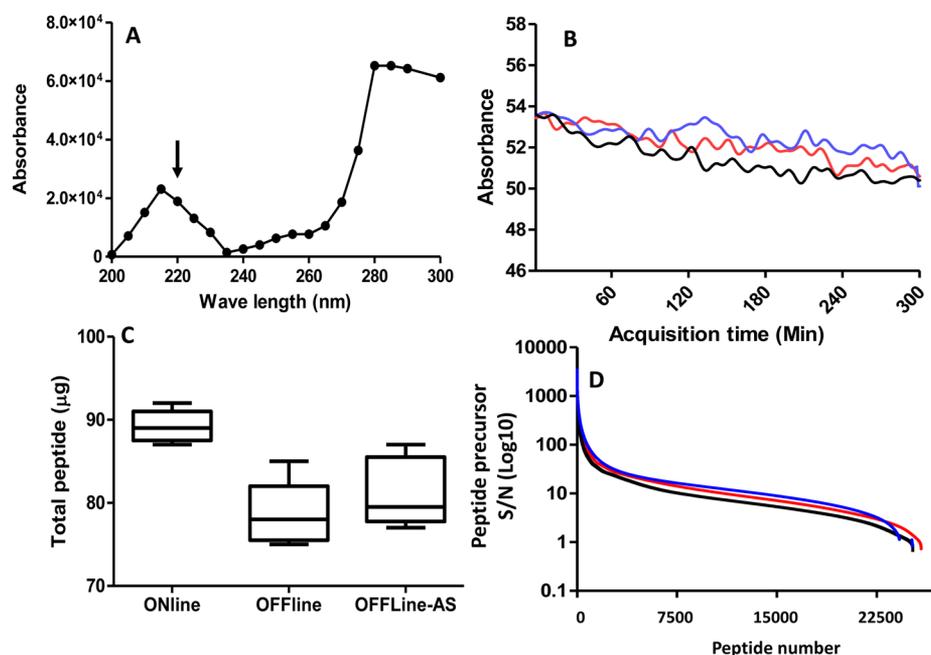
Figure 8. Impact of peptide loss and ion suppression of background noise on MudPit platforms. (A) Peptide mapping was detected more sensitive at wavelength 220 nm. (B) Real-time monitoring of peptide elution using 224 nm laser-induced native fluorescence (LINF) detector. (C) Modified bicinchoninic (BCA) method for peptide fractionated with different Mudpit platforms just before mass spectrometry. (D) Signal-to-noise (S/N) ratio represents the intensity of the signal (peak height) to the intensity of the background noise (in root-mean-square "RMS") for MudPit platforms. Red is online MudPit, blue is offline MudPit, and black is offline-AS MudPit.

platform compared with the other methods (Figure 5B). Because the major difference in the offline-AS group is that the final peptide mixture is placed in autosampler vials, it is likely that the lower spectral counts are due to adsorptive loses of analyte on the surface of the polypropylene sample vials[27,47] or that sample is lost in the flow path of the sample loop.[48] Peptide loss in the offline groups did not significantly correlate with pI, salt fraction, or peptide charge (Figure 5C−E). Although we included organic modifier (5% v/v ACN) in the IEX elution buffer that reportedly reduces surface adsorption,[49] our data suggest a modest loss of hydrophilic proteins in the offline-AS group when plotted using Bull Breese or Kyte−Doolittle scores (Figure 5F,G). This finding is in accordance with recent reports, describing a higher adsorptive tendency of soluble peptides to solid surfaces that ultimately affects peptide amount and quantification parameters.[27,48]

Label-Free Quantification Based on Normalized Spectral Abundance Factor

To further illustrate the changes in protein abundance between groups, we performed a statistical comparison of the average spectral count of triplicate runs from each experimental platform. As illustrated in Figure 6, although several proteins were quantified in high abundance between all platforms, we noticed that online proteins were more abundant and statistically more significant ($P \leq 0.01$), especially when compared with either of the offline groups (offline or offline-AS). This significance was less obvious when comparing between the offline groups (Figure 6C). Specifically, we found that ~635 and ~542 proteins were significantly higher in abundance in the online platform compared with offline and offline-AS groups by 2.2 ± 0.44 and 1.8 ± 0.52 fold, respectively.

Isobaric Tandem Mass Tag Quantification

To verify our findings with an alternate quantitation method, peptides were labeled with different amine-reactive isobaric tags

(Figure 2). The TMT experiment was designed as another way to quantitate the differences between the strategies for LC/LC. A digested HEK293 sample was aliquoted into six aliquots of 25 μg each. Each aliquot was labeled with a different mass tag. Two tagged samples each were used for online separation: two for offline with pressure bomb loading and two for off-line fractionation with autosampling. The experiments were performed as for the label-free experiments except that the outflow from the RP capillary column was collected into a single tube for each sample. The volumes were adjusted; then, all samples were combined into a single tube. The content of this tube was then analyzed by LC−MS/MS. This method ensures exact comigration with simultaneous and accurate peptide quantification using the mass tags that appears in the tandem mass spectrum.[34] We found that peptide abundance was lower ($P \leq 0.05$) in the offline groups (Figure 7A−C). Again, this was not restricted to certain peptides properties; in contrast, the majority of them showed similar trend patterns, denoting stochastic nonspecific loss. In addition, peptide ratios between online and offline groups disclosed a modest skew toward the online platform with a 13−18% increase after peptide grouping and normalization (Figure 7D). Cumulatively, the elevated ion intensity signaling for peptides detected in the automated online method corresponded to an average increase of 18% in protein abundance for more than 1100 proteins (Figure 7E). We noticed that the same proteins were underestimated in the offline groups with a significant correlation coefficient factor ($r = 0.76$, $p = 0.05$) when compared with the online platform (Figure 7E).

Influence of Peptide Loss or Ion Suppression on MudPIT Platforms

To answer the question of whether the lowered peptide intensity in the offline methods is due to peptide loss during processing or ion suppression as a result of high background noise (chemical contaminants, atmospheric sources, or electrical interference),

we utilized the high sensitivity of the LINF detector ($\geq$100 fold than UV detectors)[50] together with peptide quantification using modified BCA method.[29] This allowed us to monitor peptide changes before and after sample processing. Peptide mapping at 220 nm provided a flat baseline with better sensitivity in the current experiment (Figure 8A), and detection of SCX eluted peptides before MudPIT processing did not show any significant differences between experimental groups (Figure 8B) even at 280 nm associated with absorption of the aromatic amino acids (data not shown). Nevertheless, downstream peptide quantification, directly before mass spectrometry analysis, revealed potential low peptide yield ($P \leq 0.05$) in the offline methods (Figure 8C). This could be attributed to the nonspecific adsorption of peptides on solid surfaces. A previous report demonstrated that the adsorption of biomolecules such as peptides followed a Langmuir isotherm equation and was influenced by both solvents and the nature of the solid surfaces.[47,51] Our results support this observation. Moreover, we monitored the possible impact of sparse ion noise background on the desired peptide signal peak by calculating signal-to-noise ratio (S/N; Figure 8D) based on Gygi's method.[52] Although most peptides in our analysis were analyzed at low S/N, the consequences of low signal levels on quantitative accuracy remain to be tested. Interestingly, online MudPIT generated slightly higher background noise compared with offline and offline-AS; the S/N ratio, however, was not significantly different between all platforms (Figure 8D), and both MS and MS/MS quality was comparable. This observation could be attributed to the higher signal ion intensity of the online MudPIT method that maintained a constant S/N ratio between tested groups.

CONCLUSIONS

It has long been known that peptides and proteins can be readily lost to surface adsorption to the materials they come in contact with. This issue was particularly troublesome when trying to purify proteins or peptides to homogeneity for analysis because these methods often required much sample manipulation. Ultimately the sample losses associated with gel purification and in-gel-digestion-limited detection of proteins made these methods less attractive. One advantage of shotgun proteomics is the preparation of proteins en masse for analysis by the mass spectrometer. By preparing samples as a complex mixture, the more abundant proteins can act as carriers of the less abundant proteins. After digestion, the complex mixture of peptides needs to be separated by HPLC for introduction into the mass spectrometer. Several strategies have evolved to fractionate peptide mixtures prior to entry into the mass spectrometer.

We demonstrated that the use of automated online MudPIT results in more comprehensive peptide separation and substantially more protein and peptide identification in a label-free quantification experiment, although results were less striking for TMT quantification, where the nature of the experimental design may have complicated the comparison. Differences attributed to sample loading are alleviated by normalization correction in both experiments, together with comparable MS/MS spectra (similar S/N ratio), so it is likely that stochastic peptide loss due to adsorption could be affecting offline sample collection (such as tubes and vials). Certainly, this conclusion does not discourage using offline platforms because each format has its inherent advantages and disadvantages (i.e., flexibility of offline fraction collection and reduced labor time in online separation). However, because adsorption is a concentration-dependent surface phenomenon, one should critically consider the potential sample loss due to surface adsorption when considering offline fractionation platforms, especially when processing minute quantities of valuable samples.

ASSOCIATED CONTENT

ⓢ Supporting Information

Supplementary 1. Protein and peptide identification for different MudPIT formats. Supplementary 2. Label free quantification based on spectral counting of different MudPIT formats. Supplementary 3. TMT-sixplex quantification for MudPIT. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Authors

*S.M.: E-mail: Samehmagd@med.niigata-u.ac.jp, Samehmagd@yahoo.com. Tel: +81 (25) 227-2152. Fax: +81 (25) 227-0768.
*J.R.Y.: E-mail: jyates@scripps.edu. Tel: +1 (858)784-8862. Fax: +1 (858)784-8883.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) McCormack, A. L.; Schieltz, D. M.; Goode, B.; Yang, S.; Barnes, G.; Drubin, D.; Yates, J. R., III Direct analysis and identification of proteins in mixtures by LC/MS/MS and database searching at the low-femtomole level. *Anal. Chem.* **1997**, *69* (4), 767−776.

(2) Link, A. J.; Eng, J.; Schieltz, D. M.; Carmack, E.; Mize, G. J.; Morris, D. R.; Garvik, B. M.; Yates, J. R., 3rd Direct analysis of protein complexes using mass spectrometry. *Nat. Biotechnol.* **1999**, *17* (7), 676−682.

(3) Washburn, M. P.; Wolters, D.; Yates, J. R., 3rd. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* **2001**, *19* (3), 242−247.

(4) Giddings, J. C. Two-dimensional separations: concept and promise. *Anal. Chem.* **1984**, *56* (12), 1258A−1260A.

(5) Motoyama, A.; Yates, J. R., 3rd. Multidimensional LC separations in shotgun proteomics. *Anal. Chem.* **2008**, *80* (19), 7187−7193.

(6) Anderegg, R. J.; Wagner, D. S.; Blackburn, R. K.; Opiteck, G. J.; Jorgenson, J. W. A multidimensional approach to protein characterization. *J. Protein Chem.* **1997**, *16* (5), 523−526.

(7) Bushey, M. M.; Jorgenson, J. W. Automated instrumentation for comprehensive 2-dimensional high-performance liquid-chromatography capillary zone electrophoresis. *Anal. Chem.* **1990**, *62*, 978−984.

(8) Larmann, J. P.; Lemmo, A. V.; Moore, A. W.; Jorgenson, J. W. Two-dimensional separations of peptides and proteins by comprehensive liquid chromatography-capillary electrophoresis. *Electrophoresis* **1993**, *14* (5−6), 439−447.

(9) Moore, A. W.; Jorgenson, J. W. Comprehensive three-dimensional separation of peptides using size exclusion chromatography/reversed

phase liquid chromatography/optically gated capillary zone electrophoresis. *Anal. Chem.* **1995**, *67* (19), 3456−3463.

(10) Moore, A. W.; Larmann, J. P.; Lemmo, A. V.; Jorgenson, J. W. Two-dimensional liquid chromatography-capillary electrophoresis techniques for analysis of proteins and peptides. *Methods Enzymol.* **1996**, *270*, 401−419.

(11) Opiteck, G. J.; Jorgenson, J. W. Two-dimensional SEC/RPLC coupled to mass spectrometry for the analysis of peptides. *Anal. Chem.* **1997**, *69* (13), 2283−2291.

(12) Opiteck, G. J.; Ramirez, S. M.; Jorgenson, J. W.; Moseley, M. A. Comprehensive two-dimensional high-performance liquid chromatography for the isolation of overexpressed proteins and proteome mapping. *Anal. Biochem.* **1998**, *258* (2), 349−361.

(13) Qiu, R.; Regnier, F. E. Use of multidimensional lectin affinity chromatography in differential glycoproteomics. *Anal. Chem.* **2005**, *77* (9), 2802−2809.

(14) Slentz, B. E.; Penner, N. A.; Regnier, F. E. Protein proteolysis and the multi-dimensional electrochromatographic separation of histidine-containing peptide fragments on a chip. *J. Chromatogr. A* **2003**, *984* (1), 97−107.

(15) Hsieh, Y. F.; Gordon, N.; Regnier, F.; Afeyan, N.; Martin, S. A.; Vella, G. J. Multidimensional chromatography coupled with mass spectrometry for target-based screening. *Mol. Diversity* **1997**, *2* (4), 189−196.

(16) Kennedy, R. T.; Jorgenson, J. W. Preparation and evaluation of packed capillary liquid chromatography columns with inner diameters from 20 to 50 microns. *Anal. Chem.* **1989**, *61*, 1128−1135.

(17) Gatlin, C. L.; Kleemann, G. R.; Hays, L. G.; Link, A. J.; Yates, J. R., 3rd. Protein identification at the low femtomole level from silver-stained gels using a new fritless electrospray interface for liquid chromatography-microspray and nanospray mass spectrometry. *Anal. Biochem.* **1998**, *263* (1), 93−101.

(18) McDonald, W. H.; Ohi, R.; Miyamoto, D. T.; Mitchison, T. J.; J. R. Yates, I. Comparison of three directly coupled HPLC MS/MS strategies for identification of proteins from complex mixtures: single dimension LC/MS/MS, 2-phase MudPIT, and 3-phase MudPIT. *Int. J. Mass Spectrom.* **2002**, *219*, 245−251.

(19) Wolters, D. A.; Washburn, M. P.; Yates, J. R., 3rd. An automated multidimensional protein identification technology for shotgun proteomics. *Anal. Chem.* **2001**, *73* (23), 5683−5690.

(20) Peng, J.; Elias, J. E.; Thoreen, C. C.; Licklider, L. J.; Gygi, S. P. Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. *J. Proteome Res.* **2003**, *2* (1), 43−50.

(21) Zhou, F.; Sikorski, T. W.; Ficarro, S. B.; Webber, J. T.; Marto, J. A. Online nanoflow reversed phase-strong anion exchange-reversed phase liquid chromatography-tandem mass spectrometry platform for efficient and in-depth proteome sequence analysis of complex organisms. *Anal. Chem.* **2011**, *83* (18), 6996−7005.

(22) Dwivedi, R. C.; Spicer, V.; Harder, M.; Antonovici, M.; Ens, W.; Standing, K. G.; Wilkins, J. A.; Krokhin, O. V. Practical implementation of 2D HPLC scheme with accurate peptide retention prediction in both dimensions for high-throughput bottom-up proteomics. *Anal. Chem.* **2008**, *80* (18), 7036−7042.

(23) Wang, Y.; Yang, F.; Fu, Y.; Huang, X.; Wang, W.; Jiang, X.; Gritsenko, M. A.; Zhao, R.; Monore, M. E.; Pertz, O. C.; Purvine, S. O.; Orton, D. J.; Jacobs, J. M.; Camp, D. G., 2nd; Smith, R. D.; Klemke, R. L. Spatial phosphoprotein profiling reveals a compartmentalized extracellular signal-regulated kinase switch governing neurite growth and retraction. *J. Biol. Chem.* **2011**, *286* (20), 18190−18201.

(24) Hunt, D. F.; Henderson, R. A.; Shabanowitz, J.; Sakaguchi, K.; Michel, H.; Sevilir, N.; Cox, A. L.; Appella, E.; Engelhard, V. H. Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. [see comments]. *Science* **1992**, *255* (5049), 1261−1263.

(25) Merrell, K.; Southwick, K.; Graves, S. W.; Esplin, M. S.; Lewis, N. E.; Thulin, C. D. Analysis of low-abundance, low-molecular-weight serum proteins using mass spectrometry. *J. Biomol. Tech.* **2004**, *15* (4), 238−248.

(26) Masuda, T.; Sugiyama, N.; Tomita, M.; Ishihama, Y. Microscale phosphoproteome analysis of 10,000 cells from human cancer cell lines. *Anal. Chem.* **2011**, *83* (20), 7698−7703.

(27) Stejskal, K.; Potesil, D.; Zdrahal, Z. Suppression of peptide sample losses in autosampler vials. *J. Proteome Res.* **2013**, *12* (6), 3057−3062.

(28) Thakur, S. S.; Geiger, T.; Chatterjee, B.; Bandilla, P.; Frohlich, F.; Cox, J.; Mann, M. Deep and highly sensitive proteome coverage by LC-MS/MS without prefractionation. *Mol. Cell. Proteomics* **2011**, *10* (8), M110003699.

(29) Kapoor, K. N.; Barry, D. T.; Rees, R. C.; Dodi, I. A.; McArdle, S. E.; Creaser, C. S.; Bonner, P. L. Estimation of peptide concentration by a modified bicinchoninic acid assay. *Anal. Biochem.* **2009**, *393* (1), 138−140.

(30) Fonslow, B. R.; Stein, B. D.; Webb, K. J.; Xu, T.; Choi, J.; Park, S. K.; Yates, J. R., 3rd. Digestion and depletion of abundant proteins improves proteomic coverage. *Nat. Methods* **2013**, *10* (1), 54−56.

(31) Magdeldin, S.; Yoshida, Y.; Li, H.; Maeda, Y.; Yokoyama, M.; Enany, S.; Zhang, Y.; Xu, B.; Fujinaka, H.; Yaoita, E.; Sasaki, S.; Yamamoto, T. Murine colon proteome and characterization of the protein pathways. *BioData Min.* **2012**, *5* (1), 11.

(32) Guzzetta, A. W.; Chien, A. S. A double-vented tetraphasic continuous column approach to MuDPIT analysis on long capillary columns demonstrates superior proteomic coverage. *J. Proteome Res.* **2005**, *4* (6), 2412−2419.

(33) Schieltz, D. M.; Washburn, M. P. Analysis of Complex Protein Mixtures Using Multidimensional Protein Identification Technology (MuDPIT). *CSH Protoc.* **2006**, DOI: 10.1101/pdb.prot4555.

(34) Thompson, A.; Schafer, J.; Kuhn, K.; Kienle, S.; Schwarz, J.; Schmidt, G.; Neumann, T.; Johnstone, R.; Mohammed, A. K.; Hamon, C. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal. Chem.* **2003**, *75* (8), 1895−1904.

(35) Lohrig, K.; Wolters, D. Multidimensional protein identification technology. *Methods Mol. Biol.* **2009**, *564*, 143−153.

(36) Bonnin, C.; Matoga, M.; Garnier, N.; Debroche, C.; de Vandiere, B.; Chaminade, P. 224 nm Deep-UV laser for native fluorescence, a new opportunity for biomolecules detection. *J. Chromatogr. A* **2007**, *1156* (1−2), 94−100.

(37) Rauniyar, N.; Gao, B.; McClatchy, D. B.; Yates, J. R., 3rd. Comparison of protein expression ratios observed by sixplex and duplex TMT labeling method. *J. Proteome Res.* **2013**, *12* (2), 1031−1039.

(38) McDonald, W. H.; Tabb, D. L.; Sadygov, R. G.; MacCoss, M. J.; Venable, J.; Graumann, J.; Johnson, J. R.; Cociorva, D.; Yates, J. R., 3rd. MS1, MS2, and SQT-three unified, compact, and easily parsed file formats for the storage of shotgun proteomic spectra and identifications. *Rapid Commun. Mass Spectrom.* **2004**, *18* (18), 2162−2168.

(39) Xu, T.; Wong, C. C.; Kashina, A.; Yates, J. R., 3rd. Identification of N-terminally arginylated proteins and peptides by mass spectrometry. *Nat. Protoc* **2009**, *4* (3), 325−332.

(40) Tabb, D. L.; McDonald, W. H.; Yates, J. R., 3rd. DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. *J. Proteome Res.* **2002**, *1* (1), 21−26.

(41) Park, S. K.; Venable, J. D.; Xu, T.; Yates, J. R., 3rd. A quantitative analysis software tool for mass spectrometry-based proteomics. *Nat. Methods* **2008**, *5* (4), 319−322.

(42) Thakur, D.; Rejtar, T.; Wang, D.; Bones, J.; Cha, S.; Clodfelder-Miller, B.; Richardson, E.; Binns, S.; Dahiya, S.; Sgroi, D.; Karger, B. L. Microproteomic analysis of 10,000 laser captured microdissected breast tumor cells using short-range sodium dodecyl sulfate-polyacrylamide gel electrophoresis and porous layer open tubular liquid chromatography tandem mass spectrometry. *J. Chromatogr. A* **2011**, *1218* (45), 8168−8174.

(43) Mawuenyega, K. G.; Kaji, H.; Yamuchi, Y.; Shinkawa, T.; Saito, H.; Taoka, M.; Takahashi, N.; Isobe, T. Large-scale identification of Caenorhabditis elegans proteins by multidimensional liquid chromatography-tandem mass spectrometry. *J. Proteome Res.* **2003**, *2* (1), 23−35.

(44) Magdeldin, S.; Li, H.; Yoshida, Y.; Enany, S.; Zhang, Y.; Xu, B.; Fujinaka, H.; Yaoita, E.; Yamamoto, T. Comparison of two dimensional

electrophoresis mouse colon proteomes before and after knocking out Aquaporin 8. *J. Proteomics* **2010**, *73* (10), 2031−2040.

(45) Wang, H. L.; Shi, L. Z.; Wong, C. C. L.; Han, X. M.; Hwang, P. Y. H.; Truong, L. N.; Zhu, Q. Y.; Shao, Z. P.; Chen, D. J.; Berns, M. W.; Yates, J. R.; Chen, L. C.; Wu, X. H. The Interaction of CtIP and Nbs1 Connects CDK and ATM to Regulate HR-Mediated Double-Strand Break Repair. *PLoS Genet.* **2013**, *9* (2), e1003277.

(46) Koda, R.; Zhao, L.; Yaoita, E.; Yoshida, Y.; Tsukita, S.; Tamura, A.; Nameta, M.; Zhang, Y.; Fujinaka, H.; Magdeldin, S.; Xu, B.; Narita, I.; Yamamoto, T. Novel expression of claudin-5 in glomerular podocytes. *Cell Tissue Res.* **2011**, *343* (3), 637−648.

(47) van Midwoud, P. M.; Rieux, L.; Bischoff, R.; Verpoorte, E.; Niederlander, H. A. Improvement of recovery and repeatability in liquid chromatography-mass spectrometry analysis of peptides. *J. Proteome Res.* **2007**, *6* (2), 781−791.

(48) Warwood, S.; Byron, A.; Humphries, M. J.; Knight, D. The effect of peptide adsorption on signal linearity and a simple approach to improve reliability of quantification. *J. Proteomics* **2013**, *85*, 160−164.

(49) John, H.; Walden, M.; Schafer, S.; Genz, S.; Forssmann, W. G. Analytical procedures for quantification of peptides in pharmaceutical research by liquid chromatography-mass spectrometry. *Anal. Bioanal. Chem.* **2004**, *378* (4), 883−897.

(50) Chan, K. C.; Veenstra, T. D.; Issaq, H. J. Comparison of fluorescence, laser-induced fluorescence, and ultraviolet absorbance detection for measuring HPLC fractionated protein/peptide mixtures. *Anal. Chem.* **2011**, *83* (6), 2394−2396.

(51) Pezeshki, A.; Vergote, V.; Van Dorpe, S.; Baert, B.; Burvenich, C.; Popkov, A.; De Spiegeleer, B. Adsorption of peptides at the sample drying step: influence of solvent evaporation technique, vial material and solution additive. *J. Pharm. Biomed. Anal.* **2009**, *49* (3), 607−612.

(52) Bakalarski, C. E.; Elias, J. E.; Villen, J.; Haas, W.; Gerber, S. A.; Everley, P. A.; Gygi, S. P. The impact of peptide abundance and dynamic range on stable-isotope-based quantitative proteomic analyses. *J. Proteome Res.* **2008**, *7* (11), 4756−4765.