Article

In-Depth Comparison of Reagent-Based Digestion Methods and Two Commercially Available Kits for Bottom-Up Proteomics

Travis Uslan,[§] Baiyi Quan,[§] Ting-Yu Wang, Marion Pang, Yanping Qiu, and Tsui-Fen Chou*
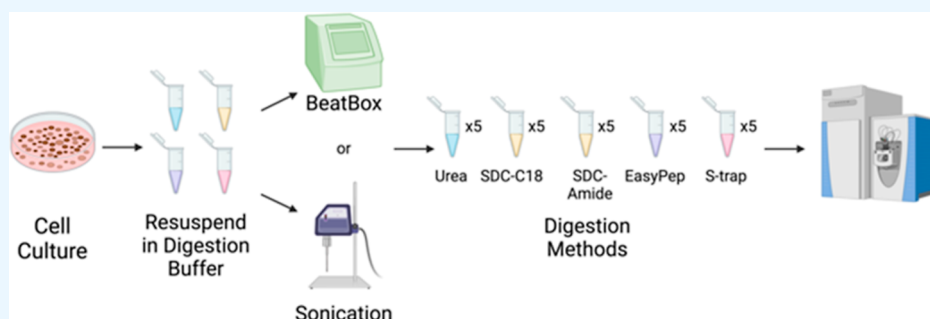
Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Proteomic analysis plays an essential role in biology with several methodologies available for sample preparation and analysis. This study evaluates and compares various cell lysis and protein digestion protocols for bottom-up proteomics using HeLa S3 cells. We assessed two physical disruption methods to homogenize cells—sonication and BeatBox—alongside four digestion protocols. Two of them are lab-reagent strategies: urea-based and sodium deoxycholate (SDC)-based in-solution digestion, and two are commercially available kits: the EasyPep kit from Thermo Fisher Scientific and S-Trap from Protifi. Each method's efficacy was evaluated based on protein recovery, peptide yield, and number of unique proteins identified through LC−MS analysis. Our results indicate that while both sonication and the BeatBox (PreOmics Inc.) methods provided comparable protein recovery and coverage, the choice of digestion method had a much bigger impact on the amount of protein IDs found. SDC digestion yielded the highest protein and peptide counts, while S-Trap exhibited the most consistent peptide recovery. Conversely, EasyPep showed higher variability in peptide recovery, with a ±10% difference in the average peptide number. Each homogenization strategy and digestion method also yielded its own list of unique proteins. These results provide several lists of proteins for biologists to select from based on experimental needs and highlight the importance of choosing appropriate protocols for comprehensive proteomic analyses.

INTRODUCTION

With the efforts to understand the proteome increasing since the human genome project,[1,2] progress in the field has been driven via the creation of different protocols utilizing liquid chromatography-tandem mass spectrometry (LC−MS/MS). Protein structures and functions vary expansively due to the multitude of duties required of them. Combining this with a wide variance of biological questions and sample types, this can create an intimidating and complex question that demands a range of tailored approaches to meet the specific qualifications of each experimental setup. Proteomics can be a complicated and multistep affair by nature, with any variation during affecting the final data collected. It is important to have a robust, easy-to-follow, and reproducible protocol. Many different preparation methods have been developed throughout the years. More classical methods utilize widely used denaturants such as urea and guanidine hydrochloride or detergents such as sodium dodecyl sulfate (SDS) and SDC. These methods use denaturants to unfold proteins prior to digestion. Newly developed methods include commercially

available kits which aim to make the sample preparation process as quick and easy as possible, offering readily made buffers, protease mixes, and cleanup devices all in one package, such as Thermo Fisher's EasyPep MS Sample Prep Kit.

This immense number of protocols from which to select can slow down the selection of a suitable method from project to project. Previous studies have been performed to compare the output of different digestion methods.[3−7] While these papers can serve as a general guide for people doing bottom-up proteomics for the first time, our aim was to provide a more in-depth look at the specifics of some of the more popular methods. Our goal was to determine what unique proteins, if any, were identified in each digestion method. The results of

our study were used to make lists of both unique proteins identified and enriched and depleted protein abundances across the methods in our study. With the introduction of new homogenization methods such as BeatBox, additional desalting columns, and added information to improve protease efficiency, we used HeLa S3 cells to evaluate two different lysis methods and four digestion methods. For the two cell lysis conditions, we selected sonication, a common method in many laboratories,[4,8] and BeatBox, a recently emerged tissue homogenizer from PreOmics Inc., which utilizes small magnetic beads to break down biological samples through a high speed motion. For the protein digest, two reagent-based methods (urea and SDC) and two commercially available methods (Thermo's EasyPep and Protifi's S-Trap) were used.[8,9,10] For reagent-based methods, we used GL Sciences MonoSpin C18 for desalting.[11] For SDC samples, an additional column, GL Sciences MonoSpin amide, was selected to compare with the C18 column. While C18 columns use an octadecyl functional group to bind peptides, the amide column uses amide, which has a high affinity for various hydrophilic acidic and basic compounds. The EasyPep kit comes with its own desalting column, and S-Trap uses suspension trapping and several wash steps to remove the need for desalting most samples.[8] Our results serve as an in-depth comparison on the differences between four different proteomics protocols and can assist in selecting the appropriate method for sample preparation using the lists of unique or enriched/depleted proteins provided.

■ MATERIALS AND METHODS

Cell Culture. HeLa S3 Cells (ATCC no. CCL-2.2) were cultured in DMEM (Sigma-Aldrich) with 10% FBS (Sigma, Cat no. F0926), 1% penicillin−streptomycin (Fisher, Cat no. 15140122) medium, cultured at 37 °C in a 5% CO_2 incubator. Cells were collected at ∼80% confluency with trypsin, washed with PBS, and pelleted down via centrifugation, before being frozen. Cells were stored at −80 °C in 10×10^6 pellets until use. The routine mycoplasma test was performed to make sure there was no contamination.

Sample Preparation. Each protocol had five replicates, all from the same cell pellet, for statistical analysis purposes. 100 μg of protein was used for each replicate during digestion, and 2 μg of peptide was used for LC−MS/MS analysis.

In-Solution Protocols. *Cell Lysate Preparation.* A total of four HeLa cell pellets were thawed to room temperature and designated for one of four digestion methods: urea, SDC, EasyPep, and S-Trap. Each pellet was resuspended in 1 mL of digestion buffer specific to that digestion method: urea buffer [8 M urea, 100 mM Tris−HCl, pH 8.5, (Sigma-Aldrich)], SDC buffer (1% SDC, 100 mM Tris−HCl, pH 8.5), EasyPep buffer (EasyPep kit lysis buffer), and S-Trap buffer [50 mM triethylammonium bromide (TEAB) (Thermo Fischer)/5% SDS (Sigma-Aldrich)]. 5 μL of Universal Nuclease (Thermo Scientific, San Jose, CA) was added, and cells were gently pipetted up and down until the entire pellet had been homogenized. The cell lysate was then split into 2 × 500 μL aliquots, one labeled for sonication, the other labeled for BeatBox.

Sonication. Sonication was performed by using a Branson sonicator in the pulse mode. The power was set to 25%, and each pulse was to last 5 s with 10 s between pulses. Samples were sonicated for 10 cycles on ice. Samples were gently vortexed and centrifuged for 10 min at 13,000g.

BeatBox. For BeatBox, samples were placed into a 2 mL BeatBox tissue kit 24× tube (P.O. 00128) and loaded onto the BeatBox. Samples were run on high for 10 min twice, after which they were removed and centrifuged for 10 min at 13,000g.

Pierce BCA Assay. Protein concentration was determined using the Thermo Fisher BCA assay (Cat no. 23225).

Urea and SDC. Each sample was then divided into 10 aliquots (5 urea sonication, 5 urea BeatBox) and 20 aliquots (10 SDC sonication, 10 SDC BeatBox) of 100 μg protein. ∼90 μL of denaturation buffer was added to each aliquot to bring the total volume up to ∼110 μL. Protein was reduced with 1 μL of 500 mM tris(2-carboxyethyl)phosphine (TCEP) for 20 min at 37 °C on a Thermo shaker, shaking at 750 rpm. After, samples were allowed to cool to room temperature before being alkylated with 3 μL of 500 mM chloroacetamide (CAA) in the dark for 15 min. 375 μL of 50 mM 2-(4-(2-hydroxyethyl)piperazin-1-yl)ethanesulfonic acid was added to dilute urea samples to a final concentration to 2 M. 5 μL portion of 100 mM $CaCl_2$ was added to assist in trypsin digestion. 16.67 μL (ratio of 1 μg of trypsin to 30 μg of protein) of Thermo trypsin/Lys-C protease mix (0.2 μg/μL) (Cat no. 1863467) was added to each aliquot and digested overnight. Digestion occurred on a Thermo Mixer at 37 °C, shaking at 750 rpm while covered with tin foil. Digestion was stopped the next day with 20 μL of 20% trifluoroacetic acid (TFA). Acidifying the solution caused SDC samples to precipitate. These samples were centrifuged at 13,000g for 10 min, and the supernatant was collected for desalting.

Urea samples were desalted using GL Science's MonoSpin C18 desalting columns and eluted with 70% acetonitrile (ACN), 0.2% formic acid (FA). The SDC samples were then divided into two sets of 10 samples (five sonication samples and 5 BeatBox samples). Set 1 was desalted using GL Science's MonoSpin C18 desalting columns and eluted with 70% ACN, 0.2% FA, while set 2 was desalted using GL Science's MonoSpin amide desalting columns and eluted with 10% ACN, 0.1% ammonia.

All samples were dried down using a Labconco refrigerated Centrivap concentrator and cold trap SpeedVac.

Commercial Kit Protocols. *EasyPep.* After protein concentration was determined, each sample was then divided into 10 aliquots of 100 μg of protein (five sonications and five BeatBox). EasyPep kit lysis buffer was added to each aliquot to make the final volume 100 μL. 50 μL portion of kit reduction solution and 50 μL of kit alkylation solution were added to each aliquot and then incubated at 95 °C in a heat block for 10 min. After incubation, samples were removed from the heat block and allowed to cool to room temperature. 50 μL of Thermo trypsin/Lys-C protease mix (0.2 μg/μL) was added to each aliquot and digested overnight. Digestion occurred on a Thermo Shaker at 37 °C, 750 rpm. Digestion was stopped with 50 μL of STOP digestion solution.

Peptide cleanup was performed on the columns provided with the kit. Column storage liquid was spun out, and each sample was loaded onto a dry desalting column. Samples were loaded 2 times. A single wash step using wash solution A and two wash steps using wash solution B were performed. Peptides were eluted with 300 μL of elution solution and dried in a SpeedVac.

S-Trap. After protein concentration was determined, each sample was then divided into 10 aliquots of 100 μg of protein (5 sonication samples and 5 BeatBox samples). Aliquots were
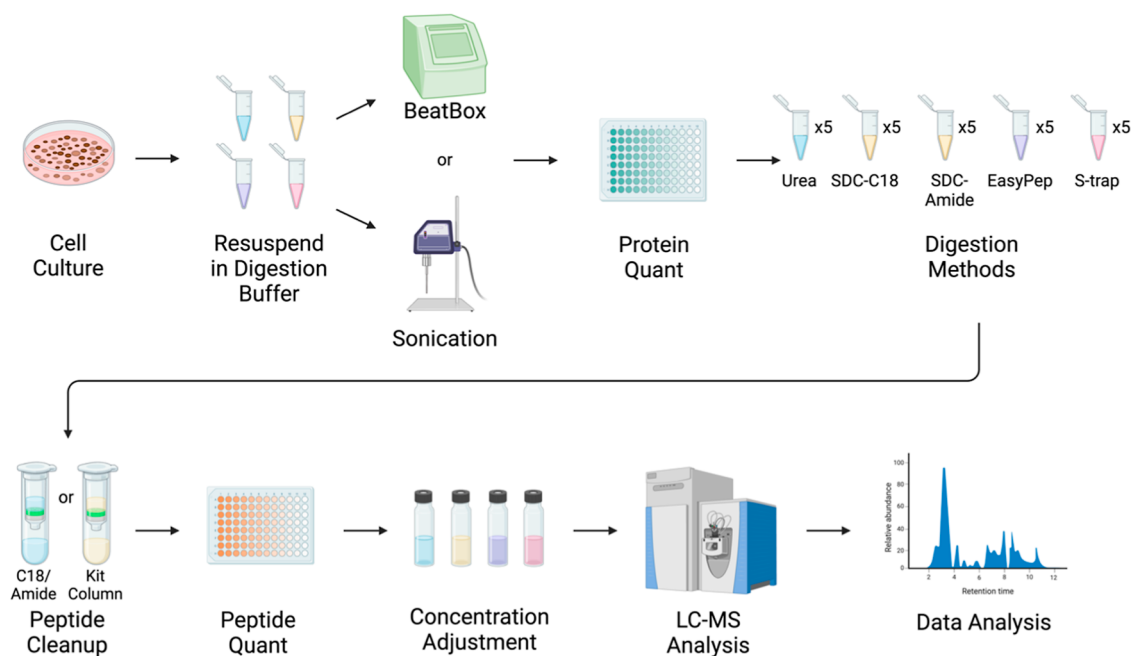
Figure 1. Experimental overview and project workflow. (A) Workflow for the experimental procedure. HeLa cells were cultured, and later resuspended in their respective digestion buffer prior to cell lysis, as indicated by sample preparation methods. Two homogenization strategies were selected for the experiment: sonication and BeatBox. Post homogenization, protein concentration for each sample is measured and normalized. Created in BioRender. Chou, T. (2025) https://BioRender.com/s94q110.

briefly placed into a SpeedVac to bring the total volume of the samples to ~25 μL. They were then reduced by incubation with 1 μL of 120 mM TCEP at 55 °C for 15 min. Proteins were alkylated with 1 μL of 500 mM CAA at room temperature in the dark for 10 min. Samples were then acidified with 5 μL of 27.5% phosphoric acid. 180 μL of S-Trap binding/washing buffer (100 mM TEAB, 90% MeOH) was added to each sample, and all contents pipetted onto a dry S-Trap. S-Traps were then centrifuged for 1 min at 4000g. Samples were washed with 150 μL of binding/washing buffer and then centrifuged for 1 min at 4000g. This was repeated three times. After the third wash, 20 μL of digestion buffer (50 mM TEAB) containing a total of 10 μg of trypsin/Lys-C mix was added, and the entire trap was placed into an incubator humidified with water bath at 37 °C overnight.

Samples were eluted with three consecutive wash steps: 40 μL of 50 mM TEAB, followed by 40 μL of 0.2% FA, and finally 40 μL of 50% ACN, 0.2% FA. In between each step was a 1 min centrifugation at 4000g. Samples were then dried in a SpeedVac.

LC−MS/MS Analysis Method. After drying down, each sample was resuspended in 50 μL of 2% ACN, 0.2% FA and sonicated for 5 min. They were then vortexed and centrifuged at 13,000g for 5 min. The peptide concentration was determined using the fluorescent peptide concentration kit (Thermo Fisher Catalog no. 23290) and diluted to 1 μg/μL and 5 μL was placed into a sample vial (Thermo Fisher). 2 μg of protein was loaded per run on a 2 h gradient.

LC−MS/MS analysis was performed by using a Thermo Q-Exactive Orbitrap attached to a Thermo Easy-nLC system. An Aurora Frontier 60 cm column (Cat no. AUR3-60075C18-TS) was used for peptide separation attached to a Thermo Easy-spray source maintained at 55 °C. Mobile phase A was 2% ACN, 0.2% FA, and mobile phase B was 80% ACN, 0.2% FA. For peptide chromatography, the gradient was increased from

6% B to 25% B to 97 min, then to 40% until 120 min; after-which a cleaning step is included, fluctuating between 98% B and 2% B to clean the column between samples (with a flow rate of 220 nL/min). Scan sequence began with an Orbitrap MS1 spectrum using the following parameters: resolution 60,000; scan range 350 to 1600 m/z; AGC set to 3 × 10^6; maximum injection time (IT) set to 15 ms; for dd-MS^2, the resolution was set to 30,000; AGC to 1 × 10^5; maximum IT set to 45 ms; and isolation window set to 1.2 m/z. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the data set identifier PXD056409.[12]

Proteome Discoverer Methods. The raw MS files were analyzed using Proteome Discoverer 2.5 (Thermo Fisher). MS^2 spectra were searched against human proteome archived by UniProt Knowledge Base[13] (archived 20221128). Enzyme was set to trypsin, with results filtered for high-confidence peptides with enhanced peptide and protein annotations. Quantitative abundances were normalized to the same total peptide amount per channel and scaled so that the average abundance per protein was 100. For precursor quantification settings, peptides to use were set to Unique + Razor, considering protein groups for peptide uniqueness were set to true along with using shared Quan results. Quan results with missing channels were not rejected in the precursor methods but were filtered out using in-house Python scripts later. Precursor quantification abundance was based on intensity with scaling for all average abundances. Normalization and protein roll-up were carried out using all peptides. No peptide modifications were excluded from the protein quantification. Maximum number of missed cleavages was set to 2, and the peptide length range was set from 6 to 144. FDR for peptide and protein discovery was set to 0.05. Precursor mass tolerance was set to 10 ppm with fragment mass tolerance set to 0.02 Da. Average precursor and fragment mass tolerances were not
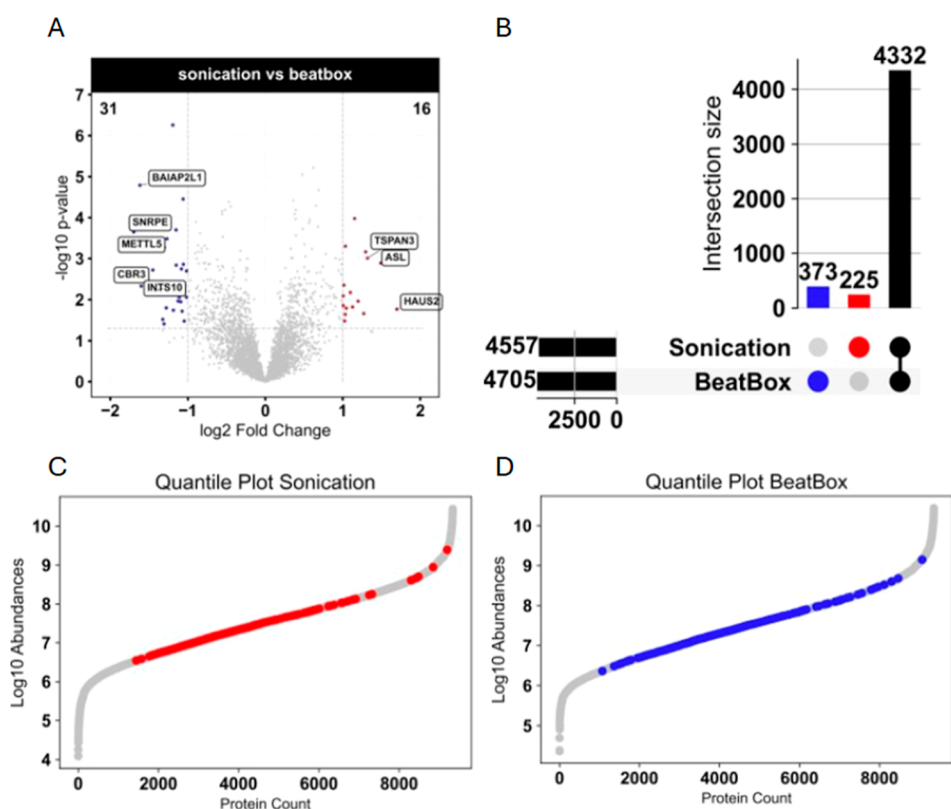
Figure 2. Comparison of homogenization strategies. (A) Volcano plot of $\log_2$ abundance values for 4332 proteins shared by sonication and BeatBox methods. (B) Upset plot of unique proteins found only in sonication or BeatBox. (C) Quantile plot of all abundances in sonication samples, with unique proteins colored red. (D) Quantile plot of all abundances in BeatBox samples, with unique proteins colored blue.

used. Oxidation (+15.995 Da), acetyl (+42.011 Da), Met-loss (−131.040 Da), and Met-loss + acetyl (−89.030 Da) were set as dynamic modifications, while carbamidomethyl (+57.021 Da) was set as a fixed modification.

FragPipe Analysis.[14] To check for protein modifications, FragPipe's open search function was used. First, all raw files were converted to mzML files using[14] MS Convert[16] with write index, use zlib compression, and TPP compatibility. Peak picking and zero samples were also included as filter parameters. The default open search parameters were used with the following modifications: enzyme was set to trypsin/ Lys-C, −150 to 500 Da precursor mass tolerance, with a 500 to 5000 peptide mass range, methionine oxidation and N-terminal acetylation as variable modifications, and cysteine carbamidomethylation as a fixed modification. Peptide length could not extend past 120, so a new peptide length of 6−120 was used. All mass changes seen by FragPipe were filtered into an output.tsv and searched for abundant and relevant modifications.

Computational Methods. All computational analyses were made using in-house python scripts (supplemental_material_scripts). Protein abundances were normalized using median normalization via python script and considered identified if they were found in all five replicates of a digestion method's samples. Differential expression (DE) criteria for selecting significance was an absolute value of $\log_2$-fold change greater than 1 and p value smaller than 0.05. P-value was generated via two-tailed student's t-test. Hydrophobicity of the proteins and peptides was calculated using LocalCider v0.1.20.[15]

Hierarchical Clustering Method. For the hierarchical clustering analysis, we performed DE analysis on each digestion preparation method versus each other digestion method. Criteria for selecting significance was absolute value of $\log_2$-fold change greater than 1 and p value smaller than 0.05. P-value was generated via two-tailed student's t-test. The resulting proteins from this DE analysis became our "hit list", with all proteins that met our criteria then subjected to hierarchical clustering using seaborn in python and mapped using z-score.[17]

GO Panther Analysis. Cellular components[18] and biological functions were determined by Gene Ontology (GO) Panther analysis.[18,19] Categories were considered significant if they passed the statistical test performed by the panther algorithm and had an FDR under 0.05. Categories were reported based on FDR value and relevance.

■ RESULTS AND DISCUSSION

For comparative analysis of MS bottom-up proteomics, we evaluated two different cell lysis conditions, four different digestion conditions, and, for one sample set, two different desalting devices. The experiment covered four commonly used digestion protocols (Figure 1). Five replicates then proceeded to their respective digestion and cleanup method as listed in sample preparation methods before peptide concentration was measured. MS analysis was performed by using a Q Exactive hybrid Quadrupole-Orbitrap MS instrument. Raw file was analyzed using Proteome Discoverer 2.5, and all outputs are shown in Table S1.

For all digestion methods, cells were lysed on ice for sonication, while the BeatBox lysed cells at room temperature.
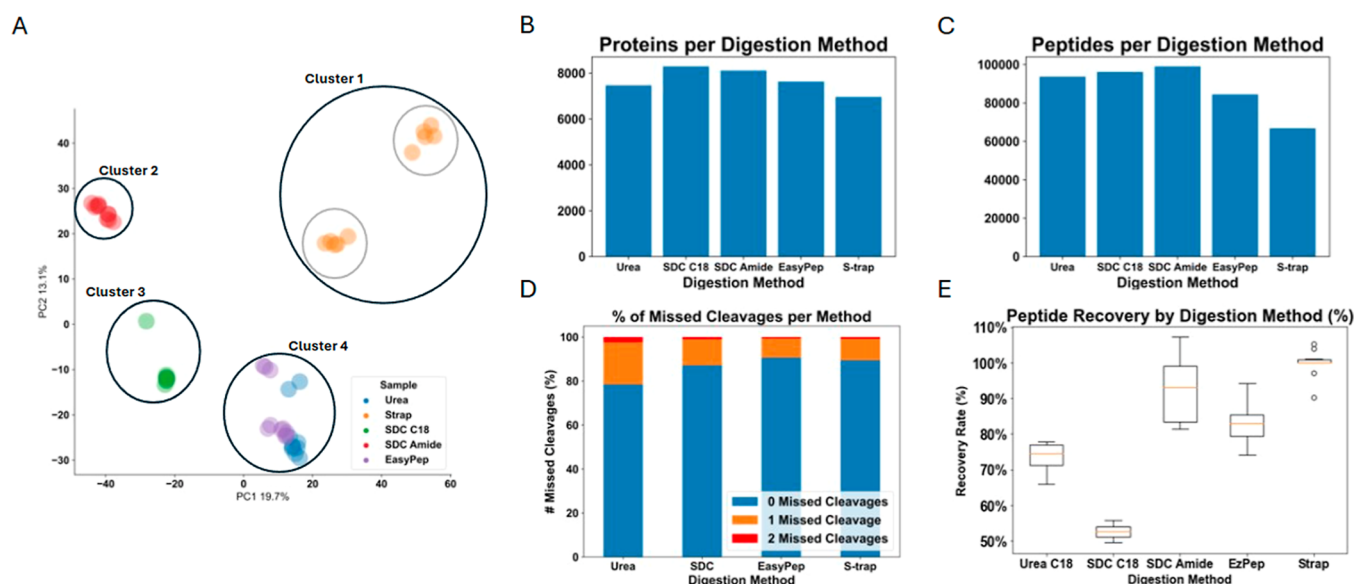
Figure 3. Comparison of general performance across digestion methods. (A) PCA analysis of all samples lysis + digestion method combinations. (B) Total proteins identified by the digestion method. (C) Total peptides identified by the digestion method. (D) Percentage of peptides identified that had 0−2 missed cleavages during digestion. (E) Box plot of peptide recovery (%) from a starting protein load of 100 μg.

We saw minor differences in the extract protein amount between the two homogenization methods (Figure S1A), but the main difference was observed between which lysis buffer was in use.

Our data analysis was broken down into five steps. First, we wanted to compare the overall performance of the four digestion methods by identifying any correlations between digestion methods via principal component analysis (PCA), evaluating the total protein and peptide identifications of each method, the number of missed cleavages, and peptide recovery rate postdigestion. Next, we wanted to see what, if any, unique proteins were identified in each of our methods or homogenization methods and whether those unique IDs were due to the stochastic sampling during the data-dependent acquisition. We then performed a hierarchical cluster analysis on the overlapping differentially extracted proteins across all other digestion methods for each digestion method. Finally, we used GO enrichment analysis tool for a pathway analysis on the subset of proteins of interest.

Comparison Results from Two Homogenization Strategies. After evaluating the total protein IDs and digestion efficiency, the next important question to ask would be what, if any, differences there would be between the two homogenization strategies we employed (Figure 2 and Tables S2 and S3). First, we generated a volcano plot using the abundances of proteins found in both sonication and BeatBox, in order to compare the log$_2$ fold change after normalization. The abundances of most proteins observed in both sonication and BeatBox samples had little to no significant changes. We found 47 proteins were differentially enriched or depleted, depending on the strategy (Figure 2A). Sonication samples have 16 proteins enriched compared to their abundances in BeatBox samples, while 31 protein abundances were found depleted. With over 4000 protein IDs shared between the two methods, this relatively small difference leads us to believe the two strategies perform equivalently.

Next, we determined whether any proteins were uniquely identified between the two homogenization strategies. To do this, we created an upset plot containing only proteins

identified in all five injections of a digestion method's samples. In addition to the shared proteins, our results display that, while sonication and BeatBox share more than 92% of their identified proteins, there were a total of 598 proteins that were only found in one of the two methods (Figure 2B). BeatBox samples had a total of 373 unique proteins, while sonication samples had a total of 225. Since both numbers are fairly similar, our methods did not observe a significant difference in protein coverage between the two homogenization strategies.

To determine if these unique proteins were only found in their respective digestion methods due to low abundance and the random nature of stochastic sampling in data-dependent acquisition, the next step in our process was to evaluate the protein abundance level of each list of unique proteins. Each colored dot represents one unique protein ID. The colors correspond to the colors of that method on the upset plot. By graphing the abundance values of all proteins found in each digestion method in a separate color from the unique proteins, we saw that the unique proteins were not all low abundance and therefore not due to being randomly detected in one digestion method. For both sonication (Figure 2C) and BeatBox (Figure 2D), the colored dots were mainly clustered in the center of both graphs, indicating to us that the sampling of these protein IDs were not due to the nature of stochastic sampling.

While the overall performance of both homogenization methods was remarkably similar across all digestion methods, the S-Trap digestion samples seem to show a slight separation between sonication and BeatBox based on the PCA (Figure 3A, cluster 1). There were some additional benefits to BeatBox during experimental procedures. First, all BeatBox samples could be carried out in parallel, while those lysed via sonication probe had to be done one at a time. The mechanical nature of BeatBox also allows the user to walk away and attend to other duties/experiments during homogenization, while sonication requires their full attention. This comes at a cost disadvantage as the BeatBox requires specially made, one time use tubes containing magnetic beads, which are consumable items. Though it takes slightly more attention, sonication requires far
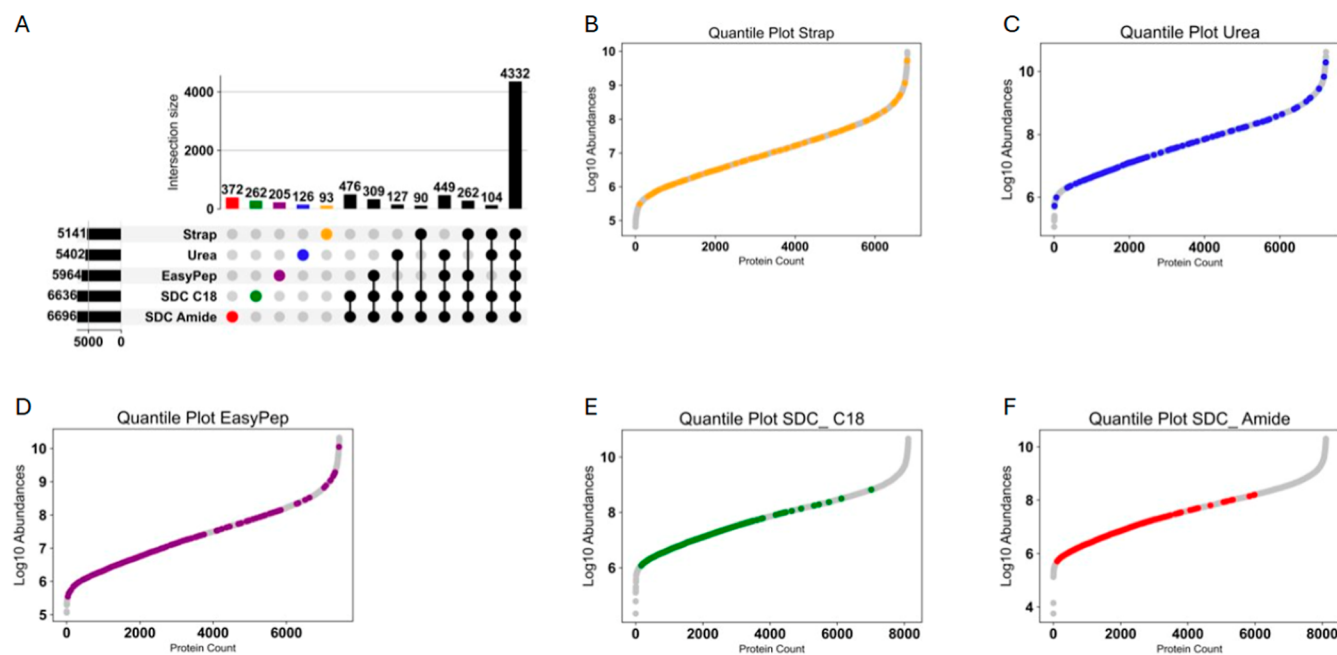
Figure 4. Unique proteins identified by digestion method and abundance of said proteins. (A) Upset plot of unique proteins identified in each digestion method. (B) Quantile plot of all abundances in S-Trap samples with unique proteins highlighted. (C) Quantile plot of all abundances in urea samples with unique proteins highlighted. (D) Quantile plot of all abundances in EasyPep samples with unique proteins highlighted. (E) Quantile plot of all abundances in SDC C18 samples with unique proteins highlighted. (F) Quantile plot of all abundances in SDC amide samples with unique proteins highlighted.

less consumables, needing only a bucket of ice, and any low-binding Eppendorf tube to store the sample. Though it has been reported in previous studies that sonication can cause aggregation of more hydrophobic proteins,[20] we used the LocalCider[15] to calculate the hydrophobicity of each uniquely identified protein and each protein that was enriched or depleted in our lysis strategies. Sonication causing aggregation of such hydrophobic proteins was not reflected in our data (Figure S2 and Table S4). Additionally, we also plotted the distribution of other biophysical properties such as isoelectric point and molecular weight of the uniquely identified proteins in both lysis strategies (Figure S3) and observed little to no differences in their distribution.

Comparison of General Performance across Digestion Methods. We compared the overall performance of each digestion method. Total number of protein IDs, peptide IDs, peptide recovery rate, and number of missed cleavages are all good indicators of a method's overall performance.[3,4] PCA of the results shows that samples fall into four distinct groups (Figure 3A). S-Trap samples are further divided into two distinct clusters. This is due to the homogenization method (Figure S1C), which demonstrates that the homogenization method had the some impact on S-Trap samples. SDC and urea maintain a tight grouping regardless of sonication or BeatBox, showing that the lysis method has little to no effect on those digestion methods. EasyPep has a few sonication samples that are different across PC2, though they are very similar across PC1.

We identified between 6000 to 8000 proteins along with 90,000+ peptides (Figure 3B) per digestion method. Both SDC-based methods showed higher protein and peptide coverage than the other methods, with SDC samples desalted with C18 having the most protein IDs and SDC amide having the most peptide IDs (Figure 3C). This is in line with the observations of other studies, which suggest that SDC has a

positive effect on trypsin activity.[21] This could also be due to the detergent extracting more proteins than other digestion methods, allowing access to a part of the proteome the other 3 methods struggle to reach, such as the endomembrane system, the peroxisome, and the microbody (Figure S4).[22] Interestingly, while S-Trap had similar numbers of protein IDs as the other methods, it also had almost 10,000 fewer peptide IDs than the urea sample.

We also examined the percentages of missed cleavages (Figure 3D). Thermo's EasyPep has the lowest percentage of missed cleavages (9%), which is in line with the results of other studies.[3−7] Conversely, urea has the highest percentage of 1−2 missed cleavages (21%), which could be due to interference with tryptic activity.[23,24] This incomplete digestion could result in lower protein IDs and lower proteome coverage overall, and we do observe urea to be below both the SDC and EasyPep methods on protein coverage (Figure 3B). Previous experiments have shown using dual-protease with trypsin and Lys-C has a positive effect on proteome coverage.[25,26] To that end, we utilized Thermo Fisher's trypsin/Lys-C mix for all our digestion methods, in order to maximize coverage and better compare each method's results. As a result, the protein and peptide counts identified in this study is higher than seen in similar studies, providing additional information that is not investigated before.[3−7] In terms of peptide recovery after desalting, S-Trap samples were the most consistent in terms of peptide recovery post desalting. EasyPep had the highest variation among samples for peptide recovery (Figure 3E). SDC C18 had the lowest overall recovery, while SDC amide, S-Trap, and EasyPep had the highest overall recovery. The C18 columns were more effective at recovering peptides from urea samples compared to SDC samples, perhaps due to not all SDC being precipitated out via pH change or possibly due to SDC extracting some components of the cell which can interfere with peptides binding to C18.
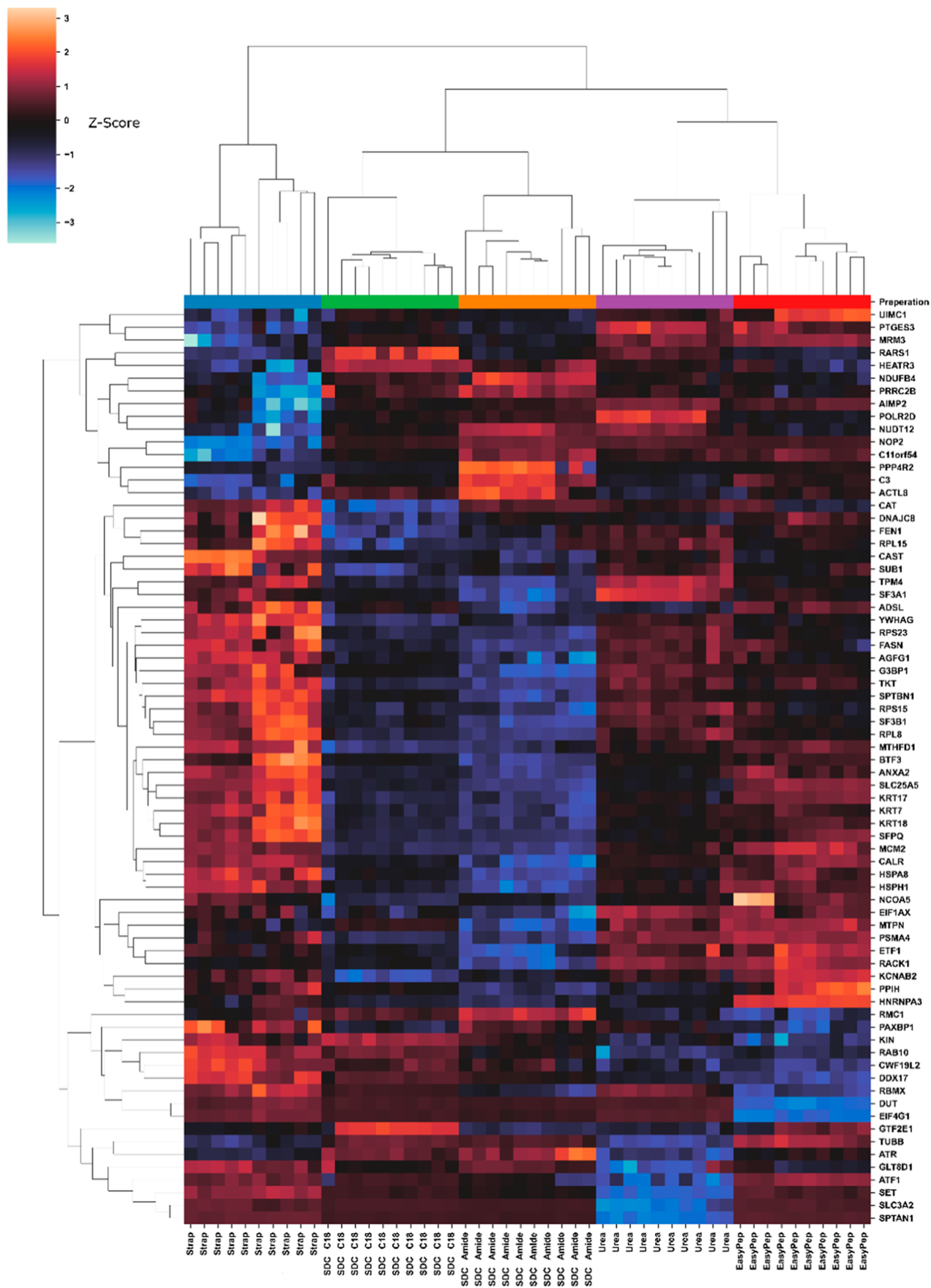
Figure 5. Analysis of uniquely depleted proteins abundance z-scores across digestion protocols. Hierarchical cluster analysis of differentially identified proteins that are both shared among all digestion methods and have a *p*-value of less than 0.05. Hierarchical cluster analysis shows four clusters of proteins uniquely depleted by that digestion method.
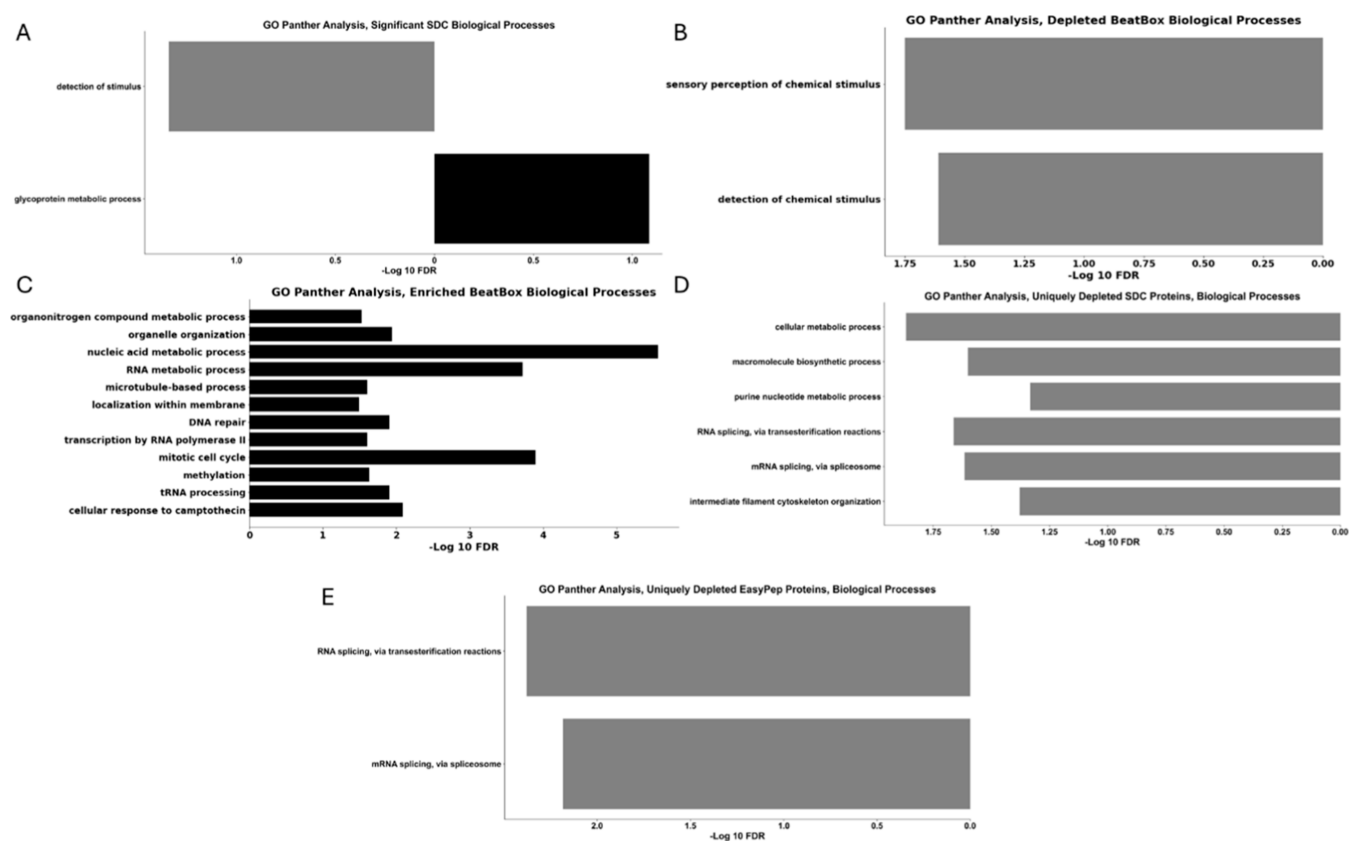
Figure 6. Pathway analysis for biological processes (BP) found in each unique protein list and the uniquely depleted protein lists. (A) GO Panther analysis of enriched SDC amide BP. (B) GO Panther analysis of depleted BeatBox BP. (C) Go Panther analysis of enriched BeatBox BP. (D) GO Panther analysis of all uniquely depleted SDC BP. (E) GO Panther analysis of uniquely depleted EasyPep BP.

Sample Preparation Artifacts. The data of each digestion method were also reanalyzed using FragPipe's open search function to identify modified peptides without the need to specify our modifications of interest.[13] This was performed to determine if any of the digestion methods created any additional modifications, which we termed sample preparation artifacts. During our sample preparation, we used TCEP as a reducing agent and CAA as an alkylator. We chose these over the more commonly used DTT and IAA, which are known to increase levels of carbamylation under urea conditions, while also needing a longer treatment time at higher temperature.[25] We filtered out common PTMs found in proteomic sample preparations, such as carbamidomethylation (+57.0215 Da), acetylation (+42.0106 Da), and oxidation (+15.9949 Da). With these three PTMs excluded, we found no other PTMs in our sample set. While the open search data do not have the sensitivity to reveal all modifications and possible sample preparation artifacts, they did provide an unbiased view of our sample peptides and showcase that little to no peptides are affected by modifications created by our sample methods.

Unique Proteins Identified by Digestion Methods. Having evaluated the differences in overall method performance, we moved on to investigating how different digestion methods affected the amount and type of proteins identified (Figure 4A). Most of the protein coverage is commonly identified across all five digestion methods. Like the previous upset plot for the homogenization strategies, only proteins identified in all five injections of a digestion method's samples were considered for this upset plot. Over 4000 proteins were shared across our five digestion methods, with S-Trap samples

having the lowest amount of total overlap. This could be due to the nature of suspension trapping versus in-solution digests (which is a category EasyPep also falls into). SDC amide has the most unique proteins, followed by SDC C18 and then EasyPep. These three methods also had the highest number of total protein IDs. Both types of SDC had nearly identical amounts of identified proteins, with the amide column identifying only 60 additional proteins over the C18 column. SDC samples encompass most proteins found across other digestion methods, missing out only on the unique protein IDs (686) of the other digestion methods. This is in line with data from Figure 3B,C, which shows SDC (either desalting method) has the highest number of identified proteins and peptides.

As we did with the unique proteins found in our homogenization strategies, we used a quantile plot to determine if the stochastic sampling had any impact on the detection of unique proteins for the lysis strategies (Figure 4B−F). As with Figure 3, the silver dots represent each abundance value for every protein identified by the lysis strategy, while the colored dots correspond to the unique proteins found in each digestion method. Proteins found only in the S-Trap digestion method were spread out across our plot (Figure 4B). This would indicate that these proteins were not selected entirely at random. The quantile plots for urea (Figure 4C) and EasyPep (Figure 4D) are very similar to those of S-Trap. All three have a few high abundant unique proteins, with the majority distributed across the entire plot. For SDC C18 (Figure 4E) and SDC amide (Figure 4F), the majority of

unique proteins are located near the bottom half of the plot, with only a few colored points appearing above the middle.

Hierarchical Clustering Shows Uniquely Depleted Proteins for Each Digestion Method. To further evaluate any differences between the digestion methods, we performed a DE analysis with each digestion method compared to each other digestion method. Volcano plots show that proteins are differentially expressed by directly comparing different digestion methods (Figure S5). This was followed by a hierarchical cluster analysis of all our hit list proteins, expressed as an overlap of all pairwise comparisons containing each digestion method. If the protein had a *p*-value in the DE of less than 0.05 and fold change larger than 2-fold, we considered it a hit. This allowed us to isolate a "hit list" of protein identities that fell into the statistically significant category for our experiments and evaluate any patterns found (Figure 5). 72 proteins followed that criterion, from which we observed four distinct groups, clustered by the digestion method. Among the proteins listed, SDC had the largest number of uniquely depleted proteins among all methods. Both urea and EasyPep were observed to have a z-score within the 0 to 2.5 range for most of these proteins, with small pockets of depletions.

Our next step was to discover what properties, if any, these groups of proteins had in common. Using GO Panther analysis, we analyzed each group of proteins to see if there were any significant enrichments in cellular components or biological processes (Figure S4 and Tables S5, S6, S7, and S8). For the list of proteins depleted in S-Trap samples, our examination found no significant enrichment in either category. The same was true of proteins depleted in the urea samples. These groups of proteins have no statistically significant ties to any biological function or cellular components. For people who are interested in performing a proteomic study, it is encouraged to check our data set to avoid using methods that can negatively influence the identification or quantitation of their proteins of interest.

Pathway Analysis. We further investigated whether the uniquely identified proteins for each method (Figure 3A) are enriched in any biological pathways. Using GO Panther analysis, each list of unique proteins was searched if there was a statistically significant number associated with a particular cellular component or biological process. Unique proteins found in SDC amide samples had an enriched number of proteins associated with glycoprotein metabolic processes. In addition, this list of proteins also had a depleted (a lower-than-expected number of proteins associated with a particular component or process) number of proteins associated with the detection of stimuli (Figure 6A). SDC amide was the only digestion method whose unique protein list generated any results for biological processes.

Analyzing cellular components utilizing GO Panther analysis led to significant enrichments for all five digestion methods. Proteins uniquely identified in urea samples were found enriched in two cellular component categories (Figure S4A), while proteins uniquely identified in urea samples contain an enriched amount from the nucleus (Figure S4B). EasyPep samples had components from all the above categories, while adding the cytoplasm and nucleoplasm as well (Figure S4C). SDC C18 (Figure S4D) and SDC amide (Figure S4E) shared enrichments from intracellular anatomical structures and cytoplasm, though SDC C18 samples had additional enrichments from the endomembrane system and intraciliary transport particle A. The list of unique proteins found in

sonication was not associated with any biological process in a relevant way. However, these proteins showed unique cellular component extractions, as they were enriched in terms associated with organelles and mitochondria (Figure S4A). For unique proteins identified only in BeatBox samples, we also found enrichments in organelle components as well as some nuclear components (Figure S4B). GO analysis of our list of unique BeatBox proteins found both depletions and enrichments in several biological processes. Synaptic and plasma membrane proteins along with proteins associated with sensory perception were lower than expected in the list of abundant proteins for BeatBox samples (Figure 6B), while several categories, such as tRNA processing, transcription, and DNA repair, were enriched (Figure 6C). Since depletions are only seen in our BeatBox samples, sonication might serve as a slightly better "all-around" lysis strategy while BeatBox may be best utilized when looking in specific areas, such as nuclear-associated proteins; the two lysis strategies are virtually identical outside the unique protein lists, which made up only 8% of the total protein IDs.

For the group of uniquely depleted proteins identified in SDC (Figure 6D) and EasyPep samples (Figure 6E), our GO panther analysis showed a statistically significant depletion in the expected amount of representation of the listed biological processes. Both SDC and EasyPep samples shared depletions in two categories: RNA transfection and mRNA splicing, though SDC samples had four additional categories related to cytoskeleton organization and various metabolic processes. In addition, we found no significant differences in the proportion of membrane proteins identified (Table S7), with all methods containing around 50% membrane proteins.

■ CONCLUSIONS

Our study provides an in-depth evaluation of the homogenization strategies of sonication and BeatBox, along with in-solution digest methods using urea and SDC, and commercially available kits such as S-Trap and EasyPep. Previous publications also comparing bottom-up techniques[3−7] noted that some methods could be improved with the addition of Lys-C during the digestion step. Since EasyPep uses a mixture of both trypsin and Lys-C, we used the same protease mixture in all of our digestions, which played a role in higher protein and peptide identifications when compared with similar studies. A comparison of overall performance for our chosen digestion methods showcased strengths and weaknesses for each. While SDC samples had the highest number of protein and peptide identifications, the need to cause SDC to precipitate out of solution using organics such as methanol or acidifying it with TFA is an extra step not needed in urea in-solution or in either kit. Similarly, the S-Trap requires a much higher amount of trypsin due to its method including precipitating proteins in order to trap them. EasyPep and S-Trap, while having the lowest number of missed cleavages, also had the lowest number of identified proteins and peptides. EasyPep was also the quickest digestion method to perform, taking roughly 4 h including incubation steps. Though our data demonstrated both homogenization strategies were overall remarkably similar in levels of protein extraction efficiency and protein identification, our PCA demonstrated it does have an effect if using Profti's S-Trap digestion method. Deeper analysis revealed that the unique proteins found in our SDC amide method had depletions and enrichments in certain biological processes; BeatBox also had significant depletions

and enrichments, like DNA repair and detection of chemical stimuli, while sonication had none. We attempted to probe the hydrophobicity of the uniquely identified proteins and peptides across each method but did not find any significant differences (Figure S2 and Table S4). We believe that the mode of action causing these uniquely identified proteins is not due to their hydrophobic properties, and there is some other force at work. More effort is required to elucidate what makes these proteins unique when utilizing these digestion methods. The commercially available kits performed remarkably similar to the lab-reagent-based protocols but with the added benefit of already prepared reagents and a simplified workflow. Similar to the unique proteins found in each of the homogenization strategies, a deeper dive is required to determine what causes the higher abundant unique proteins in each digestion method to appear more frequently and only in that particular digestion method. While the majority of identified proteins are shared across all methods, there are a few differences. If you are interested in specific proteins, we encourage you to check the provided lists of unique proteins as there is a chance one method will be better for that specific identity or project.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.4c11585.

> In-depth look at homogenization strategy differences in digestion methods; direct comparison of the hydrophobicity of unique proteins and peptides identified by digestion methods; direct comparison of the isoelectric point and molecular weight of unique proteins and peptides identified by digestion methods; GO Panther cellular component analysis using unique protein lists from digestion methods; and direct comparison of digestion methods using volcano plots (PDF)

> Protein pathway groups (XLSX)

> Sonication unique proteins and BeatBox unique proteins (XLSX)

> Urea unique proteins, SDC C18 unique proteins, SDC amide unique proteins, EasyPrep unique proteins, and S-trap unique proteins (XLSX)

> BP BeatBox and CC sonication (XLSX)

> Sonication and BeatBox peptides (XLSX)

> CC depleted SDC proteins, BP depleted SDC proteins, and BP depleted EasyPrep proteins (XLSX)

> Urea full, SDC C18, SDC amide, EasyPrep full, strap full, and membrane full (XLSX)

> CC Urea, CC SDC C18, CC SDC amide, CC EasyPrep, CC S-strap, and BP SDC amide (XLSX)

> Abundance: F5: sample, sonication_urea (XLSX)

■ AUTHOR INFORMATION

Corresponding Author

Tsui-Fen Chou − *Proteome Exploration Laboratory, Beckman Institute, California Institute of Technology, Pasadena, California 91125, United States; Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, California 91125, United States;* ⓞ orcid.org/0000-0003-2410-2186; Email: tfchou@caltech.edu

Authors

Travis Uslan − *Proteome Exploration Laboratory, Beckman Institute, California Institute of Technology, Pasadena, California 91125, United States*

Baiyi Quan − *Proteome Exploration Laboratory, Beckman Institute, California Institute of Technology, Pasadena, California 91125, United States;* ⓞ orcid.org/0000-0001-6313-4274

Ting-Yu Wang − *Proteome Exploration Laboratory, Beckman Institute, California Institute of Technology, Pasadena, California 91125, United States;* ⓞ orcid.org/0000-0002-9014-6825

Marion Pang − *Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, California 91125, United States;* ⓞ orcid.org/0000-0002-0158-2976

Yanping Qiu − *Proteome Exploration Laboratory, Beckman Institute, California Institute of Technology, Pasadena, California 91125, United States; Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, California 91125, United States;* ⓞ orcid.org/0000-0003-2948-2173

Complete contact information is available at:
https://pubs.acs.org/10.1021/acsomega.4c11585

Author Contributions

§T.U. and B.Q. contributed equally.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Zhang, Y.; Fonslow, B. R.; Shan, B.; Baek, M.-C.; Yates, J. R. Protein analysis by Shotgun/Bottom-up Proteomics. *Chem. Rev.* **2013**, *113* (4), 2343−2394.

(2) Venter, J. C.; Adams, M. D.; Myers, E. W.; Li, P. W.; Mural, R. J.; Sutton, G. G.; Smith, H. O.; Yandell, M.; Evans, C. A.; Holt, R. A.; Gocayne, J. D.; Amanatides, P.; Ballew, R. M.; Huson, D. H.; Wortman, J. R.; Zhang, Q.; Kodira, C. D.; Zheng, X. H.; Chen, L.; Skupski, M.; Subramanian, G.; Thomas, P. D.; Zhang, J.; Gabor Miklos, G. L.; Nelson, C.; Broder, S.; Clark, A. G.; Nadeau, J.; McKusick, V. A.; Zinder, N.; Levine, A. J.; Roberts, R. J.; Simon, M.; Slayman, C.; Hunkapiller, M.; Bolanos, R.; Delcher, A.; Dew, I.; Fasulo, D.; Flanigan, M.; Florea, L.; Halpern, A.; Hannenhalli, S.; Kravitz, S.; Levy, S.; Mobarry, C.; Reinert, K.; Remington, K.; Abu-Threideh, J.; Beasley, E.; Biddick, K.; Bonazzi, V.; Brandon, R.; Cargill, M.; Chandramouliswaran, I.; Charlab, R.; Chaturvedi, K.; Deng, Z.; Francesco, V. D.; Dunn, P.; Eilbeck, K.; Evangelista, C.; Gabrielian, A. E.; Gan, W.; Ge, W.; Gong, F.; Gu, Z.; Guan, P.; Heiman, T. J.; Higgins, M. E.; Ji, R.-R.; Ke, Z.; Ketchum, K. A.; Lai, Z.; Lei, Y.; Li, Z.; Li, J.; Liang, Y.; Lin, X.; Lu, F.; Merkulov, G. V.; Milshina, N.; Moore, H. M.; Naik, A. K.; Narayan, V. A.; Neelam, B.; Nusskern, D.; Rusch, D. B.; Salzberg, S.; Shao, W.; Shue, B.; Sun, J.; Wang, Z. Y.; Wang, A.; Wang, X.; Wang, J.; Wei, M.-H.; Wides, R.; Xiao, C.; Yan, C.; Yao, A.; Ye, J.; Zhan, M.; Zhang, W.; Zhang, H.; Zhao, Q.; Zheng, L.; Zhong, F.; Zhong, W.; Zhu, S. C.; Zhao, S.; Gilbert, D.; Baumhueter, S.; Spier, G.; Carter, C.; Cravchik, A.; Woodage, T.; Ali, F.; An, H.; Awe, A.; Baldwin, D.; Baden, H.; Barnstead, M.; Barrow, I.; Beeson, K.; Busam, D.; Carver, A.; Center, A.; Cheng, M. L.; Curry, L.; Danaher, S.; Davenport, L.; Desilets, R.; Dietz, S.; Dodson, K.; Doup, L.; Ferriera, S.; Garg, N.; Gl_uecksmann, A.; Hart, B.; Haynes, J.; Haynes, C.; Heiner, C.; Hladun, S.; Hostin,

D.; Houck, J.; Howland, T.; Ibegwam, C.; Johnson, J.; Kalush, F.; Kline, L.; Koduru, S.; Love, A.; Mann, F.; May, D.; McCawley, S.; McIntosh, T.; McMullen, I.; Moy, M.; Moy, L.; Murphy, B.; Nelson, K.; Pfannkoch, C.; Pratts, E.; Puri, V.; Qureshi, H.; Reardon, M.; Rodriguez, R.; Rogers, Y.-H.; Romblad, D.; Ruhfel, B.; Scott, R.; Sitter, C.; Smallwood, M.; Stewart, E.; Strong, R.; Suh, E.; Thomas, R.; Tint, N. N.; Tse, S.; Vech, C.; Wang, G.; Wetter, J.; Williams, S.; Williams, M.; Windsor, S.; Winn-Deen, E.; Wolfe, K.; Zaveri, J.; Zaveri, K.; Abril, J. F.; Guigó, R.; Campbell, M. J.; Sjolander, K. V.; Karlak, B.; Kejariwal, A.; Mi, H.; Lazareva, B.; Hatton, T.; Narechania, A.; Diemer, K.; Muruganujan, A.; Guo, N.; Sato, S.; Bafna, V.; Istrail, S.; Lippert, R.; Schwartz, R.; Walenz, B.; Yooseph, S.; Allen, D.; Basu, A.; Baxendale, J.; Blick, L.; Caminha, M.; Carnes-Stine, J.; Caulk, P.; Chiang, Y.-H.; Coyne, M.; Dahlke, C.; Mays, A. D.; Dombroski, M.; Donnelly, M.; Ely, D.; Esparham, S.; Fosler, C.; Gire, H.; Glanowski, S.; Glasser, K.; Glodek, A.; Gorokhov, M.; Graham, K.; Gropman, B.; Harris, M.; Heil, J.; Henderson, S.; Hoover, J.; Jennings, D.; Jordan, C.; Jordan, J.; Kasha, J.; Kagan, L.; Kraft, C.; Levitsky, A.; Lewis, M.; Liu, X.; Lopez, J.; Ma, D.; Majoros, W.; McDaniel, J.; Murphy, S.; Newman, M.; Nguyen, T.; Nguyen, N.; Nodell, M.; Pan, S.; Peck, J.; Peterson, M.; Rowe, W.; Sanders, R.; Scott, J.; Simpson, M.; Smith, T.; Sprague, A.; Stockwell, T.; Turner, R.; Venter, E.; Wang, M.; Wen, M.; Wu, D.; Wu, M.; Xia, A.; Zandieh, A.; Zhu, X. The sequence of the human genome. *Science* **2001**, *291* (5507), 1304−1351.

(3) Varnavides, G.; Madern, M.; Anrather, D.; Hartl, N.; Reiter, W.; Hartl, M. In search of a universal Method: A Comparative survey of Bottom-Up Proteomics sample preparation methods. *J. Proteome Res.* **2022**, *21* (10), 2397−2411.

(4) Glatter, T.; Ahrné, E.; Schmidt, A. Comparison of different sample preparation protocols reveals Lysis Buffer-Specific extraction biases in Gram-Negative bacteria and human cells. *J. Proteome Res.* **2015**, *14* (11), 4472−4485.

(5) Fic, E.; Kedracka-Krok, S.; Jankowska, U.; Pirog, A.; Dziedzicka-Wasylewska, M. Comparison of protein precipitation methods for various rat brain structures prior to proteomic analysis. *Electrophoresis* **2010**, *31* (21), 3573−3579.

(6) Wojtkiewicz, M.; Berg Luecke, L.; Kelly, M. I.; Gundry, R. L. Facile preparation of peptides for mass spectrometry analysis in Bottom-Up proteomics workflows. *Curr. Protoc.* **2021**, *1* (3), No. e85.

(7) Pang, M.; Jones, J. J.; Wang, T.-Y.; Quan, B.; Kubat, N. J.; Qiu, Y.; Roukes, M. L.; Chou, T.-F. Increasing Proteome Coverage Through a Reduction in Analyte Complexity in Single-Cell Equivalent Samples. *J. Proteome Res.* **2024**, DOI: 10.1021/acs.jproteome.4c00062.

(8) HaileMariam, M.; Eguez, R. V.; Singh, H.; Bekele, S.; Ameni, G.; Pieper, R.; Yu, Y. S-Trap, an ultrafast Sample-Preparation approach for shotgun proteomics. *J. Proteome Res.* **2018**, *17* (9), 2917−2924.

(9) Zougman, A.; Selby, P. J.; Banks, R. E. Suspension trapping (STrap) sample preparation method for bottom-up proteomics analysis. *Proteomics* **2014**, *14* (9), 1006−1010.

(10) Dupree, E. J.; Jayathirtha, M.; Yorkey, H.; Mihasan, M.; Petre, B. A.; Darie, C. C. A critical review of Bottom-Up proteomics: the good, the bad, and the future of this field. *Proteomes* **2020**, *8* (3), 14.

(11) Zhang, C.; Shi, Z.; Han, Y.; Ren, Y.; Hao, P. Multiparameter optimization of two common proteomics quantification methods for quantifying Low-Abundance proteins. *J. Proteome Res.* **2019**, *18* (1), 461−468.

(12) Perez-Riverol, Y.; Bai, J.; Bandla, C.; García-Seisdedos, D.; Hewapathirana, S.; Kamatchinathan, S.; Kundu, D. J.; Prakash, A.; Frericks-Zipper, A.; Eisenacher, M.; Walzer, M.; Wang, S.; Brazma, A.; Vizcaíno, J. A. The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* **2022**, *50* (D1), D543−D552.

(13) The UniProt Consortium. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res.* **2023**, *51* (D1), D523−D531.

(14) Kong, A. T.; Leprevost, F. V.; Avtonomov, D. M.; Mellacheruvu, D.; Nesvizhskii, A. I. MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry−based proteomics. *Nat. Methods* **2017**, *14* (5), 513−520.

(15) Holehouse, A. S.; Das, R. K.; Ahad, J. N.; Richardson, M. O. G.; Pappu, R. V. CIDER: Resources to Analyze Sequence-Ensemble Relationships of Intrinsically Disordered Proteins. *Biophys. J.* **2017**, *112* (1), 16−21.

(16) Martens, L.; Chambers, M.; Sturm, M.; Kessner, D.; Levander, F.; Shofstahl, J.; Tang, W. H.; Römpp, A.; Neumann, S.; Pizarro, A. D.; Montecchi-Palazzi, L.; Tasman, N.; Coleman, M.; Reisinger, F.; Souda, P.; Hermjakob, H.; Binz, P.-A.; Deutsch, E. W. MZML—a community standard for mass spectrometry data. *Mol. Cell. Proteomics* **2011**, *10* (1), R110.000133.

(17) Waskom, M. seaborn: statistical data visualization. *J. Open Source Softw.* **2021**, *6* (60), 3021.

(18) Mi, H.; Muruganujan, A.; Casagrande, J. T.; Thomas, P. D. Large-scale gene function analysis with the PANTHER classification system. *Nat. Protoc.* **2013**, *8* (8), 1551−1566.

(19) Thomas, P. D.; Ebert, D.; Muruganujan, A.; Mushayahama, T.; Albou, L.; Mi, H. PANTHER: Making genome-scale phylogenetics accessible to all. *Protein Sci.* **2022**, *31* (1), 8−22.

(20) Zhu, Z.; Zhu, W.; Yi, J.; Liu, N.; Cao, Y.; Lu, J.; Decker, E. A.; McClements, D. J. Effects of sonication on the physicochemical and functional properties of walnut protein isolate. *Food Res. Int.* **2018**, *106*, 853−861.

(21) Masuda, T.; Tomita, M.; Ishihama, Y. Phase transfer Surfactant-Aided Trypsin digestion for membrane proteome analysis. *J. Proteome Res.* **2008**, *7* (2), 731−740.

(22) Danko, K.; Lukasheva, E.; Zhukov, V. A.; Zgoda, V.; Frolov, A. Detergent-Assisted Protein Digestion—On the way to avoid the key bottleneck of shotgun Bottom-Up proteomics. *Int. J. Mol. Sci.* **2022**, *23* (22), 13903.

(23) Viswanatha, T.; Pallansch, M.; Liener, I. E. The inhibition of trypsin. II. The effect of synthetic anionic detergents. *J. Biol. Chem.* **1955**, *212* (1), 301−309.

(24) Gabel, D. The denaturation by urea and guanidinium chloride of trypsin and N-Acetylated-Trypsin derivatives bound to Sephadex and agarose. *Eur. J. Biochem.* **1973**, *33* (2), 348−356.

(25) Zhang, X. Less is More: Membrane Protein Digestion Beyond Urea−Trypsin Solution for Next-level Proteomics. *Mol. Cell. Proteomics* **2015**, *14* (9), 2441−2453.

(26) Betancourt, L. H.; Sanchez, A.; Pla, I.; Kuras, M.; Zhou, Q.; Andersson, R.; Marko-Varga, G. Quantitative Assessment of Urea In-Solution Lys-C/Trypsin Digestions Reveals Superior Performance at Room Temperature over Traditional Proteolysis at 37 °C. *J. Proteome Res.* **2018**, *17* (7), 2556−2561.