Published in final edited form as:

Anal Chem. 2013 March 5; 85(5): 2825–2832. doi:10.1021/ac303352n.

Segmentation of precursor mass range using ‘tiling’ approach increases peptide identifications for MS¹-based label-free quantification

Catherine E. Vincent^{1,*}, Gregory K. Potts^{1,*}, Arne Ulbrich^{1,*}, Michael S. Westphall¹, James A. Atwood III⁴, Joshua J. Coon^{1,2,§}, and D. Brent Weatherly^{3,4}

¹Department of Chemistry, University of Wisconsin, Madison, WI 53706

²Department of Biomolecular Chemistry, University of Wisconsin, Madison, WI 53706

³Complex Carbohydrate Research Center, University of Georgia, Athens, GA

⁴Ionformatics, LLC, Lawrenceville, GA

Abstract

Label-free quantification is a powerful tool for the measurement of protein abundances by mass spectrometric methods. To maximize quantifiable identifications, MS¹-based methods must balance the collection of survey scans and fragmentation spectra while maintaining reproducible extracted ion chromatograms (XIC). Here we present a method which increases the depth of proteome coverage over replicate data-dependent experiments without the requirement of additional instrument time or sample pre-fractionation. Sampling depth is increased by restricting precursor selection to a fraction of the full MS¹ mass range for each replicate; collectively, the *m/z* segments of all replicates encompass the full MS¹ range. Although selection windows are narrowed, full MS¹ spectra are obtained throughout the method, enabling the collection of full mass range MS¹ chromatograms such that label-free quantitation can be performed for any peptide in any experiment. We term this approach “binning” or “tiling” depending on the type of *m/z* window utilized. By combining the data obtained from each segment, we find that this approach increases the number of quantifiable yeast peptides and proteins by 31% and 52%, respectively, when compared to normal data-dependent experiments performed in replicate.

Introduction

Mass spectrometry methods for quantitative proteomics aim to maximize protein identifications and accurately characterize protein abundance in a cost- and time-efficient manner. MS¹-based label free methods are an attractive option for relative protein quantification, as they eliminate the expenses and sample preparation associated with isotope or mass tag labeling techniques.^{1–3} To attain quantitative data, these methods exploit the linearity of peptide spectral peak intensity and relative peptide abundance in a mixture.^{2, 4–6} Each sample individually undergoes LC-MS/MS analysis, and extracted ion chromatogram (XIC) signal intensities from identical peptides are then compared across the separate analyses, such that the relative abundance of their parent proteins within the different samples can be determined.^{7–9} The development of algorithms to facilitate chromatogram alignment has been crucial for these XIC comparisons, but highly

§to whom correspondence should be addressed: dbrentw@uga.edu.

*CEV, GKP, and AU contributed equally to this work

Supporting Information Available: This material is available free of charge via the Internet at <http://pubs.acs.org>.

reproducible separations remain essential for the acquisition of reliable quantitative data using an MS¹-based approach.^{10–14}

As in other proteomics experiments, the maximization of protein identifications using MS¹-based label-free methods becomes more daunting as sample complexity increases. Traditional data-dependent acquisition favors the highest-intensity peptides for analysis, which can preclude the sampling and identification of species present at low signal-to-noise. The reduction of sample complexity afforded by off-line fractionation facilitates an increase in attainable peptide identifications.^{15–20} The creation of multiple fractions from one complex sample disperses high-abundance peptides over multiple experiments, enabling the detection, sampling, and identification of less abundant species from reduced MS¹ complexity spectra. Unfortunately, overall sample loss and variable peptide elution across fractions are inevitable consequences of off-line fractionation, and these effects introduce additional challenges to chromatogram alignment for label-free MS¹-based quantification.²¹ Various post-acquisition data analysis strategies have been developed to correct for any systematic bias off-line fractionation introduces to MS¹-based label free analyses,^{21–22} but, the ability to increase identifications without having to devote extra time to sample preparation and post-acquisition analysis would be advantageous.

Online fractionation techniques, and the use of longer chromatography columns and/or extended chromatographic gradients, also facilitate a reduction in sample complexity, increasing overall peptide identifications.^{23–25} These methods improve chromatographic resolution over an extended LC-MS/MS analysis time to boost the number of peptides that will be detected, sampled, and identified during a single LC-MS/MS experiment. Another alternative to off-line fractionation is gas-phase fractionation (GPF) in the mass-to-charge (m/z) dimension.^{26–28} In the GPF technique, MS¹ or SIM scans are performed over a narrow m/z range (e.g. 600 – 700 instead of the full 300 – 1300 mass space). Precursor selection is restricted to this truncated scan range, which allows extension of sampling depth into lower intensity features within the m/z region. By interrogating sub-sections of the MS¹ mass range in sequential experiments, GPF increases peptide identifications compared to normal data-dependent methods.²⁷

The aforementioned strategies are useful alternatives to off-line fractionation for the maximization of peptide identifications in MS¹-based label-free analyses, since the ability to inject and analyze an unfractionated sample for each experiment limits run-to-run chromatographic variability that could compromise quantification. The major drawback of fractionation and longer gradients, however, is increased time for analysis. Every sample in an MS¹-based label-free study requires a separate LC-MS/MS experiment, and, as the number of samples involved in the comparison and/or the number of replicate experiments to perform increases, the total cost associated with analysis time also increases, compared to traditional data-dependent acquisition.

Here we report a strategy which attains a greater number of quantifiable peptide identifications than replicate data-dependent experiments, without the requirement of off-line fractionation or additional instrument analysis time. To achieve greater sampling depth than traditional data-dependent acquisition, we narrow the mass range from which a precursor can be selected for MS² analysis. This idea is similar to that employed for GPF, except the MS¹ scans, from which precursors are selected, encompass the full mass range (in our case, m/z 300 – 1300). Scherl et. al. employed a similar strategy when they limited precursor selection to truncated, continuous regions of the MS¹ mass range; their study found that the depth of proteome coverage was increased as a result.²⁹ We refer to this method as ‘binning.’ We demonstrate that, by restricting precursor selection to several narrowed m/z ranges distributed throughout the full MS¹ mass range (‘tiling’) instead of one

continuous region ('binning'), we can boost peptide identifications to an even greater extent, without the requirement of additional instrument analysis time. The 'binning' and 'tiling' methods are ideal for the maximization of protein identifications in MS¹-based label-free analyses, as they maintain acquisition of the full MS¹ mass range throughout all experiments. This not only preserves high chromatographic reproducibility between LC-MS/MS analyses for accurate XIC comparisons, but it also enables the acquisition of replicate data within one set of experiments (a peptide identified in one experiment should theoretically be quantifiable in all 'binning' or 'tiling' experiments). We apply our 'binning' and 'tiling' approaches to the MS¹-based label-free quantitative analysis of a complex mixture of tryptically-derived yeast peptides and demonstrate: 1) a significant boost in quantifiable identifications over traditional data-dependent acquisition (DDA); and 2) a reduction in analysis time, compared to the acquisition of replicate DDA analyses of off-line fractionated samples.

Materials and methods

Cell culture, differentiation, and lysis

Wild-type yeast (*Saccharomyces cerevisiae*) was grown in rich medium to an OD₆₀₀ of 0.6. Cells were collected and centrifuged at 14,200 g for 10 min at 4 °C. The resulting cell pellet was washed twice with sterile water and centrifuged at 1,100 g for 5 min. Lysis buffer of approximately three times the cell pellet volume was added. The lysis buffer contained 8 M urea, 75 mM NaCl, 50 mM Tris (pH 8.2), 1 mM sodium orthovanadate, 100 mM sodium butyrate, complete mini EDTA-free protease inhibitor (Roche Diagnostics, Indianapolis, IN) and phosSTOP phosphatase inhibitor (Roche Diagnostics, Indianapolis, IN). Yeast cells were frozen into small droplets using liquid nitrogen. The resulting pellets were placed in a pre-chilled grinder jar (Restek, Bellefonte, PA) in equal volume with acid-washed glass beads (Sigma Aldrich, St. Louis, MO). The cells were lysed using a MM4000 Mixer Mill (Restek, Bellefonte, PA), where the sample was beat at 30 Hz for 4 minutes a total of three times. The lysate was centrifuged for 10 minutes at 4,000 rpm at 4°C.

Cell digestion

Yeast proteins were subjected to cysteine residue reduction using 5 mM DTT and alkylation using 10 mM iodoacetamide. The protein sample was diluted to a final concentration of 1.5 M urea (pH 8) with a solution of 25 mM Tris and 2 mM CaCl. Sequencing-grade trypsin (Promega, Madison, WI) was added to each sample at a ratio of 1:50 (enzyme:protein) and the resulting mixture was incubated at 37°C overnight. The reaction was quenched using trifluoroacetic acid. The sample was desalted using C18 solid-phase extraction columns (SepPak, Waters, Milford, MA) and dried to completion).

Sample Preparation

Yeast peptides were re-suspended in 0.2% formic acid and split into two equal mass aliquots. One aliquot was used for LC-MS analysis (unfractionated sample). The other aliquot was subjected to reverse-phase (RP) fractionation. The digested peptides were dried to completion and dissolved in 100 µL of RP buffer A [20 mM NH₄HCO₂ (pH 10)]. The sample was injected onto a column packed with non-polar material (Gemini 5 µm C18 110 Å LC Column 250 × 4.6 mm, Phenomenex, Torrance, CA) attached to a Surveyor LC quaternary pump (Thermo Electron, West Chester, PA) running at 3.0 mL/min. Peptides were detected using a PDA detector (Thermo Electron, West Chester, PA). The eluate was collected in 30-sec intervals starting 12 minutes into the following gradient: 4 min of 95% RP buffer A and 5% RP buffer B [20 mM NH₄HCO₂, 80% ACN (pH 10)], followed by a linear gradient of 5–12% RP buffer B from 4 to 8 min, followed by a linear gradient of 12–45% RP buffer B from 8 to 34 min, followed by a linear gradient of 45% – 100% RP buffer

B from 34 to 36 min. RP buffer B was held at 100% for 5 minutes, after which there was a 1 min transition to 5% RP buffer B. The column was re-equilibrated at 5% RP buffer B for 20 minutes. A total of 40 fractions were collected into 20 vials (each vial contained two 30-sec fractions collected 10 minutes apart). The 20 vials were pooled into 5 samples such that each sample contained fractions collected 2.5 minutes apart. All samples were dried on a vacuum centrifuge and resuspended in 0.2% formic acid for LC-MS analysis.

Liquid chromatography-mass spectrometry

All experiments were performed using a NanoAcquity UPLC system (Waters, Milford, MA) coupled to a LTQ Orbitrap Elite mass spectrometer (Thermo Fisher Scientific, San Jose, CA). Samples were loaded onto a precolumn (75 μm i.d., packed with 10 cm of 5 μm C18 particles, pore size 100Å; Microm Bioresources Inc, Auburn, CA) for 10 minutes at 98:2 buffer A [0.2% formic acid]:buffer B [acetonitrile with 0.2% formic acid] at a flow rate of 1.0 $\mu\text{L}/\text{min}$. Samples were then separated on an analytical column (75 μm i.d., packed with 15 cm of 5 μm C18 particles, pore size 100Å; Microm Bioresources Inc, Auburn, CA) at a flow rate of 0.300 $\mu\text{L}/\text{min}$ using the following gradient: an initial steep rise to 8% B in 1 min, followed by a 59 min linear gradient from 8% to 30% B, followed by a final 5 min ramp to 70% B which was held for 5 minutes. The column was equilibrated with 2% buffer B for an additional 20 min. Precursor peptide cations were generated from the eluent through the utilization of a nanoESI source.

All instrument methods consisted of an MS^1 scan (300 – 1300 m/z) analyzed in the orbitrap at a resolution ($m/\Delta m$) of 60,000 followed by ten data-dependent HCD MS^2 scans analyzed in the orbitrap at a resolution ($m/\Delta m$) of 15,000. For the DDA control methods, precursors were selected from the entire MS^1 mass range for MS^2 analysis. For the binning and tiling methods, precursors were only selected for MS^2 analysis if they fell within specific m/z range(s). The m/z ranges designated for each experiment in this study are presented in Table 1. Restricted precursor selection for the tiling and binning techniques was implemented through slight modifications to the instrument control language (ITCL, access granted by Thermo Fisher Scientific). Note that while the tiling method requires ITCL modification, the binning method can be implemented using existing Xcalibur software (m/z range restriction set in the method file under data dependent settings $\rightarrow$ global $\rightarrow$ global $\rightarrow$ mass range for selecting MS dependent masses).

All MS^2 scans employed a precursor isolation window of 2 Th and an HCD normalized collision energy (NCE) setting of 30 for 0.1 ms. The automatic gain control (AGC) target settings for precursor cations were 1×10^6 charges for MS^1 scans and 5×10^4 charges for HCD-activated MS^2 scans. Precursors were subject to dynamic exclusion for 30 seconds using a 10 ppm window.

Data Analysis

Data was processed using the in-house software suite COMPASS.³⁰ OMSSA (version 2.1.8) searches were performed against the International Protein Index (IPI: <http://www.ebi.ac.uk/IPI/>) target-decoy database comprised of yeast (*Saccharomyces* Genome Database, <http://www.yeastgenome.org>, February 3, 2011, “all” version including all systematically named open reading frames (ORFs), including verified, uncharacterized, and dubious ORFs and pseudogenes) proteins.³¹ Searches were conducted using a mono-isotopic precursor mass tolerance of ± 5.0 Da and a mono-isotopic product mass tolerance of ± 0.01 Da. The fixed modification specified was carbamidomethylation of cysteine residues and the variable modification specified was the oxidation of methionine residues. A maximum of 3 missed tryptic cleavages were allowed. For each method, data was acquired over multiple

experiments and peptide identifications were collectively filtered to both 1% FDR and 10% FDR.

Comparisons between approaches were performed using a customized version (based on version 2.6) of the ProteoIQ software (NuSep Inc., Bogart, GA, www.nusep.com). Peptides from the filtered COMPASS results were loaded into ProteoIQ, matched to their corresponding proteins (using same yeast database), and organized into protein groups. Proteins were regarded as “present” in a sample if it was identified in the 1% FDR list and “absent” if it was not identified in the 10% FDR list.

Quantifiable peptides were defined as those that gave a “quant score” of at least 0.5, as determined by the ProteoIQ software. The quant score was above 0.5 when at least one observed isotopic distribution in a full MS¹ scan matched the theoretical isotopic distribution of the associated peptide identification with a product R² value of 0.5 (the product of the R²-values based on observed vs. theoretical relative intensity and observed vs. theoretical mass accuracy).

Results and Discussion

Binning and Tiling Method Implementation

Here we present a method that increases the depth of proteome coverage over replicate data-dependent experiments without the need for additional instrument analysis time or sample fractionation. To facilitate straightforward compatibility with MS¹-based label-free quantification, we designed this method to gather full MS¹ scans throughout each LC-MS/MS analysis, thereby limiting chromatographic variability and minimizing the time required for the acquisition of quantitative data in replicate.

Figure 1A illustrates three different strategies for precursor selection using a full MS¹ scan. The first approach (Fig. 1Ai) is the traditional data-dependent acquisition (DDA), a strategy in which a user-defined number of the most intense precursors from the full MS¹ mass-to-charge (m/z) range are selected for MS² analysis. When this technique is utilized for the analysis of complex samples, the number of achievable peptide identifications is often limited by sampling; in other words, selection of the most intense precursors from the full MS¹ m/z space can preclude the sampling and identification of lower abundance species. The plots in Figure 1A show the m/z distribution of precursors sampled during a DDA control run. The color of each point, and the bars on the right hand side of each panel, indicate the experiment in which each precursor would have been sampled within each of the respective methods. Note that, for visualization purposes, each point in the DDA plot was assigned a random color; any precursor can be selected in any DDA experiment, and many will be selected in more than one LC-MS/MS analysis.

Unlike DDA, the second and third strategies (Figure 1Aii–iii) restrict precursor selection to designated segments of the full MS¹ mass range. In both methods, an unfractionated sample is subjected to a series of LC-MS/MS analyses (one for each m/z segment). Full MS¹ scans are collected throughout each analysis, and ions are only eligible for precursor selection if they fall inside the designated m/z segment of the MS¹ scan. By restricting precursor interrogation in this manner, the LC-MS/MS analyses will sample deeper into each MS¹ spectra, enabling the investigation of low-intensity peptides, which wouldn't typically be sampled by DDA. At the conclusion of the experimental sets, peptide identifications obtained from all analyses provide a comprehensive list of precursors sampled across the full MS¹ mass range.

We investigate two approaches for the creation of m/z segments from the full MS^1 mass range. The first strategy, 'binning,' divides the full MS^1 range into continuous, sequential segments (Fig. 1Aii). The second strategy, 'tiling,' partitions the full MS^1 range into segments which contain several small m/z sections spanning upper, middle, and lower m/z regions of the full MS^1 mass range (Fig. 1Aiii). The m/z range(s) assigned to each segment can be determined in several ways,^{27, 29} here we use two: 1) segments are created such that they span equal m/z ranges; and 2) boundaries are set such that each segment produces an equivalent number of peptide identifications. Note that to optimally implement this second approach, a control DDA run must be performed prior to the analysis so that the m/z distribution of peptides in the sample can be determined for the LC-MS method employed. The advantage to this second approach is that the most abundant peptides in the sample will be equally distributed between each of the segments.

Unique Peptide Identifications Achievable with Binning, Tiling, and Data-dependent Methods

The ability of the binning and tiling methods to improve peptide identification was assessed using a complex peptide mixture generated through digestion of yeast cell lysate with trypsin. Binning and tiling methods were employed using three or five segments; the m/z boundaries of each segment were assigned based on the m/z distribution of precursors identified in a prior data-dependent run of the yeast sample. These methods were then compared to normal DDA performed in triplicate (control). Figure 1 presents the outcome of these experiments. Using the binning and tiling strategies illustrated in Figure 1A, we identified 17% and 22% more unique peptides, respectively, than were identified by three replicate data-dependent control methods (Fig. 1Bi).

As was observed by Scherl et. al.,²⁹ we find that precursor m/z generally increases with elution time (Figure 1A). This indicates that, although each segment in the binning method contains an equal number of peptide precursors, these precursors are distributed unevenly along the LC gradient (Figure 1Aii). This is detrimental to the acquisition of maximum peptide identifications. For example, the majority of the precursors sampled in replicate 1 (m/z 300 – 552) of Figure 1Aii elute during the first third of the experiment. At this time, there are more precursors eluting in that m/z range than can be selected for MS/MS analysis, meaning peptides in this region are being under-sampled (Supplemental Figure 1). During the last third of this analysis, however, there are very few precursors eluting that fall within the acceptable m/z range. At this point in the gradient, there is excess time to sample all precursors available for MS/MS analysis, meaning sampling depth is maximized, but analysis time is not being effectively utilized. To reduce the sampling inefficiencies observed with the binning method, we developed a tiling scheme for the creation of m/z segments (Figure 1Aiii). By incorporating low-, medium-, and high- m/z ranges within a single segment, the tiling scheme accounts for the changes in precursor m/z with elution time to ensure a more homogeneous distribution of peptide identifications across each experiment. This, presumably, led to the modest increase in peptide identifications achieved by the tiling method compared to the binning strategy, (8,930 vs. 8,570, respectively) (Figure 1, Supplemental Figure 1). Having found that the tiling method performs just as well, and even slightly better, than the binning method, we chose to focus on the comparison between tiling and DDA methods for the remainder of our study.

Crucial to the optimization of the tiling method is the balance between maximizing peptide identifications and minimizing required analysis time. Having only a few segments is time-efficient, but the sampling of low-level peptides is sacrificed, as higher intensity precursors within the large m/z segments are predominantly sampled. Conversely, increasing the total number of segments improves the probability of sampling low-level precursors, but incurs significant time costs; a greater number of segmented experiments will need to be

performed, and time which could be utilized for precursor sampling will be allotted to acquiring MS¹ scans since too few precursors may exist in the small m/z segments to fill all ten data-dependent slots. By increasing the number of segments in our tiling method from three to five, we obtained a 20% boost in unique peptide identifications; this enabled the identification of 43% more unique peptides than could be obtained by five replicates of normal DDA (Fig. 1Bii).

Tiling Method Improves Sampling of Low Abundance Peptides and Proteins

By only sampling precursors from a certain m/z segment of the MS¹ spectrum, both binning and tiling methods increase the probability of sampling low intensity precursors that would often be passed over in normal data-dependent methods, in favor of more intense peptides with m/z values outside of the designated range. Replicate data-dependent runs frequently re-sample the most abundant peptides in the mixture, so in restricting precursor selection to specified m/z ranges, the binning and tiling methods are able to use this time to sample low-intensity precursors, gaining sampling depth and maximizing the number quantifiable identifications (Supplemental Figure 2). Figure 2A compares the peak depth of precursors selected for MS² analysis throughout one of the DDA control runs and throughout one of the tiling runs (5-segment tiling analysis, i.e. 5-tiling analysis) performed in the experiments above. Peak depth is a measure of sampling depth; if all of the species observable in an MS¹ spectrum are ranked in decreasing order of spectral intensity, the rank of a precursor selected for MS² analysis is the peak depth of that precursor (e.g., the most abundant ion has a peak depth of 1). While the DDA control barely sampled precursors above a peak depth of 250, the 5-tiling method sampled precursors to a peak depth of nearly 2,000 (Fig. 2A). Both the tiling and the DDA methods identified the most abundant yeast peptides, and the boost in unique peptide identifications achievable with the 5-tiling method was the direct result of the method's ability to evaluate low-level peptides that were not even considered by the DDA control method. This is displayed in Figure 2B, which shows that the discrepancy in identifications between the two methods was the greatest for low abundance yeast proteins.

Label-free Quantification with Tiling vs. Data-dependent Methods

Our 5-tiling method consistently outperforms normal DDA in the identification of unique peptides and proteins from an unfractionated sample; however, it fails to surpass the number of identifications gained through sample fractionation. We sought to directly compare the number of quantifiable peptides and proteins achievable using off-line fractionation and using our novel 5-tiling method. Yeast cell lysate was digested with trypsin and divided in two: one half was left unfractionated, while the other half was subjected to reverse-phase fractionation to generate five samples containing approximately equal amounts of material. The unfractionated sample was analyzed using five replicate DDA runs and the 5-tiling method, while each of the fractionated samples was analyzed once using DDA. The results are illustrated in Figure 3. Once again, the 5-tiling method produced more peptide and protein identifications than were produced by replicate DDA runs (36% and 56%, respectively), but fractionation outperformed the 5-tiling method by 37% and 34%, respectively.

Just as the 5-tiling method allows for the identification of low-level peptides absent from DDA analyses, fractionation reduces sample complexity to enable the selection and identification of precursors that remain unselected in the complex MS¹ spectra of the unfractionated sample (Figure 3B). Sample fractionation is, therefore, the best strategy for achieving maximum protein identifications, as long as sufficient time and resources are available. Aside from the challenges associated with the reconstruction of chromatograms over multiple fractions, the major drawback of conducting MS¹-based quantification with fractionated samples is the amount of analysis time required for the acquisition of replicate

analyses for reproducibility. Our 5-tiling method analyzes the same unfractionated sample in all five experiments. Since full MS¹ scans are obtained throughout each analysis in the 5-tiling method, a peptide identified in one experiment can, theoretically, be quantified in all five experiments. To achieve the same quality of quantitation (number of replicates) as the 5-tiling analysis of our unfractionated sample, the DDA analysis of our fractionated sample would require 25 (5×5) experiments. Furthermore, this elevated analysis time requirement grows considerably as the number of fractions obtained, or the number of samples being compared, increases.

Figure 3A compares the number of quantifiable peptides and proteins identified in the DDA, 5-tiling, and fractionation experiments, performed in the above investigation. Additionally, each plot depicts the number of replicate experiments in which each peptide or protein group could be quantified. A peptide was only considered quantifiable if it had at least one observed isotopic distribution in an MS¹ scan that matched its theoretical isotopic distribution with a ProteoIQ quant score of 0.5 or higher. Although the increase in peptides identified by the 5-tiling method came from lower-level proteins, the MS¹ spectral quality of these less abundant species was only marginally sacrificed; 94.8% (8,342 of 8,795) of the peptides identified by the 5-tiling method were quantifiable, compared to the 98.5% (6,353 of 6,448) quantification rate of the DDA, and the 97.0% (11,666 of 12,032) quantification rate of the fractionated analyses.

As is illustrated in Figure 3A, all quantitative information gleaned from the five analyses of the fractionated sample is grouped into one replicate. Any peptide seen in two or more experiments is indicative of variable peptide elution over multiple fractions, a factor which complicates post-acquisition analysis. In the same amount of time, our 5-tiling method collected quantitative information in quintuplicate for ~72% of all quantified peptides and proteins. Although the same percentage of identifications could be quantified in quintuplicate for the DDA analyses, the 5-tiling method successfully quantified 31% more unique peptides, and 52% more unique protein groups, than the replicate DDA experiments. Figure 3C shows a characteristic example of a peptide that was not identified in any of the DDA experiments, but was identified by the 5-tiling method and quantified across the MS¹ scans of all 5-tiling runs. Overall, there were more proteins quantified in all 5-tiling replicates than were even *identified* by DDA (Fig. 3A). This establishes our 5-tiling strategy as advantageous for the maximization of peptide and protein quantification using MS¹-based label-free methods.

Conclusion

We present a method which restricts precursor selection to sub-segments – ‘bins’ or ‘tiles’ – of the full MS¹ mass range. Each m/z segment is interrogated over separate runs, but, full m/z range MS¹ spectra are acquired throughout each experiment. This enables the run-to-run preservation of MS¹-based label-free quantification capabilities for all peptides identified in the analysis, while increasing sampling depth to yield more quantifiable peptide and protein identifications than can be obtained using traditional data-dependent acquisition. Unlike typical data-dependent methods, our binning and tiling strategies require multiple experiments to interrogate peptides from the full MS¹ mass range. This additional time requirement, however, does not only lead to a boost in identifications; precursor selection is restricted to sub-sections of full MS¹ mass range scans, meaning that, even though a peptide may only be interrogated in one experiment, it can be quantified in every experiment. This indicates that the vast majority of the peptides identified by our binning and tiling methods will also be quantified in technical replicate.

Online, off-line, and gas-phase fractionation techniques reduce sample and MS¹ spectral complexity to gain more peptide identifications than can be attained by typical data-dependent acquisition. These strategies, however, are not ideally suited for MS¹-based label-free quantitative analysis, as they complicate chromatographic alignment for accurate XIC comparisons, and require substantial instrument time for the acquisition of replicate data. These binning and tiling strategies are, therefore, useful to the label-free quantification community, as they provide an effective way to simultaneously boost protein identifications and acquire quantification data in technical replicate in a cost- and time-efficient manner.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank AJ Bureta for assistance with figure illustrations, and A. E. Merrill for culturing the yeast cells. This work was supported by the National Institutes of Health Grant GM080148 to J.J.C. C.E.V. was supported by an NLM training grant to the Computation and Informatics in Biology and Medicine Training Program (NLM T15LM007359). For access to tiling method capabilities, please contact Thermo Fisher Scientific.

References

1. Patel VJ, Thalassinou K, Slade SE, Connolly JB, Crombie A, Murrell JC, Scrivens JH. *Journal of Proteome Research*. 2009; 8:3752–3759. [PubMed: 19435289]
2. Zhu W, Smith JW, Huang CM. *Journal of Biomedicine and Biotechnology*. 2010
3. Bantscheff M, Schirle M, Sweetman G, Rick J, Kuster B. *Analytical and Bioanalytical Chemistry*. 2007; 389:1017–1031. [PubMed: 17668192]
4. Wang G, Wu WW, Zeng W, Chou CL, Shen RF. *Journal of Proteome Research*. 2006; 5:1214–1223. [PubMed: 16674111]
5. Chelius D, Bondarenko PV. *Journal of Proteome Research*. 2002; 1:317–323. [PubMed: 12645887]
6. Wang W, Zhou H, Lin H, Roy S, Shaler TA, Hill LR, Norton S, Kumar P, Anderle M, Becker CH. *Analytical Chemistry*. 2003; 75:4818–4826. [PubMed: 14674459]
7. Higgs, JREJ., Jr; Randall, K.; Kaiser; Raymond, E. US Patent. 5885841. 1999.
8. Chelius D, Zhang T, Wang G, Shen RF. *Analytical Chemistry*. 2003; 75:6658–6665. [PubMed: 14640742]
9. Bondarenko PV, Chelius D, Shaler TA. *Analytical Chemistry*. 2002; 74:4741–4749. [PubMed: 12349978]
10. Higgs RE, Knierman MD, Gelfanova V, Butler JP, Hale JE. *Journal of Proteome Research*. 2005; 4:1442–1450. [PubMed: 16083298]
11. Radulovic D, Jelveh S, Ryu S, Hamilton TG, Foss E, Mao Y, Emili A. *Molecular & Cellular Proteomics*. 2004; 3:984–997. [PubMed: 15269249]
12. Norton, S. US Patent 7628914. 2009.
13. Panchaud A, Affolter M, Moreillon P, Kussmann M. *Journal of Proteomics*. 2008; 71:19–33. [PubMed: 18541471]
14. Finney GL, Blackler AR, Hoopmann MR, Canterbury JD, Wu CC, MacCoss MJ. *Analytical Chemistry*. 2008; 80:961–971. [PubMed: 18189369]
15. Hanash SM. *ELECTROPHORESIS*. 2000; 21:1202–1209. [PubMed: 10786892]
16. Peng J, Elias JE, Thoreen CC, Licklider LJ, Gygi SP. *Journal of Proteome Research*. 2002; 2:43–50. [PubMed: 12643542]
17. Fournier ML, Gilmore JM, Martin-Brown SA, Washburn MP. *Chemical Reviews*. 2007; 107:3654–3686. [PubMed: 17649983]
18. Villen J, Gygi SP. *Nat Protocols*. 2008; 3:1630–1638.

19. Dowell JA, Frost DC, Zhang J, Li L. *Analytical Chemistry*. 2008; 80:6715–6723. [PubMed: 18680313]
20. Wang H, Chang-Wong T, Tang HY, Speicher DW. *Journal of Proteome Research*. 2009; 9:1032–1040. [PubMed: 20014860]
21. Old WM, Meyer-Arendt K, Aveline-Wolf L, Pierce KG, Mendoza A, Sevinsky JR, Resing KA, Ahn NG. *Molecular & Cellular Proteomics*. 2005; 4:1487–1502. [PubMed: 15979981]
22. Lubner CA, Cox J, Lauterbach H, Fancke B, Selbach M, Tschopp J, Akira S, Wiegand M, Hochrein H, O’Keeffe M, Mann M. *Immunity*. 2010; 32:279–289. [PubMed: 20171123]
23. Thakur SS, Geiger T, Chatterjee B, Bandilla P, Fröhlich F, Cox J, Mann M. *Molecular & Cellular Proteomics*. 2011;10.
24. Washburn MP, Wolters D, Yates JR. *Nat Biotech*. 2001; 19:242–247.
25. Altelaar AFM, Heck AJR. *Current Opinion in Chemical Biology*. 2012; 16:206–213. [PubMed: 22226769]
26. Spahr CS, Davis MT, McGinley MD, Robinson JH, Bures EJ, Beierle J, Mort J, Courchesne PL, Chen K, Wahl RC, Yu W, Luethy R, Patterson SD. *PROTEOMICS*. 2001; 1:93–107. [PubMed: 11680902]
27. Yi EC, Marelli M, Lee H, Purvine SO, Aebersold R, Aitchison JD, Goodlett DR. *ELECTROPHORESIS*. 2002; 23:3205–3216. [PubMed: 12298092]
28. Davis MT, Spahr CS, McGinley MD, Robinson JH, Bures EJ, Beierle J, Mort J, Yu W, Luethy R, Patterson SD. *PROTEOMICS*. 2001; 1:108–117. [PubMed: 11680890]
29. Scherl A, Shaffer SA, Taylor GK, Kulasekara HD, Miller SI, Goodlett DR. *Analytical Chemistry*. 2008; 80:1182–1191. [PubMed: 18211032]
30. Wenger CD, Phanstiel DH, Lee MV, Bailey DJ, Coon JJ. *PROTEOMICS*. 2011; 11:1064–1074. [PubMed: 21298793]
31. Geer LY, Markey SP, Kowalak JA, Wagner L, Xu M, Maynard DM, Yang X, Shi W, Bryant SH. *Journal of Proteome Research*. 2004; 3:958–964. [PubMed: 15473683]

Figure 1. Tiling method achieves more unique peptide identifications than binning and data-dependent methods

A. Mass-to-charge (m/z) values of precursors selected for MS² analysis during data-dependent acquisition over a 55 minute LC gradient. Any restrictions imposed for precursor m/z selection in the individual methods are represented by the red, blue, and black dots (in plots) and bars (to right of plots). B. Comparison of unique peptide identifications obtained using i) different precursor selection requirements and ii) a different number of ‘tiles’ ($n = 2$). Note that data for figures (i) and (ii) were obtained on different instruments on different dates; this accounts for identification discrepancies observed between the two figures.

Figure 2. Tiling method enables the sampling and identification of low-abundance peptides which are not observable by regular data-dependent methods

A. Comparison of MS¹ peak depth for precursors selected and/or identified by DDA control vs. tiling methods over a 60 minute LC gradient. B. Distribution of yeast proteins found in the yeast proteome (blue), the tiling experiments (red), and the DDA control experiments (grey) grouped based on relative abundance (protein copies per cell) within the yeast proteome.

Figure 3. Tiling method identifies more quantifiable unique peptides and protein groups than regular data-dependent acquisition. Tiling method attains less quantifiable identifications, but significantly more replicates, than fractionation over five LC-MS/MS analyses

A. Quantifiable unique peptides and protein groups discovered in data-dependent acquisition (5 replicates), the 5-tiling method, and fractionation (5 fractions) performed on the same tryptic yeast digest. The number of replicate experiments in which each peptide or protein group could be quantified is depicted. B. Unique peptide and protein group identifications found in each of the three methods (at 1% FDR), but absent from each of the other methods (at 10% FDR). C. Extracted ion chromatograms for peptide KVVKEQTEAAK (m/z 410.91) from each of the 5-tiling experiments. This exemplifies one of many peptides which could be quantified in all five of the Tiling experiments, but could not be quantified in any of the data-dependent experiments.

Table 1

MS¹ m/z ranges from which precursors were selected for MS² analysis in ‘binning’ and ‘tiling’ methods

Binning (3)	
Experiment	m/z ranges for precursor selection
1	300–552
2	552–730
3	730–1300
Tiling (3)	
Experiment	m/z ranges for precursor selection
1	300–435, 550–605, 730–810
2	435–500, 605–665, 810–920
3	500–550, 665–730, 920–1300
Tiling (5)	
Experiment	m/z ranges for precursor selection
1	300–367, 635–702, 970–1036
2	367–434, 702–769, 1036–1102
3	434–501, 769–836, 1102–1168
4	501–568, 836–903, 1168–1234
5	568–635, 903–970, 1234–1300