NIH Public Access
Author Manuscript
J Proteome Res. Author manuscript; available in PMC 2014 July 05.

Published in final edited form as:
J Proteome Res. 2013 July 5; 12(7): 3071–3086. doi:10.1021/pr3011588.

Evaluation and Optimization of Mass Spectrometric Settings during Data-Dependent Acquisition Mode: Focus on LTQ-Orbitrap Mass Analyzers

Anastasia Kalli[¥], **Geoffrey T. Smith**, **Michael J Sweredoski**, and **Sonja Hess**
Proteome Exploration Laboratory, Division of Biology, Beckman Institute, California Institute of Technology, Pasadena, CA 91125, USA

Sonja Hess: shess@caltech.edu

Abstract

Mass-spectrometry-based proteomics has evolved as the preferred method for the analysis of complex proteomes. Undoubtedly recent advances in mass spectrometry instrumentation have greatly enhanced proteomic analysis. A popular instrument platform in proteomics research is the LTQ-Orbitrap mass analyzer. In this tutorial we discuss the significance of evaluating and optimizing mass spectrometric settings on the LTQ-Orbitrap during CID data-dependent acquisition (DDA) mode to improve protein and peptide identification rates. We focus on those MS and MS/MS parameters, which have been systematically examined and evaluated by several researchers and are commonly used during DDA. More specifically we discuss the effect of mass resolving power, preview mode for FTMS scan, monoisotopic precursor selection, signal threshold for triggering MS/MS events, number of microscans per MS/MS scan, number of MS/MS events, automatic gain control target value (ion population) for MS and MS/MS, maximum ion injection time for MS/MS, rapid and normal scan rate and prediction of ion injection time. We, furthermore, present data from the latest generation LTQ-Orbitrap system, the Orbitrap Elite, along with recommended MS and MS/MS parameters. The Orbitrap Elite outperforms the Orbitrap Classic in terms of scan speed, sensitivity, dynamic range, resolving power and resulting in higher identification rates. Several of the optimized MS parameters determined on the LTQ-Orbitrap Classic and XL were easily transferable to the Orbitrap Elite, whereas others needed to be reevaluated. Finally, the Q Exactive and HCD are briefly discussed, as well as, sample preparation, LC-optimization and bioinformatics analysis. We hope this tutorial will serve as guidance for researchers new to the field of proteomics and assist in achieving optimal results.

Keywords

shotgun proteomics; CID data-dependent acquisition mode; identification rates; LTQ-Orbitrap; MS and MS/MS parameters

INTRODUCTION

In recent years, mass spectrometry (MS) has emerged as a powerful technique to identify, characterize and quantify proteins in complex biological samples. Shotgun approaches, in which digested protein samples are analyzed by liquid chromatography-tandem mass

[*]Corresponding author: shess@caltech.edu, phone: 626-395-2339, fax: 626-449-4159.
[¥]Current address: Department of Pathology and Laboratory Medicine, Children's Hospital Los Angeles, Los Angeles, CA 90027, USA

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

spectrometry (LC-MS/MS) in a data-dependent acquisition (DDA)[1] mode, are most commonly used in proteomics analyses.

Developments and improvements in MS instrumentation have played a key role and facilitated mass spectrometry-based proteomics. Particularly, hybrid LTQ-Orbitrap[2,3] mass spectrometers provide a powerful instrument platform for protein analysis and are increasingly used in proteomics research. In the hybrid configuration the linear trap quadrupole (LTQ, known as linear ion trap) is used as the front end for ion trapping, ion selection, fragmentation reactions and low resolution ion detection, whereas the Orbitrap is used for high resolution, high mass accuracy ion detection. The Orbitrap mass analyzer uses electrostatic fields to trap and analyze ions and consists of one central spindle and one outer barrel-like electrode.[4–6] Ions injected into the Orbitrap are electrostatically trapped while rotating and oscillating along the central electrode. Oscillating ions induce an image current signal on the outer electrodes. Image current signals are converted into frequencies by Fourier transformation. The frequencies, which are characteristic of each ion m/z value, are finally converted into a mass spectrum.[4–6] These hybrid instruments combine the sensitivity and speed of the LTQ with the high mass accuracy and high resolution of Orbitrap[2,3,6–8] mass analyzers. In addition, they can operate in two different detection modes. In the first mode, the Orbitrap mass analyzer acquires the full MS spectra and the LTQ the MS/MS fragment ion spectra, also called the FT-IT mode. In the alternative mode, both the MS and MS/MS scans are analyzed in the Orbitrap detector. This mode is referred to FT-FT. Several generations of Orbitrap mass spectrometers are currently available encompassing the Classic, XL, Discovery, Velos, Velos Pro and Elite with the later generations offering also HCD (high energy collision induced dissociation). For shotgun proteomics experiments the FT-IT mode is the most widely employed. It should be noted though that the improved HCD capabilities of the latest generation of hybrid instruments might very well change this in the future.

LTQ-Orbitrap mass analyzers have enabled the analysis of thousands of proteins providing insights beyond what is currently achievable with classical biological techniques. Consequently, mass spectrometry-based proteomics has attracted considerable interest in the biological community. In fact, an increasing number of researchers, who may not necessarily be primarily interested in mass spectrometry, move into the field because they realize that mass spectrometry-based proteomics will allow them to gain a greater understanding of biological systems and cellular processes.

However, proteome analysis by LC-MS/MS can be challenging as it requires a basic understanding and experience in mass spectrometry. In fact, familiarity with the entire proteomics pipeline, including sample preparation, LC-MS/MS, and bioinformatics analysis, is essential to achieve optimal results. Therefore, optimization of proteomics experiments can be particularly demanding for proteomics researchers with limited experience in these technologies.

In addition, when reviewing the proteomics literature it becomes apparent that the MS parameters selected for global proteomics experiments can vary widely within the same and between instrument platforms. Moreover, in a large number of publications, important information about the MS methodology is not included. The heterogeneity and sometimes even lack of instrument settings' description make it even harder for proteomics researchers to decide, which parameter to choose for their analyses.

These challenges have motivated several research groups to evaluate LC and MS settings during DDA mode in order to examine their significance in the experimental outcome and optimize instrument performance for a successful analysis. For LC optimization, reports

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

have focused on the effect of the sample loading amount,[9,10] sample loading conditions,[11] flow rate, elution gradient range and length.[10] In terms of MS parameters the effect of dynamic exclusion,[12,13] signal thresholds setting to trigger an MS/MS event,[14] mass resolving power,[15] ion injection time,[13] number of MS/MS,[13] monoisotopic precursor selection,[13] preview mode for FTMS scan,[16] normalized collision energy,[13] AGC target value (ion population) for MS and MS/MS scans,[16] and number of microscans,[16] have been evaluated and discussed, aiming to examine the significance and effect of instrument settings on identification rates. These reports revealed that LC and MS settings influence identification rates and improved results can be obtained by optimizing instrument performance.

In this tutorial we focus on the optimization of MS parameters on the widely used LTQ-Orbitrap mass spectrometer for improved proteome coverage using CID as the fragmentation technique. In addition recommended MS and MS/MS settings for the Q Exactive are presented. Sample preparation, LC-performance and bioinformatics analysis are also discussed. It should be noted that instrument settings may not necessarily translate across different instrument platforms or different fragmentation techniques (i.e. ETD).

DATA-DEPENDENT ACQUISITION (DDA) SETTINGS

The DDA mode, in which the most intense peptide ions from a full MS scan are selected for fragmentation, is by far the most widely used approach in shotgun proteomics experiments. A DDA experiment consists of several MS and MS/MS parameters for which their values need to be defined by the user. These include, but are not limited to, mass resolution, monoisotopic precursor selection, preview mode for FTMS master scan, ion population (AGC target value), and maximum ion injection time for the full MS scan. MS/MS parameters include dynamic exclusion, minimal signal threshold, number of microscans per MS/MS scan, number of MS/MS events, AGC target value and maximum ion injection time for MS/MS. Even though it is advantageous to have the flexibility of deciding which of these parameters will be used and which values will be selected, at the same time, the various parameters and the wide range of values that these parameters can obtain, complicate the design of a DDA experiment. In the following sections we discuss in detail the significance of several MS and MS/MS parameters and their effect on protein and peptide identification rates. MS and MS/MS parameters discussed herein are summarized in Table 1 and suggested optimized values determined for the Orbitrap Classic and XL are shown in Figure 1. Unless otherwise noted, results presented in the following sections were obtained using an unfractionated tryptic digest of *S. cerevisiae* as the model system.

Mass Resolving Power

When the MS or MS/MS scans are detected in the Orbitrap mass analyzer, the mass resolving power to be used during the scan event needs to be defined by the user. Usually in shotgun proteomics experiments only the full MS scan is analyzed in the Orbitrap, whereas the MS/MS scans are performed and analyzed in the linear ion trap. The resolving power of the Orbitrap mass analyzer diminishes as the square root of the m/z ratio[2,5] and during DDA mode the selected setting is specified at 400 m/z. Consequently, the resolving power is not steady across the m/z scan range and ions at higher m/z ratios are measured with lower resolution as shown in Figure 2.

Acquiring data at higher resolving power results in better mass accuracy,[4] but the scan duration increases with increasing resolving power. Pandey and coworkers examined the effect of different mass resolving power settings on identification rates using the Escherichia coli proteome and a standard 48 protein mixture on an LTQ-Orbitrap XL.[15] Both the FT-FT detection mode (MS and MS/MS scans were analyzed in the Orbitrap) and the FT-IT mode

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

(MS scan was analyzed in the Orbitrap and MS/MS scans in the ion trap) were assessed. Due to the longer scan cycle observed when employing the FT-FT method, a lower number of MS/MS scans was obtained resulting in a significantly lower number of peptide spectrum matches compared with the FT-IT method. Furthermore, the FT-IT experiments resulted in a higher number of unique peptides and proteins. With the FT-IT method the number of identifications was maximized at 30,000 resolving power, followed closely by 60,000 resolving power. The authors concluded that for shotgun proteomics experiments the FT-IT method at a mass resolving power of 30,000 could maximize identification rates. In applications in which high mass accuracy was essential the FT-FT method at a mass resolving power of 15,000 (for both the MS and MS/MS scans) provided the best comprise between speed, accuracy and identification rates.[15]

Preview Mode for FTMS Master Scan

The preview mode for FTMS master scan is another parameter used during DDA for triggering MS/MS events. When the preview mode is enabled, a preview scan in the FT-Orbitrap analyzer is performed at lower resolution and precursor ions subjected to MS/MS are selected based on the preview scan. A high resolution MS spectrum and MS/MS scans are then acquired in parallel. With the preview mode disabled, the high resolution MS scan is performed first, followed by the MS/MS scans. In the latter setting a slightly longer cycle time is required. We have recently examined the effect of enabling and disabling the preview mode for FTMS scan on protein and peptide identification rates using the *S. cerevisiae* proteome as a model system.[16] Due to the faster duty cycle a higher number of acquired and identified MS/MS spectra were observed by enabling the preview mode scan. Despite the higher number of acquired and identified MS/MS spectra, comparable results in terms of identification rates were obtained regardless of whether this option was enabled or not. Therefore, we can conclude that the preview mode for FTMS scan has no significant effect on identification rates, but it allows for a faster duty cycle.[16]

Monoisotopic precursor selection

During DDA mode, monoisotopic precursor selection can be enabled or disabled. By enabling this option, only the monoisotopic peak of the overall isotopic distribution will be selected for fragmentation. By disabling this option every isotopic peak in the full MS scan becomes accessible to fragmentation. The influence of monoisotopic precursor selection on proteome coverage was evaluated by Muddiman and coworkers[13] using the design of experiments[17] approach. A tryptic digest of *S. cerevisiae* was used as the model system. The authors showed that monoisotopic precursor selection significantly influences proteome coverage and that enabling this option results in an increase in proteome coverage. In addition to the monoisotopic precursor selection, they determined that the maximum ion injection time and the number of MS/MS events also affect identification rates.[13] These two parameters are discussed in more detail in the following sections.

Minimum Signal Threshold

This setting refers to the minimum ion abundance (intensity) required for a precursor ion in the full MS scan to be automatically selected for fragmentation during DDA mode. Increasing the signal threshold will result in lower number of acquired MS/MS spectra but the overall quality of the MS/MS spectra will increase. Decreasing the level of the signal threshold will result in a large number of MS/MS events, but the quality of the MS/MS spectra arising from low intensity ions may be insufficient for identification or the MS/MS spectra may be derived from chemical noise.

The effect of signal threshold level from 1×10^1 to 1×10^7 intensity units on peptide and protein identifications was investigated by Yates and coworkers, using unfractionated yeast

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

peptide mixture, to determine the optimal value for maximizing identification rates.[14] For signal thresholds at and below the noise level of the mass analyzers comparable results were obtained for the number of acquired MS/MS spectra and identified peptides and proteins. Therefore, the authors concluded that the optimal threshold setting is around the noise level of the mass analyzer. When the MS scan is performed in the LTQ they suggested a threshold setting of ca. 1×10^4 intensity units and when the MS scan is acquired in the Orbitrap a value of ca. 1×10^5 intensity units was proposed.[14]

Number of Microscans per MS/MS scan

The number of average microscans to produce one MS/MS scan is another variable during DDA. Increasing the number of microscans improves the signal to noise ratio but also increases the scan cycle time resulting in the collection of less MS/MS spectra. We have evaluated the effect of the number of microscans in shotgun proteomics experiments for maximizing proteome coverage.[16] As expected, increasing the number of microscans resulted in an increase in the scan cycle time and a considerably lower number of acquired MS/MS spectra. A 72% decrease in the number of acquired MS/MS spectra was observed by increasing the number of microscans from one (27139 MS/MS events) to five (7597 MS/MS events). Consequently, a lower number of identified protein groups was obtained when averaging 5 microscans per MS/MS scan. Employing one microscan resulted in 1052 (±20) protein groups identification, whereas 5 microscans resulted in 584 (±6) identifications. These results suggest that one microscan provides a CID MS/MS spectrum of sufficient quality for peptide and subsequent protein identification.

Number of data-dependent MS/MS Events

The number of data-dependent MS/MS scans can also be defined by the user during a shotgun proteomics experiment. Increasing the number of MS/MS scans will generally allow more precursor ions from the MS scan to be selected for fragmentation but the actual number of MS/MS scans performed also depends on the duration of the scan cycle and on the number of peptide candidates in the MS scan.

It has been shown that the number of MS/MS scans significantly affects proteome coverage.[13,16] By evaluating the effect of data-dependent MS/MS scans at different AGC target values and different ion injection times for MS/MS (5×10^3/50 ms, 5×10^3/150 ms, and 1×10^4/50 ms) we have shown that 12 to 15 MS/MS outperformed 5 to 10 MS/MS events for all three tested settings. For instance, at an AGC target value of 1×10^4 and maximum ion injection time of 50 ms, 5 and 12 MS/MS events resulted in 806 (±47) and 1113 (±12) protein groups identifications, respectively. Similar trends were obtained for the number of identified and unique peptides. Further increasing the number of data-dependent MS/MS scans to 20, did not provide any apparent benefit in the Orbitrap Classic. Therefore, improved proteome coverage can be achieved by increasing the number of MS/MS events to 12 – 15.[16]

AGC (Automatic Gain Control) Target Value

The AGC target value refers to the ion population in the linear ion trap or Orbitrap mass analyzer. The idea behind the AGC is to regulate the number of ions in the mass analyzer in order to avoid or minimize space charge effects to improve mass accuracy. It is important to note that the AGC target value and maximum ion injection time are not independent parameters. The MS or MS/MS events are controlled by either the AGC target value or the maximum ion injection time depending on which parameter is reached first. For example, if the AGC target value is set at 1×10^4 and the maximum ion injection time at 50 ms but it takes longer than 50 ms to accumulate 1×10^4 ions then the MS/MS (or MS) events will be performed with 50 ms regardless of the set AGC target value. If it takes less time (<50 ms)

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

to accumulate 1×10^4 ions then the MS/MS (or MS) events will be performed with the selected AGC target value.

Table 2 summarizes set and actual MS and MS/MS ion injection times observed in the LTQ-Orbitrap Classic for different AGC target values during LC-MS/MS analyses of unfractionated tryptically digested yeast lysate. As illustrated in Table 2 by increasing the AGC target values the maximum ion injection time also increases since it takes longer to accumulate a higher number of ions.

The effect of AGC target values for MS and MS/MS events on identification rates has been recently evaluated.[16] AGC target values of 5×10^5 to 1×10^6 for the full MS scan result in very similar identification rates, with 1001 and 1013 protein groups identified at AGC target values of 5×10^5 and 1×10^6, respectively. Increasing the AGC target value to 2×10^6 and 3×10^6 ions results in a lower number of identified peptides, unique peptides and identified protein groups. In addition, at higher AGC target values ($\geq 2 \times 10^6$) some mass deviation on precursor ions' mass accuracy is observed due to space charge effects and the percentage of identified MS/MS spectra decreases.[16]

For the MS/MS AGC target values comparable results are obtained for $3 \times 10^3 - 1 \times 10^4$ ions resulting in 1035 to 1086 protein groups. Significantly lower identification rates for both peptides and protein groups are observed for ion populations of 5×10^4 and 1×10^5. By increasing the AGC target value for MS/MS, the MS/MS scan rates decrease as it takes longer to accumulate a higher ion population resulting in a lower number of acquired MS/MS. In addition, setting the AGC target value very low at 1×10^3 ions also results in lower identification rates, presumably due to the lower quality of MS/MS spectra. This assumption is also supported by the fact that the percentage of identified MS/MS spectra decreases at AGC target values of 1×10^3.[16] Based on these observations it was concluded that an ion population of 5×10^5 to 1×10^6 for MS scans and an ion population of 3×10^3 to 1×10^4 for MS/MS events provide optimal results.[16]

Maximum Ion Injection Time for MS/MS events

The maximum ion injection time is the maximum time that ions are allowed to accumulate in the linear ion trap or the C-trap, before being transferred to the Orbitrap, provided that the AGC target value has not been reached first as discussed above. Setting the ion injection time very high (e.g.; > 1000 ms) can result in significantly longer scan cycle times in cases of weak or no signals because in such circumstances the AGC target value is not reached and the MS/MS events will be performed based on the selected maximum ion injection time.

The effect of the maximum ion injection time on proteome coverage has been previously evaluated.[13,16] In general, it has been shown that lower ion injection times of 50–150 ms are correlated with higher identification rates. For instance, for an MS/MS AGC target value of 8×10^3 an ion injection time of 80 ms resulted in more peptide and protein identifications when compared with an ion injection time of 400 ms.[13] For ion population of 5×10^3 and 1×10^4, optimal results were obtained in the range of 50–100 and 50–150 ms, respectively.[16] Using an ion population of 5×10^3 for MS/MS and an ion injection time of 50 ms, 1049 protein groups and 7913 peptides were identified, versus 1008 proteins groups and 7236 peptides at 500 ms ion injection time.[16] Ion injection times above 100 or 150 ms did not provide any benefit and overall a decrease in identification rates was observed.

It is worth noting that electrospray conditions significantly affect ion injection times, therefore considerable attention should be given in achieving optimized and stable electrospray conditions. Electrospray emitter tips and ESI voltages should be carefully chosen. Under optimized electrospray conditions and sufficient sample concentration the

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

actual MS and MS/MS injection times should be low. As illustrated in Table 2 the actual MS and MS/MS ion injection times are low, except in cases in which the AGC target values are set at high values. Newer generation visualization tools, such as LogViewer[18] (http://pel.caltech.edu/software/) or RawMeat (http://vastsci.com/rawmeat/), which display ion injection times in addition to several other parameters, can be used to assess whether the MS/MS events were performed based on the AGC target value or the ion injection time and also assist in the optimization process.

It should be noted that in cases in which the MS/MS events are analyzed in the Orbitrap mass analyzer, injection times are considerably higher when compared to those analyzed in the ion trap. This is mainly due to the fact that a higher number of ions are required for detection in the Orbitrap and thus longer ion injection times are needed for accumulation of the desired number of ions.

Additional MS Parameters during DDA Mode

Besides the DDA settings discussed above, there are several other MS and MS/MS parameters commonly used during DDA acquisition. These parameters are briefly discussed below.

The isolation width is the m/z range that the MS detector uses to isolate the precursor ions. Setting the isolation width too narrow will result in loss of sensitivity, and in extreme cases cut out a portion of the ion packets, thus resulting in inaccurate measurements. In contrast, a wide isolation window may result in co-isolation and co-fragmentation of neighboring peptides resulting in unidentifiable or low scoring spectra. The isolation width is usually set at 2 to 4 m/z[10,14,15,19–24] and it has been suggested that that these typically used values (2 to 4 m/z) are optimal in practice.[25]

For shotgun proteomics experiments, the collision energy and activation time for CID are generally set at 35% and 30 ms,[11,14,15,19–21,23,26–30] respectively. In addition, it has been shown that the collision energy does not significantly affect identification rates.[13]

In terms of charge states selected for fragmentation, usually singly charged ions and unassigned charge states are rejected.[10,13,20,21,23,27,29–33] This is mainly due to the fact that the majority of peptides derived from proteolytic digestions (i.e; trypsin, Lys C, Glu C) are multiply charged following electrospray ionization. Although some proteolytically derived peptides are singly charged these are in general small in size. Given that during database searching a minimum peptide length of 6 or 7 amino acids is required for confident peptide identification these small size peptides will be filtered out and therefore collecting MS/MS spectra on these peptides provides no benefit. In addition, singly charged species are more likely to be contaminant ions and thus provide no useful information. By excluding them from fragmentation, all available time for MS/MS is devoted to multiply charged ions and no analysis time is wasted in collecting uninformative MS/MS scans or MS/MS spectra, which will be filtered during data analysis.

Dynamic exclusion duration is another commonly used MS parameter and it can either be enabled or disabled. If enabled, an exclusion duration value should be determined by the user. Ideally, the dynamic exclusion should be matched to the peak width of the individual peaks in a chromatogram. With dynamic exclusion enabled a precursor ion for which an MS/MS event has been triggered is placed on an exclusion list and will be excluded for repeated MS/MS analysis for a specified time, called exclusion duration. The exclusion list on all Orbitrap mass spectrometers is currently limited to 500 precursor ions. When the dynamic exclusion is disabled, repeated MS/MS scans of the most abundant precursor ions will be generated. In shotgun proteomics experiments the dynamic exclusion is normally

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

enabled to minimize repeated sequencing of peptides and allow for new precursor ions to undergo fragmentation. Commonly used values for dynamic exclusion duration are 60, 90 or 180 sec with one repeat count and 30 sec repeat duration.[3,10,11,15,19,21,25,27,34,35] Muddiman and coworkers evaluated the effect of dynamic exclusion duration by comparing 30 and 180 sec and found that dynamic exclusion duration had a minimal effect on identification rates.[13] Two additional parameters of the dynamic exclusion settings are the exclusion mass width and early expiration. In our laboratory we set the exclusion mass width to 10 or 15 ppm and enable early expiration as these settings, in our hands, result in higher identification rates.

Orbitrap Elite

The hybrid Orbitrap Elite[36] is the latest generation of the Orbitrap mass analyzers family. This instrument features higher resolving power, up to 240,000 at m/z 400, compared to the 100,000 resolving power of the earliest generation. In addition, the Orbitrap Elite offers improved dynamic range and faster scan rates than its predecessors. A detailed description of hardware and software improvements implemented in the Orbitrap Elite can be found in a recent publication.[36]

We felt that this tutorial would not be complete without covering and presenting the optimized DDA settings for the hybrid Orbitrap Elite. Therefore, we performed a similar analysis to that performed for the Orbitrap Classic and XL[13–16] to determine optimized MS and MS/MS settings for the Orbitrap Elite during DDA mode. For direct comparison we used tryptic digest of unfractionated yeast lysate. The parameters examined, with their recommended settings, are summarized in Table 3.

Several of the optimized settings determined for the Orbitrap Classic and XL are easily transferable to the Orbitrap Elite. For instance, the optimized settings for the AGC target for the MS and MS/MS scans determined for the Orbitrap Elite are identical to those determined for the Orbitrap Classic and XL. The effect of maximum ion injection time on identification rates is shown in Figure 3. These data clearly demonstrate that the observed trends are similar to those obtained for the Orbitrap Classic.[16] Although, differences in the identification rates are not statistically significant (ANOVA analysis resulted in p-values of 0.55), there is no apparent advantage of increasing the ion injection time, as discussed above. Therefore, we recommend setting the ion injection time at 30 to 100 ms for the MS/MS scans.

Based on the data obtained on the Orbitrap Classic,[16] for the number of MS/MS scans we examined only 15 and 20 MS/MS events. We have not observed any significant increase in the number of identified peptides or protein groups by increasing the number of MS/MS to 20. The number of identified peptides, unique peptides and protein groups was 14706, 13196 and 2151, respectively, with 15 MS/MS. With 20 MS/MS scans 14008 peptides, 12703 unique peptide and 2111 protein groups were identified in unfractionated yeast samples. It should be pointed out that for more complex samples, such as an unfractionated peptide mixture from human or mouse cell lines, increasing the number of MS/MS to 20 might be advantageous. In fact, in previous reports in which a peptide mixture from a tryptic HeLa digest was analyzed, 20 MS/MS events were selected for CID performed on the LTQ-Orbitrap Velos[30] and Orbitrap Elite.[36]

A major improvement of the Orbitrap Elite is the reduced trap dimension of the Orbitrap mass analyzer leading to a higher field strength. The higher field strength results in an almost twofold increase in resolving power at the same scan time compared to the previous Orbitrap generations. Therefore, the MS scans can be performed at higher resolving power without significantly increasing the cycle times. In addition to achieving better mass accuracy,[4] a higher number of isotope patterns can be detected[36] by acquiring data at higher

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

resolving power. Since the improvement in resolving power constitutes a major difference between the Orbitrap Elite and its predecessors, the effect of MS resolving power on identification rates was reevaluated to assess the optimum value for the Orbitrap Elite. Increasing the MS resolving power from 30,000 to 120,000 results in a concomitant increase in identification rates by 21% for protein groups and by 30% for identified peptides (Figure 4). Further increasing the resolving power to 240,000 causes a decrease in the number of identified peptides and protein groups. The decrease in identification rates observed at 240,000 resolving power can be attributed to the lower number of acquired MS/MS scans (37,369 ± 411 at 240,000 resolution versus 38,950 ± 275 at 120,000 resolution) and to the lower percentage of identified MS/MS spectra (38% ± 0.39 at 240,000 resolution versus 43% ± 0.39 at 120,000 resolution). ANOVA analysis revealed that differences in identification rates were statistically significant as the obtained *p*-values were 0.00083, 0.00077 and 0.00017 for the number of identified peptides, unique peptides and protein groups, respectively. Therefore, for the Orbitrap Elite a resolving power of 120,000 provides optimum results, whereas for the previous Orbitrap generation instruments optimal results are obtained at a resolving power of 30, 000 or 60,000.

It is worth pointing out that for the analysis of a tryptic HeLa digest on the Orbitrap Elite it was shown that a resolving power of 240,000 resulted in a higher number of isotope clusters being detected and therefore a 240,000 resolution survey scan was selected for subsequent analysis.[36] These observations[36] are in contrast with our observations, in which we determined that a resolving power of 120,000 resulted in higher identification rates. One plausible explanation for this discrepancy could be the different sample employed in the two investigations. The digested HeLa lysate is more complex compared to the yeast lysate we used in our investigations. It is reasonable to assume that the higher resolving power is advantageous for more complex samples as the complexity of the sample will result in an increase in the number of peptides eluting at a given time and therefore in an increase in the number of peptide features (isotope clusters). In addition to different sample complexity, different sample amount and gradient time were used and also MS settings were not identical between the two investigations. These factors might contribute to the different resolving power setting determined as optimum between the two reports.

The Orbitrap Elite features two different CID scan modes, one with rapid scan rate and the other with normal scan rate. As the name implies, the rapid scan mode is faster than the normal scan mode. With the rapid scan mode up to 12.5 MS/MS spectra can be acquired per second.[36] The data obtained by selecting the rapid and normal scan rata is displayed in Table 4A. A higher number of MS/MS spectra are collected with the rapid scan rate, as expected, but this does not translate to significantly higher identification rates. Nevertheless, there is no drawback of selecting the rapid scan mode and, in fact, it might be advantageous for very complex samples as it allows more MS/MS scans to be acquired.

Increasing the number of fragmentation events can also be achieved by enabling the prediction for ion injection time. By enabling this option the Orbitrap MS analyzer skips the data-dependent MS/MS prescan and uses the intensity of the parent ion present in the MS scan to determine the ion injection time for the MS/MS events. Therefore, more MS/MS scans can be triggered across a chromatographic peak. Table 4B displays the data obtained by enabling and disabling this option. Indeed by enabling the ion prediction time, the number of fragmentation events increases by 19 % and a higher number of identified peptides and protein groups is observed compared to the data obtained by disabling this option.

Similar to the results obtained on the Orbitrap XL[13], we observed that enabling the monoisotopic precursor selection results in higher identification rates on the Orbitrap Elite.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

More specifically by enabling this option an increase of 18 % was observed in the number of identified proteins. We also examined the effect of the exclusion mass width. By setting the exclusion mass width to 0.51–1.51 m/z, we observed a decrease of 26% in the number of identified protein groups compared to the results obtained by setting the exclusion mass width to 15 ppm.

Finally, for the Orbitrap Elite, at a resolving power of 120,000, disabling the preview mode for FTMS results in higher number of peptide and protein groups identifications. Therefore, disabling this option is recommended for the Orbitrap Elite.

In Table 5, the data obtained over three replicate LC-MS/MS analyses performed under the same experimental conditions is summarized. Small run-to-run variations are observed in terms of identification numbers indicating good repeatability. The data-dependent settings used were 1×10^6 ion population for the MS scan and ion injection time of 200 ms, 5×10^3 ion population for the MS/MS scan and ion injection time of 50 ms. The resolving power was set to 120,000 and the 15 most intense ions (15 MS/MS scans) were selected for fragmentation. The rapid scan mode was selected, the prediction of ion injection time was enabled and the preview mode for the FTMS scan was disabled. Dynamic exclusion was enabled with a repeat count of 1, repeat duration of 50, exclusion list size of 500 and exclusion duration of 90 sec. Exclusion mass width was set to 15 ppm and early expiration was enabled. CID was performed at a normalized collision energy of 35% and the isolation width was set to 2 Da. With these data-depending settings an average of 14,207 (± 1,095) peptides and 2,113 (± 78) protein groups (≥2 peptides) were identified.

A direct comparison of the Orbitrap Classic and Orbitrap Elite is displayed in Figures 5 and 6. As it is clearly illustrated in these Figures, the Orbitrap Elite outperforms the Orbitrap Classic in terms of identification rates. A two fold increase in the number of identified protein groups and identified peptides is observed with the Orbitrap Elite (Figure 5). For proteins with more than 10,000 copies per cell[37], the number of identified proteins is comparable between the two instruments but for proteins with 100–10000 copies per cell[37] the Orbitrap Elite identified 942 more proteins corresponding to a 2.7 fold increase compared to the Orbitrap-Classic (Figure 6). It is also worth pointing out that for the LC-MS/MS analysis on the Orbitrap Classic, 1.2 μg of unfractionated yeast lysate was used versus 0.2 μg for the Orbitrap Elite. Therefore, in addition to the increase in identification rates, an enhanced sensitivity is also achieved by the Orbitrap Elite.

Q Exactive and HCD

The Q Exactive[38], introduced in 2011, represents another member of the Orbitap mass analyzers family. The Q Exactive features a quadrupole mass filter at the front end and unlike the LTQ-Orbitrap mass analyzers it is not a dual detector instrument (the Orbitrap mass analyzer is the only detector on the Q Exactive configuration) and thus, it does not have CID capabilities. HCD, in which both the precursor and product ions are analyzed in Orbitrap analyzer is, currently, the only available fragmentation technique on the Q Exactive. The quadrupole mass filter allows transfer of specified mass range ions into the C-trap for accumulation improving sensitivity for MS/MS experiments. In addition, the Q Exactive provides high mass accuracy, low detection limits, resolving power up to 140,000 at m/z 200 (note that this is in contrast to the LTQ-Orbitrap instruments where the resolving power is defined at m/z 400) and allows up to 12 high-resolution MS/MS spectra per second. These features make the Q Exactive a suitable instrument platform for analysis of complex proteomes.[38–40]

The performance of the Q Exactive was evaluated using 5 μg of unfractionated tryptic digests of HeLa cell lysate over a 90 min gradient.[38] HCD was performed at collision energy

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

of 30 eV and at a resolution of 17,500 (at m/z 200). Ten data-dependent MS/MS scans were selected per MS scan. The other MS and MS/MS data-dependent settings used are displayed in Table 7. On average 12,563 unique peptides and 2,557 protein groups were identified. The total number of identified unique peptides and protein groups over a triplicate analysis were 16,255 and 2,864, respectively.[38]

The Q-Exactive was recently coupled with ultra-high pressure liquid chromatography for the analysis of unfractionated Lys-C digested yeast cell lysate.[40] HCD was performed on the 10 most abundant precursor ions with a collision energy of 25 eV, maximum ion injection time of 60 ms and AGC target value of 1×10^6. In order to increase proteome depth a long 4h gradient using a 50 cm analytical LC column and small particle sizes (1.8 μm) was used. Six LC-MS/MS analyses were performed and 4 μg of the unfractionated peptide mixture was loaded per run. An average of 4,084 ± 8 and a total of 4,206 proteins were identified with 4,026 proteins being identified by at least two peptides. These results demonstrated that the Q Exactive with its HCD capabilities is well suited for the analysis of complex peptide mixtures providing high proteome coverage.[40]

Furthermore, it has been demonstrated that peptide and protein identification rates on the Q-Exactive are dependent on sample amount, acquisition speed and data quality.[39] These observations are analogous to the results observed for the LTQ-Orbitrap mass analyzers. Two acquisition HCD methods, one optimized for sensitivity (termed sensitive method) and the other optimized for speed (termed “fast” method), were evaluated for the analysis of yeast and HeLa tryptic peptide mixtures. For the sensitive acquisition method the ion injection time was set at 120 ms and the resolution was 35,000, whereas for the fast method the ion injection time was 60 ms and the resolution was set at 17,500. For both methods the AGC target value for MS/MS was set at a very high number (5×10^5) to ensure that all MS/MS events were performed at the desired maximum ion injection time. The 12 most abundant precursor ions were selected for fragmentation. All other data-dependent acquisition settings were identical between the two methods (Table 7). The authors reasoned that the fast scanning method would be more suitable for the analysis of complex and abundant peptides, while the sensitive method would be better suited for low-complexity and low abundant samples since the longer ion injection times employed in the sensitive method will result in better quality MS/MS spectra. Indeed, at a low sample concentration (< 125 ng) of a tryptic yeast digest the sensitive method resulted in a higher number of peptide spectra matches and identified peptides, whereas at concentration ≥ 125 ng the fast method outperformed the sensitive method.[39]

The fast and sensitive acquisition methods were also evaluated for the analysis of unfractionated tryptic HeLa digests over three different gradient lengths. It was shown that for gradient lengths of 60 and 120 min the fast method resulted in a higher number of identified peptides and protein groups, whereas for a gradient length of 180 min higher identification rates were observed with the sensitive method. More specifically for the 60 min gradient, 2,109 protein groups were identified with the sensitive method and 2,814 with the fast method. For a gradient length of 120 min an average of 3,580 and 3,821 protein groups were identified with the sensitive and fast method, respectively. For the longer 180 min gradient evaluated, the sensitive method resulted in 4,097 protein groups identifications versus 3,905 protein groups identified with the fast method. Based on these observations the authors reemphasized the importance of optimizing acquisition methods based on the amount loaded and analysis time. [39]

HCD fragmentation is also available on the LTQ-Orbitrap XL, Velos and Elite. The HCD performance on the LTQ-Orbitrap Velos and Q Exactive has been compared using unfractionated HeLa lysate,[38] yeast lysate[39] and RAW 264.7 lysate[41] as the model systems.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

During a 60 min gradient of a 250 ng yeast peptide mixture, the LTQ-Orbitrap Velos identified on average 4,491 unique peptides, whereas 10,784 or 8,572 unique peptides, depending on the sensitive or fast acquisition method used (see discussion above), were identified with the Q Exactive.[39] Similar results were obtained for the analysis of 5 μg HeLa tryptic digests over a 90 min gradient.[38] Ten HCD MS/MS were selected for both instruments and the resolution was set at 17,500 (at m/z 200) on the Q Exactive and at 7,500 (at m/z 400) for the Orbitrap Velos (details on the acquisition settings are depicted on Table 7). On average 1,881 and 2,557 proteins were identified on the LTQ-Orbitrap Velos and Q Exactive, respectively. The authors pointed out though, that a detailed comparison between the two instruments requires additional experiments and also that the LTQ-Orbitrap Velos has CID and ETD capabilities, which are not available on the Q Exactive.[38] The performance of HCD between the two instrument platforms was further evaluated for different sample loading amounts (1 to 1000 ng).[41] It was shown that for sample amounts ≥ 10 ng the Q Exactive outperformed the LTQ-Orbitrap Velos. For 1 and 5 ng of RAW 264.7 cell lysate, depending on the HCD acquisition method used on the Q Exactive, identification rates were found to be higher or lower compared to the LTQ-Orbitrap Velos. These observations reemphasize the importance of optimizing acquisition settings. The authors also compared CID on the LTQ-Orbitrap Velos with HCD on the Q Exactive and reported that HCD outperformed CID.[41] Overall, these reports[38,39,41] demonstrated that HCD performed on the Q Exactive results in a higher number of identified peptides and proteins compared to HCD performed on the LTQ-Orbitrap Velos highlighting improvements in instrumentation. Identification rates obtained by HCD on the Orbitrap Elite were also reported by analyzing 400 ng of an unfractionated tryptic HeLa digest over a 2h gradient using 15 HCD MS/MS events per MS scan (see Table 8 for additional acquisition settings).[36] HCD resulted in an average of 10,740 unique peptides and 2,073 protein groups. Furthermore a comparison between CID and HCD on the Orbitrap Elite was also performed. The 20 most abundant precursor ions were selected for CID fragmentation and a slightly higher number of identified protein groups, 2,268, was obtained when compared to HCD. The authors pointed out that the peptide identification numbers obtained in CID and HCD performed on the Orbitrap Elite were significantly improved compared to the results obtained on the LTQ-Orbitrap Velos.[36]

Tables 6 and 7 summarize MS and MS/MS acquisition settings commonly employed by different groups for CID and HCD analyses performed on LTQ-Orbitrap Classic, LTQ-Orbitrap XL, LTQ-Orbitrap Velos, Orbitrap Elite and Q Exactive. These values could serve as a starting point when optimizing acquisition settings for proteomics analysis on different Orbitrap mass analyzers platforms.

SAMPLE PREPARATION, LIQUID CHROMATOCRAPHY AND BIOINFORMATICS ANALYSIS

Sample preparation, including protein extraction, protein solubilization/denaturation, reduction/alkylation, proteolytic digestion and sample cleanup prior to LC-MS/MS analysis, is a fundamental step for successful peptide and protein identifications in shotgun proteomics. In fact, the quality of the sample to be analyzed largely influences the data obtained in a proteomics experiment and thus considerable attention should be given during sample preparation. It should be pointed out that there is no single method, which is applicable to all possible samples. Therefore, the sample preparation method should be carefully and experimentally optimized based on the nature of the sample and on the biological question to be addressed. For instance, different sample processing protocols are followed for protein extraction from i.e. cell lysates, tissues, biological fluids, plants and bacteria. In addition, certain protein classes such as membrane proteins and post-translationally modified proteins require additional steps, such as enrichment procedures and

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

protein solubilization, for successful identification. The importance of optimizing sample preparation has been demonstrated by Winter and Steen, who have carefully optimized the cell lysis, protein precipitation and digestion protocols for the widely used HeLa S3.[42] The authors demonstrated that improved peptide and protein identification rates can be achieved using an optimizing protocol. For instance, by optimizing the digestion protocol the authors were able to obtain a 3.5 and 2.7 fold increase in the number of identified peptides and proteins, respectively.[42]

Description on sample preparation strategies based on the nature of the sample and the properties of the target proteins is out of the scope of this tutorial and discussion on these topics can be found in published reports and references therein.[42–60]

In this tutorial, we briefly discuss several steps during sample preparation that are generally applicable to a wide range of proteomics samples. Possible problems and ways to overcome these are also discussed. Regardless of the nature of the sample, an efficient proteolytic digestion producing peptides with high yields, preferentially for the vast majority of proteins present, undoubtedly facilitates and improves protein identification. Many proteases i.e. trypsin, Lys-C, Lys-N, Glu-C, Asp-N, chymotrypsin and pepsin, are available for protein digestion and each enzyme features its own specificity, efficiency and optimum digestion conditions. Missed cleavages, unspecific cleavages and incomplete digestion are factors that need to be considered during the digestion step and database search. Time should be devoted in optimizing digestion conditions such as enzyme to substrate ratio, temperature, reaction time, buffer and pH in order to minimize missed cleavages, unspecific cleavages and incomplete digestion.

Enzymatic missed cleavages are due to incomplete hydrolysis reactions between the substrates and enzymes. The number of missed cleavages will negatively affect peptide identification rates in cases in which mis-cleaved peptides are present but the search was performed with zero missed cleavages. Brownridge and Beynon analyzed mis-cleaved peptides in a trypsin digest of yeast extract and reported that 1,415 ± 88 mis-cleaved peptides were identified over four biological replicates. Of these 51% were N-terminal mis-cleaved peptides and 32% were C-terminal containing short extensions (up to five amino acids) at the N-terminus and C-terminus respectively.[61] Furthermore, it has been shown that the digestion temperature, reaction time, enzyme to substrate ratio and digestion buffer affect the number of missed cleavages during conventional and microwave-assisted trypsin digestion.[62] To account for the presence of mis-cleaved peptides a maximum of two missed cleavages is commonly allowed in database searches. Incomplete digestion may arise from factors such as limited protein solvation or denaturation, short reaction times, low enzyme concentration, digestion solution pH, and low protein accessibility to the enzyme.[63,64] It is highly recommended to quantify the amount of the protein to be digested in order to allow determination of the appropriate enzyme amount for protein digestion. Incomplete protein digestion can lead to generation of low abundance peptides which will not be selected for fragmentation during data-dependent acquisition mode or, if selected, will result in low MS/MS spectra quality. Longer reaction times (> 30 min) under microwave-assisted digestion,[62] addition of organic solvents, surfactants and chaotropes to increase protein solubilization and denaturation,[61,63] and use of the optimum enzyme pH range during digestion[64] have all been shown to enhance completeness of the digestion. It is important to note that considerable attention should be given to addition of surfactants and chaotropes as these may interfere with LC-MS and therefore need to be removed prior to LC-MS analysis (see discussion below).

Unspecific cleavages refer to peptides, whose termini do not contain the expected amino acid derived from specific protease cleavage. These may arise from protease impurities,

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

contamination with other proteases and nonspecific proteolysis during sample preparation.[65] In addition, incubation time, digestion buffers, digestion temperature,[65] and enzyme to protein ratio,[66] have also been shown to affect enzyme specificity. Recently, the specificity of Glu-C, chymotrypsin and CNBr digestion was assessed by Pevzner and coworkers using *Shewanella oneidensis* as the model system and the frequency of unspecific cleavages was reported and evaluated.[67] The proteolytic cleavage specificity of Lys-C and Lys-N was also examined by digesting HEK293 cell lysates under optimized digestion conditions.[68] The authors reported 90% enzyme specificity for Lys-C and 72% for Lys-N.[68]

Trypsin has gained popularity in proteomics not only due to its high specificity but also due to the generation of suitable size peptides (10–15 amino acids residues long) for MS analysis, and the fact that tryptically digested peptides produce doubly charged precursor ions, which fragment efficiently in CID.[61,64,69] Several protocols have been developed to improve and enhance trypsin digestion using the addition of denaturation and solubilization agents, high hydrostatic pressure, immobilized enzymes, infrared irradiation and microwave irradiation.[63,64,70–81] In addition, the cleavage specificity, efficiency and reproducibility of several different types of commercially available trypsin were evaluated.[82] The authors concluded that the sequencing grade modified trypsin from Promega provided optimal results.[82] Double digestion, Lys-C digestion followed by trypsin digestion, is very common in proteomics especially in cases in which high concentration of urea ~8 M is used for proteins denaturation. Trypsin has a reduced activity at 8 M urea, whereas Lys-C retains its activity at 8 M urea. Before trypsin addition, urea concentration is diluted to 2 M. This treatment generally results in good sequence coverage with limited missed cleavages.

In addition to urea, SDS is also widely used for protein denaturation and solubilization. An acid labile detergent, RapiGest SF,[83] is also gaining popularity in proteomics sample preparation as it can be easily removed from the sample and it is compatible with proteolytic enzymes. An increase in protein identifications has been reported in the presence of RapiGest SF *versus* other commonly employed detergents.[42,63] Additionally, the performance of three commercially available MS compatible surfactants (Invitrosol, PPS silent surfactant and RapiGest SF) was evaluated in both aqueous and organic-aqueous digestion solutions.[84]

Identification rates and proteome sequence coverage can be improved by the use of multiple proteases with varying specificity, as it was illustrated for the yeast proteome[85] and plasma.[86] Further increase in proteome coverage can be obtained by sample fractionation followed by a second dimension separation achieved by LC-MS/MS analysis of each individual fraction. Pre-fractionation reduces sample complexity and consequently decreases undersampling during LC-MS/MS analysis as fewer analytes are simultaneously introduced into the mass spectrometer.[87] Pre-fractionation techniques should preferentially be orthogonal to reversed phase chromatography as the latter is usually employed for on-line LC separation. Several chromatographic and electrophoretic approaches are used for sample pre-fractionation,[87–89] such as SDS page, strong cation exchange, strong anion exchange, hydrophilic interaction, and isoelectric focusing, The performance of different pre-fractionation techniques in terms of proteome coverage and reproducibility has been previously evaluated for a wide range of samples.[19,90–94] It is important to note that the effectiveness of the fractionation technique will be largely dependent on the nature of the sample. Furthermore, it should be pointed out that pre-fractionation reduces sample throughput, but at the same time improved proteome coverage is achieved.

Sample cleanup is another factor that needs to be considered for obtaining high quality proteomics data. Salts, buffers and detergent contaminants, such as Tris•HCl, NaCl, urea and sodium dodecyl sulfate (SDS) interfere with the LC-MS/MS analysis, causing column

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

clogging, decreasing quality of MS and MS/MS spectra, reducing ESI sensitivity due to ion suppression, and therefore avoidance and/or removal prior to analysis is a prerequisite for a successful experimental outcome.[95] A protocol that allows fast cleanup of samples containing salts, detergents and buffers has been developed.[96,97] This protocol utilizes spin filter devices for digestion to allow resuspension of proteins in buffers compatible with digestion and removal of detergents or salts prior to MS analysis.[96,97] In agreement with these reports we also observed that digestion utilizing spin filter resulted in higher identification rates compared to conventional in-solution digestion. Recently, Bereman et.al. introduced the use of self-packed spin columns to effectively remove SDS from complex proteomics samples (Saccharomyces cerevisiae, Caenorhabditis elegans lysate and HEK293T cells), prior to LC-MS/MS analysis.[98] Efficient removal of a wide range of commonly used detergents in proteomic, including SDS, CHAPS, lauryl maltoside, octyl and glucoside, was demonstrated for a commercially available spin column.[99] Alternatively, offline LC purification can be used for sample cleanup, allowing for independent UV quantification of the sample.[95]

Further discussion on methodologies and strategies for sample preparation in proteomics as well as advantages and disadvantages associated with these methods can be found in previously published reports.[51,64,69,76] In addition, available digestion techniques as well as recent developments to accelerate the digestion step have been very recently reviewed and discussed in details by Switzar et.al.[54] Approaches specifically designed for certain protein classes, which are difficult to analyze with conventional digestion approached, were also presented.[54]

In addition to the MS data acquisition and sample preparation, LC performance is also important in proteomics analysis. The performance of the LC system can be affected by the column material, improper fluidic connections, leakages or blockages in the system, flow rate,[10] elution gradient and length,[10] sample loading amount,[9,10] sample loading conditions[11] and sample carryover.[100] Optimization and evaluation of several of these parameters have been previously performed.[9–11,100] It has been demonstrated that sample loading amount affects the identification rates[9,10] and that the optimal gradient length depends on the amount of sample loaded.[10] In addition it was shown that sample loading conditions can affect identification rates depending on the column packing material used for the analysis.[11] Optimal flow rates and elution gradient also lead to higher identification rates.[10] In agreement with these observations we also observed an increase in identifications by optimizing the LC-gradient, column material and sample loading amount.

Given the very large amount of data generated in MS-based proteomics experiments, peptide and protein identification is almost exclusively performed through automated database searching. Several database search algorithms, including Mascot,[101] X!Tandem,[102] OMSSA,[103] Sequest,[104] Andromeda,[105] and ROCCIT (roccit.caltech.edu) are available for translating MS/MS spectra into peptides sequences. In our laboratory, due to its advanced signaling processing capabilities for Thermo raw files, we generally use MaxQuant for quantitative bottom-up data analysis.

The general principle of most search engines is the selection of candidate peptides from the sequence database whose theoretical masses match the experimental masses determined from the MS spectra. For each candidate peptide a theoretical MS/MS spectrum is created and compared to the experimental MS/MS spectrum to calculate a similarity score between the two spectra and the top scoring peptide is reported.[106–110]

Due to the widespread use and critical role of database search algorithms in proteomics, several reports have focused on systematic evaluation and comparison of frequently used

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

search engines,[106,111] performance of search algorithms under different mass tolerance settings,[107] and description of the rationale and basic concepts used by database search tools.[108,112,113] In addition, considerations and recommendations regarding the search parameters, search strategy, protein sequence database used, chosen search engine,[108,113,114] and data interpretation[115] have been reviewed.

A major challenge of automated database searching is to ensure high confidence peptide and protein identifications since the top scoring peptide assigned by the search engines does not always represent the correct identification.[107–111,116–118] A widely used method to assess the confidence in the set of peptide identifications for large-scale data sets is the use of false discovery rate. False discovery rate (FDR) is an estimation of the number of incorrect identifications among all accepted assignments and allows data filtering by selecting a desired FDR cutoff. A commonly used strategy for evaluating FDRs is the target decoy approach.[119] The use and compatibility of target decoy approach with several popular database search tools has been recently discussed and reevaluated by Pevzner and coworkers.[120]

CONCLUSIONS

In this tutorial we have described and discussed important MS and MS/MS settings widely employed during DDA mode on LTQ-Orbitrap mass analyzers. We presented recommended values for these data-dependent settings determined by systematic investigation and evaluation. [13–16] Data from the Orbitrap Classic, XL, and the newly introduced Orbitrap Elite was presented. While several settings are easily transferable to the Orbitrap Elite from the previous generation Orbitrap mass analyzers, others needed to be reexamined and reevaluated. Discussion on the Q Exactive and HCD was also given. In addition, sample preparation, LC optimization and bioinformatics analysis, which are also vital for a successful proteomics experiment, were also discussed. The discussion presented here, strongly suggests that by optimizing sample preparation, MS and LC instruments performance and bioinformatics analysis, significantly improved analysis of complex proteomic samples can be achieved. We hope that this tutorial will provide useful insights into understanding and optimizing important variables for shotgun proteomics experiments.

Acknowledgments

We thank Tonya P. Second, Rolly Singh, Anjum A. Khan and Martin Zeller at Thermo Fisher Scientific for their valuable discussions and help. The Proteome Exploration Laboratory is supported by the Gordon and Betty Moore Foundation through Grant GBMF775, the Beckman Institute and an instrumentation grant from NIH (10565784).

REFERENCES

1. Stahl DC, Swiderek KM, Davis MT, Lee TD. Data-controlled automation of liquid chromatography tandem mass spectrometry analysis of peptide mixtures. J. Am. Soc. Mass Spectrom. 1996; 7(6): 532–540.
2. Makarov A, Denisov E, Kholomeev A, Baischun W, Lange O, Strupat K, Horning S. Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. Anal. Chem. 2006; 78(7):2113–2120. [PubMed: 16579588]
3. Yates JR, Cociorva D, Liao LJ, Zabrouskov V. Performance of a linear ion trap-orbitrap hybrid for peptide analysis. Anal. Chem. 2006; 78(2):493–500. [PubMed: 16408932]
4. Scigelova M, Hornshaw M, Giannakopulos A, Makarov A. Fourier transform mass spectrometry. Mol. Cell. Proteomics. 2011; 10(7):1–19.
5. Perry RH, Cooks RG, Noll RJ. Orbitrap mass spectrometry: instrumentation, ion motion and applications. Mass Spectrom. Rev. 2008; 27(6):661–699. [PubMed: 18683895]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

6. Hu QZ, Noll RJ, Li HY, Makarov A, Hardman M, Cooks RG. The Orbitrap: a new mass spectrometer. J. Mass Spectrom. 2005; 40(4):430–443. [PubMed: 15838939]
7. Makarov A. Electrostatic axially harmonic orbital trapping: A high-performance technique of mass analysis. Anal. Chem. 2000; 72(6):1156–1162. [PubMed: 10740853]
8. Hardman M, Makarov AA. Interfacing the orbitrap mass analyzer to an electrospray ion source. Anal. Chem. 2003; 75(7):1699–1705. [PubMed: 12705605]
9. Liu KH, Zhang JY, Wang JL, Zhao LY, Peng X, Jia W, Ying WT, Zhu YP, Xie HW, He FC, Qian XH. Relationship between sample loading amount and peptide identification and Its effects on quantitative proteomics. Anal. Chem. 2009; 81(4):1307–1314. [PubMed: 19146458]
10. Xu P, Duong DM, Peng JM. Systematical optimization of reverse-phase chromatography for shotgun proteomics. J. Proteome Res. 2009; 8(8):3944–3950. [PubMed: 19566079]
11. Peterson A, Hohmann L, Huang L, Kim B, Eng JK, Martin DB. Analysis of RP-HPLC loading conditions for maximizing peptide identifications in shotgun proteomics. J. Proteome Res. 2009; 8(8):4161–4168. [PubMed: 19601632]
12. Zhang Y, Wen ZH, Washburn MP, Florens L. Effect of dynamic exclusion duration on spectral count based quantitative proteomics. Anal. Chem. 2009; 81(15):6317–6326. [PubMed: 19586016]
13. Andrews GL, Dean RA, Hawkridge AM, Muddiman DC. Improving proteome coverage on a LTQ-Orbitrap using design of experiments. J. Am. Soc. Mass Spectrom. 2011; 22(4):773–783. [PubMed: 21472614]
14. Wong CCL, Cociorva D, Venable JD, Xu T, Yates JR. Comparison of different signal thresholds on data dependent sampling in Orbitrap and LTQ mass spectrometry for the identification of peptides and proteins in complex mixtures. J. Am. Soc. Mass Spectrom. 2009; 20(8):1405–1414. [PubMed: 19467883]
15. Kim MS, Kandasamy K, Chaerkady R, Pandey A. Assessment of resolution parameters for CID-based shotgun proteomic experiments on the LTQ-Orbitrap mass spectrometer. J. Am. Soc. Mass Spectrom. 2010; 21(9):1606–1611. [PubMed: 20483638]
16. Kalli A, Hess S. Effect of mass spectrometric parameters on peptide and protein identification rates for shotgun proteomic experiments on an LTQ-orbitrap mass analyzer. Proteomics. 2012; 12(1): 21–31. [PubMed: 22065615]
17. Riter LS, Vitek O, Gooding KM, Hodge BD, Julian RK. Statistical design of experiments as a tool in mass spectrometry. J. Mass Spectrom. 2005; 40(5):565–579. [PubMed: 15880604]
18. Sweredoski JM, Smith TG, Kalli A, Graham LJR, Hess S. LogViewer: A software tool to visualize quality control parameters to optimize proteomics experiments using Orbitrap and LTQ-FT mass spectrometers. J. Biomol. Tech. 2011; 22(4):122–126. [PubMed: 22131886]
19. Slebos RJC, Brock JWC, Winters NF, Stuart SR, Martinez MA, Li M, Chambers MC, Zimmerman LJ, Ham AJ, Tabb DL, Liebler DC. Evaluation of strong cation exchange versus isoelectric focusing of peptides for multidimensional liquid chromatography-tandem mass spectrometry. J. Proteome Res. 2008; 7(12):5286–5294. [PubMed: 18939861]
20. Mesmin C, Fenaille F, Becher F, Tabet J-C, Ezan E. Identification and characterization of apelin peptides in bovine colostrum and milk by liquid chromatography-mass spectrometry. J. Proteome Res. 2011; 10(11):5222–31. [PubMed: 21939284]
21. Scherl A, Shaffer SA, Taylor GK, Hernandez P, Appel RD, Binz PA, Goodlett DR. On the benefits of acquiring peptide fragment ions at high measured mass accuracy. J. Am. Soc. Mass Spectrom. 2008; 19(6):891–901. [PubMed: 18417358]
22. Yu L, Tan Y, Tsai Y, Goodlett DR, Polfer NC. On the relevance of peptide sequence permutations in shotgun proteomics studies. J. Proteome Res. 2011; 10(5):2409–2416. [PubMed: 21413817]
23. Nagaraj N, Mann M. Quantitative analysis of the intra- and inter-individual variability of the normal urinary proteome. J. Proteome Res. 2010; 10(2):637–645. [PubMed: 21126025]
24. Poliakov A, Russell CW, Ponnala L, Hoops HJ, Sun Q, Douglas AE, van Wijk KJ. Large-scale label-free quantitative proteomics of the pea aphid-buchnera symbiosis. Mol. Cell. Proteomics. 2011; 10(6)
25. Michalski A, Cox J, Mann M. More than 100,000 detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data-dependent LC-MS/MS. J. Proteome Res. 2011; 10(4):1785–1793. [PubMed: 21309581]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

26. Lu B, Motoyama A, Ruse C, Venable J, Yates JR III. Improving protein identification sensitivity by combining MS and MS/MS information for shotgun proteomics using LTQ-Orbitrap high mass accuracy data. Anal. Chem. 2008; 80(6):2018–2025. [PubMed: 18275164]
27. Gokce E, Andrews GL, Dean RA, Muddiman DC. Increasing proteome coverage with offline RP HPLC coupled to online RP nanoLC-MS. J. Chromatogr. B. 2011; 879(9–10):610–614.
28. Tayo LL, Lu B, Cruz LJ, Yates JR III. Proteomic analysis provides insights on venom processing in Conus textile. J. Proteome Res. 2010; 9(5):2292–2301. [PubMed: 20334424]
29. Frese CK, Altelaar AFM, Hennrich ML, Nolting D, Zeller M, Griep-Raming J, Heck AJR, Mohammed S. Improved peptide identification by targeted fragmentation using CID, HCD and ETD on an LTQ-Orbitrap Velos. J. Proteome Res. 2011; 10(5):2377–2388. [PubMed: 21413819]
30. Olsen JV, Schwartz JC, Griep-Raming J, Nielsen ML, Damoc E, Denisov E, Lange O, Remes P, Taylor D, Splendore M, Wouters ER, Senko M, Makarov A, Mann M, Horning S. A dual pressure linear ion trap Orbitrap instrument with very high sequencing speed. Mol. Cell. Proteomics. 2009; 8(12):2759–2769. [PubMed: 19828875]
31. Hahne H, Wolff S, Hecker M, Becher D. From complementarity to comprehensiveness - targeting the membrane proteome of growing Bacillus subtilis by divergent approaches. Proteomics. 2008; 8(19):4123–4136. [PubMed: 18763711]
32. Thakur SS, Geiger T, Chatterjee B, Bandilla P, Frohlich F, Cox J, Mann M. Deep and highly sensitive proteome coverage by LC-MS/MS without prefractionation. Mol. Cell. Proteomics. 2011; 10(8)
33. Makawita S, Smith C, Batruch I, Zheng YY, Ruckert F, Grutzmann R, Pilarsky C, Gallinger S, Diamandis EP. Integrated proteomic profiling of cell line conditioned media and pancreatic juice for the identification of pancreatic cancer biomarkers. Mol. Cell. Proteomics. 2011; 10(10)
34. Adachi J, Kumar C, Zhang YL, Olsen JV, Mann M. The human urinary proteome contains more than 1500 proteins, including a large proportion of membrane proteins. Genome Biol. 2006; 7(9):R72. [PubMed: 16901338]
35. de Souza GA, Godoy LMF, Mann M. Identification of 491 proteins in the tear fluid proteome reveals a large number of proteases and protease inhibitors. Genome Biol. 2006; 7(8):R80. [PubMed: 16948836]
36. Michalski A, Damoc E, Lange O, Denisov E, Nolting D, Muller M, Viner R, Schwartz J, Remes P, Belford M, Dunyach JJ, Cox J, Horning S, Mann M, Makarov A. Ultra high resolution linear ion trap Orbitrap mass spectrometer (Orbitrap Elite) facilitates top down LC MS/MS and versatile peptide fragmentation modes. Mol. Cell. Proteomics. 2012; 11(3)
37. Ghaemmaghami S, Huh W, Bower K, Howson RW, Belle A, Dephoure N, O'Shea EK, Weissman JS. Global analysis of protein expression in yeast. Nature. 2003; 425(6959):737–741. [PubMed: 14562106]
38. Michalski A, Damoc E, Hauschild JP, Lange O, Wieghaus A, Makarov A, Nagaraj N, Cox J, Mann M, Horning S. Mass Spectrometry-based Proteomics Using Q Exactive, a High-performance Benchtop Quadrupole Orbitrap Mass Spectrometer. Mol. Cell. Proteomics. 2011; 10(9)
39. Kelstrup CD, Young C, Lavallee R, Nielsen ML, Olsen JV. Optimized fast and sensitive acquisition methods for shotgun proteomics on a quadrupole orbitrap mass spectrometer. J. Proteome Res. 2012; 11(6):3487–3497.
40. Nagaraj N, Kulak NA, Cox J, Neuhauser N, Mayr K, Hoerning O, Vorm O, Mann M. System-wide Perturbation Analysis with Nearly Complete Coverage of the Yeast Proteome by Single-shot Ultra HPLC Runs on a Bench Top Orbitrap. Mol. Cell. Proteomics. 2012; 11(3)
41. Sun LL, Zhu GJ, Dovichi NJ. Comparison of the LTQ-Orbitrap Velos and the Q-Exactive for proteomic analysis of 1–1000 ng RAW 264.7 cell lysate digests. Rapid Commun. Mass Spectrom. 2013; 27(1):157–162. [PubMed: 23239329]
42. Winter D, Steen H. Optimization of cell lysis and protein digestion protocols for the analysis of HeLa S3 cells by LC-MS/MS. Proteomics. 2011; 11(24):4726–4730. [PubMed: 22002805]
43. Beltran L, Cutillas PR. Advances in phosphopeptide enrichment techniques for phosphoproteomics. Amino Acids. 2012; 43(3):1009–1024. [PubMed: 22821267]
44. Thingholm TE, Jensen ON, Larsen MR. Analytical strategies for phosphoproteomics. Proteomics. 2009; 9(6):1451–1468. [PubMed: 19235172]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

45. Eyrich B, Sickmann A, Zahedi RP. Catch me if you can: Mass spectrometry-based phosphoproteomics and quantification strategies. Proteomics. 2011; 11(4):554–570. [PubMed: 21226000]
46. Faca V, Pitteri SJ, Newcomb L, Glukhova V, Phanstiel D, Krasnoselsky A, Zhang Q, Struthers J, Wang H, Eng J, Fitzgibbon M, McIntosh M, Hanash S. Contribution of protein fractionation to depth of analysis of the serum and plasma proteomes. J. Proteome Res. 2007; 6(9):3558–3565. [PubMed: 17696519]
47. Nagaraj N, Lu AP, Mann M, Wisniewski JR. Detergent-based but Gel-free method allows identification of several hundred membrane proteins in single LC-MS runs. J. Proteome Res. 2008; 7(11):5028–5032. [PubMed: 18839980]
48. Fila J, Honys D. Enrichment techniques employed in phosphoproteomics. Amino Acids. 2012; 43(3):1025–1047. [PubMed: 22002794]
49. Helbig AO, Heck AJR, Slijper M. Exploring the membrane proteome-Challenges and analytical strategies. J. Proteomics. 2010; 73(5):868–878. [PubMed: 20096812]
50. Pan S, Chen R, Aebersold R, Brentnall TA. Mass spectrometry based glycoproteomics-from a proteomics perspective. Mol. Cell. Proteomics. 2011; 10(1)
51. Bodzon-Kulakowska A, Bierczynska-Krzysik A, Dylag T, Drabik A, Suder P, Noga M, Jarzebinska J, Silberring J. Methods for sample preparation in proteomic research. J. Chromatogr. B-Analytical Technologies in the Biomedical and Life Sciences. 2007; 849(1–2):1–31.
52. Wang FJ, Song CX, Cheng K, Jiang XN, Ye ML, Zou HF. Perspectives of comprehensive phosphoproteome analysis using shotgun strategy. Anal. Chem. 2011; 83(21):8078–8085. [PubMed: 21928789]
53. Tichy A, Salovska B, Rehulka P, Klimentova J, Vavrova J, Stulik J, Hernychova L. Phosphoproteomics: Searching for a needle in a haystack. J. Proteomics. 2011; 74(12):2786–2797. [PubMed: 21839867]
54. Switzar L, Giera M, Niessen MAW. Protein digestion: An overview of the available techniques and recent developments. J. Proteome Res. 2013; 12(3):1067–1077. [PubMed: 23368288]
55. Sprenger RR, Jensen ON. Proteomics and the dynamic plasma membrane: Quo Vadis? Proteomics. 2010; 10(22):3997–4011. [PubMed: 21080491]
56. Proc JL, Kuzyk MA, Hardie DB, Yang J, Smith DS, Jackson AM, Parker CE, Borchers CH. A Quantitative Study of the Effects of Chaotropic Agents, Surfactants, and Solvents on the Digestion Efficiency of Human Plasma Proteins by Trypsin. J. Proteome Res. 2010; 9(10):5422–5437. [PubMed: 20722421]
57. Callesen AK, Madsen JS, Vach W, Kruse TA, Mogensen O, Jensen ON. Serum protein profiling by solid phase extraction and mass spectrometry: A future diagnostics tool? Proteomics. 2009; 9(6):1428–1441. [PubMed: 19235169]
58. Bladergroen MR, Derks RJE, Nicolardi S, de Visser B, van Berloo S, van der Burgt YEM, Deelder AM. Standardized and automated solid-phase extraction procedures for high-throughput proteomics of body fluids. J. Proteomics. 2012; 77:144–153. [PubMed: 22842157]
59. Cordwell SJ, Thingholm TE. Technologies for plasma membrane proteomics. Proteomics. 2010; 10(4):611–627. [PubMed: 19834916]
60. Leitner A, Sturm M, Lindner W. Tools for analyzing the phosphoproteome and other phosphorylated biomolecules: A review. Anal. Chim. Acta. 2011; 703(1):19–30. [PubMed: 21843671]
61. Brownridge P, Beynon RJ. The importance of the digest: Proteolysis and absolute quantification in proteomics. Methods. 2011; 54(4):351–360. [PubMed: 21683145]
62. Reddy PM, Hsu WY, Hu JF, Ho YP. Digestion completeness of microwave-assisted and conventional trypsin-catalyzed reactions. J. Am. Soc. Mass Spectrom. 2010; 21(3):421–424. [PubMed: 20044269]
63. Klammer AA, MacCoss MJ. Effects of modified digestion schemes on the identification of proteins from complex mixtures. J. Proteome Res. 2006; 5(3):695–700. [PubMed: 16512685]
64. Capelo JL, Carreira R, Diniz M, Fernandes L, Galesio M, Lodeiro C, Santos HM, Vale G. Overview on modern approaches to speed up protein identification workflows relying on

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

enzymatic cleavage and mass spectrometry-based techniques. Anal. Chim. Acta. 2009; 650(2): 151–159. [PubMed: 19720186]

65. Lubec G, Afjehi-Sadat L. Limitations and pitfalls in protein identification by mass spectrometry. Chem. Rev. 2007; 107(8):3568–3584. [PubMed: 17645314]
66. Hara S, Rosenfeld R, Lu HS. Preventing the generation of artifacts during peptide map analysis of recombinant human insulin-like growth factor-I. Anal. Biochem. 1996; 243(1):74–79. [PubMed: 8954527]
67. Gupta N, Hixson KK, Culley DE, Smith RD, Pevzner PA. Analyzing protease specificity and detecting in vivo proteolytic events using tandem mass spectrometry. Proteomics. 2010; 10(15): 2833–2844. [PubMed: 20597098]
68. Raijmakers R, Neerincx P, Mohammed S, Heck AJR. Cleavage specificities of the brother and sister proteases Lys-C and Lys-N. Chem. Commun. 2010; 46(46):8827–8829.
69. Canas B, Pineiro C, Calvo E, Lopez-Ferrer D, Gallardo JM. Trends in sample preparation for classical and second generation proteomics. J. Chromatogr. A. 2007; 1153(1–2):235–258. [PubMed: 17276441]
70. Vaezzadeh AR, Deshusses JMP, Waridel P, Francois P, Zimmermann-Ivol CG, Lescuyer P, Schrenzel J, Hochstrasser DF. Accelerated digestion for high-throughput proteomics analysis of whole bacterial proteomes. J. Microbiol. Meth. 2010; 80(1):56–62.
71. Hervey WJ, Strader MB, Hurst GB. Comparison of digestion protocols for microgram quantities of enriched protein samples. J. Proteome Res. 2007; 6(8):3054–3061. [PubMed: 17616116]
72. Massolini G, Calleri E. Immobilized trypsin systems coupled on-line to separation methods: Recent developments and analytical applications. J. Sep. Science. 2005; 28(1):7–21.
73. Wang S, Zhang LY, Yang PY, Chen G. Infrared-assisted tryptic proteolysis for peptide mapping. Proteomics. 2008; 8(13):2579–2582. [PubMed: 18546161]
74. Sun W, Gao SJ, Wang LJ, Chen Y, Wu SZ, Wang XR, Zheng DX, Gao YH. Microwave-assisted protein preparation and enzymatic digestion in proteomics. Mol. Cell. Proteomics. 2006; 5(4):769–776. [PubMed: 16339992]
75. Ye XX, Li L. Microwave-assisted protein solubilization for mass spectrometry-based shotgun proteome analysis. Anal. Chem. 2012; 84(14):6181–6191. [PubMed: 22708679]
76. Lill JR, Ingle ES, Liu PS, Pham V, Sandoval WN. Microwave-assisted proteomics. Mass Spectrom. Rev. 2007; 26(5):657–671. [PubMed: 17474122]
77. Pramanik BN, Mirza UA, Ing YH, Liu YH, Bartner PL, Weber PC, Bose MK. Microwave-enhanced enzyme reaction for protein mapping by mass spectrometry: A new approach to protein digestion in minutes. Protein Sci. 2002; 11(11):2676–2687. [PubMed: 12381849]
78. Freeman E, Ivanov AR. Proteomics under pressure: development of essential sample preparation techniques in proteomics using ultrahigh hydrostatic pressure. J. Proteome Res. 2011; 10(12): 5536–5546. [PubMed: 22029901]
79. Lee B, Lopez-Ferrer D, Kim BC, Bin Na H, Il Park Y, Weitz KK, Warner MG, Hyeon T, Lee SW, Smith RD, Kim J. Rapid and efficient protein digestion using trypsin-coated magnetic nanoparticles under pressure cycles. Proteomics. 2011; 11(2):309–318. [PubMed: 21204257]
80. Dycka F, Bobal P, Mazanec K, Bobalova J. Rapid and efficient protein enzymatic digestion: An experimental comparison. Electrophoresis. 2012; 33(2):288–295. [PubMed: 22170586]
81. Duan JC, Liang Z, Yang C, Zhang J, Zhang LH, Zhang WB, Zhang YK. Rapid protein identification using monolithic enzymatic microreactor and LC-ESI-MS/MS. Proteomics. 2006; 6(2):412–419. [PubMed: 16342240]
82. Burkhart JM, Schumbrutzki C, Wortelkamp S, Sickmann A, Zahedi RP. Systematic and quantitative comparison of digest efficiency and specificity reveals the impact of trypsin quality on MS-based proteomics. J. Proteomics. 2012; 75(4):1454–1462. [PubMed: 22166745]
83. Yu YQ, Gilar M, Lee PJ, Bouvier ESP, Gebler JC. Enzyme-friendly, mass spectrometry-compatible surfactant for in-solution enzymatic digestion of proteins. Anal. Chem. 2003; 75(21): 6023–6028. [PubMed: 14588046]
84. Chen EI, Cociorva D, Norris JL, Yates JR. Optimization of mass spectrometry-compatible surfactants for shotgun proteomics. J. Proteome Res. 2007; 6(7):2529–2538. [PubMed: 17530876]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

85. Swaney DL, Wenger CD, Coon JJ. Value of using multiple proteases for large-scale mass spectrometry-based proteomics. J. Proteome Res. 2010; 9(3):1323–1329. [PubMed: 20113005]
86. Choudhary G, Wu SL, Shieh P, Hancock WS. Multiple enzymatic digestion for enhanced sequence coverage of proteins in complex proteomic mixtures using capillary LC with ion trap MS/MS. J. Proteome Res. 2003; 2(1):59–67. [PubMed: 12643544]
87. Angel TE, Aryal UK, Hengel SM, Baker ES, Kelly RT, Robinson EW, Smith RD. Mass spectrometry-based proteomics: existing capabilities and future directions. Chem. Soc. Rev. 2012; 41(10):3912–3928. [PubMed: 22498958]
88. Selvaraju S, El Rassi Z. Liquid-phase-based separation systems for depletion, prefractionation and enrichment of proteins in biological fluids and matrices for in-depth proteomics analysis - An update covering the period 2008–2011. Electrophoresis. 2012; 33(1):74–88. [PubMed: 22125262]
89. Ly L, Wasinger VC. Protein and peptide fractionation, enrichment and depletion: Tools for the complex proteome. Proteomics. 2011; 11(4):513–534. [PubMed: 21241016]
90. Manadas B, English JA, Wynne KJ, Cotter DR, Dunn MJ. Comparative analysis of OFFGel, strong cation exchange with pH gradient, and RP at high pH for first-dimensional separation of peptides from a membrane-enriched protein fraction. Proteomics. 2009; 9(22):5194–5198. [PubMed: 19771557]
91. Wang H, Chang-Wong T, Tang HY, Speicher DW. Comparison of extensive protein fractionation and repetitive LC-MS/MS analyses on depth of analysis for complex proteomes. J. Proteome Res. 2010; 9(2):1032–1040. [PubMed: 20014860]
92. Dowell JA, Frost DC, Zhang J, Li LJ. Comparison of two-dimensional fractionation techniques for shotgun proteomics. Anal. Chem. 2008; 80(17):6715–6723. [PubMed: 18680313]
93. Whiteaker JR, Zhang HD, Eng JK, Fang RH, Piening BD, Feng LC, Lorentzen TD, Schoenherr RM, Keane JF, Holzman T, Fitzgibbon M, Lin CW, Zhang H, Cooke K, Liu T, Camp DG, Anderson L, Watts J, Smith RD, McIntosh MW, Paulovich AG. Head-to-head comparison of serum fractionation techniques. J. Proteome Res. 2007; 6(2):828–836. [PubMed: 17269739]
94. Wang YX, Yang F, Gritsenko MA, Wang YC, Clauss T, Liu T, Shen YF, Monroe ME, Lopez-Ferrer D, Reno T, Moore RJ, Klemke RL, Camp DG, Smith RD. Reversed-phase chromatography with multiple fraction concatenation strategy for proteome profiling of human MCF10A cells. Proteomics. 2011; 11(10):2019–2026. [PubMed: 21500348]
95. Graham RLJ, Kalli A, Smith GT, Sweredoski MJ, Hess S. Avoiding Pitfalls in Proteomics Sample Preparation. Biomacromolecular Mass Spectrometry. 2011; 2(4):261–271.
96. Manza LL, Stamer SL, Ham AJL, Codreanu SG, Liebler DC. Sample preparation and digestion for proteomic analyses using spin filters. Proteomics. 2005; 5(7):1742–1745. [PubMed: 15761957]
97. Wisniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. Nat. Methods. 2009; 6(5):359–U60. [PubMed: 19377485]
98. Bereman MS, Egertson JD, MacCoss MJ. Comparison between procedures using SDS for shotgun proteomic analyses of complex samples. Proteomics. 2011; 11(14):2931–2935. [PubMed: 21656683]
99. Antharavally BS, Mallia KA, Rosenblatt MM, Salunkhe AM, Rogers JC, Haney P, Haghdoost N. Efficient removal of detergents from proteins and peptides in a spin column format. Anal. Biochem. 2011; 416(1):39–44. [PubMed: 21640699]
100. Mitulovic G, Stingl C, Steinmacher I, Hudecz O, Hutchins JRA, Peters JM, Mechtler K. Preventing carryover of peptides and proteins in nano LC-MS separations. Anal. Chem. 2009; 81(14):5955–5960. [PubMed: 19537771]
101. Perkins DN, Pappin DJC, Creasy DM, Cottrell JS. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis. 1999; 20(18):3551–3567. [PubMed: 10612281]
102. Craig R, Beavis RC. TANDEM: matching proteins with tandem mass spectra. Bioinformatics. 2004; 20(9):1466–1467. [PubMed: 14976030]
103. Geer LY, Markey SP, Kowalak JA, Wagner L, Xu M, Maynard DM, Yang XY, Shi WY, Bryant SH. Open mass spectrometry search algorithm. J. Proteome Res. 2004; 3(5):958–964. [PubMed: 15473683]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

104. Eng JK, McCormack AL, Yates JR. An approach to correlate tandem mass-spectral data of peptides with amino-acid-sequences in a protein database. J. Am. Soc. Mass Spectrom. 1994; 5(11):976–989.
105. Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. Andromeda: A peptide search engine integrated into the MaxQuant environment. J. Proteome Res. 2011; 10(4):1794–1805. [PubMed: 21254760]
106. Balgley BM, Laudeman T, Yang L, Song T, Lee CS. Comparative evaluation of tandem MS search algorithms using a target-decoy search strategy. Mol. Cell. Proteomics. 2007; 6(9):1599–1608. [PubMed: 17533222]
107. Brosch M, Swamy S, Hubbard T, Choudhary J. Comparison of Mascot and X!Tandem performance for low and high accuracy mass spectrometry and the development of an adjusted Mascot threshold. Mol. Cell. Proteomics. 2008; 7(5):962–970. [PubMed: 18216375]
108. Eng JK, Searle BC, Clauser KR, Tabb DL. A face in the crowd: Recognizing peptides through database search. Mol. Cell. Proteomics. 2011; 10(11)
109. Nesvizhskii AI, Vitek O, Aebersold R. Analysis and validation of proteomic data generated by tandem mass spectrometry. Nat. Methods. 2007; 4(10):787–797. [PubMed: 17901868]
110. Deutsch EW, Lam H, Aebersold R. Data analysis and bioinformatics tools for tandem mass spectrometry in proteomics. Physiol. Genomics. 2008; 33(1):18–25. [PubMed: 18212004]
111. Kapp EA, Schutz F, Connolly LM, Chakel JA, Meza JE, Miller CA, Fenyo D, Eng JK, Adkins JN, Omenn GS, Simpson RJ. An evaluation, comparison, and accurate benchmarking of several publicly available MS/MS search algorithms: Sensitivity and specificity analysis. Proteomics. 2005; 5(13):3475–3490. [PubMed: 16047398]
112. Sadygov RG, Cociorva D, Yates JR. Large-scate database searching using tandem mass spectra: Looking up the answer in the back of the book. Nat. Methods. 2004; 1(3):195–202. [PubMed: 15789030]
113. Matthiesen R. Methods, algorithms and tools in computational proteomics: A practical point of view. Proteomics. 2007; 7(16):2815–2832. [PubMed: 17703506]
114. Nesvizhskii AI, Aebersold R. Analysis, statistical validation and dissemination of large-scale proteomics datasets generated by tandem MS. Drug Discovery Today. 2004; 9(4):173–181. [PubMed: 14960397]
115. Nesvizhskii AI, Aebersold R. Interpretation of shotgun proteomic data - The protein inference problem. Mol. Cell. Proteomics. 2005; 4(10):1419–1440. [PubMed: 16009968]
116. Qian WJ, Jacobs JM, Liu T, Camp DG, Smith RD. Advances and challenges in liquid chromatography-mass spectrometry-based proteomics profiling for clinical applications. Mol. Cell. Proteomics. 2006; 5(10):1727–1744. [PubMed: 16887931]
117. Choi H, Ghosh D, Nesvizhskii AI. Statistical validation of peptide identifications in large-scale proteomics using the target-decoy database search strategy and flexible mixture modeling. J. Proteome Res. 2008; 7(1):286–292. [PubMed: 18078310]
118. Cargile BJ, Bundy JL, Stephenson JL. Potential for false positive identifications from large databases through tandem mass spectrometry. J. Proteome Res. 2004; 3(5):1082–1085. [PubMed: 15473699]
119. Elias JE, Gygi SP. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nat. Methods. 2007; 4(3):207–214. [PubMed: 17327847]
120. Gupta N, Bandeira N, Keich U, Pevzner PA. Target-decoy approach and false discovery rate: When things may go wrong. J. Am. Soc. Mass Spectrom. 2011; 22(7):1111–1120. [PubMed: 21953092]

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

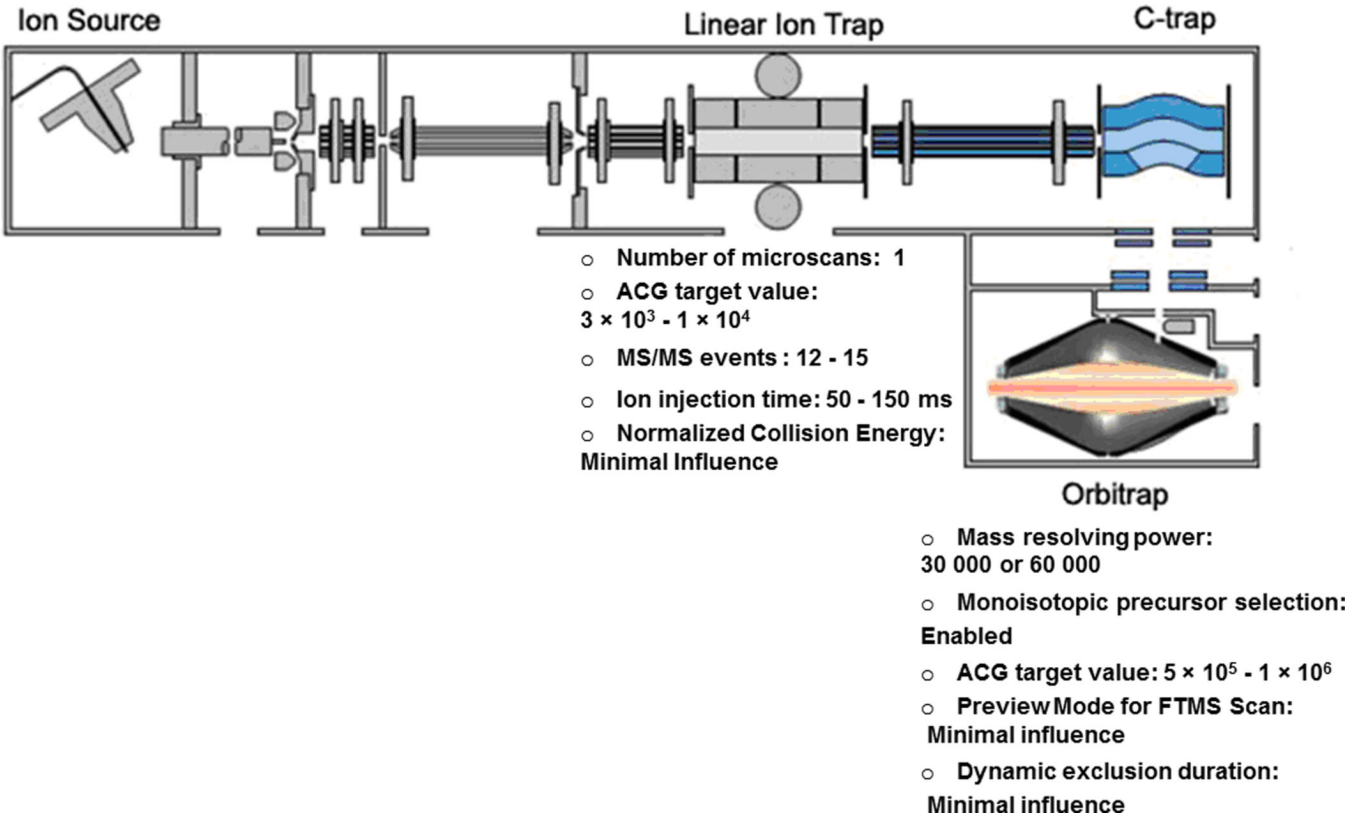

Figure 1.

Suggested optimized MS and MS/MS parameters for the LTQ-Orbitrap. Values were determined on the LTQ-Orbitrap Classic and XL and are taken from references [13], [15] and [16].

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

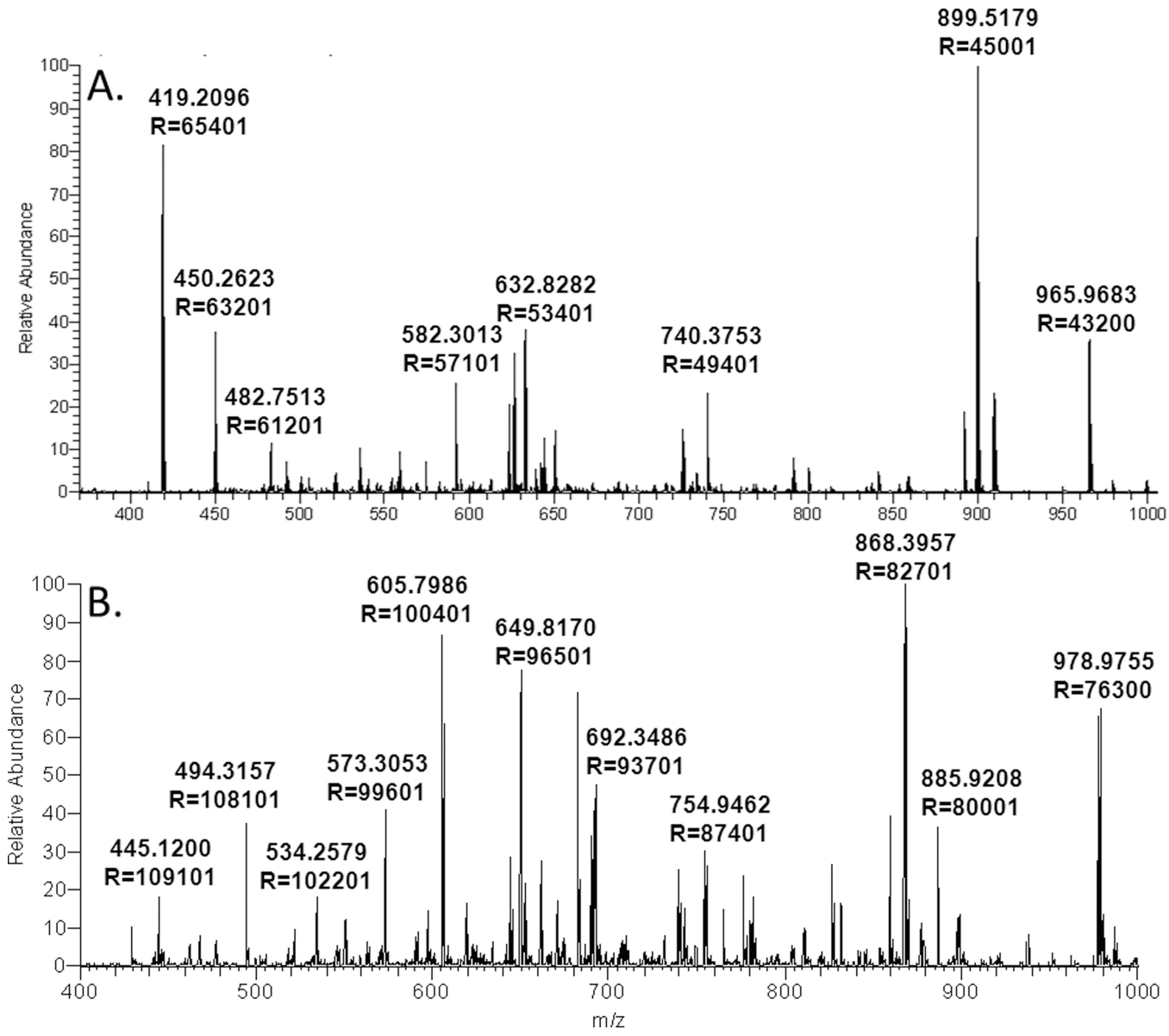

Figure 2.

Mass resolving power of the Orbitrap mass spectrometers is not steady across the m/z scan range and decreases as the square root of the m/z ratio. Data from Orbitrap Classic at a set mass resolving power of 60,000 (A) and Orbitrap Elite at a set mass resolving power of 120,000 (B).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

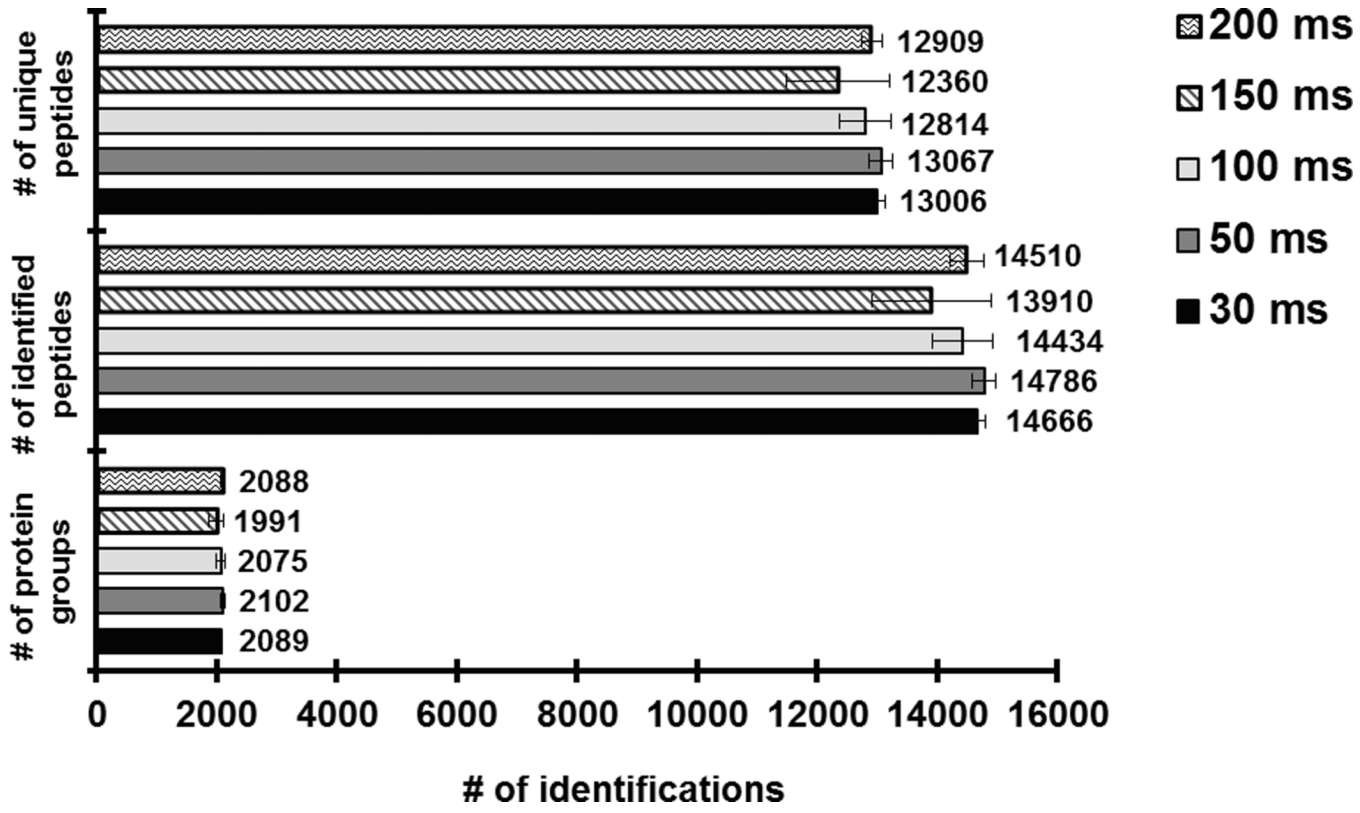

Figure 3.
Effect of MS/MS maximum ion injection time on identification rates observed for the Orbitrap Elite.

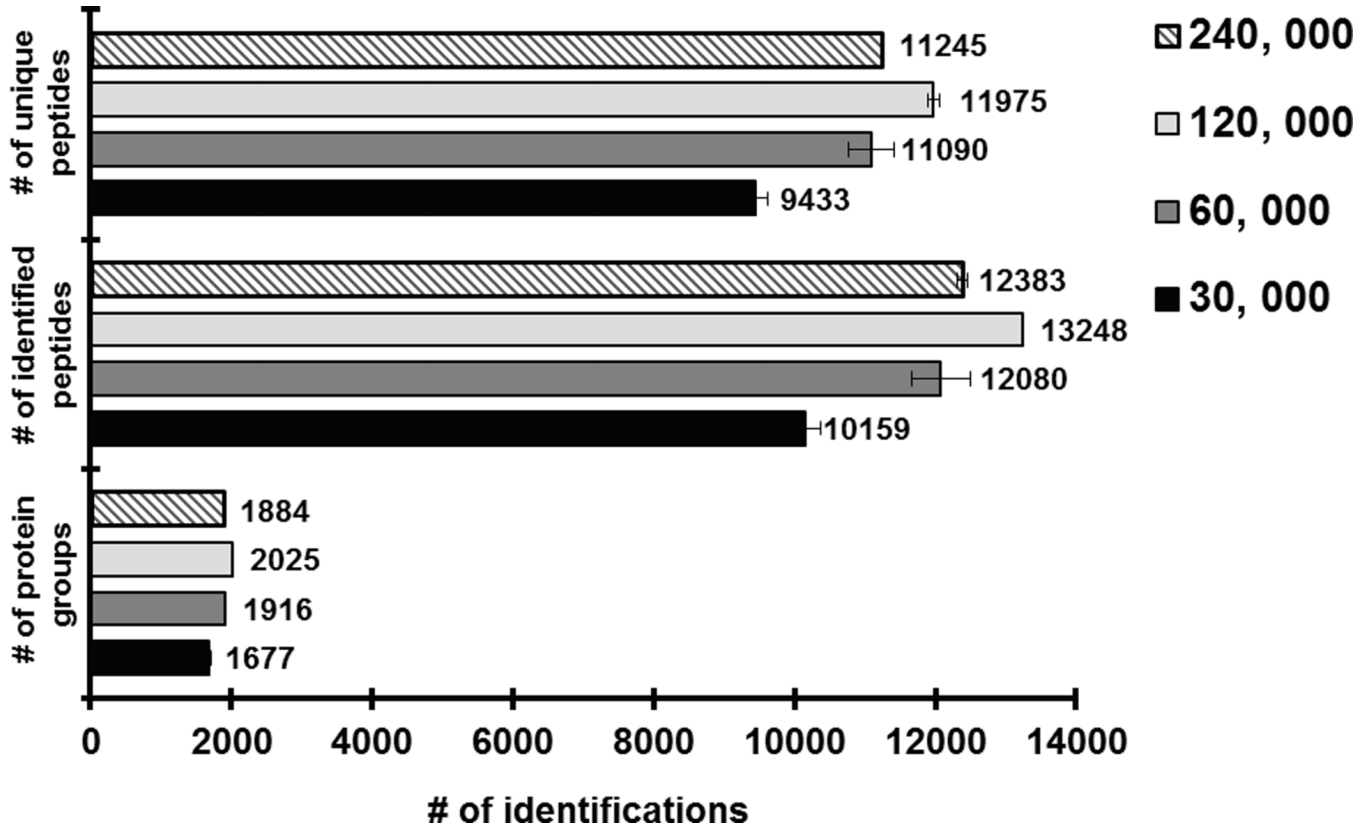

Figure 4.
Effect of MS resolving power on identification rates observed for the Orbitrap Elite.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

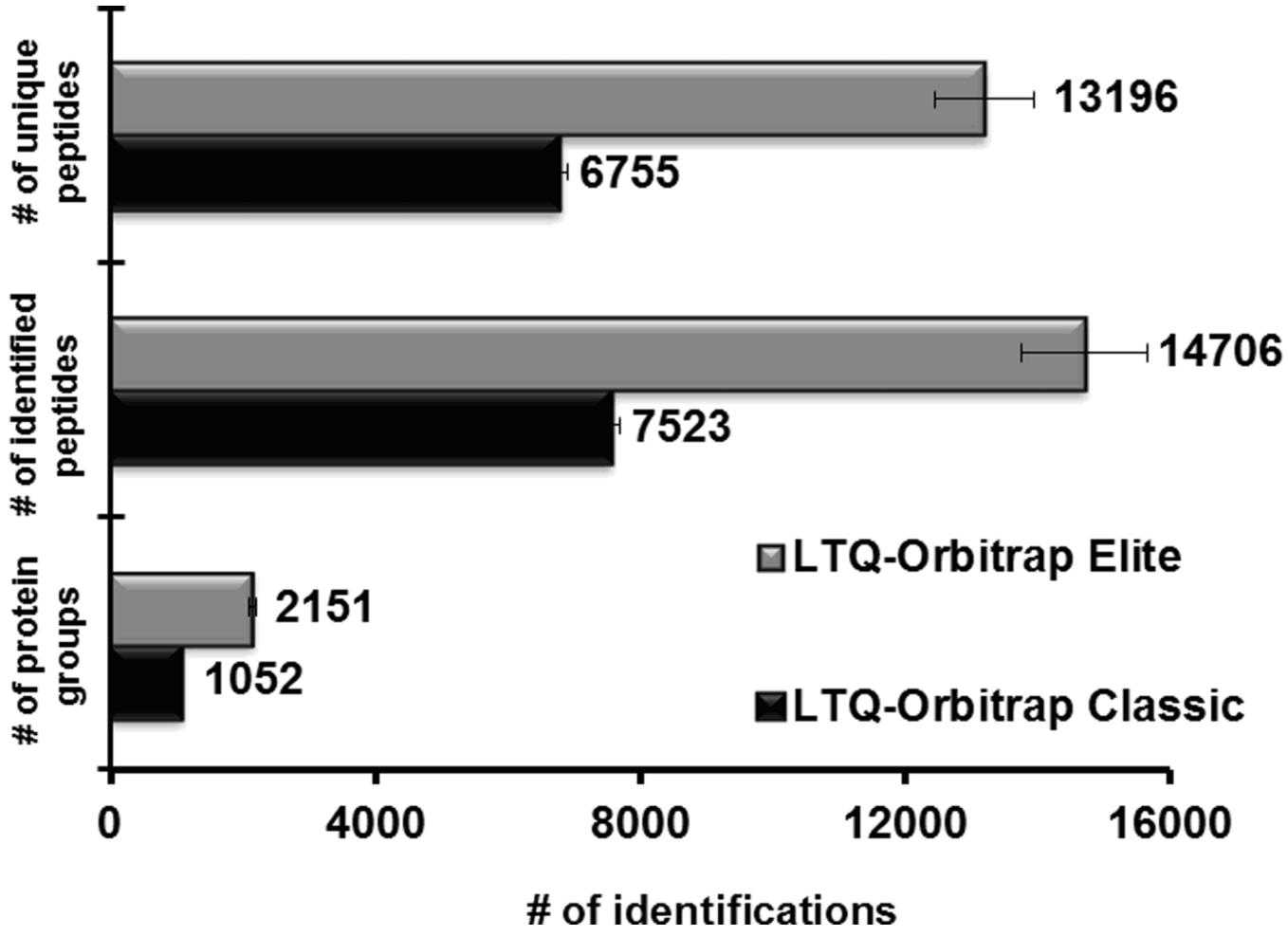

Figure 5.

Identification rates obtained in the LTQ-Orbitrap Classic and Orbitrap Elite under optimized settings. (LTQ-Orbitrap Classic: 12 MS/MS events, 1 microscan, preview mode for FTMS scan enabled, ion injection time for MS/MS 50 ms, AGC target value for MS/MS 5×10^3, AGC target value for MS 5×10^5, mass resolving power 60, 000, Orbitrap Elite: 15 MS/MS events, 1 microscan, preview mode for FTMS scan not enabled, ion injection time for MS/MS 50 ms, AGC target value for MS/MS 5×10^3, AGC target value for MS 1×10^6, mass resolving power 120, 000).

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

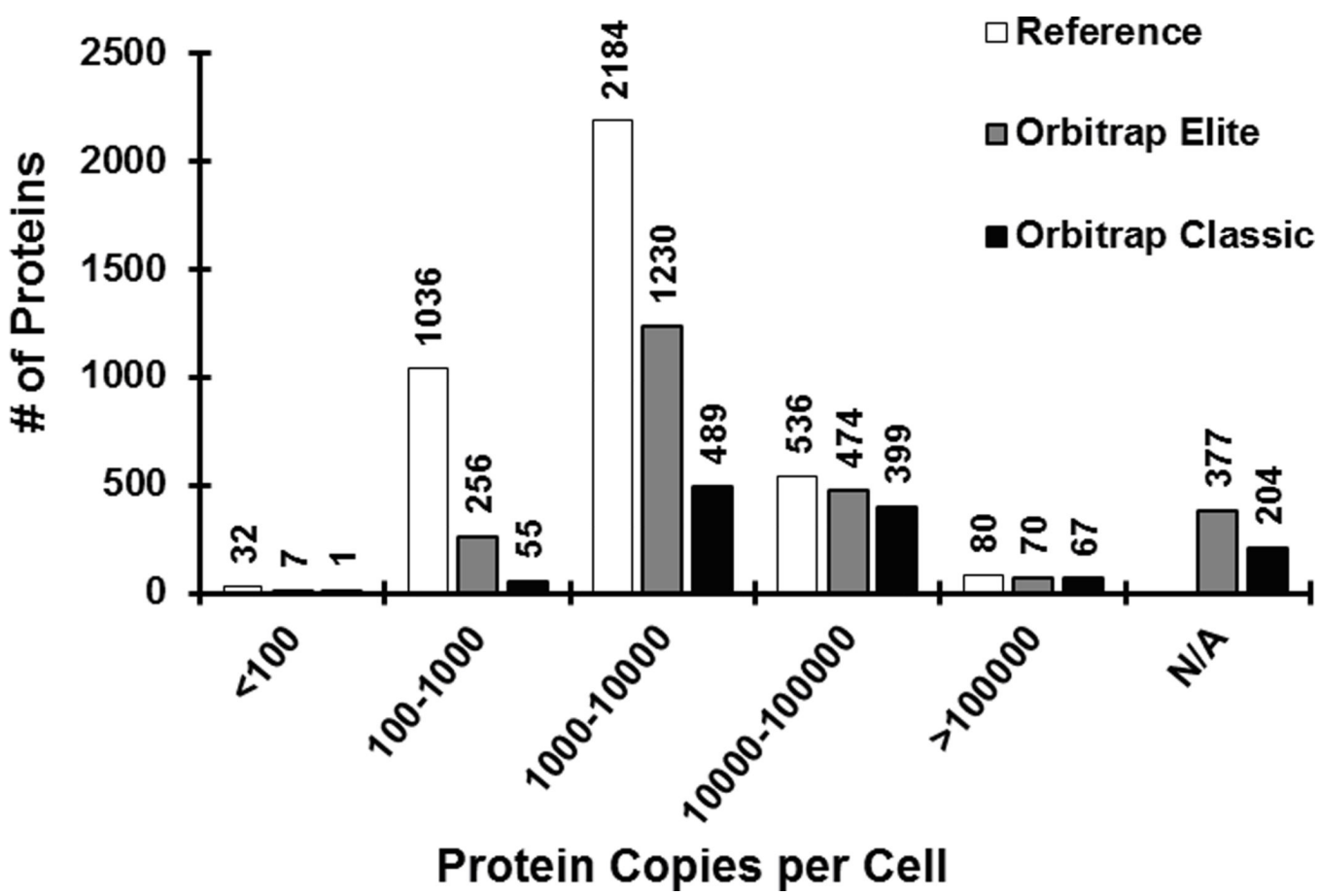

Figure 6.

Identified proteins as a function of their cellular expression levels observed on the LTQ-Orbitrap Classic and Orbitrap-Elite using the *Saccharomyces cerevisiae* proteome as the model system. Protein cellular expression levels were obtained from reference [37]. The number of identified proteins displayed in the Figure is a sum over two replicate analyses.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 1

Summary of MS and MS/MS parameters discussed and values tested for each parameter with their associated references.

MS parameter	Values/Settings Examined	Reference
Resolving power	7 500, 15 000, 30 000, 60 000, 100 000	[15]
Monoisotopic precursor selection	Enabled/Disabled	[13]
Preview mode for FTMS master scan	Enabled/Disabled	[16]
Minimum signal threshold for MS/MS	1×10^1, 1×10^2, 1×10^3, 1×10^4, 1×10^5, 1×10^6, 1×10^7	[14]
Number of microscans	1, 2, 3	[16]
Number of data dependent MS/MS	3, 5, 7, 8, 10, 12, 15, 20	[13, 16]
Automatic gain control (AGC) target value for MS	5×10^5, 8×10^5, 1×10^6, 2×10^6, 3×10^6	[16]
Automatic gain control (AGC) target value for MS/MS	1×10^3, 3×10^3, 5×10^3, 8×10^3, 1×10^4, 5×10^4, 1×10^5	[16]
Maximum ion injection time for MS/MS	50, 80, 100, 150, 200, 250, 400, 500 (ms)	[13, 16]
Normalized collision energy	28% and 35%	[13]
Dynamic exclusion duration	30 sec and 180 sec	[13]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 2

Set and actual ion injection time for different (A) MS and (B) MS/MS AGC target values.

A.

AGC target value for MS	Set MS ion injection time (ms)	Actual MS ion injection time (ms)[a]
5×10^5	700	18 (±3)
8×10^5	700	35 (±6)
1×10^6	700	64 (±10)
2×10^6	700	97 (±10)
3×10^6	700	150 (±10)

B.

AGC target value for MS/MS	Set MS/MS ion injection time (ms)	Actual MS/MS ion injection time (ms)[a]
1×10^3	500	11 (±1)
3×10^3	500	33 (±2)
5×10^3	500	51 (±6)
8×10^3	500	85 (±9)
1×10^4	500	115 (±9)
5×10^4	500	361 (±10)
1×10^5	500	440 (±9)

[a] Actual ion injection times were obtained from LogViewer.[18]

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 3

MS and MS/MS parameters examined for the Orbitrap Elite. Recommended settings are also indicated.

MS parameter	Values/Settings Examined	Recommended Settings
Automatic gain control (AQC) target value for MS	5×10^5, 8×10^5, 1×10^6	5×10^6–1×10^6
Automatic gain control (AGO) target value for MS/MS	3×10^3, 5×10^3, 8×10^3, 1×10^4	3×10^3–1×10^4
Maximum Ion Injection time for MS/MS	30, 50, 100, 150, 200 (ms)	30 – 100 (ma)
Number of data dependent MS/MS	15 and 20	15–20
Resolving power	30 000, 60 000, 120 000, 240 000	120 000
Scan rate for MS/MS events	Rapid and Normal	Rapid
Predict Ion Injection time	Enabled/Disabled	Enabled
Preview mode for FTMS master aaan	Enabled/Disabled	Disabled

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

J Proteome Res. Author manuscript; available in PMC 2014 July 05.

Table 4

Data obtained on the Orbitrap Elite by selecting rapid and normal CID scan mode (A) and by enabling/disabling the prediction of ion injection time (B).

A

Experiment	# Of MS scans	# Of MS/MS scans	# of Identified MS/MS spectra	% of Identified MS/MS spectra	# of identified peptides	# of unique peptides	# of identified proton groups (≥ 2 peptides)
Rapid CIO scan mode	5524 (±88)	41151 (±934)	17150 (± 1548)	42 (±3.0)	14207 (± 1095)	12827 (± B25)	2113 (± 78)
Normal CID scan mode	4911 (±48)	38887 (±326)	17334 (±329)	45 (±1.1)	14240 (± 560)	12855 (± 462)	2093 (± 42)

B

Experiment	# of MS scans	# of MS/MS scans	# of Identified MS/MS spectra	% of Identified MS/MS spectra	# of identified peptides	# of unique peptides	# of Identified protein groups (≥ 2 peptides)
Prediction Injection time enabled	5488 (±88)	41570 (±831)	17418 (±2089)	42 (±4.2)	14706 (±951)	13196 (±738)	2151 (±59)
Prediction injection time disabled	4308 (+51)	34937 (+184)	t4495 (+93)	42 (+0.05)	13013 (±30)	11790 (+61)	1977 (+23)

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 5

Data obtained on the Orbitrap Elite over three replicate LC-MS/MS analyses under the same experimental conditions.

Experiment	# of MS scans	# of MS/MS scans	# of Identified MS/MS spectra	% of Identified MS/MS spectra	# of Identified peptides	# of unique peptides	# of identified protein groups (≥ 2 peptides)
Experiment 1	5426	42158	18895	45	15378	13717	2192
Experiment 2	5550	40983	15941	39	14033	12675	2109
Experiment 3	5597	40312	18613	41	13209	12088	2037

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 6

MS and MS/MS settings employed during CID analysis on different Orbitrap mass analyzers.

	LTQ-Orbitrap	LTQ-Orbitrap	LTQ-Orbitrap Velos	LTQ-Orbitrap Velos	Orbitrap Elite	Orbitrap Elite
Model system	Yeast lysate	Yeast lysate	HeLa lysate	Yeast lysate, HEK-293	HeLa lysate	Yeast lysate
Amount loaded	1 μg	1.2 μg	5 μg	2 μg	400 ng	200 ng
LC gradient duration	160 min	160 min	120 min	140 min or 480 min	120 min	160 min
Resolving power MS	60,000	60,000	60,000	60,000	240,000	120,000
Minimum signal threshold for MS/MS	500	500	500	500	500	1000
Number of data dependent MS/MS	10	12	20	25	20	15
AGC target value for MS	1×10^6	5×10^5	1×10^6	1×10^6	1×10^6	1×10^6
AGC target value for MS/MS	5×10^3	5×10^3	5×10^3	5×10^3	5×10^3	5×10^3
Maximum ion injection time for MS	Not specified	700 ms	500 ms	250 ms[#]	100 ms[#]	200 ms
Maximum ion injection time for MS/MS	Not specified	50 ms	25 ms	25 ms[*]	25 to 200 ms	50 ms
Normalized collision energy	35 %	35%	35 %	40 %	33 %[#]	35 %
Dynamic exclusion duration	60 sec	90 sec	30 sec[*]	90 sec	60 sec	90 sec
Isolation width	2.0 Da	2.0 Da	2.0 Da[*]	2.0 Da[*]	2.0 Da	2.0 Da
Reference	[10]	[16]	[30]	[32]	[36]	This tutorial

[#]Stevan Horning, personal communication

[*]Matthias Mann, personal communication

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 7

MS and MS/MS settings employed during HCD analysis on different Orbitrap mass analyzers

MS and MS/MS settings	LTQ-Orbitrap Velos	Q Exactive	Q Exactive	Orbitrap Elite
Model system	HeLa lysate	Yeast lysate HeLa lysate	Yeast lysate	HeLa lysate
Amount loaded	5 μg	5 – 250 ng (yeast) 1 μg (HeLa)	4 μg	400 ng
LC gradient duration	120 min	60 min, 120 min, 180 min	240 min	120 min
Resolving power MS[a]	30,000	70,000	70,000	120,000
Resolving power MS/MS[a]	7,500	17,500[b] 35,000[c]	17,500	15,000
Minimum signal threshold for MS/MS	5×10^3	1×10^5	Not specified	$5 \times 10^{3\#}$
Number of data dependent MS/MS	10	12	10	15
AGC target value for MS	1×10^6	1×10^6	1×10^6	1×10^6
AGC target value for MS/MS	3×10^4 or 5×10^4	5×10^5	1×10^6	4×10^4
Maximum ion injection time for MS	500 ms	120 ms	20 ms	100 ms[#]
Maximum ion injection time for MS/MS	250 ms	60 ms[b] 120 ms[c]	60 ms	150 ms[#]
Normalized collision energy	40%	25 %	25 %	35%[#]
Dynamic exclusion duration	120 sec[*]	30 sec	40 sec	60 sec
Isolation width	5 Da[*]	1.2 Da (yeast) 2.0 Da (HeLa)	1.6 Da	2.0 Da
Reference	[30]	[39]	[40]	[36]

[a] Mass resolving power on the Orbitrap is specified at m/z = 400 and on the Q Exactive at m/z = 200

[b] Mass resolving power of 17, 5000 was combined with a maximum injection time of 60 ms

[c] Mass resolving power of 35, 5000 was combined with a maximum injection time of 120 ms

[#] Stevan Horning, personal communication

[*] Matthias Mann, personal communication