
Increasing Peptide Identification in Tandem Mass Spectrometry Through Automatic Function Switching Optimization

Brian Carrillo, Kossi Lekpor, Corey Yanofsky, Alexander W. Bell, Daniel Boismenu, and Robert E. Kearney

Department of Biomedical Engineering, McGill University, Montreal, Quebec, Canada

Comprehensive proteomic studies that employ MS directed peptide sequencing are limited by optimal peptide separation and MS and tandem MS data acquisition routines. To identify the optimal parameters for data acquisition, we developed a system that models the automatic function switching behavior of a mass spectrometer using an MS-only dataset. Simulations were conducted to characterize the number and the quality of simulated fragmentation as a function of the data acquisition routines and used to construct operating curves defining tandem mass spectra quality and the number of peptides fragmented. Results demonstrated that one could optimize for quality or quantity, with the number of peptides fragmented decreasing as quality increased. The predicted optimal operating curve indicated that significant improvements can be realized by selecting the appropriate data acquisition parameters. The simulation results were confirmed experimentally by testing 10 LC MS/MS data acquisition parameter sets on an LC-Q-TOF-MS. Database matching of the experimental fragmentation returned peptide scores consistent with the predictions of the model. The results of the simulations of mass spectrometer data acquisition routines reveal an inverse relationship between the quality and the quantity of peptide identifications and predict an optimal operating curve that can be used to select an optimal data acquisition parameter for a given (or any) sample. (J Am Soc Mass Spectrom 2005, 16, 1818–1826) © 2005 American Society for Mass Spectrometry

A fundamental goal of proteomics is the systematic, simultaneous analysis of large numbers of proteins in biological samples [1]. Automated, high-throughput analyses of entire or subcellular fractions are presently routine (made possible by the application of soft-ionization methods to mass spectrometry), as seen from the completion of an ever increasing number of genome projects [2, 3]. These innovations have allowed for the identification and characterization of both known and unknown proteins with unprecedented sensitivity, while boasting shorter analysis times, consistency in the analysis process, and the flexibility of multiple assays. Global analyses such as these will provide a comprehensive framework within which more traditional studies directed toward “well characterized” proteins, as well as those with no known function, can be carried out.

In shotgun proteomics, protein samples are generally enzymatically digested into smaller peptide fragments, making them amenable to sequence analysis by mass spectrometry [4]. Since the processing of a single protein

results in many peptides, and since mass spectrometers cannot efficiently analyze very complex samples, a complex peptide sample is typically first separated by using liquid chromatography (LC). In this technique, a gradient created by an aqueous and organic phase separate peptides based on their hydrophobic properties. The length of the gradient can be adjusted to achieve the desired time spacing between the elution of different peptides. The output of the LC can be coupled to a tandem mass spectrometer, allowing peptides to be detected and selected for fragmentation as they elute [2].

Tandem mass spectrometry uses two mass analyzers. The first selects a single peptide mass from the initial mass spectrum (MS) by filtering out all other masses. The single peptide is then fragmented in a collision cell. The second mass analyzer acquires the resulting fragmentation spectra (MS/MS). Since peptides typically fragment along the polypeptide backbone rather than in the side chains, the series of ions generated can be used to determine the amino acid sequence of the peptide [5]. Protein database searches can then be used to match the mass of the parent ion to all candidate peptides of *in silico* protein digests, and to rank the candidates based on the matching of theoretical and experimental fragmentation spectra [6]. Mascot, one of the more commonly used protein database

Published online September 28, 2005

Address reprint requests to Dr. B. Carrillo, Department of Biomedical Engineering, McGill University, 3640 University Street, Room M5, Montreal, Quebec H3X 2B3, Canada. E-mail: Brian.Carrillo@McGill.ca

search tools, finds the most suitable peptide for each spectra by matching peaks in each fragmentation spectrum to the calculated ions in the *in silico* digest. Mascot also calculates an identity score which is a measure of the probability that the match was a random error [7].

Mass spectra are necessary to identify potential peptides as target ions for tandem MS and to identify the molecular weight of peptides. Tandem mass spectra are necessary to confirm the presence of a peptide and to potentially identify its sequence [8]. A strategy termed automatic function switching is used to automatically switch between these types of mass spectral data acquisitions [9]. During automatic function switching, the mass spectrometer collects mass spectra by default. As each spectrum is collected, online software analyzes it to identify potential target ions for sequencing via MS/MS. Once target ions have been chosen, the mass spectrometer switches to MS/MS mode and collects one or more fragmentation spectra on each of the target ions in turn. The mass spectrometer then returns to the default MS mode and attempts to identify new target ions. Before data acquisition, the operator chooses the maximum number of target ions (n) and the number of tandem mass spectra to be collected per target ion (r). The last parameter required is ms , defined as the minimum number of mass spectra to be collected before the mass spectrometer switches to fragmentation mode. These three variables, ms , n , and r , comprise the data acquisition routine (DAR) parameters. Unfortunately, while tandem mass spectra are being collected, valuable sample is being consumed, allowing other possible target ions to go undetected and unfragmented. Since LC elution limits the total number of spectra that can be acquired, a trade-off exists between the number of peptides analyzed and the quality of the resulting data.

There has been considerable investigation into the effects of mass spectrometer parameters on the quality of the resulting spectra. Vaidyanathan et al. [10], for example, investigated 14 different variables including several flow rates, temperatures, and voltages within the mass spectrometer. We have investigated the effects of two DAR parameters, the number of target ions (n) and the repetition rate (r), on the quality and quantity of detected peptides. In doing so we were able to maximize the number of peptides being accurately identified. Mass spectrometer data were first collected using only MS scans on a complex sample of rat proteins. Automated function switching was simulated to model the selection behavior of the mass spectrometer for a given DAR. The number of peptides fragmented and the quality of the fragmentation spectra as indicated by the intensity of the precursor ion in the MS scan was then determined [11]. The effects of the DAR parameters were examined by simulating the behavior of the mass spectrometer model while varying the parameter combinations. The results predicted by the model were then verified experimentally.

Methods

The behavior of the mass spectrometer was simulated to avoid the time consuming and resource taxing alternative of running thousands of nearly identical experiments. Simulation allowed the variation of DAR parameters and the calculation of all possible outcomes in the order of minutes, whereas experimentation would have required weeks or months. Algorithms that mimic the behavior of the mass spectrometer were created to predict the outcome of the automatic function switching. The rules that govern the mass spectrometer's online target ion selection software were as follows:

1. Only ions determined to be doubly and triply charged are considered for fragmentation as these are considered most likely to be tryptic peptide ions.
2. If multiple ions are detected within a single scan, the fragmentation order is ion intensity dependant from the most to the least intense ion.
3. An ion will not be considered for fragmentation if an ion with the same mass and charge has been previously fragmented within an operator specified time retention time window.
4. An ion will not be considered if its current peak intensity is less than a threshold specified by the operator.

Algorithms that adhere to these rules were created to mimic the ion selection behavior of the mass spectrometer. The raw data used to simulate LC-MS/MS experiments were obtained from actual LC-MS analyses ($N = 16$) of a sample estimated to contain $\sim 10,000$ peptides (unpublished Q-TOF MS/MS studies of this sample have identified 100+ proteins, all from the 50 kDa band providing roughly 1 peptide/kDa with the possibility of one missed cleavage). Surface intensity analysis, which scans the intensity map of the sample for isotopically distributed 3D peak sets, detected ~ 2100 ions available to be fragmented [12]. Surface intensity analysis is a more robust method of detecting potential peptide ions because it uses adjacent spectra to mitigate high-frequency noise. Simulations were graded via two criteria, the number of peptides fragmented and the parent ion intensity of the peptide selected for fragmentation. The number of peptides fragmented is a measure of the completeness of the selection scheme. The parent ion intensity is an indicator of the expected quality of the fragmentation spectra. Figure 1 shows the relationship between the parent ion intensity and the total MS/MS ion intensity for the dataset used in the following experiments. The maximum likelihood best fit is shown and the 95% confidence area is shown in dashed lines. While the lower limit is flat, the upper limit shows a strong correlation between parent ion intensity and the maximum total MS/MS ion intensity. Strong parent ion intensities do not necessarily relate to high quality fragmentation spectra, but weak parent ion intensities do indicate that there are few molecules to fragment,

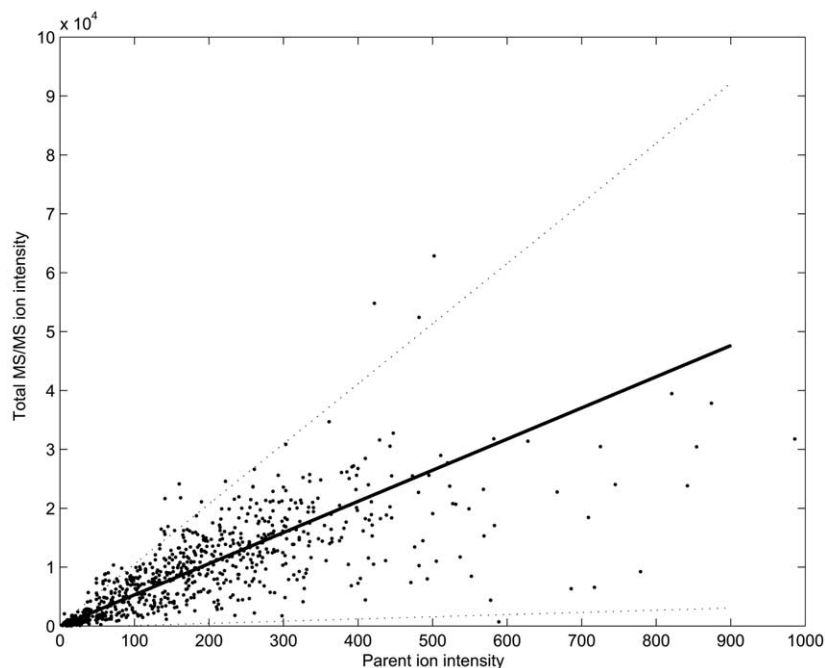


Figure 1. Relationship between total MS/MS ion intensity and parent ion intensity of 800+ spectra. The maximum likelihood fit (solid line) and the 95% confidence interval (dashed lines) are superimposed.

and that the corresponding fragmentation spectra will, in general, be of lower quality [13].

In simulating the mass spectrometer's data collection, our approach was to analyze each spectrum using optimal methods so that any differences could be attributed to data acquisition routines.

Protein Collection

Enriched preparations of endoplasmic reticulum, rough [14] and smooth microsomes [15], derived from rat liver tissue were resolved by SDS-PAGE on 7 to 15% gradient gels. The gel was Coomassie blue stained and the region previously found to be enriched in cytochrome P450 proteins with apparent molecular weight ~ 46 to 57 kDa (manuscript in preparation, J. J. M. Bergeron) was excised and diced into 1 mm cubes for robotic in-gel digestion. For robotic treatment, the gel pieces were destained, followed by a 30 min reduction with 10 mM dithiothreitol, a 20 min alkylation with 55 mM iodoacetamide, in-gel digested with trypsin (6 ng/ μ l in 50 mM $[\text{NH}_4]_2\text{CO}_3$) for 5 h at 37 °C and peptides extracted in 1% formic acid/2% acetonitrile [17].

Mass Spectrometry

MS analysis was performed as described previously [16, 17] by employing a CapLC (Waters, Milford, MA) linked to a Q-TOF MS (Micromass, Manchester, UK) equipped with a trapping guard column (PepMap C18, LC Packings, Sunnyvale, CA), and an analytical column (PicoFrit BioBasic C18, New Objective, Woburn, MA) as

described by Blondeau et al. [16]. Mass spectrometric data were acquired using either MS scans only, or automatic function switching [9] employing the data directed analysis feature available on MassLynx (Micromass) operating software with an array of data acquisition parameters (ms, n, r) whereby for each ms second MS scan, n precursor ions were selected for up to r fragmentation scans of 1 s, with an intensity threshold of 6.

Filtering

The spectra were preprocessed using a piecewise linear filtering technique [18]. This algorithm sliced each spectrum into 50 m/z wide segments. Each segment was then filtered using a Butterworth low pass filter with an order and cutoff frequency dependent on the m/z range of the segment (Butterworth, S. On the Theory of Filter Amplifiers. *Experimental Wireless and the Radio Engineer*, 1930, 7, 536–554.). These values were precomputed based on the modeled peak shape of calibration data. The spectra were rectified to remove any negative intensities (meaningless in mass spectrometry) introduced by the impulse response of the filter. Figure 2a and b show a segment of a typical mass spectrum before and after this filtering operation.

Peak Picking

Peak picking was simulated using a custom heuristic algorithm that used empirically determined peak proper-

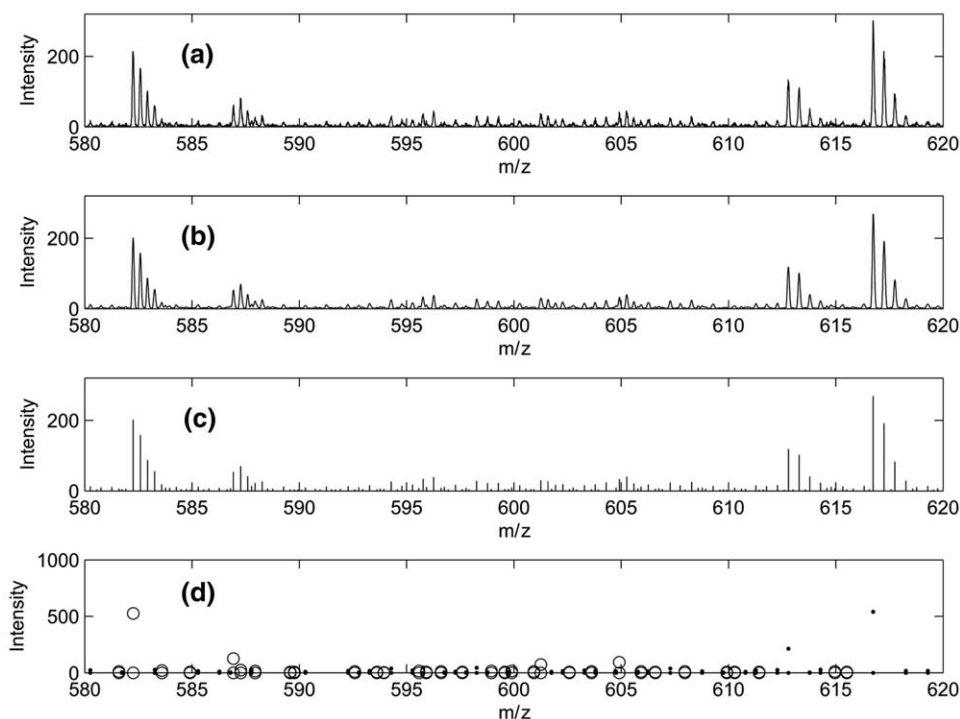


Figure 2. Sample processing of spectrum. (a) Raw spectrum. (b) Filtered spectrum. (c) Peak picked spectrum. (d) Deisotoped Spectrum. Dots represent doubly charged ions, circles represent triply charged ions

ties (particular to our Q-TOF instrument) to aid in peak finding. They include:

1. The width of an ion peak varies linearly with increasing m/z .
2. Weak intensity peaks close to high intensity peaks cannot be resolved.
3. The most intense data point within the peak defines the m/z location.

The peak picking algorithm applied this methodology utilizing a divide-and-conquer approach as described below to reduce processing time.

The algorithm begins by finding the most intense data point in the spectrum and computes its expected peak width based on its m/z (Rule 1). The peak intensity and location are recorded, and the peak and all data points within the calculated peak width are removed from the spectrum (Rule 2). The spectrum is split into higher and lower m/z portions at the point where the peak is detected, and these sub-spectra are then processed separately in the same manner, to find the next most intense peak, as if they were individual spectra. The algorithm concludes when all data points have been processed. All of the recorded peaks are combined to form a complete peak list. Figure 2c shows the results of peak-picking the filtered spectra.

Deisotoping

Deisotoping was achieved using an extension of the Wehofsky [11] algorithm. The deisotoping algorithm

scans the m/z axis sequentially, taking the first peak it encounters, and determines the spacing of consecutive peaks from the same isotopic cluster and then calculates the charge state of the ion. The set of isotopically distributed peaks indicates the presence of potential peptide ions.

For a given m/z , the relative intensity and spacing of the isotopic peaks of peptide ions are reasonably predictable. We employed the Breen method to model the isotopic distribution of peptide ions [19]. Small deviations (<20%) of isotopic peak heights for peptides from these average distributions can be attributed to molecules containing atomic compositions that vary from the average amino acid [20]. Larger deviations indicate that the considered mass/charge state does not adequately describe a peptide, or that more than one peptide ion is present in the considered peak set. If a set of isotopic peaks was within the tolerances of the isotopic pattern predicted by the model, the peak set was removed from the spectrum, recorded, and the processing of the spectrum continued. This list of recorded peaks was the deisotoped peak list. Figure 2d shows the results of the deisotoping operation.

Ion Selection Simulation

Simulating ion selection mimicked how the mass spectrometer ranks ions by intensity to select target ions for fragmentation. The target selection was guided by the DAR parameters which were varied such that each

feasible parameter combination was simulated (Figure 3).

The deisotoped peak lists were searched, one spectrum at a time, to identify peaks whose deisotoped intensity were above the mass spectrometer minimum threshold. If no suitable ions were found, the next scan was processed. If suitable ions were present, they were sorted by intensity and individually added (starting with the most intense) to the list of fragmented ions, up to the maximum specified by the DAR parameters. The repetition rate dictated how many times each ion was fragmented, and accordingly how many times it was replicated on the fragmented ion list. The number of ions (including repetitions) added to the list in a single scan determined how many mass spectra would not be collected as a consequence of acquiring MS/MS data. Ions were not considered for fragmentation if they were on the exclusion list, which was defined as the ions (i.e., m/z values) that have been previously fragmented within an operator-specified time (60 s). The exclusion list prevents the same peptide from being chosen for fragmentation twice in a row. All replicate tandem MS from the same parent ion were collapsed into a single entry in the list of fragmented ions. The list of fragmented ions contained the scan number and the m/z of the ions.

Database Searching

MS/MS raw data (150–350MB Micromass “raw” files per sample) were transferred from the LC-Q-TOF computer to a 40 terabyte server and automatically manipulated by TOMAS[21] for generation of peak lists by employing the Peak Detection Algorithm, of Matrix Science (<http://www.matrixscience.com/distiller.html>). The following peak detection algorithm parameters were employed. Maximum interaction: 500; correlation threshold: 0.7; minimum peak: 50 m/z ; maximum peak: 100,000 m/z ;

minimum S/N: 2; minimum peak width: 0.01; maximum peak width: 1; expected peak width: 0.1; maximum allowed precursor difference: 3 Da. Baseline correction and reject width outliers were activated. Resampling was done with a value of 20 data points per Da. The peak listed data (250–900 kB Mascot Distiller “mgf” files per sample) were then submitted to Mascot Cluster version 1.9.03 (<http://www.matrixscience.com/cluster.html>) by TOMAS[21] and searched against a copy of the National Center for Biotechnology Information (NCBI) nonredundant database (<ftp://ftp.ncbi.nih.gov/blast/db/FASTA/nr.gz>) (NCBI nrdb March 18, 2004). The search was restricted to up to 1 missed (trypsin) cleavage, fixed carbamidomethyl alkylation of cysteines, variable oxidation of methionine, 0.5 Da tolerance on parent and fragment ions, monoisotopic [16,21], and limited to the *rattus* taxonomy (30,949 sequences). Each peak list took ~10 s to search on a 4 cpu Mascot cluster.

Results

The results of simulated automatic function switching (Figure 4) show that varying the DAR parameters (n and r) had a significant impact on both the predicted intensity and predicted quantity of peptide fragmentation spectra. Predicted intensity is an average of the parent ion intensity. Each curve in the figure represents the results of ten experiments where the number of target ions (n) remained constant (as depicted in the legend), and the number of repetitions (r) increased from one to ten from right to left. The simulations predict that there is an optimal operating curve (grey line in Figure 4) that defines a set of DAR parameters that are Pareto optimal [22], each parameter set maximizing the parent ion intensity for a given number of fragmentations. This curve is optimal since any choice of DAR parameters below the curve will yield data of lower quality and/or quantity. The predicted optimal operating curve shows that there is an inverse relationship between predicted quality and quantity. This curve further indicates that most DAR parameter sets are suboptimal, indicating that improvements can be realized by operating on the optimal curve.

These predictions were tested experimentally using the same sample as for the MS-only data collection. A representative set of DAR parameters that define the simulated optimal operating curve in Figure 4 was selected and the corresponding experimental data for these DARs was collected in triplicate. Figure 5 shows that the experimental average parent ion intensities and the number of ions fragmented for the nine DARs that are predicted to be on the optimal operating curve conform to the simulated trends. The transformation between the simulated optimal operating curve and the experimental optimal operating curve is highly linear in the x -axis ($r^2 = 0.9512$), and somewhat linear in the y -axis ($r^2 = 0.7498$) (figures not shown). For comparison, an additional data acquisition routine (4,5) was

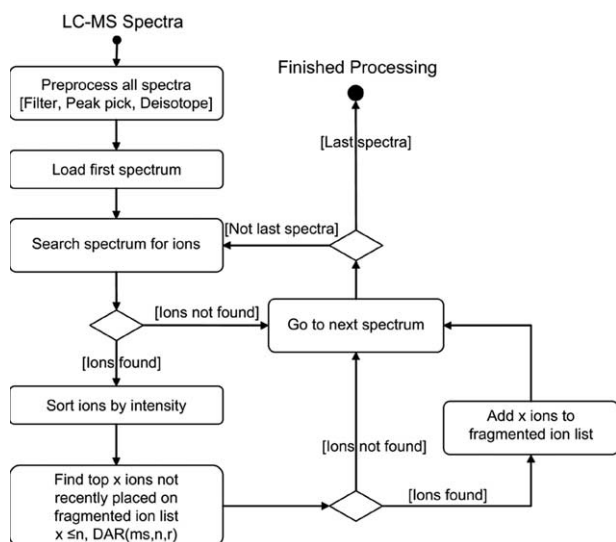


Figure 3. Ion selection simulation flowchart

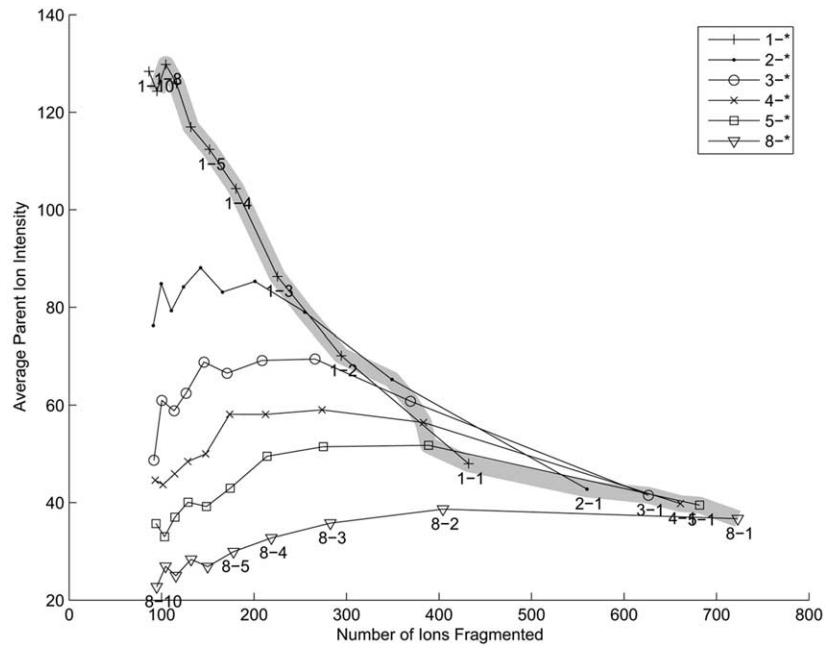


Figure 4. Simulated results of various DAR parameters: number of target ions selected and the number of repetitions were varied. (Repetitions increase from 1 to 10 moving right to left on each line, the number of target ions represented by different symbols, inset). The thick grey line represents the optimal operating curve. The DAR parameters are indicated as $n-r$ where n is the number of target ions and r is the number of seconds each target ion was fragmented.

also tested, as it represents the routines employed in previous proteomics experiments.

To evaluate the utility of the fragmentation data, it was submitted to database search; the Mascot scores were used as a measure of utility. The results of these

database searches and the average peptide scores versus the number of peptides are shown in Figure 6. The new shape is similar to the previous curve and the extreme data acquisition parameters of 8-1 deviates from the predicted trends. However, the general trend

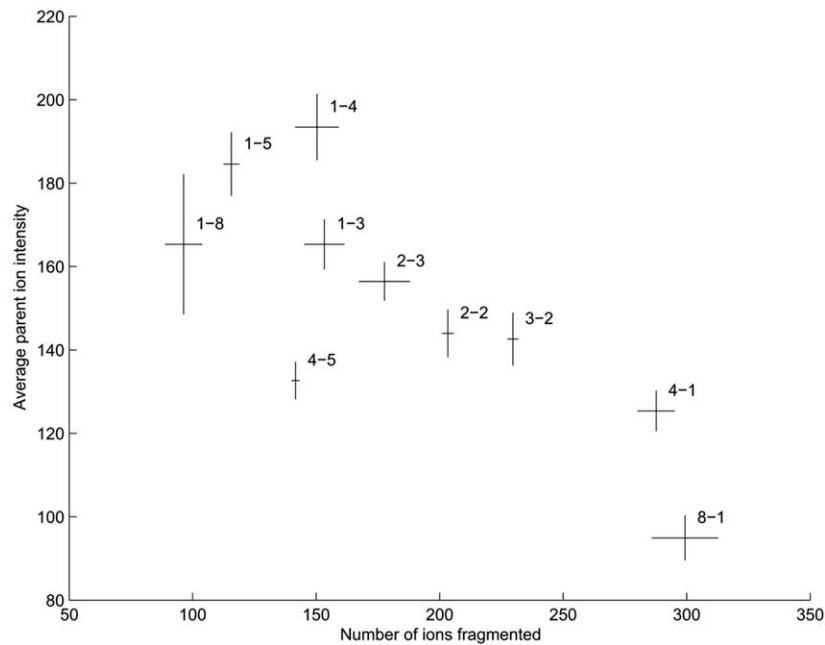


Figure 5. Experimental results of the average number of peptides fragmented and the average parent ion intensity for selected DAR parameters for the three replicates. Averages are plotted with standard error bars.

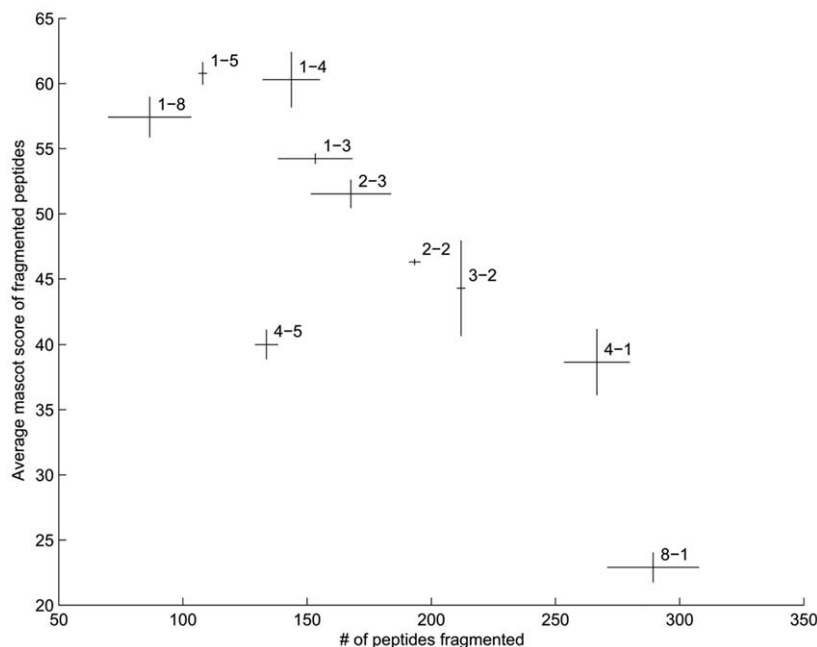


Figure 6. Experimental results showing the relationship between average assigned Mascot peptide scores and the average number of peptides fragmented for the three replicates.

between predicted quality and predicted quantity is realized. The transformation between the simulated optimal operating curve and the curve generated by the average Mascot score is highly linear in the x-axis ($r^2 = 0.9498$), and somewhat linear in the y-axis ($r^2 = 0.7930$) (figures not shown).

The average Mascot score for a given DAR is an indicator of the average quality of the mass spectra;

however, the average spectral quality does not reflect the total amount of *useful* information. Accordingly, the fragmentation data that were assigned to a peptide with a Mascot score at or greater than identity score was compiled and plotted in Figure 7. The resulting curves are dramatically different from the prior figures; both quality and quantity are maximized simultaneously, and there is no longer a trade-off between them. The

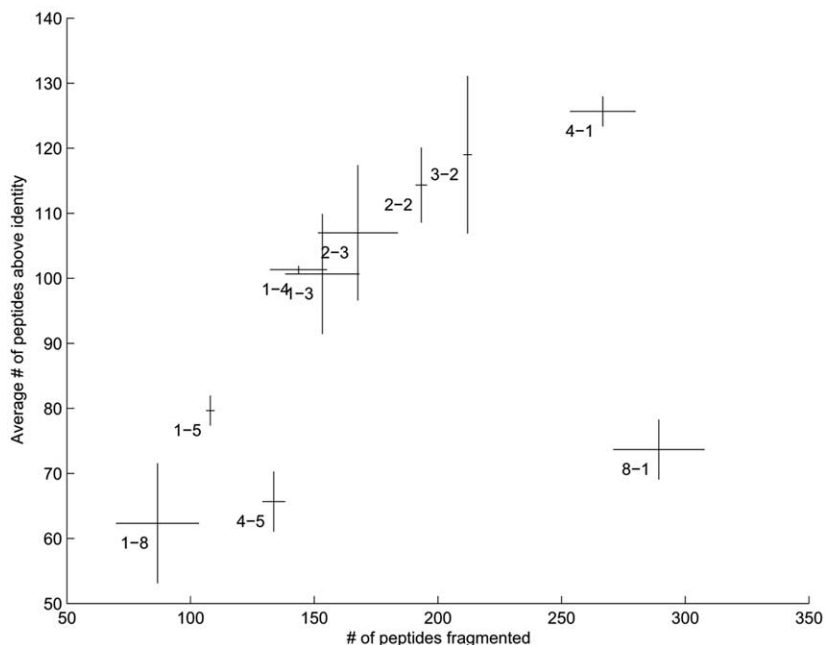


Figure 7. Experimental results showing the relationship between the number of peptides with Mascot scores greater than the identity score and the average number of peptides fragmented.

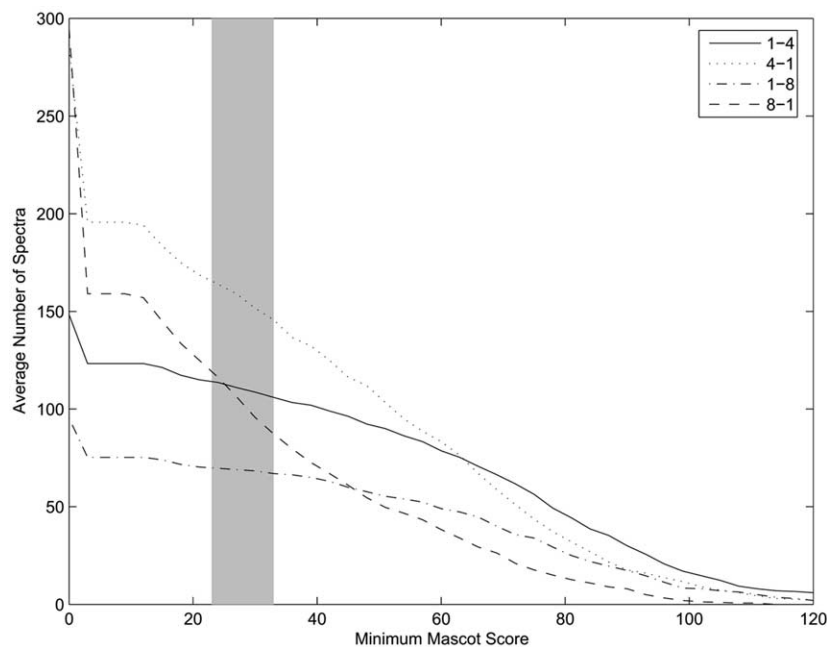


Figure 8. The plot shows the number of spectra that meet or exceed particular mascot scores. The 4-1 DAR parameter has a fewer number of high scores, but has sufficient scores above identity to provide the largest number of identifications. Grey area denotes Mascot identity threshold.

number of peptides above the Mascot identity threshold increases with the number of peptides selected for fragmentation in all except the extreme switching parameter of 8-1. The distribution of Mascot scores in Figure 8 shows that the 4-1 data acquisition parameter has fewer high scores (>80) than 1-4, but has significantly more scores that are at/near the identity threshold. The distribution of scores also gives insight into the dramatic drop of the 8-1 data acquisition parameter as a large proportion of the spectra result in Mascot scores below the identity threshold. This is likely a cause of the falloff of the peak intensity as the gradient changes; it is most dramatic for the 8-1 data acquisition parameter as many seconds have passed since the ion was detected and selected for fragmentation.

Conclusions

We have presented a method of modeling the peptide selection behavior of the mass spectrometer and to characterize the relationship between the quality and the quantity of peptides fragmented. Simulations using this method defined an optimal operating curve that specifies the data acquisition parameters required to produce the maximum number of peptides for a particular quality (or vice versa). This optimal operating curve predicts that a Pareto efficient operating point can be reached; a significant improvement in either quality, quantity, or both can be realized by selecting appropriate DAR parameters. The optimal operating curve can be realized from a set of MS-only data and allows an operator to make analysis decisions based on sample content. Experimental curves generated by testing 10

switching parameters validated the shape and trends predicted by the model. The experimental data were subsequently analyzed by Mascot and the resulting peptide scores also confirmed the predictions of the model. The results indicate that for this sample, selecting and fragmenting four target ions for one second each maximized the number of peptide identifications.

The complexity and concentration of the sample, however, will have a significant impact on the choice of the optimal operating conditions. Consequently, for each new biological preparation, it would be beneficial to perform preliminary MS-only experiments to determine sample complexity, and the availability of target ions for fragmentation. Based on this, the methods described in this paper could then be used to define the optimal set of data acquisition parameters. Alternatively, a DAR with $n = 1$ and $r = 1$ could be used to acquire some fragmentation data and, thus, determine some sample components, while still retaining sufficient MS data to reconstruct (through interpolation) a MS-only profile suitable for determining optimal data acquisition parameters.

Acknowledgments

This work was supported by grants from Genome Quebec, Genome Canada, the Natural Sciences and Engineering Research Council of Canada, and the Canadian Institute of Health Research. The authors are indebted to John Bergeron for securing funding (GQ/GC, CFI, VRQ) for the creation and operation of the Montreal Proteomics Network (RPMPN).

References

1. Cahill, D. J. Bridging genomics and proteomics. Proteomics: From protein sequence to function; BIOS: Oxford, 2001; pp 1–22.
2. Cho, A.; Normile, D. Nobel Prize in Chemistry. Mastering macromolecules. *Science* **2002**, *298*, 527–528.
3. Wilkins, M. R. The automation of proteomics: Technical and informatic solutions for high-throughput protein analysis; in *Proteomics: From protein sequence to function*; BIOS: Oxford, 2001; pp. 171–192.
4. Aebersold, R.; Mann, M. Mass spectrometry-based proteomics. *Nature* **2003**, *422*, 198–207.
5. De Hoffman, E.; Stroobant, V. Mass spectrometry: Principles and applications. Wiley: Chichester, 2001; pp. 255–259.
6. Eng, J. K.; McCormack, A. L.; Yates, J. T., III. 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*, 976–989.
7. Perkins, D. N.; Pappin, D. J.; Creasy, D. M.; Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **1999**, *20*, 3551–3567.
8. Dancik, V.; Addona, T. A.; Clauser, K. R. Vath, J. E.; Pevzner, P. A. De novo peptide sequencing via tandem mass spectrometry. *J. Comput. Biol.* **1999**, *6*, 327–342.
9. Hoyes, E.; Gaskell, S. J. Automatic function switching and its usefulness in peptide and protein analysis using direct infusion microspray quadrupole time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 1802–1806.
10. Vaidyanathan, S.; Broadhurst, D. I.; Kell, D. B.; Goodacre, R. Explanatory optimization of protein mass spectrometry via genetic search. *Anal. Chem.* **2003**, *75*, 6679–6686.
11. Wehofskey, M.; Hoffmann, R. Automated deconvolution and deisotoping of electrospray mass spectra. *J. Mass Spectrom.* **2002**, *37*, 223–229.
12. Carrillo, B.; Lekpor, K.; Yanofsky, C.; Bell, A.; Boismenu, D.; Kearney, R. E. Intensity surface analysis for peptide counting. *IEEE Eng. Med. Biol. Soc.* **2004**, *1*, 220–223.
13. Katta, V.; Chowdhury, S. K.; Chait, B. T. Use of a single-quadrupole mass spectrometer for collision-induced dissociation studies of multiply charged peptide ions produced by electrospray ionization. *Anal. Chem.* **1991**, *63*, 174–178.
14. Paiement, J.; Bergeron, J. J. Localization of GTP-stimulated core glycosylation to fused microsomes. *J. Cell Biol.* **1983**, *96*, 1791–1796.
15. Lavoie, C.; Lanoix, J.; Kan, F. W.; Paiement, J. Cell-free assembly of rough and smooth endoplasmic reticulum. *J. Cell Sci.* **1996**, *109*, 1415–1425.
16. Blondeau, F.; Ritter, B.; Allaire, P. D.; Wasiak, S.; Girard, M.; Hussain, N. K.; Angers, A.; Legendre-Guillemain, V.; Roy, L.; Boismenu, D.; Kearney, R. E.; Bell, A. W.; Bergeron, J. J.; McPherson, P. S. Tandem MS analysis of brain clathrin-coated vesicles reveals their critical involvement in synaptic vesicle recycling. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 3833–3838.
17. Wasiak, S.; Legendre-Guillemain, V.; Puertollano, R.; Blondeau, F.; Girard, M.; de Heuvel, E.; Boismenu, D.; Bell, A. W.; Bonifacino, J. S.; McPherson, P. S. Enthoprotin: A novel clathrin-associated protein identified through subcellular proteomics. *J. Cell Biol.* **2002**, *158*, 855–862.
18. Lekpor, K. *Time-varying filtering of time-of-flight mass spectra for proteomics*; Masters in Engineering Thesis, McGill University, Canada, 2004.
19. Breen, E. J.; Hopwood, F. G.; Williams, K. L.; Wilkins, M. R. Automatic poisson peak harvesting for high throughput protein identification. *Electrophoresis* **2000**, *21*, 2243–2251.
20. Gay, S.; Binz, P. A.; Hochstrasser, D. F.; Appel, R. D. Modeling peptide mass fingerprinting data using the atomic composition of peptides. *Electrophoresis* **1999**, *20*, 3527–3534.
21. Morales, F.; Kearney, R. E.; Bencsath-Makkai, Z.; Bergeron, J. J. M. TOMAS: Toolbox for mass spectrometry data analysis. *Mol. Cell. Proteom.* **2003**, *2*, 832.
22. Feldman, A. M. *Welfare economics and social choice theory*; Kluwer: Boston, 1980.