

# Post-Acquisition ETD Spectral Processing for Increased Peptide Identifications

David M. Good,<sup>a</sup> Craig D. Wenger,<sup>a</sup> Graeme C. McAlister,<sup>a</sup> Dina L. Bai,<sup>b</sup> Donald F. Hunt,<sup>b,c</sup> and Joshua J. Coon<sup>a,d</sup>

<sup>a</sup> Department of Chemistry, University of Wisconsin, Madison, Wisconsin, USA

<sup>b</sup> Department of Chemistry, University of Virginia, Charlottesville, Virginia, USA

<sup>c</sup> Department of Pathology, University of Virginia, Charlottesville, Virginia, USA

<sup>d</sup> Department of Biomolecular Chemistry, University of Wisconsin, Madison, Wisconsin, USA

Tandem mass spectra (MS/MS) produced using electron transfer dissociation (ETD) differ from those derived from collision-activated dissociation (CAD) in several important ways. Foremost, the predominant fragment ion series are different: *c*- and *z*<sup>\*</sup>-type ions are favored in ETD spectra while *b*- and *y*-type ions comprise the bulk of the fragments in CAD spectra. Additionally, ETD spectra possess charge-reduced precursors and unique neutral losses. Most database search algorithms were designed to analyze CAD spectra, and have only recently been adapted to accommodate *c*- and *z*<sup>\*</sup>-type ions; therefore, inclusion of these additional spectral features can hinder identification, leading to lower confidence scores and decreased sensitivity. Because of this, it is important to pre-process spectral data before submission to a database search to remove those features that cause complications. Here, we demonstrate the effects of removing these features on the number of unique peptide identifications at a 1% false discovery rate (FDR) using the open mass spectrometry search algorithm (OMSSA). When analyzing two biologic replicates of a yeast protein extract in three total analyses, the number of unique identifications with a ~1% FDR increased from 4611 to 5931 upon spectral pre-processing—an increase of ~28.6%. We outline the most effective pre-processing methods, and provide free software containing these algorithms. (J Am Soc Mass Spectrom 2009, 20, 1435–1440) © 2009 American Society for Mass Spectrometry

Collision-activated dissociation (CAD) is the most commonly employed method of peptide fragmentation. During CAD, peptide ions are collided with inert gas atoms, imparting energy, and causing preferential cleavage of the weakest, most energetically favorable bonds. For unmodified peptides, the weakest bonds tend to be the amide linkages between amino acid residues. Cleavage of these bonds results in the formation of *b*- and *y*-type fragment ions. A fundamentally distinct, electron-based fragmentation method was discovered 10 years ago by Zubarev et al. [1] In electron capture dissociation (ECD), the capture of thermal electrons by peptide cations also induces cleavage of backbone bonds. However, in contrast to CAD, N–C<sub>α</sub> bond is favored with ECD, thus forming *c*- and *z*<sup>\*</sup>-type fragment ions. Recently, an ion/ion analog to ECD was discovered. In electron-transfer dissociation (ETD), the ability to produce ECD-like spectra was achieved within an ion trap mass spectrometer through reaction of analyte cations with radical anion molecules [2, 3]. Subsequent studies have suggested that ETD is

better suited than CAD for analysis of certain peptides, particularly those that are large, highly basic, and/or contain post-translational modifications [4, 5], and is considered a complementary dissociation method [6, 7]. With the advent of commercially available bench-top ion trap instruments capable of electron-based dissociation, its inclusion in large-scale proteomics experiments has become increasingly widespread. Due to this increased use, several of the more common database search algorithms have been adapted to allow for searching of *c*- and *z*<sup>\*</sup>-type ions. However, these algorithms had originally been designed for searching CAD-generated data, typically using trypsin-digested proteins as the standard for the construction of the search algorithm. These algorithms therefore do not take into account several of the major spectral differences between CAD- and ETD-produced spectra.

In addition to producing different fragment ion types, ETD spectra differ from those generated by CAD in several important ways. As well as containing significantly more unreacted precursor than is typical for a CAD spectra generated via resonant excitation, charge-reduced species—those precursor cations that received one or more electrons but did not dissociate (ETnoD)—often constitute a significant fraction of the total ion

Address reprint requests to Professor J. J. Coon, Department of Chemistry, University of Wisconsin, 1101 University Avenue, Madison, WI 53706, USA. E-mail: jcoon@chem.wisc.edu

current of the spectrum. Another major difference between ETD and CAD spectra are the types of neutral losses observed as a result of the fragmentation. While CAD spectra frequently contain fragment ions with neutral losses of immonium ions or water molecules, these are not often seen in ETD spectra. Recently, several groups have reported on those neutral losses which are specific to ECD spectra [8–10]. While it has been suggested that neutral losses produced from the electron-based methods may be used to increase the confidence of peptide identifications, this has been shown only within the context of a spectral database, and has not yet been extended to database search algorithms [11, 12]. Therefore, while many algorithms account for several of the well-documented neutral losses observed in CAD, they do not currently take advantage of the extra information available within ETD-generated spectra. On the contrary, those database search algorithms that do allow for searching of ETD or ECD spectra typically produce lower scoring identifications when these species are present within the spectrum, ultimately leading to fewer identifications or a lower confidence threshold (data not shown).

Here, we demonstrate the effects of removing these features on both the confidence and number of identifications using a commonly employed database search algorithm, the open mass spectrometry search algorithm (OMSSA) [13]. We outline the most effective pre-processing methods, and provide free software containing these algorithms to aid in automated workflows.

## Materials and Methods

### Cell Culture and Protein Harvesting

Wild-type *Saccharomyces cerevisiae* were grown and proteins harvested as previously described [14]. Briefly, yeast was grown in rich media to an  $OD_{600}$  of 0.97, spun down, washed with sterile water, and pelleted via centrifugation at 10,000 *g* for 5 min. The cell pellet was added to a volume of lysis buffer containing protease inhibitors. The sample was French-pressed three times and centrifuged for 15 min at 30,000 *g* at 4 °C. This procedure was performed twice, using biologic replicates.

### Digestion

To reduce and alkylate cysteine residues, ~4.2 mg of protein (as determined by BCA assay) was incubated in 2.5 mM DTT for 25 min at 60 °C followed by incubation in 7 mM iodoacetamide in the dark at room temperature for 30 min. Alkylation was capped by incubation in 2.5 mM DTT for 15 min at room temperature. The samples were digested with endoproteinase lysine-C, desalted, and the eluent lyophilized following established protocol [14].

### Fractionation

SCX fractionation of all samples was performed as previously described by Villen et al. [15] Peptides were re-dissolved in 500  $\mu$ L SCX buffer A (5 mM  $KH_2PO_4$ , pH 2.6/30% acetonitrile) and separated using a  $9.4 \times 200$ -mm polysulfoethyl aspartamide column (5 mm particle size, 200 Å pore; PolyLC, Columbia, MD) with a surveyor HPLC pump and PDA detector (Thermo Fisher Scientific, San Jose, CA). Buffer A was flowed over the column for 3 min, and peptides were then separated with a linear gradient from 0% to 21% buffer B (5 mM  $KH_2PO_4$ , pH 2.6/30% acetonitrile/350 mM KCl) over 35 min. The column was equilibrated with multiple washes of 100% buffer B and 100% buffer C (20 mM Tris, pH 8.5). Fractions were collected over 3 min intervals, yielding 12 total fractions. These fractions were lyophilized, re-suspended in 0.5% TFA, and desalted using 100 mg tC18 SepPak cartridges (Waters, Milford, MA). The resulting eluent was lyophilized and stored at –20 °C until further use.

### Chromatography

SCX fractions were loaded onto a pre-column via a Waters nanoAcquity autosampler (Waters, Framingham, MA). Columns were prepared in-house as described previously [16], with the exception that the analytical column was packed to 12 cm. Peptides were chromatographically separated using a vented column setup on a Waters nanoAcquity and employing a 40 min linear gradient of 1.4% to 49% acetonitrile in 0.2% formic acid.

### Mass Spectrometry

Mass spectrometry experiments were performed in an on-line manner, in which peptides eluting from the above described nHPLC method were sampled via an integrated electrospray emitter for peptide ionization [16]. Analysis was carried out using a hybrid QLT-Orbitrap (Thermo Scientific, Bremen Germany), which had been modified in-house to perform ETD [17, 18].  $MS^1$  analysis was executed in the Orbitrap at 30,000 resolving power, followed by six data-dependent MS/MS events with product ion analysis performed in the QLT. The list of included precursor ion targets was determined by descending intensity, using charge state rejection (precursors with two or more charges only) and dynamic exclusion (one fragmentation event led to exclusion for 30 s with a maximum peak list of 500). Each SCX fraction was analyzed using ETD, employing a reaction time of 63 ms, precursor cation target value of 10,000 ions, and an anion target value of 250,000 reagent ions. Duplicate analysis was performed on the first biologic replicate to ensure reproducibility of the data, while the other biologic replicate was analyzed a single time (three total analyses).

## Spectral Processing

The software used to generate .dta files (DTA Generator) is freely available on the Coon Lab website at <http://www.chem.wisc.edu/~coon/>. This software was written in-house, developed in C# with Microsoft Visual Studio 2005 and 2008, using the Microsoft .NET Framework 2.0. The XRawfile COM library (XRawfile2.dll) provided with Xcalibur 2.0.7 was used for access to the proprietary Thermo Scientific .raw binary data file format.

Input files for OMSSA (.dta) were generated directly from Thermo Scientific .raw LC-MS/MS data files. For each fragmentation spectrum, DTA Generator first attempts to determine the precursor charge state using the information recorded in the scan header. If this information was available, only a single .dta file was written using that precursor charge. If it was not available (could not be determined by the instrument firmware), a .dta file was generated for every charge state in a user-specified range. The precursor  $[M + H]^+$  value was written with its corresponding charge state, followed by every centroid  $m/z$  and intensity pair from the MS<sup>2</sup> scan. DTA Generator takes multiple input .raw files, automatically distinguishes between CAD- and ETD-generated spectra, allows the user to specify which spectral features should be removed (i.e., precursor, charge-reduced precursor(s), and/or neutral losses), and generates either individual .dta files or merges 10,000 .dta files into batch text files (for searching with OMSSA).

The software supports three different options: (1) removal of the precursor—all peaks  $\pm 3.1 m/z$  from the precursor  $m/z$  were removed; (2) removal of those charge-reduced precursor cations that were the result of ETnoD—all peaks  $\pm 3.1 m/z$  from the charge-reduced precursor  $m/z$  for each charge state from +1 to the precursor charge state -1; and (3) removal of those peaks that were the result of neutral losses from charge-reduced precursors—all peaks -60 Da (scaled to charge) to the charge-reduced precursor  $m/z$  for each charge state from +1 to the precursor charge state -1.

**Table 1.** Common neutral losses produced by electron capture dissociation, which are removed by the spectral processing algorithm [8–10]

Loss (Da)	Formula
17.027	NH <sub>3</sub>
18.011	H <sub>2</sub> O
27.995	CO
32.026	CH <sub>3</sub> OH
34.053	N <sub>2</sub> H <sub>6</sub> (2 × NH <sub>3</sub> )
35.037	H <sub>4</sub> NO
36.021	H <sub>4</sub> O <sub>2</sub> (2 × H <sub>2</sub> O)
44.037	CH <sub>4</sub> N <sub>2</sub>
45.021	CH <sub>3</sub> NO
46.006	CH <sub>2</sub> O <sub>2</sub>
46.042	C <sub>2</sub> H <sub>6</sub> O
59.037	C <sub>2</sub> H <sub>5</sub> NO
59.048	CH <sub>5</sub> N <sub>3</sub>

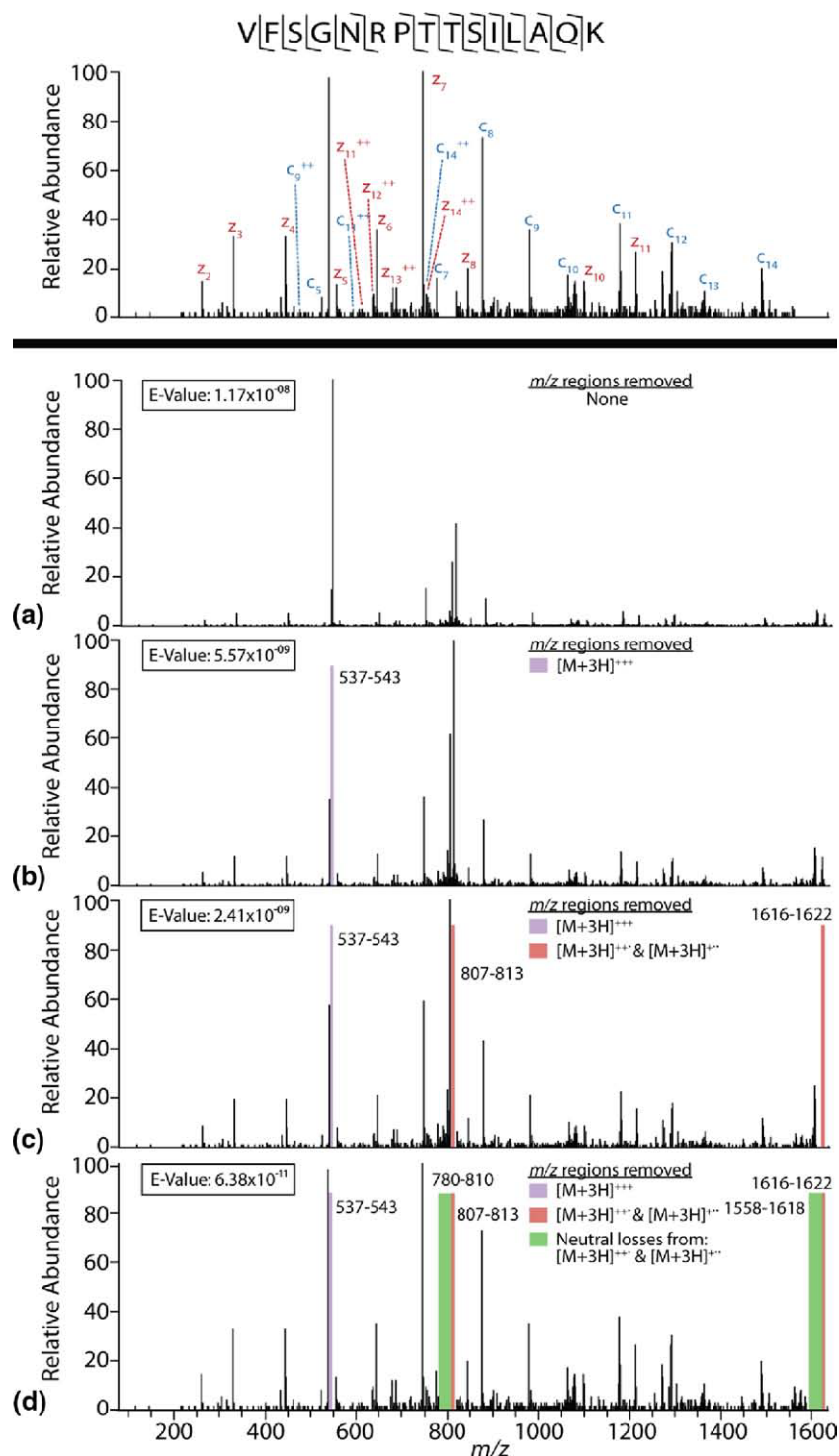
Table 1 contains a list of neutral losses which occur in the  $m/z$  region that is removed. For comparison of the impact that removal of different spectral features has on the probability of identification, spectra were processed in five ways: (1) performing all spectral processing steps; (2) performing the first two processing steps, but not removing neutral losses; (3) removing only the neutral losses, but not the precursor or charge-reduced precursors; (4) eliminating only the precursor; and (5) not processing the spectra in any way, thus providing a baseline for comparison.

## Database Searching

MS/MS spectra were searched against a concatenated target-decoy version of the *Saccharomyces* Genome Database (database reversed using a modified version of the open-source software available from the Gygi lab at: <http://gygi.med.harvard.edu/gygilab/index.php?html=software>) using OMSSA (Open Mass Spectrometry Search Algorithm; freely available from NCBI). The search algorithm parameters were set to consider the static modification of +57 Da on cysteine residues (carbamidomethylation), a variable modification of +16 Da on methionine residues (oxidation), a precursor mass tolerance of 4.0 Da, and a fragment mass tolerance of 0.5 Da. Search results were filtered to a 1% false discovery rate using software written in-house for optimization of identification expectation value and precursor mass error thresholds [19].

## Results and Discussion

A view onto how spectral processing affects the data submitted to a protein database search is provided in Figure 1a–d. Here, we display the step-wise, in silico removal of specified  $m/z$  regions from the original spectrum (Figure 1a), with the complexity of the spectrum clearly decreasing as more  $m/z$  regions are removed (Figure 1b–d). As is evident comparing Figure 1a and b, the precursor  $m/z$  is generally the greatest contributor to the total ion current within the spectrum. This, of course, is a function of ion/ion reaction duration. Note the precursor can be entirely removed by extending the ETD reaction duration; however, doing so also results in the destruction of the newly formed product ions as they are likewise subject to reaction with the ETD reagent. For a triply protonated precursor, the optimal reaction duration is one that removes ~2/3 of the initial amount of precursor cation; thus, a substantial amount of unreacted precursor remains. Figure 1c displays the spectrum that results from removal of the precursor  $m/z$  as well as the removal of all the charge-reduced precursors (ETnoD). Finally, Figure 1d illustrates the fully processed spectrum, where neutral losses from the charge-reduced precursors have also been removed. This MS/MS event provides an interesting example of how processing affects the possibility for identification, as only the fully processed spectrum provided an identification when searched



**Figure 1.** The effect of removing interfering peaks from the MS/MS scan #3118 from SCX fraction 6, biological replicate #1, first analysis. (a) Shows the unprocessed scan; (b)–(d) show the step-wise removal of the precursor, charge-reduced precursors, and neutral losses, respectively. Note how while all scans were identified by OMSSA, as the interfering peaks were removed, the confidence of the identification increased, and only that which was completely processed remained after trimming to a 1% FDR.

using the OMSSA database search algorithm and filtered for 1% FDR. Though OMSSA provided high-quality hits for all spectra, regardless of processing, the e-value for peptides identified from processed spectra

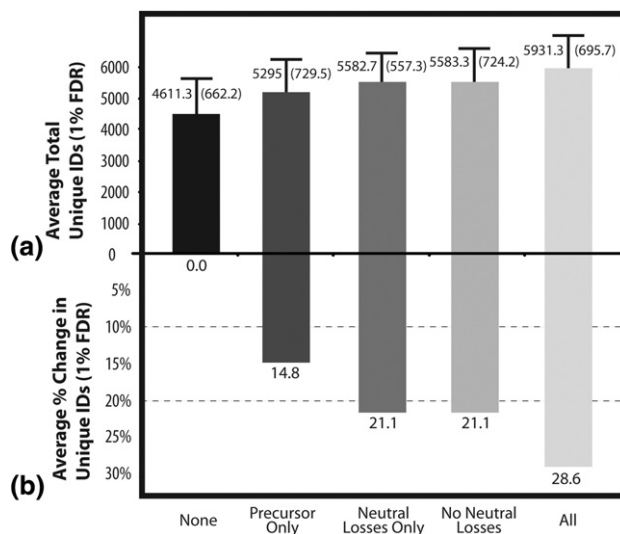
decreased (thus the probability of a correct match increased) with more extensive processing. In this case, the e-value decreased with each successive processing step, from  $1.17 \times 10^{-8}$  for no pre-submission processing



to  $6.38 \times 10^{-11}$  for the fully processed spectrum. In each case, OMSSA identified 21 of the 28 possible fragments for this peptide, resulting from 13 of the 14 possible bonds being cleaved. The only bond which was not dissociated was the N-terminal Arg to Pro bond, one which is not observed in ETD due to the ring structure of proline. With each successive processing step, non-matching peaks are removed which, in turn, decreases the chance that the match is a random occurrence. Interestingly, the more high confidence identifications there are in a dataset, the more low confidence identifications that can be included at a 1% false discovery rate. So, as a result of vigorous spectral processing, the standards for inclusion of an identification at the 1% FDR level are loosened—it is this trend that allowed this particular match to be included when it was subjected to all forms of cleaning.

Referring back to Figure 1, one can see the effect of removing the relatively high intensity  $m/z$  peaks that result from unreacted precursor and ETnoD products on the identification of a peptide. The precursor is the highest intensity peak within the spectrum, with the doubly charged, charge-reduced precursor being the second most intense peak. As OMSSA removes “noise” peaks dynamically within a range relative to the most intense peak, removal of these two most intense peaks drops the initial value for the maximum intensity within the spectrum by several-fold. Subsequent removal of the other interfering ions further reduces the theoretical maximum intensity within the spectrum, with the removal of each successively intense ion dropping the maximum by a relatively large percentage. It is therefore probable that in spectra like that illustrated in Figure 1, true  $c$ - and  $z^*$ -type ion peaks are not considered by OMSSA because of their low intensity relative to the precursor, ETnoD, and neutral loss peaks. This is especially problematic for ETD-generated spectra, as production of fragment ions is typically less efficient than in CAD (i.e., due to charge neutralization), thus causing these ions to have small fraction of the total intensity of the precursor and charge-reduced precursors. The observations made with this particular example will not necessarily hold true for every spectrum generated using ETD; precursor characteristics such as charge state can greatly affect the partitioning between ETD and ETnoD.

To assess the impact of spectral processing on a large amount of ETD data, we analyzed spectra produced from a large-scale experiment [14]. Yeast protein extracts (two replicates) were digested and fractionated by SCX into 12 pools. An nHPLC-MS/MS experiment was performed on each of the fractions, twice for the first replicate and once for the second for three sets of 12 analyses. Figure 2 presents the overall effects of removing varying  $m/z$  regions from these datasets before submission to a protein database search. All identifications were determined to be within a 1% FDR. The greatest increases in the quality of matches produced came from removal of the precursor (due to its contribution to the total ion current of the spectrum;  $\sim 15\%$



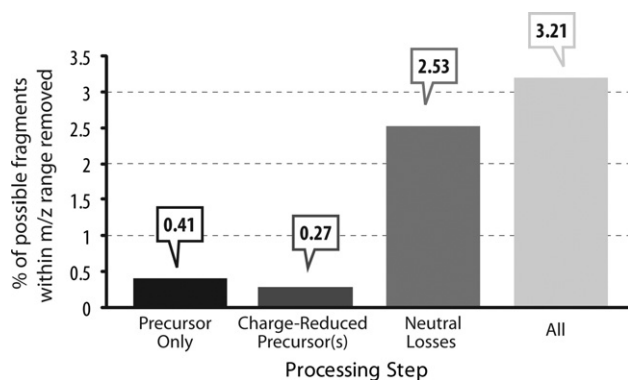
**Figure 2.** The effect of spectral pre-processing on total IDs from a proteomics experiment. (a) Provides the average number of unique (sequence exclusive) identifications at a 1% FDR for the three analyses based upon the amount of spectral processing; (b) illustrates the percent increase for these identifications as compared to no spectral processing.

average increase in unique IDs) and removal of neutral losses (likely an effect of removing multiple interfering peaks within a single step;  $\sim 21\%$  average increase in unique IDs). Subsequent removal of other interfering species showed gradual improvement of the total number of unique (sequence exclusive) identifications, with the combined removal of precursor, ETnoD, and neutral loss  $m/z$  ranges producing the greatest number of matches (5931). The collective effect of removing all of the interfering species listed above resulted in a substantial improvement in the total number of identifications,  $\sim 26.3\%$  increase.

An obvious concern for any type of spectrum pre-processing is the amount of information being removed. Figure 3 addresses this question by providing a measure of the average percentage of matching  $c$ - and/or  $z^*$ -type fragment ion peaks per peptide that fell within the  $m/z$  ranges that were removed during the spectral processing steps. In each step, a minute percentage of the total possible fragment ions was contained within the removed  $m/z$  region, and even when performing all processing steps, only  $\sim 3.2\%$  of all possible  $c$ - and/or  $z^*$ -type fragment ions per peptide, relative to the number of theoretical fragments that could be observed for each peptide, were removed during processing.

## Conclusions

We have illustrated the importance of removing interfering peaks from ETD-generated spectra before submission to a database search using the OMSSA algorithm. Note that the OMSSA algorithm has recently been updated to remove charge-reduced precursors



**Figure 3.** Analysis of sequence-informative peaks present in ETD filtering regions of tandem mass spectra. For precursors and charge-reduced precursors, only a fraction of a percent (0.41% and 0.27%, respectively) of the possible theoretical fragments are observed within the processed range. For neutral losses from charge-reduced precursors, this increases to 2.53%. Altogether, a relatively minor 3.21% of fragments are lost due to ETD processing.

and some of their associated neutral losses. This command was not employed during the OMSSA searches, and the algorithm presented here varies from that employed in OMSSA. Whereas OMSSA assumes a precursor charge of four and scales the windows removed by mass tolerance, our algorithm determines the windows based upon the charge contained within the .dta file (this information is obtained through collection of the MS data with the high-resolution Orbitrap mass analyzer). We anticipate the conclusions reached here are applicable for all peptide search algorithms, although the impact of spectral preprocessing could be less significant for algorithms designed specifically for ETD. Removal of these peaks led to a marked increase in total identifications and also identifications trimmed to a 1% false discovery rate. The step-wise removal of these categories of peaks was characterized and the relative impact of each was analyzed. The removal of the precursor from the spectra produced the greatest increase in total identifications, while subsequent removal of the charge-reduced precursor(s) and neutral loss  $m/z$  regions also provided gains in the number of identifications.

The software used to generate .dta files and remove the above-specified spectral features is freely available for download: <http://www.chem.wisc.edu/~coon/index.html>.

As this work focuses solely on the impact of removal of certain extraneous ETD-generated features on results produced from the OMSSA search algorithm, future analyses will be performed using multiple, commonly employed search algorithms, and will also investigate the effect of feature intensity on resulting IDs.

## Acknowledgments

The authors dedicate this manuscript to Alexander Makarov and to his wonderful achievement—the development of the Orbitrap mass analyzer. Alexander's innovation has transformed not only the authors' own research, but the entire field of proteomics. The

authors are honored to participate in this issue and to celebrate his considerable contribution. Congratulations Alexander!

D.M.G. and G.C.M. are grateful for support from NIH predoctoral fellowships (Biotechnology Training Program, NIH 5T32GM08349).

The authors acknowledge the University of Wisconsin, the Beckman Foundation, Eli Lilly, and the National Institutes of Health (1R01GM080148) for providing financial support for this work. The authors also thank Danielle Swaney and Heather Coon for generating the large-scale yeast dataset.

## References

- Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. Electron capture dissociation of multiply charged protein cations. A nonergodic process. *J. Am. Chem. Soc.* **1998**, *120*(13), 3265–3266.
- Coon, J. J.; Ueberheide, B.; Syka, J. E.; Dryhurst, D. D.; Ausio, J.; Shabanowitz, J.; Hunt, D. F. Protein identification using sequential ion/ion reactions and tandem mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*(27), 9463–9468.
- Syka, J. E.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*(26), 9528–9533.
- Molina, H.; Horn, D. M.; Tang, N.; Mathivanan, S.; Pandey, A. Global proteomic profiling of phosphopeptides using electron transfer dissociation tandem mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*(7), 2199–2204.
- Chi, A.; Huttenhower, C.; Geer, L. Y.; Coon, J. J.; Syka, J. E. P.; Bai, D. L.; Shabanowitz, J.; Burke, D. J.; Troyanskaya, O. G.; Hunt, D. F. Analysis of phosphorylation sites on proteins from *Saccharomyces cerevisiae* by electron transfer dissociation (ETD) mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*(7), 2193–2198.
- Good, D. M.; Wirtala, M.; McAlister, G. C.; Coon, J. J. Performance characteristics of electron transfer dissociation mass spectrometry. *Mol. Cell. Proteom.* **2007**, *6*(11), 1942–1951.
- Huang, Y. Y.; Triscari, J. M.; Tseng, G. C.; Pasa-Tolic, L.; Lipton, M. S.; Smith, R. D.; Wysocki, V. H. Statistical characterization of the charge state and residue dependence of low-energy CID peptide dissociation patterns. *Anal. Chem.* **2005**, *77*(18), 5800–5813.
- Cooper, H. J.; Hakansson, K.; Marshall, A. G.; Hudgins, R. R.; Haselmann, K. F.; Kjeldsen, F.; Budnik, B. A.; Polfer, N. C.; Zubarev, R. A. Letter: The diagnostic value of amino acid side-chain losses in electron capture dissociation of polypeptides. Comment on: "Can the (M(-)X) region in electron capture dissociation provide reliable information on amino acid composition of polypeptides?" *Eur. J. Mass Spectrom.* **2002**, *8*, 461–469 (2002). *Eur. J. Mass Spectrom.* (Chichester) **2003**, *9*(3), 221–222.
- Cooper, H. J.; Hudgins, R. R.; Hakansson, K.; Marshall, A. G. Characterization of amino acid side chain losses in electron capture dissociation. *J. Am. Soc. Mass Spectrom.* **2002**, *13*(3), 241–249.
- Fung, Y. M. E.; Chan, T. W. D. Experimental and theoretical investigations of the loss of amino acid side chains in electron capture dissociation of model peptides. *J. Am. Soc. Mass Spectrom.* **2005**, *16*(9), 1523–1535.
- Falsh, M.; Savitski, M. M.; Nielsen, M. L.; Kjeldsen, F.; Andren, P. E.; Zubarev, R. A. Analytical utility of small neutral losses from reduced species in electron capture dissociation studied using SwedECD database. *Anal. Chem.* **2008**, *80*(21), 8089–8094.
- Savitski, M. M.; Nielsen, M. L.; Zubarev, R. A. Side-chain losses in electron capture dissociation to improve peptide identification. *Anal. Chem.* **2007**, *79*(6), 2296–2302.
- Geer, L. Y.; Markey, S. P.; Kowalak, J. A.; Wagner, L.; Xu, M.; Maynard, D. M.; Yang, X. Y.; Shi, W. Y.; Bryant, S. H. Open mass spectrometry search algorithm. *J. Proteome Res.* **2004**, *3*(5), 958–964.
- Swaney, D. L.; McAlister, G. C.; Coon, J. J. Decision tree-driven tandem mass spectrometry for shotgun proteomics. *Nat. Methods* **2008**, *5*(11), 959–964.
- Villen, J.; Beausoleil, S. A.; Gerber, S. A.; Gygi, S. P. Large-scale phosphorylation analysis of mouse liver. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*(5), 1488–1493.
- Martin, S. E.; Shabanowitz, J.; Hunt, D. F.; Marto, J. A. Subfemtomole MS and MS/MS peptide sequence analysis using nano-HPLC micro-ESI Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* **2000**, *72*(18), 4266–4274.
- McAlister, G. C.; Berggren, W. T.; Griep-Raming, J.; Horning, S.; Makarov, A.; Phanstiel, D.; Stafford, G.; Swaney, D. L.; Syka, J. E. P.; Zabrouskov, V.; Coon, J. J. A proteomics grade electron transfer dissociation-enabled hybrid linear ion trap-Orbitrap mass spectrometer. *J. Proteome Res.* **2008**, *7*(8), 3127–3136.
- McAlister, G. C.; Phanstiel, D.; Good, D. M.; Berggren, W. T.; Coon, J. J. Implementation of electron-transfer dissociation on a hybrid linear ion trap-orbitrap mass spectrometer. *Anal. Chem.* **2007**, *79*(10), 3525–3534.
- Elias, J. E.; Gygi, S. P. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat. Methods* **2007**, *4*(3), 207–214.