We would like to thank the reviewers for their comments. The resulting changes will provide a richer and clearer discussion of how alternative fragmentation modes affect peptide identification.

# Reviewer 1

COMMENTS FOR THE AUTHOR:

The authors present a review of peptide identifications via MS/MS using fragmentation techniques beyond CID. The manuscript is well written and covers the topic with sufficient detail and with appropriately extensive citations of the relevant literature, which will aid readers in finding the primary works in this field. However, before it is suitable for publication, I recommend that the authors address these comments:

1) Figure 1 depicts a single interesting example of complementary CID and ETD spectra, but it is quite difficult to see the break indicators in the amino acid series as they are very thin and faint. They were not discernable on my printed copy, although can be seen when blown up in the electronic version. It would be helpful to make this figure more easy to read.

We have updated the figure with larger peptide break indicators:



2) The reader is left wondering if the complementarity in figure 1 is merely a rare, lovely example or if this is quite prevalent. It would be interesting to calculate and show a distribution of a metric of complementarity between two fragmentation types in the datasets that are analyzed here.

Although the original manuscript had statistics measuring the complementarity between CID, HCD, and ETD over all identified spectra (in Figure 2 and Table SM-1), they were not well integrated into the main text. We have discussed in the “Peptide fragmentation modes” section (at the end of the first paragraph) how these statistics demonstrate complementarity between alternative fragmentation modes:

“… Of peptides that can be identified by all three acquisition modes, the *breaks* (observed cleavages along the peptide backbone, supported by either N- or C-terminal fragments) captured by ETD tend to complement those captured by CID and HCD, especially for precursors of charge 3 or higher. This can be seen in statistics from Figure 2, which show how the union of observed peptide breaks increases by 24-72% from CID/HCD to CID/HCD/ETD for precursors of charge 3 or higher. Supplementary table SM-1 details how many breaks are unique to every possible combination of fragmentation modes: ETD alone accounts for 19% of all possible peptide breaks in CID/HCD/ETD triplets, the intersection of breaks seen in CID and HCD accounts for 17%, and the intersection of breaks seen in CID, HCD, and ETD accounts for 30% … (Page 4-5)”

3) The terminology in the legends of figures 1 and 2 seem incongruent. The figure 1 legend refers to “breaks” (which might be a bit too terse). The figure 2 title refers to “N/C-Term Breaks”, and the figure 2 legends refers to “N/C-terminal peptide cleavages”. I infer that these all refer to the same concept, but it reduced clarity to use different terms. In particular, “cleavage” is usually used in a different context. Or perhaps I have misunderstood and these are different somehow. In either case, more consistent use of terminology would help clarify the situation.

We streamlined the terminology such that “breaks” is the only term used: “ … *breaks* (observed cleavages along the peptide backbone, supported by either N- or C-terminal fragments) … (Page 5)“

4) References 83 and 84 probably need to be updated. 83 is no longer in press and 84 doesn't even mention Tranche, but rather just a Java framework. Perhaps a more apt reference would be: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2292060/>

The appropriate reference has been updated as suggested.

5) Overall the manuscript would benefit from some additional editing to produce a tidier manuscript. The occasional use of “-“ terminate sentences reduces readability. There also appears to be excessive capitalization for terms that are not proper nouns throughout the manuscript. For example on page 2, “collision induced dissociation” is differently capitalized and differently punctuated and has its acronym twice defined. Other fragmentation types are needlessly capitalized. There are several other such distractions throughout the manuscript that should be corrected.

Inappropriate capitalizations and uses of “-“ have been revised throughout the text.

6) the authors comment in the conclusion that the situation is improving as more tools support alternative fragmentation modes. It might be worth adding the point that more intelligent, integrated analysis of multiple fragmentation mode datasets by more tools would perhaps be even more beneficial.

The issue of too few tools supporting alternative fragmentation modes has been raised in the “Conclusion and outlook” section:

“… Although the concept of combining fragmentation modes is over a decade old49, most existing scoring functions for peptide identification process CID, HCD, and/or ETD spectra individually … (Page 9)”

# Reviewer 2

COMMENTS FOR THE AUTHOR:

This article does a nice job of outlining the current state of affairs of CID, ETD, and HCD for MS/MS. The emphasis is placed on informatics tools for integrating these scan modes for maximized results. This reviewer agrees strongly that full integration of all these technologies is currently limited by software. Overall I think the review is very well written and quite straightforward to follow. I think this will be very valuable for the field and is timely. I recommend its publication and only have one very minor comment. Page 2 first paragraph, “Current protocols can regularly identify thousands of proteins…….per experiment and can deliver very high levels of reproducibility”. I am not sure what is meant by “reproducibility” but I think most would consider that the overlap of identifications in one shotgun experiment to the next. In that view I would say most believe that shotgun methods are not highly reproducible.

The reviewer is correct in that peptide identifications are not commonly reproducible from one bottom-up shotgun experiment to the next. Although our statement was not clear, we were referring to targeted approaches as reproducible by citing reference 3, which is a review on targeted proteomics protocols. The statement has been revised:

“… Current protocols regularly identify thousands of proteins and post-translational modifications (PTMs2) per experiment and can deliver very high levels of reproducibility3 (e.g. using targeted approaches) …”

# Reviewer 3

COMMENTS FOR THE AUTHOR:

This article reviews current efforts to make use of the combination of different fragmentation approaches to provide complementary information for peptide identification. It is a thorough review and could be published as is. However, I would suggest expanding on a couple of points. The effect of enzymes could be discussed more. The authors state that trypsin contributes to peptide ID by cleaving C-term of K/R, but they don’t really explain why this is particularly beneficial. The use of different enzymes to get optimal information from different fragmentation approaches has been proposed; e.g. many groups have recommended the use of Lys-C prior to ETD analyisis and the Heck group has promoted the combination of Lys-N and ETD, especially for de novo sequencing. Secondly, the authors should discuss potential differences in performance of search engines at analyzing different data types; i.e. are software that were primarily developed for analyzing ion trap CID data going to perform as well with HCD data? How simple is it to convert a search engine to analyze ETD instead of CID data? With this in mind, are there advantages/disadvantages to using different search engines to analyze the different data types?

Given the popularity of using non-tryptic enzymes with alternative fragmentation modes, a discussion on the effect of different enzymes is certainly appropriate. We have added to the “Introduction” section a high-level overview of enzyme digestion strategies (e.g. using a single specific enzyme, non-specific enzyme(s), or multiple specific enzymes) and how they affect both peptide fragmentation as well as the complexity of peptide identification. We have also mentioned the particular Lys-C/ETD and Lys-N/ETD combinations.

Regarding the performance of search engines with multiple fragmentation modes, we have revised the second paragraph of the “Peptide identification” section to include such a discussion. The point was made that CID scoring functions should be re-designed for HCD and ETD. It is difficult to speculate on how challenging it would be to update a search engine, so we could only comment on the search engine we are most familiar with (MS-GFDB):

“… Although database search tools developed for CID can be used for HCD and easily adapted to process ETD (by simply examining c/z\_ ion offsets instead of b/y), they would likely perform much worse than tools that are sensitive to unique features of HCD and ETD. Features of HCD that are not captured by CID scoring models include peaks in the low m/z range (including immonium ions), high fragment mass resolution (most CID spectra have low fragment mass resolution) and the presence of internal ions. Aside from c/z\_ ions, ETD spectra typically contain charge-reduced precursor peaks with high intensity, characteristic losses from charge-reduced precursors and additional related ions at offsets ±H from c and z\_ ions. Thus, CID scoring models should be re-designed for HCD and ETD in order to be most effective. The difficulty of the adaptation depends on the algorithm being considered. For example, MS-GFDB36 can be automatically re-trained for new types of spectra with only 1,000 PSMs from unique peptides (per precursor charge state) … (Page 6)”

At the end of the “Estimation of false discovery rates” section, we have also added a comment on using different search engines to analyze different fragmentation modes:

“… It is expected that combining search engines for CID, HCD and/or ETD also improves results, but it is important to combine these approaches with appropriate estimation of FDR (such as iProphet68)… (Page 8)”