Fragmentation Patterns of EThcD Spectra of Phosphopeptides and Glycopeptides

Marshall Bern[1] Yong J. Kil[1] Wilfred Tang[1] Christopher Becker[1] Christian K. Frese[2] Maarten Altelaar[2] Shabaz Mohammed[2] Albert J. R. Heck[2] John Syka[3] Ryan Bomgarden[3] Rosa Viner[3]

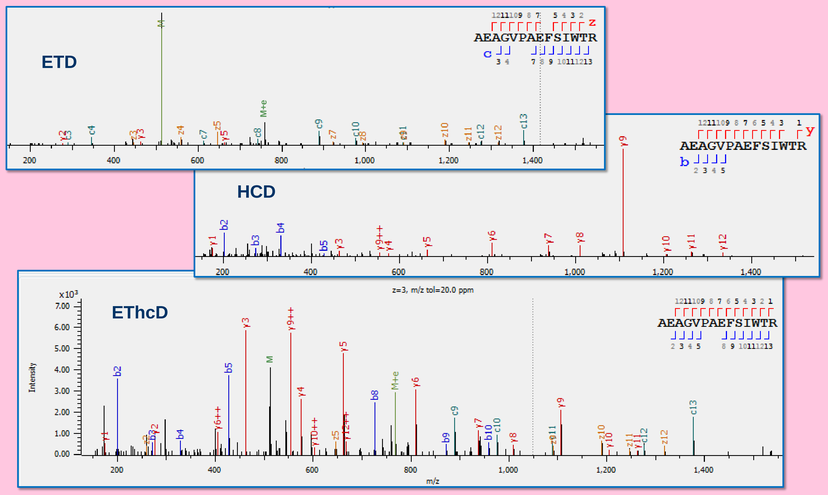
[1]Protein Metrics Inc., San Carlos, CA [2]Biomolecular Mass Spectrometry and Proteomics, Utrecht University, Utrecht, Netherlands [3]Thermo Fisher Scientific, San Jose, CA

Contact: bern@proteinmetrics.com

Introduction

Electron transfer dissociation (ETD) and beam-type collision-induced dissociation (HCD) are combined in a new fragmentation method called EThcD on Thermo Scientific Orbitrap instruments. This method first applies ETD in the linear ion trap and then transfers all the ions, both precursors and products, to the HCD collision cell for further fragmentation. An EThcD spectrum includes b-, c-, y-, and z-ions, and is roughly equivalent to the union of an ETD and an HCD spectrum. EThcD has been shown to give more complete fragmentation of unmodified peptides than either ETD or HCD alone [Frese et al, 2012], and to give more confident site localization of phosphorylations [Frese et al, 2013], yet many basic questions about the fragmentation method remain. In this poster we give statistical and anecdotal answers to these questions:

- Which fragmentation mode (ETD, HCD, or EThcD) gives the most identifications?
- How complete is EThcD fragmentation? Which ions predominate?
- How do EThcD spectra vary with precursor charge and HCD collision energy?
- Does EThcD give good fragmentation even on multiply phosphorylated peptides?
- Does an EThcD spectrum of an N-glycopeptide show the glycan fragmentation of an HCD spectrum along with the peptide fragmentation of an ETD spectrum?
- Does an EThcD spectrum of an O-glycopeptide allow for modification site localization?



Example spectra of a triply charged unmodified peptide from human Filamin-A

ETD spectra often have large amounts of unfragmented precursor (denoted M or M+e for singly charge-reduced) and small fragment peaks. HCD spectra tend to have few b-ions above m/z 800 because larger b-ions continue to fragment with the higher collision energy. EThcD converts uninformative unfragmented precursor to b and y ions, while retaining the c- and z-ions from ETD.

Sample Preparation and Data Acquisition

HeLa. The sample is HeLa cell lysate, digested with trypsin, fractionated into 50 fractions by SCX, and then analyzed by LC-MS/MS using ETD with supplemental activation, HCD and EThcD on an ETD-equipped Thermo Scientific Orbitrap Velos, as described previously [Frese et al, 2012]. All mass analysis was performed in the Orbitrap with resolution 30,000 for MS and 7500 for MS/MS. Normalized collision energy was 30% for both HCD and EThcD. This sample contains over 4000 identifiable proteins, and hence is a good test of EThcD for large-scale proteomics.

Ti-IMAC. The sample is HeLa cell lysate, digested with trypsin, and enriched for phosphopeptides with Ti^{4+}-IMAC beads as described previously [Frese et al, 2013]. The sample was split into three equal amounts and directly analyzed by single run LC-MS/MS using, respectively, ETD with supplemental activation, HCD and EThcD on an ETD-equipped Thermo Scientific Orbitrap Velos. All mass analysis was performed in the Orbitrap.

Serum. The sample is human blood serum, enriched for glycopeptides using OasisMAX SPE column (Waters) in HILIC mode, dried and resuspended in water before LC-MS analysis. LC-MS analysis was performed using an Easy nLC1000 nano-flow system with 75 µm x 20 cm spray tip Magic C18 column and a Thermo Scientific Orbitrap Elite mass spectrometer. Equal quantities were analyzed with (1) HCD-pd-ETD fragmentation [Saba et al, 2012] with ETD triggering on oxonium ions, and (2) EThcD. Both (1) and (2) gathered MS/MS spectra on the top 10 most intense precursors. Several different HCD collision energies were tried for EThcD in order to study the effect on the spectra. A number of major serum proteins contain abundant and well-studied N-linked glycosylation.

Glycophorin-A. A sample containing mostly human glycophorin-A (along with small quantities of about 100 other proteins) was purchased from Sigma, reduced, alkylated, and trypsin-digested. Equal quantities were analyzed by, respectively, ETD, HCD, and EThcD on an Orbitrap Elite. One additional aliquot was labeled with TMT^2 to increase charge and analyzed by EThcD with HCD collision energy 30%. All mass analysis was performed in the Orbitrap. Glycophorin-A is the primary glycoprotein in red blood cell membranes, and includes a heavily O-glycosylated extracellular mucin domain.

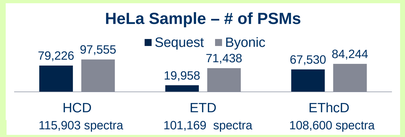
Data Analysis

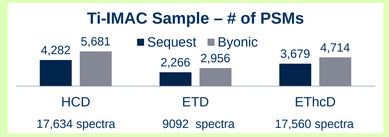
We added EThcD as a fifth fragmentation option in the Byonic [Bern et al, 2007] search engine: low-energy CID, beam-type CID (meaning QTOF and HCD), ECD/ETD, MALDI-TOF-TOF, and now EThcD. We did not train Byonic on EThcD spectra, but simply predicted a theoretical EThcD spectrum to be the sum of theoretical HCD and ETD spectra. We wrote special-purpose software to compile statistics of annotated fragment peaks in Byonic's top-scoring identification; these peaks are roughly the top 200 peaks by "local intensity" that matched theoretical ions within 0.5 Da for ion-trap mass analysis, and within 20 ppm for Orbitrap mass analysis. Byonic's peak annotations closely agree with human expert annotations.

Previous studies (Frese et al, 2012; Frese et al, 2013) on the same HeLa and Ti-IMAC data used Sequest, and directed Sequest to predict b, c, y, and z ions. For these data sets, we show both Byonic and Sequest results.

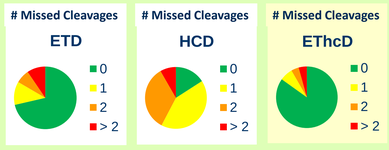
Results

Which fragmentation mode gives the most identifications. Sequest analysis used Percolator to give a list with 1% PSM FDR and unknown protein FDR. Byonic analysis used the 2D target/decoy strategy [Bern & Kil, 2012] to give a list with 1% protein FDR and PSM FDR ranging from 0.1% to 0.5%, depending on the sample.

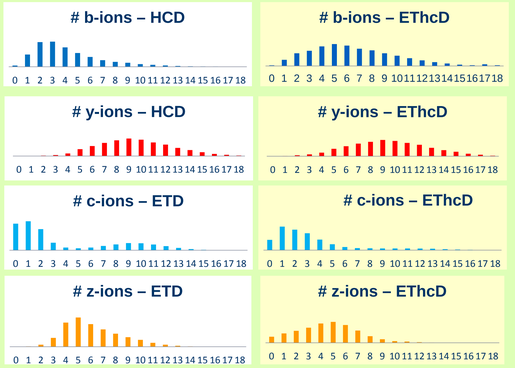




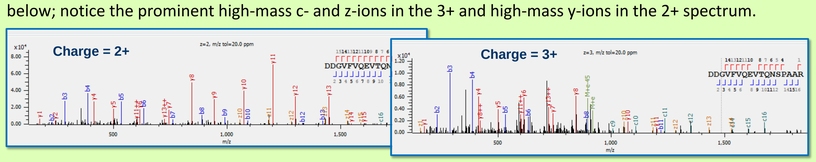
Fragmentation Completeness. As in [Frese et al, 2012] we addressed this question by plotting the number of missed cleavages (no c, c-1, z, z+1, b, y, or y++ ion) over all well-identified PSMs in the HeLa data set. HCD frequently misses the first and/or last cleavage, but EThcD gives all cleavages on 85% of spectra.



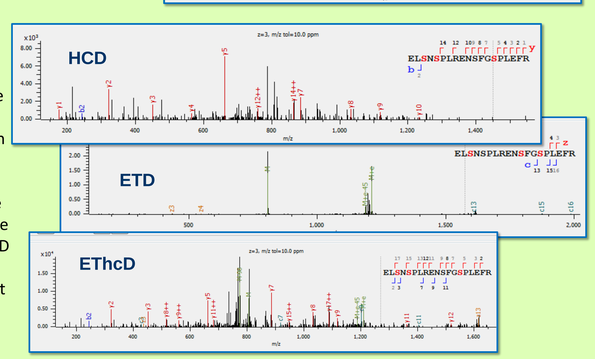
Which Ions Predominate. We compiled histograms of the number of identifiable c, z, b, y, (but not c-1, z+1, or y++) peaks in the identified spectra. The vertical axis shows fraction of total, so that bars sum to one. As shown here, the identifiable EThcD spectra in this data set have as many y-ions as do HCD spectra, and have more b-ions than do HCD spectra, but tend to have fewer z-ions than do ETD spectra. One anomaly: the ETD spectra gave a bimodal pattern for c-ions, because charge-2 precursors give more c-1 than c-ions. Interestingly, EThcD does not show this anomaly.



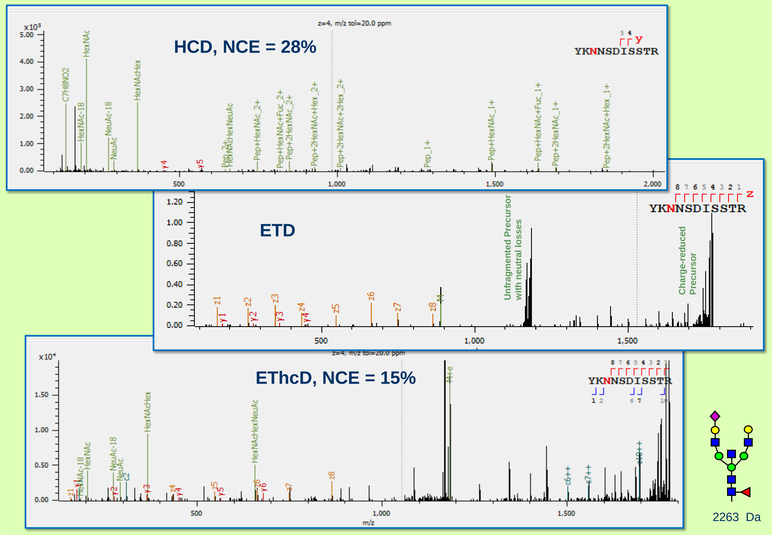
Precursor Charge and Collision Energy. EThcD spectra vary from almost pure ETD to almost pure HCD, depending upon the amount of initial ETD fragmention (more charge → more fragmentation) and subsequent HCD fragmentation (more collision energy → more fragmentation). We show the charge effect below; notice the prominent high-mass c- and z-ions in the 3+ and high-mass y-ions in the 2+ spectrum.



Multiple Phosphorylations. As reported previously [Frese et al, 2013], EThcD gives more complete fragmentation and more reliable phosphorylation site localization than either HCD or ETD alone. On this triply phosphorylated peptide from serine/arginine repetitive matrix protein (SRRM2), EThcD and HCD could place all three phosphorylations, but ETD left ambiguity due to poor fragmentation.



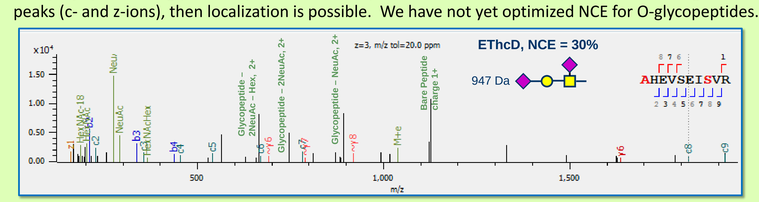
N-Glycopeptides. EThcD does indeed fragment both peptide and glycan for precursors with charge at least 4+. We tried normalized collision energies from 5 – 30%, and found that 10 – 15% gave the best results. EThcD tends to give better Orbitrap spectra than ETD.



Example spectra of a 4+ N-glycopeptide from human Ig mu chain C

HCD spectra of N-glycopeptides tend to show singly charged glycan fragments (oxonium ions) and intact peptide plus base monosaccharides in various charge states, with little to no peptide backbone fragmentation. ETD spectra show peptide backbone fragmentation along with abundant unfragmented precursor. EThcD spectra show both glycan and peptide fragmentation. The glycan in the glycopeptide shown has mass 2262.8 Da, corresponding to 5 HexNAc, 5 Hex, 1 Fuc, and 1 NeuAc, with a possible cartoon depicted.

O-Glycopeptides. O-glycans stay on in ETD but fly off in HCD; if the EThcD spectrum has enough ETD peaks (c- and z-ions), then localization is possible. We have not yet optimized NCE for O-glycopeptides.



Example spectrum of a 3+ O-glycopeptide from Glycophorin-A

In this spectrum, the O-glycan is well localized by the c-ion series. The O-glycan tends to fly off the b- and y-ions, and we denote by ~y6, ~y7, etc., the ions with all O-glycans off. In this spectrum, we see a small peak annotated y6; this is the y6 ion with the O-glycan still on. With EThcD, we obtained only a few PSMs with multiple O-glycans. With ETD-pd-HCD (data not shown) we obtained more.

Discussion

Although very new, EThcD is already proving to be a valuable method, giving in one spectrum roughly the same information as an HCD/ETD pair. Especially interesting is the observation that EThcD on the HeLa data is actually giving more b- and y-ions than is HCD, perhaps because the HCD part of EThcD is fragmenting precursors in more than one charge state. EThcD appears especially advantageous for de novo sequencing, phosphopeptide site localization, and N-glycopeptides. We cannot reach any firm conclusions about O-glycopeptides yet. Site localization in this case depends almost entirely upon the ETD peaks, so we expect that normalized collision energies of 5 – 10% will be needed to preserve intact glycosylation, but have not yet tried energies this low. More research into EThcD method optimization is needed, due to the dependence of EThcD spectra on precursor charge, collision energy, modifications, and sequence.

Acknowledgment: This research was supported in part by NIH grant R43 GM103352.

References

1. Frese CK, Altelaar AF, van den Toom H, Nolting D, Griep-Raming J, Heck AJ, Mohammed S. Towards full peptide sequence coverage… *Anal. Chem.* 2012 84(22), 9668-9673.
2. Frese CK, Zhou H, Tau T, Altelaar AF, Mechtler K, Heck AJ, Mohammed S. Unambiguous phosphosite localization using electron-transfer/higher-energy collision dissociation (EThcD). *J. Proteome Res* 2013, 12(3), 1520-1525.
3. Saba J, Dutta S, Hemenway E, Viner R. Increasing the productivity of glycopeptides analysis by using higher-energy collision dissociation-accurate mass-product-dependent electron transfer dissociation. *Int J. Proteomics 2012.*
4. Bern M, Cai Y, Goldberg D. Lookup peaks: a hybrid of de novo sequencing and database search for protein identification by tandem mass spectrometry. *Anal. Chem.* 2007, 79 (4), 1393 – 1400.
5. Bern M, Kil YJ. Two-dimensional target decoy strategy for shotgun proteomics. J. Proteome Res. 2011, 10(12), 5296-5301.