



Published in final edited form as:

Proteomics. 2011 June ; 11(12): 2568–2572. doi:10.1002/pmic.201000547.

Systematic evaluation of alternating CID and ETD fragmentation for phosphorylated peptides

Min-Sik Kim^{1,2}, Jun Zhong¹, Kumaran Kandasamy^{1,4}, Bernard Delanghe⁵, and Akhilesh Pandey^{1,2,3}

¹McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21205

²Department of Biological Chemistry, Johns Hopkins University School of Medicine, Baltimore, MD 21205

³Department of Pathology and Oncology, Johns Hopkins University School of Medicine, Baltimore, MD 21205

⁴Institute of Bioinformatics, International Technology Park, Bangalore, 560066, India

⁵Thermo Fisher Scientific (Bremen) GmbH, Bremen, Germany

Abstract

CID has become a routine method for fragmentation of peptides in shotgun proteomics while electron transfer dissociation (ETD) has been described as a preferred method for peptides carrying labile PTMs. Though both of these fragmentation techniques have their obvious advantages, they also have their own drawbacks. By combining data from CID and ETD fragmentation, some of these disadvantages can potentially be overcome because of the complementarity of fragment ions produced. To evaluate alternating CID and ETD fragmentation, we analyzed a complex mixture of phosphopeptides on an LTQ-Orbitrap mass spectrometer. When the CID and ETD-derived spectra were searched separately, we observed 2504, 491, 2584 and 3249 phosphopeptide-spectrum matches from CID alone, ETD alone, decision tree-based CID/ETD and alternating CID and ETD, respectively. Combining CID and ETD spectra prior to database searching should, intuitively, be superior to either method alone. However, when spectra from the alternating CID and ETD method were merged prior to database searching, we observed a reduction in the number of phosphopeptide-spectrum matches. The poorer identification rates observed after merging CID and ETD spectra are a reflection of a lack of optimized search algorithms for carrying out such searches and perhaps inherent weaknesses of this approach. Thus, although alternating CID and ETD experiments for phosphopeptide identification are desirable for increasing the confidence of identifications, merging spectra prior to database search has to be carefully evaluated further in the context of the various algorithms before adopting it as a routine strategy.

Keywords

Calyculin A; Combination of CID and ETD spectra; Phosphoproteomics

Correspondence: Dr. Akhilesh Pandey, McKusick-Nathans Institute of Genetic Medicine and Departments of Biological Chemistry, Pathology and Oncology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, pandey@jhmi.edu, Fax: +1-410-502-7544.

MSK, JZ, KK and AP have declared no conflict of interest. B.D. is an employee of Thermo Fisher Scientific.

Collision induced dissociation (CID), one of the key modes of fragmenting peptides ions in the gas phase, is widely used for shotgun proteomics although it is not the ideal method for localizing labile post-translational modifications (PTMs) such as serine/threonine phosphorylation and O-GlcNAc. In recent years, other methodologies for fragmentation have been developed to obtain more informative MS/MS spectra. These newer methods to improve PTM site identification include gas phase reaction of positive peptides with electron species (i.e. electron capture dissociation (ECD) and electron transfer dissociation (ETD)), resulting in the production of c and z ions [1–3]. Only recently, it has become possible to employ different fragmentation methods on a single mass spectrometer that allows one to generate MS/MS spectra from the same peptide ion species in an alternating mode. There are several types of alternating acquisition modalities including CID and ETD, CID and pulsed-Q dissociation (PQD), CID and HCD [4–6]. However, both alternating CID and PQD and alternating CID and HCD produce the same type of fragment ions (i.e. b and y ions), while alternating CID and ETD generate complementary types of ions (i.e. b, c, y, and z ions). It has been shown that there is a complementarity between the fragment ions produced by CID and ETD [7]. In spite of the knowledge regarding the complementarity of ions produced, the effect of combining the information from the different types of fragment ions prior to searching databases has not been evaluated in the context to global PTM analysis.

Because of the labile nature of serine/threonine phosphorylation, a complex mixture of phosphorylated peptides is a good model system to examine the effects of combining fragment ions on peptide identification. In order to generate a complex phosphopeptide mixture, we treated a pancreatic cancer cell line (Panc198) with Calyculin A, a serine/threonine phosphatase inhibitor, for 30 min [8] and lysed cells directly in 9 M urea-containing buffer. Twenty milligrams of extracted protein were processed for trypsin digestion and 10 mg of the tryptic digest was fractionated on a strong cation exchange column [9]. One early fraction and a middle fraction were chosen for phosphopeptide enrichment (roughly equivalent to 1 mg of protein). Approximately 100 mg of TiO₂ material (Titansphere, 10 μm, GL Sciences Inc., Japan) was preincubated with 3% 2,5-dihydroxybenzoic acid (DHB, Sigma) in 80% acetonitrile containing 1% TFA for 2 hours. 30% of DHB solution in 80% acetonitrile, 1% TFA was added to the fractions. DHB-prebound TiO₂ material was transferred to the sample tubes containing ~1 mg peptides and rotated for 2 hours at room temperature. Bound peptides were washed three times with 1% TFA in 80% acetonitrile and eluted with 2% ammonium hydroxide solution in 40% acetonitrile [10]. The eluates were desalted on Sep-Pak (Waters) cartridges and analyzed.

To evaluate the alternating CID and ETD method, we used four fragmentation modes - CID alone, ETD alone, decision tree-based CID/ETD and alternating CID and ETD. This analysis was done on an LTQ-Orbitrap XL ETD, which is a hybrid instrument combining a linear ion-trap with an Orbitrap analyzer such that CID, ETD and PQD can be carried out in the linear ion-trap while higher energy collision induced dissociation (HCD) is performed in the HCD cell. The mass spectrometer was set to acquire MS/MS scans using the three or six most abundant ions from each precursor scan in different fragmentation modes. A total of 30 LC-MS/MS runs were carried out (i.e. five experimental conditions for 2 fractions in triplicate, $5 \times 2 \times 3 = 30$), which generated ~140,000 MS/MS spectra. For precursor scans, peptide ions were accumulated in the C trap with a target value of 1×10^6 and a maximum injection time of 500 ms, followed by detection in the Orbitrap at a resolution of 30000 [9]. For CID experiments, fragment ions were accumulated in the ion trap with a target value of 1×10^4 with a maximum injection time of 200 ms. Multistage activation acquisition mode was enabled with neutral loss masses of 32.66, 48.99 and 97.97. For ETD experiments, fragment ions were accumulated in the ion trap with a target value of 2×10^5 and a maximum injection time of 200 ms. The reagent ion source emission current, reagent ion

electron energy and reagent ion source CI pressure were set to 35 μ A, -70 V and 20 psi, respectively. Activation time was set to 100 ms and supplemental activation mode was enabled. For alternating CID and ETD experiments, the same precursor ion was fragmented successively by CID and ETD. Ions selected for fragmentation were dynamically excluded within a ± 7 ppm m/z window for 90 seconds. For decision tree-based CID/ETD experiments, parameters for CID and ETD modes were same as that for CID only and ETD only. In addition, precursors lower than 650.00 m/z (+3), 900.00 m/z (+4) or 950.00 ($+5$) were subjected to ETD or else to CID for fragmentation. Phosphopeptides were separated on a 12-cm analytical column packed with C_{18} -bound material (5 μ m size, 100 \AA pore). The linear gradient of solvent B (90% acetonitrile with 0.1% formic acid) was used from 3% to 7% for 3 min, from 7% to 27% for 57 min, from 27% to 100% for 15 min, 100% for 4 min, 0% for 10 min to regenerate the columns. For data analyses, mass spectral data were processed, exported and searched against a human protein database using SEQUEST and Mascot by Proteome Discoverer (beta version 1.2.0.208). Reverse database searches were carried out to estimate false discovery rates (FDRs). Search criteria used were: tryptic enzyme specificity with up to 2 missed cleavages, carbamidomethylation at cysteine residues as a fixed modification, phosphorylation at serine, threonine and tyrosine, acetylation at N-termini of protein and oxidation at methionine residues as variable modifications, 20 ppm for precursor ion tolerance and 0.8 Da for fragment ion tolerance. Exported mgf files were used for X!Tandem searches, where 0.1 of e-values at both protein and peptide level were used as a cutoff ($\sim 2\%$ FDR). CID-driven data were searched for b and y-ions, while ETD-driven data were first processed (see [11] for details.) and then searched for c, y, z+1 and z+2-ions in Mascot, for c, z+1 and z+2-ions in X!Tandem and for c and z-ions in SEQUEST.

At 1% FDR using SEQUEST, 2504, 491, 2584 and 3249 phosphopeptide-spectrum matches (phosphoPSMs) were identified by CID alone, ETD alone, decision tree-based CID/ETD and alternating CID and ETD, respectively. These data reveal that very similar numbers of phosphoPSMs were obtained by CID alone or decision tree methods, while alternating CID and ETD showed about 30% higher number of phosphoPSMs as compared to that from CID (Figure 1A). A relatively low number of phosphoPSMs was obtained by ETD alone. Although the number of phosphoPSMs in alternating CID and ETD was higher, the total number of unique phosphopeptides was quite similar (Figure 1B). Overall, we conclude that phosphopeptides identified by the 3 modes involving CID do not differ greatly from each other. We also carried out alternating CID and ETD experiments where only the top 3 ions were fragmented. However, we observed a similar number of phosphoPSMs and phosphopeptides when top 3 ions were fragmented as for the top 6 method as shown in Figure 1A and 1B, respectively and is in agreement with that the total number of acquired MS/MS spectra in both modes was similar (i.e. 32124 for the top 3 method and 33308 for the top 6 method).

One of the goals of employing alternating CID and ETD method is to take maximum advantage of the spectral data generated from two complementary fragmentation methods for large scale PTM analysis. However, the current approach for searching MS/MS spectra from alternating CID and ETD experiments is to extract the alternating CID and ETD spectra and to search them separately followed by merging of the search results. Intuitively, one might expect a higher identification rate and better coverage of peptide sequences if the two types of fragment ions were combined prior to searching. To test this, we combined the CID and ETD spectra from each precursor ion to generate a mixed spectrum using Proteome Discoverer. The alternating spectra were first normalized using 'Spectrum Normalizer' node and subsequently merged using the 'Spectrum Grouper' node. The criteria for merging two spectra were 10 ppm mass error tolerance and elution within a 0.1 min retention time window. These normalized and merged spectra were searched using SEQUEST, X!Tandem

and Mascot which match b, y, c and z ions in the same search. Merged spectra from the alternating CID and ETD methods using top 3 ions (Figure 2A, 2C and 2E) by SEQUEST, X!Tandem or Mascot led to 1875, 1715 or 3652 phosphoPSMs, respectively. Similar results were observed using the top 6 method (data not shown). Surprisingly, this number of phosphoPSMs from SEQUEST and X!Tandem was lower as compared to separate searches of CID spectra and ETD spectra, while the number of phosphoPSMs from Mascot was quite close to that from CID alone. The number of unique phosphopeptides identified by all three algorithms was also lower when the spectra were merged prior to search (Figure 2B, 2D and 2F).

This overall lack of improved performance in phosphopeptide identification by the search algorithms might be attributed to two main reasons: First, because the b/y ion series cannot be distinguished from the c/z ion series after merging, any fragment ion has a higher likelihood of being assigned incorrectly. Second, the higher number of non-c/z fragments (e.g. charge reduced precursors, neutral losses, unexplained peaks) produced by ETD again increases the chances of incorrect assignments.

Although we did not observe any significant increase in phosphopeptide identification after merging CID and ETD spectra, it is still possible that the confidence of identification is increased because of the increased information content of the merged spectra and the orthogonal nature of CID and ETD. One way to assess the confidence of peptide identification is by examining the normalized delta score [7, 12], where the difference in scores of the 1st and the 2nd ranked peptides is divided by the score of the 1st rank peptide. A PSM is likely to be a high confidence match when the normalized delta score is close to 1. Conversely, a low confidence of a PSM might result in a normalized delta score close to 0. We compared normalized Mascot delta scores from CID alone, ETD alone and merged spectra in order to see whether the confidence of phosphopeptide identification is increased after merging spectra. As an illustration, CID, ETD and normalized merged spectra from two phosphorylated peptides are shown in Figure 3A and 3B. Note that the intensity of CID spectra before normalization was ~3 times higher in Figure 3A and ~10 times higher in Figure 3B than ETD spectra. As shown in Figure 3C, we observed that the normalized Mascot delta scores from merged spectra were slightly increased compared to CID or ETD alone as marked by arrows even though the number having the normalized Mascot delta score of 1 was decreased. This implies that the confidence levels of CID and ETD spectra are merely being averaged and that the confidence of identified merged spectra is largely not improved, but is only marginally better.

One of the key goals of mass spectrometry-based phosphoproteomics is to obtain comprehensive identification of phosphopeptides in a sample. Using our alternating CID and ETD data, we examined the relative increase in phosphopeptide identification from duplicate runs using CID with that obtained by adding ETD fragmentation to CID and found a comparable increase in both strategies (~19% from replicate CID runs and ~17% by adding ETD to CID). Overall, we conclude that the conventional database search algorithms are not optimized for searching mixed spectra from CID and ETD fragmentation. Development of newer algorithms or modifications of existing ones could allow us to take maximal advantage of the complementary types of ions produced by CID and ETD. Until such time, the coverage of phosphopeptides can be improved by employing a number of complementary methodologies including subcellular fractionation (e.g. nuclear or membrane proteome), enzymes other than trypsin (e.g. Lys-C or Asp-N), fractionation methods (e.g. SCX, RPLC), phosphopeptide enrichment methods (e.g. antibodies, TiO₂ or IMAC) and multiple search algorithms.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the anonymous reviewer for constructive suggestions regarding additional analyses. This study was supported in part by a grant S10RR023025 from the High End Instrumentation Program of the National Institutes of Health, a Department of Defense Era of Hope Scholar award (W81XWH-06-1-0428), an NIH roadmap grant for Technology Centers of Networks and Pathways (U54RR020839) and a contract N01-HV-00240 from the National Heart Lung and Blood Institute.

A list of abbreviations

ETD	electron transfer dissociation
DHB	2,5-dihydroxybenzoic acid
PSM	peptide-spectrum match
phosphoPSM	phosphopeptide-spectrum match

References

1. Syka JE, et al. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc Natl Acad Sci U S A*. 2004; 101(26):9528–9533. [PubMed: 15210983]
2. Zubarev RA, et al. Electron capture dissociation for structural characterization of multiply charged protein cations. *Anal Chem*. 2000; 72(3):563–573. [PubMed: 10695143]
3. Molina H, et al. Global proteomic profiling of phosphopeptides using electron transfer dissociation tandem mass spectrometry. *Proc Natl Acad Sci U S A*. 2007; 104(7):2199–2204. [PubMed: 17287340]
4. Guo T, et al. Hybridization of pulsed-Q dissociation and collision-activated dissociation in linear ion trap mass spectrometer for iTRAQ quantitation. *J Proteome Res*. 2008; 7(11):4831–4840. [PubMed: 18837533]
5. Wu J, et al. Integrating titania enrichment, iTRAQ labeling, and Orbitrap CID-HCD for global identification and quantitative analysis of phosphopeptides. *Proteomics*. 2010; 10(11):2224–2234. [PubMed: 20340162]
6. Catalina MI, et al. Electron transfer dissociation of N-glycopeptides: loss of the entire N-glycosylated asparagine side chain. *Rapid Commun Mass Spectrom*. 2007; 21(6):1053–1061. [PubMed: 17311219]
7. Molina H, et al. Comprehensive comparison of collision induced dissociation and electron transfer dissociation. *Anal Chem*. 2008; 80(13):4825–4835. [PubMed: 18540640]
8. Gronborg M, et al. A mass spectrometry-based proteomic approach for identification of serine/threonine-phosphorylated proteins by enrichment with phospho-specific antibodies: identification of a novel protein, Frigg, as a protein kinase A substrate. *Mol Cell Proteomics*. 2002; 1(7):517–527. [PubMed: 12239280]
9. Kim MS, et al. 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]
10. Larsen MR, et al. Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol Cell Proteomics*. 2005; 4(7):873–886. [PubMed: 15858219]
11. Good DM, et al. Post-acquisition ETD spectral processing for increased peptide identifications. *J Am Soc Mass Spectrom*. 2009; 20(8):1435–1440. [PubMed: 19362853]
12. Savitski MM, et al. Confident phosphorylation site localization using the Mascot Delta Score. *Mol Cell Proteomics*. 2010 In press.

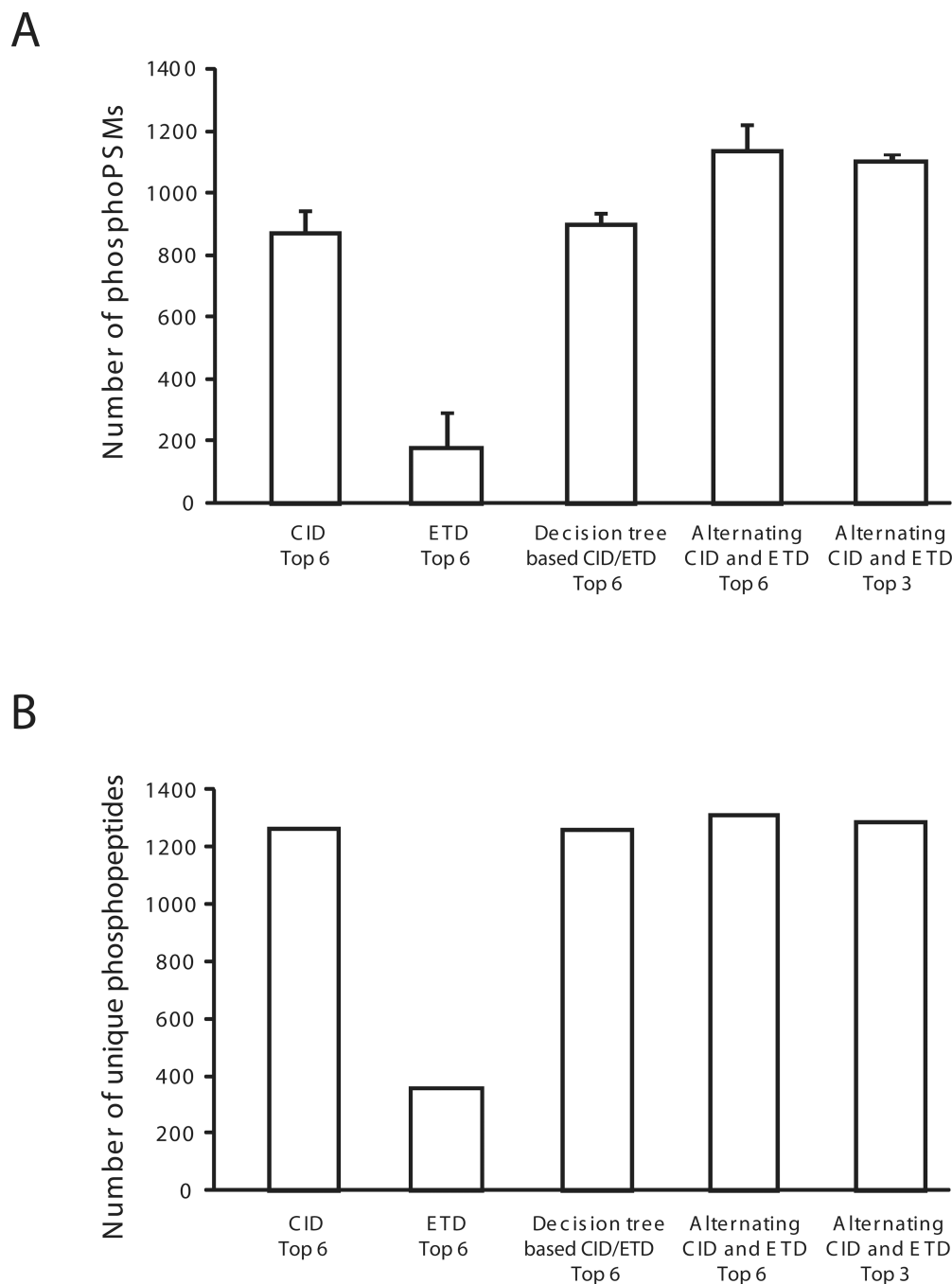


Figure 1. Comparison of five different acquisition conditions using four different fragmentation methods. Phosphopeptide-spectrum matches (phosphoPSMs) were identified by SEQUEST at 1% FDRs. The average number of phosphoPSMs (**A**) and total number of unique phosphopeptides (**B**) from triplicate LC-MS/MS runs are shown. ‘Top 6’ refers to picking of the six most abundant ions for fragmentation, while ‘Top 3’ refers to the three most abundant ions being picked for fragmentation.

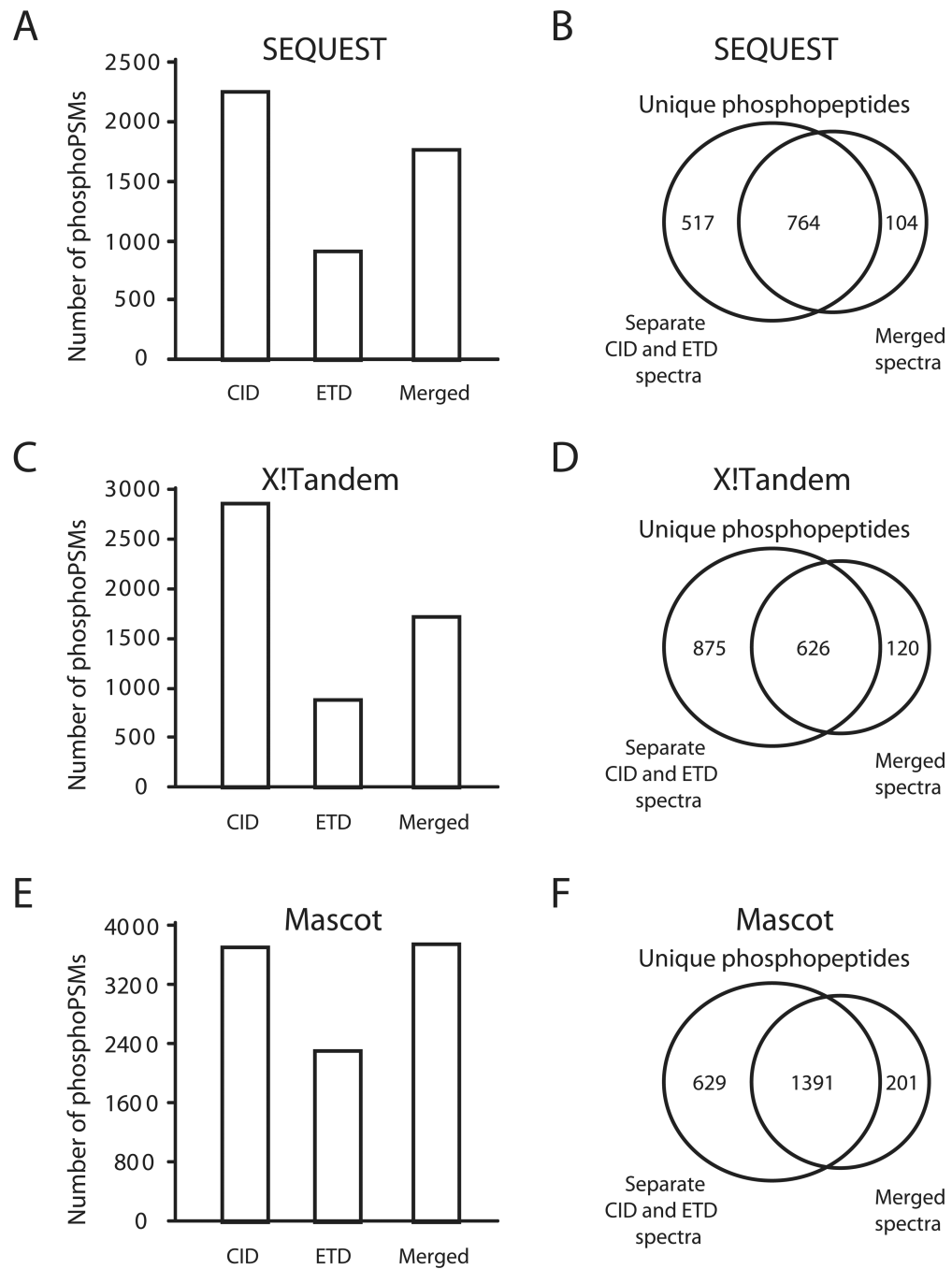


Figure 2. Comparison of merged versus separate searches of data from alternating CID and ETD experiments. CID, ETD and merged spectra generated from alternating CID and ETD method (Top 3) were searched using SEQUEST (**A** and **B**), X!Tandem (**C** and **D**) or Mascot (**E** and **F**) algorithms. Panels **A**, **C** and **E** show the number of phosphoPSMs, while panels **B**, **D** and **F** show the number of unique phosphopeptides and its Venn diagram.

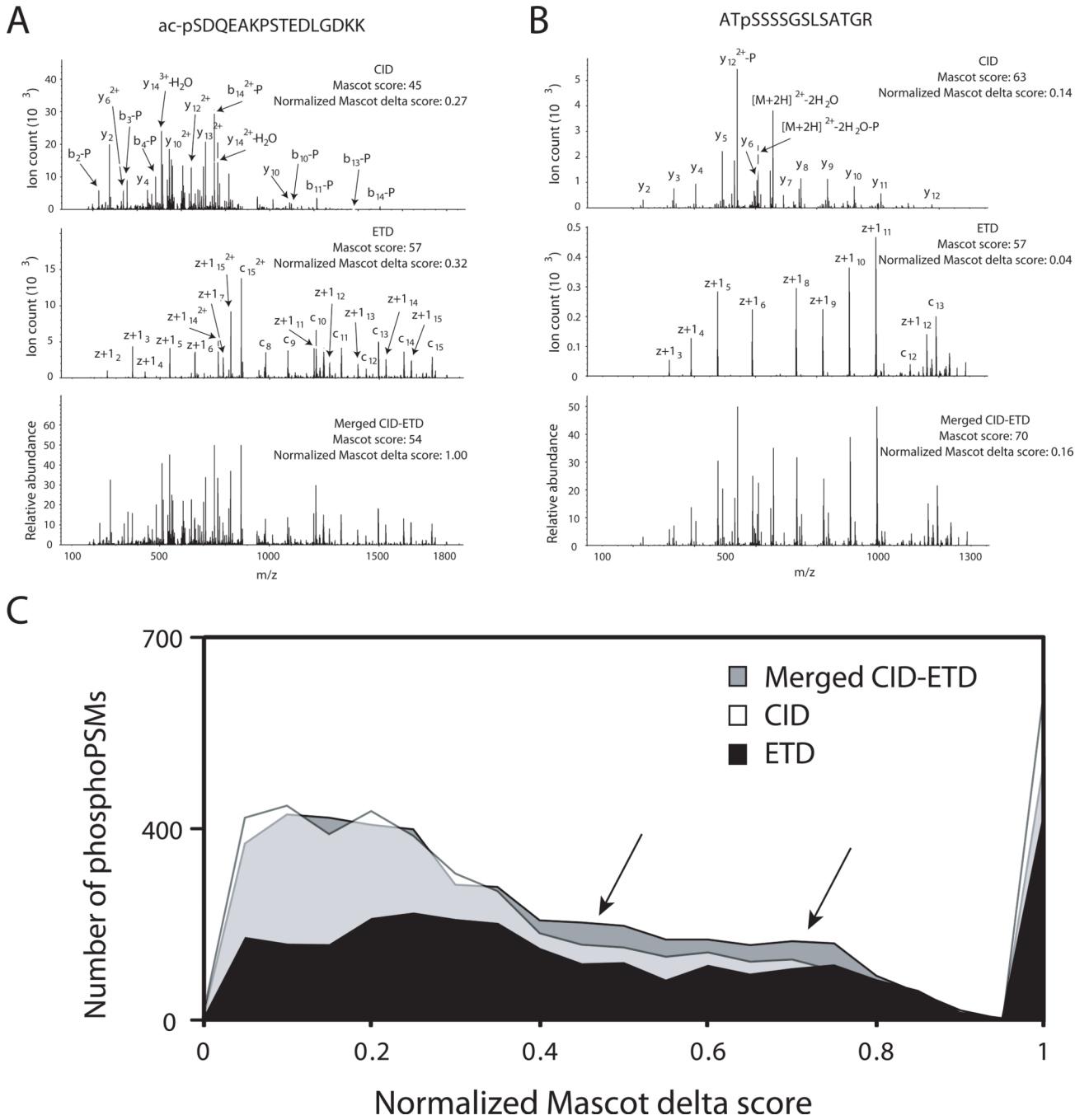


Figure 3. The effect of merging CID and ETD spectra on the confidence level of phosphoPSMs. CID, ETD and their normalized merged spectra from two precursors (**A** and **B**) were shown. The distribution (**C**) of normalized Mascot delta scores from CID alone (white), ETD alone (black) or merged spectra (gray) search were plotted using data acquired from alternating CID and ETD method (top 6 mode).