⌘ *Author's Choice*

Large Scale Localization of Protein Phosphorylation by Use of Electron Capture Dissociation Mass Spectrometry*[S]

Steve M. M. Sweet‡§, Christopher M. Bailey‡§, Debbie L. Cunningham‡§, John K. Heath‡§, and Helen J. Cooper§¶

We used on-line electron capture dissociation (ECD) for the large scale identification and localization of sites of phosphorylation. Each FT-ICR ECD event was paired with a linear ion trap collision-induced dissociation (CID) event, allowing a direct comparison of the relative merits of ECD and CID for phosphopeptide identification and site localization. Linear ion trap CID was shown to be most efficient for phosphopeptide identification, whereas FT-ICR ECD was superior for localization of sites of phosphorylation. The combination of confident CID and ECD identification and confident CID and ECD localization is particularly valuable in cases where a phosphopeptide is identified just once within a phosphoproteomics experiment. *Molecular & Cellular Proteomics 8: 904–912, 2009.*

Phosphorylation is an important protein post-translational modification. Phosphoproteomics experiments have successfully identified thousands of phosphorylation sites, predominantly using CID with relatively low resolution ion trap detection (1–3). For phosphoproteomics data to be of most use to the wider biology community, two key criteria should be met: the phosphopeptide identifications should be correct, and the sites of phosphorylation should be correctly localized. In practice it is not possible to guarantee the accuracy of identifications and site localizations; however, it is possible to include some measure of the confidence in both these results for each phosphopeptide. Database search algorithms give output such as expectation values or similar scores, which can be used to gauge the strength of an identification (4, 5). Recently algorithms have been introduced that give similar confidence scores for phosphorylation site localization (2, 3, 6–8).

Electron capture dissociation (ECD)[1] is a radical-driven fragmentation technique that provides an alternative to slow heating CID fragmentation (9). In contrast to CID fragmentation, labile modifications are usually retained intact upon ECD peptide backbone cleavage. ECD has therefore been applied to the analysis of post-translationally modified proteins almost since its inception (10–13). ECD efficiency has improved to the point where on-line LC-MS/MS experiments are feasible; however, the amount of precursor required for such analyses remains significantly greater than for corresponding CID experiments (14–16). ECD is effectively restricted to FT-ICR mass spectrometers. Although ECD has been demonstrated in ion trap instruments, to date these demonstrations have consisted only of direct infusion of known standards (17, 18). The recently developed technique of electron transfer dissociation (ETD) has allowed similar radical fragmentation to be obtained in ion trap mass spectrometers (19). Large scale comparisons of CID and ETD have been carried out for unmodified peptides (20, 21). The increase in speed and sensitivity, albeit at the expense of resolution, has also allowed ETD to be used in phosphoproteomics analyses (22, 23). These studies demonstrated that ETD can efficiently identify phosphopeptides on a chromatographic time scale. Molina *et al.* (22) compared alternating CID and ETD fragmentation of the same precursors (for an unspecified number of precursors) and found that ETD gave greater peptide sequence coverage. This was assumed to result in improved phosphorylation site localization; however, neither study calculated confidence scores for site localization.

The benefits of ECD fragmentation have been demonstrated for various single phosphoproteins and simple mixtures (24–29); however, no large scale comparison of CID and ECD for phosphoprotein analysis has been carried out. In this context, we present a phosphoproteomics data set obtained using both CID and ECD tandem mass spectrometric techniques. Particular attention was paid to the confident localization of sites of phosphorylation within the identified phosphopeptides.

From the ‡Cancer Research UK Growth Factor Group, §School of Biosciences, College of Life and Environmental Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom

⌘ *Author's Choice*—Final version full access.

Received, September 26, 2008, and in revised form, December 19, 2008

Published, MCP Papers in Press, January 8, 2009, DOI 10.1074/mcp.M800451-MCP200

[1] The abbreviations used are: ECD, electron capture dissociation; DD, data-dependent; CID, collision-induced dissociation; ETD, electron transfer dissociation; LTQ, linear ion trap (Thermo Fisher Scientific); NL-ECD, neutral loss-dependent ECD; OMSSA, open mass spectrometry search algorithm; SCX, strong cation exchange; SLoMo, site localization of modifications; NL, neutral loss.

© 2009 by The American Society for Biochemistry and Molecular Biology, Inc.
This paper is available on line at http://www.mcponline.org

EXPERIMENTAL PROCEDURES

Cell Culture—Mouse fibroblast NIH 3T3 cells were cultured at 37 °C, 5% CO_2 in Dulbecco's modified Eagle's medium (Invitrogen) supplemented with 2 mM L-glutamine (Invitrogen), 0.1 mg/ml streptomycin, 0.2 units/ml penicillin (Sigma), and 10% (v/v) donor bovine serum (Invitrogen). Following serum starvation in medium containing 0.1% serum for 18 h, cells were treated with 2 mM sodium pervanadate for 20 min prior to lysis. Cells were trypsinized and pelleted by centrifugation at 3300 rpm using a benchtop centrifuge for 5 min at 4 °C. The pellet was washed with PBS prior to addition of 1 ml of ice-cold lysis buffer (17 mM HEPES, pH 8, 7.65 M urea, 1 mM Na_3VO_4, 50 mM NaF, 25 mM β-glycerophosphate, and one tablet of Complete mini protease inhibitor mixture (Roche Diagnostics) for every 10 ml of buffer). The cells were lysed by sonication on ice. Lysates were subsequently cleared by centrifugation at 14,000 $\times$ *g* for 10 min at 4 °C. Total protein concentrations of the cleared lysates were then determined by Coomassie (Bradford) Protein Assay kit (Pierce) according to the manufacturer's instructions.

Protein Digestion—The lysates were reduced (8 mM DTT) and alkylated (20 mM iodoacetamide) in 50 mM ammonium bicarbonate. The lysates were diluted to 4 M urea, acetonitrile (10% by volume) and endoproteinase Lys-C were added (Sigma; 1:400 enzyme:protein), and digestion was allowed to proceed at 37 °C for 5 h. The lysates were then further diluted to 1 M urea and trypsin (Trypsin Gold; Promega, Madison, WI) was added (1:100 enzyme:protein) prior to overnight digestion at 37 °C.

Digested lysates were acidified by addition of trifluoroacetic acid (0.5% final volume), and acetonitrile was removed by vacuum centrifugation. Peptides were desalted (C_8 cartridge; Michrom) and dried by vacuum centrifugation.

Strong Cation Exchange (SCX) Chromatography—Desalted peptide from 1 mg of lysate (as measured prior to digestion) was resuspended in 100 μl of mobile phase A and loaded onto a 100 $\times$ 2.1-mm polysulfoethyl aspartamide column (5-μm particle size, 20-nm pore size; PolyLC) at a flow rate of 200 μl/min. Separation used a gradient elution profile that started with 100% mobile phase A (5 mM KH_3PO_4, 25% acetonitrile, pH 3), increased from 0 to 30% mobile phase B (5 mM KH_3PO_4, 25% acetonitrile, 250 mM KCl, pH 3) over 30 min, increased to 50% B over 5 min, and then returned to 100% A. Fractions (750 μl) were collected throughout the run.

Phosphopeptide Enrichment—Phosphopeptides were enriched from desalted SCX fractions approximately as described previously (30, 31). TiO_2 Titansphere beads (5-μm diameter) were obtained from GL Sciences. Peptides were loaded onto TiO_2 microcolumns in 2% TFA. Columns were washed with saturated phthalic acid, 80% MeCN, 2% TFA and then with the same buffer omitting phthalic acid. Peptides were eluted in a two-step procedure with 50 mM Na_2HPO_4 followed by dilute NH_4OH solution. Eluates were desalted using C_{18} ZipTips according to the manufacturer's instructions (Millipore). The resulting peptide mixtures were analyzed by LC-MS/MS.

LC-MS/MS—On-line liquid chromatography was performed by use of a Micro AS autosampler and Surveyor MS pump (Thermo Fisher Scientific, Bremen, Germany). Peptides were loaded onto a 75-μm (internal diameter) Integrafrit (New Objective) C_8 resolving column (length, 10 cm) and separated over a 40-min gradient from 0 to 40% acetonitrile (J. T. Baker Inc.). Peptides eluted directly ($\sim$350 nl/min) via a Triversa nanospray source (Advion Biosciences) into a 7-Tesla LTQ FT mass spectrometer (Thermo Fisher Scientific) where they were subjected either to data-dependent CID and ECD or neutral loss (NL)-dependent ECD.

Data-dependent CID and ECD (DD-CID-ECD)—The mass spectrometer alternated between a full FT-MS scan (*m/z* 400–1600) and subsequent CID and ECD MS/MS scans of the most abundant ion above a threshold of 40,000. Survey scans were acquired in the ICR cell with a resolution of 100,000 at *m/z* 400. Precursor ions, as identified by a lower resolution preview of the full FT-MS scan, were isolated and subjected to CID in the linear ion trap in parallel with the completion of the full FT-MS scan. The width of the precursor isolation window was 6 *m/z*. Only multiply charged precursor ions were selected for MS/MS. CID was performed with helium gas at a normalized collision energy of 35%. Automated gain control was used to accumulate sufficient precursor ions (target value, 5×10^4; maximum fill time, 0.2 s). Precursor ions were activated for 30 ms. For the ECD event precursor ions were isolated in the ion trap and transferred to the ICR cell. Isolation width was 6 *m/z*. Automated gain control was used to accumulate sufficient precursor ions (target value, 1×10^6/microscan; maximum fill time, 1 s). The electrons for ECD were produced by an indirectly heated barium tungsten cylindrical dispenser cathode (5.1-mm diameter, 154 mm from the cell, 1 mm off-axis). The current across the electrode was $\sim$1.1 A. Ions were irradiated for 60 ms at 5% energy (corresponding to a cathode potential of -2.775 V). Each ECD scan comprised four co-added microscans acquired with a resolution of 25,000 at *m/z* 400. Dynamic exclusion was used with a repeat count of 1 and an exclusion duration of 60 s. Data acquisition was controlled by Xcalibur 2.0 and Tune 2.2 software (Thermo Fisher Scientific).

Neutral Loss-dependent ECD (NL-ECD)—The mass spectrometer alternated between a full FT-MS scan (*m/z* 400–1600) and subsequent CID MS/MS scan of the most abundant ion. If a neutral loss of 98 Da (49, 32.67, 24.5 *m/z*) or 98 + 18 Da from a 4+ ion (29 *m/z*) from the precursor ion was observed in the five most abundant ions in the CID mass spectrum, an ECD scan of the precursor ion was also acquired. All other settings were as above.

Data Analysis—DTA files were created from the raw data using Bioworks 3.3.1 (Thermo Fisher Scientific) (parameters: no scan grouping; minimum ion threshold of 15; absolute intensity threshold of 50). The OMSSA Browser 2.1.1. was used to search the DTA files against a concatenated database consisting of the mouse International Protein Index (IPI) database (Version 3.40) supplemented with common contaminants (including keratins, trypsin, and BSA) and the reversed sequence version of the same database. The final database contained 107,688 protein entries (53,844 of which were reversed sequence versions). CID and ECD DTAs were searched separately, resulting in two database searches per experiment.

OMSSA settings for the CID search were as follows: enzyme, trypsin; peptide *m/z* tolerance, $\pm$0.02; MS/MS *m/z* tolerance, $\pm$0.8; miscleavages allowed, 2; maximum number of variable modifications per peptide, 32; fixed modification, carbamidomethyl-Cys; variable modifications, acetylation of protein N terminus, oxidation of Met, phosphorylation of Ser, Thr, or Tyr; product ion types to search, b and y; precursor search type, monoisotopic; product search type, monoisotopic; E-value cutoff, 50; at least one peak must match among the top 20 most intense peaks; lower bound of precursor charge, 2; upper bound of precursor charge, 6; charge at which to start considering multiply charged products, 3; maximum charge state allowed for product ions, 2; allow N-terminal Met cleavage, yes; allow correlation correction to probability score, no; threshold above which the mass of a neutron will be added in the exact mass search in Da, 5000; allow elimination of charge-reduced precursors in spectrum, no; precursor charge state detection, read from input file data; correction for precursor charge dependence at *m/z* tolerance, linear correction; allow first N-terminal product ion, *e.g.* b_1, to be searched, no; allow first C-terminal product ion, *e.g.* y_1, to be searched, yes. OMSSA settings for the ECD search were as above with the following exceptions: peptide *m/z* tolerance, $\pm$1.1 (see supplemental text and supplemental Fig. 1 for peptide tolerance selection rationale); MS/MS *m/z* tolerance, $\pm$0.02; product ion types to search: c, y, and z; allow

TABLE I
Summary of phosphopeptides identified from both DD-CID-ECD and NL-ECD experiments

Identified peptides	Phosphorylated	Non-phosphorylated	Percent phosphorylated
DD-CID-ECD experiment			
CID	773	1113	41.0
ECD	551	676	44.9
NL-ECD experiment			
CID	1319	2066	39.0
ECD	536	69	88.6
Total	3179	3924	
Total forward hits	7103		
Total reverse hits	40		
Distinct phosphopeptides	906		

correlation correction to probability score, yes; probability of consecutive ions for correlation correction, 0.25; allow elimination of charge-reduced precursors in spectrum, yes; correction for precursor charge dependence at *m/z* tolerance, no correction; allow first N-terminal product ion, *e.g.* b_1, to be searched, yes.

OMSSA results were filtered to allow only the top scoring identification (sequence and site of modification) per DTA. The results were then filtered by precursor mass error (in ppm) and E-value to obtain a false discovery rate (FDR) for phosphopeptides lower than 1% for each search (FDR = reverse hits/forward hits × 100) (3). The precursor mass error cutoff values for each experiment were as follows: DD-CID-ECD >1.42 and <18.81 ppm; NL-ECD >−2.52 and <16.93 ppm. The OMSSA E-value cutoffs for CID database searches were 1 and 5 (DD-CID-ECD and NL-ECD experiment, respectively). The OMSSA E-value cutoffs for ECD database searches were 0.4 and 0.068 (DD-CID-ECD and NL-ECD experiment, respectively).

Phosphorylation site localization from CID mass spectra was assessed using the SLoMo algorithm described in detail in Bailey *et al.* (7). SLoMo is based on the AScore algorithm (3) but allows both CID and ECD/ETD mass spectra to be localized. Specifically SLoMo recognizes c, $z^{\bullet}$,and z′ ions in addition to b and y ions. The fragment mass accuracy is specified in ppm, allowing localization from both high resolution ECD mass spectra and lower resolution ion trap CID mass spectra. The precursor ion and reduced precursor ion peaks are removed from ECD mass spectra. The SLoMo settings used for CID data were as follows: H_3PO_4 and H_2O neutral loss peaks were removed; singly and doubly charged fragment ions were considered; fragment ion types were b and y; fragment mass tolerance was 350 ppm. The settings used for ECD data were as follows: precursor and reduced precursor peaks were removed; singly and doubly charged fragment ions were considered; fragment ion types were c, y, $z^{\bullet}$, and z′; fragment mass tolerance was 12 ppm. A SLoMo score of >19 constitutes confident localization.

For phosphopeptide backbone bond cleavage efficiency calculation, data are from 1114 (ECD) and 2262 (CID) sites of phosphorylation. Phosphopeptides with a SLoMo score of 0 are not included. Fragment ions indicating bond cleavage were identified at a peak depth of 10 peaks per 100 *m/z* window with detection parameters as for SLoMo (b and y for CID; c, y, $z^{\bullet}$, and z′ for ECD). Cleavages N-terminal to proline are not included for ECD data. To assess random matches due to background noise, data were searched for unexpected ions (z for CID and b for ECD). The values shown are corrected for random matches (CID cleavage from b and y ions, therefore two potential random matches per bond cleavage (% cleaved − 2 × z ions); ECD cleavage from c, y, $z^{\bullet}$, and z′, therefore four potential random matches per bond cleavage (% cleaved − 4 × b ions)).

RESULTS

To assess the utility of ECD in phosphopeptide identification and site localization, a complex mixture, enriched for phosphopeptides, was analyzed by LTQ CID and FT-ICR ECD. Mouse whole cell lysate was fractionated by SCX chromatography with subsequent TiO_2-based phosphopeptide enrichment. Equal amounts of the resulting mixture were analyzed either by DD-CID-ECD or by data-dependent CID with NL-ECD. Each analysis required ~12 h of instrument time (11 SCX fractions; 65-min LC MS/MS). CID and ECD mass spectra acquired in each analysis were searched independently, resulting in a total of four database searches. The use of a concatenated forward-reverse database allowed the false discovery rate to be estimated and controlled at <1% for each search after filtering according to precursor *m/z* error and OMSSA database search algorithm E-value. The results in terms of peptides identified are presented in Table I. A complete list of the individual phosphopeptides identified with references to individual MS/MS mass spectra is given in supplemental Table 1 together with the results of the SLoMo phosphorylation localization analysis (see below). Individual MS/MS mass spectra with fragment annotations are in supplemental Fig. 2.

Neutral Loss-dependent ECD—The results from the DD-CID-ECD experiment (Table I) indicate that phosphopeptides constituted 40–45% of the starting mixture. Neutral loss-dependent ECD ensured that ECD events were predominantly restricted to phosphorylated species (89% of identifications), thereby increasing the time available for CID events. This resulted in an increased number of identifications in the NL-ECD experiment, both phosphorylated (1855 *versus* 1324 identifications) and non-phosphorylated. It should be noted that not every phosphopeptide identified by CID triggered an ECD event. Specifically 441 of the (redundant) phosphopeptides identified by CID in the NL-ECD experiment did not trigger an ECD event. 200 of these CID mass spectra were examined manually to determine why an ECD event was not triggered. The reasons, in order of significance, were as follows: combined neutral loss of

phosphate and water (36%), no dominant neutral loss (25%), tyrosine phosphorylation (21%), and neutral loss peak outside the five most intense peaks (but within the top 10; 17%).

From here on, to directly compare CID and ECD performance, we will concentrate on CID/ECD paired events. 6080 CID/ECD pairs were acquired, 1892 of which resulted in phosphopeptide identification. Phosphopeptide identifications from paired events gave 796 distinct phosphopeptide identifications, *i.e.* 88% of the overall total (906 distinct phosphopeptides). The difference between the larger (906) and smaller (796) number of distinct phosphopeptides corresponds to 110 distinct phosphopeptides that were identified only from CID mass spectra where no paired ECD mass spectrum was acquired. Distinct phosphopeptides are defined as those of unique peptide sequence and phosphorylation site and number, *i.e.* methionine oxidation variants are not counted as distinct. Phosphopeptides differing only in the site of phosphorylation are considered as distinct if both sites of phosphorylation are confidently localized ($p \leq 0.0126$).

FIG. 1. **Overlap between DTAs leading to identifications by ECD and CID.** All 4763 identifications (*IDs*) are from paired CID/ECD events (43).

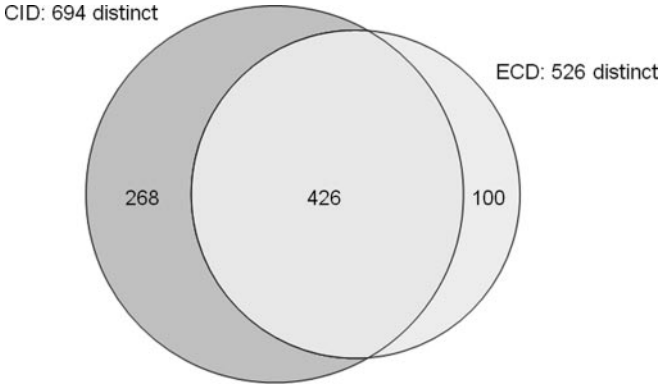

FIG. 2. **Overlap between distinct phosphopeptides identified by ECD and CID.** All 1220 identifications are from paired CID/ECD events.

Paired CID/ECD Events Allow Internal Validation of Identifications—3208 paired CID/ECD events led to peptide identification (from one or both events) for a total of 4763 separate identifications. In just under 50% of these pairs (1555 pairs) *both* CID and ECD mass spectra gave identifications (Fig. 1). If both CID and ECD identifications are correct, we expect agreement within each pair. In fact five pairs of 1555 do not agree (supplemental Table 2). This is well within the expected maximum number of conflicts predicted by the decoy search (19 incorrect identifications).

Three of the five conflicts are between homologous, isobaric peptides with either an amino acid order inversion or an isoleucine/leucine substitution. This observation illustrates the previously identified difficulty in confidently identifying peptides when homologous sequences exist (32). Manual analysis of these conflicts showed that in one case (conflict 1, supplemental Table 2) the CID and ECD mass spectra agreed, but the OMSSA algorithm misscored the ECD mass spectrum, *i.e.* assigned the higher score to the peptide with fewer fragments matched (supplemental Fig. 3). In the case of conflict 2, isoleucine and leucine could not be distinguished as the relevant w/d ions were not observed (33). In conflict 3, the

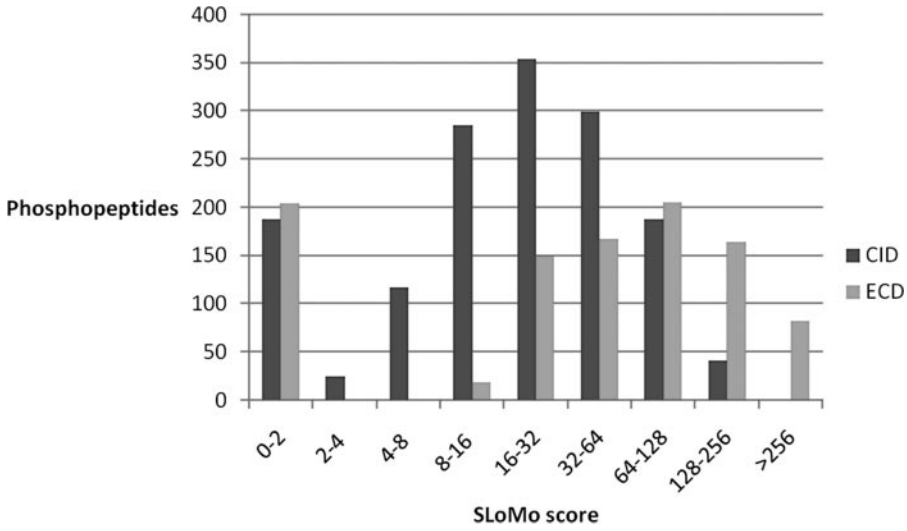

FIG. 3. **Binned distribution of SLoMo scores.** The *x axis* has log scale. All identifications are from paired CID/ECD events. Identifications with only one possible site of localization are not included.

CID mass spectrum showed evidence for the presence of both identified forms (supplemental Fig. 4). Manual analysis of the remaining conflicts indicated that in conflict 4 the ECD mass spectrum could better be explained by the sequence identified from the CID search, *i.e.* the OMSSA ECD identification was incorrect (supplemental Fig. 5**)**. The reason for the misidentification is that OMSSA is currently unable to consider z′ and c• ions (resulting from hydrogen transfer between classical ECD fragments) (34). In the final case, neither CID nor ECD identification appeared convincing (supplemental Fig. 6).

Phosphopeptide Identification—Fig. 2 shows the overlap between CID and ECD distinct phosphopeptide identifications. (See also supplemental Table 1.) The data indicate that CID was more efficient for phosphopeptide identification than ECD; however, a significant number (13%) of identifications were by ECD only. The majority of phosphopeptides were identified by both CID and ECD (53%) with obvious benefits for reducing false positives. CID identifications make up the final third of distinct phosphopeptides identified (34%). It should be noted that the comparison being made is between a CID event in the linear ion trap and an ECD event in the FT-ICR cell.

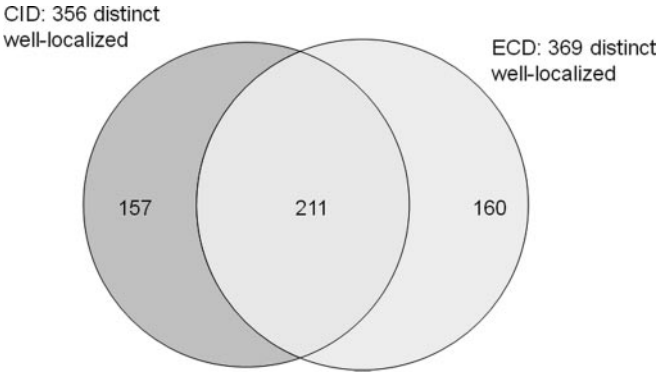

FIG. 4. **Overlap between distinct, well localized (SLoMo >19) phosphopeptides identified by ECD and CID.** All 725 identifications are from paired CID/ECD events. Identifications with only one possible localization are not included.

Site Localization of Phosphorylation—Phosphorylation site localization was determined using the SLoMo algorithm (7). The SLoMo algorithm assigns a localization confidence score to each phosphopeptide. This score is analogous to the previously described AScore (3). Scores >19 indicate confident localizations ($p \leq 0.0126$; SLoMo Score $= -10 \times \log(p)$). Fig. 3 shows the distribution of SLoMo scores for CID and ECD phosphopeptide data. (See also supplemental Table 1.) In both cases the distributions appear to be bimodal with a fraction of identifications having scores of 0, *i.e.* no site-determining ions that distinguish between the top two possibilities were detected. This division is particularly clear for the ECD data where the higher mass accuracy of site-determining ions (12 ppm compared with 350 ppm) and the lower level of background noise ensures that in most cases if a site-determining ion is detected the phosphorylation site will be confidently localized. The higher confidence localizations from ECD data are reflected in the relative contributions of CID and ECD data to the site localization of distinct phosphopeptides (Fig. 4). The higher rate of phosphopeptide backbone bond cleavage by ECD compared with CID is shown in supplemental Fig. 7. Despite contributing fewer phosphopeptide identifications (Fig. 2), the ECD data contribute more confident localizations than the CID data.

As discussed above for phosphopeptide identification, when paired CID and ECD mass spectra both give confident site localization we expect the two localizations to agree. A total of 340 pairs of mass spectra (non-distinct phosphopeptides) fall into this category. Of these pairs, 336 agree (99% of localizations), one pair disagrees in identification (one of the five cases listed in supplemental Table 2), and three pairs disagree on the site of phosphorylation (Table II). Although this is within the expected error rate for a SLoMo score of 19, it seems rather high given that many of the SLoMo scores are significantly greater than 19. The three conflicts shown in Table II were analyzed manually. This revealed evidence in two cases for co-elution of both phosphoisoforms, *i.e.* both localizations were correct. An example of co-elution of isobaric phosphopeptides is shown in Fig. 5. Co-elution of isobaric

TABLE II
CID/ECD localization conflicts

In total, 340 pairs gave confident localization (SLoMo > 19) from both CID and ECD data. A SLoMo score of 19 corresponds to a 1.26% error rate, which would predict nine errors from the 680 identifications. In fact we saw three localization conflicts and one identification (ID) conflict. Of the localization conflicts, 1 and 2 are cases where both isoforms are present, *i.e.* both are correct. pT and pS indicate phosphothreonine and phosphoserine, respectively. g indicates *N*-acetylation.

Conflict	Peptide sequence	SLoMo score	Charge	Manual analysis
1	TVpTPASSAKTpSPAKQQAPPVR	24.3 CID	3	Both correct: co-eluting isoforms
	TVTPApSSAKTpSPAKQQAPPVR	21.5 ECD	3	Both correct: co-eluting isoforms
2	TVpTPASSAKTpSPAKQQAPPVR	41.2 CID	3	Both correct: co-eluting isoforms
	TVTPASpSAKTpSPAKQQAPPVR	22.3 ECD	3	Both correct: co-eluting isoforms
3	AKPAAQSEEETATpSPAASPTPQSAER	22.3 CID	3	CID correct
	AKPAAQSEEETApTSPAASPTPQSAER	27.0 ECD	3	ECD incorrect
4	RNpSLTGEEGELVK	36.4 CID	2	ID conflict
	gLSRWWWRpSR	89.2 ECD	2	ID conflict

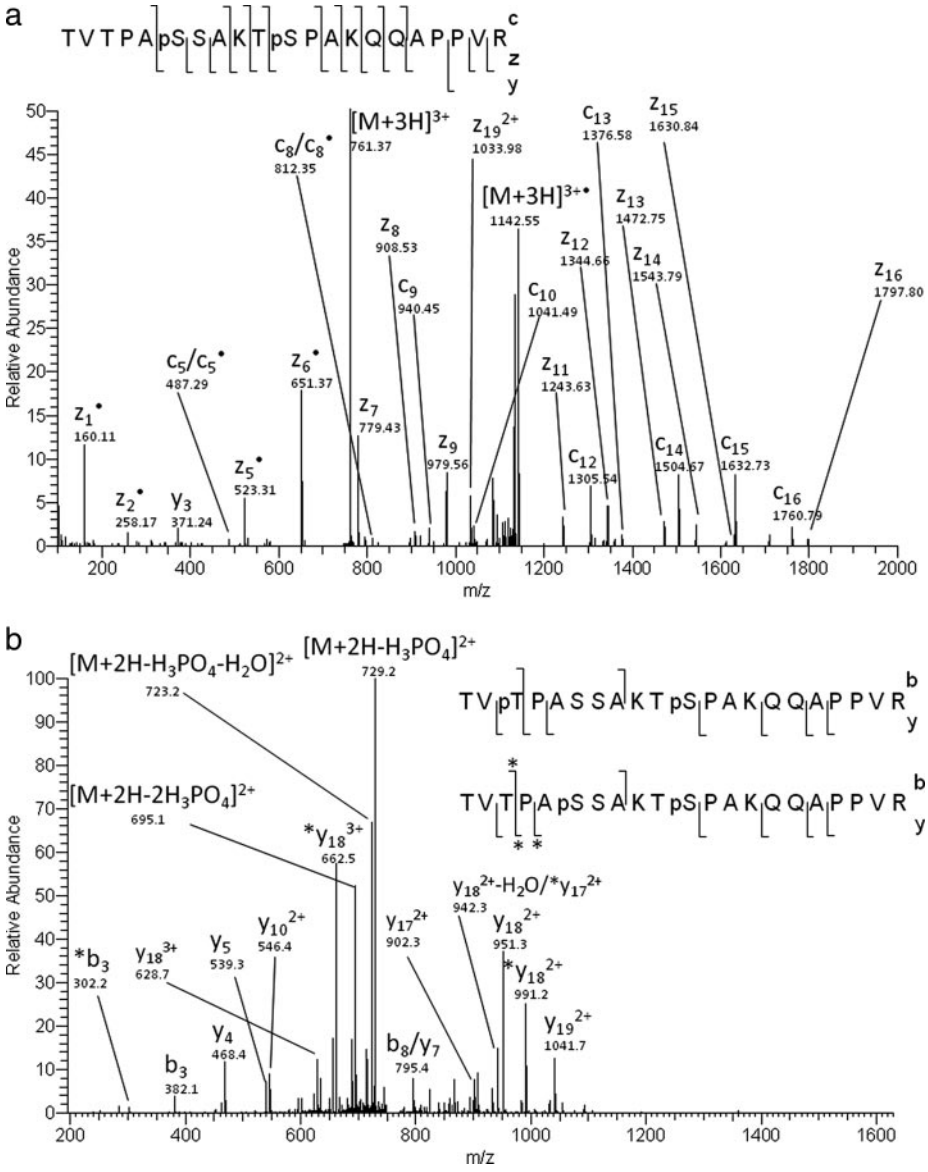

FIG. 5. **Example of isomeric phosphopeptide co-elution.** *a*, ECD mass spectrum showing first and third serine phosphorylation (SLoMo score, 21.5). *b*, CID mass spectrum showing evidence for threonine phosphorylation (SLoMo score, 24.3). The site of phosphorylation indicated by SLoMo is shown *uppermost inset* in *b*.

phosphopeptides will generally suppress the SLoMo localization scores. This is demonstrated by the example shown in supplemental Fig. 8 where the SLoMo score for the CID mass spectrum is only 11.7. In this example, both ECD and CID mass spectra contain clear evidence of two distinct phosphoisoforms.

Multiply Phosphorylated Peptides—ECD identified fewer multiply phosphorylated peptides than CID (Fig. 6). However, more than 60% of the multiply phosphorylated peptides identified by ECD were well localized as opposed to only 34% of the CID identifications. The localization score applies to the entire phosphopeptide rather than individual sites of phosphorylation; therefore a score >19 indicates that all of the sites of phosphorylation are well localized.

High Confidence Identification and Site Localization from Single Phosphopeptides—A total of 844 pairs of CID and ECD mass spectra both identified the same phosphopeptide. That corresponds to 434 distinct phosphopeptides or over half of the phosphopeptides identified from paired events. Given that the same sequence was identified from orthogonal CID and ECD mass spectra in separate database searches, additional confidence can be placed on these identifications. This feature is of particular significance when the phosphopeptide in question is only sequenced once from the multiple SCX fractions and NL-ECD and DD-CID-ECD experiments. 96 of the 434 phosphopeptides, nearly one-quarter, fall into this category.

Within the 844 agreeing pairs, the sites of phosphorylation from 409 pairs (including 73 pairs with only one possible localization) were unambiguously localized from both CID and ECD mass spectra. This corresponds to 235 distinct phosphopeptides or ~30% of the distinct phosphopeptides identified from paired events.

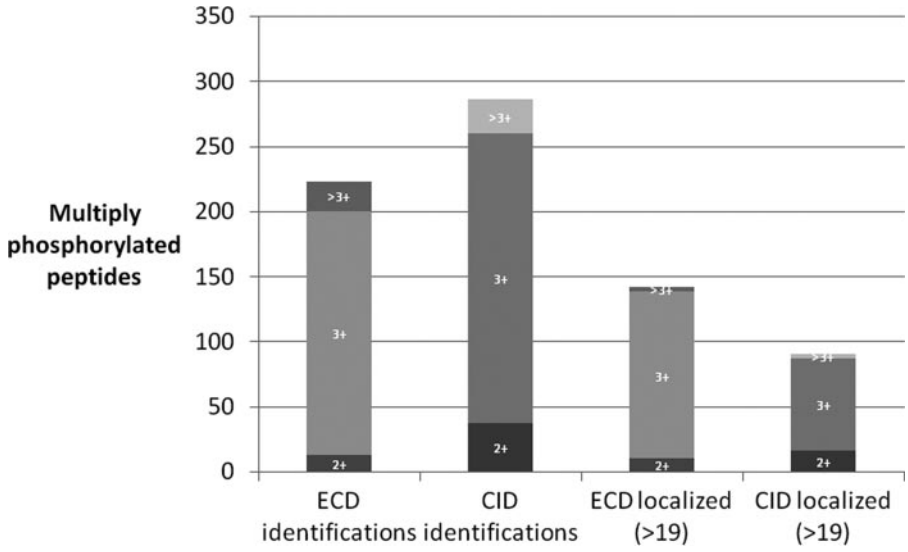

FIG. 6. **Multiply phosphorylated peptides identified and localized from ECD and CID mass spectra and sorted by charge state.** All identifications are from paired events. Identifications with only one possible localization are not included.

DISCUSSION

In this study we analyzed 6080 paired CID/ECD mass spectra, 1892 of which led to phosphopeptide identifications by CID, ECD, or both. This data set allowed a comparison of the relative merits of LTQ CID and FT-ICR ECD mass spectra for phosphopeptide identification and site localization. CID was shown to be more efficient in this experiment for phosphopeptide identification (Fig. 1). This is likely to be related to the higher sensitivity of CID in the ion trap over ECD in the FT-ICR cell. The overall target number of charges for each ECD mass spectrum was 4×10^6 as opposed to just 5×10^4 for CID. To accumulate these large numbers of ions for ECD in a reasonable time (4 s), the threshold to select a peak for fragmentation was set at 4×10^4, which is significantly higher than the 1000–3000 threshold routinely used in CID-only experiments (3). It therefore follows that all CID mass spectra will achieve the target number of ions, and the majority will be of high quality. This is indicated by the overall success rate of CID identifications: 51% (5271 identifications from 10,285 mass spectra). This success rate is significantly higher than that from the average proteomics analysis. Additional factors that may impact the efficiency of CID *versus* ECD for phosphopeptide identification are the choice of trypsin as proteolytic enzyme and the database search algorithm used to assign identifications. Trypsin tends to produce abundant doubly charged peptides, which have been shown to be suboptimal for ETD fragmentation (21). All the major search engines (Mascot, Sequest, OMSSA, X!Tandem, etc.) were designed for CID data usually acquired using either ion traps or time-of-flight mass spectrometers. An additional complication of ECD/ETD fragmentation, highlighted by conflict 4 in Table II (supplemental Fig. 5), is the potential for hydrogen rearrangement, resulting in z′ and c• ions (differing in mass by 1 Dalton from the conventional c and z fragments) (34). OMSSA does not consider these fragments in the database search.

Although ECD was less efficient at phosphopeptide identification, ECD mass spectra gave more confident site localization than CID mass spectra. Several factors are likely to contribute to this: the phosphoamino acid remains intact upon ECD backbone fragmentation; ECD tends to cleave a greater proportion of backbone bonds (15); ECD fragment intensities are more uniform than those of CID (*e.g.* CID gives dominant cleavages N-terminal to proline) (35); ECD mass spectra are acquired at higher mass accuracy, and therefore fragment identification is less likely to be ambiguous. The advantage of ECD over CID for site localization is particularly clear for multiply phosphorylated peptides. Multiply phosphorylated peptides present a greater analytical challenge than singly phosphorylated peptides due to both decreased ionization in the positive mode (when the number of phosphorylations exceeds the number of basic residues) and multiple neutral losses upon CID fragmentation (36, 37). Using the NetPhos phosphorylation prediction algorithm (38), we calculated that approximately one-fifth of tryptic peptides in the human proteome are multiply phosphorylated (supplemental text). This illustrates the limitations of using analytical techniques that are optimized for the analysis of singly phosphorylated peptides.

We also demonstrated that NL-ECD effectively restricts ECD events to phosphopeptides and therefore allows a greater number of identifications overall. The proportion of phosphopeptides selected for ECD could be increased by considering the combined neutral loss of phosphate and water or by extending the NL peak depth beyond the five most intense peaks, a finding of relevance to phosphoproteomics studies relying on NL-dependent MS^3 CID. However, we note that recent results suggest that MS^3 CID mass spectra may not be suitable for characterizing phosphorylation site localization (39).

Paired spectra where both CID and ECD give identifications (and confident localization) can be used to give additional

confidence in the identification (and localization). This complementarity provides greater confidence than repeated CID identifications of the same phosphopeptide (*e.g.* by multiple injections of the same sample) and reduces the need to look for alternative missed cleavage versions of phosphopeptides and use multiple enzymes to confirm "one-hit wonder" identifications (40). The strategy of analyzing complementary CID and ECD data separately can be contrasted with that used by Zubarev and co-workers (41, 42) where CID and ECD data are merged prior to database searching. Merging the data increases both the information available for peptide identification and the score of the resulting identification; however, the increased confidence given by two independent identifications is lost.

In conclusion, our results indicate that combined ECD and CID analysis results in high confidence phosphopeptide identifications and phosphorylation site localization. Hybrid mass spectrometers, such as the LTQ-FT used in this work, are primarily used to measure precursor ions with high mass accuracy while carrying out rapid CID (at lower resolution). We showed that there are potential advantages to using both parts of the hybrid instrument during peptide fragmentation to acquire orthogonal MS/MS data.

* This work was supported by European Union FP6 Endotrack (to S. M. M. S.), Cancer Research UK (to C. M. B., D. L. C., and J. K. H.), and Wellcome Trust Grant 074131 (to H. J. C.).

S The on-line version of this article (available at http://www.mcponline.org) contains supplemental material.

¶ To whom correspondence should be addressed. Tel.: 44-121-414-7527; Fax: 44-121-414-5925; E-mail: H.J.Cooper@bham.ac.uk.

REFERENCES

1. Ficarro, S. B., McCleland, M. L., Stukenberg, P. T., Burke, D. J., Ross, M. M., Shabanowitz, J., Hunt, D. F., and White, F. M. (2002) Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae. *Nat. Biotechnol.* **20,** 301–305
2. Olsen, J. V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. (2006) Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* **127,** 635–648
3. Beausoleil, S. A., Villen, J., Gerber, S. A., Rush, J., and Gygi, S. P. (2006) A probability-based approach for high-throughput protein phosphorylation analysis and site localization. *Nat. Biotechnol.* **24,** 1285–1292
4. Geer, L. Y., Markey, S. P., Kowalak, J. A., Wagner, L., Xu, M., Maynard, D. M., Yang, X., Shi, W., and Bryant, S. H. (2004) Open mass spectrometry search algorithm. *J. Proteome Res.* **3,** 958–964
5. Perkins, D. N., Pappin, D. J. C., Creasy, D. M., and Cottrell, J. S. (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20,** 3551–3567
6. Ruttenberg, B. E., Pisitkun, T., Knepper, M. A., and Hoffert, J. D. (2008) PhosphoScore: an open-source phosphorylation site assignment tool for MSn data. *J. Proteome Res.* **7,** 3054–3059
7. Bailey, C. M., Sweet, S. M. M., Cunningham, D. L., Zeller, M., Heath, J. K., and Cooper, H. J. (2009) SLoMo: automated site localisation of modifications from ETD/ECD mass spectra. *J. Proteome Res.* **8,** 1965–1971
8. Wan, Y., Cripps, D., Thomas, S., Campbell, P., Ambulos, N., Chen, T., and Yang, A. (2008) PhosphoScan: a probability-based method for phosphorylation site prediction using MS2/MS3 pair information. *J. Proteome Res.* **7,** 2803–2811
9. Zubarev, R. A., Kelleher, N. L., and McLafferty, F. W. (1998) Electron capture dissociation of multiply charged protein cations. A nonergodic process. *J. Am. Chem. Soc.* **120,** 3265–3266
10. Kelleher, R. L., Zubarev, R. A., Bush, K., Furie, B., Furie, B. C., McLafferty, F. W., and Walsh, C. T. (1999) Localization of labile posttranslational modifications by electron capture dissociation: the case of γ-carboxyglutamic acid. *Anal. Chem.* **71,** 4250–4253
11. Mirgorodskaya, E., Roepstorff, P., and Zubarev, R. A. (1999) Localization of O-glycosylation sites in peptides by electron capture dissociation in a Fourier transform mass spectrometer. *Anal. Chem.* **71,** 4431–4436
12. Stensballe, A., Jensen, O. N., Olsen, J. V., Haselmann, K. F., and Zubarev, R. A. (2000) Electron capture dissociation of singly and multiply phosphorylated peptides. *Rapid Commun. Mass Spectrom.* **14,** 1793–1800
13. Cooper, H. J., Håkansson, K., and Marshall, A. G. (2005) The role of electron capture dissociation in biomolecular analysis. *Mass Spectrom. Rev.* **24,** 201–222
14. Cooper, H. J., Akbarzadeh, S., Heath, J. K., and Zeller, M. (2005) Data-dependent electron capture dissociation FT-ICR mass spectrometry for proteomic analyses. *J. Proteome Res.* **4,** 1538–1544
15. Creese, A. J., and Cooper, H. J. (2007) Liquid chromatography electron capture dissociation tandem mass spectrometry (LC-ECD-MS/MS) versus liquid chromatography collision-induced dissociation tandem mass spectrometry (LC-CID-MS/MS) for the identification of proteins. *J. Am. Soc. Mass Spectrom.* **18,** 891–897
16. Nielsen, M. L., Savitski, M. M., and Zubarev, R. A. (2005) Improving protein identification using complementary fragmentation techniques in Fourier transform mass spectrometry. *Mol. Cell. Proteomics* **4,** 835–845
17. Ding, L., and Brancia, F. L. (2006) Electron capture dissociation in a digital ion trap mass spectrometer. *Anal. Chem.* **78,** 1995–2000
18. Satake, H., Hasegawa, H., Hirabayashi, A., Hashimoto, Y., and Baba, T. (2007) Fast multiple electron capture dissociation in a linear radio frequency quadrupole ion trap. *Anal. Chem.* **79,** 8755–8761
19. Syka, J. E., Coon, J. J., Schroeder, M. J., Shabanowitz, J., and Hunt, D. F. (2004) Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 9528–9533
20. Molina, H., Matthiesen, R., Kandasamy, K., and Pandey, A. (2008) Comprehensive comparison of collision induced dissociation and electron transfer dissociation. *Anal. Chem.* **80,** 4825–4835
21. Good, D. M., Wirtala, M., McAlister, G. C., and Coon, J. J. (2007) Performance characteristics of electron transfer dissociation mass spectrometry. *Mol. Cell. Proteomics* **6,** 1942–1951
22. Molina, H., Horn, D. M., Tang, N., Mathivanan, S., and Pandey, A. (2007) Global proteomic profiling of phosphopeptides using electron transfer dissociation tandem mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 2199–2204
23. 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., and Hunt, D. F. (2007) Analysis of phosphorylation sites on proteins from Saccharomyces cerevisiae by electron transfer dissociation (ETD) mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 2193–2198
24. Kjeldsen, F., Savitski, M. M., Nielsen, M. L., Shi, L., and Zubarev, R. A. (2007) On studying protein phosphorylation patterns using bottom-up LC-MS/MS: the case of human alpha-casein. *Analyst* **132,** 768–776
25. Woodling, K. A., Eyler, J. R., Tsybin, Y. O., Nilsson, C. L., Marshall, A. G., Edison, A. S., Al-Naggar, I. M., and Bubb, M. R. (2007) Identification of single and double sites of phosphorylation by ECD FT-ICR/MS in peptides related to the phosphorylation site domain of the myristoylated alanine-rich C kinase protein. *J. Am. Soc. Mass Spectrom.* **18,** 2137–2145
26. Sweet, S. M. M., Creese, A. J., and Cooper, H. J. (2006) Strategy for the identification of sites of phosphorylation in proteins: neutral loss triggered electron capture dissociation. *Anal. Chem.* **78,** 7563–7569
27. Sweet, S. M. M., and Cooper, H. J. (2007) Electron capture dissociation in the analysis of protein phosphorylation. *Exp. Rev. Proteomics* **4,** 149–159
28. Sweet, S. M. M., Mardakheh, F. K., Ryan, K. J. P., Langton, A. L., Heath, J. K., and Cooper, H. J. (2008) Targeted online liquid chromatography electron capture dissociation mass spectrometry for the localization of sites of in vivo phosphorylation in human Sprouty2. *Anal. Chem.* **80,** 6650–6657
29. Creese, A. J., and Cooper, H. J. (2008) The effect of phosphorylation on the electron capture dissociation of peptide ions. *J. Am. Soc. Mass Spectrom.* **19,** 1263–1274
30. Akbarzadeh, S., Wheldon, L. M., Sweet, S. M. M., Talma, S., Mardakheh,

F. K., and Heath, J. K. (2008) The deleted in brachydactyly B domain of ROR2 is required for receptor activation by recruitment of Src. *PLoS ONE* **3,** e1873
31. Thingholm, T. E., Jorgensen, T. J. D., Jensen, O. N., and Larsen, M. R. (2006) Highly selective enrichment of phosphorylated peptides using titanium dioxide. *Nat. Protoc.* **1,** 1929–1935
32. Shen, Y., Tolic, N., Hixson, K. K., Purvine, S. O., Pasa-Tolic, L., Qian, W. J., Adkins, J. N., Moore, R. J., and Smith, R. D. (2008) Proteome-wide identification of proteins and their modifications with decreased ambiguities and improved false discovery rates using unique sequence tags. *Anal. Chem.* **80,** 1871–1882
33. Savitski, M. M., Nielsen, M. L., and Zubarev, R. A. (2007) Side-chain losses in electron capture dissociation to improve peptide identification. *Anal. Chem.* **79,** 2296–2302
34. Savitski, M. M., Kjeldsen, F., Nielsen, M. L., and Zubarev, R. A. (2007) Hydrogen rearrangement to and from radical z fragments in electron capture dissociation of peptides. *J. Am. Soc. Mass Spectrom.* **18,** 113–120
35. Zubarev, R. A., Zubarev, A. R., and Savitski, M. M. (2008) Electron capture/transfer versus collisionally activated/induced dissociations: solo or duet? *J. Am. Soc. Mass Spectrom.* **19,** 753–761
36. Steen, H., Jebanathirajah, J. A., Rush, J., Morrice, N., and Kirschner, M. W. (2006) Phosphorylation analysis by mass spectrometry: myths, facts and the consequences for qualitative and quantitative measurements. *Mol. Cell. Proteomics* **5,** 172–181
37. Thingholm, T. E., Jensen, O. N., Robinson, P. J., and Larsen, M. R. (2008) SIMAC (sequential elution from IMAC), a phosphoproteomics strategy for the rapid separation of monophosphorylated from multiply phosphorylated peptides. *Mol. Cell. Proteomics* **7,** 661–671
38. Blom, N., Gammeltoft, S., and Brunak, S. (1999) Sequence and structure-based prediction of eukaryotic phosphorylation sites. *J. Mol. Biol.* **294,** 1351–1362
39. Palumbo, A. M., and Reid, G. E. (2008) Evaluation of gas-phase rearrangement and competing fragmentation reactions on protein phosphorylation site assignment using collision induced dissociation-MS/MS and MS3. *Anal. Chem.* **80,** 9735–9747
40. Rush, J., Moritz, A., Lee, K. A., Guo, A., Goss, V. L., Spek, E. J., Zhang, H., Zha, X. M., Polakiewicz, R. D., and Comb, M. J. (2005) Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. *Nat. Biotechnol.* **23,** 94–101
41. Lind, S. B., Molin, M., Savitski, M. M., Emilsson, L., Astrom, J., Hedberg, L., Adams, C., Nielsen, M. L., Engstrom, A., Elfineh, L., Andersson, E., Zubarev, R. A., and Pettersson, U. (2008) Immunoaffinity enrichments followed by mass spectrometric detection for studying global protein tyrosine phosphorylation. *J. Proteome Res.* **7,** 2897–2910
42. Savitski, M. M., Nielsen, M. L., Kjeldsen, F., and Zubarev, R. A. (2005) Proteomics-grade de novo sequencing approach. *J. Proteome Res.* **4,** 2348–2354
43. Littlefield, K., and Monroe, M. (2008) *Venn Diagram Plotter*, Pacific Northwest National Laboratory, Richland, WA