Strategy for the Identification of Sites of Phosphorylation in Proteins: Neutral Loss Triggered Electron Capture Dissociation

ARTICLE *in* ANALYTICAL CHEMISTRY · DECEMBER 2006

Impact Factor: 5.64 · DOI: 10.1021/ac061331i · Source: PubMed

3 AUTHORS, INCLUDING:

Steve M M Sweet
University of Sussex
21 PUBLICATIONS **815** CITATIONS
SEE PROFILE

Andrew James Creese
University of Birmingham
37 PUBLICATIONS **433** CITATIONS
SEE PROFILE

Strategy for the Identification of Sites of Phosphorylation in Proteins: Neutral Loss Triggered Electron Capture Dissociation

Steve M. M. Sweet, Andrew J. Creese, and Helen J. Cooper*

School of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

We have previously demonstrated the suitability of data-dependent electron capture dissociation (ECD) for incorporation into proteomic strategies. The ability to directly determine sites of phosphorylation is a major advantage of electron capture dissociation; however, the low stoichiometry associated with phosphorylation means that phosphopeptides are often overlooked in data-dependent ECD analyses. In contrast, collision-induced dissociation (CID) tends to result in loss of the labile phosphate group, often at the expense of sequence fragments. Here, we demonstrate a novel strategy for the characterization of phosphoproteins which exploits the neutral loss feature of CID such that *focused* ECD of phosphopeptides is achieved. Peptides eluting from a liquid chromatograph are first subjected to CID, and if a neutral loss of 98 Da (corresponding to H_3PO_4) from the precursor is observed, ECD of that same precursor is performed; i.e., the method comprises neutral loss triggered ECD (NL-ECD-MS/MS). The method was applied to tryptic digests of β-casein and α-casein. For α-casein, four sites of phosphorylation were identified with NL-ECD-MS/MS compared with a single site identified by ECD-MS/MS. The method also resulted in ECD of a doubly phosphorylated peptide. A further benefit of the method is that overall protein sequence coverage is improved. Sequence information from non-phosphorylated peptides is obtained as a result of the CID step.

Electron capture dissociation (ECD)[1] Fourier transform ion cyclotron resonance (FT-ICR)[2] mass spectrometry offers numerous advantages for the analysis of peptides and proteins and, consequently, has many potential benefits for the field of proteomics. ECD is a tandem mass spectrometry (MS/MS) technique in which trapped ions are irradiated with low-energy electrons. For peptides and protein ions, electron capture is accompanied by cleavage of the N−Cα bond to form c and $z^{\bullet}$ (or $c^{\bullet}$ and z) backbone fragments. ECD results in extensive cleavage along the peptide backbone and therefore provides excellent sequence tags.[3,4] The retention of labile posttranslational modifications (including phosphorylation, glycosylation, acylation, ubiquitination, and sumoylation)[5−11] on backbone fragments has been widely reported. This feature enables direct unambiguous assignment of the sites of modification. ECD is also unique in that disulfide bonds are preferentially cleaved.[12] Traditional tandem mass spectrometry techniques, also known as "slow-heating" techniques, such as collision-induced dissociation (CID), result in slow activation of the precursor ions and fragments form via the lowest energy pathway(s). For peptide and protein ions, cleavage of the peptide bond ($N-C_O$) occurs resulting in b and y backbone fragments.[13] For posttranslationally modified peptides, the foremost fragmentation channel is cleavage of the labile modification. While this observation confirms the presence of the modification, it does not necessarily allow localization of the site.

The introduction of the heated dispenser cathode as a source of electrons resulted in reduced time scales (ms) for ECD experiments.[14] Consequently, it was possible to couple on-line liquid chromatography (LC) with ECD mass spectrometry.[15,16] We have recently demonstrated that ECD can be incorporated into data-dependent LC MS/MS strategies for the analysis of proteomic samples: LC ECD MS/MS was applied to the analysis of a tryptic

* To whom correspondence should be addressed. Telephone: +44 (0)121 414 7527. Fax: +44 (0)121 414 5925. Email: H.J.Cooper@bham.ac.uk.

(1) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265−3266.
(2) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. *Mass Spectrom. Rev.* **1998**, *17*, 1−35.
(3) Zubarev, R. A.; Horn, D. M.; Fridriksson, E. K.; Kelleher, N. L.; Kruger, N. A.; Lewis, M. A.; Carpenter, B. K.; McLafferty, F. W. *Anal. Chem.* **2000**, *72*, 563−573.
(4) Creese, A. J.; Cooper, H. J. *J. Am. Soc. Mass Spectrom.*, submitted.
(5) Stensballe, A.; Jensen, O. N.; Olsen, J. V.; Haselmann, K. F.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 1793−1800.
(6) Shi, S. D.-H.; Hemling, M. E.; Carr, S. A.; Horn, D. M.; Lindh, I.; McLafferty, F. W. *Anal. Chem.* **2001**, *73*, 19−22.
(7) Mirgorodskaya, E.; Roepstorff, P.; Zubarev, R. A. *Anal. Chem.* **1999**, *71*, 4431−4436.
(8) Håkansson, K.; Cooper, H. J.; Emmett, M. R.; Costello, C. E.; Marshall, A. G.; Nilsson, C. L. *Anal. Chem.* **2001**, *73*, 4530−4536.
(9) Guan, Z. Q. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 1443−1447.
(10) Cooper, H. J.; Heath, J. K.; Jaffray, E.; Hay, R. T.; Lam, T. T.; Marshall, A. G. *Anal. Chem.* **2004**, *76*, 6982−6988.
(11) Cooper, H. J.; Tatham, M. H.; Jaffray, E.; Heath, J. K.; Lam, T. T.; Marshall, A. G.; Hay, R. T. *Anal. Chem.* **2005**, *77*, 6310−6319.
(12) Zubarev, R. A.; Kruger, N. A.; Fridriksson, E. K.; Lewis, M. A.; Horn, D. M.; Carpenter, B. K., McLafferty, F. W. *J. Am. Chem. Soc.* **1999**, *121*, 2857−2862.
(13) Roepstorff, P.; Fohlman, J. *Biol. Mass Spectrom.* **1984**, *11*, 601.
(14) Tsybin, Y. O.; Hakansson, P.; Budnik, B. A.; Haselmann, K. F.; Kjeldsen, F.; Gorshkov, M.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 1849−1854.
(15) Davidson, W.; Frego, L. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 993−998.
(16) Palmblad, M.; Tsybin, Y. O.; Ramstrom, M.; Bergquist, J.; Hakansson, P. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 988−992.

digest of the protein Fc-ROR2 isolated from chondrocytes.[17] A comparison of the LC ECD MS/MS approach with the more traditional LC CID MS/MS approach for protein identification suggests that the CID approach results in greater protein sequence coverage—a consequence of the speed of analysis—whereas ECD provides longer peptide sequence tags and higher confidence in peptide assignment.[4] Zubarev and co-workers have developed methods for protein identification that combine alternating CID and ECD scans in single LC MS/MS experiments.[18−22]

Phosphorylation and dephosphorylation of proteins play a key role in numerous regulatory processes, e.g., cell signaling. To understand these events, it is important that the sites and extent of phosphorylation can be characterized. Tandem mass spectrometry for the sequencing of phosphopeptides is well-established.[23−25] Generally, phosphorylation occurs on serine, threonine, and tyrosine residues. As mentioned above, CID of phosphopeptides results in loss of phosphoric acid (H_3PO_4, −98 Da) from phosphoserine and phosphothreonine and loss of the phosphate group (HPO_3, −80 Da) from phosphotyrosine and phosphohistidine.[26−30] These neutral losses are particularly useful in determining the presence of the modification. Loss of phosphoric acid from phosphoserine and phosphothreonine produces dehydroalanine (69 Da) and dehydroaminobutyric acid (83 Da), respectively. Those residue masses are unique and so sites of modification may be assigned by MS/MS; however, loss of the modification is favored over backbone cleavage, thus limiting determination of the site of phosphorylation. Characterization of sites of phosphorylation is further complicated by the fact that signaling proteins exist in low abundance and the modification typically occurs at low stoichiometry. To circumvent these problems, phosphopeptide enrichment is performed. Methods include use of phospho-specific antibodies, chromatographic enrichment, e.g., immobilized metal affinity chromatography or metal dioxide affinity chromatography, and chemical modification.[25,31−33]

Zubarev and co-workers have applied their alternating CID/ECD FT-ICR approach to the characterization of phosphopeptides.[22] Peptide digests of bovine serum albumin spiked with a synthetic phosphopeptide, osteopontin, and ovalbumin were analyzed. Independent CID and ECD were performed in the ICR cell on all eluting peptides of sufficient abundance and sites of phosphorylation identified. This method takes advantage of the complementary benefits of ECD and CID; however, the speed of analysis is relatively slow. The duty cycle of the ECD event is by its nature relatively long and ICR detection of all events (survey scan, ECD and CID scans) further lengthens analysis time. Consequently, it is likely that many peptides are overlooked.

Hunt and co-workers recently introduced the technique of electron transfer dissociation (ETD).[34] Electrons are transferred from anions, e.g., fluoranthracene, to multiply charged peptide cations. The resulting fragmentation is analogous to that observed following ECD, i.e., production of c and z ions with retention of labile posttranslational modifications. Data-dependent on-line LC ETD was applied to phosphopeptides enriched from a digest of purified nuclear proteins enabling identification of the sites of phosphorylation in the peptide ERpSLpSRER.[34] In ETD, fragment formation and detection occurs in an ion trap mass spectrometer, and the cycle time is therefore much reduced; however, there are limitations. ETD is inefficient for doubly charged peptide ions. Such ions are observed in abundance in the electrospray mass spectra of tryptic peptides. This drawback has been circumvented by use of proteases, such as Glu-C,[35] which produce longer and more highly charged peptide ions by electrospray. However, the more highly charged the precursor, the more complex the ETD mass spectrum. Interpretation of complex ETD spectra is hampered by the limited resolving power of the ion trap mass spectrometer and consequently, a further step—proton-transfer charge reduction—has been included.[35] The mass spectrum is thus simplified; however, larger backbone fragments are charge-reduced such that they fall outside of the m/z range of the instrument, constraining the depth of observed coverage and unambiguous localization of modifications.

Recently, McLuckey and co-workers have combined ion/ion reactions inducing charge inversion with electron-transfer dissociation in the analysis of phosphopeptide anions.[36] A 1 μM solution of α-casein tryptic digest was subjected to direct infusion negative mode electrospray. $[M - H]^-$ ions of peptide EQLpSTpSEENpSKK were isolated and subjected to ion/ion reaction with 7+ cations of diaminobutane core dendrimer generation 4 (DAB-g4), resulting in singly and doubly protonated phosphopeptide cations. The $[M + 2H]^{2+}$ ions were then subjected to ETD, enabling the assignment of pSer4 and pSer10. Phosphorylation at Ser10 has not been observed previously. The third phosphorylation site was localized to either Thr5 or Ser6. This is an important development as many tryptic phosphopeptides are acidic and do not readily ionize with positive electrospray.

The ideal proteomic experiment would enable full characterization of the protein in question, i.e., full sequence coverage would be obtained and all sites of modification would be localized. To that end, we demonstrate here neutral loss triggered ECD for the characterization of phosphoproteins. The work was performed

(17) Cooper, H. J.; Akbarzadeh, S.; Heath, J. K.; Zeller, M. *J. Proteome Res.* **2005**, *4*, 1538−1544.
(18) Zubarev, R. A. *Curr. Opin. Biotechnol.* **2004**, *15*, 12−16.
(19) Nielsen, M. L.; Savitski, M. M.; Zubarev, R. A. *Mol. Cell. Proteomics* **2005**, *4*, 835−845.
(20) Savitski, M. M.; Nielsen, M. L.; Kjeldsen, F.; Zubarev, R. A. *J. Proteome Res.* **2005**, *4*, 2348−2354.
(21) Savitski, M. M.; Nielsen, M. L.; Zubarev, R. A. *Mol. Cell. Proteomics* **2005**, *4*, 1180−1188.
(22) Kocher, T.; Savitski, M. M.; Nielsen, M. L.; Zubarev, R. A. *J. Proteome Res.* **2006**, *5*, 659−668.
(23) Carr, S. A.; Huddleston, M. J.; Annan, R. S. *Anal. Biochem.* **1996**, *239*, 180−192.
(24) McLachlin, D. T.; Chait, B. T. *Curr. Opin. Chem. Biol.* **2001**, *5*, 591−602.
(25) Mann, M.; Ong, S.-E.; Gronburg, M.; Steen, H.; Jensen, O. N.; Pandey, A. *Trends Biotechnol.* **2002**, *20*, 261−268.
(26) Hunter, A. P.; Games, D. E. *Rapid Commun. Mass Spectrom.* **1994**, *8*, 559−570.
(27) Wilm, M.; Neubauer, G.; Mann, M. *Anal. Chem.* **1996**, *68*, 527−533.
(28) Annan, R. S.; Carr, S. A. *Anal. Chem.* **1996**, *68*, 3413−3421.
(29) Medzihradszky, K. F.; Phillips, N. J.; Senderowicz, L.; Wang, P.; Turck, C. W. *Protein Sci.* **1997**, *6*, 1405−1411.
(30) Neubauer, G.; Mann, M. *Anal. Chem.* **1999**, *71*, 235−242.
(31) Pinkse, M. W. H.; Uitto, P. M.; Hilhorst, M. J.; Ooms, B.; Heck, A. J. R. *Anal. Chem.* **2004**, *76*, 3935−3943.
(32) Larsen, M. R.; Thingholm, T. E.; Jensen, O. N.; Roepstorff, P.; Jorgensen, T. J. D. *Mol. Cell. Proteomics* **2005**, *4*, 873−886.
(33) Kweon, H. K.; Hakansson, K. *Anal. Chem.* **2006**, *78*, 1743−1749.
(34) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 9528−9533.
(35) Coon, J. J.; Ueberheide, B.; Syka, J. E. P.; Dryhurst, D. D.; Ausio, J.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 9463−9468.
(36) Gunawardena, H. P.; Emory, J. F.; McLuckey, S. A. *Anal. Chem.* **2006**, *78*, 3788−3793.

on a hybrid linear ion trap FT-ICR mass spectrometer (Thermo Finnigan LTQ FT). CID in the ion trap was combined in a controlled manner with ECD in the ICR cell. The method takes advantage of the loss of phosphoric acid from phosphoserine/threonine residues observed in CID to initiate ECD of phosphopeptides only. The method thereby exploits the advantages of ECD, while tempering the disadvantages of a full data-dependent ECD analysis. Providing the precursor phosphopeptide ions are sufficiently abundant to trigger CID, the neutral loss triggered (NL)-ECD-MS/MS method allows direct determination by ECD of sites of phosphorylation regardless of their relative abundance. The relatively long duty cycle for each ECD event, and the accompanying possibility that a low-abundance phosphopeptide ion may not be picked for ECD, is overcome. Moreover, sequence information derived from CID mass spectra is obtained for nonphosphorylated peptides. CID analysis time is fast; therefore, overall protein sequence coverage is improved.

EXPERIMENTAL METHODS

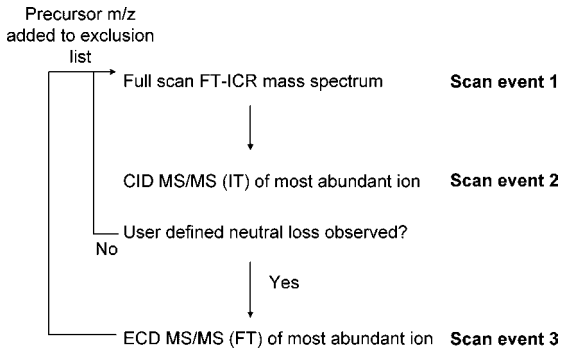
Preparation of Samples for FT-ICR Mass Spectrometry. A 200-μg aliquot of β-casein (Sigma-Aldrich, Dorset, UK) was digested overnight with trypsin Gold (Promega, Madison, WI) in ammonium bicarbonate (50 mM) at a ratio of 1:50 (w/w) at 37 °C. The digest was diluted to a final concentration of 40 fmol/μL with 0.1% formic acid (Fisher Scientific, Loughborough, UK).

A 400-μg aliquot of α-casein (Sigma-Aldrich) was digested with modified trypsin (Promega) in ammonium bicarbonate (50 mM) at a ratio of 1:50 (w/w) for 16 h at 37 °C. The reaction was quenched by addition of formic acid to 0.5%. The digest mixture was diluted to 100 fmol/μL with 0.1% formic acid.

Liquid Chromatography. On-line liquid chromatography was performed by use of a Micro AS autosampler and Surveyor MS pump (Thermo Electron). Aliquots of tryptic digests (β-casein, 10 μL of 40 fmol/μL sample; α-casein, 10, 4, and 2 μL of 100 fmol/μL sample) were loaded onto a 75-μm (internal diameter) Integrafrit (New Objective) C18 resolving column (length 10 cm) and separated over a 50 min gradient from 5 to 90% acetonitrile (Baker). Peptides eluted directly (~400 nL/min) via the nanospray source into a Thermo Finnigan LTQ FT mass spectrometer.

Neutral Loss Triggered ECD FT-ICR Tandem Mass Spectrometry. The mass spectrometer alternated between a full FT-MS scan (m/z 200−2000), a subsequent CID MS/MS scan of the most abundant ion, and, if a neutral loss of 98 Da from the precursor ion was observed in the CID mass spectrum, an ECD MS/MS scan of the same precursor ion. Data acquisition was controlled by Xcalibur 2.0 software with Developer's Kit (build 13). Survey scans were acquired in the ICR cell with a resolution of 100 000 at m/z 400. Precursor ions were isolated and subjected to CID in the linear ion trap. Isolation width was 10 Th. Only doubly and triply charged precursor ions were selected for subsequent MS/MS. All other charge states were rejected. Automatic gain control (AGC) was used to accumulate sufficient precursor ions (target value 100 000 (corresponding to the number of charges), maximum fill time 350 ms). CID was performed with helium gas at normalized collision energy of 35%. Precursor ions were activated for 30 ms. If one of the five most abundant fragment ion peaks was observed at Δ m/z −49.0 or −32.7 from the precursor m/z, corresponding to a neutral loss of 98 Da from a doubly or triply charged peptide respectively, ECD of the same precursor ion was initiated. Precursor ions were isolated in the linear ion trap and transferred to the ICR cell for ECD. Isolation width was 10 Th. AGC was used to accumulate sufficient precursor ions (target value 1 000 000, maximum fill time 4 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 ~1.1 A. Ions were irradiated for 70 ms at 15% energy (corresponding to a cathode potential of −12.55 V). Each ECD scan comprised four coadded microscans, acquired with a resolution of 25 000 at m/z 400. Dynamic exclusion was enabled either after ECD or after CID if no ECD was triggered. The exclusion window was set to 5 ppm with an exclusion time of 90 s.

Scheme 1. Summary of the Neutral Loss Triggered ECD Tandem Mass Spectrometry Method



Data were analyzed by use of the Xcalibur 2.0 software. Data were searched against the nonredundant NCBI bovine database using the SEQUEST algorithm within the Bioworks 3.2 software package (Thermo Electron Corp., Bremen, Germany).

Data-Dependent ECD FT-ICR Tandem Mass Spectrometry. In this experiment, the mass spectrometer alternated between a full FT-MS scan (m/z 200−2000) and two subsequent ECD MS/MS scans of the two most abundant precursor ions. Survey scan and ECD scan parameters were as described for the NL-ECD-MS/MS method.

RESULTS AND DISCUSSION

Scheme 1 outlines the NL-triggered ECD experiment. As peptides elute from the column into the mass spectrometer they are mass analyzed in the FT-ICR. The most abundant ion is subjected to CID in the ion trap. If a neutral loss of 98 Da (corresponding to H_3PO_4) is not observed, the precursor m/z is added to the dynamic exclusion list, and the analysis continues with a full FT-MS scan. If, on the other hand, the neutral loss is observed, ECD in the ICR cell of the same precursor ion is triggered. (This is not MS^3 of the neutral loss ion.) Following ECD, the precursor m/z is added to the dynamic exclusion list and the analysis resumes with scan event 1. The speed of the survey scan (~1 s) and the CID event (~350 ms) are such that nonphosphorylated species are quickly rejected and the next most abundant peptide investigated, thus addressing the problem of phosphopeptides coeluting with other peptides. A phosphopeptide eluting with a chromatographic peak width of 40 s is unlikely to be overlooked for ECD analysis (m/z values remain on the exclusion list

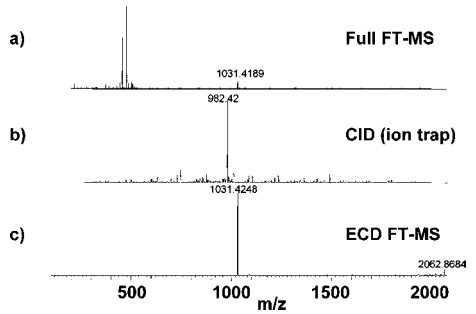


Figure 1. Subsequent mass spectra obtained during the NL-ECD-MS/MS analysis of β-casein: (a) full FT-MS survey scan at retention time 23.7 min; (b) CID MS/MS scan of precursor ion m/z 1031.4, obtained in the linear ion trap; (c) ECD MS/MS scan of precursor ion m/z 1031.4 obtained in the ICR cell.

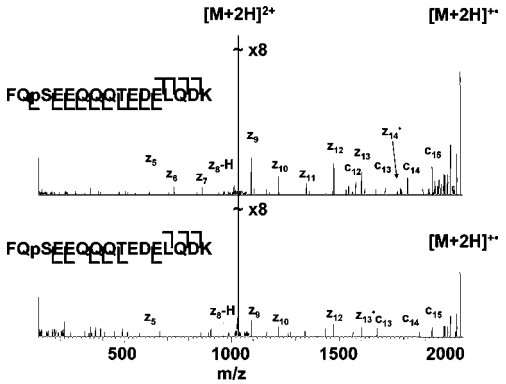


Figure 2. ECD mass spectra obtained for β-casein peptide [FQpSEEQQQTEDEQDK]$^{2+}$ precursor ions (m/z 1031.41) when 400 (top) and 200 fmol (bottom) of the tryptic digest were loaded on column. Observed cleavages are shown inset.

for 90 s), providing, of course, that it is of sufficient abundance to trigger CID.

Figure 1 shows subsequent mass spectra in the NL-triggered ECD analysis of a tryptic digest of β-casein. In the first scan event (Figure 1a), a full FT-ICR mass spectrum is obtained. The mass spectrum contains a peak m/z 1031.4189. This peak corresponds to the doubly protonated, singly phosphorylated[48−63] FQpSEEQQQTEDELQDK tryptic peptide ion (m/z_{calc} 1031.4181; $\Delta = 0.8$ ppm). In the subsequent scan event (Figure 1b), this ion was selected for CID in the linear ion trap. A peak at m/z 982.42 was observed corresponding to a neutral loss of 98 Da. Consequently, the precursor ion (m/z_{meas} 1031.4248; m/z_{calc} 1031.4181, $\Delta = 6$ ppm) was selected for ECD in the final scan event (Figure 1c). The charged-reduced species $[M + 2H]^{+\bullet}$ is observed at m/z 2062.8684 (m/z_{calc} 2062.8368, $\Delta = 15$ ppm). Note that the mass accuracies are lower in the ECD mass spectrum than in the full FT-MS spectrum. This is a consequence of the relationship between AGC targets and mass calibration. The full FT-MS scan comprises a single microscan, an AGC target of 500 000 and a maximum injection time of 1 s. The scan shown in Figure 1a had an actual injection time of 52 ms and the AGC target was achieved. The ECD scan comprises four microscans, has an AGC target of 1 000 000 and a maximum injection time for each microscan of 4 s. For the scan shown in Figure 1c, the maximum injection time was reached and the AGC target was not achieved (see below). The A and B parameters used in frequency-to-m/z calibration[37,38] are calculated for particular ion populations defined by the AGC targets. If the ICR cell is underfilled, i.e., the AGC target is not reached, inappropriate A and B parameters for the actual ion population will be applied. The above observation demonstrates the need to use the precursor mass measured in the full scan in peptide assignments and database searching.

Figure 2 compares the ECD mass spectra obtained from precursor ion m/z 1031.42 when 400 fmol (top) and 200 fmol (bottom) were loaded on the column. In both cases, the AGC target was set to 1 000 000; however, in the 400-fmol case, 31% of the target was reached; i.e., ECD was performed on ~155 000 precursor ions (ions are doubly charged), and in the 200-fmol case, 20% of the target was reached; i.e., ECD was performed on ~100 000 precursor ions. Nevertheless, for the 400-fmol case, virtually complete peptide sequence coverage was obtained − 14 out of 15 N−Cα bonds were cleaved—and of the two possibilities (Ser3 or Thr9), the site of phosphorylation was identified unambiguously as Ser3 (Ser50 on full-length protein). In the ECD mass spectrum obtained from the 200-fmol sample, 9 out of 15 N−Cα bonds were cleaved and it was possible to assign the site of phosphorylation based on the fragmentation pattern. Note that β-casein contains five sites of phosphorylation, four of which reside on the tryptic peptide ELEELNVPGEIVEpSLpSpSpSEESITRINK. This peptide is highly acidic and is rarely seen in positive electrospray mass spectra. It is unsurprising therefore that identification of those sites of phosphorylation was not achieved. In their top-down activated ion ECD analysis of intact β-casein, McLafferty and co-workers[6] were able to unambiguously confirm the first of these sites and to localize the three remaining sites to five possible locations. The fifth site, which was confirmed unambiguously by the NL-ECD method, was located to one of four possible sites. Although, this demonstrates the benefits of top-down analysis, it should be noted that the intact β-casein was electrosprayed at a concentration of 10 μM, which is outside the realms of most proteomics experiments.

Neutral loss triggered ECD analysis was also performed on tryptic digests of α-casein. The sample contained both S1 and S2 variants of α-casein. The sequences of the two variants are shown in Chart 1. The first 15 residues of each represent the signal sequence. Known sites of phosphorylation are shown highlighted. Three experiments were performed in which 200 fmol, 400 fmol, and 1 pmol were loaded on the column. In the 200-fmol experiment, ECD was triggered for precursor ions of m/z 733.8099 (corresponding to doubly protonated α-S2-casein phosphopeptide [138−149], m/z_{calc} 733.8099), m/z 830.9022 (doubly protonated α-S1-casein phosphopeptide [106−119], m/z_{calc} 830.9010, $\Delta = 1.4$ ppm), m/z 964.3481 (doubly protonated α-S1-casein doubly phosphorylated peptide [43−58], m/z_{calc} 964.3497, $\Delta = 1.6$ ppm), and m/z 976.4809 (doubly protonated α-S1-casein phosphopeptide [104−116], m/z_{calc} 976.4801, $\Delta = 0.8$ ppm). In addition, ECD was triggered for precursor ions of m/z 744.8016 and 841.8920. Those precursor ions correspond to the $[M + H + Na]^{2+}$ ions of phosphopeptides S2[138−149] and S1[106−119]. Figure 3 shows the ECD mass spectra obtained for $[M + 2H]^{2+}$ ions of TVDME-

(37) Ledford, E. B., Jr.; Rempel, D. L.; Gross, M. L. *Anal. Chem.* **1984**, *56*, 2744−2748.
(38) Shi, S. D.-H; Drader, J. J.; Freitas, M. A.; Hendrickson, C. L.; Marshall, A. G. *Int. J. Mass Spectrom.* **2000**, *195/196*, 591−598.

Chart 1. Sequences of α-S1 and α-S2 Casein[a]

α-S1-casein

mklliltclvavala	RPKHPIKHQG	LPQEVLNENL	LRFFVAPFPE	VFGKEKVNEL	SKDIGSESTE
DQAMEDIKQM	EAESISSSEE	IVPNSVEQKH	IQKEDVPSER	YLGYLEQLLR	LKKYKVPQLE
IVPNSAEERL	HSMKEGIHAQ	QKEPMIGVNQ	ELAYFYPELF	RQFYQLDAYP	SGAWYYVPLG
TQYTDAPSFS	DIPNPIGSEN	SEKTTMPLW			

α-S2-casein

mkffiftcllavala	KNTMEHVSSS	EESIISQETY	KQEKNMAINP	SKENLCSTFC	KEVVRNANEE
EYSIGSSSEE	SAEVATEEVK	ITVDDKHYQK	ALNEINQFYQ	KFPQYLQYLY	QGPIVLNPWD
QVKRNAVPIT	PTLNREQLST	SEENSKKTVD	MESTEVFTKK	TKLTEEEKNR	LNFLKKISQR
YQKFALPQYL	KTVYQHQKAM	KPWIQPKTKV	IPYVRYL		

[a] Known sites of phosphorylation are highlighted.

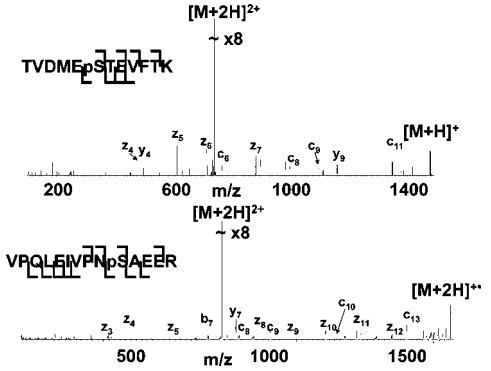


Figure 3. ECD mass spectra obtained for $[M + 2H]^{2+}$ ions of α-S2-casein peptide TVDMEpSTEVFTK (*m/z* 733.81) (top) and α-S1-casein peptide VPQLEIVPNpSAEER (*m/z* 830.90) (bottom) when 200 fmol of the tryptic digest was loaded on column. Observed cleavages are shown inset.

pSTEVFTK (S2[138−149], *m/z* 733.81) (top) and VPQLEIVPNpSAEER (S1[106−119], *m/z* 830.90) (bottom). Four possible phosphorylation sites exist in peptide TVDMESTEVFTK. The ECD mass spectrum obtained reveals cleavage of 6 out of 11 N−Cα bonds. Nonetheless, cleavage either side of Ser6 and Thr7, and C-terminal to Thr11, enables assignment of the site of phosphorylation to Ser6 (Ser143 of full-length protein). In addition to the c and z fragments, y_4 and y_9 fragments were observed. Production of a• and y ions is a minor fragmentation channel in ECD.[12] ECD of $[VPQLEIVPNpSAEER]^{2+}$ resulted in cleavage of 10 out of 13 N−Cα bonds, providing both sequence information and confirmation of the site of phosphorylation (Ser115 of full-length protein). The b_7 and y_7 backbone fragments were also observed. These correspond to cleavage of the peptide bond N-terminal to proline. For the ECD mass spectra shown, 29% (S2-[138−149]) and 39% (S1[106−119]) of the AGC target was achieved. The ECD mass spectra of $[M + 2H]^{2+}$ ions of peptides S1[43−58] and S1[104−116] showed peaks corresponding to the charge-reduced species only; i.e., no backbone fragmentation was observed. In the former mass spectrum, 11% of the AGC target was reached, and in the latter, 1% of the AGC target was reached. It is unsurprising, therefore, that these spectra yielded no useful structural information.

In the 400-fmol experiment (data not shown), ECD was triggered for $[M + Na + H]^{2+}$ ions of peptides TVDMEpSTEVFTK (S2[138−149], m/z_{meas} 744.8019, m/z_{calc} 744.8009, $\Delta = 1.3$ ppm) and VPQLEIVPNpSAEER (S1[106−119], m/z_{meas} 841.8927, m/z_{calc} 841.8920, $\Delta = 0.8$ ppm) and for $[M + 2H]^{2+}$ ions of peptides VPQLEIVPNpSAEER (S1[106−119], m/z_{meas} 830.9014, m/z_{calc} 830.9010, $\Delta = 0.5$ ppm) and YKVPQLEIVPNpSAEER (S1[104−119], m/z_{meas} 976.4804, m/z_{calc} 976.4801, $\Delta = 0.3$ ppm). ECD of $[M + Na + H]^{2+}$ ions of TVDMEpSTEVFTK resulted in cleavage of 4 out of 11 N−Cα bonds. Only 3% of the AGC target was reached. ECD of $[M + 2H]^{2+}$ ions of VPQLEIVPNpSAEER resulted in cleavage of all accessible N−Cα bonds. (Fragments resulting from cleavage within proline are rare[39] due to its cyclic structure.) Approximately 67% of the AGC target was achieved. The peptide sequence and site of phosphorylation are unequivocal. ECD of the doubly protonated YKVPQLEIVPNpSAEER ions resulted in cleavage of only two of the N−Cα bonds (35% AGC target attained).

Loading of 1 pmol of α-casein tryptic digest on the column and performing the experiment resulted in neutral loss triggered ECD of phosphopeptides S2[138−149] TVDMEpSTEVFFK ($[M + 2H]^{2+}$ m/z_{meas} 733.8105, m/z_{calc} 733.8099, $\Delta = 0.8$ ppm; $[M + Na + H]^{2+}$ m/z_{meas} 744.8013, m/z_{calc} 744.8009, $\Delta = 0.5$ ppm); S2[137−149] KTVDEpSTEVFK ($[M + 2H]^{2+}$ m/z_{meas} 797.8571,m/z_{calc} 797.8574, $\Delta = 0.3$ ppm); S1[106−119] VPQLEIVPNpSAEER ($[M + 2H]^{2+}$ m/z_{meas} 830.9017, m/z_{calc} 830.9010, $\Delta = 0.8$ ppm; $[M + Na + H]^{2+}$ m/z_{meas} 841.8929, m/z_{calc} 841.8920, $\Delta = 1$ ppm); S1[104−119] YKVPQLEIVPNpSAEER ($[M + 2H]^{2+}$ m/z_{meas} 976.4807, m/z_{calc} 976.4801, $\Delta = 0.6$ ppm) and doubly phosphorylated S1[43−58] DIGpSEpSTEDQAMEDIK ($[M + 2H]^{2+}$ m/z_{meas} 964.3517, m/z_{calc} 964.3497, $\Delta = 2$ ppm). Figure 4a shows the ECD mass spectra obtained from S2[138−149] TVDMEpSTEVFTK $[M + 2H]^{2+}$ ions (top), $[M + Na + H]^{2+}$ ions (middle), and the tryptic missed-cleavage peptide S2[137−149] KTVDMEpSTEVFTK $[M + 2H]^{2+}$ ions (bottom). The AGC target, i.e., 1 000 000, was reached for ECD of $[M + 2H]^{2+}$ ions of S2[138−149]. Nine out of 11 N−Cα bonds were cleaved and the site of phosphorylation can be confirmed unambiguously. For ECD of $[M + Na + H]^{2+}$ ions of S2[138−149], 87% of the AGC target was reached, and 8 out of 11 N−Cα bonds were cleaved. Again, the site of phosphorylation was confirmed unequivocally. The c ions dominate the fragmentation spectrum, and along with z_7 and z_{11}, all retained Na^+. This phenomenon has been noted previously[40] and is in agreement with the findings of Williams and co-workers,[41] who showed that electron capture occurs preferentially at protonated rather than metalated sites. Conversely, the smaller z ions, z_5 and z_6, retained the hydrogen atom. Only 19% of the AGC target was

(39) Cooper, H. J.; Hudgins, R. R.; Hakansson, K.; Marshall, A. G. *Int. J. Mass Spectrom.* **2003**, *228*, 723−728.
(40) Cooper, H. J.; Hudgins, R. R.; Marshall, A. G. *Int. J. Mass Spectrom.* **2004**, *234*, 23−35.
(41) Iavarone, A. T.; Paech, K.; Williams, E. R. *Anal. Chem.* **2004**, *76*, 2231−2238.

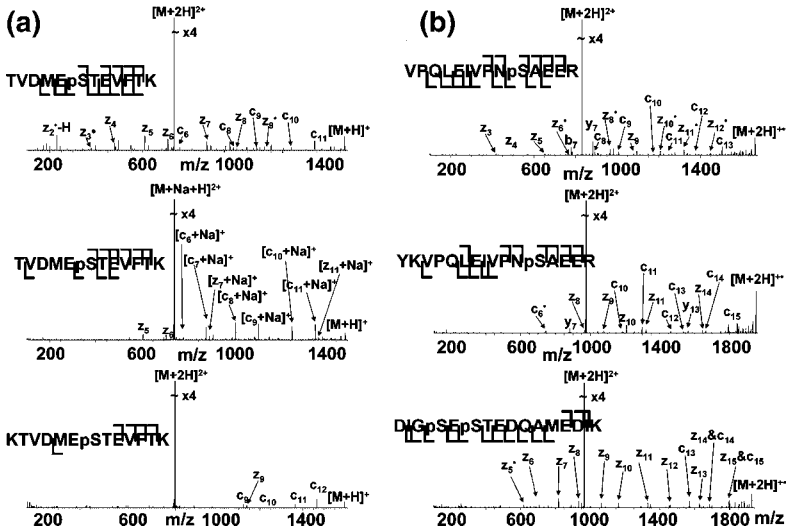


Figure 4. ECD mass spectra obtained when 1 pmol of tryptic digest was loaded on column. Observed cleavages are shown inset. (a) ECD mass spectra obtained from α-S2-casein peptide TVDMEpSTEVFTK $[M + 2H]^{2+}$ ions (top), $[M + Na + H]^{2+}$ ions (middle), and the tryptic missed-cleavage peptide KTVDMEpSTEVFTK $[M + 2H]^{2+}$ ions (bottom). (b) ECD mass spectra obtained from $[M + 2H]^{2+}$ ions of α-S1-casein peptide VPQLEIVPNpSAEER (top), its sister peptide YKVPQLEIVPNpSAEER (middle), and $[M + 2H]^{2+}$ ions of doubly phosphorylated α-S1-casein peptide DIGpSEpSTEDQAMEDIK (bottom).

achieved for ECD of $[M + 2H]^{2+}$ ions of peptide S2[137−149] resulting in cleavage of five N−Cα bonds. Thus, the phosphorylation site was localized to either Ser7 or Thr8. Figure 4b shows the ECD of $[M + 2H]^{2+}$ ions of VPQLEIVPNpSAEER (top) and its sister peptide YKVPQLEIVPNpSAEER (middle). In both cases, the AGC target was achieved. Cleavage of all available N−Cα bonds was observed for VPQLEIVPNpSAEER, confirming the peptide sequence and site of phosphorylation. For ECD of YKVPQLEIVPNpSAEER, all but two available N−Cα bonds were cleaved, again confirming the site of phosphorylation. The ECD mass spectrum of doubly phosphorylated DIGpSEpSTEDQAMEDIK $[M + 2H]^{2+}$ ions is shown in Figure 4b (bottom). A total of 61% of the AGC target was reached. Three possible sites of phosphorylation exist in this peptide (DIGSESTEDQAMEDIK). Virtually complete sequence coverage was obtained (14 out of 15 N−Cα bonds were cleaved). The fragmentation pattern pinpoints Ser4 and Ser6 (Ser46 and Ser48 in the full-length protein) as the sites of modification.

Overall, for α-casein, three sites of phosphorylation were identified in the S1 variant and a further site identified in the S2 variant. The remaining five known sites in S1 are located within the tryptic peptide QMEAEpSIpSpSpSEEIVPNpSVEQK. As in the β-casein example, this peptide is highly acidic and was not observed in survey scans. Similarly, the remaining sites in the S2 variant reside in the following tryptic peptides: NTVpSpSpSEESIIpSQETYK (four sites), NANEEEYSIGpSpSpSEEpSAEVATEEVK (four sites), and EQLpSTpSEENSK (two sites). A possible strategy for overcoming this limitation is the use of alternative protease schemes, e.g., Glu-C or Lys-C, to produce different peptides. Work in this area is ongoing in our laboratory.

An interesting aspect of the new methodology is the potential for false positives; i.e., is ECD triggered for peptides for which a "real" neutral loss was not observed? In the experiment in which 400 fmol of β-casein digest was analyzed, ECD was triggered for four precursor ions. In addition to precursor ions of m/z 1031, ECD was performed on precursor ions of m/z 535.33, 390.70, and 329.01. In the 200-fmol experiment, ECD was also triggered on ions of m/z 390.70. In each case, the precursor peaks were of very low intensity and the proposed neutral loss peak actually corresponded to the isotope of another fragment peak. For the α-casein 200-fmol experiment, in addition to the ECD events described above, ECD was triggered for precursor ions of m/z 234.21, 989.42, and 607.44. The peaks at m/z 234.21 and 607.44 are of very low abundance and the "neutral loss" peaks are noise. A SEQUEST search of the data against the NCBI bovine database suggests the ion of m/z 989.42 is the PLC-β phosphopeptide VDEFGFFLpTWRpSEGK based on the precursor mass; however, neither the CID or ECD data are of sufficient quality to confirm this. In the α-casein 400-fmol experiment, ECD was triggered for ions of m/z 924.37 (corresponding to doubly protonated, singly phosphorylated S1[43−58] DIGSESTEDQAMEDIK), 857.36, and 990.81, in addition to those described above. Inspection of the mass spectra of the latter two species suggests that the neutral loss is real; however, a SEQUEST search against the bovine database did not return any hits. In the α-casein 1-pmol experiment, ECD was also additionally triggered for ions of m/z 924.37 and 990.81; for ions of m/z 752.80 and 852.88 corresponding to doubly sodiated S2[138−149] and S1[106−119], respectively; for ions of m/z 796.36, 658.65, 684.35, and 501.18, which were of low abundance and whose neutral losses fell within the isotope patterns of other fragments; and for ions of m/z 805.86 (corresponding to methionine oxidized S2[137−149]), 932.37 (unknown origin), and 972.35 (corresponding to methionine oxidized S1-[43−58]).

Conversely, the question arises, did ECD fail to trigger for any known phosphopeptides? The known α-casein variant phosphopeptide[32] YLGEYLIVPNpSAEER was observed doubly protonated in the survey scan of the 1-pmol experiment (m/z_{meas} 916.9282, m/z_{calc} 916.9272, $\Delta = 1$ ppm). The ions were selected for CID in the ion trap and a neutral loss peak was observed at m/z 867.83; however, the neutral loss was the seventh most abundant fragment and the experiment was set to trigger if the neutral loss peak was

within the top five fragments. b and y fragments observed in the CID mass spectrum confirm the identity of the peptide. This was the only known phosphopeptide that was observed in the survey scan but did not trigger ECD as a result of the neutral loss.

For comparison, a straightforward LC ECD MS/MS analysis of 400 fmol of the β-casein digest was performed. In this experiment, ECD was not triggered by observation of a neutral loss; rather, the experiment comprised a full FT-MS scan followed by ECD of the two most abundant ions in that scan. ECD of the m/z 1031.42 precursor ion was observed (data not shown). A total of 51% of the AGC target was reached. Twelve out of 15 of the N−Cα bonds were cleaved. Cleavage either side of the serine residue was not observed; however, cleavage either side of the threonine residue was observed allowing assignment of the site of phosphorylation. This result is perhaps not surprising as the β-casein digest is a relatively simple mixture. The equivalent experiment was therefore performed on 400 fmol of the α-casein digest. The S2[138−149] and S1[104−119] phosphopeptides, which triggered ECD in the neutral loss experiment, were not subjected to ECD in the straightforward LC ECD MS/MS experiment. Only doubly protonated ions of S1[106−119] were subjected to ECD. A total of 28% of the AGC target was achieved and 7 out of 13 N−Cα bonds were cleaved, although cleavage was not observed adjacent to the modified amino acid. These results clearly demonstrate that the neutral loss triggered ECD method is a significant improvement on unfocused data-dependent ECD resulting in characterization of a greater number of phosphorylation sites.

The two data sets (LC NL-ECD-MS/MS and LC ECD-MS/MS) obtained from each experiment (400-fmol β-casein digest and 400-fmol α-casein digest) were searched against the nonredundant bovine NCBI database using the SEQUEST algorithm within Bioworks 3.2. In the LC ECD MS/MS data set, three β-casein peptides were identified—a sequence coverage of 13%. For the neutral loss triggered ECD experiment, five peptides were identified—a sequence coverage of 25%. The search of the LC ECD MS/MS of α-casein identified two S1 peptides (sequence coverage 14%) and three S2 peptides (sequence coverage 13%), whereas the LC NL-ECD MS/MS of α-casein identified six S1 peptides (sequence coverage 33%) and five S2 peptides (sequence coverage 23%). These results can be attributed to the greater speed of analysis of the NL-ECD method. In the straightforward ECD method, ECD is triggered for the two most abundant peaks in the survey scan. Each ECD scan comprises four microscans, each with a maximum injection time of 4 s. Clearly, it is possible that less abundant peptide ions will elute without being subjected to tandem mass spectrometry. As mentioned above, phosphopeptides are typically of lower abundance than their nonphosphorylated counterparts. The NL-ECD method subjects (virtually) all eluting peptide cations to CID (each CID scan comprises one microscan with a maximum injection time of 350 ms), thereby obtaining sequence information for a larger proportion of the protein.

CONCLUSION

The results demonstrate the applicability of neutral loss triggered ECD for the characterization of phosphoproteins. NL-ECD-MS/MS enables direct determination of sites of phosphorylation through application of focused ECD and in addition provides sequence information, through use of CID, for a greater proportion of the protein than the standard data-dependent ECD method. The method described exploits the hybrid nature of our instrument, i.e., the combination of a linear ion trap with an FT-ICR. It is important to note, however, the wider applicability of the method. It is possible that NL-ECD-MS/MS could be performed on an instrument comprising a single mass analyzer, e.g., an FT-ICR equipped with CID or infrared multiphoton dissociation capabilities. The duty cycle would be longer in such an experiment. Nevertheless, the benefits of focused ECD would still be realized. Moreover, the NL-ECD-MS/MS approach is not limited to the analysis of Ser- and Thr-phosphopeptides. The method could be adapted to investigate other labile modifications, e.g., loss of HPO_3 from phosphotyrosine, loss of SO_3 from tyrosine sulfation, and carbohydrate losses from O- and N-linked glycopeptides. Finally, in the present case, we have investigated neutral losses from doubly and triply charged peptides (which form the major component of electrosprayed tryptic digests). It is entirely feasible to investigate neutral losses of labile modifications from other charge states via the NL-ECD-MS/MS approach.

ACKNOWLEDGMENT

The authors thank Drs. Martin Zeller and Bernard Delanghe for technical assistance. The Wellcome Trust (074131) (H.J.C.), EU (FP6 EndoTrack, www.endotrack.org) (S.M.M.S.), and the Biotechnology and Biological Sciences Research Council (A.J.C.) are acknowledged for funding.

Received for review July 21, 2006. Accepted August 24, 2006.

AC061331I