Anal. Chem. **2001,** *73,* 19−22

Phosphopeptide/Phosphoprotein Mapping by Electron Capture Dissociation Mass Spectrometry

Stone D.-H. Shi, Mark E. Hemling,* and Steven A. Carr*

Department of Physical and Structural Chemistry, SmithKline Beecham Pharmaceuticals, King of Prussia, Pennsylvania 19406

David M. Horn, Ingemar Lindh, and Fred W. McLafferty*

Baker Chemistry Laboratory, Cornell University, Ithaca, New York 14853

Of methods for dissociation of multiply charged peptide and protein ions, electron capture dissociation (ECD) has the advantages of cleaving between a high proportion of amino acids, without loss of such posttranslational modifications as glycosylation and carboxylation. Here this capability is successfully extended to phosphorylation, for which collisionally activated dissociation (CAD) can cause extensive loss of H_3PO_4 and HPO_3. As shown here, these losses are minimal in ECD spectra, an advantage for measuring the degree of phosphorylation. For phosphorylated peptides, ECD and CAD spectra give complementary backbone cleavages for identifying modification sites. For a 24-kDa heterogeneous phosphoprotein, bovine *β*-casein, activated ion ECD cleaved 87 of 208 backbone bonds that identified a phosphorylation site at Ser-15, and localized three more among Ser-17,-18, -19, and -22 and Thr-24, and the last among four other sites. This is the first direct site-specific characterization of this key posttranslational modification on a protein without its prior degradation, such as proteolysis.

Phosphorylation is the most common and important reversible regulatory modification of proteins, as recognized since the mid-1950s.[1] Protein kinases catalyze the addition of phosphate from a nucleoside triphosphate donor, while protein phosphatases catalyze the removal of specific phosphate groups. A complex network of these reactions controls many aspects of cell growth, metabolism, division, motility, and differentiation. Genes for protein kinases may constitute as much as 3% of the entire eukaryotic genome.[1] Characterization of the sites of phosphorylation and the extent of phosphorylation at each specific site is critical for understanding how these modifications modulate the target activity. Recognition of the disadvantages of traditional phosphopeptide mapping by Edman degradation of ^{32}P-labeled peptides[2] has led to the development of valuable alternative methodologies based on mass spectrometry (MS).[3] These methods generally utilize proteolysis of the phosphoprotein and subsequent identification of the phosphorylated peptide by metal affinity[3j] or characteristic loss of phosphate on treatment with phosphatase[3i] or by collisionally activated dissociation (CAD).[3c,e] If multiple phosphorylation sites are possible on an identified peptide, tandem MS data are used to locate the amino acid residue on which the phosphorylation occurs. Ionized peptides phosphorylated on Ser or Thr typically show partial to complete loss of H_3PO_4 ($m = 98$) in their CAD spectra for the peptide fragment ions containing the phosphorylated amino acid residue. This higher cleavage tendency of the phosphate moiety on phosphoserine and phosphothreonine relative to cleavage of the peptide backbone complicates MS/MS mapping of the peptide phosphorylation sites from their CAD spectra because loss of H_3PO_4 produces the same mass as that of H_2O loss from the nonphosphorylated peptide. Similarly, phosphotyrosine-containing peptides show loss of HPO_3, which produces the same mass as the nonphosphorylated peptide. Further, CAD sequence information often becomes increasingly incomplete for phosphopeptides as the size of the peptide increases. Finally, accuracy in measurement of the extent of phosphorylation is compromised by any partial side-chain loss. These problems have been described in detail recently.[3g,h]

A newly developed ion fragmentation technique, electron capture dissociation (ECD), appears promising for these problems.[4] In comparison to the CAD spectra of peptide and protein ions, ECD spectra show far less side-chain loss of posttranslational modifications such as carboxyl,[5] sulfonate,[4c] and glycosyl[6] groups, but usually with a higher proportion of backbone cleavage to provide sequence information. In this work, we demonstrate for phosphopeptides and a 24-kDa phosphoprotein that ECD also gives more sequence-informative backbone fragmentation with negligible loss of phosphate and therefore gives promise to be a

(1) Hubbard, M. J.; Cohen, P. *Trends Biochem. Sci.* **1993**, *18*, 172−177. Kennelly, P. J.; Potts, M. *J. Bacteriol.* **1996**, *178*, 4759−4764.

(2) Wang, Y. H.; Fiol, C. J.; DePaoli-Roach, A. A.; Bell, A. W.; Hermodson, M. A.; Roach, P. J. *Anal. Biochem.* **1988**, *174*, 537−547. Yan, J. X.; Packer, N. H.; Gooley, A. A.; Williams, K. L. *J. Chromatogr., A* **1998**, *808*, 23−41.

(3) (a) Gibson, B. W.; Cohen, P. *Methods Enzymol.* **1990**, *193*, 480−501. (b) Covey, T. In *Methods in Protein Sequence Analysis*; Jornvall, H., Ed.; Advances in Life Sciences: Birkhauser-Verlag: Basel, 1991; pp 249−257. (c) Huddleston, M. J.; Annan, R. S.; Bean, M. F.; Carr, S. A., *J. Am. Soc. Mass Spectrom.* **1993**, *4*, 710−717. (d) Affolter, M.; Watts, J. D.; Krebs, D. L.; Aebersold R. *Anal. Biochem.* **1994**, *223*, 74−81. (e) Carr, S. A.; Huddleston, M. J.; Annan R. S. *Anal. Biochem.* **1996**, *239*, 180−192. (f) Resing, K. A.; Ahn, N. G. *Methods Enzymol.* **1997**, *283*, 29−44. (g) Dass, C. In *Mass Spectrometry of Biological Materials*; Larsen, B. S., McEwen, C. N., Eds.; Dekker: New York, 1998; pp 247−280. (h) Kassel, D. B.; Blackburn, R. K.; Antonsson, B. In *Mass Spectrometry of Biological Materials*; Larsen, B. S., McEwen, C. N., Eds.; Dekker: New York, 1998; pp 137−158. (i) Zhang, X.; Herring, C. J.; Romano, P. R.; Szczepanowska, J.; Brzeska, H.; Hinnebusch, A. G.; Qin, J. *Anal. Chem.* **1998**, *70*, 2050−2059. (j) Posewitz, M. C.; Tempst, P. *Anal. Chem.* **1999**, *71*, 2883−2892.

10.1021/ac000703z CCC: $20.00
Published on Web 11/30/2000

superior method for phosphopeptide/protein mapping and for measurement of the degree of phosphorylation.

EXPERIMENTAL SECTION

Atrial natriuretic peptide substrate (ANPS) and platelet-derived growth factor (PDGF) receptor substrate were obtained from AnaSpec (San Jose, CA). PY8, a synthetic peptide corresponding to a sequence in the cytoplasmic domain of insulin-like growth factor I receptor (IGFIR-CD) was obtained as a gift from Dr. Susan L. Chen. Bovine β-casein was obtained from Sigma (St. Louis, MO). Samples were dissolved at a concentration of 10 μM in methanol/water/acetic acid (49:49:2) for analysis. Microelectrosprayed (μ-ESI; 0.5 μL/min) ions were fragmented by SORI[7] to generate CAD spectra on a 7-T Fourier transform (FT) MS (Finnigan T70; ThermoQuest, Bremen, Germany) with argon as the collision gas. ECD spectra were obtained on the Cornell FTMS (6 T, resolving power 10^5) equipped with a nanospray source.[4c] Activated ion (AI) ECD[8] is performed in the "in-beam" fashion, which is to turn on the electron gun during the internal ion accumulation (effected by ion collisions with a pulse of N_2 gas). Thus, the only experimental difference between this method and normal ECD is the time lapse between ion collection and electron exposure. Automatic data reduction for isotopic cluster identification and fragment mass assignment of MS/MS spectra were performed by the Cornell THRASH software.[9] For comparison, MS/MS data for all phosphopeptide samples were also obtained on a hybrid quadrupole/time-of-flight mass spectrometer (Q-ToF, Micromass, Beverly, MA).

RESULTS AND DISCUSSION

Phosphopeptide Analysis. The electrospray mass spectrum of the 28-mer phosphopeptide, ANPS,[10] SLRRSpSCFGGRIDRIGAQSGLGCNSFRY, $M_r = 3142.47$, gave molecular ions corresponding to $M_r = 3142.54$ with no evidence (peaks at ±79.97 Da) for phosphorylation heterogeneity. The CAD and ECD spectra of the 4+ ion formed by ESI are compared in Figure 1.

In the SORI CAD spectrum obtained by ESI-FTMS, the $[M + 4H - 98]^{4+}$ peak provides valuable confirmatory evidence for at least one phosphorylation. The CAD spectrum shows 9 b and 17 y fragments (Figure 2, some in multiple charge states and/or with phosphate loss) representing cleavages of 20 of the 27 peptide bonds (74% sequence coverage). However, as expected,[3] the corresponding peaks representing subsequent loss of H_3PO_4 are also abundant; of the nine backbone cleavages yielding fragment ions that originally contained phosphoserine, only two yield undissociated product ions, and four are accompanied by complete H_3PO_4 loss. Despite the extensive phosphate loss observed in the CAD data, with prior knowledge of the peptide sequence, the phosphorylation site can be correctly assigned to Ser-6 among six possible sites.

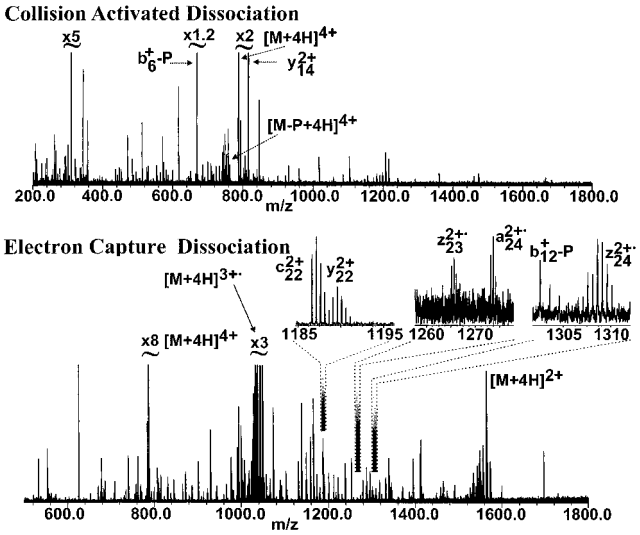

Figure 1. CAD (top) and ECD (bottom) MS/MS spectra of atrial natriuretic peptide substrate (4+).

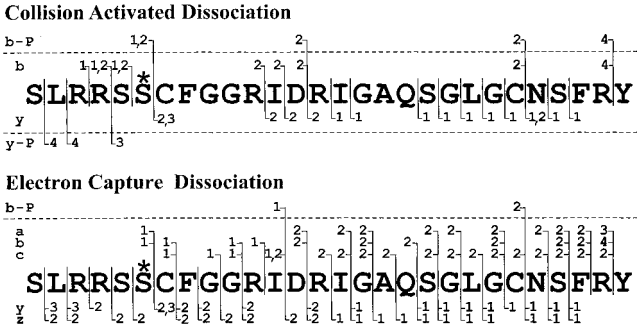

Figure 2. Fragmentation maps derived from the Figure 1 spectra. Asterisk in the sequence denotes phosphorylation site. Numbers are charge states of the ions observed. "−P" denotes fragments with phosphate (H_3PO_4) loss. HPO_3 losses are observed for y_{27} and b_{23}, but not mapped in the figure.

In the ECD spectrum, only two of the 81 fragment ions result from H_3PO_4 loss. These probably result from adventitious CAD; both of these fragment ions are of the b type, which cannot be derived from the a−y and c−z pairs that have been established as the ECD fragmentation pathways.[4] The b ions in ECD experiments might come from collisions caused by the combination of SWIFT isolation followed by a gas pulse for ion cooling. The negligible phosphate loss in the ECD data is consistent with the nonergodic mechanism of ECD[4] and with previous ECD results on other labile posttranslational modifications such as γ-carboxyglutamic acid,[5] sulfonation,[4c] and O-glycosylation.[6]

The 81 fragment ions observed in the ECD experiment represent cleavages between all of the amino acids (100% sequence coverage). The c and $z^{\bullet}$ ions are represented by 14 complementary

(4) (a) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265−3266. (b) 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. (c) 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. (d) Kruger, N. A.; Zubarev, R. A.; Carpenter, B. K.; Kelleher, N. L.; Horn, D. M.; McLafferty, F. W. *Int. J. Mass Spectrom.* **1999**, *182/183*, 1−5. (e) Kruger, N. A.; Zubarev, R. A.; Horn, D. M.; McLafferty, F. W. *Int. J. Mass Spectrom.* **1999**, *185/186/187*, 787−793.

(5) Kelleher, N. L.; Zubarev, R. A.; Bush, K.; Furie, B. C.; McLafferty, F. W.; Walsh, C. T. *Anal. Chem.* **1999**, *71*, 4250−4253.

(6) Mirgorodskaya, E.; Roepstorff, P.; Zubarev, R. *Anal. Chem.* **1999**, *71*, 4431−4436.

(7) Gauthier, J. W.; Trautman, T. R.; Jacobsen, D. B. *Anal. Chim. Acta* **1991**, *246*, 211−225. Senko, M. W.; Speir, J. P.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 2801−2808.

(8) Horn, D. M.; Ge, Y.; McLafferty, F. W. *Anal. Chem.* **2000**, *72*, 4778-4784.

(9) Horn, D. M.; Zubarev, R. A.; McLafferty, F. W. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 320−332. Horn, D. M.; Zubarev, R. A.; McLafferty, F. W. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 10313−10317.

(10) Dautzenberg, F. M.; Muller, D.; Richter, D. *Eur. J. Biochem.* **1993**, *211*, 485−490.

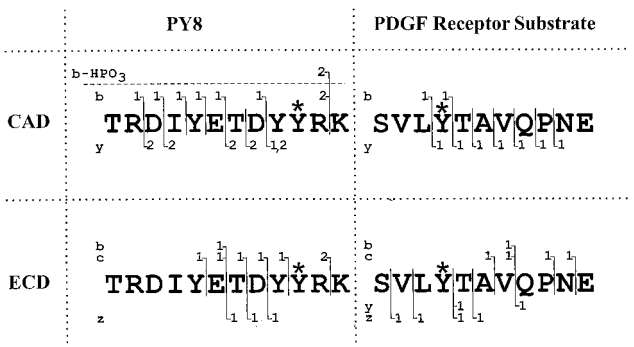

Figure 3. Comparison of fragmentation maps obtained from CAD and ECD MS/MS of a dodecapeptide and an undecapeptide. Numbers are charge states of the ions observed.

pairs, those whose masses sum to that of the molecular ion.[11] Thus, the complete sequence and the site of phosphorylation are easily identified.

ECD and CAD results from two smaller phosphorylated peptides, PDGF receptor substrate and PY8, are also compared (Figure 3). Both peptides contain phosphotyrosine, which is more stable under CAD conditions compared to phosphoserine and phosphothreonine because of the lack of β-elimination pathways.[3d] The less favorable HPO_3 loss is observed for one b ion in the CAD spectrum of PY8 but not in its ECD spectrum. However, one of the peptides (PY8) shows fewer backbone cleavages by ECD than by CAD. For PDGF, fragmentation between the tyrosine and threonine in both spectra (b_4 and y_7 in the CAD spectrum, y_7 and z_7 in the ECD spectrum) localizes the phosphorylation site. In the case of PY8, although the total sequence coverage is lower for ECD, cleavage between the two tyrosines (c_9) permits the localization of the phosphate group. No fragmentation between the two tyrosines is observed in the CAD spectrum. Demonstrating their complementary capabilities, CAD cleaves between 7 pairs of amino acids, ECD cleaves between 8 pairs, and together they cleave between all 10 pairs. Other ECD spectra of phosphorylated peptides have been reported since the submission of this paper.[12]

To compare the CAD and ECD data obtained on the FTMS to that from conventional low-energy dissociation MS/MS instruments, CAD spectra for all the phosphopeptides were obtained on a hybrid quadrupole/time-of-flight instrument. For ANPS, CAD cleaved 25 of the 27 peptide bonds, resulting in a higher sequence coverage of 93% (vs 74% by CAD-FTMS). However, extensive phosphate loss was observed for fragments containing the phosphoserine (Figure 4). Similar CAD results were obtained for the other two phosphopeptides (results not shown). Nonetheless, the location of the phosphate group could be uniquely defined by skillful interpretation of the data. It is evident that CAD on a hybrid quadrupole/time-of-flight instrument produces more fragments (especially b ions) than on a FTMS, presumably because FTMS (a trapping instrument) can only detect long-lived fragments and the CAD fragmentation efficiency of the Q-ToF is higher. However, data interpretation is less prone to interference from background chemical noise on the FTMS because of the higher resolving power. This advantage is more evident for higher molecular weight peptides that yield more fragment ions and therefore have a higher probability of producing overlapping isotopic clusters.

Phosphoprotein Analysis. In the "top down" sequencing approach,[11] posttranslational modifications can be located directly

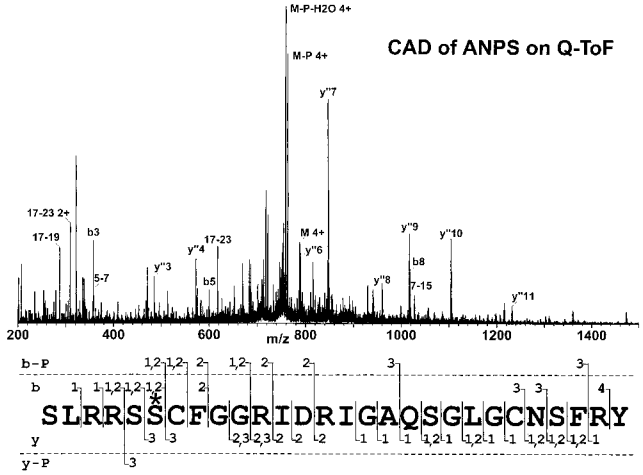

Figure 4. CAD spectrum and the corresponding fragmentation map of atrial natriuretic peptide substrate (4+) on Q-ToF. See Figure 2 legend for symbol definition.

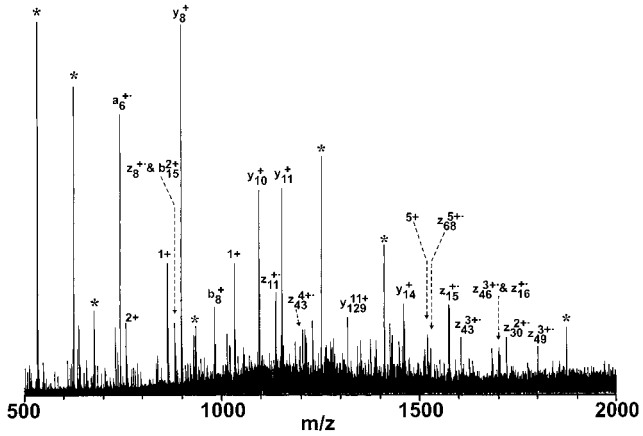

Figure 5. AI ECD spectrum (not expanded) of β-casein. Peaks labeled with an asterisk are due to electronic noise (no isotopic peaks). Examples of isotopic clusters in an expanded spectrum are shown in Figure 1.

by fragmentation of the protein molecular ion.[13] The ECD spectrum of β-casein, a 24-kDa phosphoprotein mixture of at least three variants (A1, M_r = 24 008.2; A2, M_r = 23 968.2; and B, M_r = 24 077.2) with phosphorylation at Ser-15, -17, -18, -19, and -35,[14] is shown in Figure 5. "Activated ion" ECD[8] was necessary to effect ECD on a protein this large; because of its nonergodic nature, ECD cleavage of the backbone does not increase appreciably the internal energy of the rest of the molecule. Thus, ECD backbone cleavage only produces lower mass ionic products if the noncovalent bonds between these products are already activated and cleaved. AI ECD of β-casein produced fewer N-terminal fragments (Figure 6) than AI ECD of 29-, 30-, and 42-kDa proteins reported in recent studies,[8] possibly because of tertiary interaction of the

(11) Kelleher, N. L.; Lin, H. Y.; Valaskovic, G. A.; Aaserud, D. J.; Fridriksson, E. K.; McLafferty, F. W. *J. Am. Chem. Soc.* **1999**, *121*, 806−812.
(12) Stenballe, A.; Jensen, O. N.; Olsen, J. V.; Haselmann, K. F.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 1793−1800.
(13) Kelleher, N. L.; Nicewonger, R. B.; Begley, T. P.; McLafferty, F. W. *J. Biol. Chem.* **1997**, *272*, 32215−32220. Wood, T. D.; Guan, Z.; Borders,C. L., Jr.; Chen, L. C.; Kenyon, G. L.; McLafferty, F. W. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 3362−3365. Fridriksson, E. K.; Beavil, A.; Holowka, D.; Gould, H. J.; Baird, B.; McLafferty, F. W. *Biochemistry* **2000**, *39*, 3369−3376.
(14) Rabideau Dumas, B.; Brignon, G.; Grosclaude, F.; Mercier, J. C. *Eur. J. Biochem.* **1972**, *25*, 505−514.

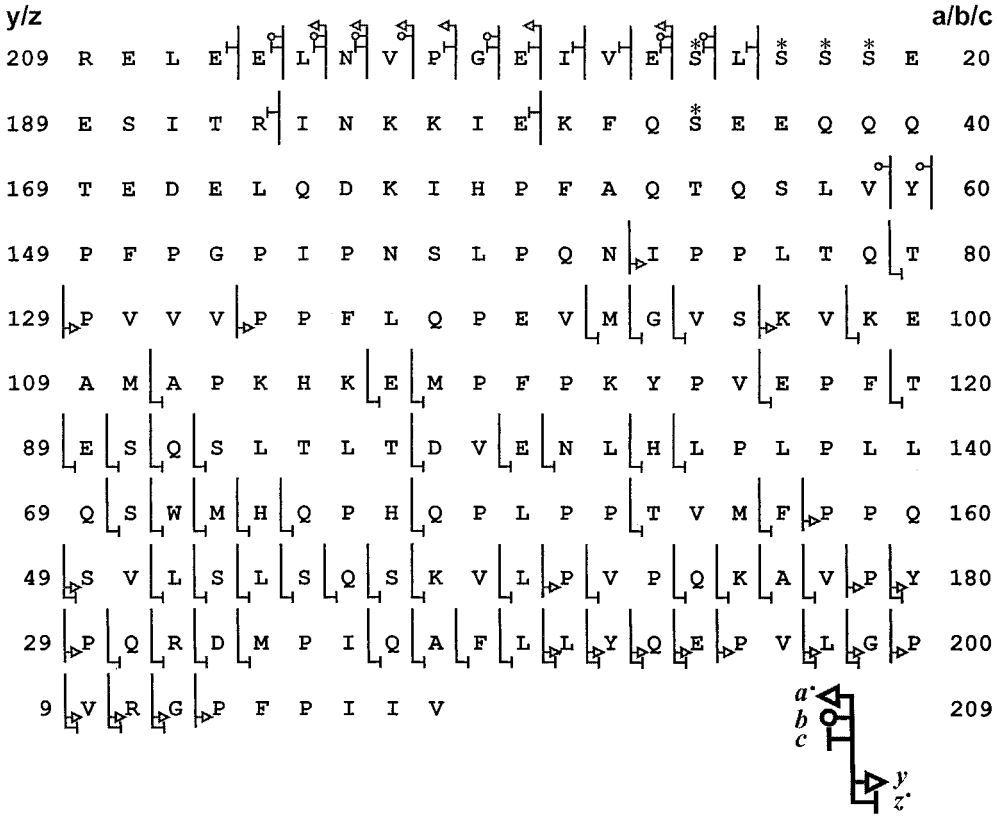

Figure 6. Fragmentation map of AI ECD of β-casein, derived by THRASH from the Figure 5 spectrum.

negative phosphoryl groups with other parts of the protein structure. However, signal/noise levels were also substantially reduced by the extensive heterogeneity.

Despite these difficulties, AI ECD fragmented 87 of the 208 backbone bonds (42% sequence coverage). Extensive cleavage near both carbon and nitrogen termini (especially the C-terminus) was observed. However, the central region yields few fragment ions, a common phenomenon in AI ECD that can be due to insufficient ion activation in this region. Also, multiple electron-capture/ion-cleavage processes were used to enhance the yield of smaller products; further fragmentation of each primary ion can occur at many bonds, so that individual internal ions are of far lower abundance.[8]

No phosphate loss is observed for any of the fragment ions generated in the ECD experiment of β-casein. With the sequence known, these fragment ions clearly identify phosphorylation at Ser-15. They also indicate that there are three more phosphorylation sites among the cluster of four serines (Ser-17, -18, -19, and -22) and one threonine (Thr-24), and one more between Lys-32 and Val-59, among two serines (Ser-35 and -57) and two threonines (Thr-41 and -55).

Other methods for ion activation effective for ECD of unmodified proteins,[8] such as laser infrared multiphoton dissociation[15] and blackbody infrared dissociation,[16] could provide more extensive fragmentation and thus more detailed information on post-translational modifications. Note that if the sequence had not been known, CAD fragment ions that both have not, and have, lost H_3PO_4 would be valuable in restricting sites of phosphorylation.[3c,a] In none of the ECD spectra of these phosphorylated peptides was a corresponding nonphosphorylated peptide found; thus, ECD should be valuable for quantitative determination of the degree of phosphorylation.

CONCLUSIONS

For the MS/MS characterization of posttranslational phosphorylation, ECD minimizes phosphate loss and therefore simplifies data interpretation. In general, ECD provides more sequence data for phosphopeptide mapping than do CAD and other MS-based methods, but these can supply complementary sequence data. ECD has provided the first direct characterization of a phosphoprotein, restricting the location of five phosphorylation sites in a 24-kDa protein without the need for enzymatic digestion and subsequent chromatographic separation. Thus, ECD appears to have a broad capability for the "top down" characterization[11] of the major types[4c,5,6] of posttranslational modifications in proteins.

ACKNOWLEDGMENT

The authors thank Michael J. Huddleston and Susan L. Chen for samples along with Roland S. Annan, Kathrin Breuker, Blas Cerda, Barry Carpenter, and Ying Ge for helpful discussions, Roman Zubarev for communication of phosphopeptide spectra before publication, and the National Institutes of Health (Grant GM16609) for generous financial support.

Received for review June 16, 2000. Accepted October 19, 2000.

AC000703Z

(15) Little, D. P., Speir, J. P.; O'Connor, P. B.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 2809−2815.

(16) Price, W. D.; Schnier, P. D.; Williams, E. R. *Anal. Chem.* **1996**, *68*, 859−866.