Journal Pre-proof

Are Internal Fragments Observable in Electron Based Top-Down Mass Spectrometry?

Neven N. Mikawy, Carolina Rojas Ramírez, Steven A. DeFiglia, Carson W. Szot, Jessie Le, Carter Lantz, Benqian Wei, Muhammad A. Zenaidee, Greg T. Blakney, Alexey I. Nesvizhskii, Joseph A. Loo, Brandon T. Ruotolo, Jeffrey Shabanowitz, Lissa C. Anderson, Kristina Håkansson

PII: S1535-9476(24)00104-X

DOI: https://doi.org/10.1016/j.mcpro.2024.100814

Reference: MCPRO 100814

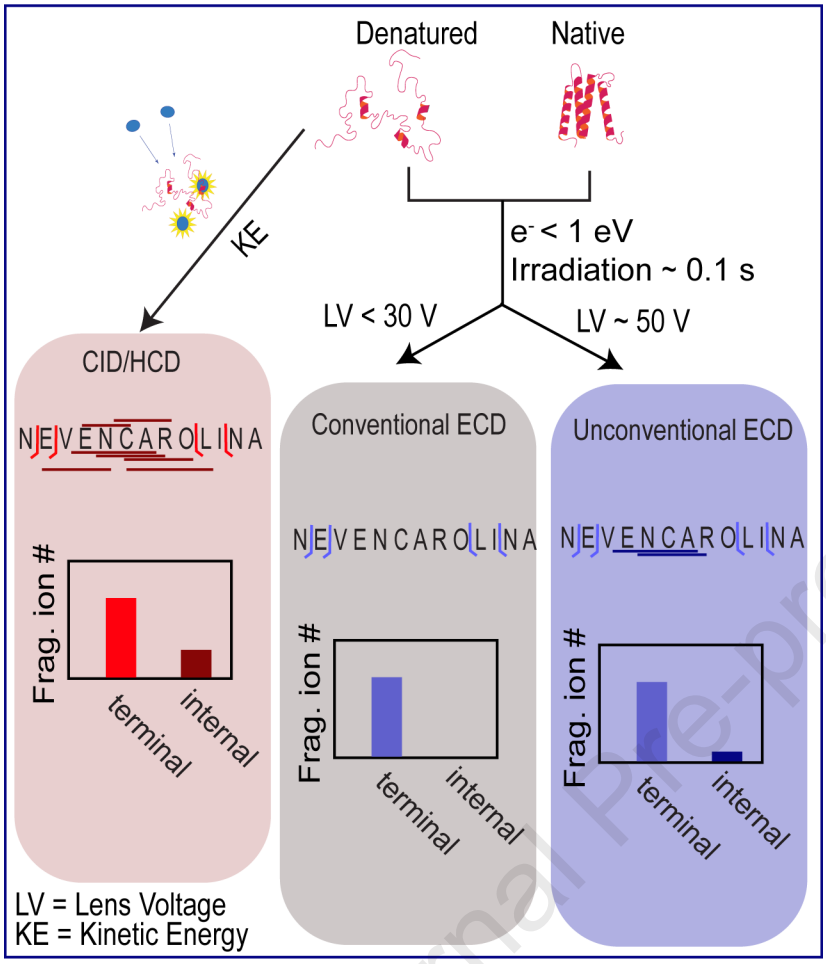
To appear in: *Molecular & Cellular Proteomics*

Received Date: 26 April 2024

Revised Date: 26 June 2024

Accepted Date: 12 July 2024

Please cite this article as: Mikawy NN, Rojas Ramírez C, DeFiglia SA, Szot CW, Le J, Lantz C, Wei B, Zenaidee MA, Blakney GT, Nesvizhskii AI, Loo JA, Ruotolo BT, Shabanowitz J, Anderson LC, Håkansson K, Are Internal Fragments Observable in Electron Based Top-Down Mass Spectrometry?, *Molecular & Cellular Proteomics* (2024), doi: https://doi.org/10.1016/j.mcpro.2024.100814.



Are Internal Fragments Observable in Electron Based Top-Down Mass Spectrometry?

Neven N. Mikawy,[1,2] Carolina Rojas Ramírez,[1,3] Steven A. DeFiglia,[1] Carson W. Szot,[1] Jessie Le,[4] Carter Lantz, [4] Benqian Wei, [4] Muhammad A. Zenaidee, [5] Greg T. Blakney,[6] Alexey I. Nesvizhskii,[3,7] Joseph A. Loo,[4] Brandon T. Ruotolo,[1] Jeffrey Shabanowitz,[8] Lissa C. Anderson,[6] Kristina Håkansson[1*]

[1]***Department of Chemistry, University of Michigan, Ann Arbor, MI 48109-1055, United States***

[2]***Department of Pharmaceutical Analytical Chemistry, Faculty of Pharmacy, Ain-Shams University, Cairo, Egypt***

[3]***Department of Pathology, University of Michigan, Ann Arbor, MI 48109-5602, United States***

[4]***Department of Chemistry and Biochemistry, University of California-Los Angeles, Los Angeles, CA 90095, United States***

[5]***Australian Proteome Analysis Facility, Macquarie University, Sydney, NSW, 2109, Australia***

[6]***National High Magnetic Field Laboratory, Florida State University, Tallahassee, FL 32310, United States***

[7]***Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI 48109-2218, United States***

[8]***Department of Chemistry, University of Virginia, Charlottesville, VA 22904, United States***

*****Correspondence: kicki@umich.edu

Running title: Internal Fragments are Elusive in Top-Down ECD/ETD

Abbreviations – The abbreviations used are: AmAc, ammonium acetate; AGC, automated gain control; CID, collision induced dissociation; ClipsMS, Comprehensive Localization of Internal Protein Sequences; dTD, denaturing top-down; ECD, electron capture dissociation; EInD, electron induced dissociation; EIoD, electron ionization dissociation; ESI, electrospray ionization; ETD, electron transfer dissociation; FT-ICR, Fourier-transform ion cyclotron resonance; HCD, higher energy collision dissociation; HESI, heated ESI; IM, ion mobility; LC, liquid chromatography;

m/z, mass-to-charge ratio; MS, mass spectrometry; MS/MS, tandem mass spectrometry; millisecond; nESI, nanoelectrospray ionization; nTD, native top-down; PTMs, posttranslational modifications; PTR, proton transfer reaction; RDD, radical-directed dissociation; S/N, signal-to-noise ratio; T, tesla; V, voltage

ABSTRACT:

Protein tandem mass spectrometry (MS/MS) often generates sequence-informative fragments from backbone bond cleavages near the termini. This lack of fragmentation in the protein interior is particularly apparent in native top-down MS. Improved sequence coverage, critical for reliable annotation of posttranslational modifications (PTMs) and sequence variants, may be obtained from internal fragments generated by multiple backbone cleavage events. However, internal fragment assignments can be error prone due to isomeric/isobaric fragments from different parts of a protein sequence. Also, internal fragment generation propensity depends on the chosen MS/MS activation strategy. Here, we examine internal fragment formation in electron capture dissociation (ECD) and electron transfer dissociation (ETD) following native and denaturing MS, as well as liquid chromatography (LC)/MS of several proteins. Experiments were undertaken on multiple instruments, including Q-ToF, Orbitrap, and high-field FT-ICR across four laboratories. ECD was performed at both ultrahigh vacuum and at similar pressure to ETD conditions. Two complementary software packages were used for data analysis. When feasible, ETD-higher-energy collision dissociation (ETD-HCD) MS^3 was performed to validate/refute potential internal fragment assignments, including differentiating MS^3 fragmentation behavior of radical vs. even-electron primary fragments. We show that, under typical operating conditions, internal fragments cannot be confidently assigned in ECD, nor ETD. On the other hand, such fragments, along with some *b*-type terminal fragments (not typically observed in ECD/ETD spectra) appear at atypical ECD operating conditions, suggesting they originate from a separate ion-electron activation process. Furthermore, atypical fragment ion types, e.g., *x* ions, are observed at such conditions as well as upon EThcD, presumably due to vibrational activation of radical *z*-type ions.

INTRODUCTION

In top-down (tandem) mass spectrometry (MS/MS) (1, 2, 3), intact proteins are transferred into the gas phase, typically via electrospray ionization (4, 5) to yield a distribution of multiply charged ions. These multiply charged proteins are then activated to generate sequence-informative fragments from backbone bond cleavages. N-terminus-containing fragments are referred to as *a*, *b*, and *c*-type ions which correspond to cleavage of inter-residue C_α-C, C(=O)-N, and N-C_α bonds, respectively. The complementary C-terminus-containing fragments are referred to as *x*, *y*, and *z*-type ions (6, 7). One drawback of protein MS/MS is that, often (and depending on the activation/dissociation method employed), mainly backbone bond cleavages close to the protein termini are matched to the protein sequence, thus limiting sequence coverage. However, recent work has proposed to also include assignments of fragments originating from multiple cleavage events, i.e., internal fragments (8). Consideration of such internal fragments has been demonstrated to significantly enhance sequence coverage from collision induced dissociation (CID) (9). However, Julian and co-workers showed that terminal fragments are heavily favored, independent of protein size (10). Electron-based dissociation methods such as electron capture dissociation (ECD) (11) and electron ionization dissociation (EIoD) (12, 13) have also been shown to generate internal fragments under certain operating conditions (14, 15). Note that the acronym (EIoD) was introduced by Baba *et al.* (13) to differentiate electron ionization dissociation from electron induced dissociation, which proceeds through a different mechanism not necessarily involving ionization (16, 17). However, internal fragment assignments can be error prone due to isomeric/isobaric fragments from different parts of a protein sequence (9). The ambiguity of assigning internal fragments scales significantly as the size of the protein increases, and the number of posttranslational modifications (PTMs) increases. Thus false discovery rates increase in a manner similar to the consideration of additional backbone fragment ion types (18) or additional proteoforms (19).

While EIoD involves sufficiently high electron energies to generate charge-increased, oxidized species from multiply charged peptide and protein cations along with rich fragmentation chemistry (12, 20), ECD involves low-energy electron capture to yield charge-reduced radical cations (11). The latter charge-reduced species preferentially fragment at N-C_α backbone bonds to generate even-electron c' and radical $z\bullet$ product ions (Zubarev nomenclature (21)) along with a less prominent fragmentation pathway yielding less abundant $a^\bullet$ and y'-type ions (22). In addition,

hydrogen atom migration between complementary *c′*/*z*• fragment ion pairs to yield *c*•/*z′*-type ions is also common, particularly for more compact precursor ions containing intramolecular hydrogen bonds or salt bridges (23). ECD has limited fragmentation efficiency because fragment ions can continue capturing electrons to form low abundance secondary fragment ion signals that cannot be distinguished from noise, that are charge-reduced to m/z ratios outside the scan range, or that may be charge-neutralized (24). Also, unlike CID (25), there are no strong cleavage preferences at certain amino acids, i.e., available signal is spread over a larger number of fragmentation channels in ECD. Thus, observation of secondary dissociation to form internal fragments should be considerably less favorable compared with CID, which can proceed with near 100% fragmentation efficiency. On the other hand, ECD-type fragmentation is often desired because labile PTMs can be retained to a larger extent in fragment ions, thus improving PTM site determination (26, 27).

Top-down MS has two main implementations (28, 29); the denaturing top-down (dTD) approach and the native top-down (nTD) approach. The former implementation focuses on protein identification and sequence characterization by maximizing the number of fragment ions and cleavage sites, providing more precise PTM localization. In this approach, the precursor protein is ionized from a denaturing solution, resulting in an extended gas-phase conformation and corresponding high charge states. By contrast, the nTD approach (28, 30) relies on nanoelectrospray ionization of monomeric proteins and noncovalent assemblies, typically from near neutral pH ammonium acetate-containing solutions. This implementation can provide information on protein higher order structure by fragmenting the protein from a folded state. Lack of fragmentation in the protein interior is particularly apparent in nTD MS.

An additional challenge in top-down MS is the broad isotope distributions of large fragment ions with the concomitant decrease in the relative abundance of the monoisotopic peak as fragment ion mass increases. Here, we examine the detection and assignment of internal fragments after ECD and electron transfer dissociation (ETD) (31) following denaturing MS, liquid chromatography (LC)/MS, and native MS, including MS^3 experiments of internal fragment candidates. We also explore two different strategies for fragment ion monoisotopic peak assignment as well as two software packages for automated assignment of both terminal and internal fragment ions.

EXPERIMENTAL PROCEDURES

Experimental Design and Statistical Rationale

ECD and ETD data from four different laboratories were examined with two different software packages, designed to mine for internal fragments. The utilized instrumentation included 7 T and 15 T FT-ICR mass spectrometers, an ion mobility-Q-ToF instrument equipped with an e-MSion ExD cell, and a 21 T FT-ICR with front-end ETD. Data were shared between laboratories and analyzed by different individuals. Manual interpretation was performed by at least two individuals. All spectra are averages over multiple scans for improved statistical representation.

Materials

Melittin from honeybee venom (~2.8 kDa; M2272-1MG), bovine calmodulin (~14 kDa; SRP6310-1MG), apomyoglobin (~17 kDa; SKU A8673-1VL) from equine skeletal muscle, bovine carbonic anhydrase II (~29 kDa; SKU C2624-100MG), enolase I (~46 kDa; E6126-500UN) from baker's yeast, and ammonium acetate were purchased from Millipore Sigma. All other chemicals were obtained at LC grade from Thermo Fisher Scientific and used without further purification unless stated otherwise.

Sample Preparation and Liquid Chromatography

Calmodulin and enolase I were dissolved in water to 1 mg/mL and purified with Biospin gel filtration columns (6 kDa MWCO) and (10 kDa MWCO), respectively, into 1 M ammonium acetate (AmAc) three times followed by 200 mM AmAc three times. For direct infusion electrospray ionization (ESI) all proteins, except carbonic anhydrase, were dissolved in water:acetonitrile (50:50, v/v) with 0.1% formic acid to a concentration of 1 μM. Carbonic anhydrase was dissolved in water:methanol:acetic acid (49:49:2; v/v/v) to a concentration of 1 μM. For nanoelectrospray ionization (nESI), calmodulin was reconstituted in 50 mM AmAc. LC/ECD MS/MS of calmodulin was performed with an Agilent 1290 HPLC using an Agilent PLRP-s, 2.1 mmx50 mm, 5μm, 1000 Å stainless steel or an AdvanceBio RP-mAb SB-C8, 2.1x50 mm, 3.5 μm column with an acetonitrile:water/0.1% formic acid solvent system at a flow rate of 0.4 mL/min. The autosampler, column, and drying gas were operated at 10, 40, and 200-250 °C, respectively. Gradient elution was employed from 5 (or 20) to 60% acetonitrile over 6-7.5 min.

Mass Spectrometry

ECD-Fourier transform ion cyclotron resonance (FT-ICR) MS experiments were conducted on a 7 T Bruker SolariX instrument equipped with a ParaCell (32) (Bruker Daltonics) and a hollow dispenser cathode electron source (33). A schematic diagram of the ECD configuration is shown in Supplementary Figure 1. Direct infusion ESI and LC-MS were performed with a capillary voltage of 5 kV. For calmodulin, native-like ECD was performed via a CaptiveSpray nESI source with a voltage 1.4 kV. Quadrupole isolation of charge states of interest was performed with a 5-30 m/z window. For melittin, calmodulin, and carbonic anhydrase II ECD, the irradiation time was 0.01-0.15 s, with a lens voltage of 15-20 V and a bias voltage of – 0.1-0.3 V. For apomyoglobin, +15 and +16 charge states, the lens voltage was 15 V and the bias voltage was - 0.1 V. For enolase I, the irradiation time was 0.02 s, with a lens voltage of 15 V and a bias voltage of - 0.3 V. Each spectrum was averaged over 16 scans except for enolase I, which was acquired over 32 and 64 scans, respectively.

ECD was also performed with an ion mobility (IM)-Q-ToF (Agilent 6560c) mass spectrometer, equipped with an e-MSion ExD cell (34). Direct infusion ESI of melittin and apomyoglobin was performed with a Jet Stream ion source operating at 2.5-3.5 kV, 325 °C with a sheath gas temperature of 275-350 °C. Quadrupole isolation was performed with the wide window setting. ECD was acquired for 3 and 60 min with a collision gas pressure of 28 and 25 psi for melittin and apomyoglobin, respectively. The ECD heater was at 2.5 A. For calmodulin native MS, 10 µM in 200 AmAc was introduced via a nESI source with gold-coated borosilicate emitters at a capillary voltage of 1400 V under ambient temperature. Sulfur hexafluoride was used as drying gas with a flow rate of 2 L/min at 25 °C. The front funnel and trap funnel were operated at 4−4.5 torr, while the drift tube with <18.5 V/cm was operated at 3.95 torr under high purity nitrogen. Transmission was tuned in 'Extended Mass Range' and 'Sensitivity Mode'. The instrument was mass calibrated using Agilent Tune Mix sprayed with the Agilent Jet Stream Source. Broadband ECD was performed without quadrupole isolation with an acquisition time of at least 10 min.

ETD-higher energy collision dissociation (HCD) MS^3 and EThcD experiments were conducted on an Orbitrap Fusion Lumos (Thermo Fisher Scientific). Proteins were infused via a heated ESI (HESI) ion source, operating at 3,800 V, with a sweep gas of 3 arbitrary unit, an ion transfer tube temperature of 280 ˚C, and a flow rate of 5 µl/min. Quadrupole isolation window, maximum ETD injection time, and normalized HCD energy were 1.5-2 m/z, 50 ms, and 7 or 42% for melittin and 2-5 m/z, 40 ms, and 15-33% for calmodulin, respectively. For EThcD, the

normalized HCD energy was 25-50%. ETD and EThcD spectra were acquired using the calibrated charge-dependent ETD parameters with a normalized automated gain control (AGC) target of 100%. Precursor ion isolation for MS^3 was performed in the linear ion trap with a 2-10 m/z window. Detection was performed in the Orbitrap with 120K or 500K resolution at 200 m/z, maximum injection time = 2000 ms and normalized AGC = 100%.

Data Analysis

The Solarix data was deisotoped with the SNAP 2.0 algorithm at a signal-to-noise (S/N) ratio of 3 in the Bruker DataAnalysis software and internally calibrated with five relatively abundant, confidently assigned fragment ions. Resulting lists of fragment ion m/z values, charge states, and abundances were transferred to Microsoft Excel and saved as .csv files. Also, ASCII files were generated directly from the Bruker DataAnalysis software. For Agilent data, Agilent MassHunter Qualitative Analysis Navigator B.08.00 was used to generate m/z vs. intensity files. For native ECD data using IM separation, IM-MS Browser 10.0 was used along with mMass v. 5.5.0 for manual peak picking. Orbitrap data were manually annotated with FreeStyleTM 1.8 SP2 software. Theoretical fragment ion masses were computed with ProteinProspector (https://prospector.ucsf.edu/). Sequence coverage maps were generated using a custom in-house script.

Two software packages were used to mine all spectra for both terminal and internal fragment ions: Comprehensive Localization of Internal Protein Sequences (ClipsMS (14)) and Fragariyo (35). For determination of theoretical internal fragment ion m/z ratios, ClipsMS does not differentiate between potential *a*-*x*, *b*-*y*, and *c*-*z*-type ions but allows for addition or subtraction of hydrogen atoms by including H• as an unlocalized modification in a separate analysis. By contrast, Fragariyo uses unique masses for *b*-*y* vs. *c*-*z* internal fragments as *b*-*y* ions should be even-electron species whereas *c*-*z* ions should be radical species, thus differing in mass by ~1 Da (Supplementary Figure 2). For Fragariyo analysis .csv files were uploaded and fragments were assigned with an error tolerance of 10 ppm. For SolariX data, ASCII files were also uploaded for the internal fragment search. For ClipsMS analysis, fragment ion lists must first be deconvolved to the corresponding singly-charged m/z values. Such deconvolution, when required (e.g., following SNAP deisotoping), was performed in Excel and saved as .csv files, which were uploaded to ClipsMS. The error was set to 10 ppm for terminal fragments and 5 ppm for internal

fragments, with the smallest internal fragment size set at 5 amino acid residues. Protein sequences were derived from UniProt: P01501 (melittin, residues 43-69; C-terminal amidation); P62157 (calmodulin, residues 2-149, N-terminal acetylation and lysine 116 trimethylation); P68082 (myoglobin, residues 2-154), P00921 (carbonic anhydrase II, residues 2-260, N-terminal acetylation), and P00924 (enolase I, residues 2-437). As needed, fixed (localized) protein modifications were included in the Fragariyo and ClipsMS input parameters. Biased search, which preferentially annotates terminal fragments over internal fragments was used. A highly complex ETD-FT-ICR spectrum, acquired at 21 T (36), was deisotoped and deconvolved with the Xcalibur™ QualBrowser-embedded 'Xtract' algorithm (Thermo Fisher Scientific) using default parameters and an S/N ratio threshold of 5 to generate monoisotopic $[M + H]^+$ values. Artifactual signals assigned to a charge state of zero were removed before Fragariyo/ClipsMS analysis of these data.

The 21 T ETD-FT-ICR data were also subjected to manual interpretation, aided by custom software, 'Predator Protein Fragment Calculator'. This software breaks each fragment into its elemental composition, based on amino acid sequence plus any chemical modifications, ion type, and charge state. The neutral mass of all corresponding isotopologue masses and their abundances above a reporting threshold are then calculated. The abundance weighted m/z average and abundance of all isotopologues of the desired fragment ion are plotted and displayed in a table for comparison with raw data. Fragments were assigned with a 10 ppm mass tolerance.

RESULTS

Melittin ECD, ETD, and ETD-HCD MS^3

ECD experiments of the quadrupole-isolated melittin 4+ charge state were performed under denaturing ESI conditions in two different configurations; an e-MSion ExD cell (34) installed on an Agilent 6560c IM-Q-ToF mass spectrometer and conventional ECD in the ICR cell of a 7 T Bruker SolariX Q-FT-ICR instrument. The ECD MS/MS spectrum from the 6560c is shown in Figure 1A. Similar to previously published ECD FT-ICR MS/MS (37, 38), 100% sequence coverage is observed from c' and $z\bullet$-type fragment ions with the exception of the N-terminal side of proline (not observed due to its cyclic structure). However, a y'-type ion (y'_{13}) is observed in that position. As previously noted, y'-type ions can result from ECD; however, it is difficult to

eliminate collisional activation in the ExD cell geometry. Thus, this y'_{13} ion may result from CID. Upon further manual analysis, many *b* and even electron *a*-type ions, characteristic of CID, are observed in this ECD spectrum along with several *y′*-type ions. In addition to these expected fragment ions, a number of *w*-type fragments can be assigned. Significant *w* ion formation has previously been reported with the ExD cell (39). Following manual spectral annotation, only 13 out of 89 isotopic clusters remained unannotated. Of these 13 observed signals nine do not match with any theoretical internal fragment ion m/z values. For the remaining four isotopic clusters, one is close in mass to an internal *b*-*y* fragment; however, the corresponding mass measurement error of 14 ppm at m/z ~ 680 is too high for confident assignment. Likewise, two isotopic clusters match closely with theoretical internal *c*-*z*-type fragments. Note that such internal fragments would contain one additional hydrogen atom compared with the corresponding internal *b*-*y*-type ion (Supplementary Figure 2). However, again, the mass measurement errors for two of the potential *c*-*z* fragments are too high (16 and 17 ppm, respectively). On the other hand, one observed doubly charged isotopic cluster (out of 89) matches the internal fragment, AVLKVLTTGLPALISWIKRKRQ (observed m/z 1237.77), corresponding to melittin residues 4-25. Even if this assignment is correct, this fragment does not add sequence information as terminal fragments already provided 100% coverage. All observed isotopic clusters and their annotation are shown in Supplementary Table 1.

To avoid unintended collisional activation, ECD was also performed on an FT-ICR-MS instrument. The resulting ECD MS/MS spectrum is shown in Supplementary Figure 3. No signals identified following SNAP deisotoping could be assigned as internal fragments; however, the AVLKVLTTGLPALISWIKRKRQ potential internal *c*-*z* fragment is observed at low abundance (m/z 1237.78) upon manual inspection of the spectrum. The corresponding signal (of too low quality to be identified by SNAP) matches this internal fragment within 1.1 ppm. However, its low abundance precludes confirmation via, e.g., an MS^3 experiment. For comparison, a melittin ETD spectrum (Supplementary Figure 4) was also acquired on an Orbitrap Fusion Lumos (Thermo Fisher Scientific) instrument. No internal fragments were noted. However, because internal *c*-*z*-type fragments should be radical ions, we were curious how such ions would behave upon MS^3. Thus, we performed ETD-HCD MS^3 of one even-electron *c′*-type ion ($c'_{25}{}^{2+}$) and one radical *z•*-type ion ($z_{15}{}^{2+\bullet}$). The HCD MS^3 spectrum of $c'_{25}{}^{2+}$ (Figure 1B) shows typical mobile proton-driven fragmentation, resulting in *a*, *b*, and *y′*-type fragments along with associated ammonia loss. By

contrast, HCD MS^3 of the radical $z_{15}^{2+\bullet}$ showed a mixture of *a/b/y′*-type fragments as well as *c′/c•/z•/z′/x/x•* ions, typical of radical-directed dissociation (RDD) (40) (Figure 1C). Similar results have been reported by McLuckey and co-workers (41).

Calmodulin ECD from Native Solvent, LC-ECD, and ETD-HCD MS^3

Previous ECD experiments following native FT-ICR MS of proteins have shown that ECD fragments appear from unstructured protein regions whereas folded regions are refractory to ECD (42, 43). We performed ECD of native calmodulin following nanoESI from 200 mM AmAc solvent on the Agilent 6560c instrument. Due to low signal and low ECD efficiency, these data were collected in broadband mode, i.e., all observed charge states (6+ to 9+) were fragmented together. Calcium binding was insignificant in these experiments. The resulting ECD spectrum (Supplementary Figure 5) was subjected to automated analysis with the Fragariyo software against the bovine calmodulin UniProt sequence. This analysis annotated a series of *z•*-type ions from the calmodulin C-terminus as well as three potential internal fragments at m/z 1336.65 (2+), 1490.68 (1+), and 1759.81 (1+). The larger, doubly charged fragment matches the internal *b-y*-type fragment NGYISAAELRHVMTNLGEKLTDEE (residues 98-121) within 2.1 ppm. However, it also matches two isomeric internal *b-y* fragments from a very different protein region: **E**QIAEFKEAFSLFDKDGDGTITTK (residues 8-31) or QIAEFKEAFSLFDKDGDGTITTK**E** (residues 9-32) within 6.1 ppm. The m/z 1490.68 (1+) fragment matches the two internal, isomeric *c-z* fragments **L**GEKLTDEEVDEM (residues 113-125) and GEKLTDEEVDEM**I** (residues 114-126) within <0.5 ppm. While *b-y*-type internal fragments appear unlikely from ECD, perhaps hydrogen atom loss could result in a mass matching such internal fragments. The final, m/z 1759.81 (1+) ion matches the internal *b-y*-type fragment LTDEEVDEMIREADI (residues 117-131) with an error of 2.6 ppm. After adding the known calmodulin N-terminal acetylation, *c′* fragments were also annotated (Fig. S5). Specifically, the m/z 1336.65 (2+) ion also matches the N-terminally acetylated, terminal c'_{23} fragment within 7.5 ppm. Despite the three internal fragment candidates matching with lower error, the N-terminally acetylated, terminal c'_{23} annotation is more likely. Specifically, ClipsMS has a biased search version that assigns terminal fragments over internal ones when there is ambiguity.

Because the Q-ToF ECD experiment did not employ quadrupole isolation (i.e., significant chemical noise is likely present) and to further assess the identity of the remaining annotated

internal fragments, nano-ESI with a CaptiveSpray source was performed on the 7 T SolariX FT-ICR instrument. As this mass spectrum (Supplementary Figure 6A) showed a bimodal charge state distribution, we will refer to this analysis as "native-like". ECD of the 9+ charge state, which was also abundant in the Q-ToF native MS experiment, was analyzed with both Fragariyo (Supplementary Table 2) and ClipsMS (Supplementary Table 3), resulting in a sequence coverage of 20% (Figure 2A and Supplementary Figure 6B), based on terminal fragment ions. Neither software annotated any internal fragments for these data with the previously observed fragment at m/z 1336.64 (2+) annotated as acetylated c'_{23} within 46 ppb. The 1759.80 (1+) fragment was also observed and annotated as a C-terminal z'_{15} ion. Also, while the 1490.68 (1+) fragment was not observed in the FT-ICR data, we note that it matches the calmodulin y'_{12} fragment within 9.8 ppm in the Q-ToF data, thus an alternative, terminal fragment assignment exists for this potential internal fragment as well. The lack of this fragment ion in the FT-ICR data suggests that it was formed via low level collisional activation in the ExD cell. Finally, another known calmodulin PTM, lysine 116 (UniProt sequence) trimethylation, further supports that the initial internal fragment assignments NGYISAAELRHVMTNLGE*K*LTDEE and **L**GE*K*LTDEEVDEM/GE*K*LTDEEVDEM**I** are erroneous as they contain unmodified lysine 116 (italicized).

To examine whether differences in internal fragment formation exist for native vs. denatured calmodulin, top-down LC-ECD MS/MS of the same 9+ charge state was performed on the 7 T SolariX FT-ICR instrument. This experiment provided similar fragment ion S/N ratio as the nanoESI direct infusion ECD experiments. Without including N-terminal acetylation (which changes the mass of N-terminal fragments), three internal fragment candidates are noted: two doubly charged ions (m/z 1336.7 and 1422.7) and one quadruply charged ion (m/z 1314.6) (Supplementary Figure 7). The first doubly charged fragment ion (m/z 1336.7) is the same one observed from native ECD on the Q-ToF instrument and native-like ECD on the FT-ICR. After adding the two known calmodulin PTMs (N-terminal acetylation and lysine 116 trimethylation), this ion was again reassigned as acetylated c'_{23}. The latter two ions were reannotated as an N-terminally acetylated $c'_{25}{}^{2+}$ fragment (m/z 1422.7) and a C-terminal $z_{45}{}^{4+\bullet}$ ion with K116 trimethylation (m/z 1314.6). To further validate these assignments as PTM-including terminal fragments, ETD-HCD MS^3 experiments on an Orbitrap Fusion Lumos were performed (Supplementary Figure 8). For both the doubly-charged fragment ions, HCD MS^3 spectra showed

typical *a/b/y′*-type fragments, confirming the assignments as even-electron *c′*-type ions. By contrast, similar to the melittin ETD-HCD MS^3 experiment (Fig. 1C), the quadruply charged fragment showed a mixture of mobile proton- and radical-driven dissociation (Supplementary Fig. 8C). Overall, top-down LC-ECD-MS/MS of the calmodulin 9+ charge state provided the same sequence coverage with only minor differences in observed fragment ions compared with the native-like direct infusion experiment (Figure 2B). The corresponding ClipsMS and Fragariyo analyses are shown in Supplementary Tables 4 and 5.

We hypothesized that LC-ECD MS/MS of a higher charge state may result in an improved probability of observing internal fragment ions. Thus, the 16+ calmodulin charge state was also examined (Supplementary Figure 9). Both Fragariyo and ClipsMS analysis were performed on the resulting data (Supplementary Tables 6 and 7). ClipsMS only annotated terminal fragments, whereas Fragariyo proposed several *c*-*z*-type internal fragments. We attempted to confirm the sequence of these internal fragment candidates; however, due to their low abundance, such experiments were unsuccessful. The observed sequence coverage based on only terminal fragments was significantly higher for the 16+ charge state, 61% (Figure 2C). All acquired LC-ECD spectra showed similar S/N ratio as previously published direct infusion ECD data (14).

Apomyoglobin ECD

Apomyoglobin, electrosprayed from denaturing conditions, was subjected to ECD with both the 7 T FT-ICR (15+ and 16+ charges states, Supplementary Figure 10) and the 6560c (16+ charge state only, Supplementary Figure 11) configurations. In contrast to previously published broadband (i.e., no precursor ion isolation) ECD on a 15 T FT-ICR (14), no internal fragments were annotated by Fragariyo, or an initial ClipsMS analysis from our data obtained following quadrupole isolation (Supplementary Tables 8-12). Because user settings in ClipsMS may affect annotation outcomes, data were shared between the authors' research groups, recalibrated and reannotated. In the latter analysis, four internal fragments were annotated for the myoglobin 15+ charge state: m/z 1069.57, 5801.02, 5886.07, and 2175.16 (all singly charged as prior deconvolution is required). The following additional, singly-charged, internal fragments were annotated for the 16+ charge state: m/z 4356.25, 8925.79, 1577.84, and 6025.14. The different outcomes between different users is attributed to which fragment ion types are considered as well as what mass tolerance is accepted.

For example, with ClipsMS, the addition of modifications is needed to consider $a\bullet$ ions from ECD. Furthermore, deisotoping errors, common for larger fragment ions (36), are not considered.

For the potential internal fragments annotated by ClipsMS for the 15+ charge state, we note that the myoglobin y'_9 (1+) fragment is isomeric with the proposed m/z 1069.57 internal fragment. Thus, there is an alternative, more likely explanation for this fragment. We also note that the myoglobin c'_{51} and $a_{52}\bullet$ fragments have calculated m/z values of 5800.04 (1+) and 5885.07 (1+), respectively, corresponding to an ~1 Da mass difference compared with the annotated potential internal fragments. This discrepancy may be attributed to a deisotoping error as these terminal fragment assignments may be more likely. For the 2175.16 (1+) ion we did not find any alternative, terminal fragment ion assignment; however, we note that the assignment (myoglobin residues 40-57) would be an internal *b*-*y* ion rather than a *c*-*z* ion, which should be more likely from ECD. The myoglobin $a_{40}\bullet$ (1+) fragment has a calculated m/z value of 4354.25, which is off by ~2 Da from the internal fragment assignment, thus it may be a less likely assignment. Nevertheless, the annotated internal fragment is a *b*-*y*-type rather than a *c*-*z*-type ion. For the m/z 8925.79 fragment, it matches the alternative assignment $a_{81}\bullet$ (1+) within 3.6 ppm. We did not find an alternative assignment for the m/z 1577.84 fragment; however, again it would be a *b*-*y*-type ion which is unlikely to result from electron-mediated fragmentation chemistry alone. Finally, for the m/z 6025.14 assignment, we did not see any signals from other isotopologues, thus this peak likely corresponds to electronic noise erroneously included by the SNAP algorithm.

Apomyoglobin sequence coverage was 71% and 80% for the 15+ and 16+ charge states, respectively, from the FT-ICR data (Supplementary Figure 12), and 65% from the 6560c data (Supplementary Figure 13) based on observed terminal fragment ions. The FT-ICR data were collected under typical ECD conditions (100 ms irradiation, -0.1 V cathode bias voltage, and 15 V lens voltage). We noted that the previously published 15 T FT-ICR broadband ECD data (14) were acquired with an unusually high lens voltage (50 V) and, thus, we also examined ECD with various lens voltages up to 50 V for the 15+ to 20+ charge states. The latter experiments used a bias voltage of -2 V, which is also common in “typical” ECD. The irradiation time was optimized to not deplete the precursor ion to a level below the highest abundance fragment ion. For the 15+ charge state at 9 ms irradiation/30 V lens and 8 ms irradiation/45 V lens, no internal fragments were annotated (Supplementary Table 13). However, at 7 ms/50 V lens, two potential internal *b*-*y*-type fragments were observed at low abundance. Increasing the precursor ion charge state resulted in detection of

a few potential internal *b*-*y*-type fragments at lower lens voltage: 45 V for the 16+ charge state, 30 V for the 17+ charge state, and 15 V for the 18+-20+ charge states (Supplementary Table 13). However, as expected, all potential internal fragments are of low abundance. ECD spectra for the 19+ charge state at 50 V lens voltage are shown in Figure 3A-C, including a low abundance potential internal *b*-*y* fragment (inset). Terminal *b*-type ions are scarce in these data.

Melittin High Lens Voltage ECD and EThcD

After finding that the ECD lens voltage may have a significant effect on fragmentation outcomes, we revisited the smaller polypeptide, melittin, under such ECD conditions. Figure 4 shows an ECD spectrum of the melittin 4+ charge state with 55 ms irradiation, -2 V cathode bias voltage, and a 50 V lens voltage. Notably, a plethora of *b* and potential internal *b*-*y* fragments are observed along with the expected *a*•/*y*′ and *c*′/*z*•-type ions. Two minor *w*-type side-chain fragments observed following conventional ECD showed significantly higher abundance with high lens voltage and four additional *w*-type fragments were observed under the latter conditions (Supplementary Table 14 for a complete list). By contrast, ECD of melittin 4+ under more typical conditions (150 ms irradiation, - 0.3 V cathode bias voltage, and 15 V lens voltage, Supplementary Figure 3) showed only one minor *b*-type ion and no internal fragments. Ions observed from both typical ECD conditions and ECD with 50 V lens voltage are highlighted with asterisks in Table S14. These ions include c′, z•, z′, *a*•, and y′-type terminal fragments as well as two side-chain *w*-type fragments. The minor *b*-type ion (b_{10}^{+}) has significantly higher (~10-fold) abundance at 50 V lens voltage.

Upon further inspection of the melittin ECD spectra at typical vs. high lens voltage we noticed that some terminal fragments showed a higher relative decrease than others at the high lens voltage, suggesting that they were subjected to preferential secondary fragmentation under such conditions. Evidence in the literature suggests that vibrational activation can occur upon low energy electron bombardment (17). To further examine this hypothesis, we compared the melittin ETD spectrum (Supplementary Figure 4) to melittin EThcD (Figure 3D) in which all ETD fragment ions are subjected to supplemental vibrational activation. As shown in the insets of Fig. 3D, the radical $z_{23}^{2+\bullet}$ fragment undergoes a significantly higher abundance decrease upon supplemental HCD compared with the even-electron c_{24}^{2+} fragment. This disparate response to collisional activation is consistent with the lower activation barrier for the radical fragment ion. In addition, O'Connor and co-workers showed that *z*•-type ions can undergo secondary charge remote

fragmentation to yield amino acid side chain losses and internal fragments in peptide ECD (44). Based on our ETD-HCD-MS^3 experiments (Fig. 1C), RDD-type fragmentation occurs for radical *z*-type ions. Thus, RDD-type fragments, including *x* ions, may be present in EThcD spectra. Upon inclusion of *x* ions in our analysis, the previously annotated internal *c*-*z* fragment at m/z 1237.78 can be reannotated as a terminal x_{21}^{2+} ion within 12 ppm in the Q-TOF data. It is difficult to assess the mass accuracy in the FT-ICR data due to poor signal quality. As noted above, supplemental collisional activation is likely in the ExD cell. Of the annotated potential *b*-*y* fragments observed from high lens voltage ECD, one fragment at m/z 1322.3 (2+) was also observed in the EThcD data (Fig. 3D). We note that an alternative assignment may be ($z_{24}^{2+\bullet}$ - NH_3). Consistent with this hypothesis, ETD-HCD MS^3 of this fragment shows several shared fragment ions with ETD-HCD MS^3 of $z_{24}^{2+\bullet}$ (Supplementary Figure 14).

Carbonic Anhydrase ECD and ETD-High Field-FT-ICR MS

Carbonic anhydrase II was electrosprayed from denaturing conditions into the 7 T FT-ICR instrument. The 34+ charge state was subjected to ECD under typical conditions (10 ms, -0.1 bias voltage, 20 V lens voltage) and the resulting spectrum (Supplementary Figure 15) was analyzed by Fragariyo and ClipsMS. No internal fragment ions were annotated (Supplementary Tables 15 and 16). The observed sequence coverage from annotated terminal fragment ions was 65% (Supplementary Figure 16). We also performed ECD with 50 V lens voltage at different electron irradiation times. At 5 ms irradiation (-0.3 V bias voltage, 50 V lens voltage, Supplementary Figure 17A), five internal *b*-*y*-type fragment ions were annotated by Fragariyo (Supplementary Table 17). At 3 ms irradiation (-1 V bias voltage, 50 V lens voltage, Supplementary Figure 17B), fewer (two) internal *b*-*y*-type fragments were observed (Supplementary Table 18).

Because ECD spectra on the 7 T FT-ICR instrument are incredibly complex for higher mass analytes, particularly with higher lens voltage, we compared the ECD data to an ETD spectrum of the same 34+ charge state acquired on a 21 T FT-ICR instrument (36). The ETD data are shown in Figure 5. The raw spectrum was manually interpreted by two coauthors, each independently confirming the others' assignments. A total of 1,239 isotopic peak clusters were identified based on comparison with isotope distributions generated by 'Predator Protein Fragment Calculator'. Terminal fragment annotations include 492 *c'*, 548 *z*•, 99 *a*•, 92 *y'*, and 8 *b* ions. No internal fragments were assigned. The identified fragments accounted for ~98% of the total ion

current in the spectrum, which yielded 91% sequence coverage of the protein (Supplementary Figure 18). Annotated mass scale expanded segments of the spectrum are shown in Supplementary Figure 19. The Xtract deconvolved data were also analyzed by Fragariyo (Supplementary Table 19) with no internal fragments annotated. By contrast, ClipsMS annotated 13 potential internal fragments (Supplementary Table 20). Four of these potential assignments have m/z values close to the precursor ion m/z value, thus they may correspond to co-isolated chemical noise. Other annotated internal fragment ions also have alternative explanations. Three fragments are off by ~1 Da from the c'_{30}, y'_{40}, and y'_{96} terminal fragments. As discussed above, such discrepancies may correspond to errors from the deisotoping algorithm. Two annotated internal fragments match with the terminal $a\bullet_{49}$ (<0.6 ppm) and $a\bullet_{102}$ (3.4 ppm) fragments, and three annotated internal fragments match with known hydrogen atom migration (23) to form $c\bullet_{53}$, z'_{87}, and z'_{97} terminal fragments.

Enolase ECD

Enolase I was electrosprayed from denaturing conditions. The 41+ charge state was subjected to ECD on the 7 T FT-ICR instrument. The resulting ECD spectrum is shown in Figure 6A. These data were analyzed by ClipsMS, which did not annotate any internal fragments (Supplementary Table 21). By contrast, Fragariyo annotated one potential internal *c*-*z*-type fragment (Fig. 6a, inset and Supplementary Table 22). However, fragment ion S/N ratio in this experiment was lower than for direct infusion ECD of smaller proteins (Supplementary Figures 6, 9, 10, 15). This decrease in S/N ratio is expected at larger molecular weight (~46 kDa for enolase) as available signal is spread over a higher number of fragmentation pathways, charge states, and isotopologues. Consequently, the observed sequence coverage (Supplementary Figure 20) is low (5%) with assigned fragments localized to the protein termini.

In order to improve enolase data quality, ECD data acquisition was lengthened from 32 to 64 scans (Figure 6B). As expected, fragment ion S/N ratio increased; however, the low abundance, potential internal fragment annotated by Fragariyo in the lower quality data (Fig. 6A) was not observed, suggesting it was not a real signal. Furthermore, even if this signal would have corresponded to a true fragment ion, there are two closely isobaric potential assignments, differing by only 4 ppm, SLMKRYPIVSIEDP (residues 285-298) and PTGAKTFAEALRIGSE (residues 174-189). These assignments again would correspond to different portions of the enolase sequence.

DISCUSSION

The data presented here show, as expected, that top-down ECD and ETD spectra are incredibly complex, particularly as protein mass increases. On the other hand, with high resolution mass analyzers, e.g., Orbitrap and FT-ICR, most of the fragment ion signals are isotopically resolved and, thus, their charge states can be directly assigned for confident annotation although some overlapping signals are observed. In such cases, PTR experiments can resolve ambiguities for overlapping isotopic distributions of different charge states (45, 46, 47). Alternatively, ion mobility spectrometry coupled with MS can add another dimension for separating such overlapping signals (48, 49). Mass accuracy is also tremendously important as, for intact proteins, many potential isobaric annotations exist. For example, isobaric potential internal fragment assignments, corresponding to different regions of the enolase sequence, were noted. For calmodulin, five potential internal fragment assignments have isobaric terminal fragments when known PTMs were taken into consideration. However, mass accuracy should not take precedence over terminal fragment ion annotation as long as both possibilities have acceptable mass measurement error. Also, because we have not confidently annotated any internal fragments from our data, the structure/mass of such ions in ECD/ETD is currently unclear, i.e., whether they would be radical or even-electron fragments, thus differing by 1 Da.

We did not observe any major difference in fragmentation behavior between native-like ECD and LC-ECD of the same charge state with the FT-ICR instrument. A direct comparison with the native Q-ToF ExD experiment was difficult due to the lack of quadrupole isolation and moderate spectral resolution. We hypothesized that higher charge states may be more likely to yield internal fragments from ECD because more electron capture events occur, thus increasing the likelihood of multiple bond cleavages. As expected, LC-ECD of a higher calmodulin charge state, 16+ vs. 9+, yielded a more complex ECD spectrum with some low abundance signals that could correspond to internal *c*-*z*-type internal fragments. However, due to the many challenges in their confident assignment, we do not believe it is advisable to include such fragments in sequence coverage analysis unless other data are available, e.g., MS^3 experiments. Similar caution has been advised for UVPD data (50). The data shown in Fig. 1B, C provide insight into how radical vs. even-electron fragments behave in collision-activated MS^3 analysis and thus could provide guidance towards interpretation of such spectra for potential internal fragment assignments. In

particular, RDD-type fragments such as *x* ions may appear with supplemental vibrational activation, e.g., at high electron flux in ECD or in EThcD. Such supplemental activation may explain the previously noted high abundance of *w*-type ions with the high pressure ExD cell from secondary fragmentation of *z*-type radical fragments.

Higher mass accuracy analysis may aid more confident internal fragment assignment; however, previously published ECD data from a 15 T FT-ICR (14) did not include quadrupole isolation and manually interpreted ETD data from a 21 T FT-ICR (Fig. 5) annotated 98% of the ion current without invoking internal fragments. The high resolving power of such instruments facilitates detection and assignment of low abundance, high-mass fragments with wider isotopic distributions. However, misassignment of the monoisotopic mass is more likely for larger ions and, thus, may introduce another source of ambiguity. Furthermore, we observed that divergent user settings in the available annotation programs can cause differences in internal vs. terminal fragment assignments, suggesting that further development is necessary. Table 1 summarizes our recommendations for terminal fragment ion types to preferentially assign over any internal fragments.

A notable finding in this study was the observation that the ECD lens voltage can have a dramatic effect on fragmentation outcomes with many additional fragments observed at 50 V. Because previous ECD data from a 15 T FT-ICR instrument (14) were generated at this unusually high lens voltage, annotation of several internal fragments is not surprising; however, it is noted that these fragments are ~1 Da lighter than expected for *c*-*z*-type internal fragments. Thus, these annotated internal fragments are not likely a result of "pure" ECD but rather other ion-electron processes, currently under further investigation. In conclusion, our experiments show no evidence that internal fragments are formed at appreciable levels from typical ECD/ETD operating conditions, i.e., at ECD lens voltages < ~ 30 V. Even under atypical conditions, they should be assigned with great caution due to their innate potential for high false discovery rates.

Acknowledgements **–** We thank John E. P. Syka, Joshua D. Hinkle, Chad R. Weisbrod, and Christopher L. Hendrickson for valuable discussions. We also acknowledge Joshua P. Salem, Hye Kyong Kweon, and Scott Daniels for invaluable technical support.

DATA AVAILABILITY

All direct infusion ECD, LC-ECD, and ETD-HCD MS^3 data are available in MassIVE (MSV000094542). The 21 T FT-ICR ETD MS/MS spectrum (.raw) of carbonic anhydrase II is included in the Supporting Information of reference 36 (Weisbrod et al.) and is available free of charge at http://doi.org/10.1007/s13361-017-1702-3. The native calmodulin IM-Q-TOF data can be accessed upon request.

Supplemental Data – this article contains supplemental data.

Funding and Additional Information

This work was supported by National Science Foundation grant CHE2004043 to K.H. and an Agilent Thought Leader Award to K. H. and B. T. R. The Thermo Scientific Orbitrap Fusion Lumos mass spectrometer was acquired via National Institutes of Health (NIH) grant S10OD021619 to K. H. C.R.R. was supported by NIH grants T32CA140044, RO1GM094231, and U24CA271037. S. A. D was partially supported by a George Ashworth Summer Fellowship. J.A.L. acknowledges support from the US National Institutes of Health (R35GM145286) and the US Department of Energy (DE-FC02-02ER63421). J.S. was supported by NIH grant GM037537. A portion of this work was performed at the Ion Cyclotron Resonance User Facility at the National High Magnetic Field Laboratory at Florida State University, supported by the National Science Foundation Divisions of Materials Research and Chemistry (DMR-1644779 & DMR-2128556) and by the State of Florida.

Author Contributions – N. N. M., C. R. R., S. A. D., C. W. S., B. T. R., L. C. A., and K. H. methodology and experiments; N. N. M., C. R. R., J. L., G. T. B., J. S., L. C. A., and K. H. data analysis; N. N. M., C. R. R., J. A. L., C. L., B. W., M. A. Z., L. C. A., and K. H. writing–editing; J. A. L., B. T. R., L. C. A., and K. H. project administration; N. N. M. and K. H. writing–original draft.

REFERENCES

1. Reid, G. E., McLuckey, S. A. (2002) 'Top Down' Protein Characterization via Tandem Mass Spectrometry. *J. Mass Spectrom.* **37,** 663-675

2. Cui, W., Rohrs, H. W., Gross, M. L. (2011) Top-Down Mass Spectrometry: Recent Developments, Applications and Perspectives. *Analyst* **136,** 3854
3. Fornelli L., Toby, T. K., Schachner, L. F., Doubleday, P. F., Srzentić, K., DeHart, C. J., Kelleher, N. L. (2018) Top-Down Proteomics: Where We Are, Where We Are Going? *J Proteomics* **175,** 3–4
4. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. (1990) Electrospray Ionization - Principles and Practice. *Mass Spectrom. Rev.* **9,** 37-70
5. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., Whitehouse, C. M. (1989) Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *Science* **246,** 64-71
6. Roepstorff, P., Fohlman, J. (1984) Proposal for a Common Nomenclature for Sequence Ions in Mass Spectra of Peptides. *Biomed. Mass Spectrom.* **11,** 601-601
7. Biemann, K. (1990) Appendix 5. Nomenclature for Peptide Fragment Ions (Positive Ions). *Methods Enzymol.* **193,** 886-887
8. Zenaidee, M. A., Wei, B., Lantz, C., Wu, H. T., Lambeth, T. R., Diedrich, J. K., Ogorzalek Loo, R. R., Julian, R. R., Loo, J. A. (2021) Internal Fragments Generated from Different Top-Down Mass Spectrometry Fragmentation Methods Extend Protein Sequence Coverage. *J. Am. Soc. Mass Spectrom.* **32,** 1752–1758
9. Schmitt, N. D., Berger, J. M., Conway, J. B., Agar, J. N. (2021) Increasing Top-Down Mass Spectrometry Sequence Coverage by an Order of Magnitude through Optimized Internal Fragment Generation and Assignment. *Anal. Chem.* **93,** 6355–6362
10. Lyon, Y. A., Riggs, D., Fornelli, L.; Compton, P. D., Julian, R. R. (2018) The Ups and Downs of Repeated Cleavage and Internal Fragment Production in Top-Down Proteomics. *J. Am. Soc. Mass Spectrom*. **29**, 150-157
11. Zubarev, R. A., Horn, D. M., Fridriksson, E. K., Kelleher, N. L., Kruger, N. A., Lewis, M. A., Carpenter, B. K., McLafferty, F. W. (2000) Electron Capture Dissociation for Structural Characterization of Multiply Charged Protein Cations. *Anal. Chem.* **72,** 563-573
12. Fung, Y. M. E., Adams, C. M., Zubarev, R. A. (2009) Electron Ionization Dissociation of Singly and Multiply Charged Peptides. *J. Am. Chem. Soc.* **131,** 9977-9985
13. Baba, T., Rajabi, K., Liu, S. Y., Ryumin, P., Zhang, Z., Pohl, K., Causon, J., Le Blanc, J. C. Y., Kurogochi, M. (2022) Electron Impact Excitation of Ions from Organics on Singly Protonated Peptides with and without Post-Translational Modifications. **33,** 1723-1732
14. Lantz, C., Zenaidee, M. A., Wei, B., Hemminger, Z., Ogorzalek Loo, R. R., Loo, J. A. (2021) ClipsMS: An Algorithm for Analyzing Internal Fragments Resulting from Top-Down Mass Spectrometry. *J. Proteome Res.* **20,** 1928−1935
15. Zenaidee, M. A., Lantz, C., Perkins, T., Jung, W., Loo, R. R. O., Loo, J. A. (2020) Internal Fragments Generated by Electron Ionization Dissociation Enhance Protein Top-Down Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **31,** 1896-1902
16. Budnik, B. A., Haselmann, K. F., Elkin, Y. N., Gorbach, V. I., Zubarev, R. A. (2003) Applications of Electron-Ion Dissociation Reactions for Analysis of Polycationic Chitooligosaccharides in Fourier Transform Mass Spectrometry. *Anal. Chem.* **75,** 5994-6001
17. Gord, J. R., Horning, S. R., Wood, J. M., Cooks, R. G., Freiser, B. S. (1993) Energy Deposition during Electron Induced Dissociation. *J. Am. Soc. Mass Spectrom.* **4,** 145-151
18. Fornelli, L., Srzentić, K., Toby, T. K., Doubleday, P. F., Huguet, R., Mullen, C., Melani, R. D., dos Santos Seckler, H., DeHart, C. J., Weisbrod, C. R., Durbin, K. R., Greer, J. B.,

Early, B. P., Fellers, R. T., Zabrouskov, V., Thomas, P. M., Compton, P. D., Kelleher, N. L. (2020) Thorough Performance Evaluation of 213 nm Ultraviolet Photodissociation for Top-down Proteomics. *Mol. Cell. Proteomics* **19,** 405-420

19. LeDuc, R. D., Fellers, R. T., Early, B. P., Greer, J. B., Shams, D. P., Thomas, P. M., Kelleher, N. L. (2019) Accurate Estimation of Context-Dependent False Discovery Rates in Top-Down Proteomics. *Mol. Cell. Proteomics* **18,** 796-805
20. Zubarev, R. A., Yang, H. (2010) Multiple Soft Ionization of Gas-Phase Proteins and Swift Backbone Dissociation in Collisions with >99 eV Electrons. *Angew. Chem. Int. Ed.* **49,** 1439 –1441
21. Kjeldsen, F., Haselmann, K., Budnik, B. A., Jensen, F., Zubarev, R. A. (2002) Dissociative Capture of Hot Electrons by Polypeptide Polycations: an Efficient Process Accompanied by Secondary Fragmentation. *Chem. Phys. Lett.* **356,** 201-206
22. Zubarev, R. A., Kruger, N. A., Fridriksson, E. K., Lewis, M. A., Horn, D. M., Carpenter, B. K., McLafferty, F. W. (1999) Electron Capture Dissociation of Gaseous Multiply-charged Proteins is Favored at Disulfide Bonds and Other Sites of High Hydrogen Atom Affinity. *J. Am. Chem. Soc.* **121,** 2857-2862
23. O'Connor, P. B., Lin, C., Cournoyer, J. J., Pittman, J. L., Belyayev, M., Budnik, B. A. (2006) Long-Lived Electron Capture Dissociation Product Ions Experience Radical Migration via Hydrogen Abstraction *J. Am. Soc. Mass Spectrom.* **17,** 576-585
24. Zubarev, R. A., Haselmann, K. F., Budnik, B., Kjeldsen, F., Jensen, F. (2002) Towards an Understanding of the Mechanism of Electron Capture Dissociation: A Historical Perspective and Modern Ideas. *Eur. Mass Spectrom.* **8,** 337-349
25. Huang, Y., Triscari, J. M., Tseng, G. C., Pasa-Tolic, L., Lipton, M. S., Smith, R. D., Wysocki, V. H. (2005) Statistical Characterization of the Charge State and Residue Dependence of Low-Energy CID Peptide Dissociation Patterns. *Anal. Chem.* **77,** 5800-5813
26. Cooper, H. J., Hakansson, K., Marshall, A. G. (2005) The Role of Electron Capture Dissociation in Biomolecular Analysis. *Mass Spectrom. Rev.* **24,** 201-222
27. Lermyte, F., Valkenborg, D., Loo, J. A., Sobott, F. (2018) Radical Solutions: Principles and Application of Electron-Based Dissociation in Mass Spectrometry-Based Analysis of Protein Structure. *Mass Spectrom. Rev.* **37,** 750-771
28. Williams, J. P., Morrison, L. J., Brown, J. M., Beckman, J. S., Voinov, V. G., Lermyte, F. (2020) Top-Down Characterization of Denatured Proteins and Native Protein Complexes Using Electron Capture Dissociation Implemented within a Modified Ion Mobility-Mass Spectrometer. *Anal. Chem.* **92,** 3674–3681
29. Lermyte, F., Tsybin, O. Y., O'Connor, P. B., Loo, J. A. (2019) Top or Middle? Up or Down? Toward a Standard Lexicon for Protein Top-Down and Allied Mass Spectrometry Approaches. *J. Am. Soc. Mass Spectrom.* **30,** 1149-1157
30. Liu, R., Xia, S., Li, H. (2023) Native Top-Down Mass Spectrometry for Higher-Order Structural Characterization of Proteins and Complexes. *Mass Spectrom. Rev.* **42,** 1876–1926
31. Syka, J. E. P., Coon, J. J., Schroeder, M. J., Shabanowitz, J., 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
32. Boldin, I. A., Nikolaev, E. N. (2011) Fourier Transform Ion Cyclotron Resonance Cell with Dynamic Harmonization of the Electric Field in the Whole Volume by Shaping of

the Excitation and Detection Electrode Assembly. *Rapid Commun. Mass Spectrom.* **25,** 122-126

33. Tsybin, Y. O., Witt, M., Baykut, G., Kjeldsen, F., Hakansson, P. (2003) Combined Infrared Multiphoton Dissociation and Electron Capture Dissociation with a Hollow Electron Beam in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **17,** 1759-1768
34. Voinov, V. G., Deinzer, M. L., Barofsky, D. F. (2008) Electron Capture Dissociation in a Linear Radiofrequency-free Magnetic cell *Rapid Commun. Mass Spectrom.* **22,** 3087-3088
35. Rojas-Ramirez, C., Murtada, R., Gao, J., Ruotolo, B. T. (2022) Free Radical-Based Sequencing for NativeTop-Down Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **33,** 2283-2290
36. Weisbrod, C. R., Kaiser, N. K., Syka, J. E. P., Early, L., Mullen, C., Dunyach, J.-J., English, A. M., Anderson, L. C., Blakney, G. T., Shabanowitz, J., Hendrickson, C. L., Marshall, A. G., Hunt, D. F. (2017) Front-End Electron Transfer Dissociation Coupled to a 21 Tesla FT-ICR Mass Spectrometer for Intact Protein Sequence Analysis. *J. Am. Soc. Mass Spectrom.* **28,** 1787-1795
37. Kjeldsen, F., Savitski, M. M., Adams, C. M., Zubarev, R. A. (2006) Determination of the Location of Positive Charges in Gas-Phase Polypeptide Polycations by Tandem Mass Spectrometry. *Int. J. Mass Spectrom.* **252,** 204-212
38. Kweon, H. K., Hakansson, K. (2006) Site-Specific Amide Hydrogen Exchange in Melittin Probed by Electron Capture Dissociation Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Analyst* **131,** 275-280
39. Shaw, J. B., Malhan, N., Vasil'ev, Y. V., Lopez, N. I., Makarov, A., Beckman, J. S., Voinov, V. G. (2018) Sequencing Grade Tandem Mass Spectrometry for Top-Down Proteomics Using Hybrid Electron Capture Dissociation Methods in a Benchtop Orbitrap Mass Spectrometer. *Anal. Chem.* **90,** 10819-10827
40. Oh, H. B., Moon, B. (2015) Radical-Driven Peptide Backbone Dissociation Tandem Mass Spectrometry. *Mass Spectrom. Rev.* **34,** 116-132
41. Han, H., Xia, Y., McLuckey, S. A. (2007) Ion Trap Collisional Activation of c and z• Ions Formed via Gas-Phase Ion/Ion Electron Transfer Dissociation. *J. Proteome Res*. **6**, 3062-3069
42. Zhang, H., Cui, W. D., Wen, J. Z., Blankenship, R. E., Gross, M. L. (2010) Native Electrospray and Electron-Capture Dissociation in FTICR Mass Spectrometry Provide Top-Down Sequencing of a Protein Component in an Intact Protein Assembly. *J. Am. Soc. Mass Spectrom.* **21,** 1966-1968
43. Yin, S., Loo, J. A. (2011) Top-Down Mass Spectrometry of Supercharged Native Protein-Ligand Complexes. *Int. J. Mass Spectrom.* **300,** 118-122
44. Li, X., Lin, C., Han, L., Costello, C. E., O'Connor, P. B. (2010) Charge Remote Fragmentation in Electron Capture and Electron Transfer Dissociation. *J. Am. Soc. Mass Spectrom.* **21,** 646-656
45. Dunham, S. D., Brodbelt, J. S. (2024) Enhancing Top-Down Analysis of Proteins by Combining Ultraviolet Photodissociation (UVPD), Proton-Transfer Charge Reduction (PTCR), and Gas-Phase Fractionation to Alleviate the Impact of Nondissociated Precursor Ions. *J. Am. Soc. Mass Spectrom.* **35,** 255-265

46. Kline, J. T., Mullen, C., Durbin, K. R., Oates, R. N., Huguet, R., Syka, J. E. P., Fornelli, L. (2021) Sequential Ion-Ion Reactions for Enhanced Gas-Phase Sequencing of Large Intact Proteins in a Tribrid Orbitrap Mass Spectrometer. *J. Am. Soc. Mass Spectrom.* **32,** 2334-2345
47. Weisbrod, C. R., Anderson, L. C., Hendrickson, C. L., Schaffer, L. V., Shortreed, M. R., Smith, L. M., Shabanowitz, J., Hunt, D. F. (2021) Advanced Strategies for Proton-Transfer Reactions Coupled with Parallel Ion Parking on a 21 T FT-ICR MS for Intact Protein Analysis. *Anal. Chem.* **93,** 9119-9128
48. Gadkari, V. V., Rojas Ramírez, C., Vallejo, D. D., Kurulugama, R. T., Fjeldsted, J. C., Ruotolo, B. T. (2020) Enhanced Collision Induced Unfolding and Electron Capture Dissociation of Native-Like Protein Ions. *Anal. Chem.* **92,** 15489-15496
49. Zinnel, N. F., Pai, P.-J., Russell, D. H. (2012) Ion Mobility-Mass Spectrometry (IM-MS) for Top-Down Proteomics: Increased Dynamic Range Affords Increased Sequence Coverage. *Anal. Chem.* **84,** 3390-3397
50. Dunham, S. D., Wei, B., Lantz, C., Loo, J. A., Brodbelt, J. S. (2022) Impact of Internal Fragments on Top-Down Analysis of Intact Proteins by 193 nm UVPD. *J. Proteome Res.* **22,** 170-181

FIGURE CAPTIONS

FIG. 1

ECD MS/MS spectrum of Melittin (4+) on an Agilent 6560c equipped with an e-MSion ExD cell (A). ETD-HCD MS^3 of an even-electron $c'_{25}{}^{+2}$ ion (B) and a radical $z_{15}{}^{2+\bullet}$ ion (C) on an Orbitrap Fusion Lumos. Fragment ions highlighted in red in (C) correspond to radical driven dissociation of the radical precursor ion.

FIG. 2

Sequence coverage of calmodulin on the SolariX Q-FT-ICR instrument. ECD of native-like calmodulin (9+) (A), LC-ECD of calmodulin (9+) (B), and LC-ECD of calmodulin (16+) (C). Fragment ions labeled in red contain the known calmodulin PTMs.

FIG. 3

ECD MS/MS spectra of apomyoglobin 19+ on the SolariX FT-ICR-MS instrument with different ECD lens voltages: 30 V (A), 45 V (B), and 50 V (C). Potential internal *b-y* fragments appear at 50 V (insets). EThcD of melittin 4+ on the Orbitrap Fusion Lumos (D) with zoomed-in view of one radical and one even-electron fragment (right inset). The radical *z* ion shows preferential secondary fragmentation as compared to conventional ETD (left inset).

FIG. 4

ECD MS/MS spectrum of melittin 4+ on the SolariX Q-FT-ICR-MS instrument with an ECD lens voltage of 50 V. Low m/z region (A), high m/z region (B). Fragments labeled in green correspond to *b*- and *w*-type ions that are virtually absent at lower lens voltage.

FIG. 5

21 T FT-ICR MS/MS spectrum of carbonic anhydrase II (34+, 29 kDa) following 6 ms ETD (36). The signal was summed over 1500 acquisitions (600,000 resolving power at m/z 400) with use of 16 fills of the multipole storage device per transient acquisition (3.2E6 cumulative ion target). The mass scale-expanded segment (A) of the spectrum (B) is shaded red.

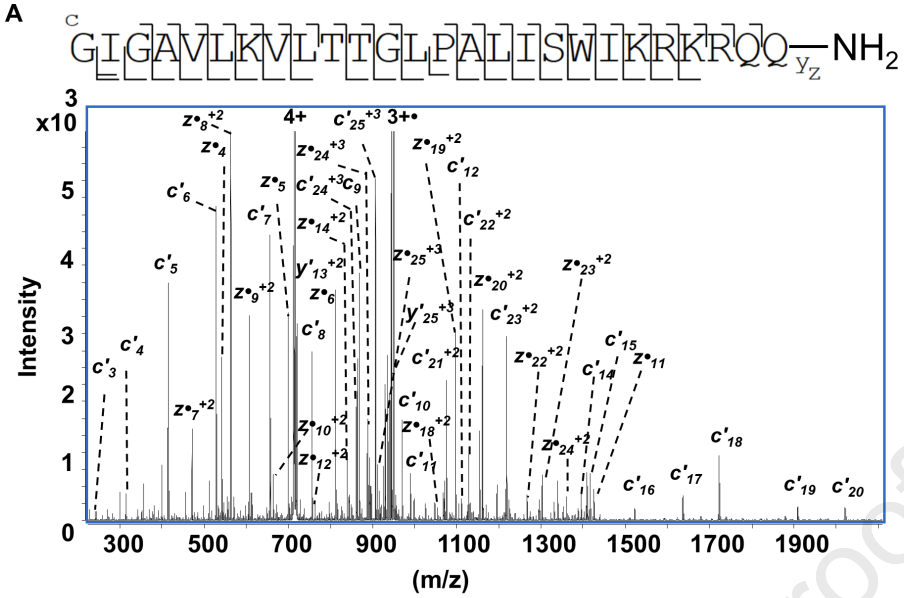
FIG. 6

ECD MS/MS spectra of enolase I, 41+ with 32 scans (A) and 64 scans (B) on the SolariX Q-FT-ICR-MS instrument. Insets show that a putative internal fragment corresponds to background noise.

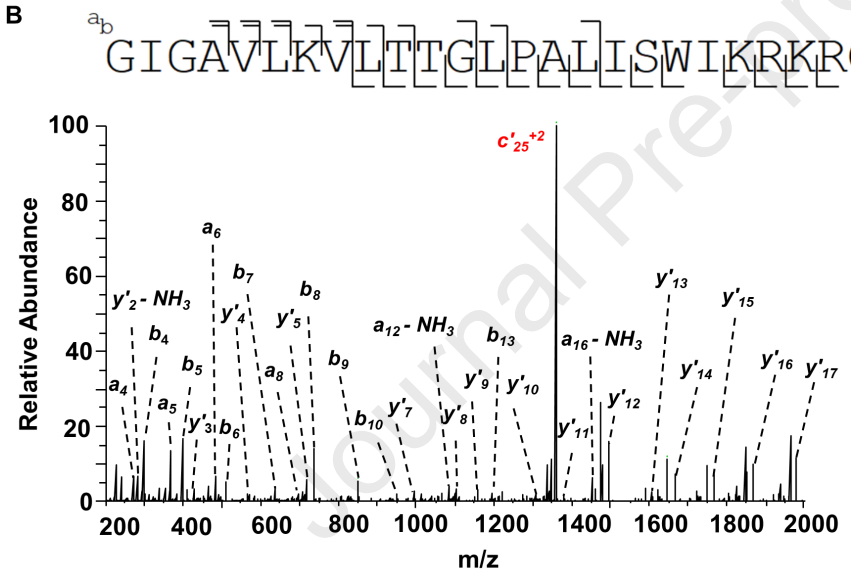
Table 1. Terminal fragment ion types expected in electron based top-down MS

Activation Technique	Expected Terminal Fragments
ECD	*c′, z•, c•, z′, a•, y′,* including NH_3 and H_2O loss
ECD with vibrational activation	*c′, z•, c•, z′, a•, y′, b, a, x, x•, d, w,* including NH_3 and H_2O loss
ETD	*c′, z•, c•, z′, a•, y′,* including NH_3 and H_2O loss
EThcD	*c′, z•, c•, z′, a•, y′, b, a, x, x•, d, w,* including NH_3 and H_2O loss

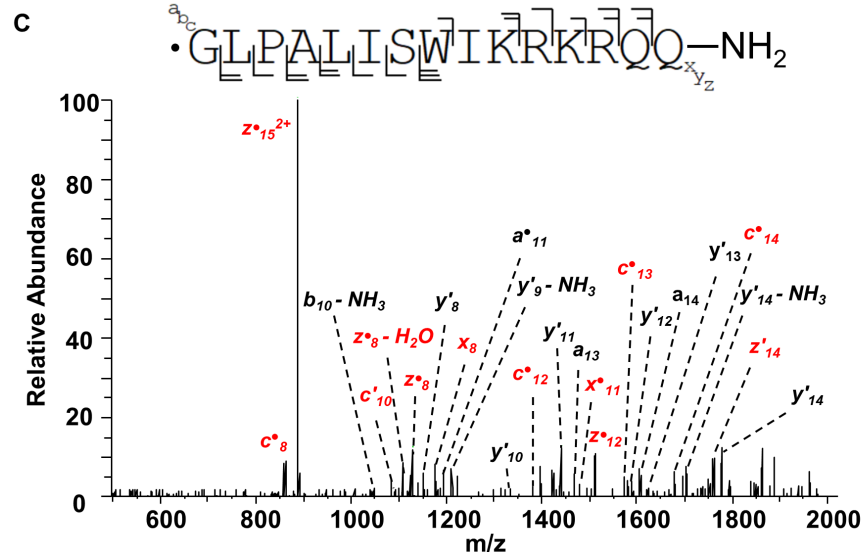
A



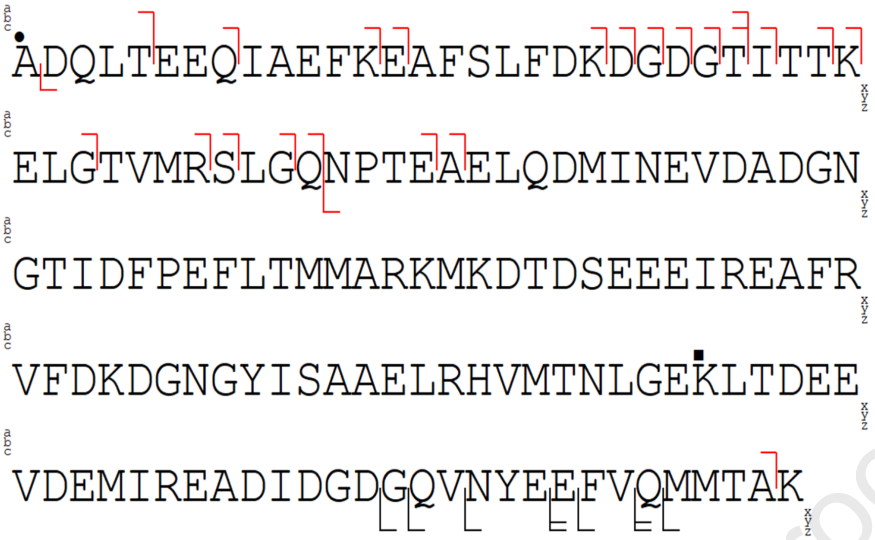
B



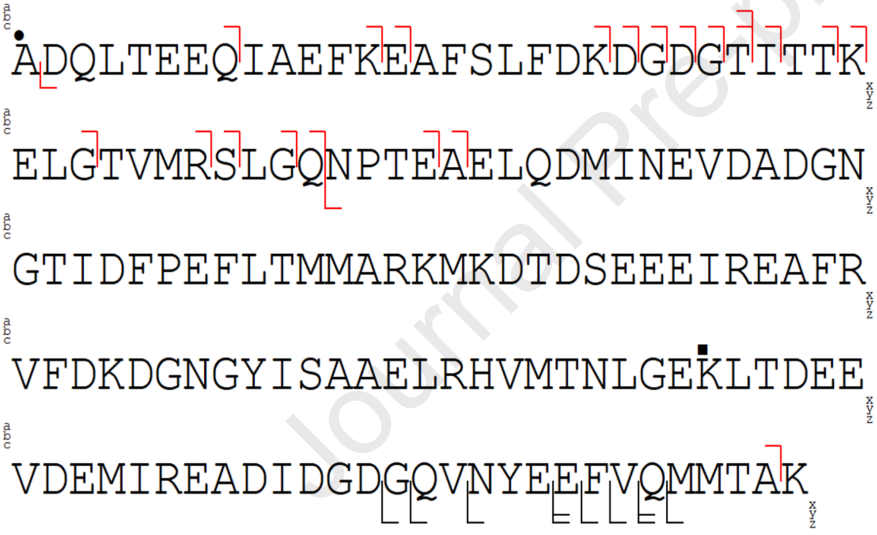
C



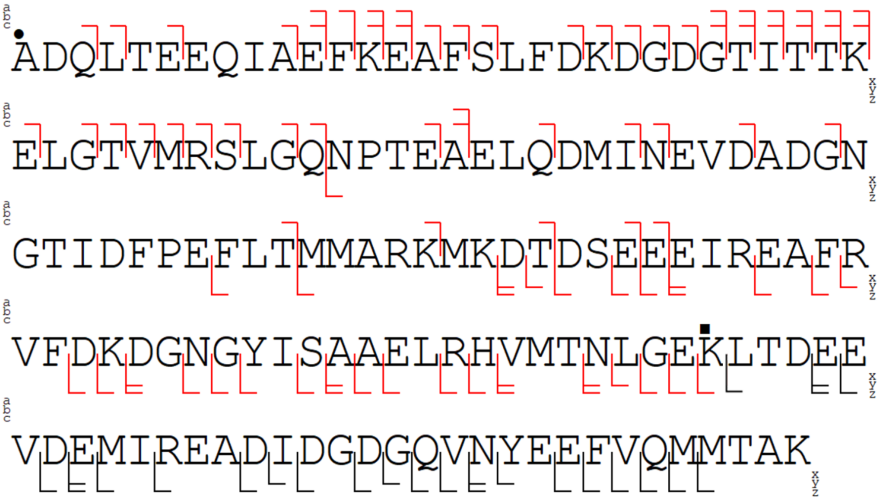
A

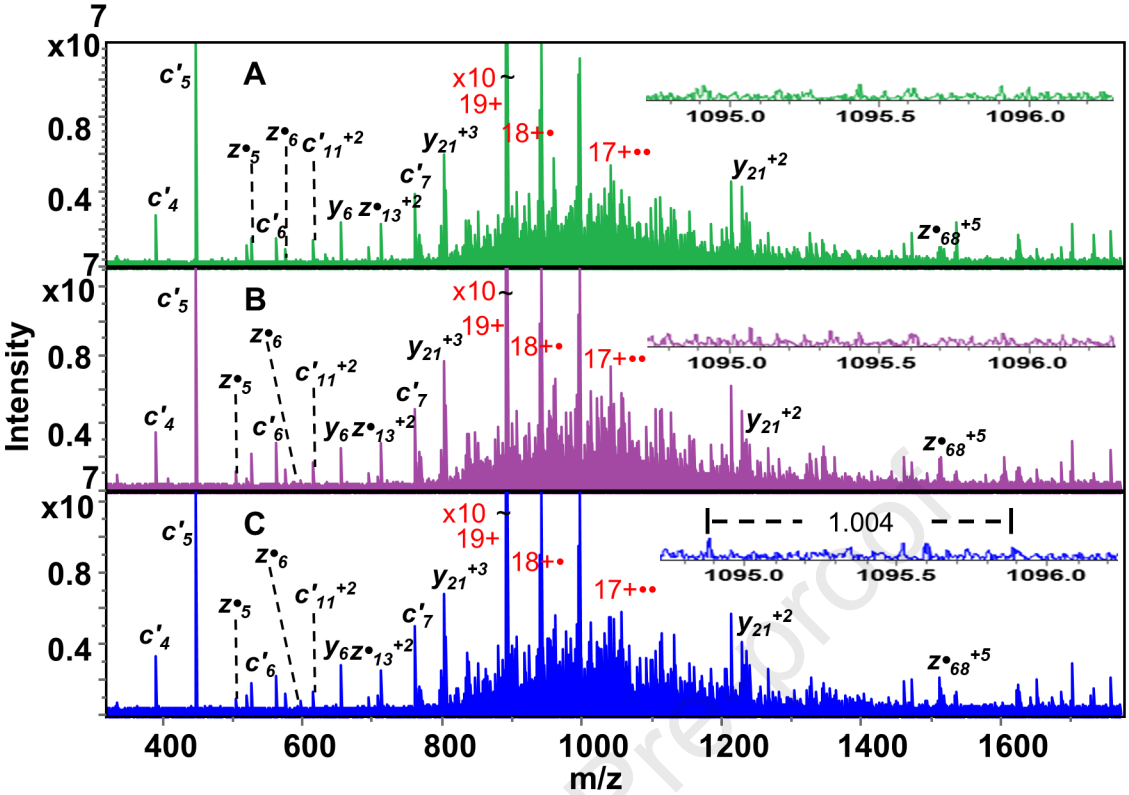


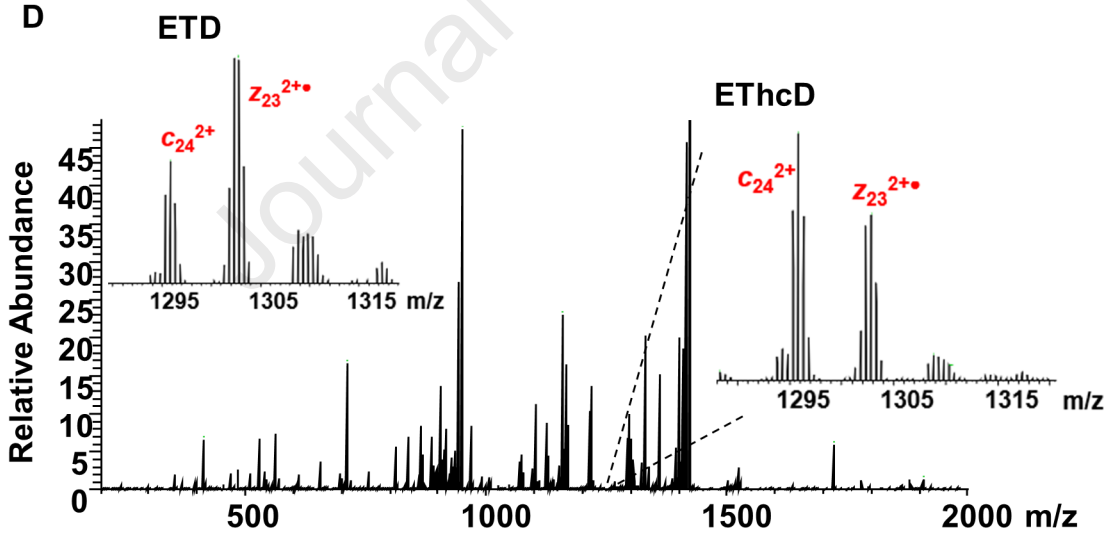
B

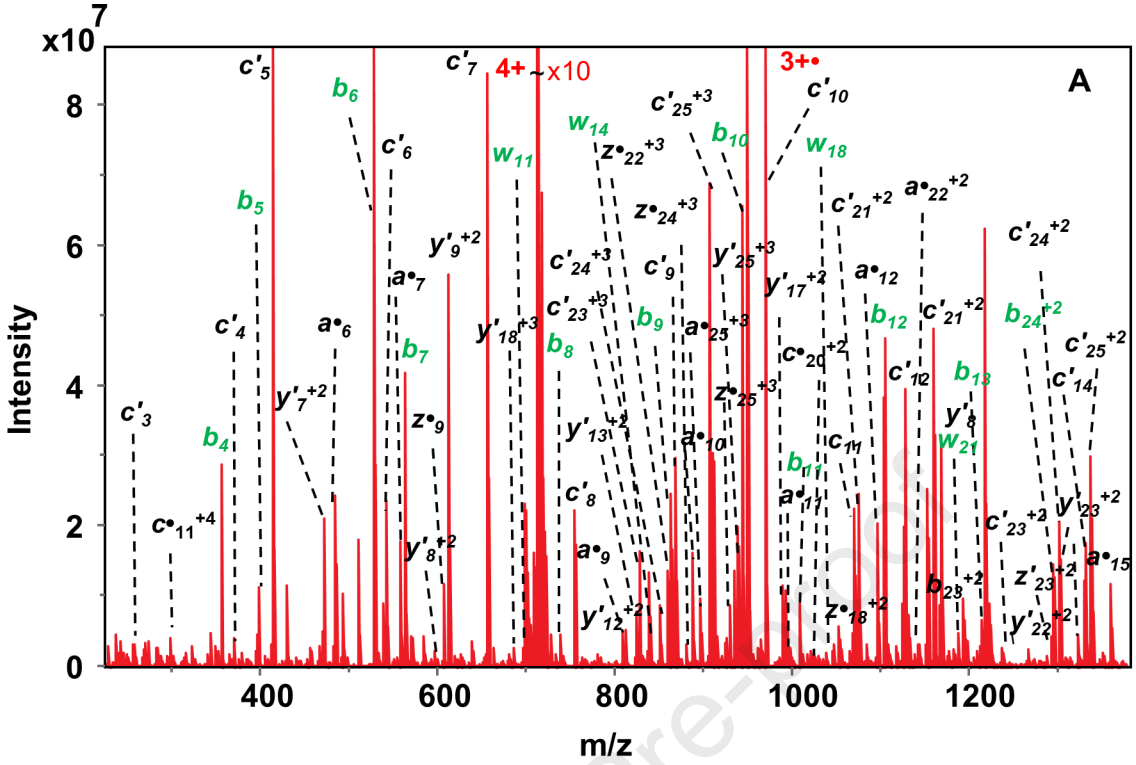


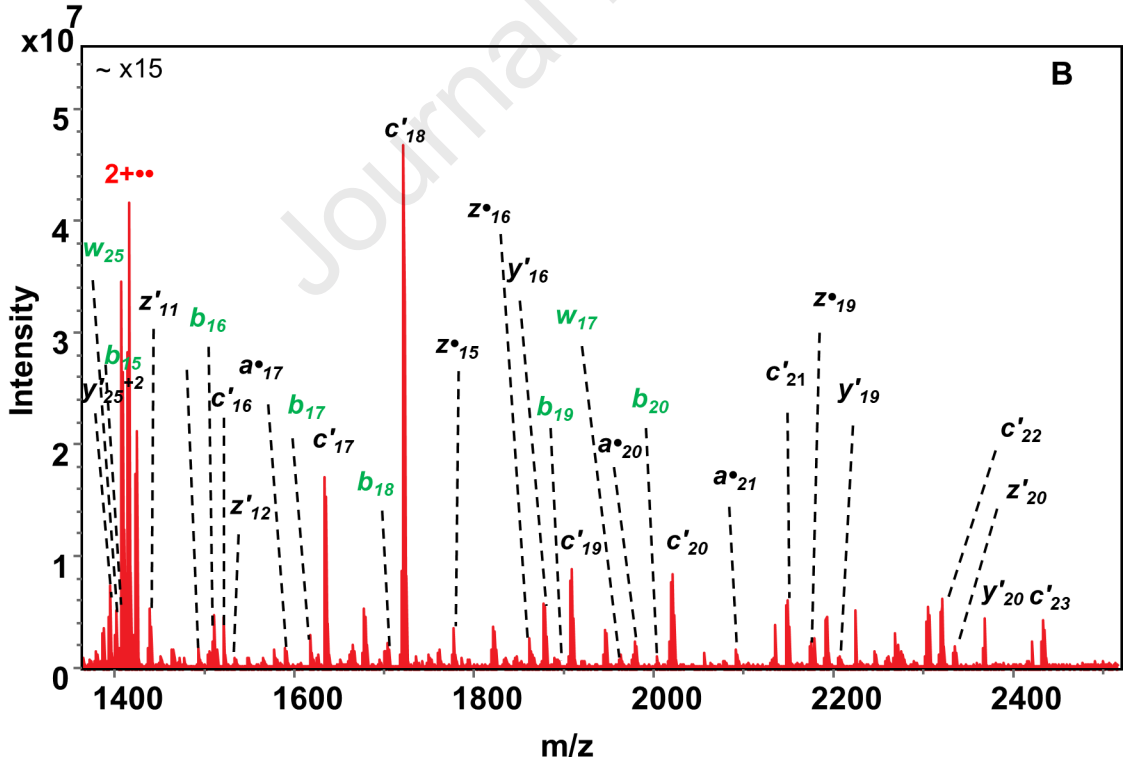
C

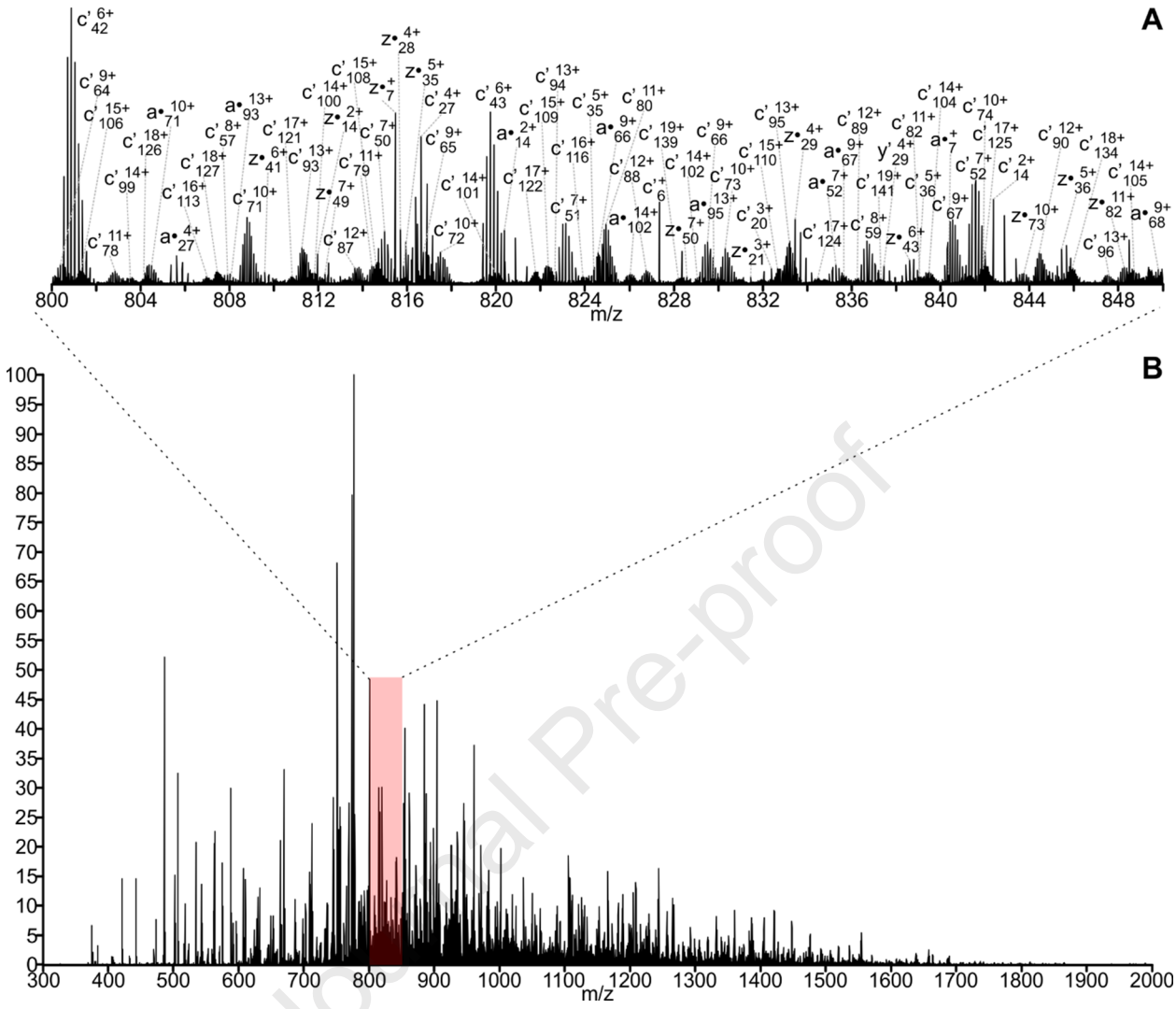


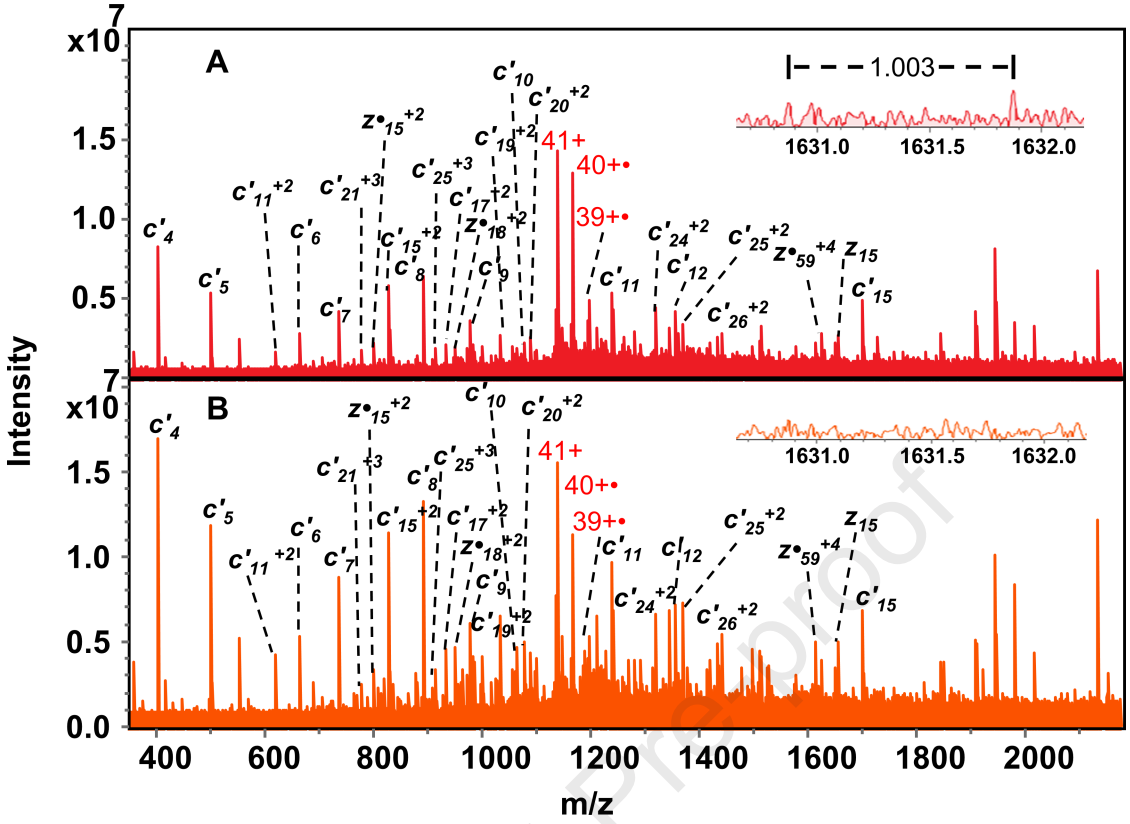












Highlights

- Internal fragments were not confidently assigned at typical operating conditions.
- ClipsMS and Fragariyo software showed differences in terminal/internal fragment annotations, presumably because different ion types were included in the searches.
- Even-electron fragments show typical *a/b/y′* fragmentation behavior, whereas radical fragments show *a/b/y* as well as *c′/c•/z/z′/x/x•* ions upon ETD-HCD-MS^3.

In Brief

Internal fragments may boost sequence coverage in top-down MS. We examined top-down electron-based MS/MS spectra from multiple instruments across four laboratories with two software packages for automated fragment ion annotation. Internal fragment candidates are virtually absent and do not show sufficient abundance for confident assignment. However, at some ECD operating parameters, such fragments are more abundant. Nevertheless, due to many isomeric/closely isobaric assignment possibilities, internal fragments should be assigned with great caution unless additional data are available, e.g., MS^3 confirmation.

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐ The author is an Editorial Board Member/Editor-in-Chief/Associate Editor/Guest Editor for *[Journal name]* and was not involved in the editorial review or the decision to publish this article.

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: