Supporting Information

for

Internal fragments generated by electron ionization dissociation enhance protein top-down mass spectrometry

Muhammad A. Zenaidee,[1] Carter Lantz,[1] Taylor Perkins,[1] Wonhyuek Jung,[1] Rachel R. Ogorzalek Loo,[2] Joseph A. Loo[1,2]

[1] Department of Chemistry and Biochemistry, University of California-Los Angeles, Los Angeles, CA 90095

[2] Department of Biological Chemistry, University of California-Los Angeles, Los Angeles, CA 90095

Corresponding Author
Joseph A. Loo
University of California-Los Angeles, Los Angeles, CA, United States
Email: jloo@chem.ucla.edu

Table of Contents - Supporting information for internal fragments generated by electron ionization dissociation enhances protein top-down mass spectrometry.

Table S1. Ion abundances of assigned fragments formed from EID of $[CAII, 25H]^{25+}$ (*m*/*z* between ~ 500-700).

Fragment	Charge	Ion abundance
c_3	2+	7.03×10^6
c_2-z_{13}	3+	3.99×10^6
c_{12}	3+	6.19×10^6
c_8	2+	4.24×10^6
c_2-z_{25}	4+	5.26×10^6
c_{127}-z_{173}	5+	3.23×10^6
z_4	3+	1.04×10^7
z_5	2+	2.17×10^7
z_9	2+	2.01×10^7
c_{15}	3+	7.46×10^6
c_{16}	3+	5.17×10^6
c_3-z_{25}	4+	3.65×10^6
z_{10}	2+	1.00×10^7
z_{13}	3+	4.16×10^6
c_{173}-z_{258}	4+	6.10×10^6
c_{18}	3+	2.89×10^6
c_{21}	3+	3.36×10^7
c_{156}-z_{178}	3+	1.35×10^7
c_{56}-z_{86}	2+	1.14×10^7

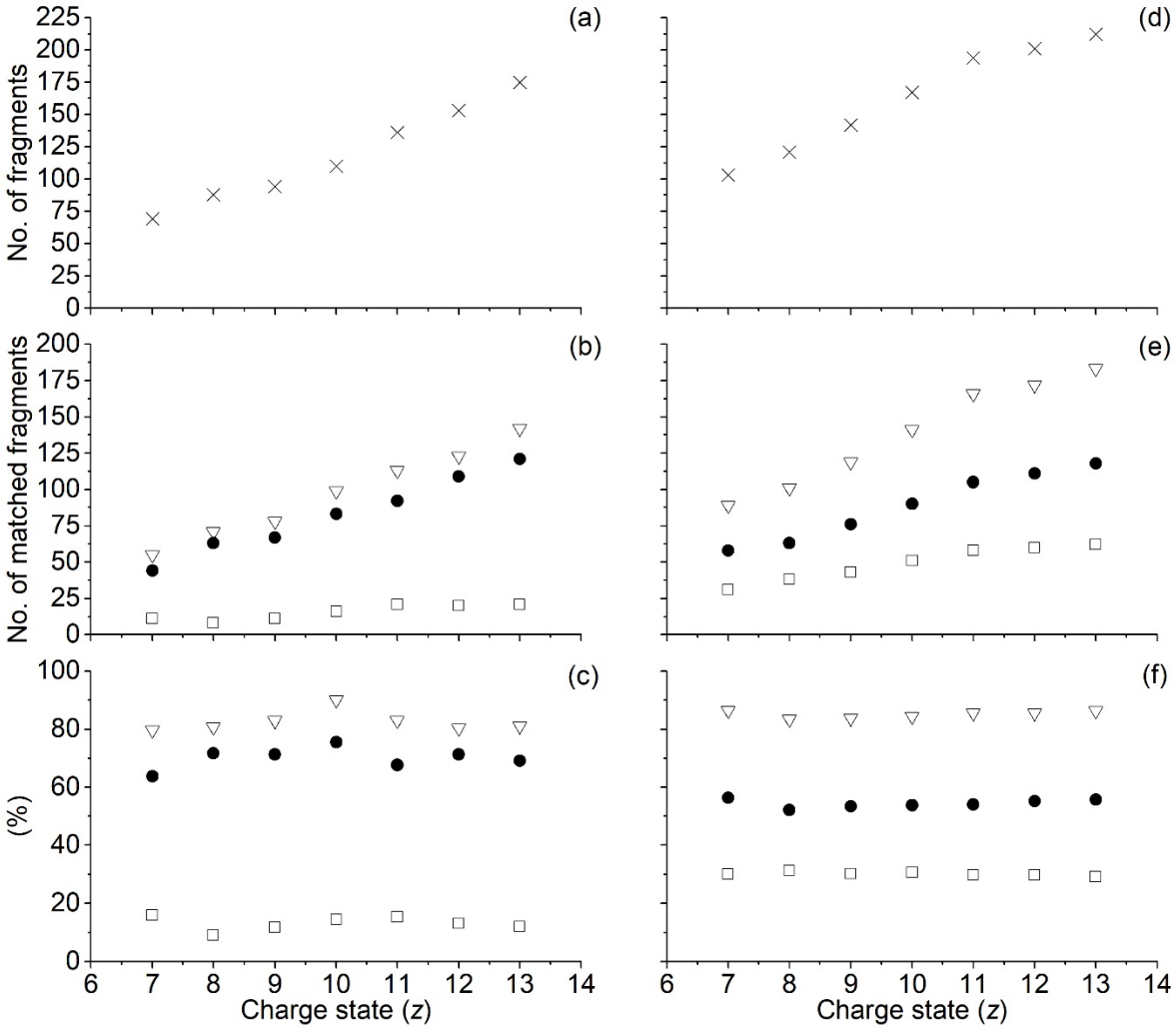

Figure S1. Performance of ECD-MS for isolated [Ubq, *z*H]$^{z+}$ (*z* = 7+ to 13+) formed from a solution of 10μM Ubq diluted in 49.5% water, 49.5% methanol, and 1% formic acid (*v/v*), where (a) is the number of fragments automatically deconvoluted, and (b) shows the total number of fragments identified (open triangles), total number of terminal fragments identified (closed circles), and the total number of internal fragments identified (open squares). The percentage of the fragments matched is shown in (c), with the percentage of all fragments identified (open triangles), percentage of terminal fragments identified (closed circles), and the percentage of internal fragments identified (open squares) depicted. The performance of EID-MS for isolated [Ubq, *z*H]$^{z+}$ (*z* = 7+ to 13+) is shown, where (d) is the number of fragments automatically deconvoluted, and (e) shows the total number of fragments identified (open triangles), total number of terminal fragments identified (closed circles), and the total number of internal fragments identified (open squares). (f) The percentage of the fragments matched, with the percentage of all fragments identified (open triangles), percentage of terminal fragments identified (closed circles), and the percentage of internal fragments identified (open squares) shown.

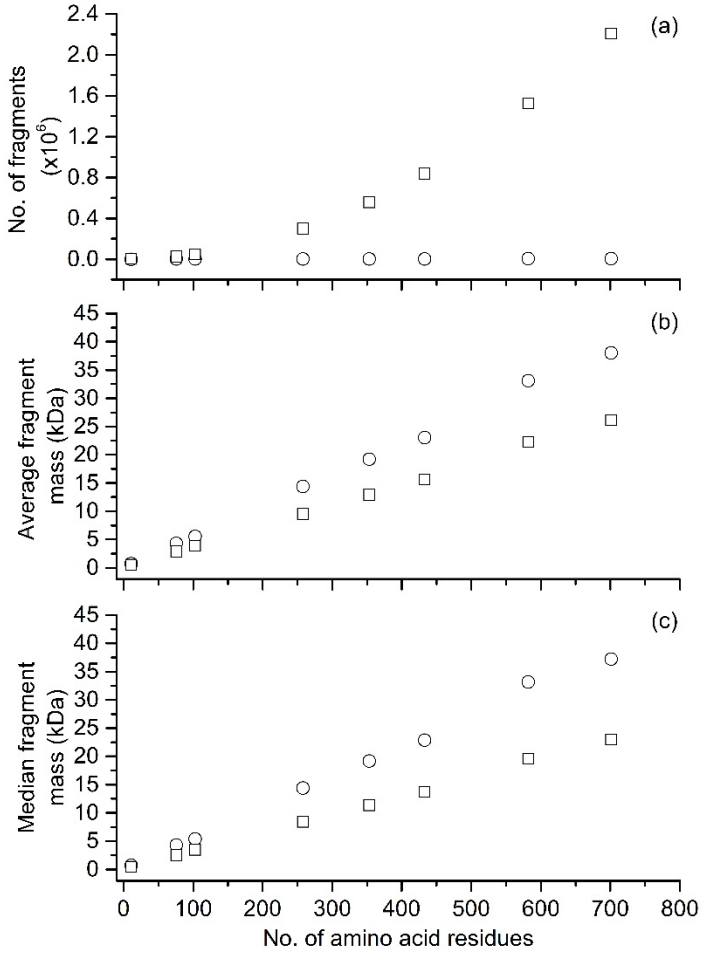

Figure S2. For a protein of *n* residues, the (a) number of theoretical deisotoped fragment ion masses, and the (b) average and (c) median theoretical fragment ion masses as calculated by the python script described in the Experimental are plotted in the Figure. Terminal fragment ions are denoted by open circles, and internal fragment ions are denoted by open squares.

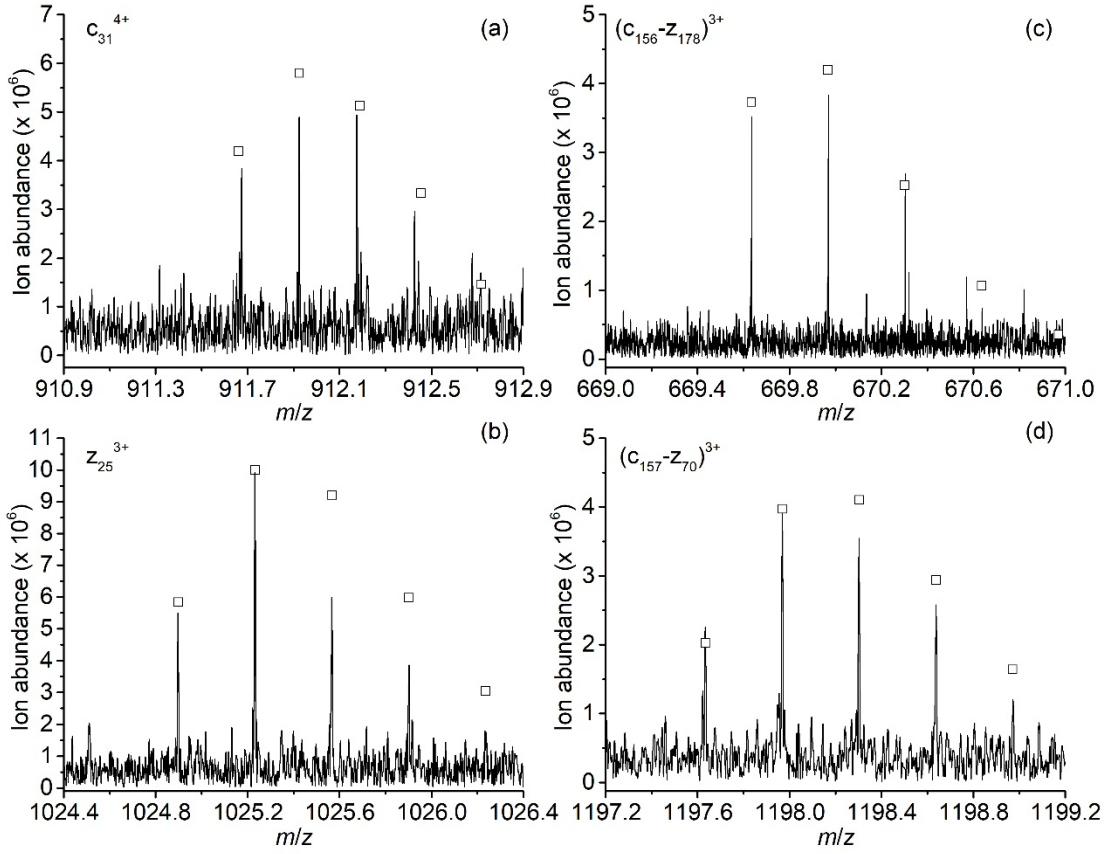

Figure S3. Isotopic distributions for (a-b) terminal fragments, and (c-d) internal fragments formed by EID of $[CAII, 25H]^{25+}$. Squares represent the theoretical isotopic distributions of the given sequence.

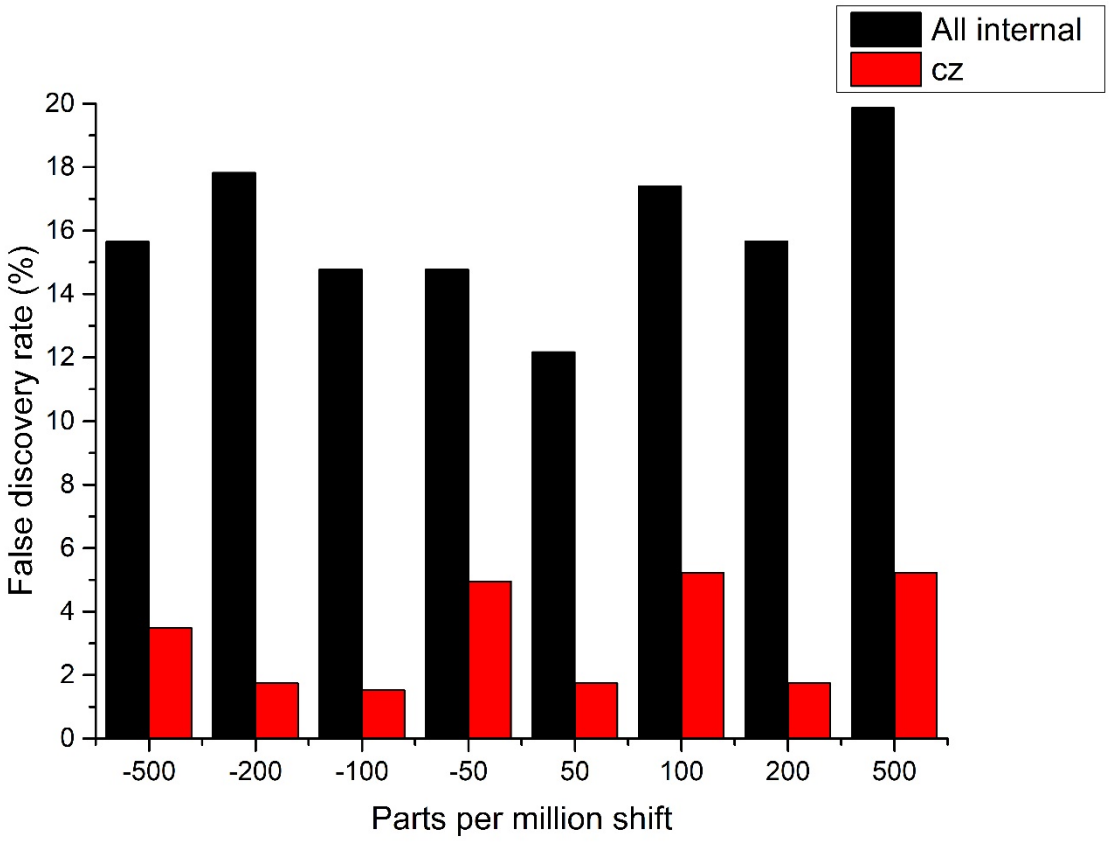

Figure S4. Bar graph showing the percentage of false discovery hits for $[CAII, 25H]^{25+}$ internal fragments for all internal fragments (black bars) and for *cz* internal fragments (red bars)

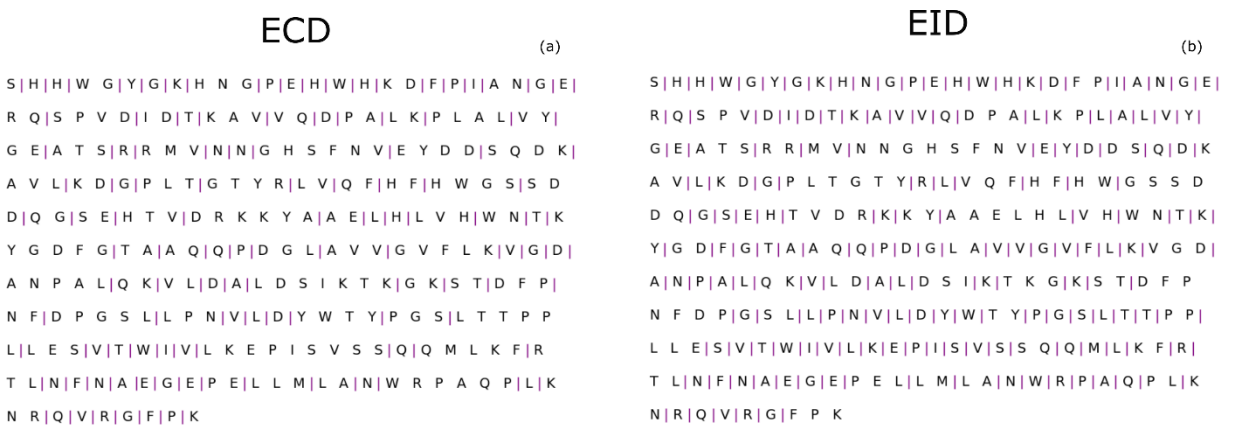

Figure S5. Location of fragmentation sites on the protein backbone for [CAII, 25H]$^{25+}$ for both internal and terminal fragments formed by (a) ECD, and (b) EID.

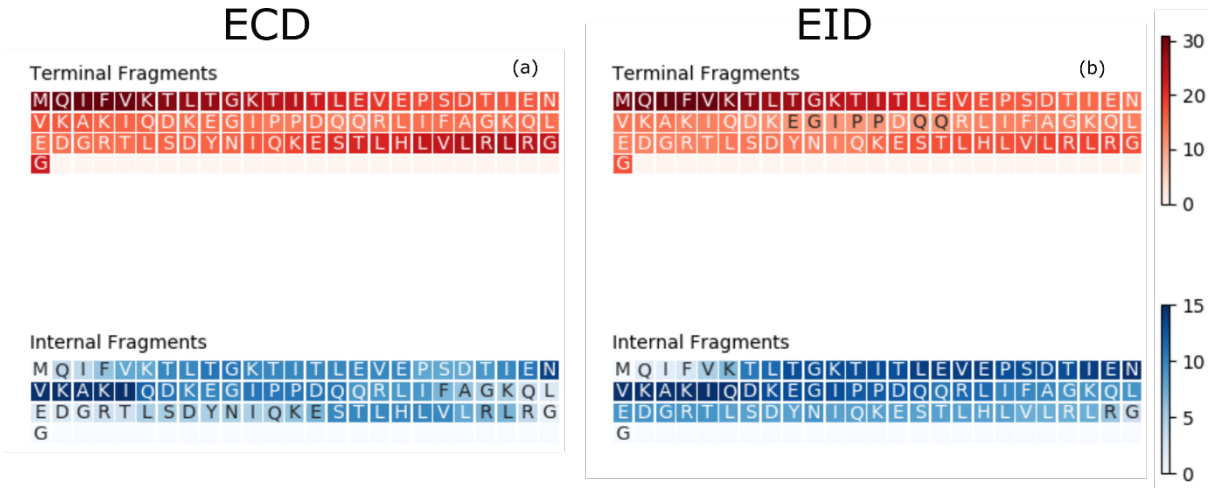

Figure S6. Heatmap depicting the number of times each residue is covered by a terminal fragment (top) or an internal fragment (bottom) for (a) ECD of $[Ubq, 10H]^{10+}$, and (b) EID of $[Ubq, 10H]^{10+}$. Darker colors indicate greater coverage.

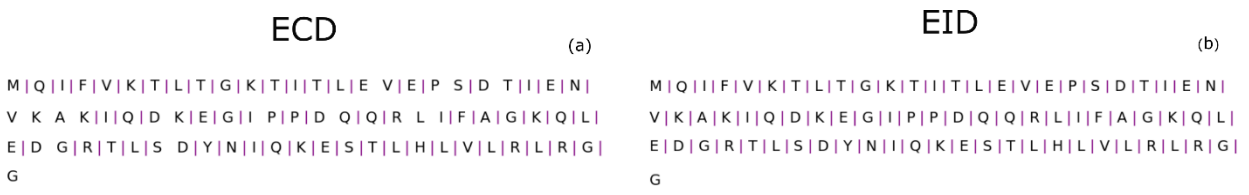

Figure S7. Location of fragmentation sites on the protein backbone for [Ubq, 10H]$^{10+}$ for both internal and terminal fragments formed by (a) ECD, and (b) EID.