Electronic Supporting Information

Influence of protein ion charge state on 213 nm top-down UVPD

Simon Becher[a], Huixin Wang[b], Michael G. Leeming[c], William A. Donald[b] and Sven Heiles[a]*

[a] Institute of Inorganic and Analytical Chemistry, Justus Liebig University Giessen, 35392 Giessen, Germany E-mail: sven.heiles@anorg.chemie.uni-giessen.de **[b] Mark Wainwright Analytical Centre, University of New South Wales, New South Wales, Australia [c] Bio21 Molecular Science and Biotechnology Institute, The University of Melbourne, Victoria, Australia [d] School of Chemistry, University of New South Wales, Sydney, New South Wales**

*Corresponding author

Dr. Sven Heiles

Institute of Inorganic and Analytical Chemistry

Justus Liebig University Giessen

35392 Giessen, Germany

E-mail: sven.heiles@anorg.chemie.uni-giessen.de

Tel.: +49 641 99-34812

Table 1. GB intrinsic values of amino acid residues used for protonation site prediction, electrostatic potential and repulsive column energy calculation for ubiquitin, cytochrome c and myoglobin using the freeware predictPrPlus (v1.0.0). Dielectric constant of 1.0, residue spacing distance of 3.6 Å, optimization steps of 10000 and iterations of 5000 are used.

	(kJ mol^{-1})
Arg	1006.6
Lys	951.0
His	941.0
Trp	916.3
Backbone	881.8

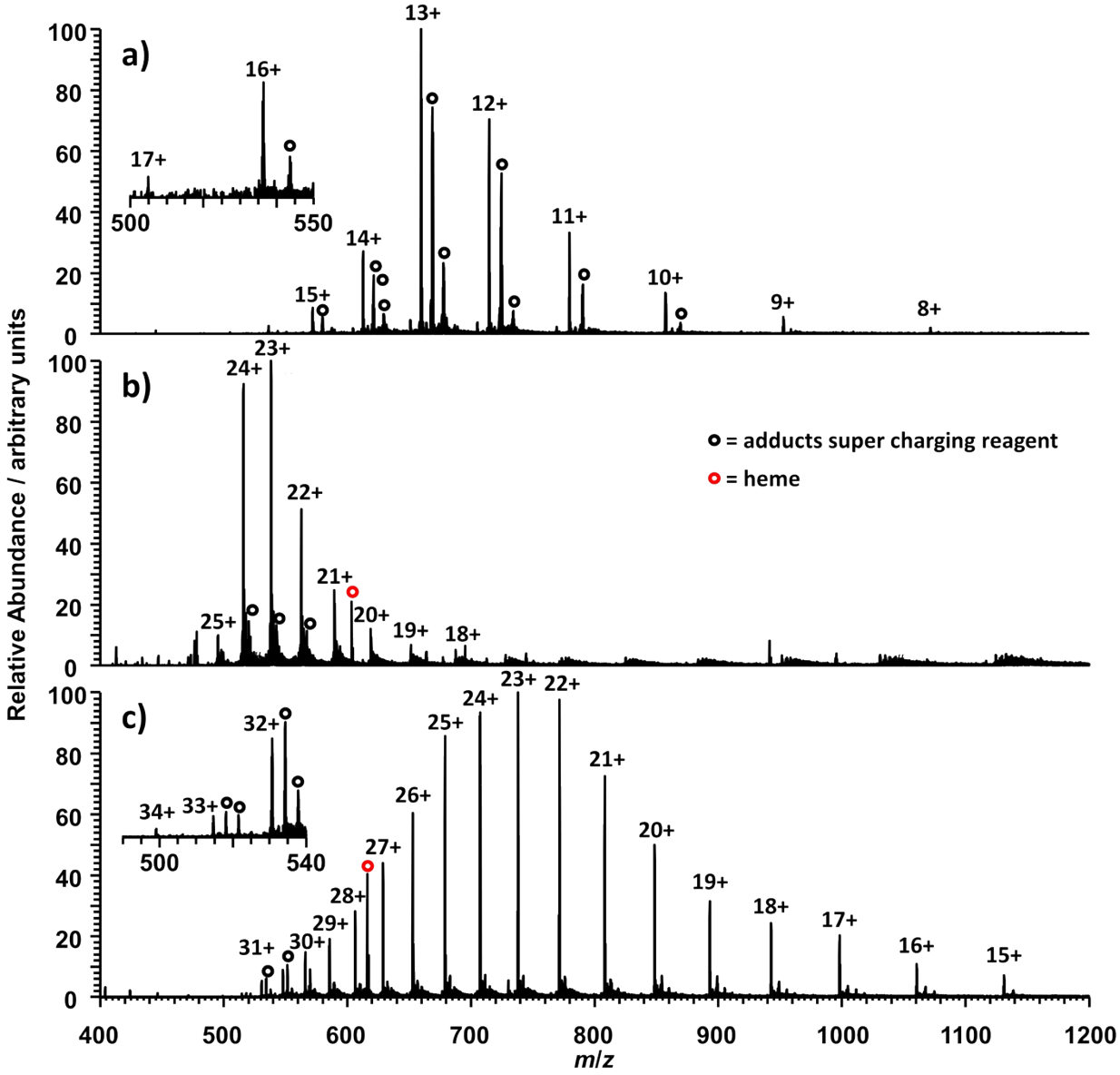


Figure S 1. Overview spectra of a) ubiquitin with 1 % vol VEC, b) cytochrome c with 20 % vol PC and c) myoglobin with 1 % vol VEC. Adducts of the supercharging reagent are marked with black dots, the heme group in cytochrome c and myoglobin with red dots.

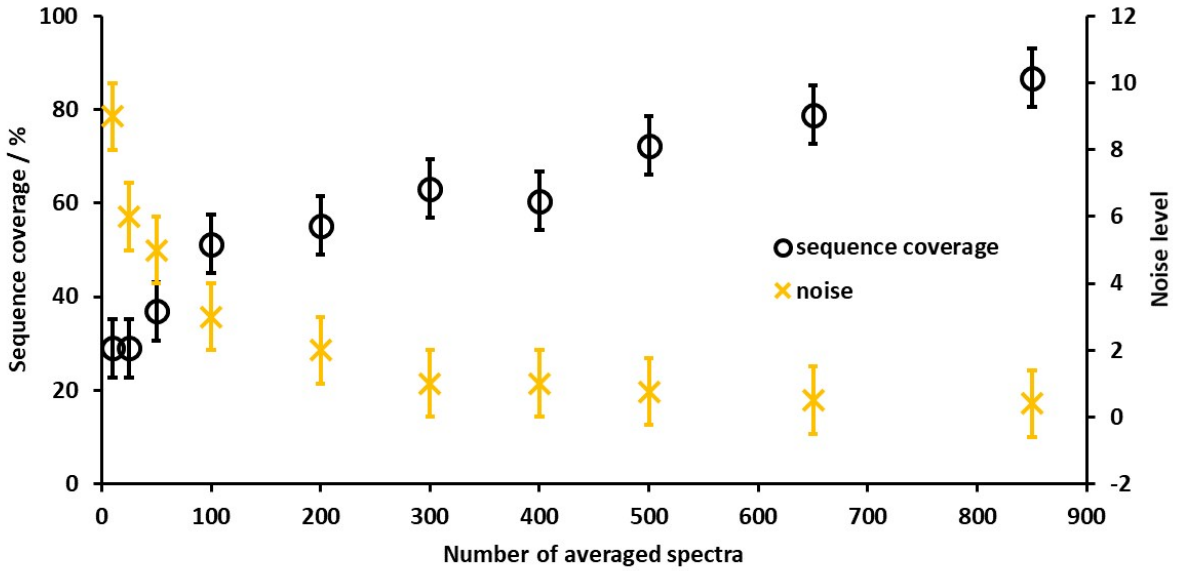


Figure S2. Dependence of sequence coverage (black circles) and the noise level (yellow crosses) of [ubiquitin+13H]$^{13+}$ at NL 2 · 10^4 from the number of averaged spectra.

Noise levels in Figure S2 were manually identified. Sequence coverages reaches over 80 % for charge state 13+, which is in the range of the sequence coverages reported in the literature.[1,2]

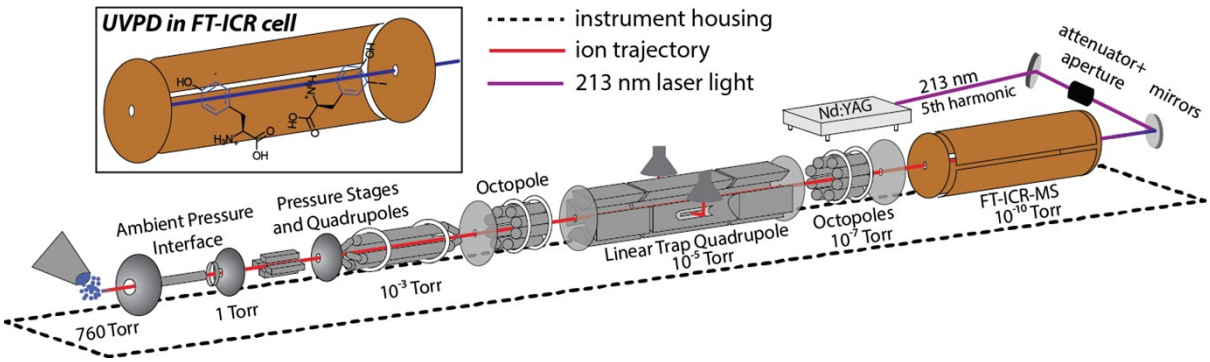


Figure S3. Schematic of the ''LTQ FT Ultra'' linear ion trap/FT ICR hybrid instrument modified for UVPD experiments. The instrument housing is indicated by a dashed black line. Typical pressure values under operation conditions are included. The ion trajectory, 213 nm laser light are shown as red and purple lines, respectively. Instrument components are not to scale.

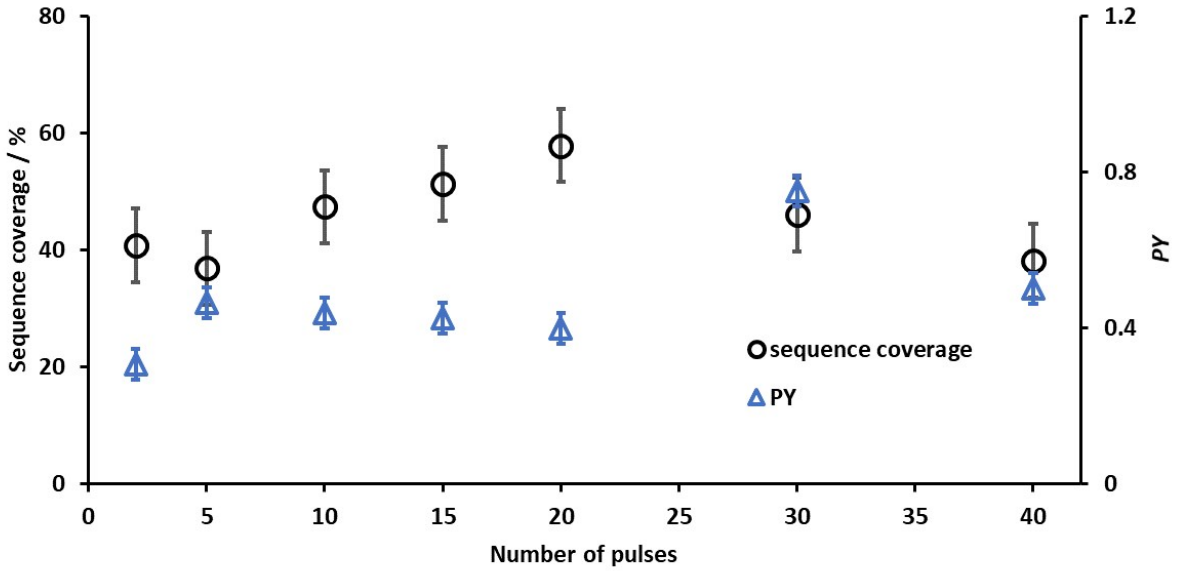


Figure S4. Sequence coverage (black circles) and PY (blue triangles) of [ubiquitin+13H]$^{13+}$ as a function of the number of laser shots per spectrum.

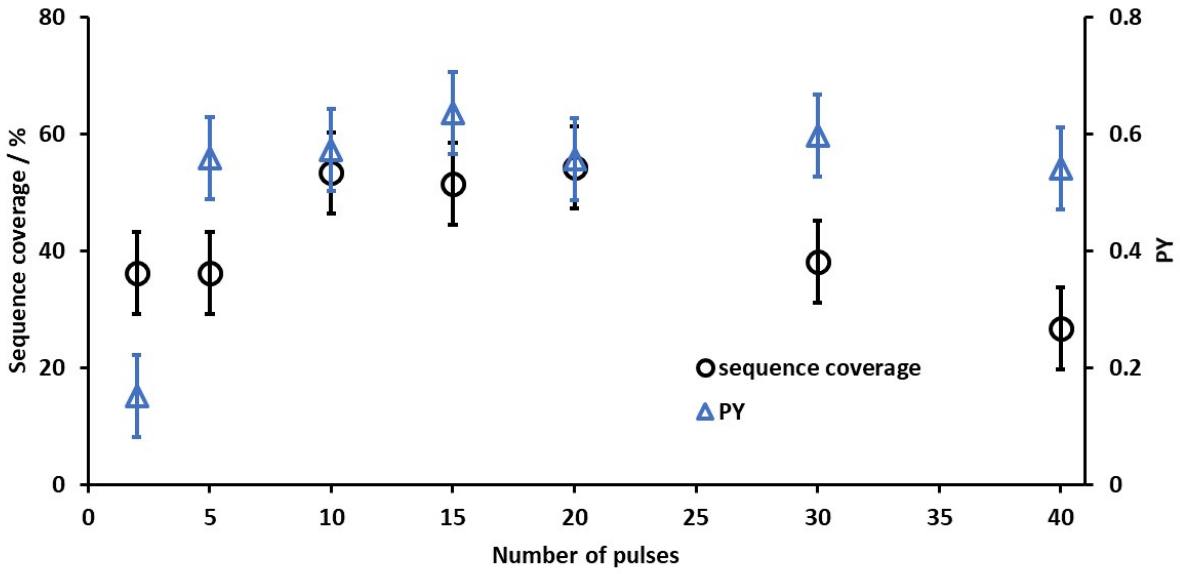


Figure S5. Sequence coverage (black circles) and PY (blue triangles) of [cytochrome c+19H]$^{19+}$ as a function of the number of laser shots per spectrum.

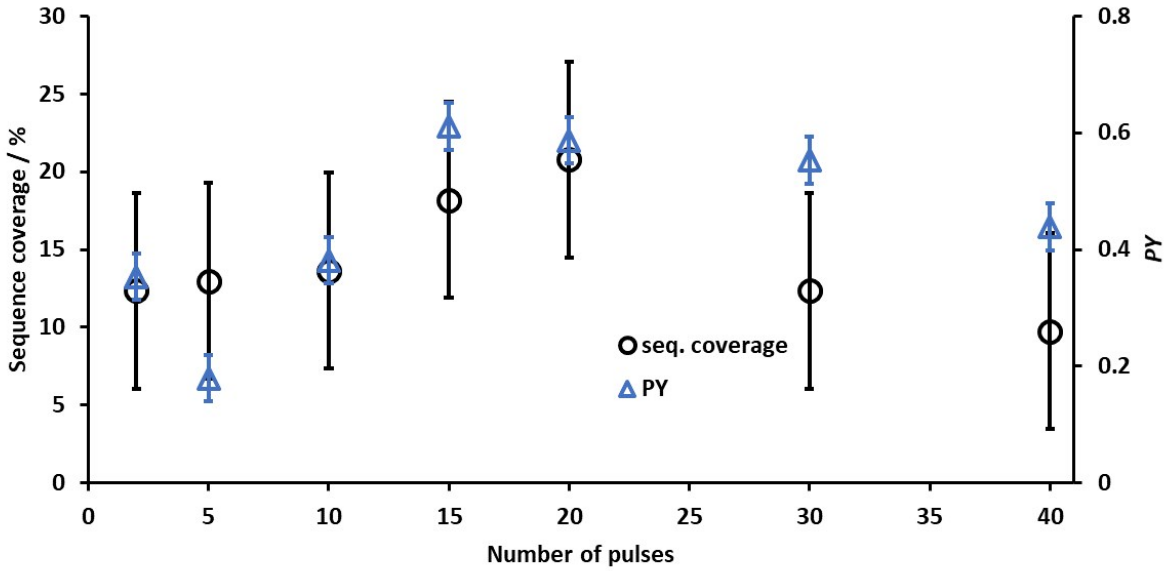


Figure S6. Sequence coverage (black circles) and PY (blue triangles) of [myoglobin c+21H]$^{21+}$ as a function of the number of laser shots per spectrum.

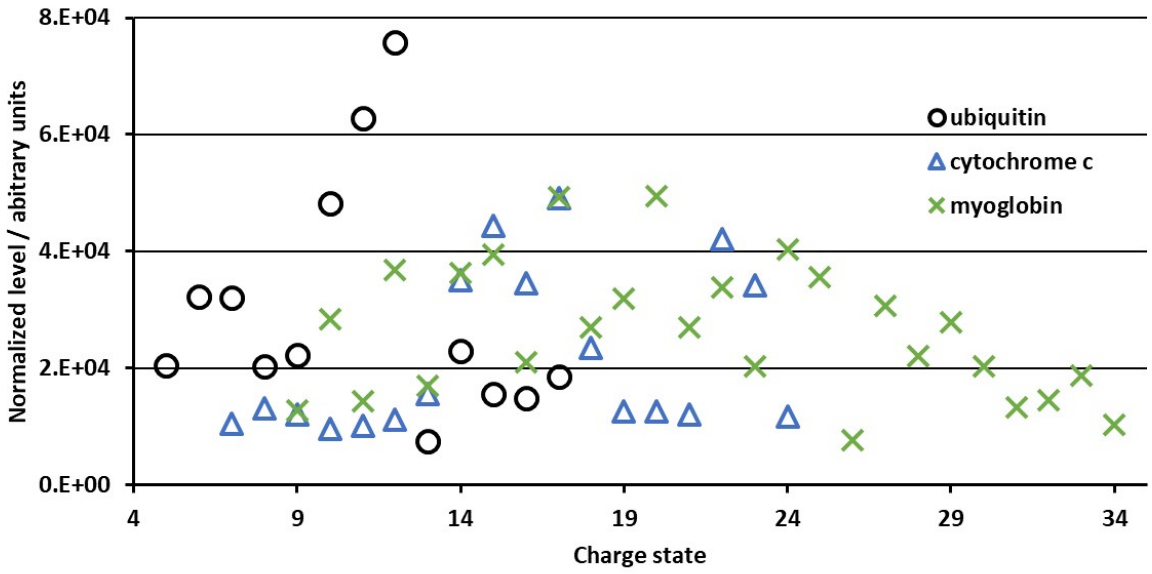


Figure S 7. Normalized levels of the isolated precursor signals of ubiquitin (black circles), cytochrome c (blue triangles) and myoglobin (green crosses).

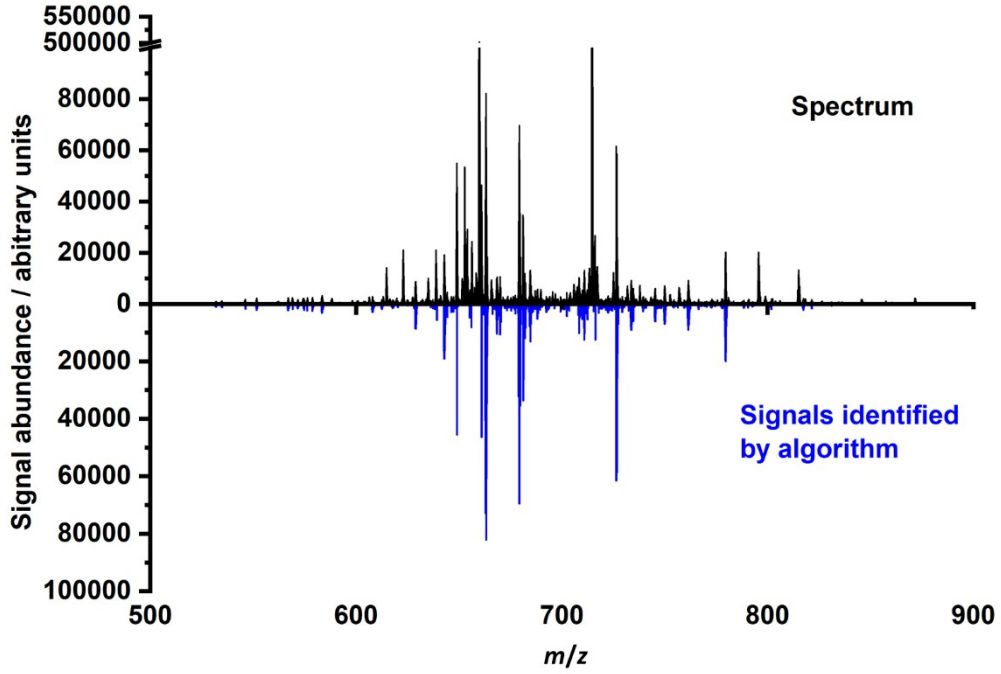


Figure S 8. Spectrum of [ubiquitin+13H]$^{13+}$ published by Halim et al (black) and the signals that ware taken in account for data analysis by the algorithm used in this study.

For the 13+ charge state of ubiquitin state Halim et al. reported a sequence coverage of 76 % using data evaluation by hand after deconvolution with Xtract.[3] The algorithm used in this study yielded 83 % sequence coverage for the same data. The algorithm omits deconvolution of the data.

Instead of a deconvolution algorithm the spectral data is used for the analysis. Each theoretical isotopic pattern for all a, b, c, x, y, z, the +/-1 fragments as well as a+2, b+2 and y-2 of all charge states are calculated upon the sequence. Finally, the theoretical predicted and measured isotopic patterns are compared. The isotopes taken in account are limited to all isotopes with a theoretical abundance of 65 % relative to the most intense isotope. Only fragments that fulfill mass accuracy and isotopic pattern pattern selection rules are taken into account for the data analysis.

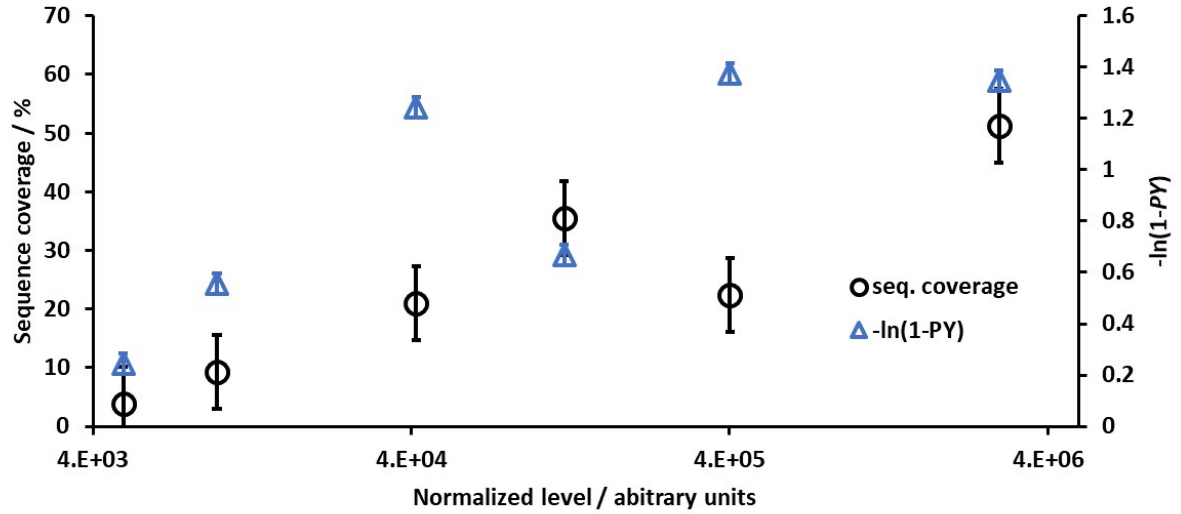


Figure S9. Sequence coverage (black circles) and PY (blue triangles) of [ubiquitin+13H]$^{13+}$ as a function of the number of trapped ions represented by the normalized level (NL).

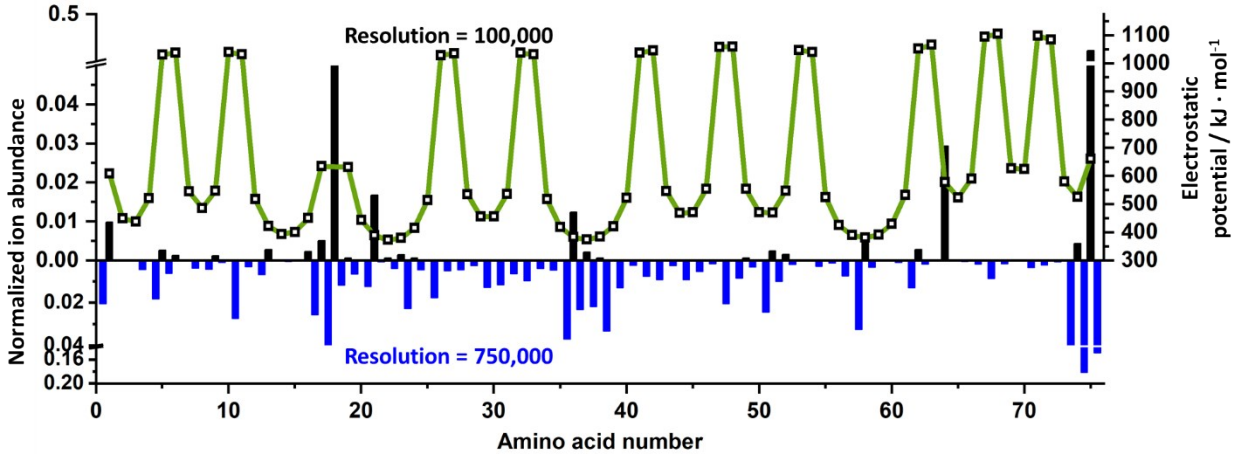


Figure S 10. Comparison of [ubiquitin+13H]$^{13+}$ fragmentation patterns measured using a resolution of 750,000 (blue bars) and 100,000 (black bars). For 750,000 mass resolution the sequence coverage is increased to 92%. The calculated electrostatic potential at each residue is shown with white squares, which are connected with green lines to guide to the eye.

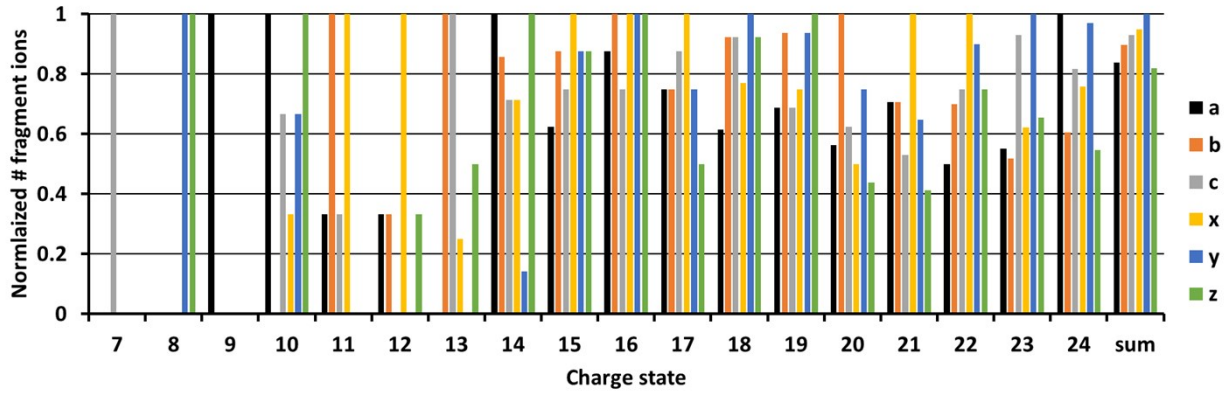


Figure S 11. Normalized number of fragment ions of cytochrome c charge states 7+ to 24+. Fragment ion of a, b, c, x , y, z are shown in black, orange, grey, yellow, blue and green, respectively.

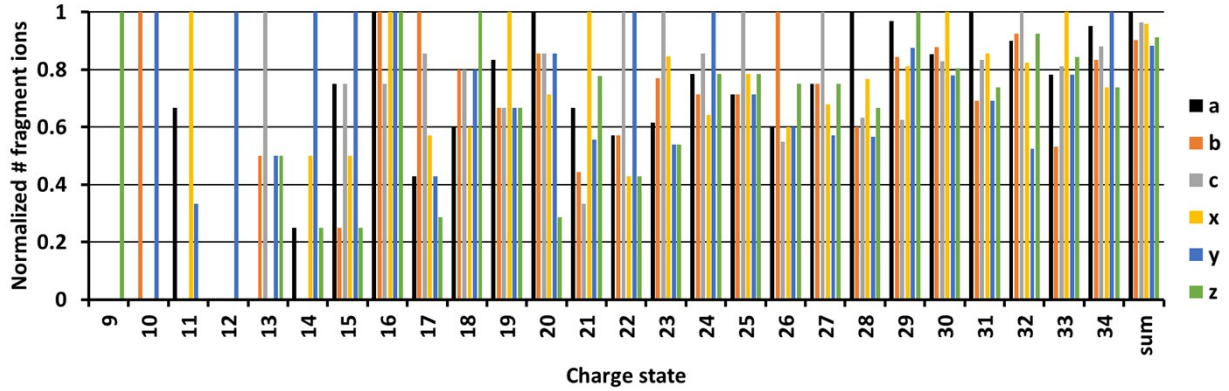


Figure S 12. Normalized number of fragment ions of myoglobin charge states 9+ to 34+. Fragment ion of a, b, c, x , y, z are shown in black, orange, grey, yellow, blue and green, respectively.

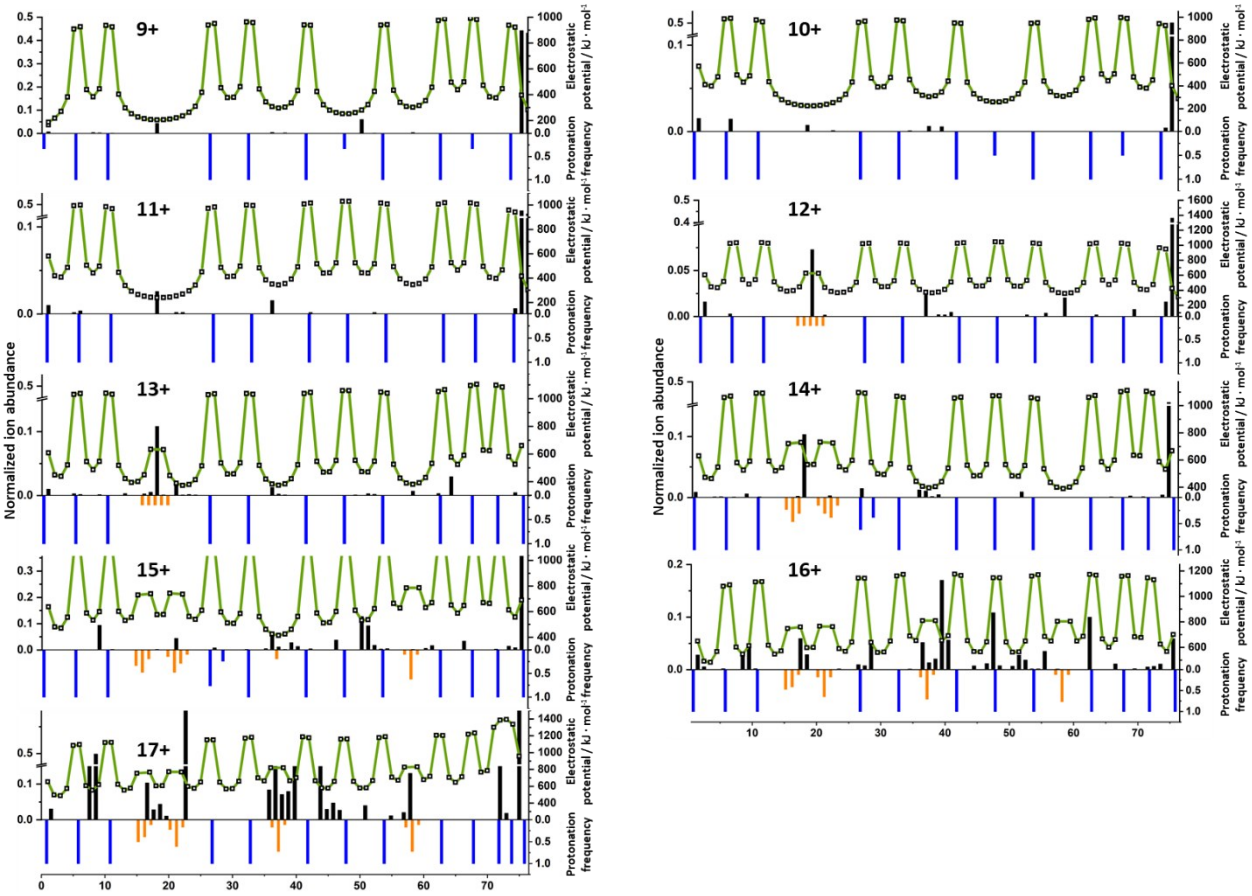


Figure S 13. Normalized experimental fragment ion abundances of ubiquitin 9+ to 17+ are shown as black bars. The calculated electrostatic potential at each residue is shown with white squares, which are connected with green lines to guide to the eye. Predicted protonation frequencies at basic side-chain amino acid residue (or N/C-terminus) and the amide backbone are shown in blue and orange bars, respectively. Ubiquitin charge states ≤ 8 are not shown, because the prediction of electrostatic potentials and protonation frequencies should be more accurate for elongated structures.

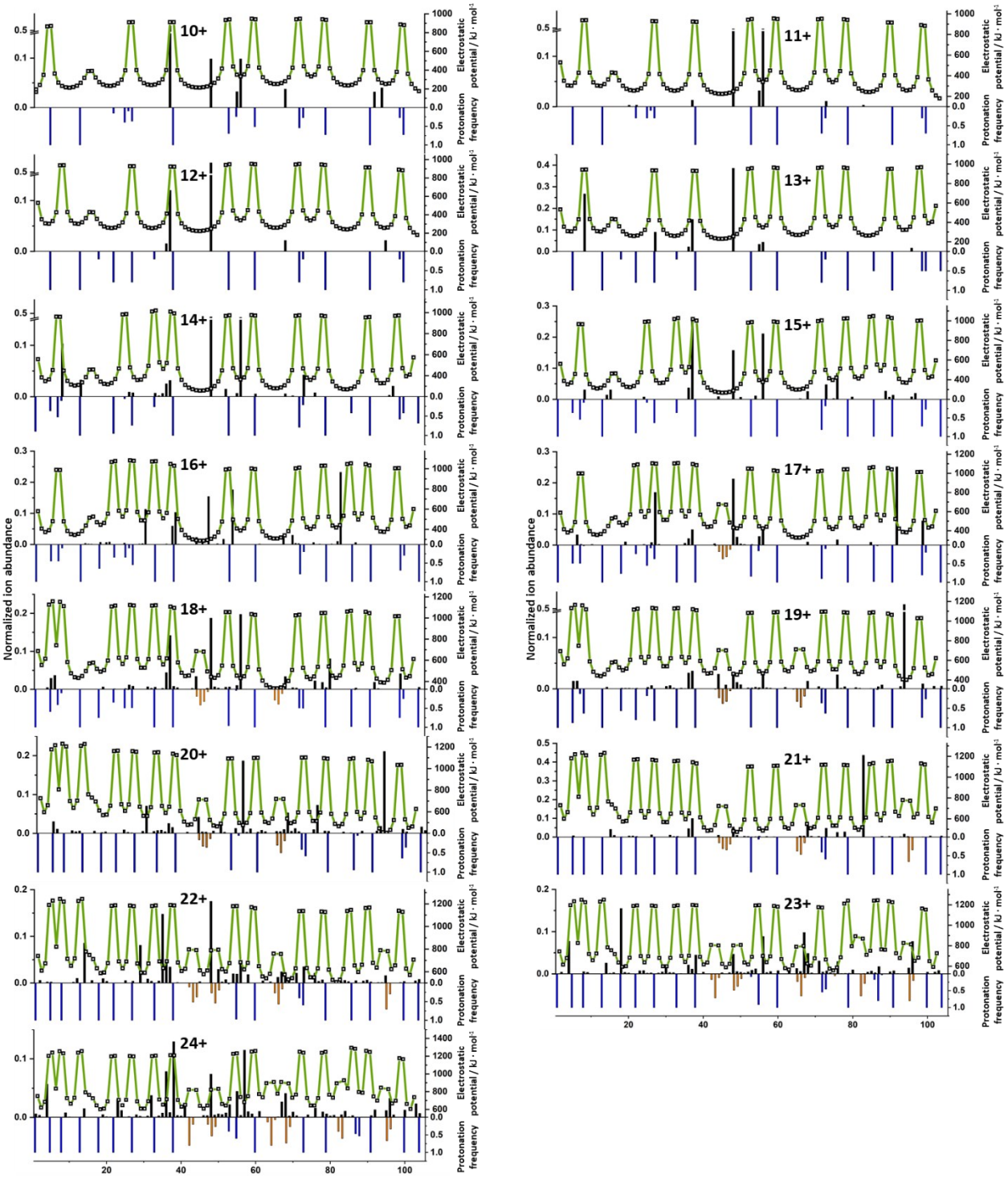
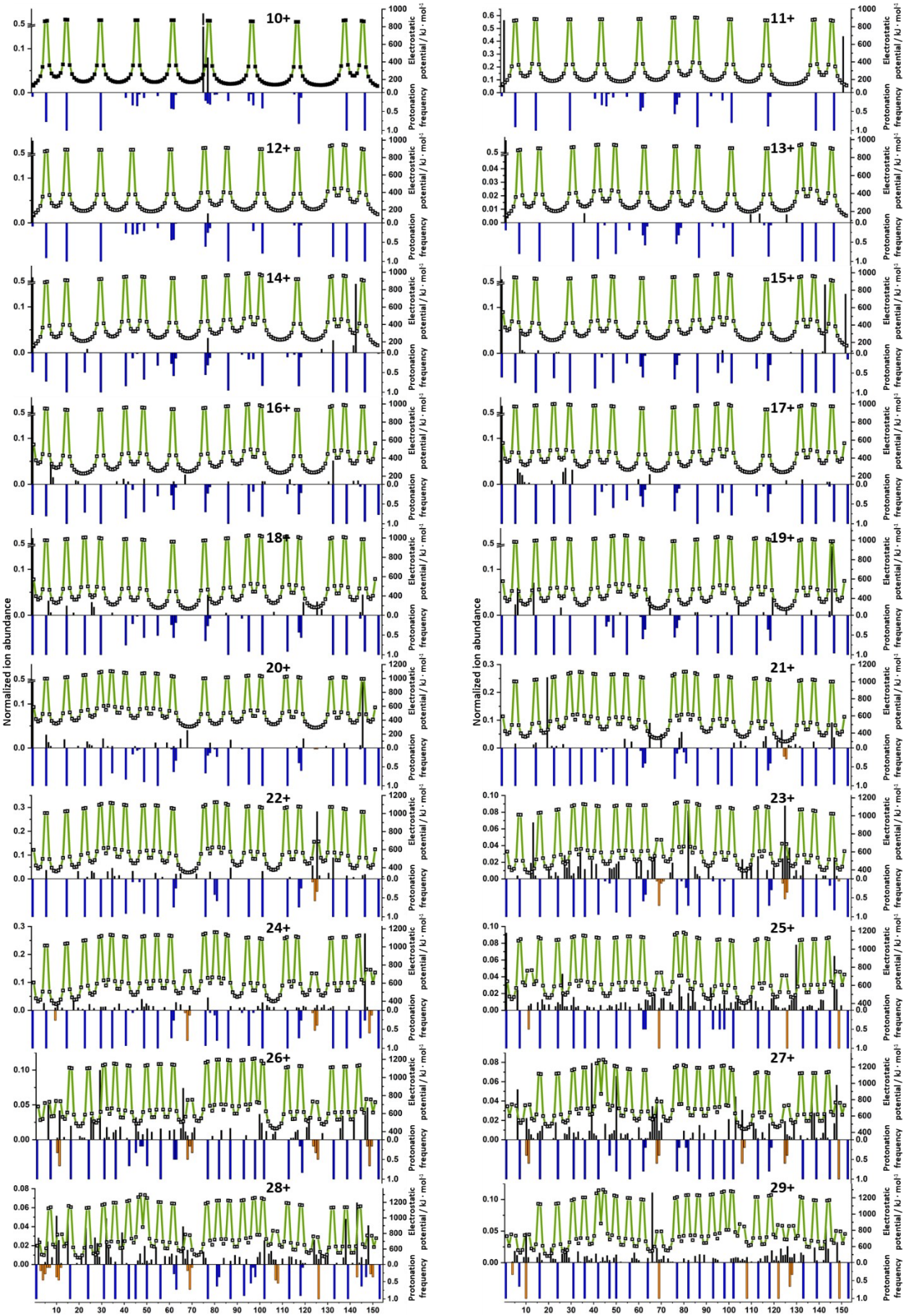


Figure S 14. Normalized experimental fragment ion abundances of cytochrome c 10+ to 24+ are shown as black bars. The calculated electrostatic potential at each residue is shown with white squares, which are connected with green lines to guide to the eye. Predicted protonation frequencies at basic side-chain amino acid residue (or N/C-terminus) and the amide backbone are shown in blue and orange bars, respectively. Cytochrome c charge states ≤ 9 are not shown, because the prediction of electrostatic potentials and protonation frequencies should be more accurate for elongated structures.


10+
11+
12+
13+
14+
15+
16+
17+
18+
19+
20+
21+
22+
23+
24+
25+
26+
27+
28+
29+
Normalized ion abundance
Normalized ion abundance
Electrostatic potential / kJ · mol⁻¹
Protonation frequency

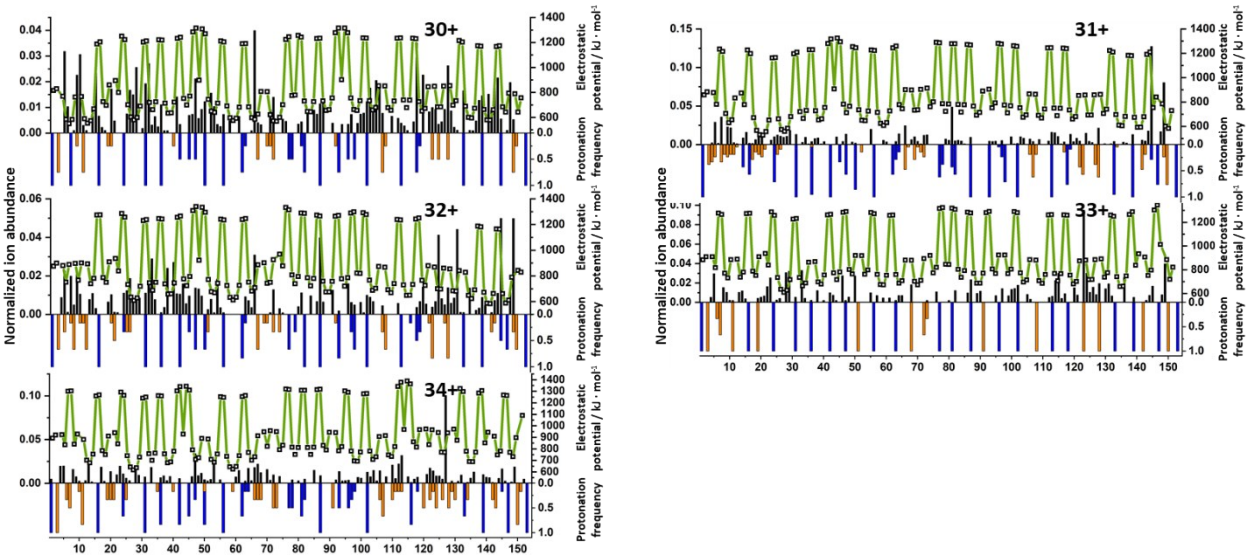


Figure S 15. Normalized experimental fragment ion abundances of myoglobin 10+ to 34+ are shown as black bars. The calculated electrostatic potential at each residue is shown with white squares, which are connected with green lines to guide to the eye. Predicted protonation frequencies at basic side-chain amino acid residue (or N/C-terminus) and the amide backbone are shown in blue and orange bars, respectively. Myoglobin charge states ≤ 9 are not shown, because the prediction of electrostatic potentials and protonation frequencies should be more accurate for elongated structures.

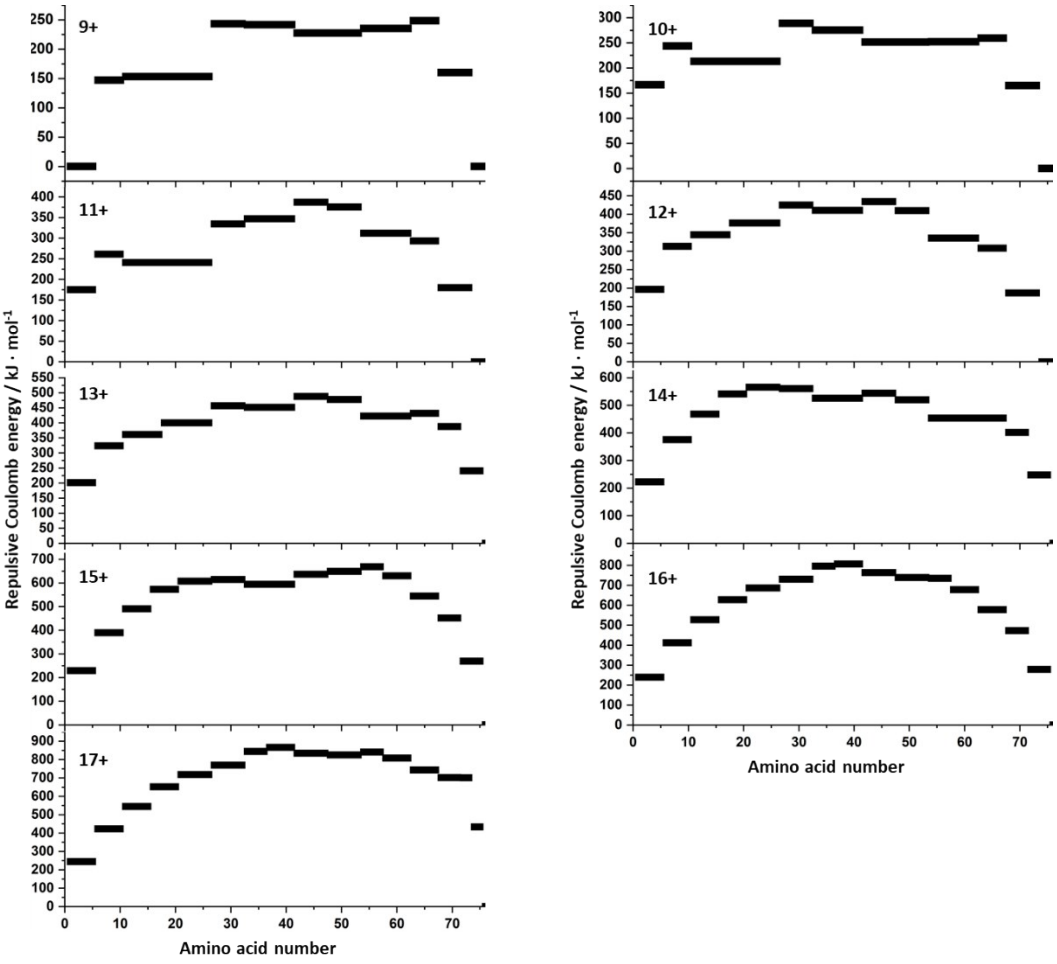


Figure S 16 Calculated repulsive Coulomb energy between hypothetical fragment ions as a function of cleavage site for [ubiquitin+9H]$^{9+}$ to [ubiquitin+17H]$^{17+}$.

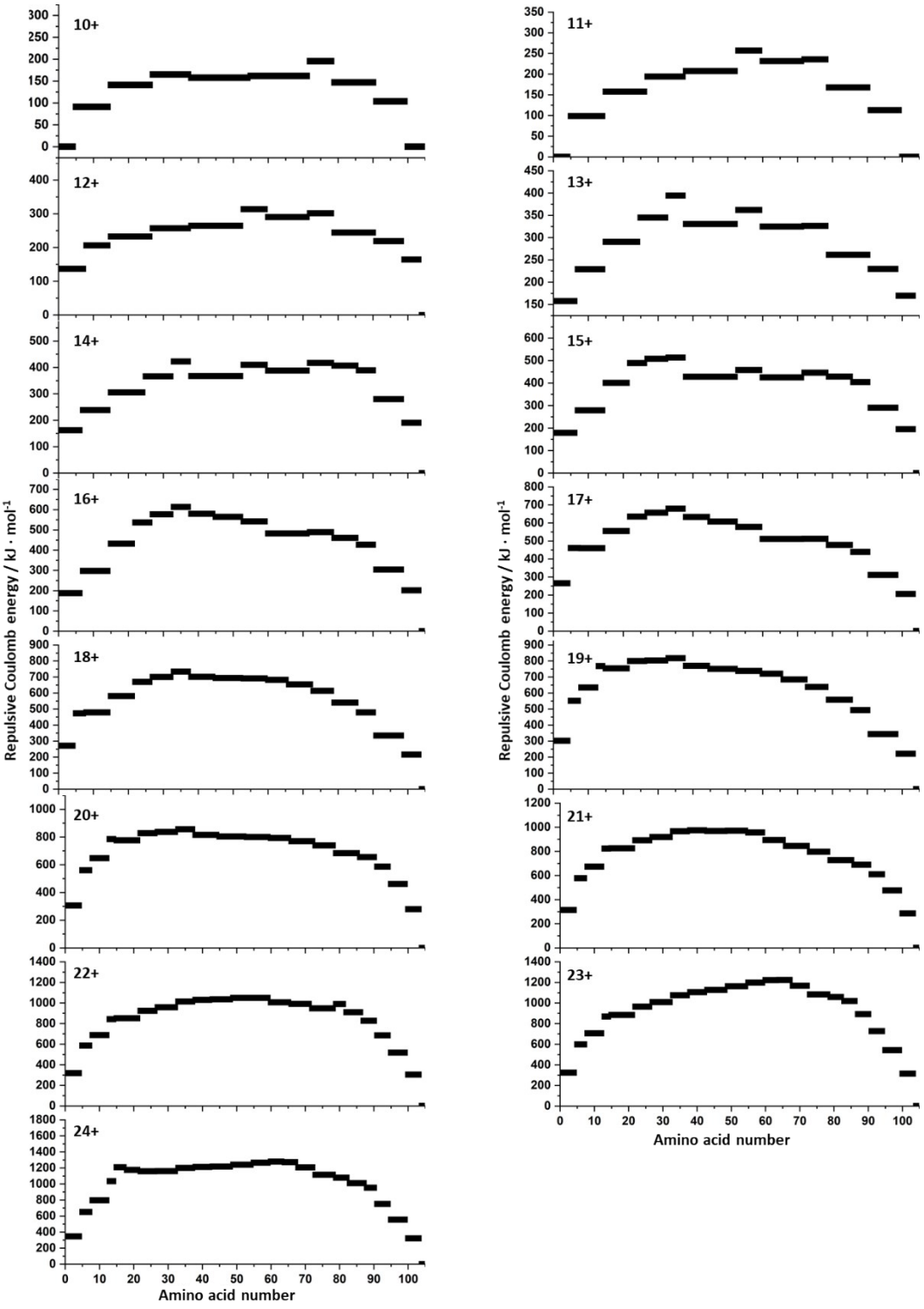
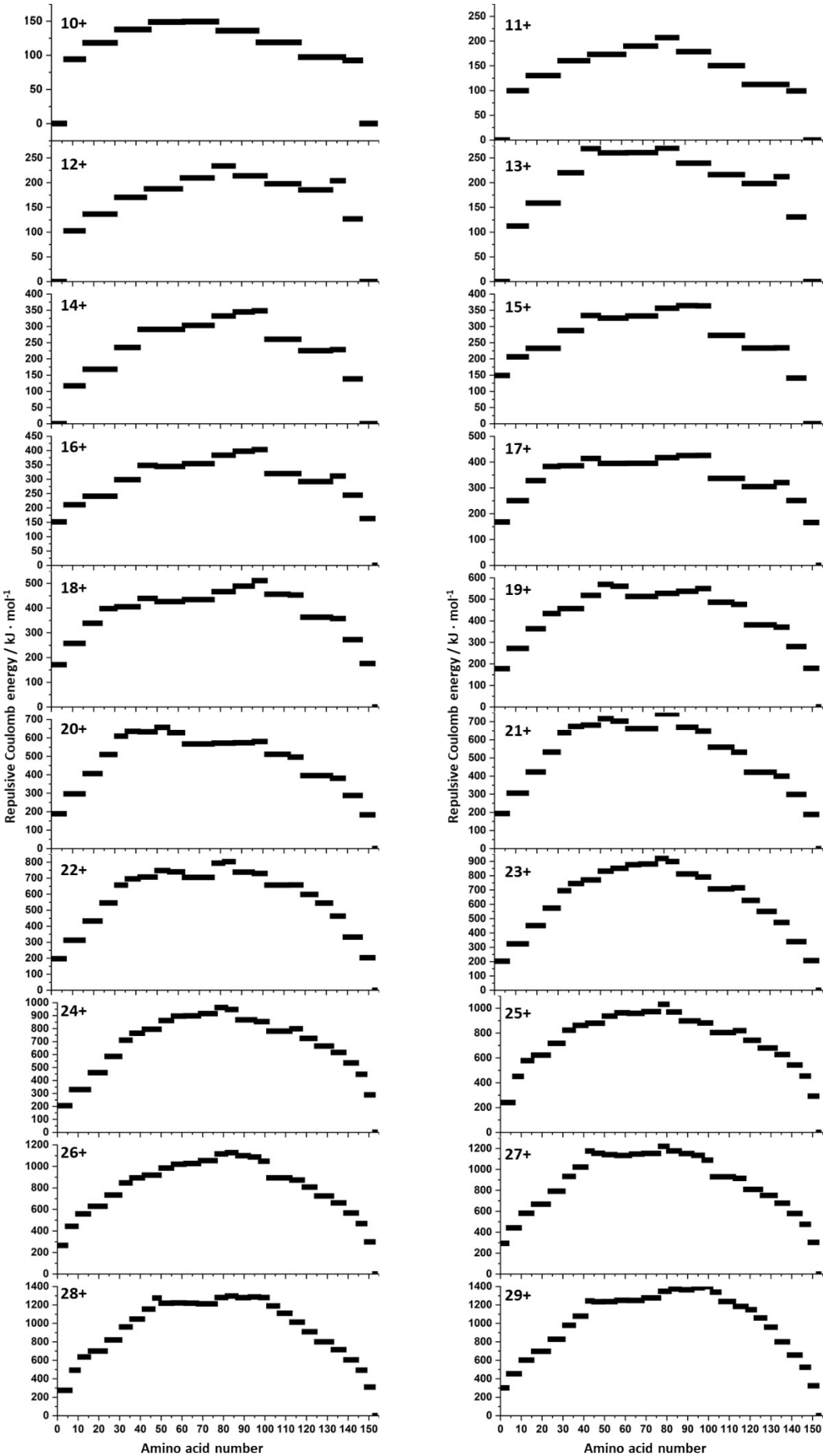


Figure S 17. Calculated repulsive Coulomb energy between hypothetical fragment ions as a function of cleavage site for [cytochrome c+9H]$^{10+}$ to [cytochrome c+23H]$^{24+}$.


10+
11+
12+
13+
14+
15+
16+
17+
18+
19+
20+
21+
22+
23+
24+
25+
26+
27+
28+
29+
Repulsive Coulomb energy / kJ · mol-1
Amino acid number

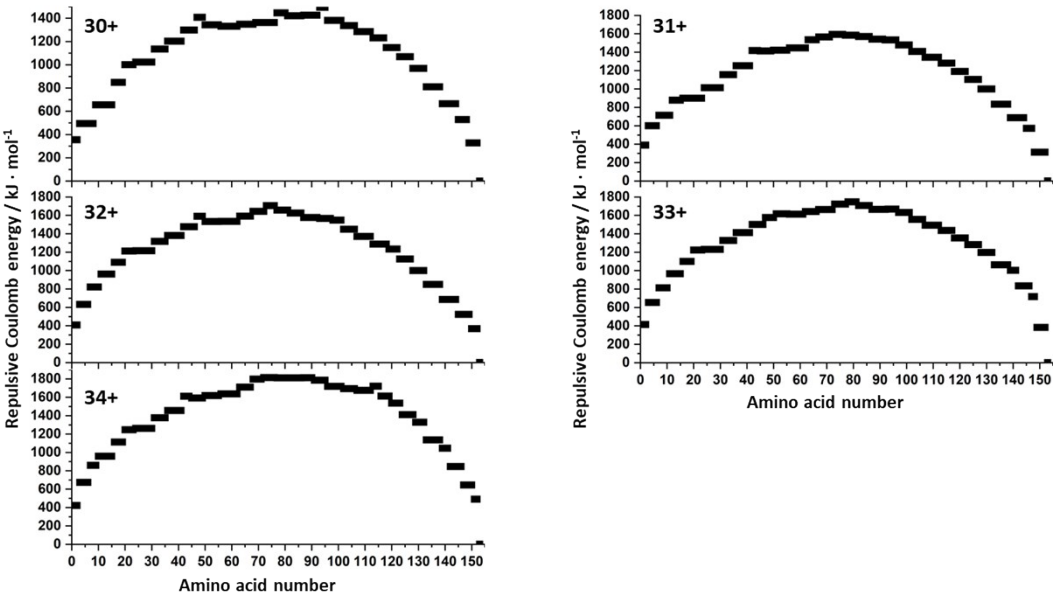


Figure S 18 Calculated repulsive Coulomb energy between hypothetical fragment ions as a function of cleavage site for [myoglobin+10H]$^{10+}$ to [myoglobin+34H]$^{34+}$.

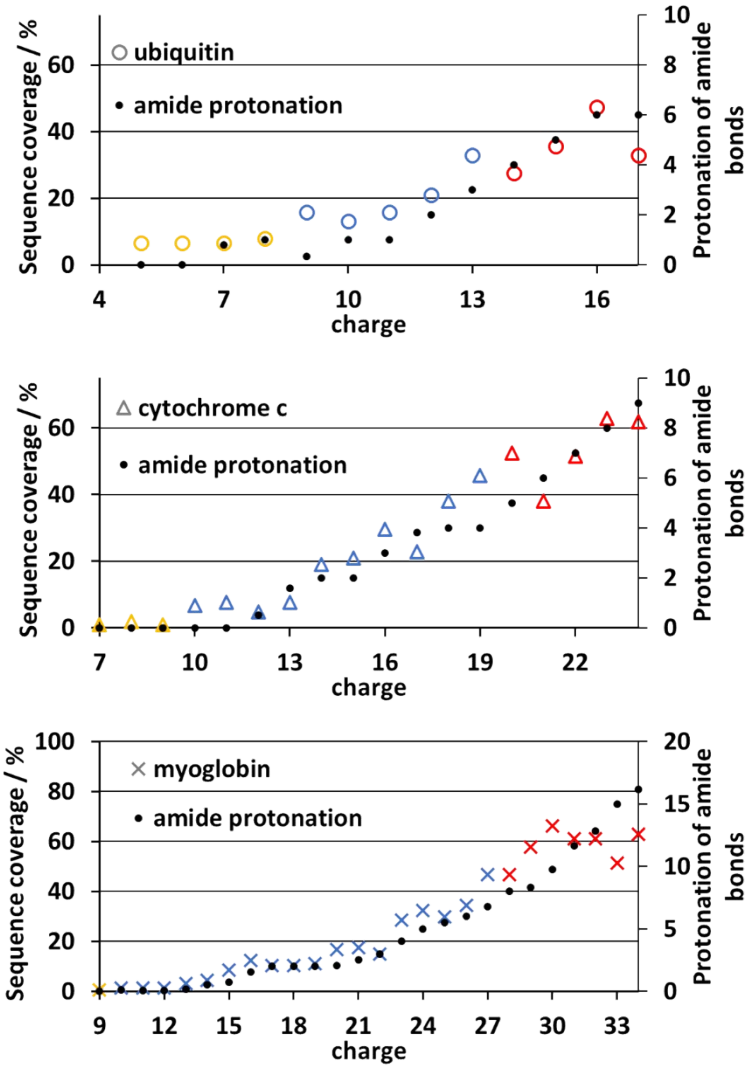


Figure S 19. Sequence coverage for ubiquitin, cytochrome c and myoglobin ions and charges at the amide backbone as a function of overall charge state. Ubiquitin, cytochrome c and myoglobin are represented by circles, triangles and crosses, respectively. Proteins in native, denaturing and unfolded conformation are color-coded by yellow, blue and red, respectively.

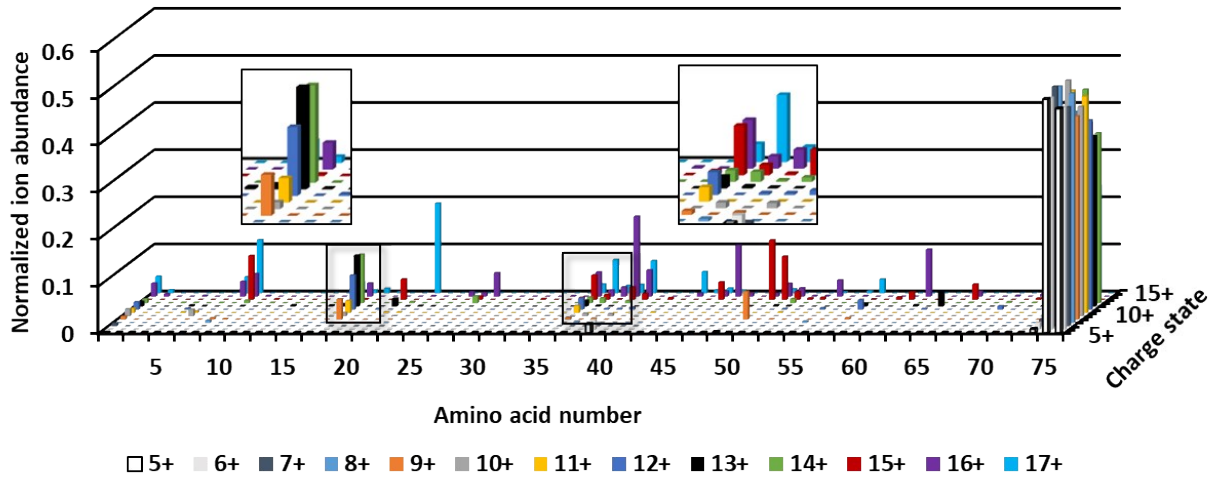


Figure S 20. Normalized fragment ion abundances of ubiquitin charge states 5+ to 17+ mapped onto the amino acid number. Charge states 5 to 17 are shown in white, light grey, blue grey, light grey blue, orange, medium grey, yellow, blue, black, green, red, purple and light blue respectively

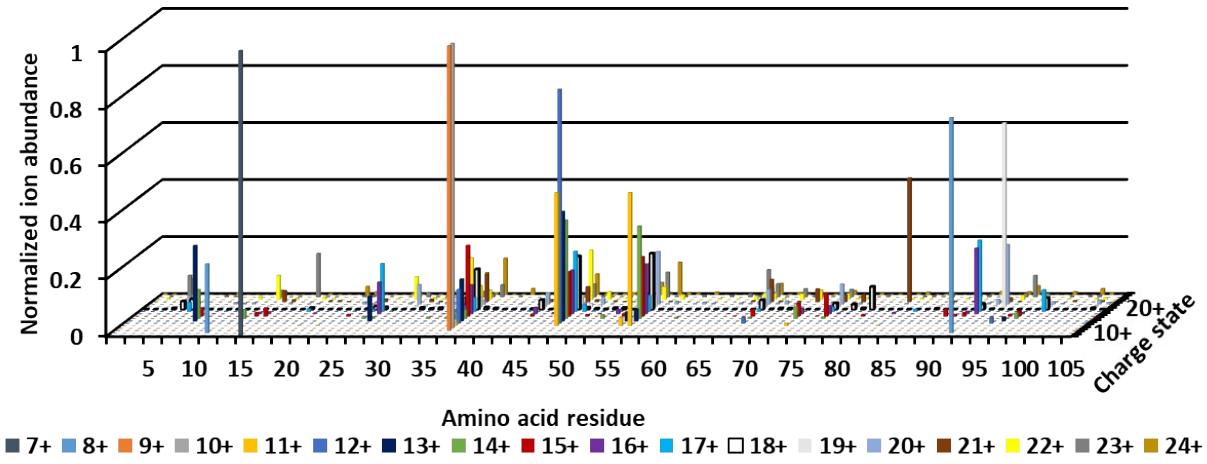


Figure S 21. Normalized fragment ion abundances of cytochrome c charge states 7+ to 24+ mapped onto the amino acid number. Charge states 7 to 24 are shown in blue grey, light grey blue, orange, medium grey, yellow, blue, black, green, red, purple, light blue, white, light grey, light grey blue, brown, yellow, dark grey and light brown respectively.

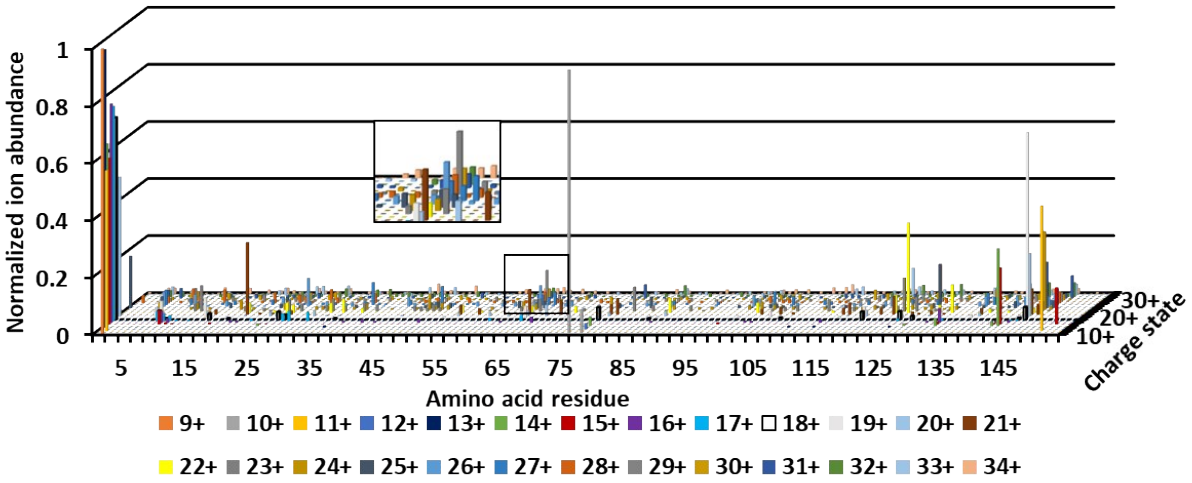


Figure S 22. Normalized fragment ion abundances of myoglobin charge states 9+ to 34+ mapped onto the amino acid number. Charge states 9 to 34 are shown in orange, medium grey, yellow, blue, black, green, red, purple, light blue, white, light grey, light grey blue, brown, yellow, dark grey, light brown, blue grey, light grey blue, blue, dark orange, dark grey, light brown, dark blue, dark green, light blue and skin color respectively.

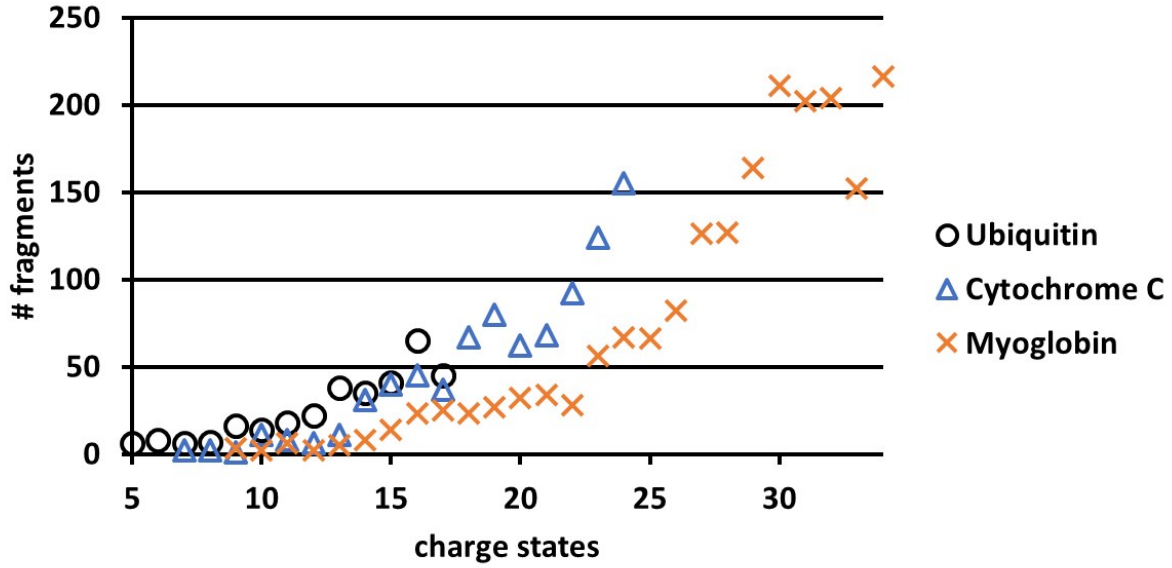


Figure S 23 Number of fragment ions for ubiquitin, cytochrome c and myoglobin vs. charge state.

References

(1) Basak, P.; Kundu, N.; Pattanayak, R.; Bhattacharyya, M. Denaturation properties and folding transition states of leghemoglobin and other heme proteins. *Biochemistry. Biokhimiia* **2015**, *80*, 463–472.

(2) 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.*; et al.* Thorough Performance Evaluation of 213 nm Ultraviolet Photodissociation for Top-down Proteomics. *Mol. Cell. Proteomics* **2020**, *19*, 405–420.

(3) Halim, M. A.; Girod, M.; MacAleese, L.; Lemoine, J.; Antoine, R.; Dugourd, P. Combined Infrared Multiphoton Dissociation with Ultraviolet Photodissociation for Ubiquitin Characterization. *J. Am. Soc. Mass Spectr.* **2016**, *27*, 1435–1442.