Associated Content

Exploring the Conformations and Binding Location of HMGA2•DNA Complexes Using Ion Mobility Spectrometry and 193 nm Ultraviolet Photodissociation Mass Spectrometry

Sarah N. Sipe,[1] Kevin Jeanne Dit Fouque,[2] Alyssa Garabedian,[2] Fenfei Leng,[2,3] Francisco Fernandez-Lima,[2,3] Jennifer S. Brodbelt[1*]

[1] Department of Chemistry, University of Texas, Austin, TX 78712 United States
[2] Department of Chemistry and Biochemistry, Florida International University, Miami, FL 33199, United States
[3] Biomolecular Sciences Institute, Florida International University, Miami, FL 33199, United States

*Correspondence to: jbrodbelt@cm.utexas.edu

Table of Contents:

Ion Mobility Spectrometry – Orbitrap Mass spectrometry With a Modular Drift Tube

An ambient-pressure drift tube was constructed based on previous specifications[1] and appended to the front-end of the Orbitrap Elite mass spectrometer with a 3D-printed source adapter. The module includes a 10-cm desolvation region and 10-cm drift region. Two three-grid shutters acted as the ion gates at the entrance and exit to the drift region. The gates were operated simultaneously for Fourier transform (FT) multiplexing experiments in which the gates are opened and closed using a stepped frequency waveform swept from 5 Hz to 7,005 Hz in 8 min as dictated and controlled by a custom LabVIEW program (National Instruments, Austin, TX). The gates were electrically isolated from the National Instruments controller with pulsers from GAA Custom Engineering (Benton City, WA). Extracted ion chromatograms (XICs) were generated for each charge state of HMGA2 and averaged across four replicates. Frequency spectra were generated for each *m/z* species and converted to arrival time distributions using a custom Matlab (2016a) program developed in-house.[2] ATDs were converted to collision cross section (CCS) distributions to allow direct comparison of the sizes of structures adopted by HMGA2.

Collision Cross Section Measurements based on Transient Decay Analysis in the Orbitrap Analyzer

The decay of ion signal in an Orbitrap mass analyzer is directly related to the pressure (which dictates the collisional rate) and the size/shape of ions (which determines their collisional frequency). The CCS of molecules can, therefore, be calculated from the time constant of the exponential decay of ions observed from the time-domain transient.[3] Experiments herein were carried out as described previously.[3] Briefly, after generation of ions using the nESI parameters described in the main text, ions were transferred to the HCD cell for 2 ms prior to transfer to the Orbitrap for measurement of the ion decay profiles. The transfer to the HCD cell is included to mimic the path ions take during photoactivation experiments in order to ensure the CCSs that are measured are indicative of the conformers that are probed during UVPD. Nitrogen flow into the HCD cell was optimized for readback of 0.40 x 10^{-10} torr from the Pirani ion gauge in the Orbitrap vacuum chamber, which was estimated to correspond to 0.96 x 10^{-10} torr in the Orbitrap itself. This estimation was based on the calibration using the 9+ charge state of ubiquitin, which is known to have a narrow CCS distribution, as described by Sanders *et al*.[3] Acquisition of transients was made possible by a custom license provided by Thermo Fisher Scientific. Time-domain transients, shown in **Figure S1B**, were processed using a custom Matlab script, in which a "filtered" time domain transient of each HMGA2 precursor, bound or unbound to DNA, was generated from the peaks of interest that included the top seven most abundant isotopic peaks of the desired precursor. The filtered transient was then fitted with an exponential decay function, demonstrated in **Figure S1C**. This method was employed for the 6+ through 13+ charge states of apo-HMGA2 and for the 8+ through 10+ and 12+ charge states of

the HMGA2•DNA_{22} complex. All experiments were replicated four times. The 11+ charge state of the HMGA2•DNA_{22} complex (*m/z* 1664.55) was omitted due to *m/z* overlap with the 7+ charge state of apo-HMGA2 (*m/z* 1664.03) demonstrated in **Figure 2B**. The abundance of the HMGA2•DNA_{50} complex was not sufficient for transient decay CCS analysis.

The CCS of the 6+ charge state of HMGA2 (1250 ± 30 $Å^2$) is slightly decreased relative to the $^{DT}CCS_{N2}$ value (~1390 $Å^2$, **Figure 3**), which may indicate deviation from the assumption that a single collision with the background gas in the Orbitrap analyzer is enough to destabilize an ion to induce signal decay, attributed to the decreased kinetic energy of high *m/z* ions.[3] Instances in which the $^{TDA}CCS_{N2}$ results are greater than the $^{DT}CCS_{N2}$ values, as is the case for the 8+ charge state apo-HMGA2, may result from protein unfolding occurring during transmission through the mass spectrometer. However, because increased $^{TDA}CCS_{N2}$ values relative to $^{DT}CCS_{N2}$ results are not observed across all charge states, the ion optics are believed to be sufficiently cool to prevent extensive protein unfolding. Instead, the deviation in CCS for 8+ apo-HMGA2 is believed to be a result of accelerated decay in the Orbitrap analyzer caused by the presence of a low abundance species with a more extended CCS relative to the abundant conformer at ~1540 $Å^2$. Transient decay measurements can only report a single CCS value, which may not necessarily reflect an average CSS of the distribution. In the case of an analyte with two distinct conformations, the more extended conformation will decay more rapidly in the Orbitrap analyzer; however, as this conformer is depleted, the more compact conformer that remains may decay at and artificially faster rate due to lack of ion coherence, resulting in an overall accelerated decay rate that is not representative of the average CCS of the distribution.

The $^{TDA}CCS_{N2}$ values of HMGA•DNA_{22} fall between 1620 and 1970 $Å^2$, which is fairly compact relative to the broad distribution adopted by HMGA2 alone (1400 to 3000 $Å^2$) and considering the additional mass of the hairpin. Comparison to $^{DT}CCS_{N2}$ results was not possible to due insufficient ion abundance. The low abundances of the HMGA2•DNA_{50} complexes and prohibited both $^{TDA}CCS_{N2}$ and $^{DT}CCS_{N2}$ measurements. Additionally, the high *m/z* and increased order (i.e., degrees of freedom) of the complex restrict accurate TDA as a single collision is insufficient to remove the ion from its coherently orbiting packet in the Orbitrap.[3]

Electron Transfer Dissociation of HMGA2 and HMGA2•DNA Complexes

Electron transfer dissociation (ETD) involves exothermic electron transfer between peptide or protein cations and reagent anions resulting in backbone cleavage to produce *c*- and *z*-ions. Noncovalent interactions and post-translational modifications (PTMs) remain intact during ETD, which is beneficial for structural characterization and PTM localization.[4] To explore its utility to provide complementary

structural information to UVPD, ETD was performed for two representative charge states of apo-HMGA2, 8+ and 12+, shown in **Figure S11**, in which charge reduction is the dominant activation product. The reagent ion (fluoranthene anions) was accumulated in the high-pressure cell of the dual linear ion trap and allowed to react with the isolated 8+ or 12+ precursor cations of HMGA2 for 20 to 100 ms. ETD using a 100 ms activation time yielded maximum sequence coverage for the 12+ and 8+ charge states of apo-HMGA2, amounting to an average of 52 ± 2% and 34 ± 2% sequence coverage, respectively. The locations of backbone cleavages that result in ETD fragment ions of HMGA2 and their relative abundance in the mass spectrum are shown in **Figure S12C,D** with comparison to equivalent UVPD results in **Figure S12A,B**. Because backbone cleavages promoted by ETD are more localized to surface exposed regions of proteins, it is tempting to make inferences of protein structure; that is, the additional backbone cleavages going from the compact, 8+ protein to the more extended 12+ protein may suggest that new regions of the protein are becoming exposed as the protein unfolds. However, the charge dependent nature of ETD confounds the ability to make these structural inferences with confidence.

ETD was also performed for HMGA2•DNA complexes, shown in **Figure S13**. Charge reduction of the precursor complex is again prevalent. Sparse fragmentation was achieved for the 9+ charge state of the HMGA2•DNA_{22} complex (**Figure S14A)**, resulting in an average of 13 ± 2% sequence coverage of the protein, mainly limited to the N-terminus. A total of three holo-fragments (protein sequence ions retaining the entirety of the DNA hairpin) were identified from replicate ETD experiments, including one *c*-ion (c_{105}) and two *z*-ions (z_{103} and z_{106}) shown in **Figure S14B**. The few holo-fragments support the UVPD results but does not provide additional insight into DNA-binding. ETD of the 10+ charge state of the HMGA2•DNA_{50} complex resulted in only charge reduction with no observed sequence ions (**Figure S13B**). Owing to the limited sequence coverage provided by ETD, it was not highlighted extensively in the manuscript.

UVPD of DNA in the Positive Mode

The laser fluence for all UVPD experiments was maintained at 3 mJ per pulse for all HMGA2 and HMGA2•DNA complexes in all charge states. However, the difference in photoabsorptivity of a protein like HMGA2 versus DNA warranted consideration of the impact of their chromophores on UVPD. In proteins, photoabsorptivity scales directly with the number of amide chromophores and, owing to the established charge independence of UVPD, there is no normalization of laser fluence based on the charge state of a single protein, unlike collisional activation that must be normalized for protein charge and degrees-of-freedom owing to the prevalence of vibrational energy redistribution.[5–7] One measure of photoabsorptivity for proteins during UVPD is precursor depletion, which represents the decrease in the abundance of the protein precursor owing to all photoactivation processes including dissociation and

charge stripping. To that end, the DNA hairpins were sprayed in positive mode using the nESI methods described for analysis of HMGA2 and HMGA2•DNA complexes. To estimate the photoabsorptivity of the DNA, the precursor depletion was calculated from the UVPD mass spectra of apo-DNA in **Figure S15**. As shown in **Figure S15**, DNA itself absorbs 193 nm photons and undergoes UVPD, although not quite as efficiently as proteins based on the precursor ion depletion measurements shown in **Figure S16**. Precursor depletion for DNA_{22} and DNA_{50} using 3 mJ UVPD 25% and 15%, respectively, as shown in **Figure S16**. For reference, the average precursor depletion of the apo-protein using 1 pulse, 3 mJ UVPD was 54 ± 7% for the 7+ through 12+ charge states. Given that the DNA demonstrated considerable absorptivity of 193 nm photons (same order of magnitude), UVPD of the HMGA2•DNA complexes are comparable to UVPD of the apo-protein. Additionally, it is not known how much the higher order structure of the hairpin limits photodissociation of the DNA, making any efforts to normalize laser fluence futile. As most UVPD of proteins occurs by production of *a/x* ions via direct dissociation from excited electronic states, and because the magnitude of photoabsorptivity of DNA is similar to that of the protein, specific steps to correct for intramolecular vibrational energy redistribution among all degrees of freedom were not taken.

References

(1) Reinecke, T.; Clowers, B. H. Implementation of a Flexible, Open-Source Platform for Ion Mobility Spectrometry. *HardwareX* **2018**, *4*, e00030. https://doi.org/10.1016/j.ohx.2018.e00030.

(2) Sanders, J. D.; Shields, S. W.; Escobar, E. E.; Lanzillotti, M. B.; Butalewicz, J. P.; James, V. K.; Blevins, M. S.; Sipe, S. N.; Brodbelt, J. S. Enhanced Ion Mobility Separation and Characterization of Isomeric Phosphatidylcholines Using Absorption Mode Fourier Transform Multiplexing and Ultraviolet Photodissociation Mass Spectrometry. *Anal. Chem.* **2022**, *94* (10), 4252–4259. https://doi.org/10.1021/acs.analchem.1c04711.

(3) Sanders, J. D.; Grinfeld, D.; Aizikov, K.; Makarov, A.; Holden, D. D.; Brodbelt, J. S. Determination of Collision Cross-Sections of Protein Ions in an Orbitrap Mass Analyzer. *Anal. Chem.* **2018**, *90* (9), 5896–5902. https://doi.org/10.1021/acs.analchem.8b00724.

(4) Lermyte, F.; Sobott, F. Electron Transfer Dissociation Provides Higher-Order Structural Information of Native and Partially Unfolded Protein Complexes. *Proteomics* **2015**, *15* (16), 2813–2822. https://doi.org/10.1002/pmic.201400516.

(5) Shaw, J. B.; Li, W.; Holden, D. D.; Zhang, Y.; Griep-Raming, J.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. Complete Protein Characterization Using Top-Down Mass Spectrometry and Ultraviolet Photodissociation. *J. Am. Chem. Soc.* **2013**, *135* (34), 12646–12651. https://doi.org/10.1021/ja4029654.

(6) Sipe, S. N.; Brodbelt, J. S. Impact of Charge State on 193 Nm Ultraviolet Photodissociation of Protein Complexes. *Phys. Chem. Chem. Phys.* **2019**, *21* (18), 9265–9276. https://doi.org/10.1039/C9CP01144G.

(7) Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A. Mobile and Localized Protons: A Framework for Understanding Peptide Dissociation. *J. of Mass Spectrom.* **2000**, *35* (12), 1399–1406. https://doi.org/10.1002/1096-9888(200012)35:12<1399::AID-JMS86>3.0.CO;2-R.

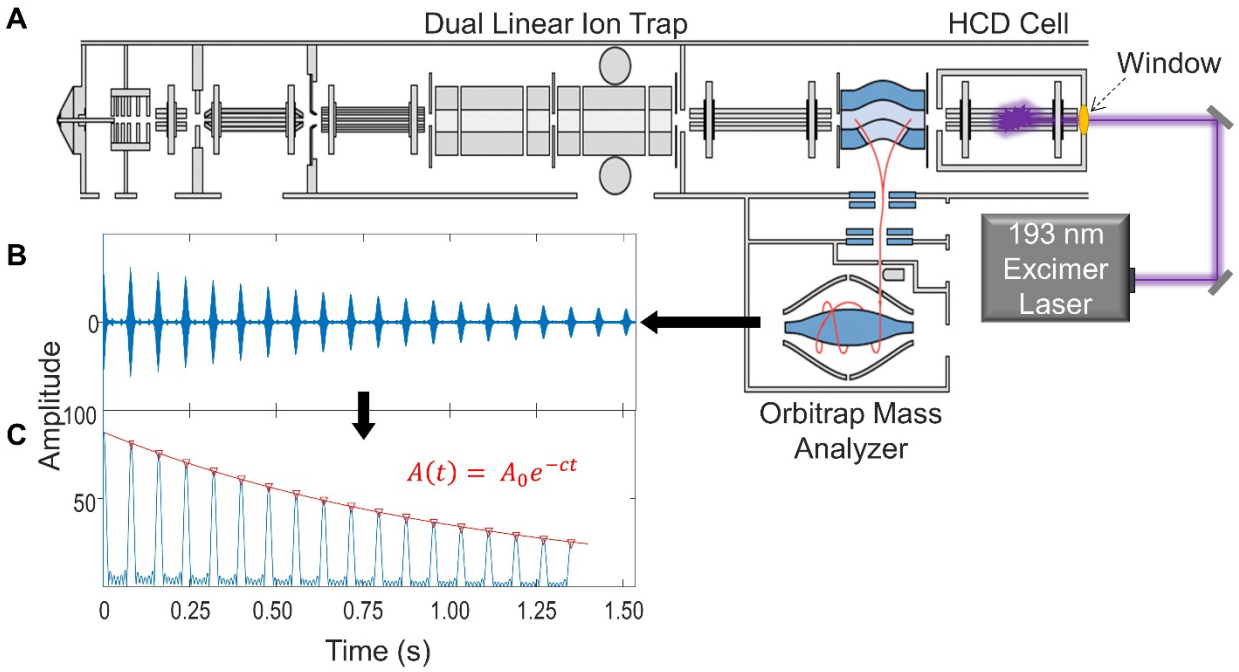

Figure S1. Schematic of A) a Thermo Scientific Orbitrap Elite mass spectrometer equipped with a 193 nm excimer laser for UVPD in the HCD cell. For CCS measurements, signal from the Orbitrap mass analyzer is saved as B) time-domain transients then C) filtered for the desired ion isotopes and fitted with an exponential decay function.

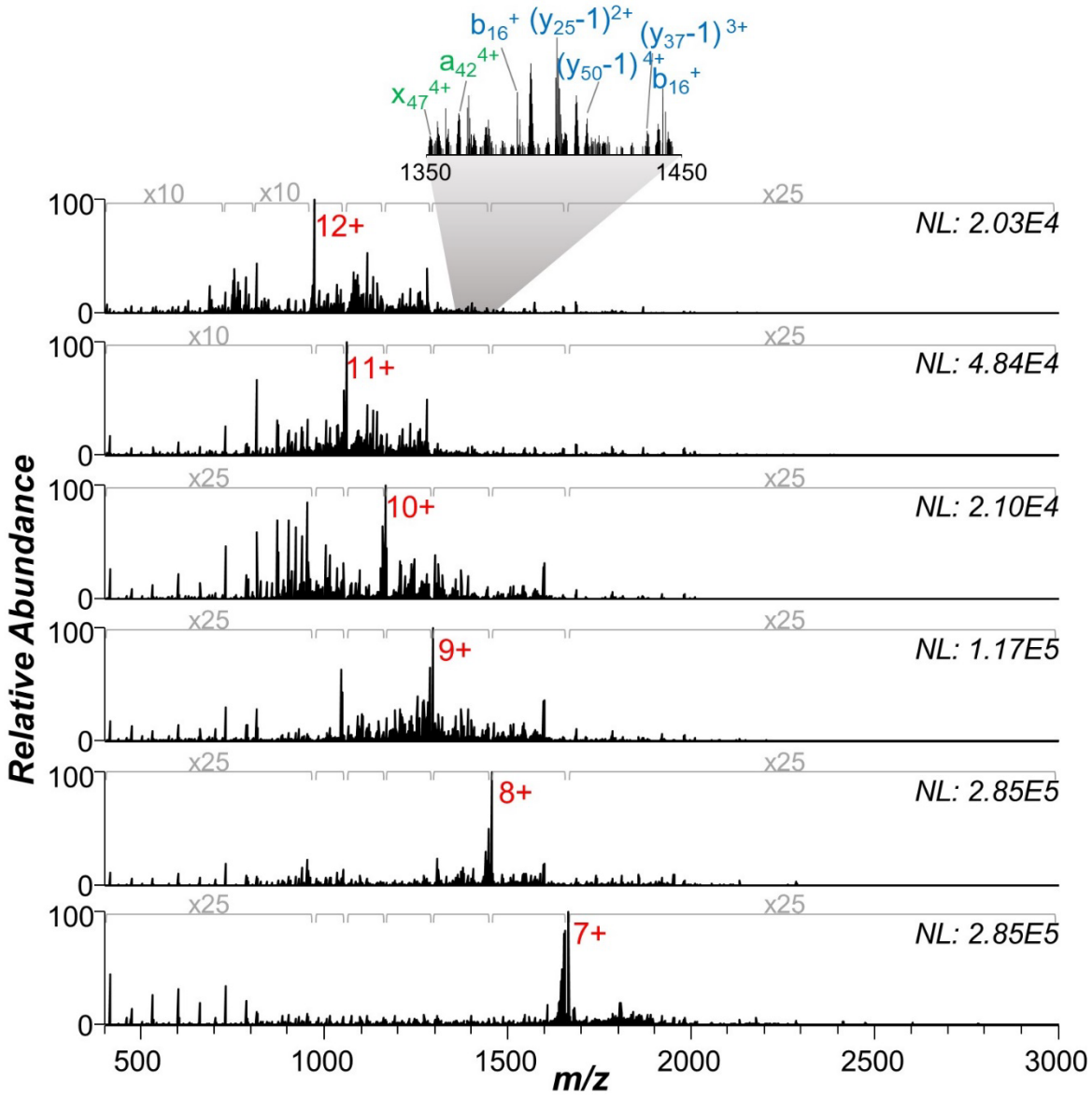

Figure S2. UVPD mass spectra for the 7+ to 12+ charge states of HMGA2 obtained using a single 3 mJ laser pulse. The inset in the top spectrum shows the rich density of fragments generated from UVPD.

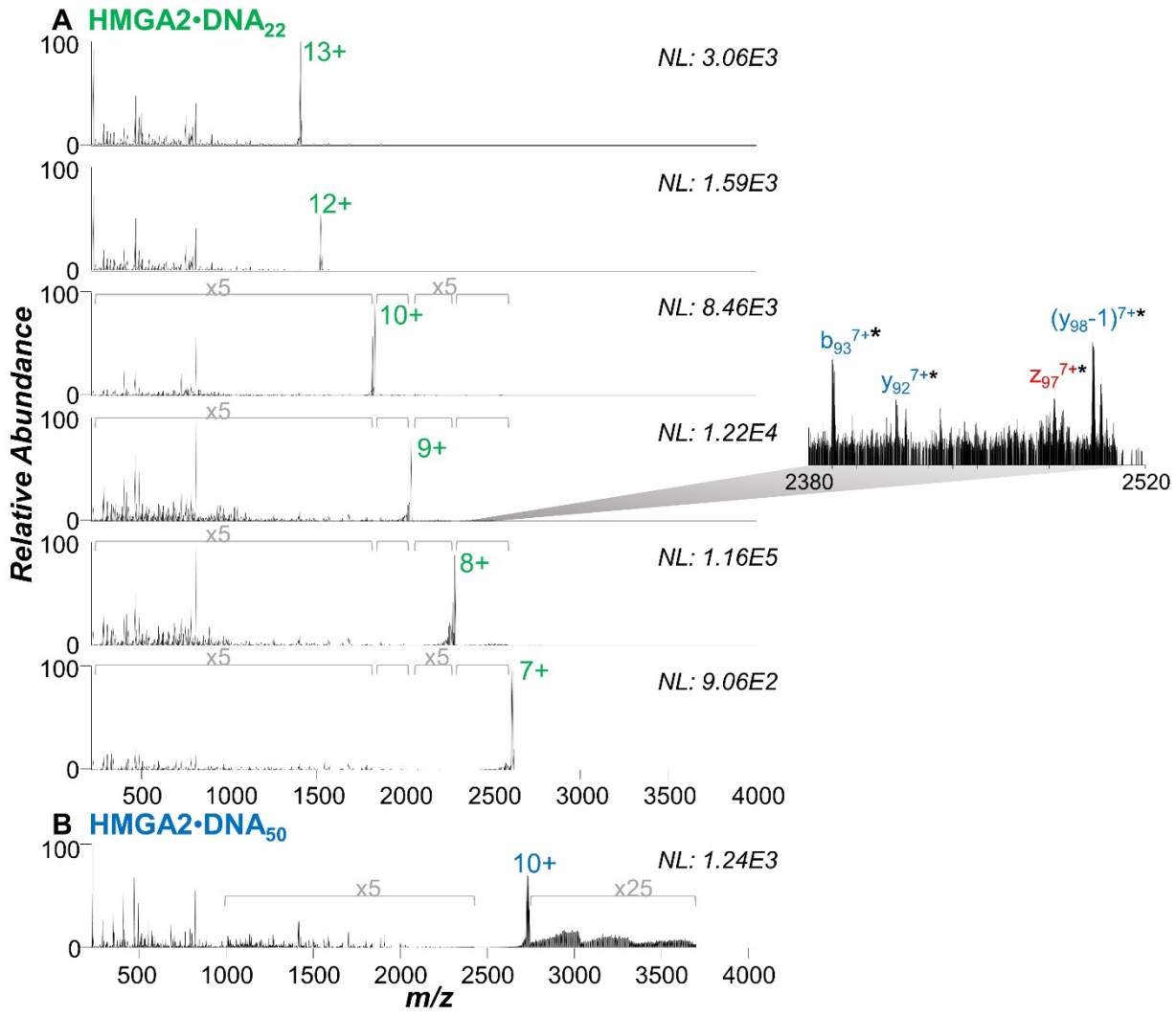

Figure S3. UVPD mass spectra for A) the 7+ to 13+ charge states of HMGA2●DNA_{22}, and B) 10+ charge states of HMGA2●DNA_{50}. All spectra were obtained using a single 3 mJ laser pulse. The 11+ charge state of HMGA2●DNA_{22} was omitted owing to its *m/z* overlap with the 7+ charge state of apo-HMGA2, preventing isolation of the complex. The inset shows labelled fragment ions with stars denoting holo-fragments (i.e. retaining the entire DNA_{22} hairpin).

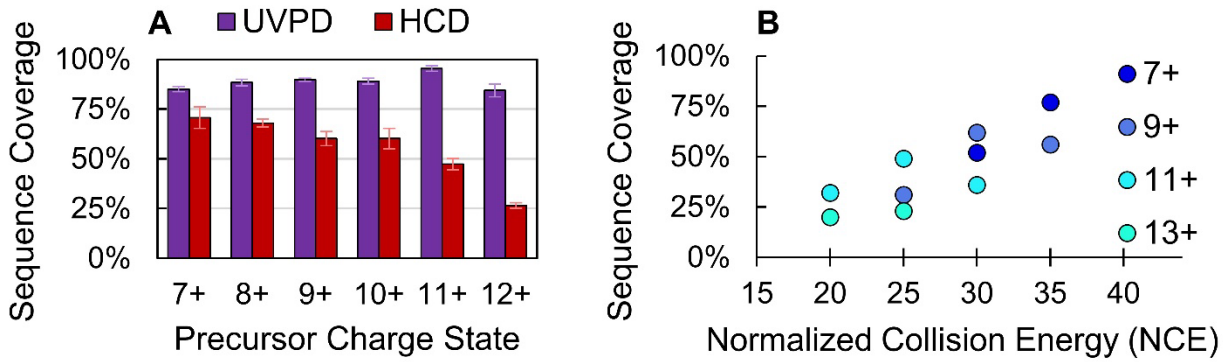

Figure S4. A) Percent sequence coverage of multiply charged HMGA2 (7+ through 12+). A single 3 mJ laser pulse was used for all UVPD experiments. HCD energy was optimized for 50% precursor depletion for each charge state as follows: 34 NCE for 7+, 34 NCE for 8+, 29 NCE for 9+, 28 NCE for 10+, 25 NCE for 11+, and 20 NCE for 12+ HMGA2. B) Sequence coverages for selected charge states of HMGA2 subjected to the same normalized collision energy settings. The sequence coverage decreases for higher charge states.

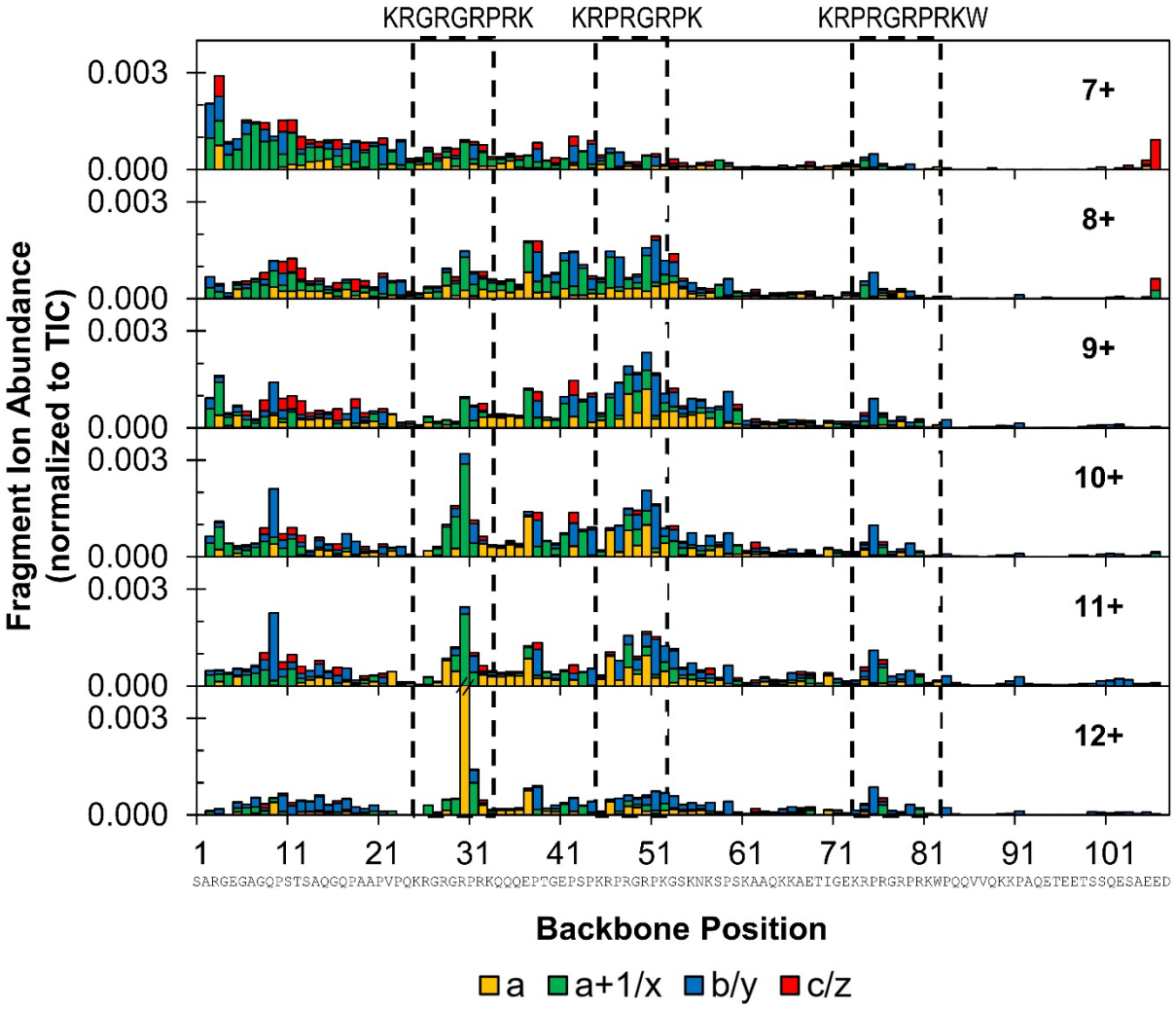

Figure S5. Distribution of backbone cleavages at each residue in the primary sequence of HMGA2 for the 7+ through 12+ charge states showing the specific contributions from a, a+1/x, b/y, and c/z ions generated from UVPD using a single 3 mJ laser pulse. Residues encompassed in dashed black boxes denote the three AT-hook regions.

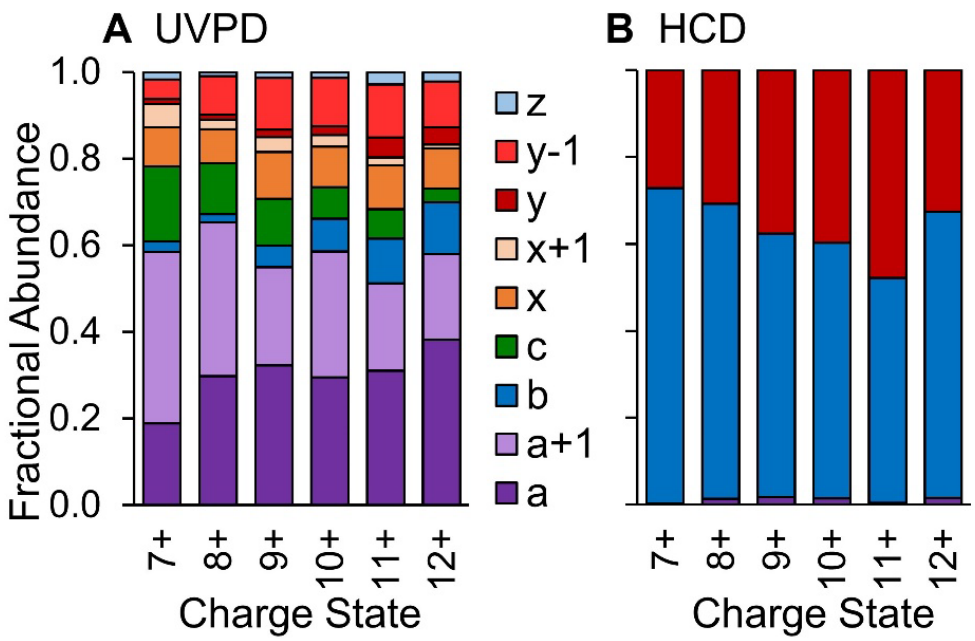

Figure S6. Fractional abundance of A) all nine ion types generated from a single pulse of 3 mJ UVPD and B) *a*/*b*/*y* ions generated from HCD, variable NCE, for each charge state of HMGA2. An increase in fractional abundance of IVR-type fragments (*a*/*b*/*y* ions) is observed from UVPD for increasing precursor charge. HCD energy for each charge state as follows: 34 NCE for 7+, 34 NCE for 8+, 29 NCE for 9+, 28 NCE for 10+, 25 NCE for 11+, and 20 NCE for 12+ HMGA2.

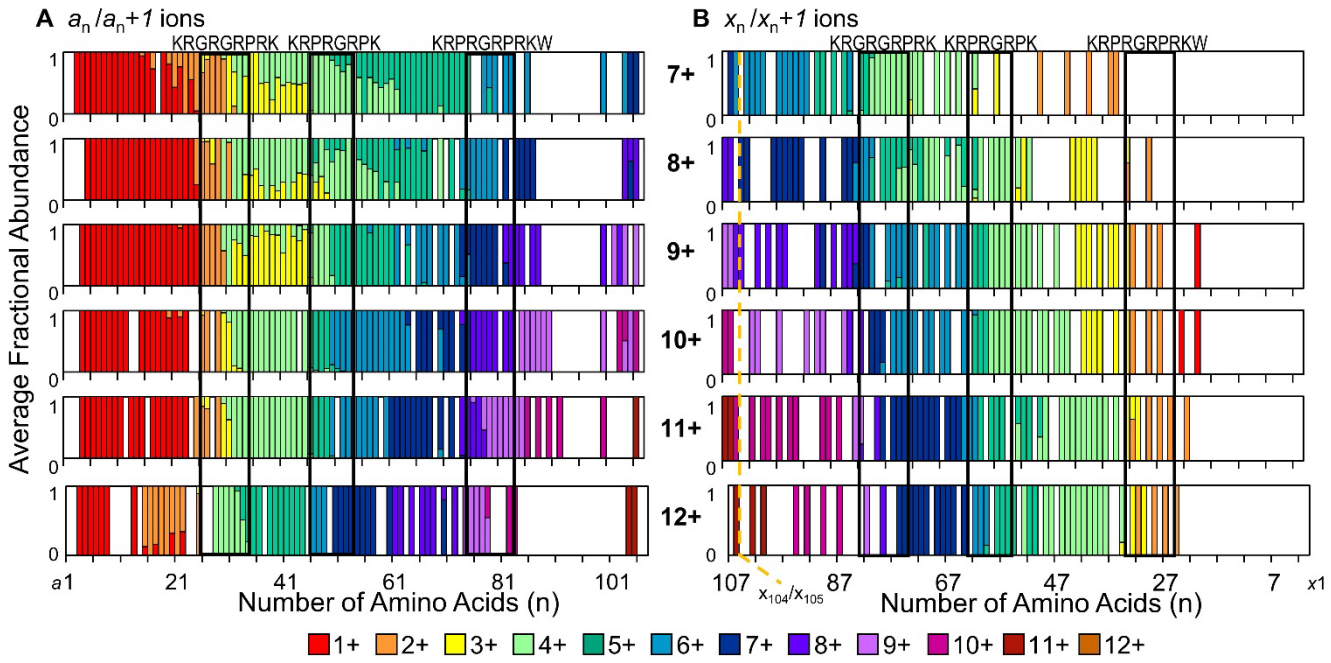

Figure S7. Charge state analysis of a/x ions based on normalized fractional abundance of the charge states of: A) a_n ions (all summed a_n and a_n+1 ions) and B) x_n ions (all summed x_n and x_n+1) produced by UVPD of HMGA2 (7+ through 12+ charge states) using a single 3 mJ laser pulse, where n is the number of amino acids contained in the fragment ion. The resulting charge site locations are summarized in **Figure S9**. Residues encompassed in black boxes denote the three AT-hook regions. Tailing is observed in the 7+, 8+, and 9+ charge states of HMGA2 as sequence ions detected in more than one charge state (*e.g.*, overlap of the 3+ and 4+ charge states of a_{31} to a_{46} ions).

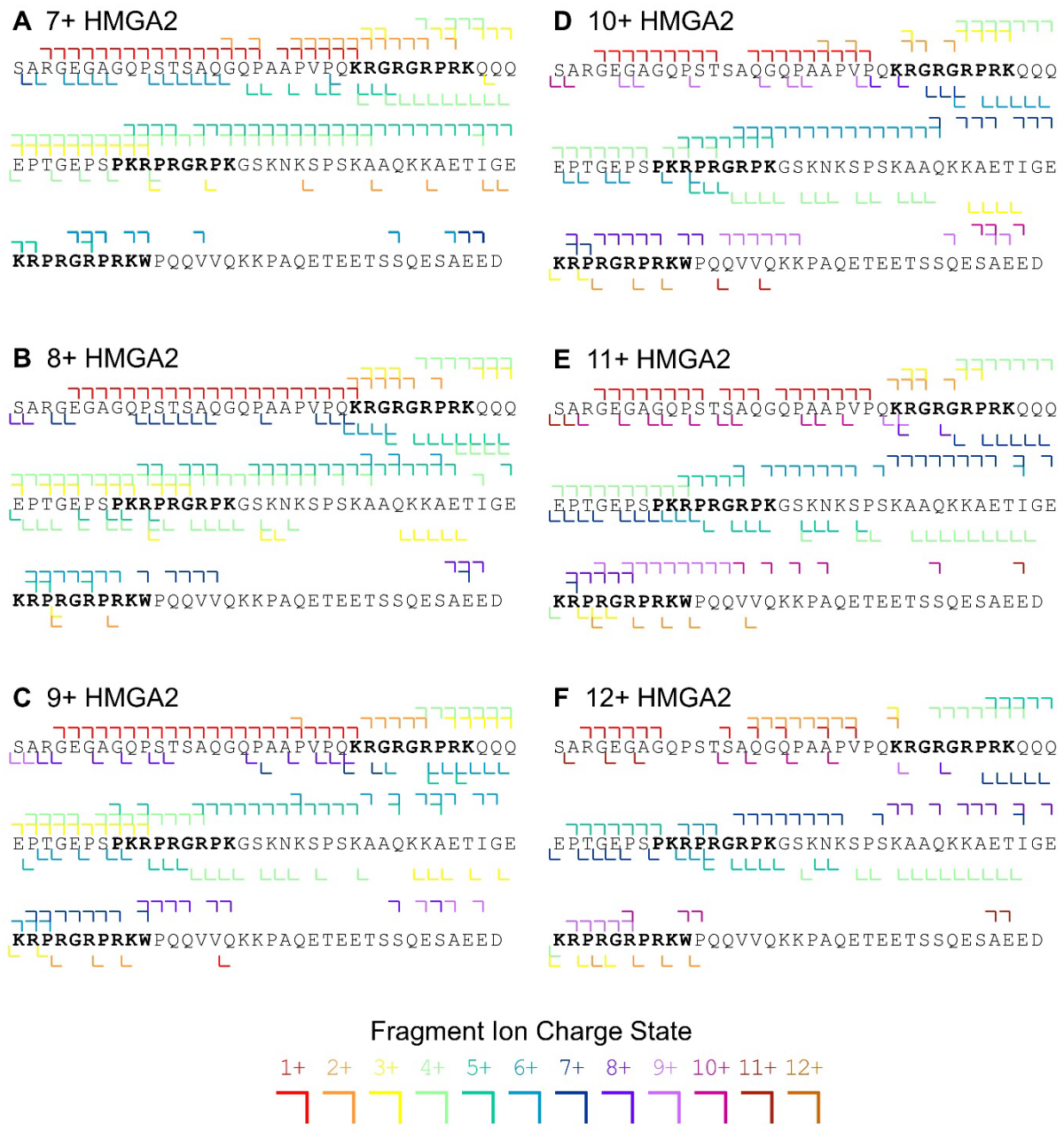

Figure S8. Locations of backbone cleavages along the primary sequence of HMGA2 and their corresponding charge states generated from 1 pulse, 3 mJ UVPD of multiply charged apo-protein. N-terminal fragments include *a* and *a+1* ions (left facing flags) and C-terminal fragments include *x* and *x+1* ions (right-facing flags). Residues in bolt font indicate the AT-hook regions.

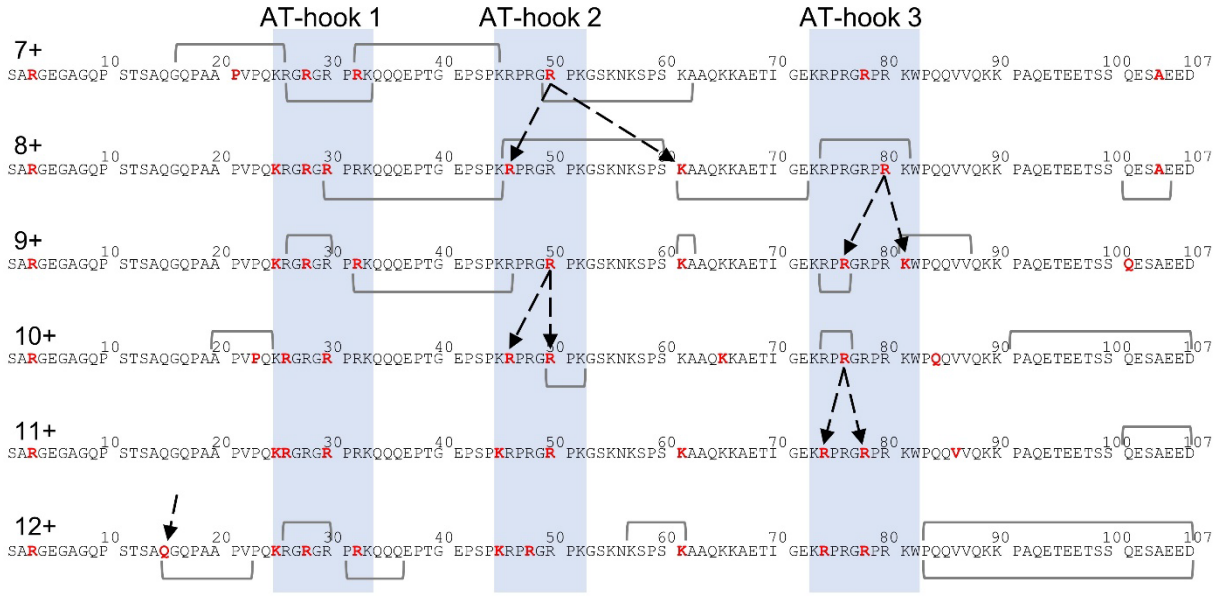

Figure S9. Summary of charge site analysis of multiply charged apo-HMGA2 (7+ through 12+) using one 3 mJ pulse for UVPD showing suggested protonation sites in red with brackets indicating regions where a proton may be mobile or indicative of more than one conformer/proton configuration. The presence of double brackets for the 12+ charge state represents two unlocalized protons on the C-terminus of the protein resulting from low fragment abundance. Arrows denote potential charge migration from one charge state to the next.

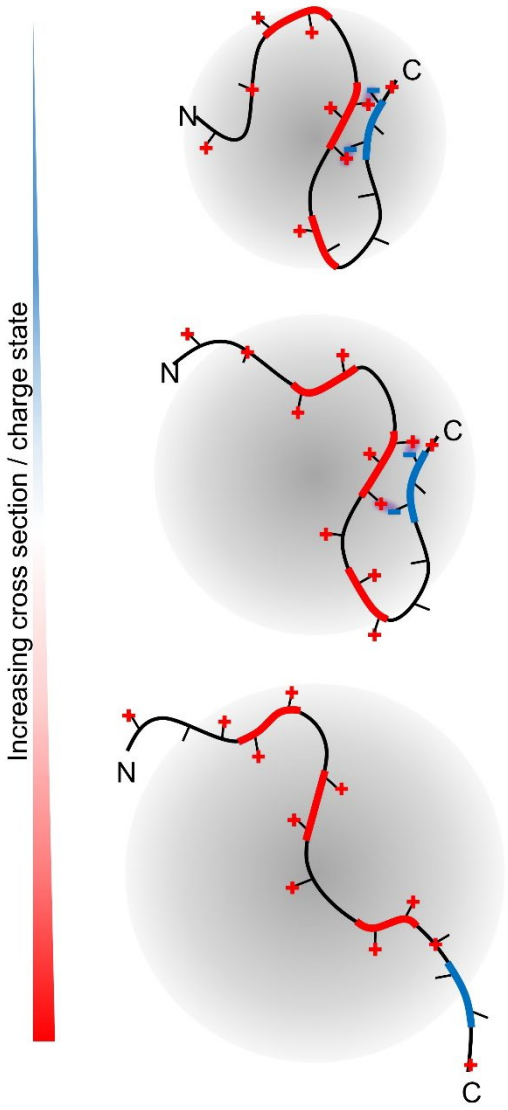

Figure S10. Proposed schematic of the unfolding of HMGA2 from a compact conformer with few charges (top) to an extended conformer with many charges (bottom). The N-terminus is believed to be exposed and/or more flexible (fewer stabilizing non-covalent interactions) based on elevated fragment ion abundance in this region (most notably in the 7+ charge state). The red regions along the backbone denote AT-hooks with the red charges (+) indicating protonated residues. The blue region along the backbone symbolizes the acidic C-terminal tail with blue charges (-) indicating possible negatively charge residues that may form salt bridges (purple). The salt bridges shown between the C-terminus and AT-hook 2 are one possibility of many (*e.g.*, C-terminus/AT-hook 3, *etc.*). The gray circles demonstrate the increasing size of the collision cross-sections with increased charge. This is just one set of proposed structures while many others are possible and likely exist in equilibrium.

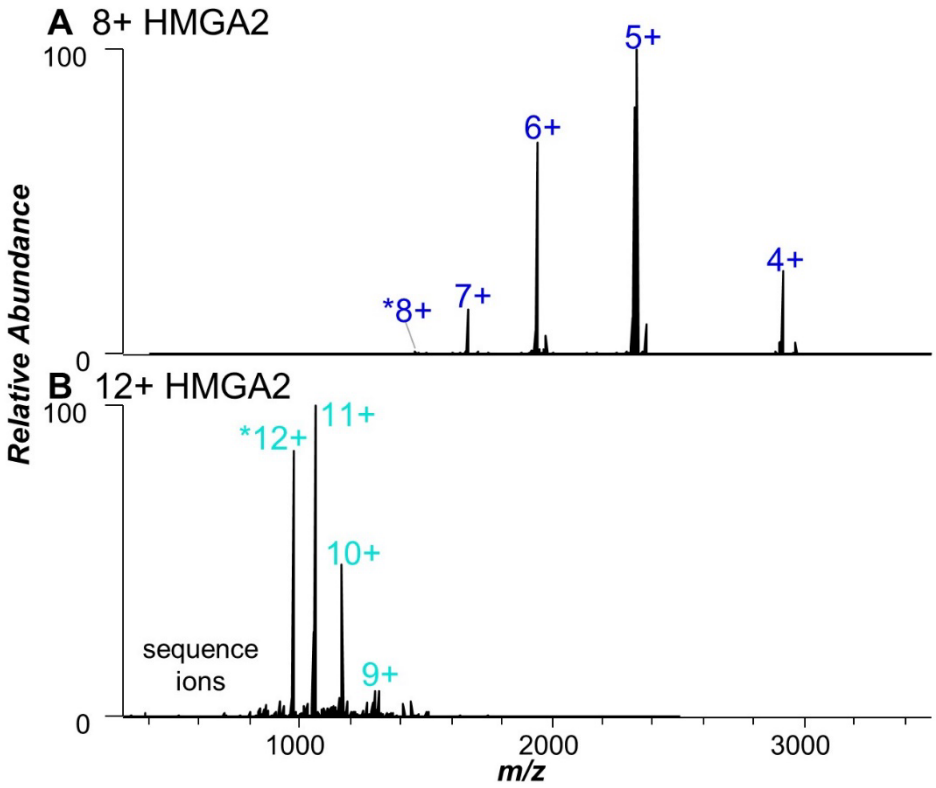

Figure S11. ETD mass spectra of A) HMGA2 (8+ charge state) and B) HMGA2 (12+ charge state) utilizing a 100 ms activation period.

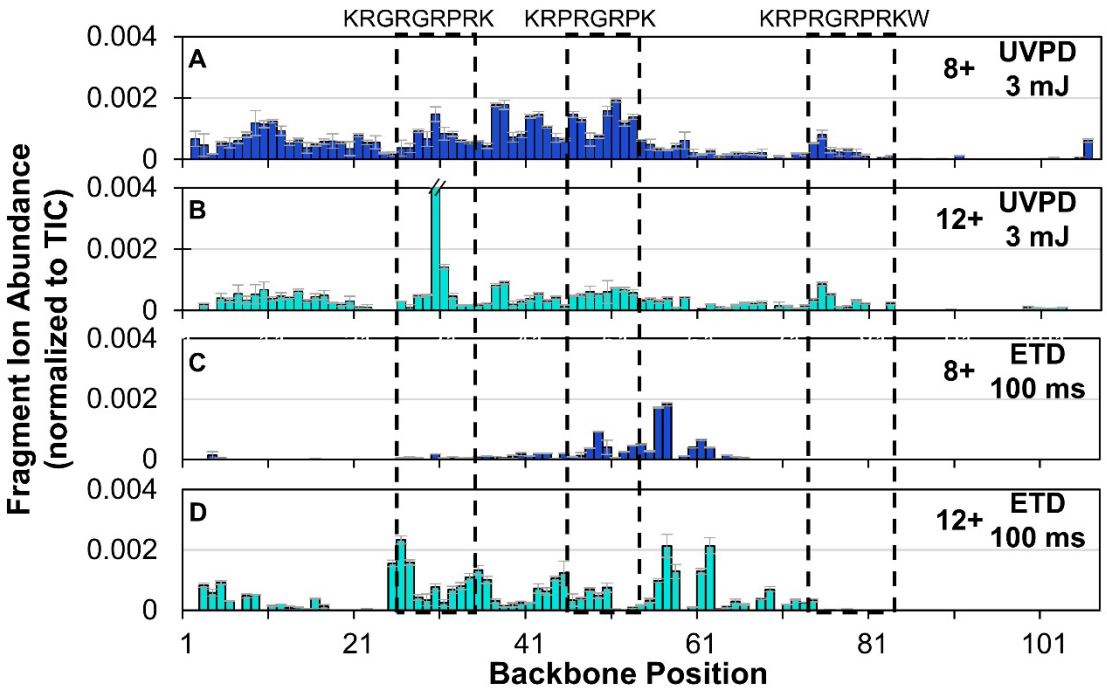

Figure S12. Sequence maps of HMGA2 for A,C) 8+ charge state and B,D) 12+ charge state activated using A,B) 1 pulse, 3 mJ UVPD or C,D) 100 ms ETD.

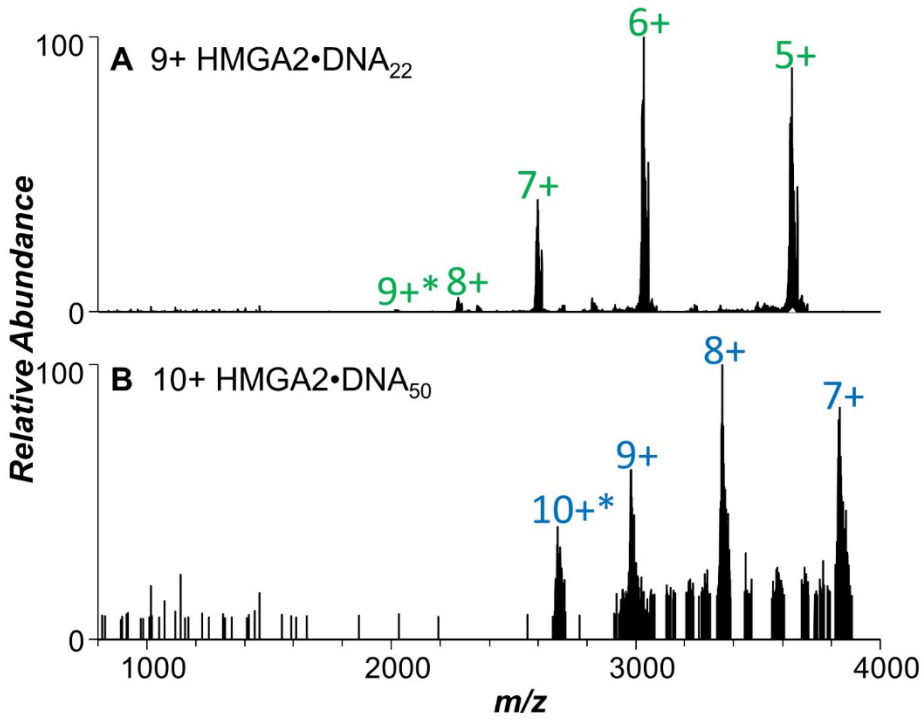

Figure S13. ETD mass spectra of A) HMGA2•DNA_{22} (9+ charge state) utilizing a 50 ms activation period and B) HMGA2•DNA_{50} (10+ charge state) utilizing a 20 ms activation period. Asterisks denote precursor ions.

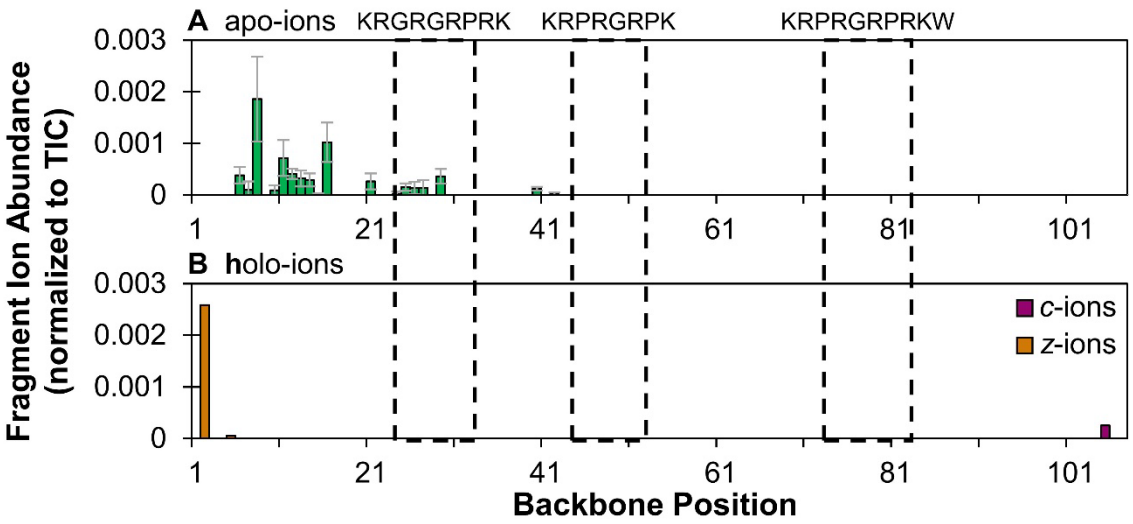

Figure S14. Sequence maps of the A) apo-fragments and B) holo-fragments generated from ETD of HMGA2•DNA_{22} (9+ charge state) using a 100 ms activation period. The dashed black boxes encompass the locations of the AT-hook peptides. Error bars are equivalent to the standard deviation of replicate measurements

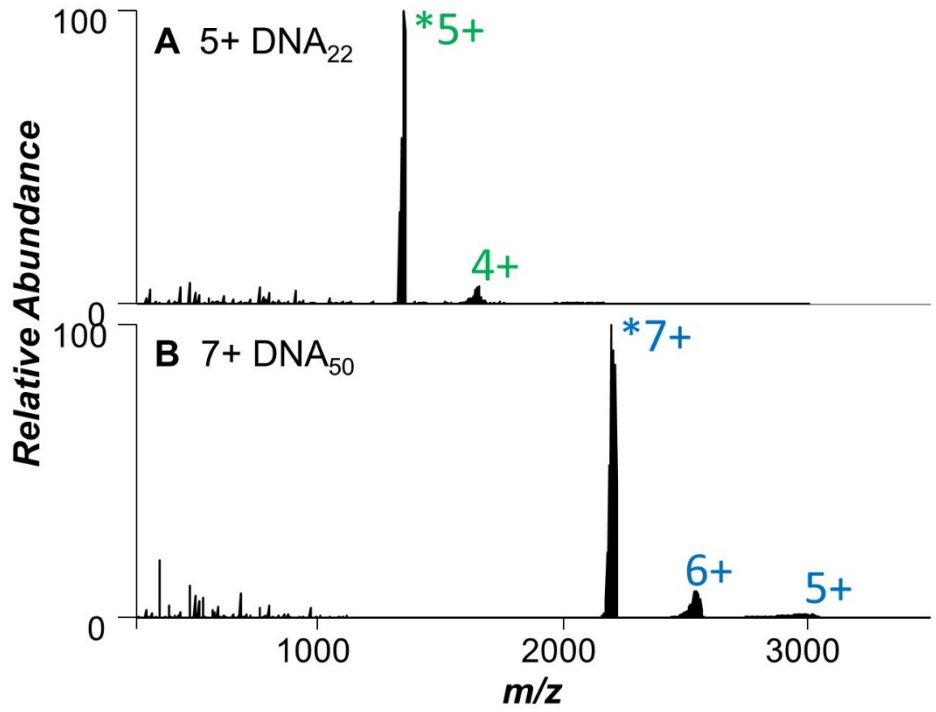

Figure S15. UVPD mass spectra of A) DNA_{22} (5+ charge state) and B) DNA_{50} (7+ charge state) utilizing 1 pulse, 3 mJ. The selected precursor is labelled with an asterisk.

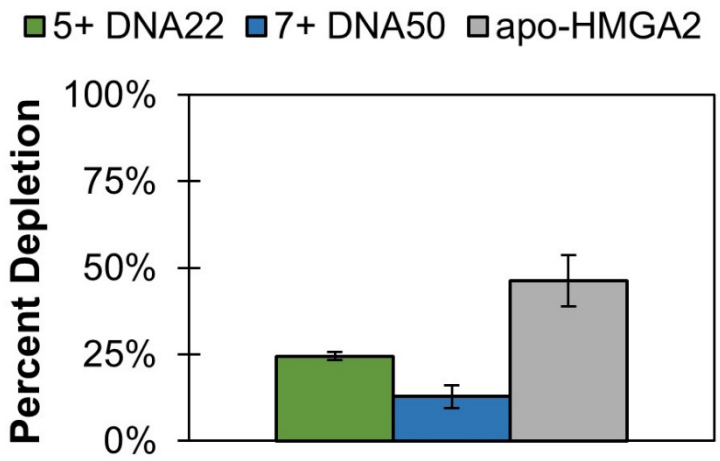

Figure S16. Percent depletion of the precursor ion following UVPD (1 pulse, 3 mJ) of DNA_{22}, (5+), DNA_{50} (7+), and all charge states (7+ through 12+) of apo-HMGA2, averaged. Error bars equal to standard deviation.

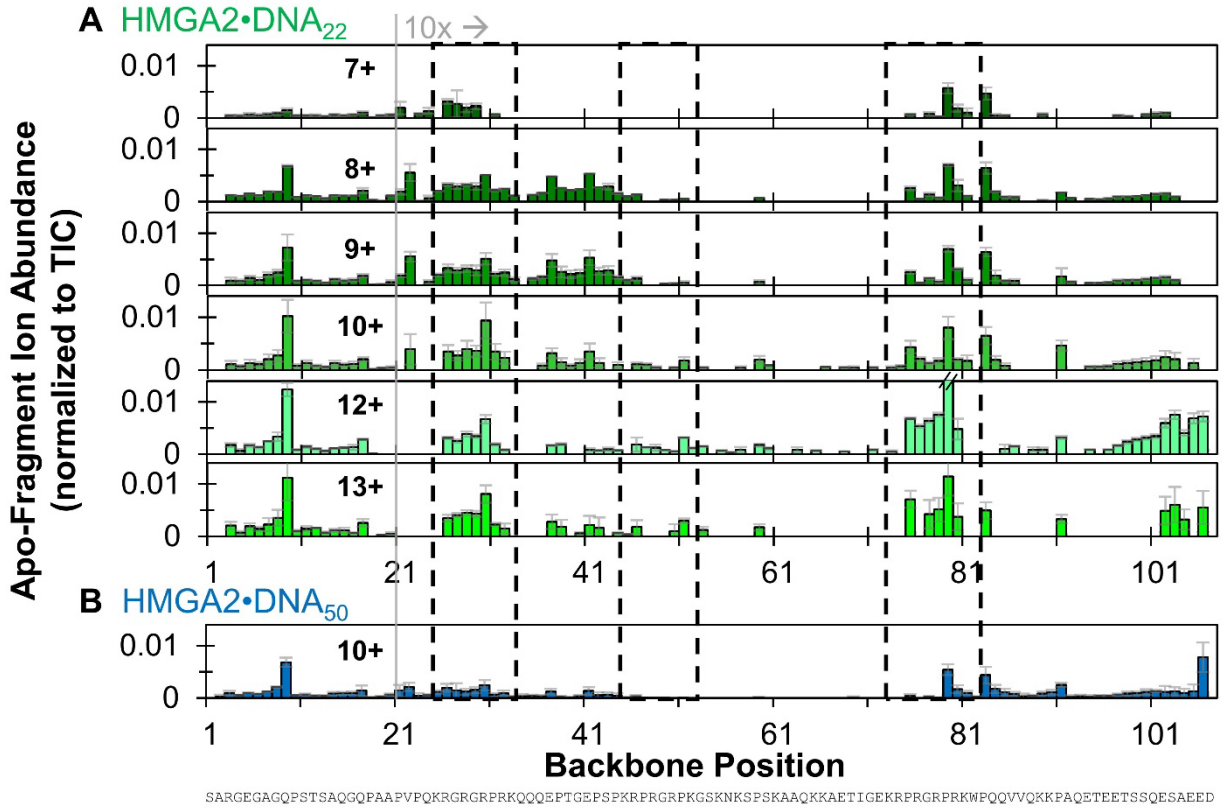

Figure S17. Distribution of backbone cleavages at each residue in the primary sequence of A) HMGA2•DNA_{22} (7+ through 13+) and B) the HMGA2•DNA_{50} complex (10+) based on sum of N- and C-terminal apo fragment ions (ones that do not retain the DNA) generated by UVPD using a single 3 mJ laser pulse. Residues encompassed in dashed black boxes denote the three AT-hook regions. Error bars are equivalent to the standard deviation of replicate measurements. The abundances between residues 21 and 107 are magnified 10-fold. The sparse fragmentation and low abundance of sequence ions generated from the 7+ charge state of HMGA2•DNA_{22} complex is attributed to the low abundance of the precursor and subsequent poor signal-to-noise of fragment ions (**Figure S3A)**. The corresponding distributions of backbone cleavages resulting in holo fragment ions (ones that retain the DNA) are shown in **Figure 5**. The abundance scale in the distributions shown here relative to those in **Figure 5** indicates that the apo fragment ions are generally >10X more abundant that the holo fragment ions.

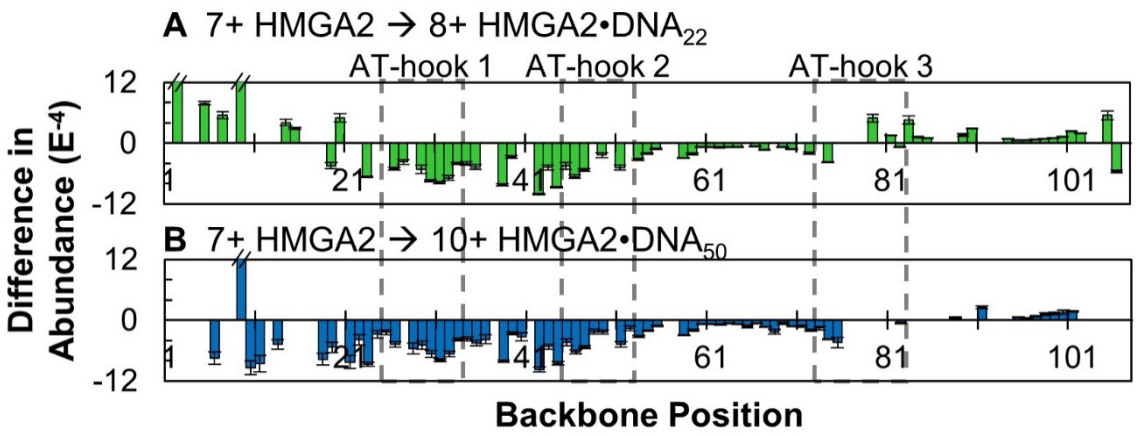

Figure S18. Difference in normalized abundances of sequence ions arising from backbone cleavages generated upon UVPD of HMGA2 (7+) relative to all UVPD sequence ions (i.e. the sum of apo- and holo-fragments) of A) HMGA2•DNA_{22} (8+) and B) HMGA2•DNA_{50} (10+). Positive values indicate regions of enhanced backbone cleavages for the HMGA2•DNA complex relative to the apo-protein; negative values indicate a suppression of backbone cleavages. Residues encompassed in dashed black boxes denote the three AT-hook regions. Only differences that are statistically different at a 98% confidence level are shown and insignificant difference are omitted, represented as "zero" values. Overall, trends are consistent with those depicted in Figure 6 with suppressed fragmentation through the center of the protein and enhavned fragmentation of the termini following binding to both DNA hairpins.

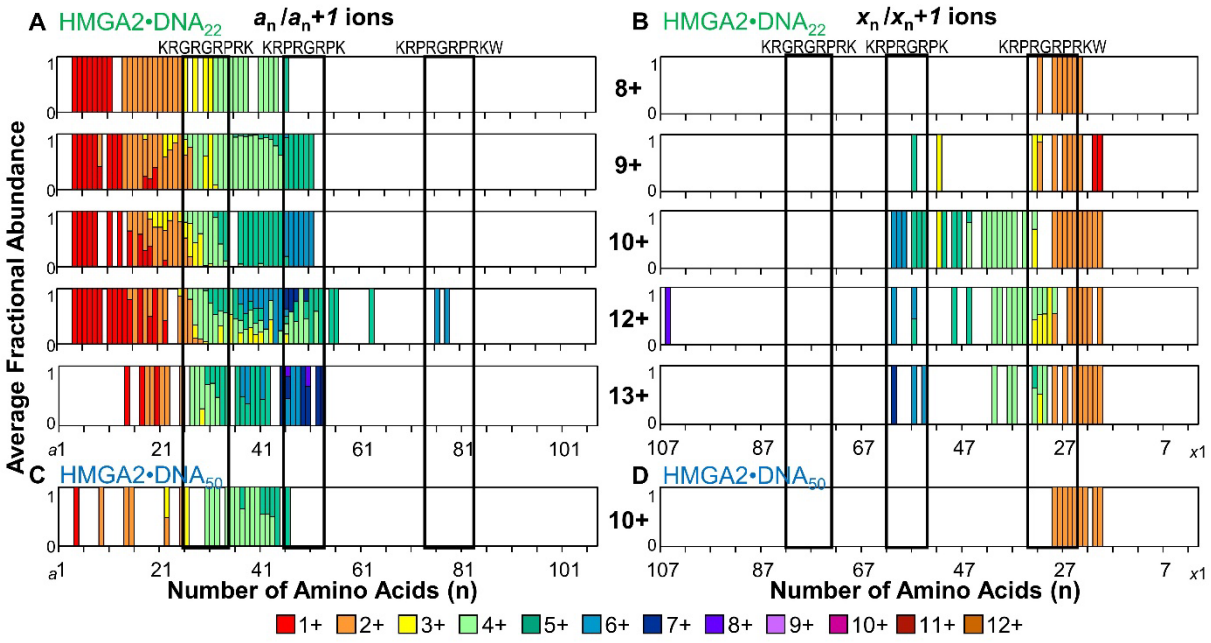

Figure S19. Normalized fractional abundance of A,C) a_n and a_n+*1* and B,D) x_n and x_n+*1* fragments of multiply charged HMGA2•DNA$_{22}$ (A,B) and HMGA2•DNA$_{50}$ (C,D) generated by UVPD using a single 3 mJ laser pulse. Data is summarized in **Figure S19**. Residues encompassed in black boxes denote the three AT-hook regions.

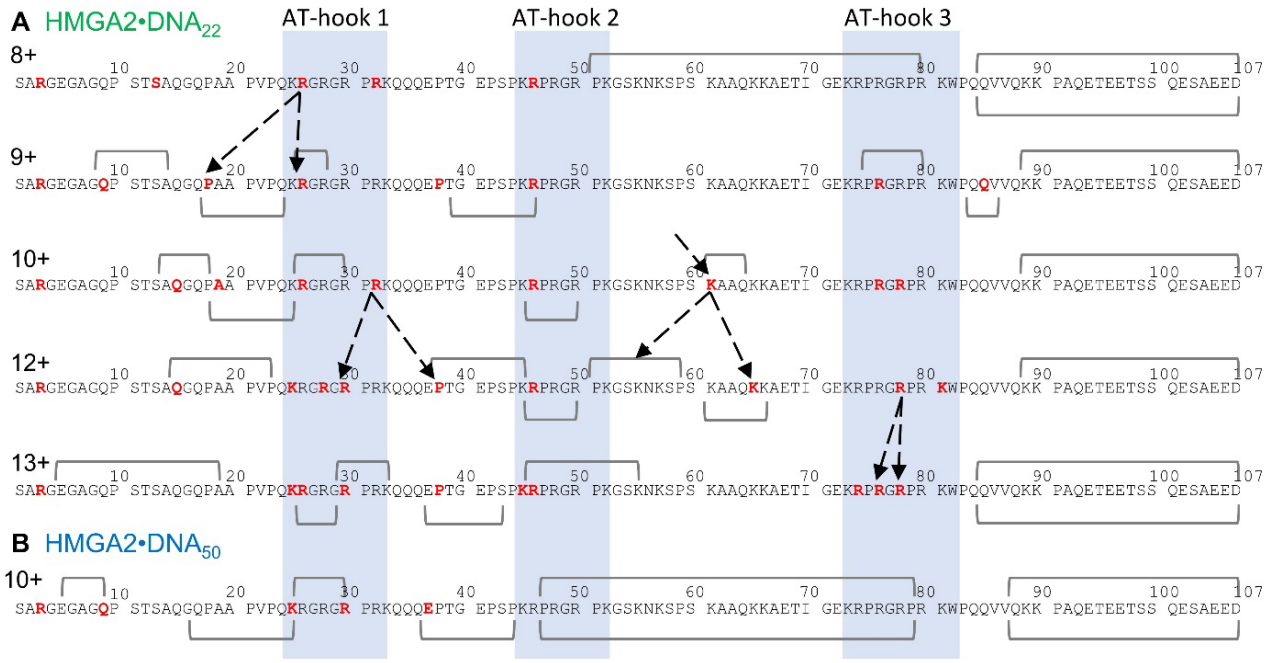

Figure S20. Summary of charge site analysis of multiply charged A) HMGA2•DNA_{22} (7+ through 13+) and B) HMGA2•DNA_{50} (10+) upon UVPD using one 3 mJ laser pulse showing suggested protonation sites in red with brackets indicating regions with insufficient fragmentation, or where a proton may be mobile or indicative of more than one conformer/proton configuration. The presence of double brackets for the 10+ charge state of HMGA2•DNA_{50} and the 8+ and 13+ charge states of HMGA2•DNA_{22} indicate uncertainty in two proton locations owing to insufficient fragment abundance. Arrows in (A) denote potential charge migration from one charge state to the next.