Supplementary Figures for:
CIDer: a statistical framework for interpreting differences in CID and HCD fragmentation

Damien B. Wilburn[1,2] Alicia L. Richards[3,4,5], Danielle L. Swaney[3,4,5], and Brian C. Searle[1,*]

[1] Institute for Systems Biology, Seattle, WA, USA
[2] Department of Genome Sciences, University of Washington, Seattle, WA, USA
[3]University of California San Francisco, Quantitative Biosciences Institute (QBI), San Francisco, CA, 94158, USA.
[4]J. David Gladstone Institutes, San Francisco, CA 94158, USA.
[5]University of California San Francisco, Department of Cellular and Molecular Pharmacology, San Francisco, CA, 94158, USA.

SUPPORTING INFORMATION:

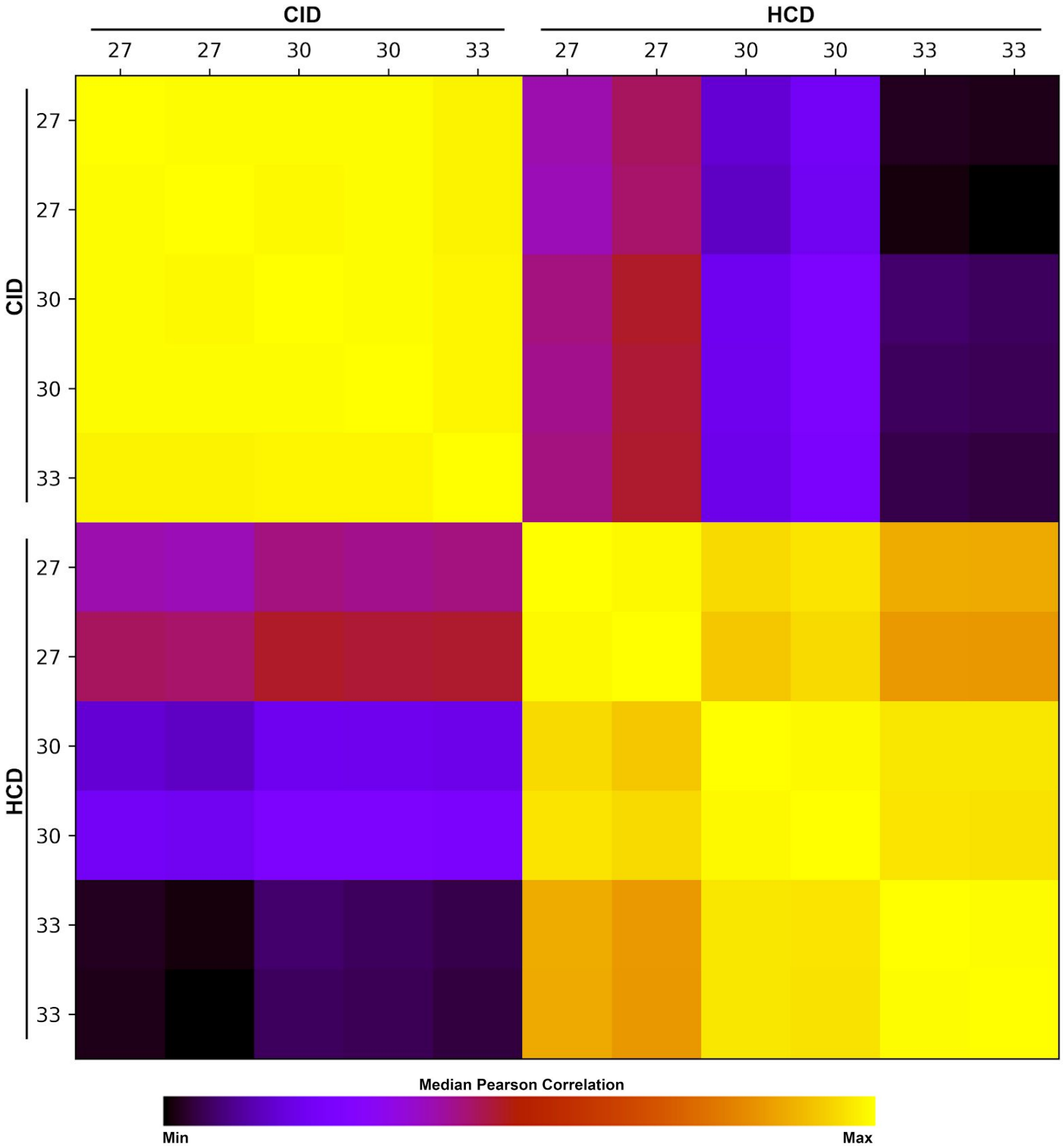

Figure S1. Comparison of HCD and CID spectra correlation by NCE. Heatmap of median Pearson correlation coefficients between spectra from multiple independently acquired datasets on the same human cell line sample, fragmented by both HCD and CID at multiple NCE settings (normalized to min/max for visualization purposes). Evaluation Dataset 1 was used to generate this figure.

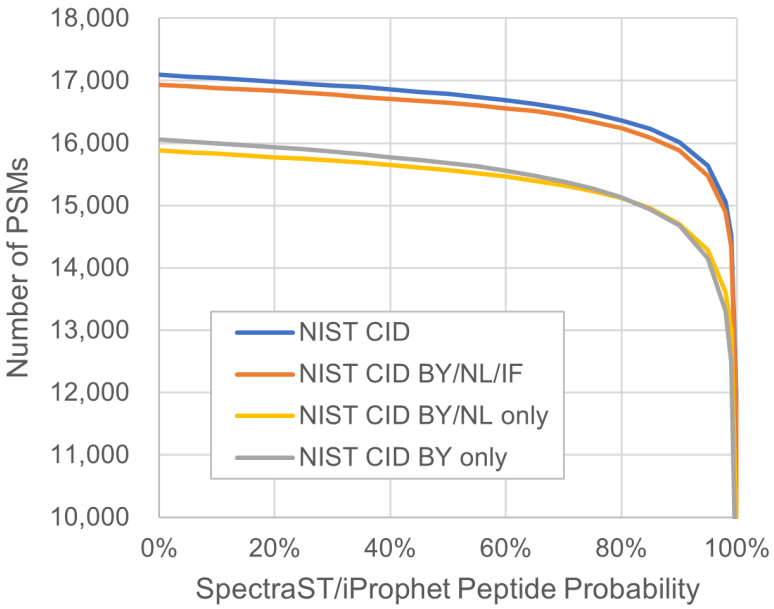

Figure S2. Comparison of library filtering on the number of detected PSMs. Evaluation Dataset 4 was searched with SpectraST using the NIST CID library, either unfiltered (NIST CID), or filtered to retain b- and y-type ions (BY), a-type ions and $-H_2O$ and $-NH_3$ neutral losses (NL), and internal fragments and other annotated ions (IF). Evaluation Dataset 2 was used to generate this figure.

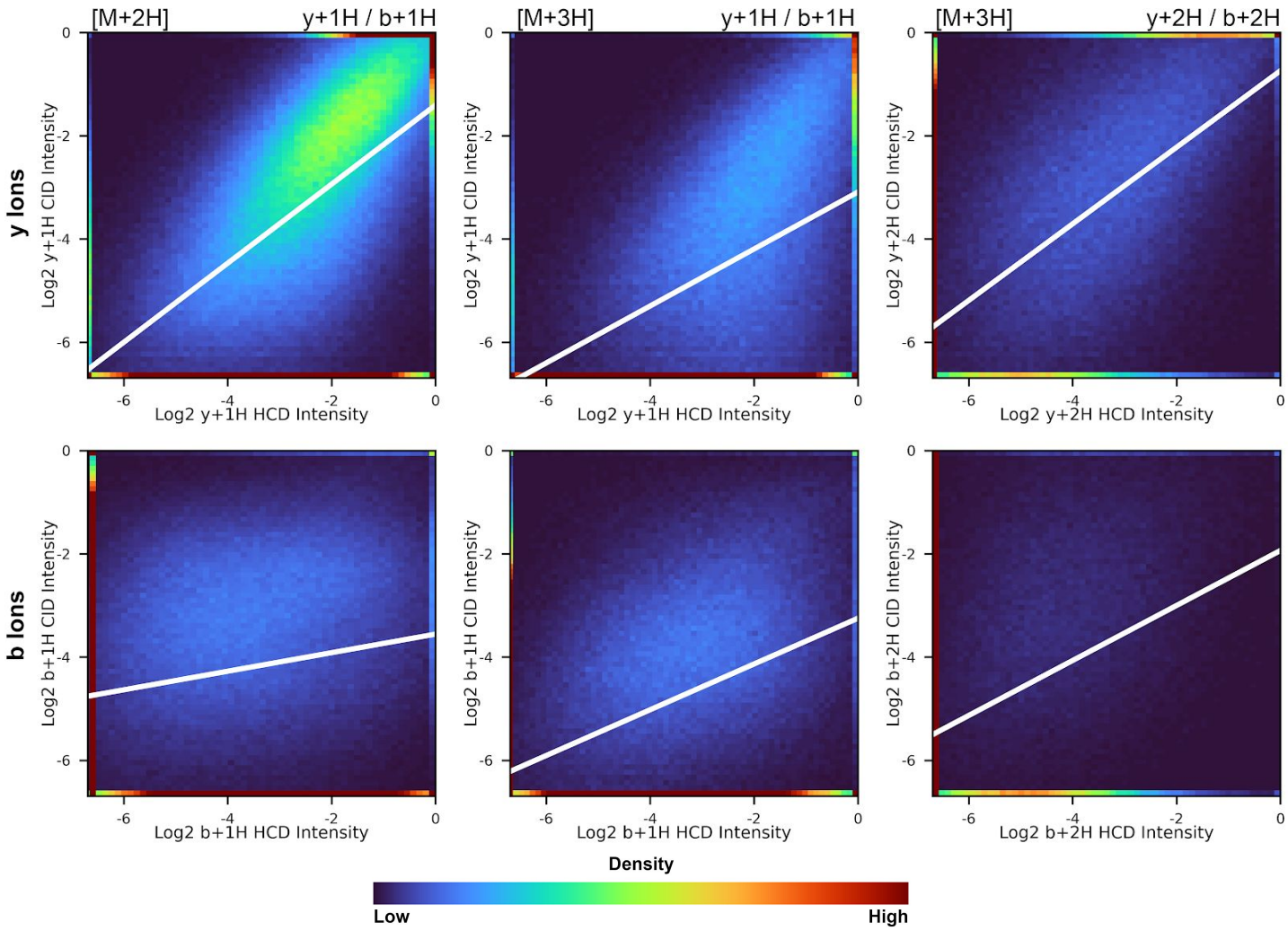

Figure S3. Comparison of y- and b-ions between fragmentation methods. Density maps of intensities of y and b ions in CID versus HCD, with white lines based on the linear regression of the data. The NIST library was used to generate this figure.

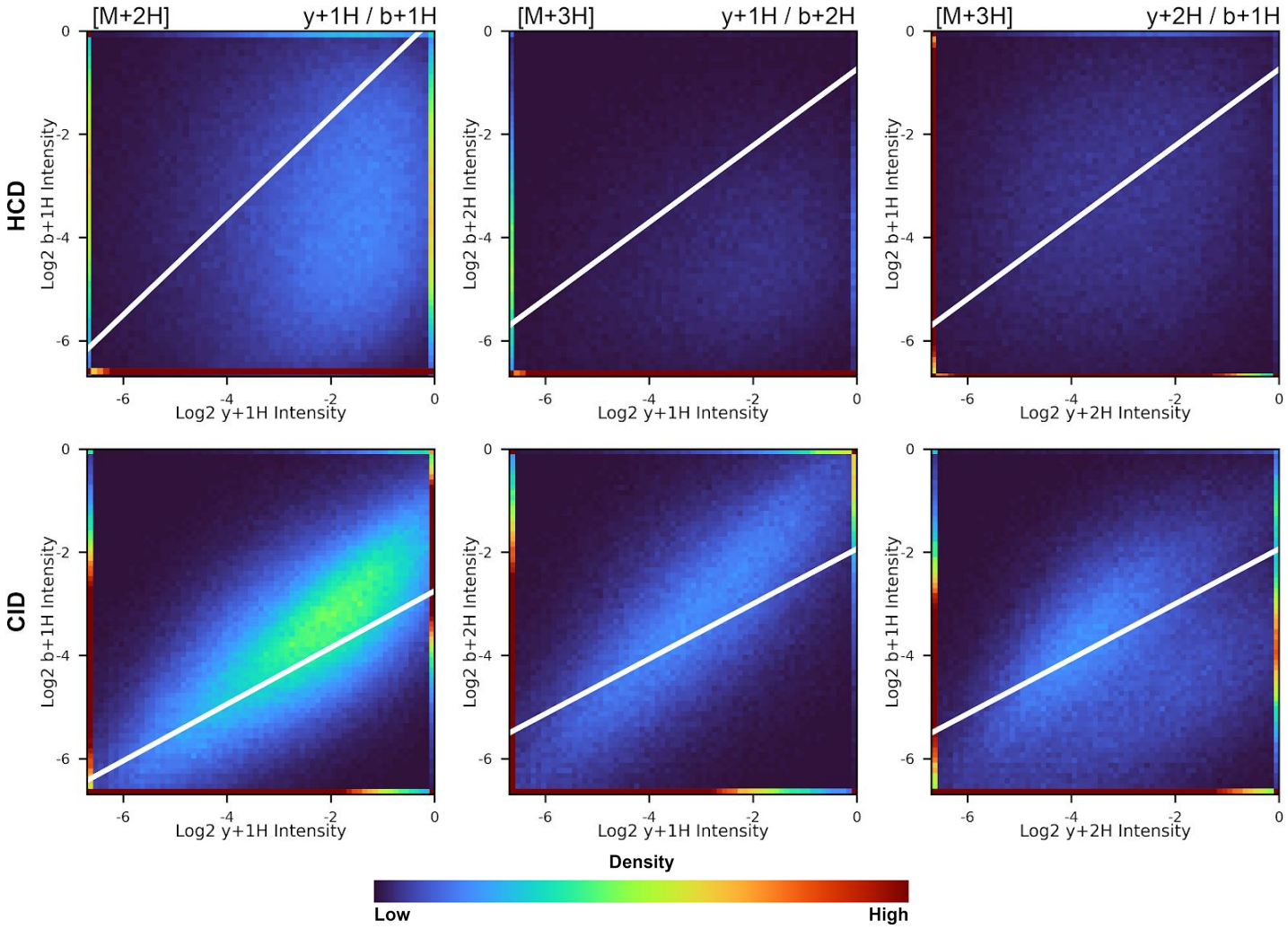

Figure S4. Comparison of y- and b-ions within fragmentation methods. Density maps of intensities of y and b ions in CID versus HCD, with white lines based on the linear regression of the data. The NIST library was used to generate this figure.

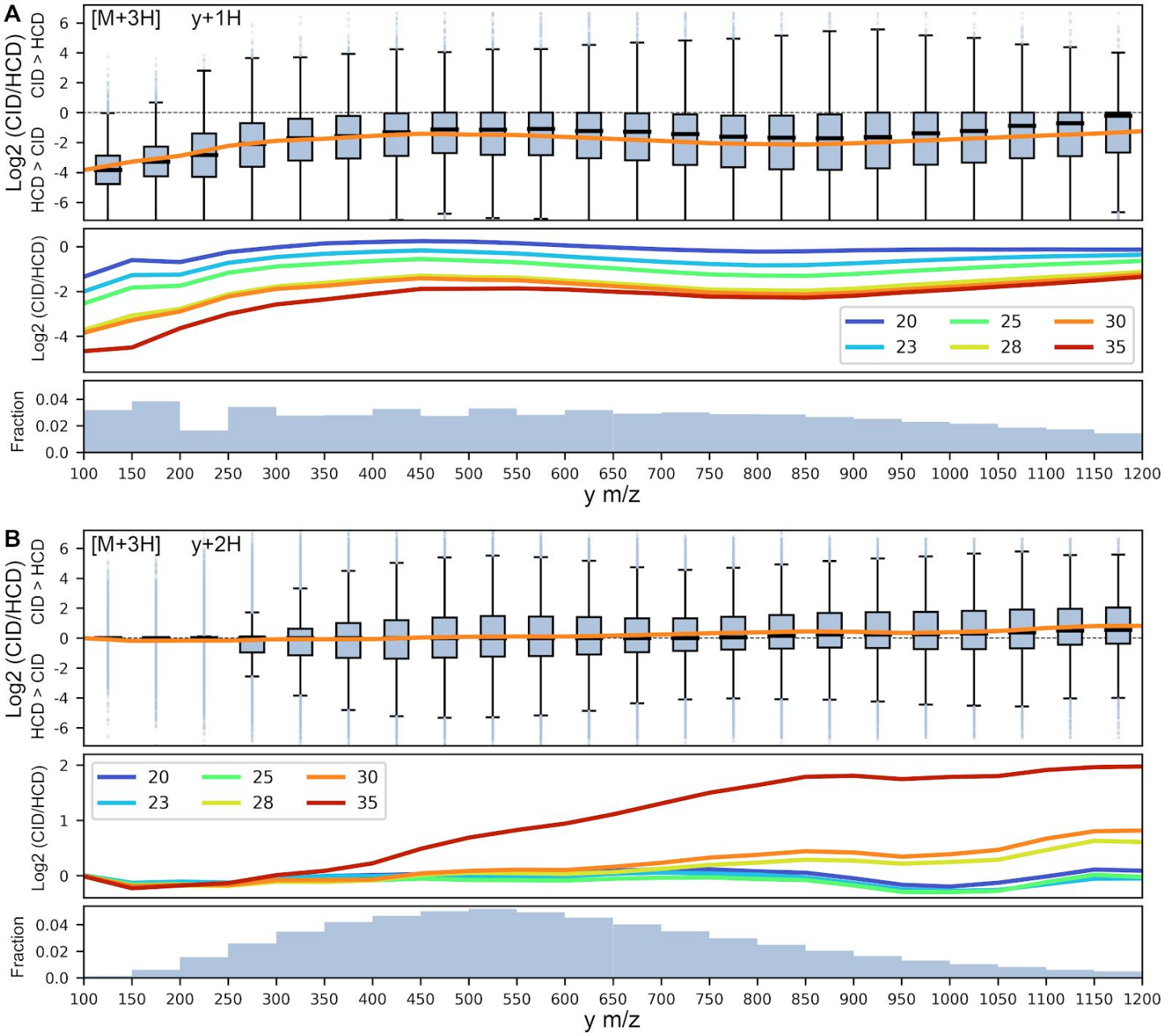

Figure S5. Effect of y-ion m/z on CID/HCD ratio of y-ions in M+3H peptides. Box plot representation of the $\log_2$ transformed ratio of y-ion intensities in CID compared to HCD at NCE 30 as a function of y m/z (in 50 m/z bins), with positive and negative values implying higher CID and HCD intensities, respectively, for (A) y+1H ions and (B) y+2H ions. Boxes reflect the range between the 25 and 75 percentiles, with dark lines denoting the mean and outliers beyond the mean ± 1.5 times the interquartile range. The secondary panels show regression lines at different NCE values, and the tertiary panels denote the fraction of fragments in each m/z bin. The ProteomeTools library was used to generate this figure.

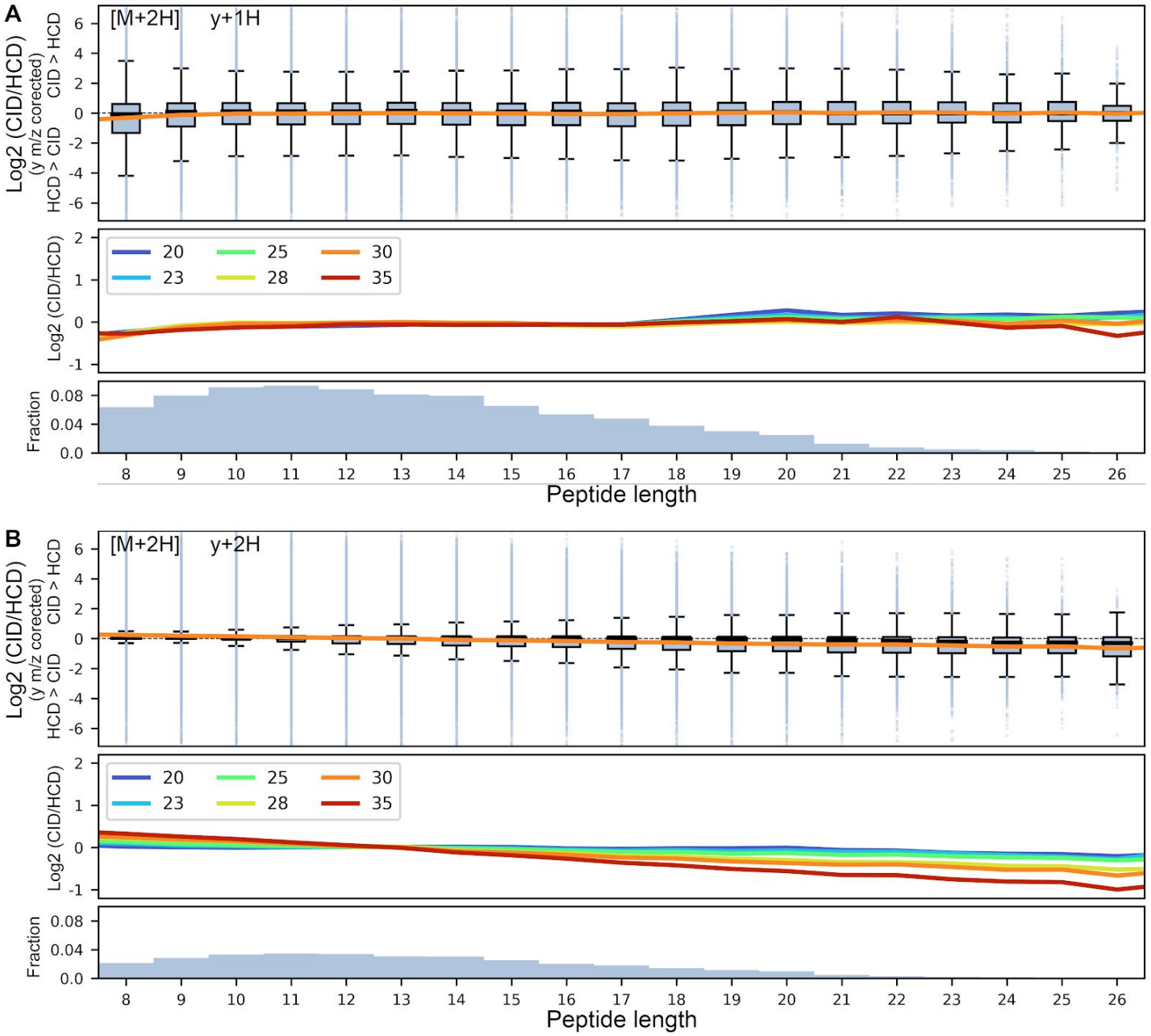

Figure S6. Effect of peptide length on CID/HCD ratio of y-ions in M+2H peptides. Box plot representation of the $\log_2$ transformed ratio of y-ion intensities in CID compared to HCD at NCE 30 (corrected for y m/z) as a function of peptide length, with positive and negative values implying higher CID and HCD intensities, respectively, for (A) y+1H ions and (B) y+2H ions. Boxes reflect the range between the 25 and 75 percentiles, with dark lines denoting the mean and outliers beyond the mean ± 1.5 times the interquartile range. The secondary panels show regression lines at different NCE values, and the tertiary panels denote the fraction of fragments in each m/z bin. The ProteomeTools library was used to generate this figure.

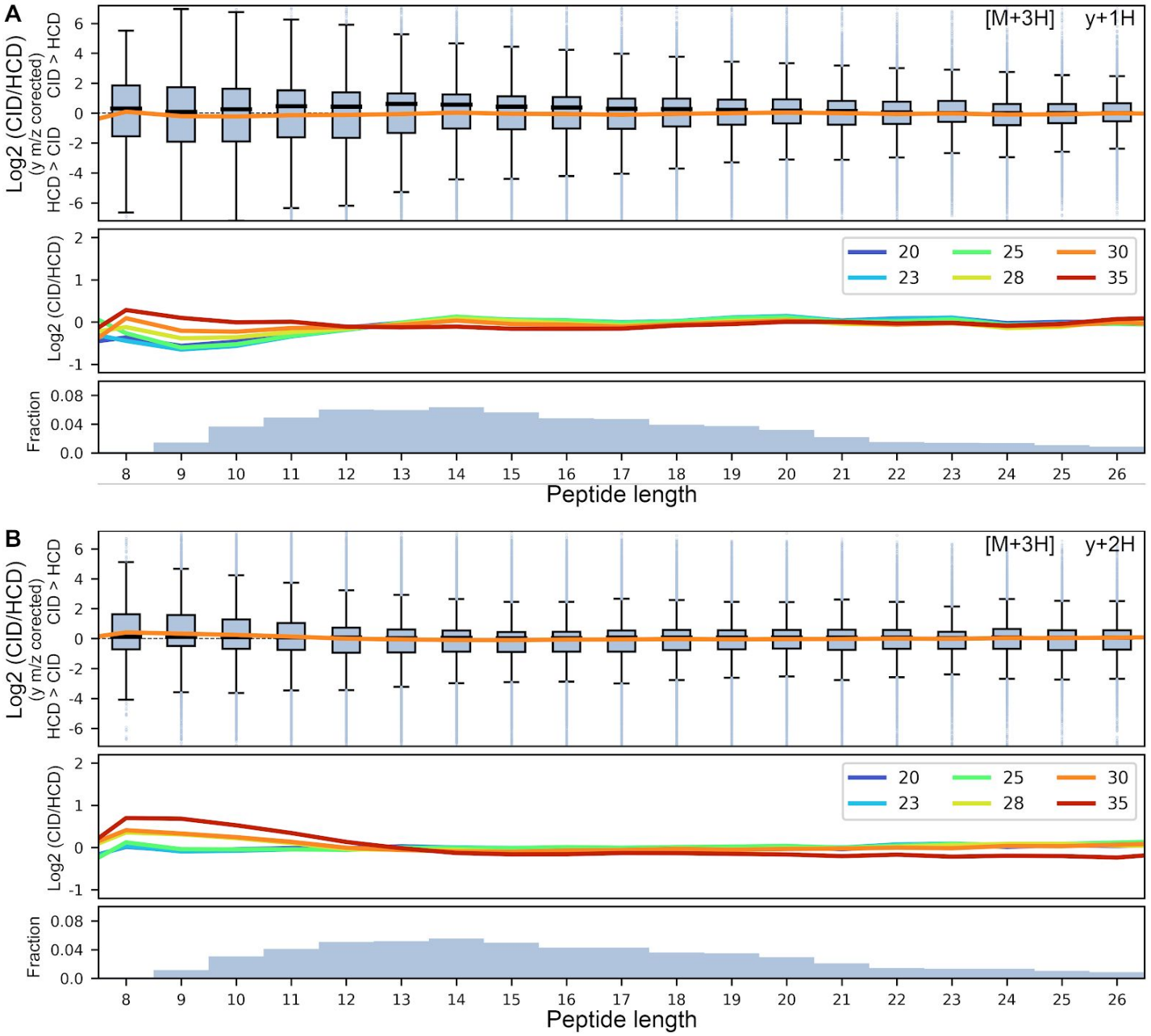

Figure S7. Effect of peptide length on CID/HCD ratio of y-ions in M+3H peptides. Box plot representation of the $\log_2$ transformed ratio of y-ion intensities in CID compared to HCD at NCE 30 (corrected for y m/z) as a function of peptide length, with positive and negative values implying higher CID and HCD intensities, respectively, for (A) y+1H ions and (B) y+2H ions. Boxes reflect the range between the 25 and 75 percentiles, with dark lines denoting the mean and outliers beyond the mean ± 1.5 times the interquartile range. The secondary panels show regression lines at different NCE values, and the tertiary panels denote the fraction of fragments in each m/z bin. The ProteomeTools library was used to generate this figure.

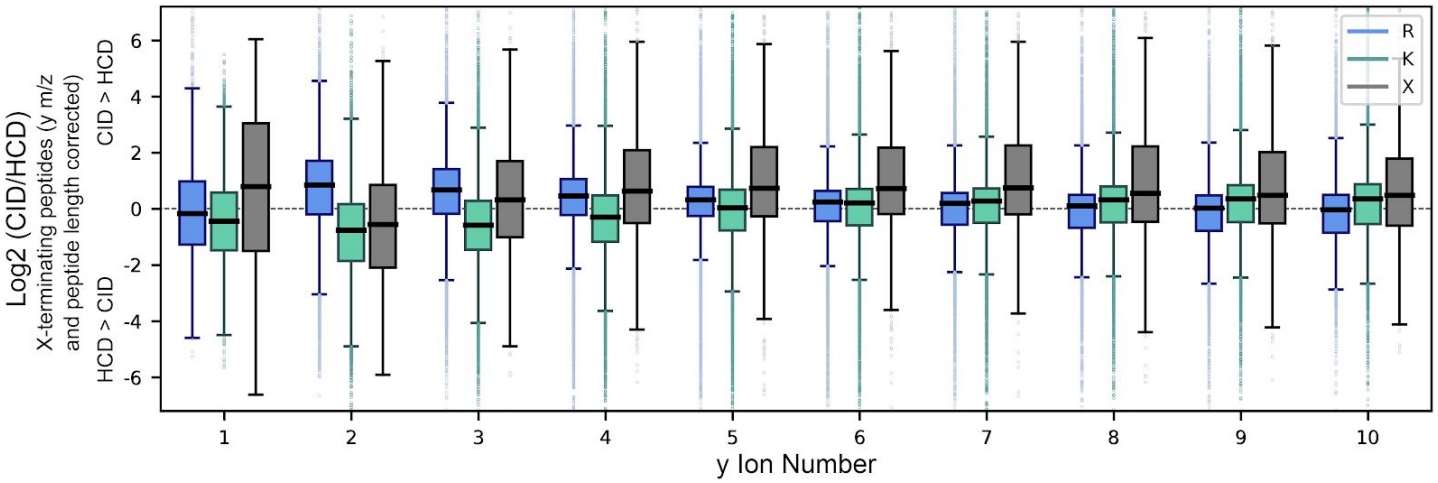

Figure S8. Terminal residue effect on y-ion intensity in HCD and CID. Examination of CID/HCD ratio (at NCE 30) of y-ions that possess a C-terminal Arg, Lys, or other (X) residue as a function of y ion length (after correcting for y m/z effect). Boxes reflect the range between the 25 and 75 percentiles, with dark lines denoting the mean and outliers beyond the mean ± 1.5 times the interquartile range. The ProteomeTools library was used to generate this figure.

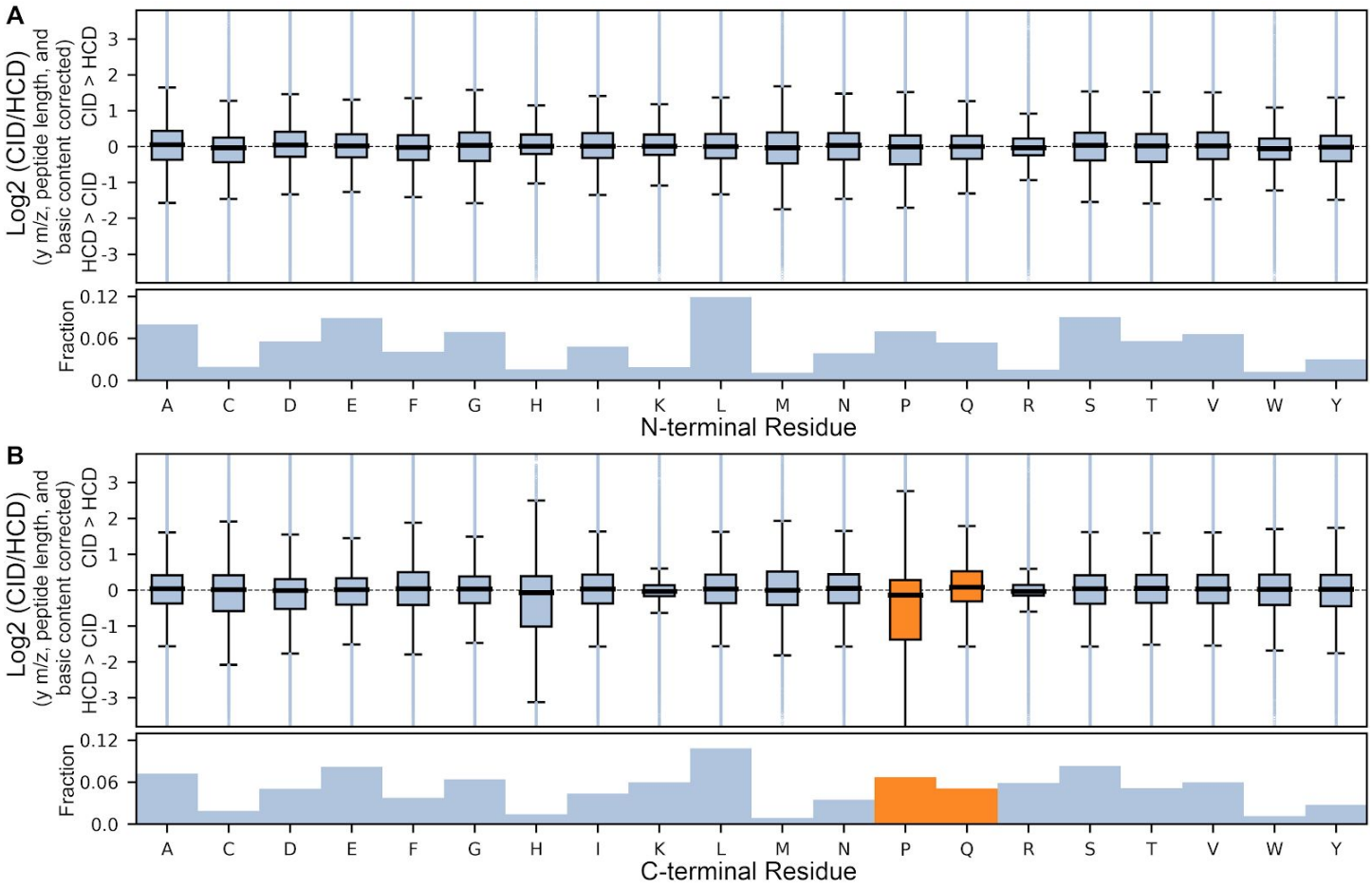

Figure S9. N- and C-terminal bond cleavage residue effects on CID/HCD ratio for y+2H ions of M+2H peptides. Box plot representation of bond cleavage residues on CID vs HCD intensity (at NCE 30) after normalizing for mass (y m/z and peptide length) and basic residue contributions for (A) the N-terminal side of the bond cleavage and (B) the C-terminal side. Boxes reflect the range between the 25 and 75 percentiles, with dark lines denoting the mean and outliers beyond the mean ± 1.5 times the interquartile range, with orange boxes reflecting effects that explain >0.1% of the variance and included in the model. The ProteomeTools library was used to generate this figure.

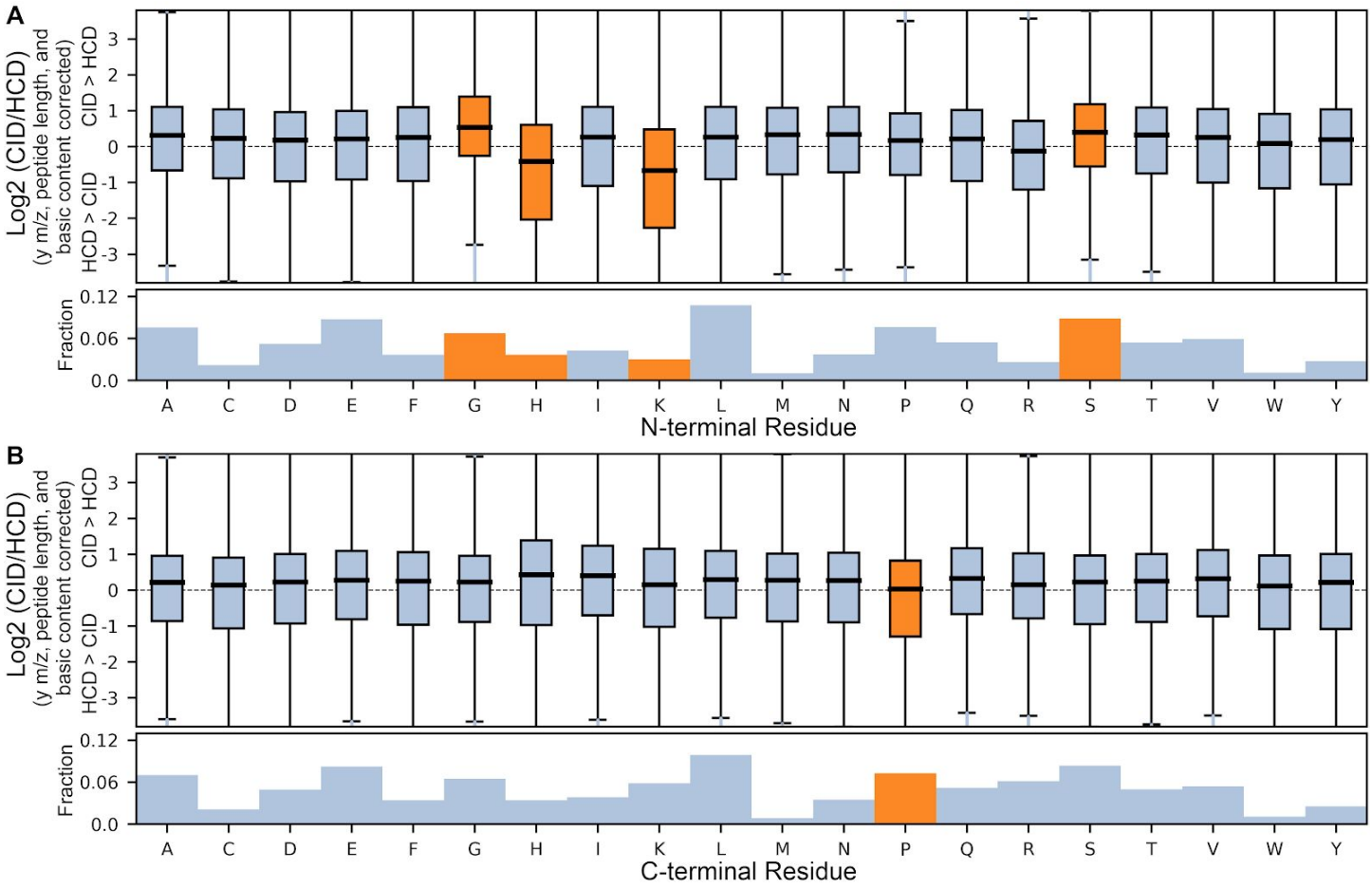

Figure S10. N- and C-terminal bond cleavage residue effects on CID/HCD ratio for y+1H ions of M+3H peptides. Box plot representation of bond cleavage residues on CID vs HCD intensity (at NCE 30) after normalizing for mass (y m/z and peptide length) and basic residue contributions for (A) the N-terminal side of the bond cleavage and (B) the C-terminal side. Boxes reflect the range between the 25 and 75 percentiles, with dark lines denoting the mean and outliers beyond the mean ± 1.5 times the interquartile range, with orange boxes reflecting effects that explain >0.1% of the variance and included in the model. The ProteomeTools library was used to generate this figure.

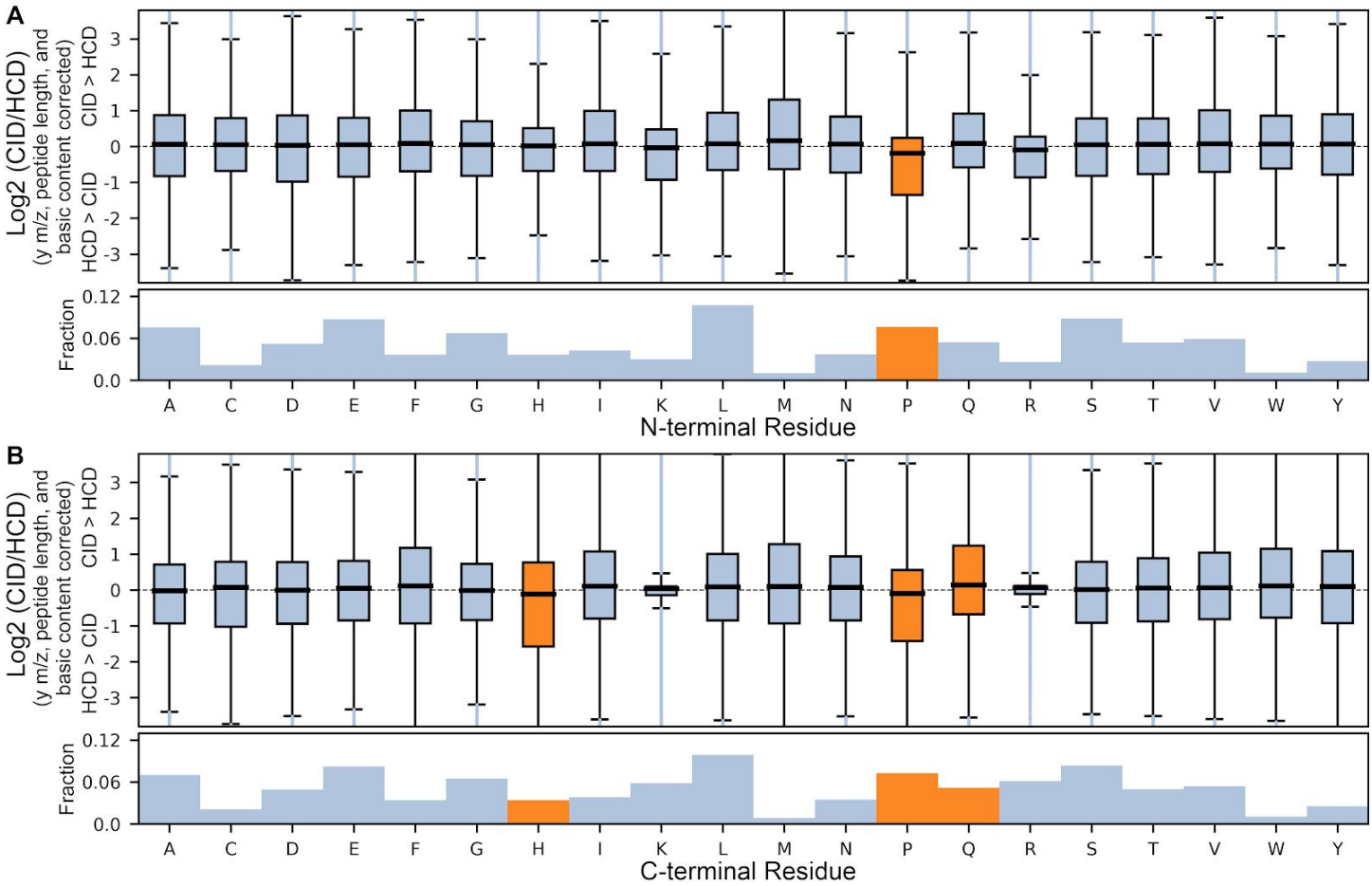

Figure S11. N- and C-terminal bond cleavage residue effects on CID/HCD ratio for y+2H ions of M+3H peptides. Box plot representation of bond cleavage residues on CID vs HCD intensity (at NCE 30) after normalizing for mass (y m/z and peptide length) and basic residue contributions for (A) the N-terminal side of the bond cleavage and (B) the C-terminal side. Boxes reflect the range between the 25 and 75 percentiles, with dark lines denoting the mean and outliers beyond the mean ± 1.5 times the interquartile range, with orange boxes reflecting effects that explain >0.1% of the variance and included in the model. The ProteomeTools library was used to generate this figure.

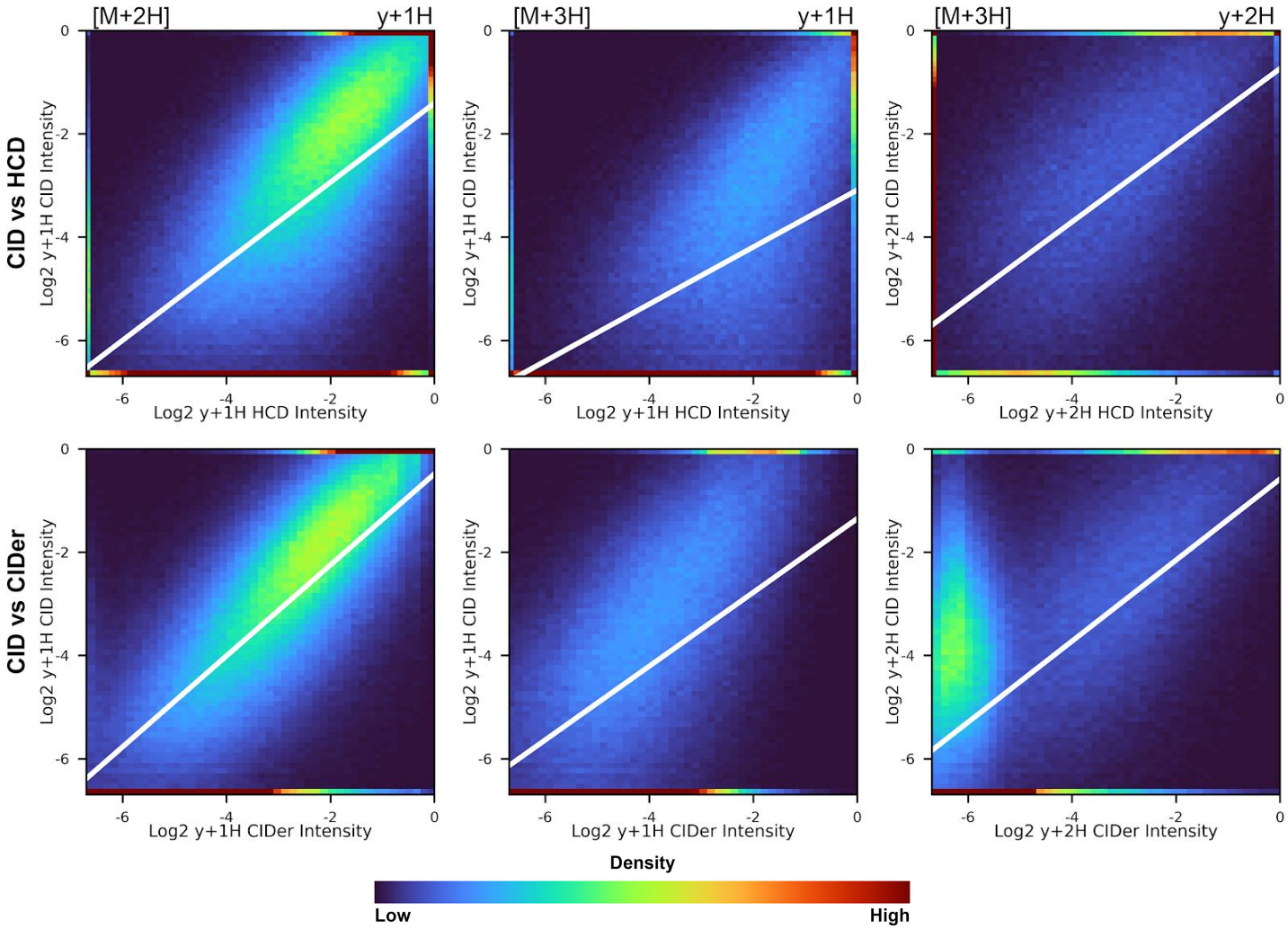

Figure S12. Comparison of CID y-ion correlation with HCD values or CIDer corrected estimates. Density maps of intensities of y ion intensities, with white lines based on the linear regression of the data. The NIST library was used to generate this figure.

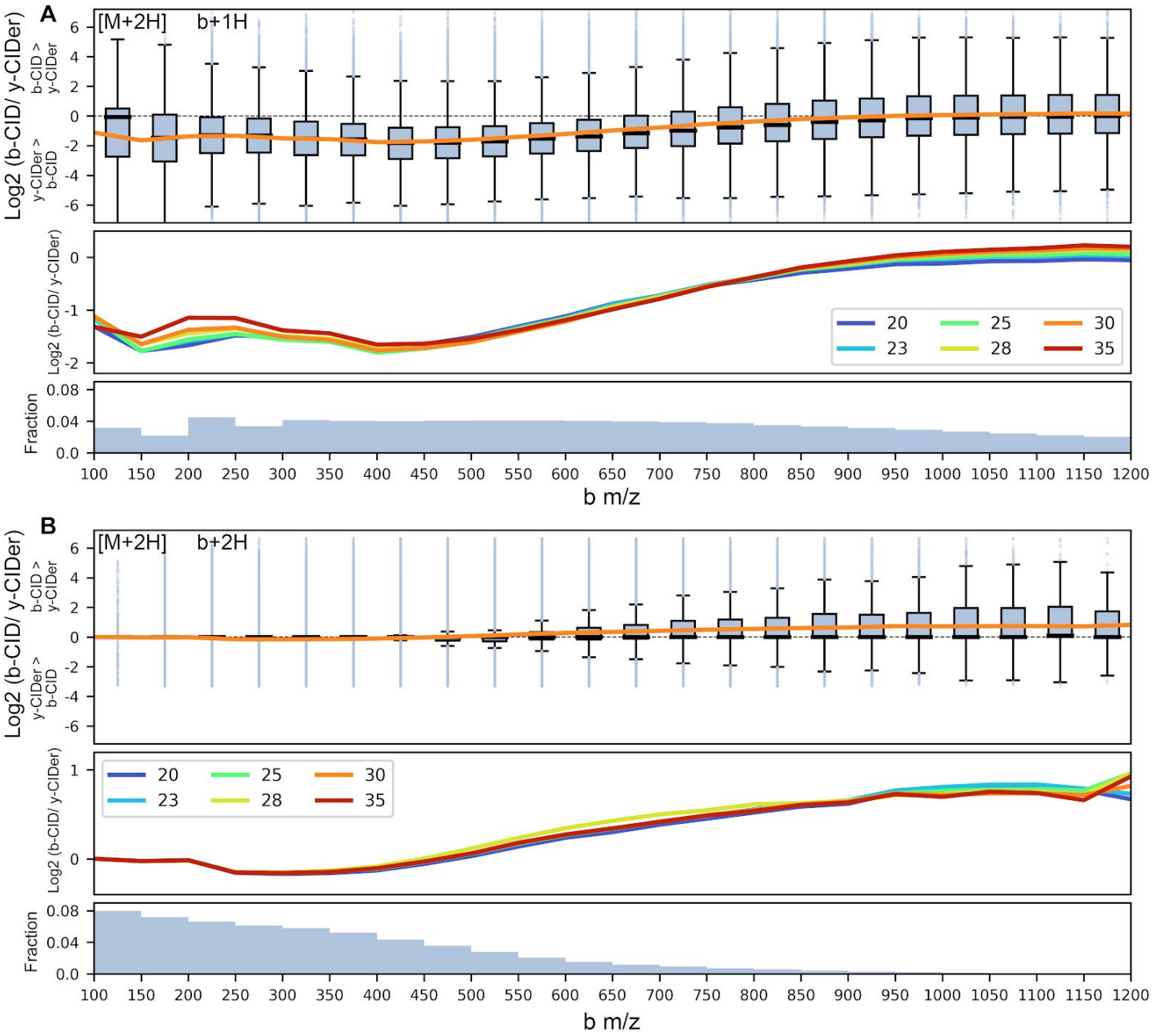

Figure S13. Effect of b-ion m/z on b-CID/y-CIDer ratio for M+2H peptides. Box plot representation of the $\log_2$ transformed ratio of b-ion intensities in CID compared to y CIDer estimates as a function of b m/z (in 50 m/z bins), with positive and negative values implying higher b-CID and y-CIDer intensities at NCE 30, respectively, for (A) b+1H ions and (B) b+2H ions. Boxes reflect the range between the 25 and 75 percentiles, with dark lines denoting the mean and outliers beyond the mean ± 1.5 times the interquartile range. The secondary panels show regression lines at different NCE values, and the tertiary panels denote the fraction of fragments in each m/z bin. The ProteomeTools library was used to generate this figure.

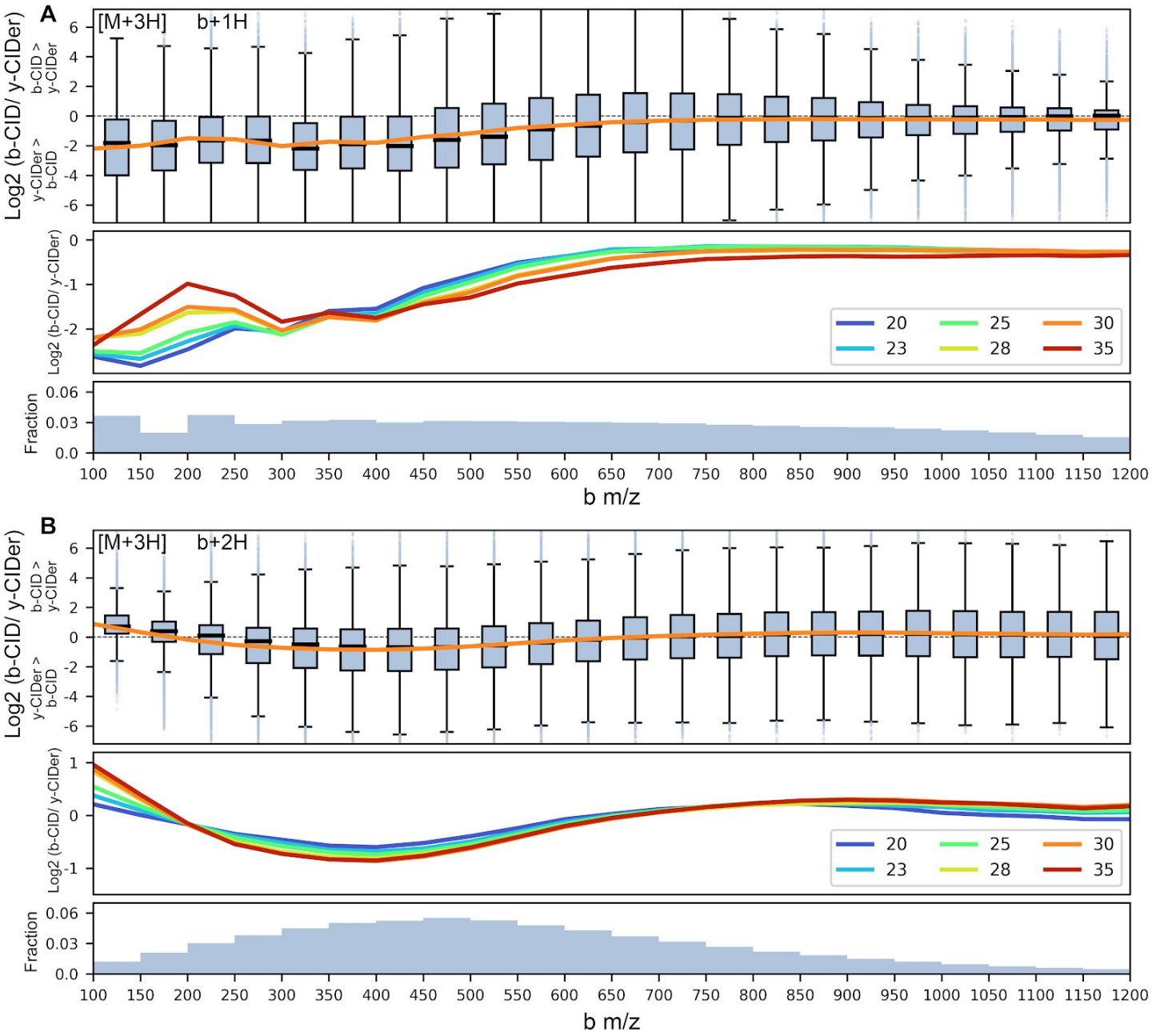

Figure S14. Effect of b-ion m/z on b-CID/y-CIDer ratio for M+3H peptides. Box plot representation of the $\log_2$ transformed ratio of b-ion intensities in CID compared to y-CIDer estimates as a function of b m/z (in 50 m/z bins), with positive and negative values implying higher b-CID and y-CIDer intensities at NCE 30, respectively, for (A) b+1H ions and (B) b+2H ions. Boxes reflect the range between the 25 and 75 percentiles, with dark lines denoting the mean and outliers beyond the mean ± 1.5 times the interquartile range. The secondary panels show regression lines at different NCE values, and the tertiary panels denote the fraction of fragments in each m/z bin. The ProteomeTools library was used to generate this figure.

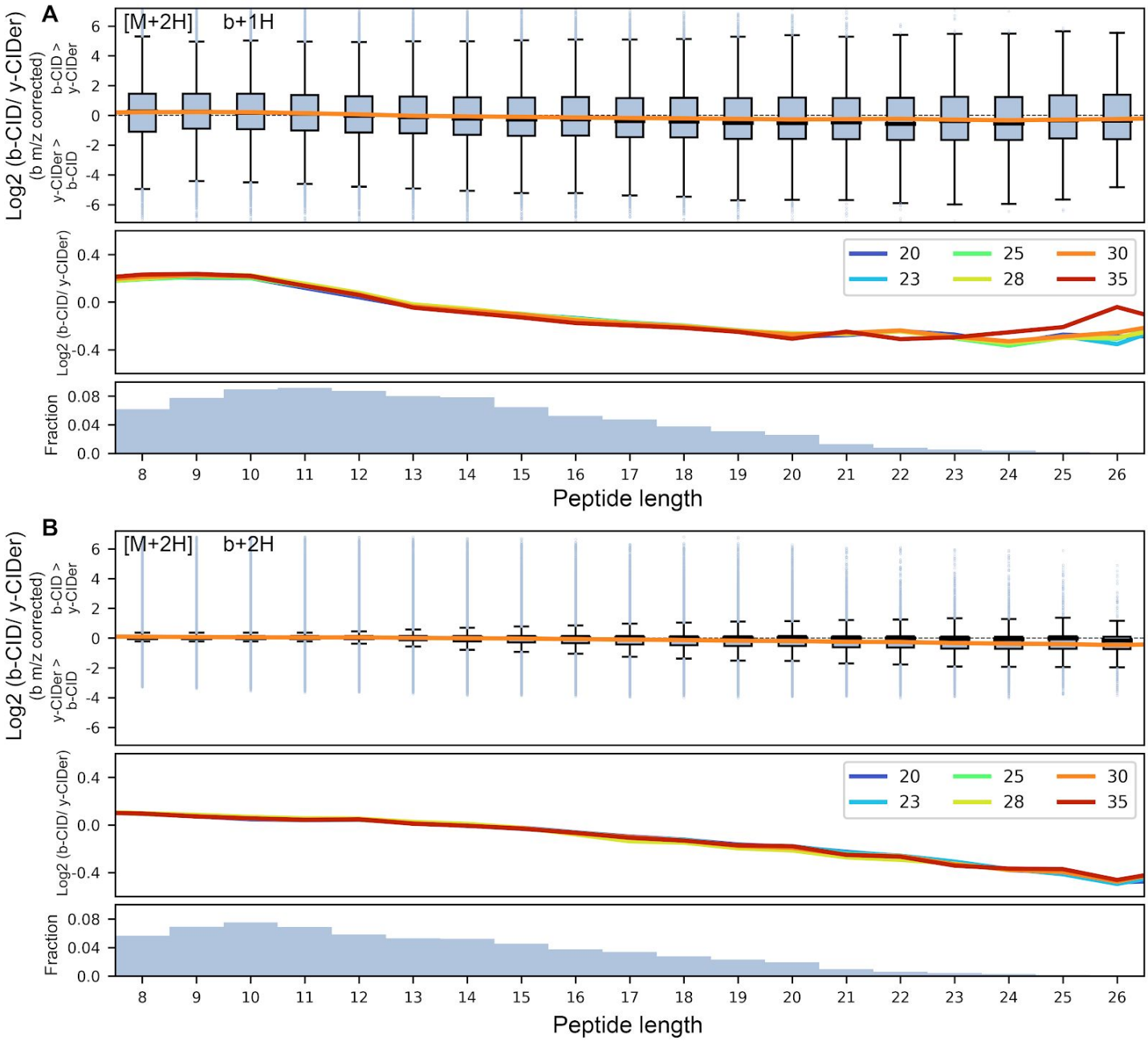

Figure S15. Effect of peptide length on b-CID/y-CIDer ratio for M+2H peptides. Box plot representation of the $\log_2$ transformed ratio of b-ion intensities in CID compared to y-CIDer (corrected for b m/z) as a function of peptide length, with positive and negative values implying higher b-CID and y-CIDer intensities at NCE 30, respectively, for (A) b+1H ions and (B) b+2H ions. Boxes reflect the range between the 25 and 75 percentiles, with dark lines denoting the mean and outliers beyond the mean ± 1.5 times the interquartile range. The secondary panels show regression lines at different NCE values, and the tertiary panels denote the fraction of fragments in each m/z bin. The ProteomeTools library was used to generate this figure.

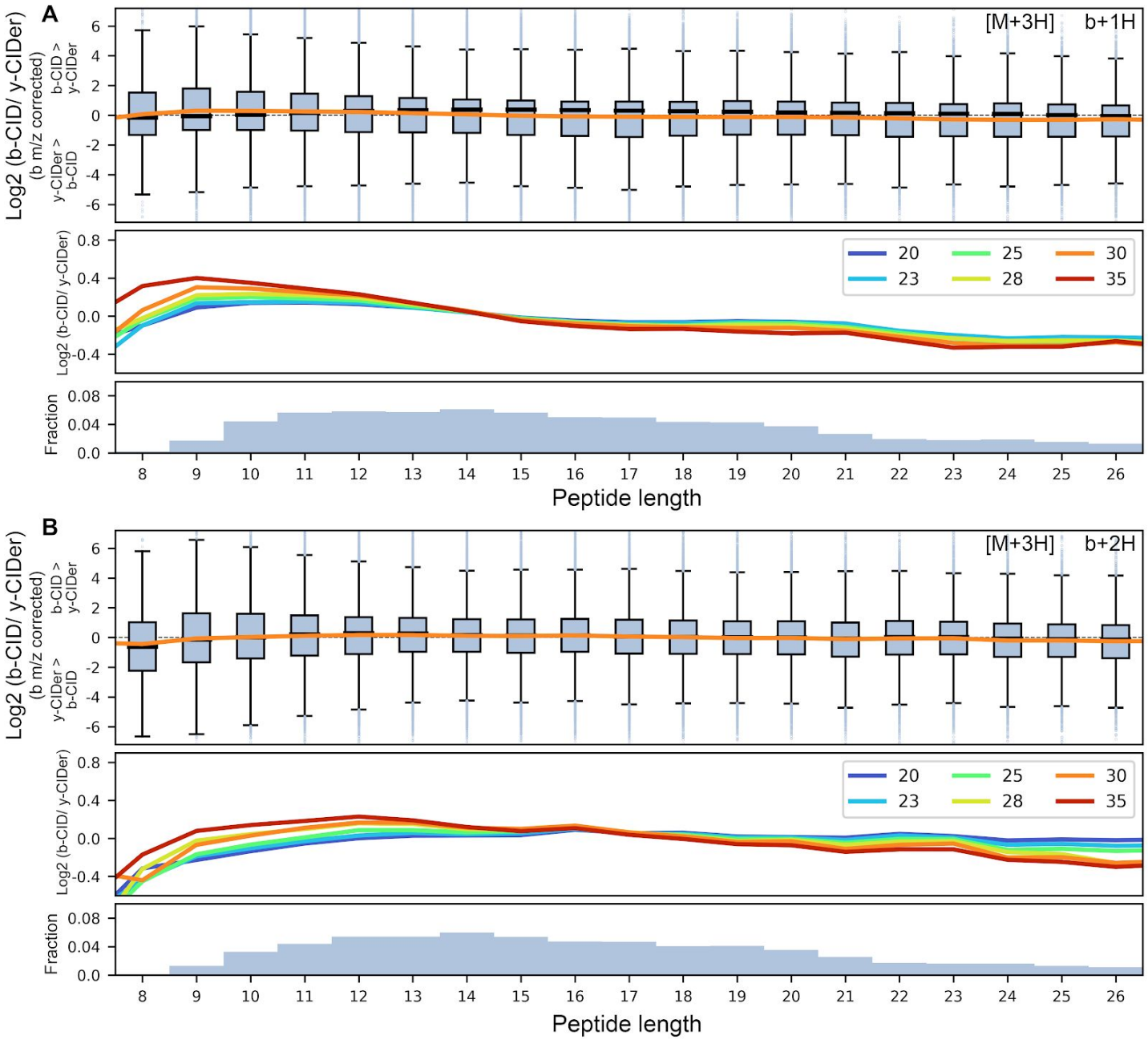

Figure S16. Effect of peptide length on b-CID/y-CIDer ratio for M+3H peptides. Box plot representation of the $\log_2$ transformed ratio of b-ion intensities in CID compared to y-CIDer (corrected for b m/z) as a function of peptide length, with positive and negative values implying higher b-CID and y-CIDer intensities at NCE 30, respectively, for (A) b+1H ions and (B) b+2H ions. Boxes reflect the range between the 25 and 75 percentiles, with dark lines denoting the mean and outliers beyond the mean ± 1.5 times the interquartile range. The secondary panels show regression lines at different NCE values, and the tertiary panels denote the fraction of fragments in each m/z bin. The ProteomeTools library was used to generate this figure.

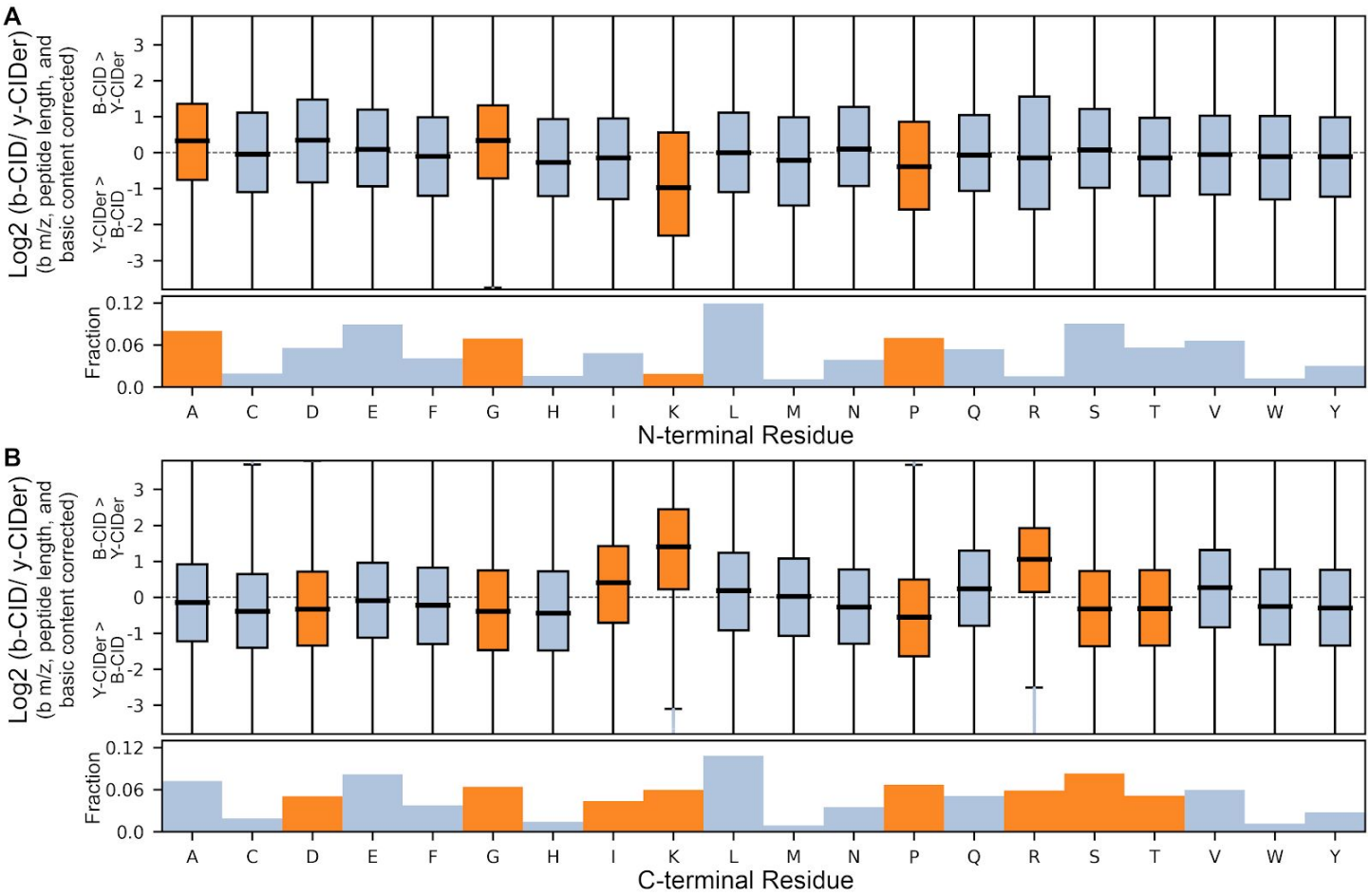

Figure S17. N- and C-terminal bond cleavage residue effects on b-CID/y-CIDer ratio for b+1H ions of M+2H peptides. Box plot representation of bond cleavage residues on b-CID vs y-CIDer intensity (at NCE 30) after normalizing for mass (b m/z and peptide length) and basic residue contributions for (A) the N-terminal side of the bond cleavage and (B) the C-terminal side. Boxes reflect the range between the 25 and 75 percentiles, with dark lines denoting the mean and outliers beyond the mean ± 1.5 times the interquartile range, with orange boxes reflecting effects that explain >0.1% of the variance and included in the model. The ProteomeTools library was used to generate this figure.

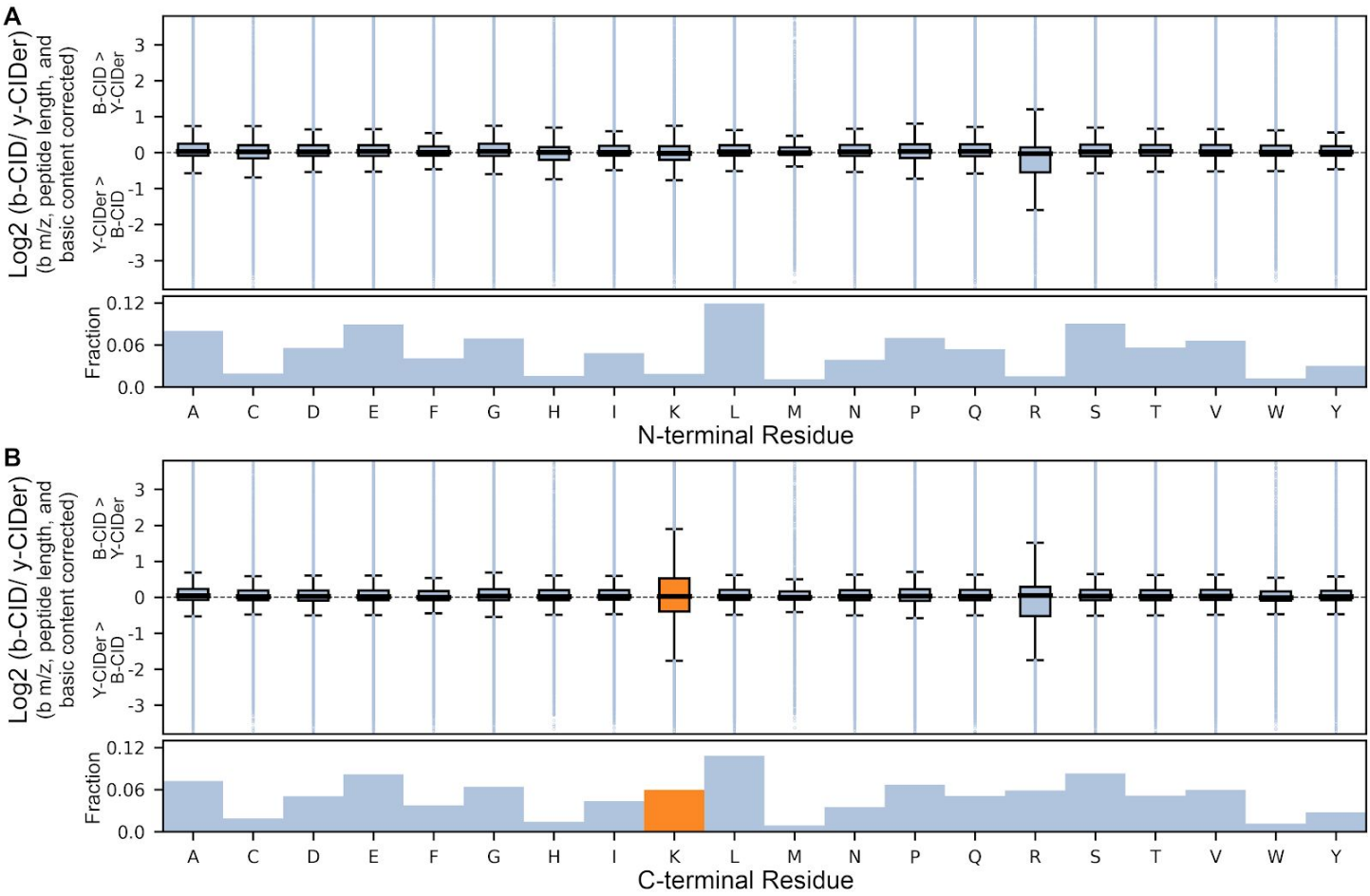

Figure S18. N- and C-terminal bond cleavage residue effects on b-CID/y-CIDer ratio for b+2H ions of M+2H peptides. Box plot representation of bond cleavage residues on b-CID vs y-CIDer intensity (at NCE 30) after normalizing for mass (b m/z and peptide length) and basic residue contributions for (A) the N-terminal side of the bond cleavage and (B) the C-terminal side. Boxes reflect the range between the 25 and 75 percentiles, with dark lines denoting the mean and outliers beyond the mean ± 1.5 times the interquartile range, with orange boxes reflecting effects that explain >0.1% of the variance and included in the model. The ProteomeTools library was used to generate this figure.

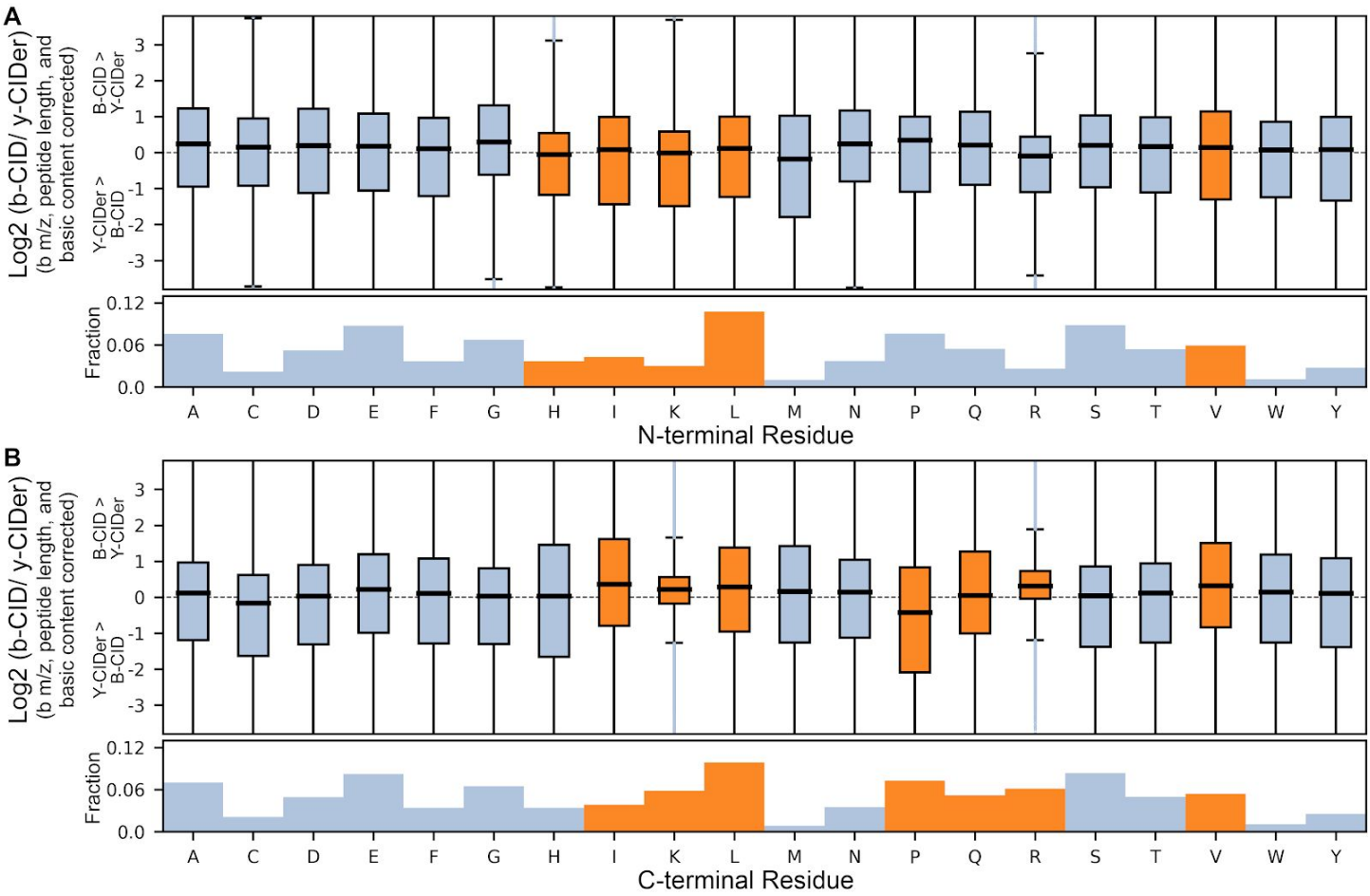

Figure S19. N- and C-terminal bond cleavage residue effects on b-CID/y-CIDer ratio for b+1H ions of M+3H peptides. Box plot representation of bond cleavage residues on b-CID vs y-CIDer intensity (at NCE 30) after normalizing for mass (b m/z and peptide length) and basic residue contributions for (A) the N-terminal side of the bond cleavage and (B) the C-terminal side. Boxes reflect the range between the 25 and 75 percentiles, with dark lines denoting the mean and outliers beyond the mean ± 1.5 times the interquartile range, with orange boxes reflecting effects that explain >0.1% of the variance and included in the model. The ProteomeTools library was used to generate this figure.

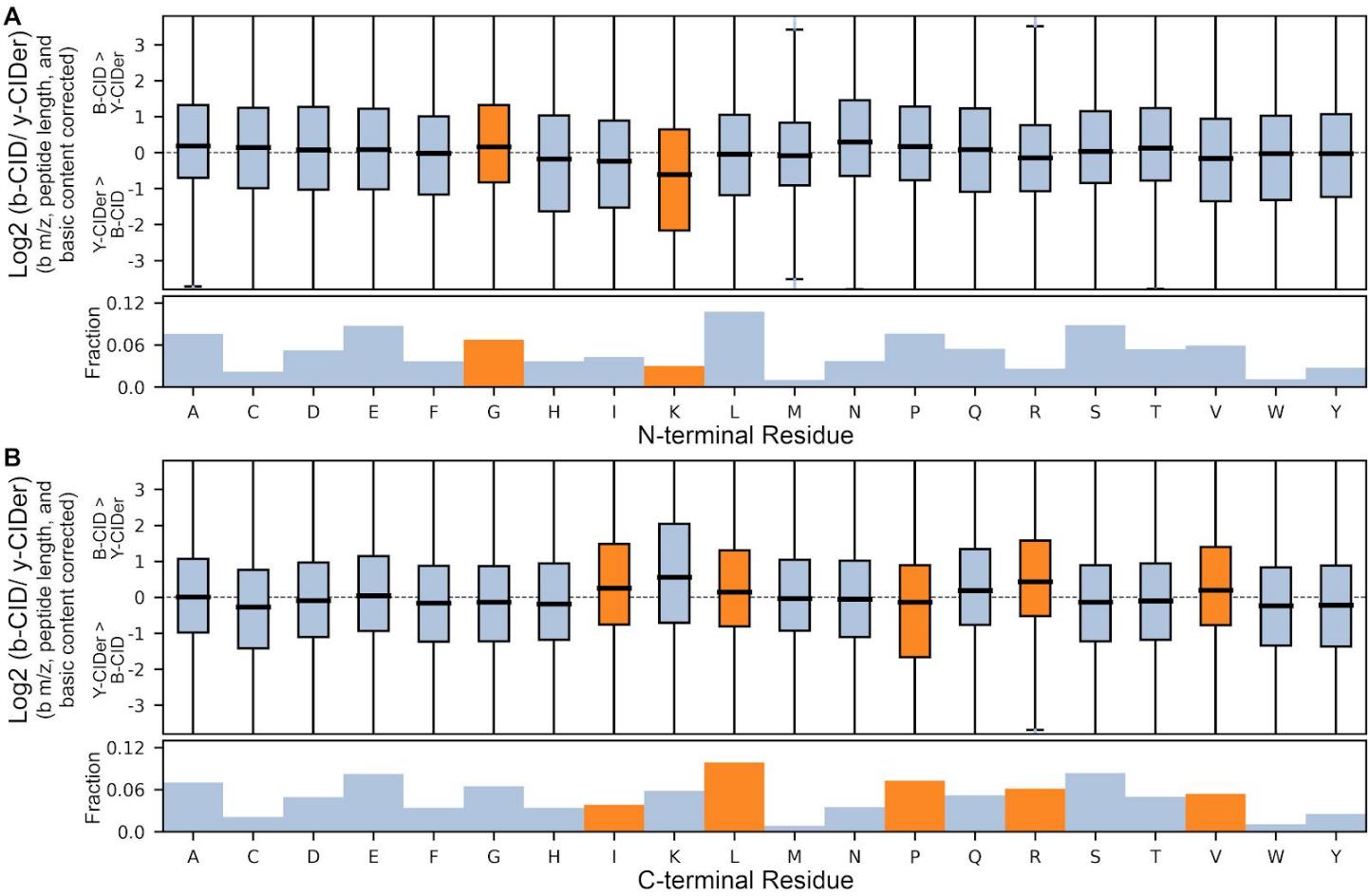

Figure S20. N- and C-terminal bond cleavage residue effects on b-CID/y-CIDer ratio for b+2H ions of M+3H peptides. Box plot representation of bond cleavage residues on b-CID vs y-CIDer intensity (at NCE 30) after normalizing for mass (b m/z and peptide length) and basic residue contributions for (A) the N-terminal side of the bond cleavage and (B) the C-terminal side. Boxes reflect the range between the 25 and 75 percentiles, with dark lines denoting the mean and outliers beyond the mean ± 1.5 times the interquartile range, with orange boxes reflecting effects that explain >0.1% of the variance and included in the model. The ProteomeTools library was used to generate this figure.

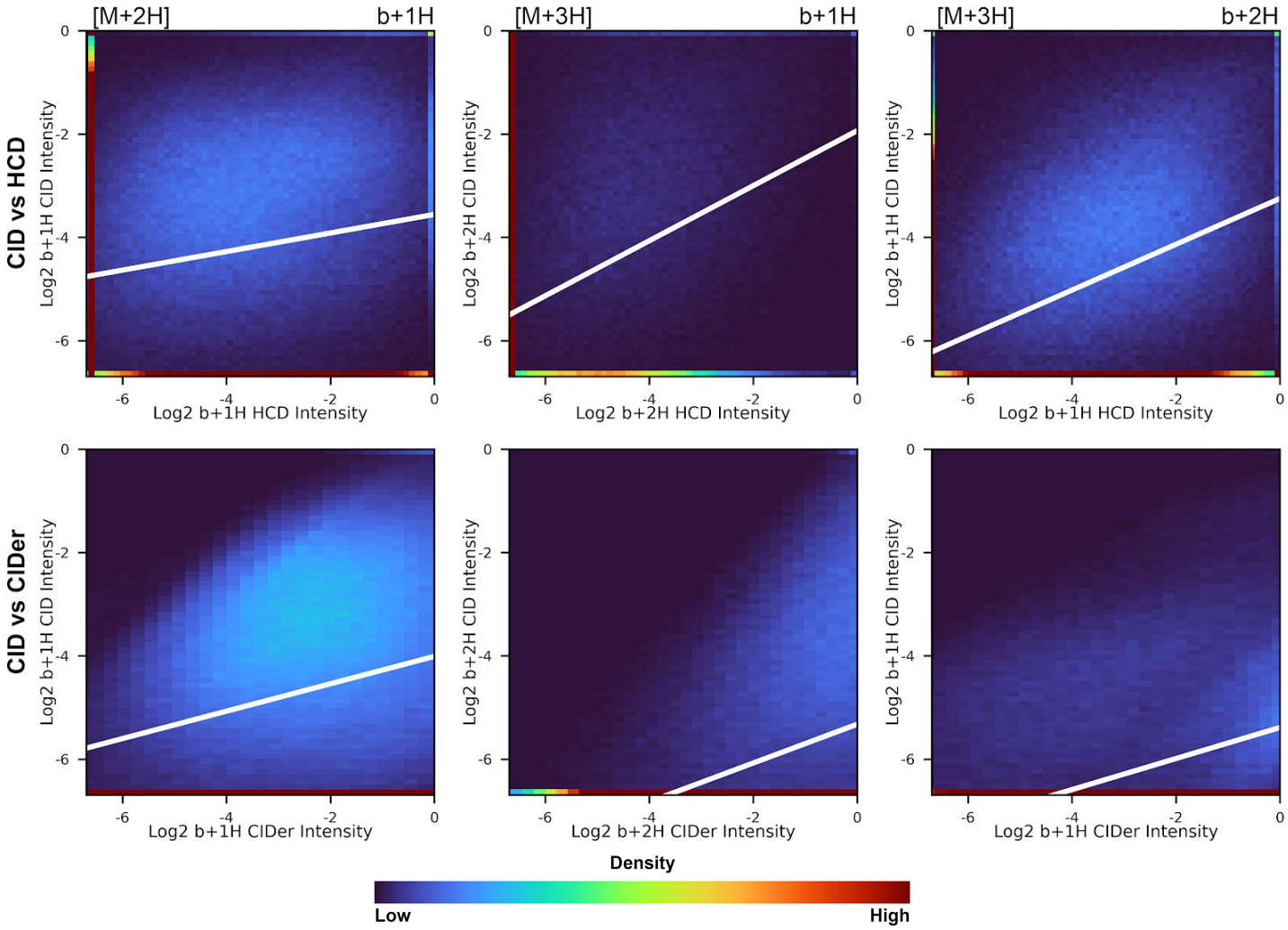

Figure S21. Comparison of CID b-ion correlation with HCD values or CIDer corrected estimates. Density maps of intensities of b ion intensities, with white lines based on the linear regression of the data. The NIST library was used to generate this figure.

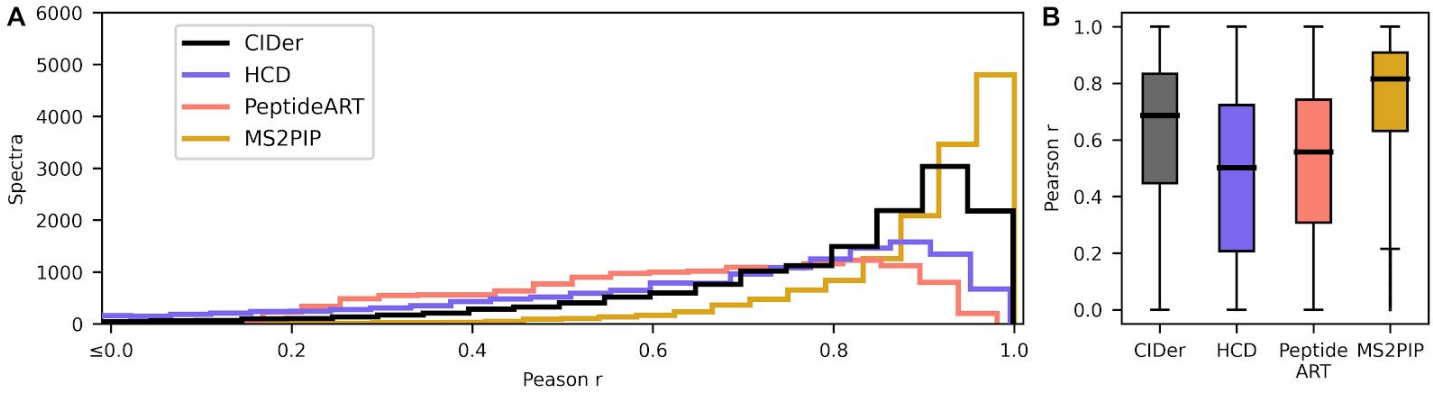

Figure S22. Comparison of CIDer performance to HCD libraries and machine learning tools. (A) Density plot of Pearson coefficients for CID correlation to each of the ascribed methods. (C) The Pearson coefficient values represented by box plot, with the box representing the range between the 25 and 75 percentiles, and the dark line demarcating the median. Evaluation Dataset 2 was used to generate this figure.

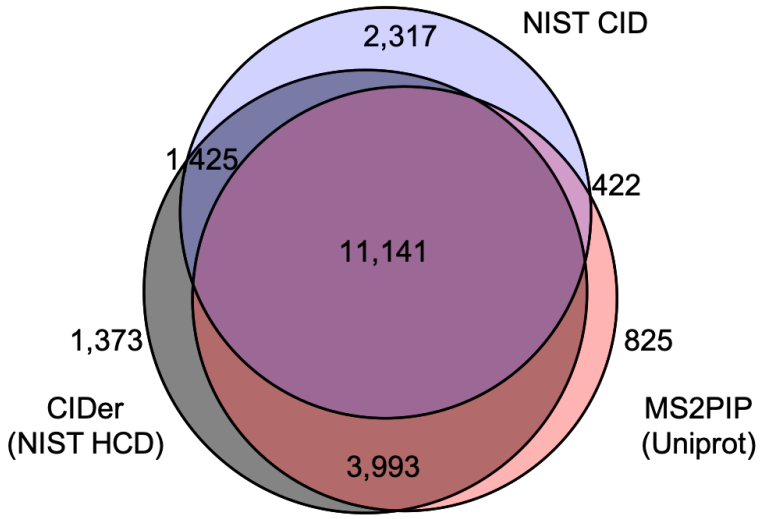

Figure S23. Overlap between unique peptide sequences reported after library searching. Searches were performed using either the NIST CID library, or libraries generated from CIDer (predicted from NIST HCD library entries) or MS2PIP (predicted from *in silico* digested peptides from a Uniprot FASTA database). Search results were filtered to a 1% peptide-level FDR. Evaluation Dataset 2 was used to generate this figure.