

Interactive Peptide Spectral Annotator: A Versatile Web-based Tool for Proteomic Applications

Authors

Dain R. Brademan, Nicholas M. Riley, Nicholas W. Kwiecien, and Joshua J. Coon

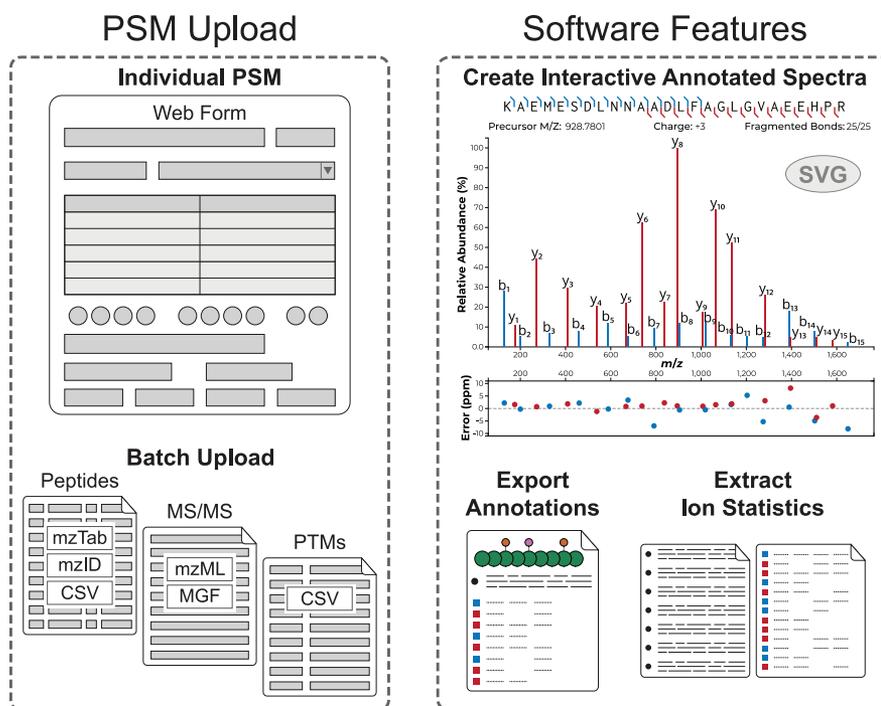
Correspondence

jcoon@chem.wisc.edu

In Brief

We have developed an innovative web-based spectrum annotator which visualizes and characterizes peptide tandem mass spectra. The generated annotated spectra are fully interactive and customizable, and any spectrum can be exported as a scalable vector graphic for publication-ready figures. All uploaded data can optionally be batch processed to rapidly extract ion statistics for method development. Our platform additionally supports the annotation of spectra collected in the negative mode, providing a much-needed resource for the proteomics community.

Graphical Abstract



Highlights

- Generation of interactive annotated spectra using scalable vector graphics.
- Upload data using generic or standard file formats mzTab, mzIdentML, or mzML.
- Annotation support for peptides collected using negative-mode ionization.
- Batch processing of uploaded identifications to report detailed ion statistics.

Interactive Peptide Spectral Annotator: A Versatile Web-based Tool for Proteomic Applications*

Dain R. Brademan^{‡¶},  Nicholas M. Riley^{‡¶}, Nicholas W. Kwiecien^{‡¶}, and  Joshua J. Coon^{‡§¶||**}

Here we present IPSA, an innovative web-based spectrum annotator that visualizes and characterizes peptide tandem mass spectra. A tool for the scientific community, IPSA can visualize peptides collected using a wide variety of experimental and instrumental configurations. Annotated spectra are customizable via a selection of interactive features and can be exported as editable scalable vector graphics to aid in the production of publication-quality figures. Single spectra can be analyzed through provided web forms, whereas data for multiple peptide spectral matches can be uploaded using the Proteomics Standards Initiative file formats mzTab, mzIdentML, and mzML. Alternatively, peptide identifications and spectral data can be provided using generic file formats. IPSA provides supports for annotating spectra collecting using negative-mode ionization and facilitates the characterization of experimental MS/MS performance through the optional export of fragment ion statistics from one to many peptide spectral matches. This resource is made freely accessible at <http://interactivepeptidespectralannotator.com>, whereas the source code and user guides are available at <https://github.com/coongroup/IPSA> for private hosting or custom implementations. *Molecular & Cellular Proteomics* 18: S193–S201, 2019. DOI: 10.1074/mcp.TIR118.001209.

Tandem mass spectrometry (MS/MS)¹ is the centerpiece of modern proteome analysis. Advances in instrument design and acquisition software have enabled collection of well over 100,000 MS/MS scans in less than an hour of analysis (1–9). Researchers have developed a wide variety of search algorithms and related computational tools to rapidly translate this large volume of experimental data to peptide spectral matches (PSMs), where peptide sequences are assigned to spectra to identify the proteins present in a sample (10–16). An important component to this process is matching expected product ions to those observed in the experimental spectra. Annotation of spectra in this sense usually involves

labeling observed *m/z* features with matched fragment ion designations (e.g. *a/x*-, *b/y*-, or *c/z*-type product ions) derived from the reported peptide sequence. Expert manual annotation is a valuable but greatly time-consuming process—unfeasible for the large volume of spectra generated in modern proteomic experiments.

Proteomic field guidelines have increasingly emphasized the importance of providing access to annotated MS/MS spectra for publication, which allows others to inspect reported PSMs and validate their assignment to a given sequence (17–20). Many software tools have been created to aid researchers annotating individual PSMs contained in bulk datasets. Most such tools are downloadable and often integrated directly into data-analysis suites, although a handful have been developed as web browser-based platforms (21–23). Lorikeet (<https://uwpr.github.io/Lorikeet/>) is a well-established web-based spectral annotator which has been integrated into several online mass spectrometry resources to visualize routine shotgun and cross-linked proteomics data (24–27). However, Lorikeet does not render generated annotated spectra in scalable vector graphics (SVG) format, limiting the flexibility of exported visualizations with regards to figure creation.

Although powerful for the platforms for which they were designed, many of these tools are inseparable from their respective analytical pipelines; data visualization in MaxQuant is only available following processing with the integrated Andromeda search engine, for example. Their purview is therefore limited, and facile spectral annotation is restricted to only those search algorithms packaged in a pipeline with a developed annotator. This restriction poses a problem for numerous applications, especially for alternative peptide fragmentation methods such as ultraviolet photodissociation (UVPD), collisionally supplemented electron-transfer dissociation (ET_hcD), or activated-ion electron-transfer dissociation (AI-ETD) (28–30). Often these methods can be integrated into established analytical pipelines adopted by the field over the

From the Departments of [‡]Chemistry and ^{||}Biomolecular Chemistry, University of Wisconsin–Madison, Madison, Wisconsin 53706; [§]Morgridge Institute for Research, Madison, Wisconsin 53715; [¶]Genome Center of Wisconsin, Madison, Wisconsin 53706

Received November 14, 2018, and in revised form, March 21, 2019

Published, MCP Papers in Press, May 14, 2019, DOI 10.1074/mcp.TIR118.001209

course of several years. But flexible annotation tools are largely unavailable in the beginning stages of method development—arguably when they are needed most. For example, Lorikeet bundles annotation calculations directly with its spectrum viewer. This requires in-depth knowledge of Lorikeet's architecture to add functionality for new technologies. However, separating the annotation process from the spectrum renderer is amenable toward a more stable platform for spectral annotation as the components can be maintained and implemented independently.

Here we present the Interactive Peptide Spectra Annotator (IPSA) to provide a standalone web platform for annotation and interpretation of peptide tandem mass spectra independent of instrumental platform, identification pipeline, and peptide fragmentation technique. IPSA provides flexibility to annotate spectra containing any of the six common peptide fragment ion types. Importantly, it can export annotated data in a tabular format, which enables the rapid culmination of fragment ion statistics for individual or multiple peptide tandem mass spectra, a useful tool in a wide range of proteomic experiments. We have also built in compatibility with spectra collected in the negative mode, providing a much-needed resource for the continued development of negative-mode proteomic approaches. Further, IPSA offers a platform for the generation and exportation of figure-ready annotated spectra in an editable format. In all, IPSA expands spectral annotation capabilities to all types of shotgun proteomic data regardless of how data was collected or processed.

EXPERIMENTAL PROCEDURES

Software Development—IPSA is composed of two major components: a client-facing interactive web visualizer and a server-side data processor which handles the data processing required for spectral annotation. Client-side visualization software was developed using AngularJS. The D3.js library is leveraged to generate interactive annotated spectra using SVG from annotated data returned from the server after analysis (31).

Server-side software consists of a set of modular PHP scripts, which perform form validation, data processing and annotation, file upload handling, and data export. A MySQL database is incorporated to securely cache parsed peptide identifications and spectral information extracted from uploaded data. MySQL integration facilitates data storage and retrieval when annotation requests are submitted to the server.

Example Data Sets—Cell pellets of *Saccharomyces cerevisiae* (strain BY4742) containing $\sim 1 \times 10^8$ cells were harvested from liquid culture by centrifugation ($3000 \times g$, 3 min, 4 °C). The supernatant was removed, and the cell pellet was resuspended in 8 M urea, 100 mM tris (pH 8.0). Methanol was added to 90% by volume and vortexed to lyse the cells and induce protein precipitation. The resulting solution was centrifuged ($14,000 \times g$, 3 min) to form a protein pellet. The supernatant was removed, and the pellet was resuspended in 8 M urea, 100 mM tris (pH 8.0), 10 mM tris(2-carboxyethyl)phosphine, and 40 mM chloroacetamide. The solution was then diluted to 1.5 M urea with 50 mM tris. Trypsin (Promega, Madison, WI) was added (1:50 enzyme/protein) and was allowed to digest overnight (22 °C). The resultant peptides were acidified (pH < 2.0) using 0.1% trifluoroacetic acid (TFA) and were desalted using polymeric reverse phase Strata-X columns. Columns were equilibrated using one bed volume of 100% acetonitrile (ACN), then one bed volume of 0.1% TFA. Peptides were loaded onto the column and washed with two bed volumes of 0.1% TFA. Peptides were eluted by an addition of 500 μ l 40% ACN, 0.1% TFA followed by an addition of 650 μ l 70% ACN, 0.1% TFA and were then dried and resuspended in 0.2% formic acid. Peptide concentration was determined using a Pierce quantitative colorimetric peptide assay (Thermo Fisher Scientific, Rockford, IL).

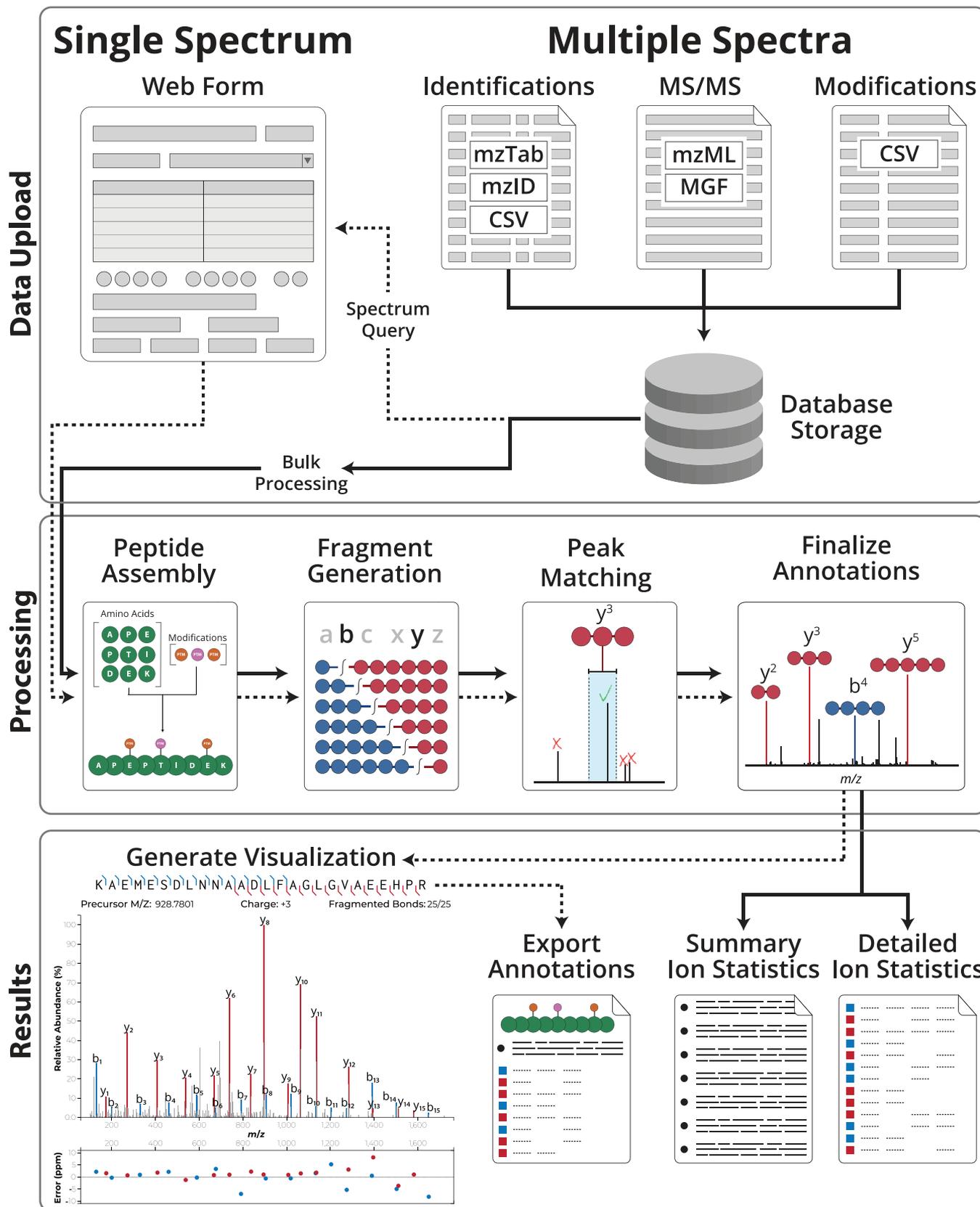
Low pH reverse-phase liquid chromatography was conducted using a Dionex UltiMate 3000 UPLC as described previously (1, 2). Eluting peptides were analyzed using a Q Exactive HF hybrid quadrupole Orbitrap mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) and were fragmented at HCD at 25% normalized collisional energy. Survey scans were taken at a resolution of 60,000 at 200 m/z , whereas tandem mass spectra were collected using a resolution of 15,000 at 200 m/z . The resulting tandem mass spectra were searched using the Coon OMSSA Proteomic Analysis Software Suite (v1.4.1) (32, 33). A precursor mass tolerance of ± 150 ppm was used, whereas fragment ions were searched using a mass tolerance of ± 0.01 Da. A maximum of 3 missed tryptic cleavages were permitted. Carbamidomethylation of cysteine was set as a fixed modification, whereas oxidation of methionine was set as a variable modification. Data was searched against a canonical and isoform *Saccharomyces cerevisiae* database (UniProt, June 10, 2016) concatenated with the reverse protein sequence for decoy generation. A 1% FDR threshold was used at the peptide level, using both e-value and precursor mass accuracy to filter results.

Additional peptide identifications and spectral data were acquired from the previous work of Riley *et al.* to demonstrate IPSA's ability to process PSMs fragmented using alternative dissociation techniques. These include ETD; collisionally supplemented ETD (ETcD and EThcD); AI-ETD; AI-NETD; and AI-ETD with supplemental infrared photon irradiation post-reaction (AI-ETD+) (34, 35).

RESULTS

Design of IPSA—IPSA was developed as a versatile web-based spectral analysis tool capable of individual or *en masse* annotation of PSMs generated from experiments that produce any of the six common peptide fragment ion types (Fig. 1). Single spectra can be annotated by entering peptide and spectral data into an intuitive web form, whereas multiple spectra can be uploaded directly to the website to be individually queried or batch processed. Single annotations are conducted using the metrics provided by the user through the web form and are returned client-side to generate an exportable, annotated spectrum. Exported spectra can easily be shared or integrated into figures. Because the individual interrogation of large numbers of PSMs can quickly become

¹ The abbreviations used are: MS/MS, tandem mass spectrometry; CSV, comma-separated value; MGF, mascot generic format; IPSA, interactive peptide spectral annotator; PSM, peptide spectral match; m/z , mass to charge ratio; ETD, electron transfer dissociation; EThcD, electron transfer and higher-energy collision dissociation; ETciD, electron transfer and collision-induced dissociation; AI-ETD, activated ion electron transfer dissociation; AI-NETD, activated ion negative electron transfer dissociation; UVPD, ultraviolet photodissociation; SVG, scalable vector graphic; ACN, acetonitrile; TFA, trifluoroacetic acid; PPM, parts per million; PTM, post-translational modification; FDR, false discovery rate; JSON, JavaScript Object Notation.



Downloaded from <http://www.mcponline.org/> by guest on October 18, 2019

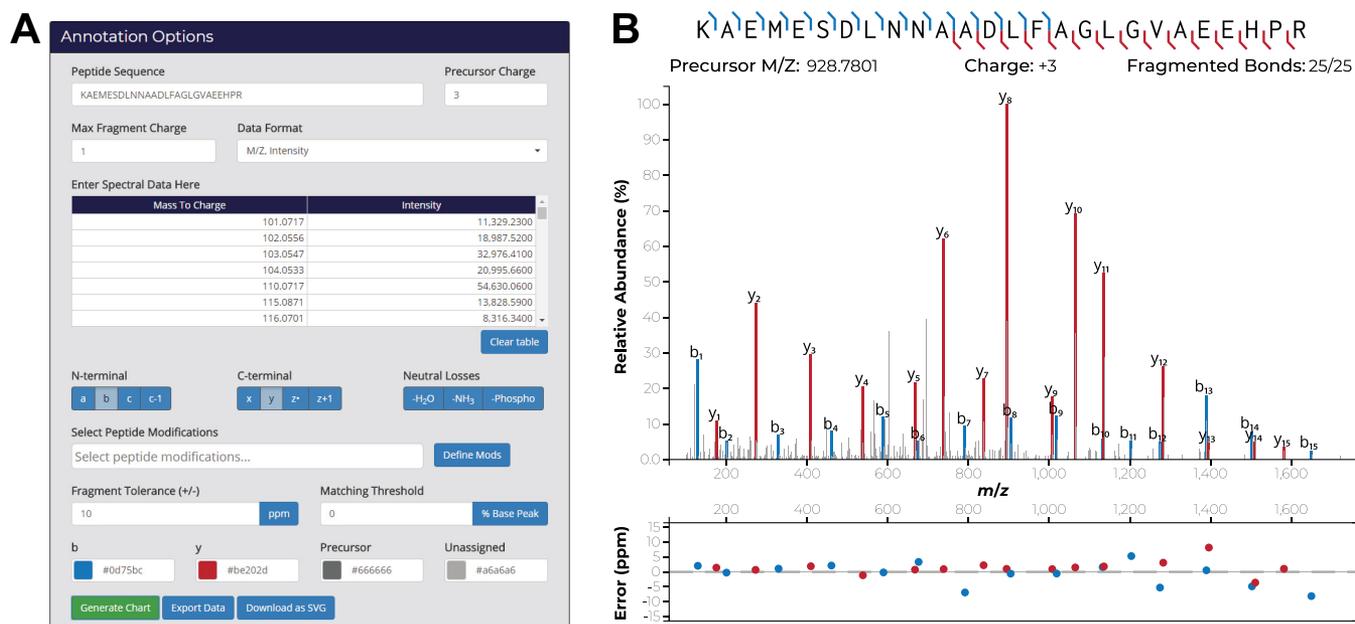


FIG. 2. **Spectrum Annotation.** The peptide KAEMESDLNNAADLFAGLGVAEEHPR and its spectral information provide an example PSM. *A*, Peptide information, fragment ion characteristics, tolerances, and chart colors can be easily set using the provided form. *B*, The generated interactive visualization after server processing containing the peptide sequence marked with the locations of matched fragment ions, an annotated mass spectrum, and visualization of mass error in either parts-per-million or daltons for all matched fragment ions.

tedious, we added functionality to batch process all uploaded PSMs and export the annotations in a tabular format. This feature permits the rapid characterization of tens of thousands of tandem mass spectra.

Single Spectrum Annotation—A single peptide spectrum can be annotated by providing the peptide's sequence, precursor charge, maximum allowed fragment charge, and spectral data to the user interface shown in Fig. 2A. Expected fragmentation patterns and neutral losses can be selected to specify which theoretical peptide fragment ions are generated during data processing (36, 37). The mass tolerance for matching experimental features to theoretical fragment ions can be set in either ppm or Daltons. A relative intensity, raw intensity, or S/N (if supplied with spectral data) cutoff can be defined to ignore low-abundance or insignificant features during matching. Visualization colors can additionally be customized.

A predefined list of common protein post-translational modifications (PTMs) can be queried and selected using a searchable dropdown below the fragmentation options. Available PTMs for a peptide are intelligently filtered to only show PTMs relevant to the entered peptide sequence. If a desired PTM is

not included in the predefined modification list, new PTMs can be defined and are stored locally in the user's web browser. The user can provide a new modification name, target site, and mass shift to create a custom PTM option.

When the server receives an annotation request, data entered into the user interface is validated and sent for processing. The peptide sequence is parsed and assembled into an intact peptide *in-silico*. Theoretical peptide fragment ions are created from the intact peptide using the fragmentation schema selected by the user. Each fragment is matched to *m/z* peak within the specified mass tolerance. To address the case that multiple theoretical fragments are mapped to the same experimental feature, only the theoretical fragment that matches with the smallest mass error is reported. Once annotation mapping has been finalized, annotated spectral data is formatted into JSON and is returned to the client for visualization.

Immediately upon this return, IPSA generates the interactive annotated spectrum (Fig. 2B). This visualization consists of three portions: a peptide sequence marked with detected fragment ion locations and summary statistics, an interactive annotated spectrum, and an interactive scatterplot of the

Fig. 1. **IPSA Software Flowchart.** Single spectra can be annotated by entering peptide and spectral data into provided web forms. Files containing multiple peptide identifications and spectra can be uploaded either in PSI-supported or generic text-based formats to be individually annotated or to be bulk processed for ion statistic extraction. Theoretical peptides are assembled *in-silico*, fragmented, and matched to the experimental spectrum. The annotated experimental spectrum is then returned and visualized client-side. This visualization can be exported as an SVG image for figure generation or as a CSV file containing ion statistics for the single spectrum. Alternatively, ion statistics for all uploaded peptide spectral matches can be calculated and exported through bulk processing, returning two files containing summary and detailed metrics for each uploaded PSM.

matched fragment-ion mass errors. The visualization supports many interactive features to help facilitate data interpretation. Both axes allow contextual zooming for deeper investigation of congested sections of annotated spectra, whereas tooltips provide exact values for any highlighted plotted experimental features. Highlighted fragments are mirrored in each section of the visualization to emphasize all aspects of the feature of interest. Additionally, annotation labels can be dragged to clearer locations to declutter busy regions. The generated visualization can be exported as an SVG file for figure creation as it appears on screen or in a tabular format at any time.

Bulk Data Upload—If many spectra need to be rapidly interrogated, IPSA provides functionality to serially process multiple PSMs by directly uploading files containing peptide identifications and spectral data to the server. Identifications can be provided either in the Proteomics Standards Initiative file formats mzTab or mzIdentML, or in a generic CSV format (18, 19, 38). Each row in the generic CSV lists a scan number, peptide sequence, precursor charge, and all PTM names and locations for each peptide identification. We chose this architecture for its simplicity; peptide identifications produced from a wide variety of search algorithms can easily be converted into this format. Spectral data can be uploaded as a Mascot Generic Format (MGF) or mzML file (12, 39). Finally, a modifications file can be uploaded to link peptide modification names to their respective masses. We provide a set of example files on IPSA's file upload page to demonstrate how each of these files should be structured. MGF and mzML files can easily be generated from vendor or open file formats using conversion tools such as MSConvert (40).

Data parsed from bulk identification and spectral data uploads are stored securely server-side in a MySQL database. On data upload, a unique identifier is assigned to the user's browser which is used to exclusively access the uploader's data. After data extraction, uploaded files are deleted to reduce server footprint. Only one data set can be stored at a time.

Negative Mode Annotation—Proteomic analyses are typically conducted using low-pH separations and positive-mode electrospray ionization to create peptide cations. This tendency leads to a systematic underrepresentation of acidic peptide species, which preferentially ionize as anions (41–43). High-pH separations using negative-mode ionization can be used to better study these acidic species, but the complexity of tandem mass spectra generated using traditional collision-based activation methods has precluded the widespread adoption of this mode. This spectral complexity arises in part from a multitude of neutral losses originating from precursor and fragment ions (44). Alternative fragmentation techniques such as UVPD or AI-NETD, producing *a/x*-, *b/y*-, *c/z*-type and *a*[•]/*x*-type product ions respectively, have recently demonstrated their utility in producing informative tandem mass spectra from peptide anions (35, 43). However, many spectral

annotators do not support these data types. IPSA is capable of annotating PSMs collected using negative-mode electrospray ionization. Fig. 3 demonstrates an IPSA-annotated spectrum of the triply deprotonated peptide LIPSDFILAAQSH-NPIENK dissociated using AI-NETD (35).

Ion Statistics—Obtaining fragment ion statistics in an automated fashion for an entire mass spectrometry experiment is no trivial task. Fragment ion statistics can be greatly informative during method optimization and can be used to monitor MS/MS performance by providing information on what ion types (and in what amounts) are being generated. Additional informative metrics include the sequence coverage of all detected peptide fragments, fragment ion mass errors, and the percent of the total ion current (TIC) that can be explained by annotated fragment ions.

IPSA provides a unique utility among web-based spectral annotators to compute and export all detected fragment ions for an uploaded experiment in a tabular format. The server extracts the fragment ion series, mass tolerances, and any intensity threshold from the provided user interface and serially processes every uploaded peptide identification. The annotation results are continuously written to a set of two downloadable CSVs. The first file contains summary statistics for the matched fragment ions for each uploaded PSM. This file reports the number of matched fragment ions, unique peptide bonds broken, and the percent of the total ion current explained by matched fragment ions. The second file contains detailed information concerning every detected fragment ion for all uploaded identifications; more specifically, the raw intensity, theoretical *m/z*, experimental *m/z*, mass error, percent of base peak, and percent of the total ion current explained is reported.

A series of experiments were previously described by Riley *et al.* to examine the efficacy of ETD, ETcaD, EThcD, AI-ETD, and AI-ETD+ fragmentation on a liquid chromatography timescale (34). The authors found AI-ETD+ to be the optimal supplemental ETD fragmentation technique. Using the authors' reported peptide identifications and spectral data, we created a set of detailed comparisons similar to those made in the referenced manuscript using the ion statistics files directly exported from IPSA (Fig. 4). No further programming was required to extract these data or make this figure, and all data manipulation postexport was performed in a spreadsheet using basic arithmetical functions.

In summary, IPSA is capable of both cleanly annotating peptide spectra collected using a wide variety of dissociation techniques in both positive and negative mode and of exporting the generated annotated spectra in the editable SVG format. Additionally, IPSA allows the bulk analysis of detected fragment ions for any number of uploaded spectra, permitting in turn the deep interrogation of data without requiring programming experience.

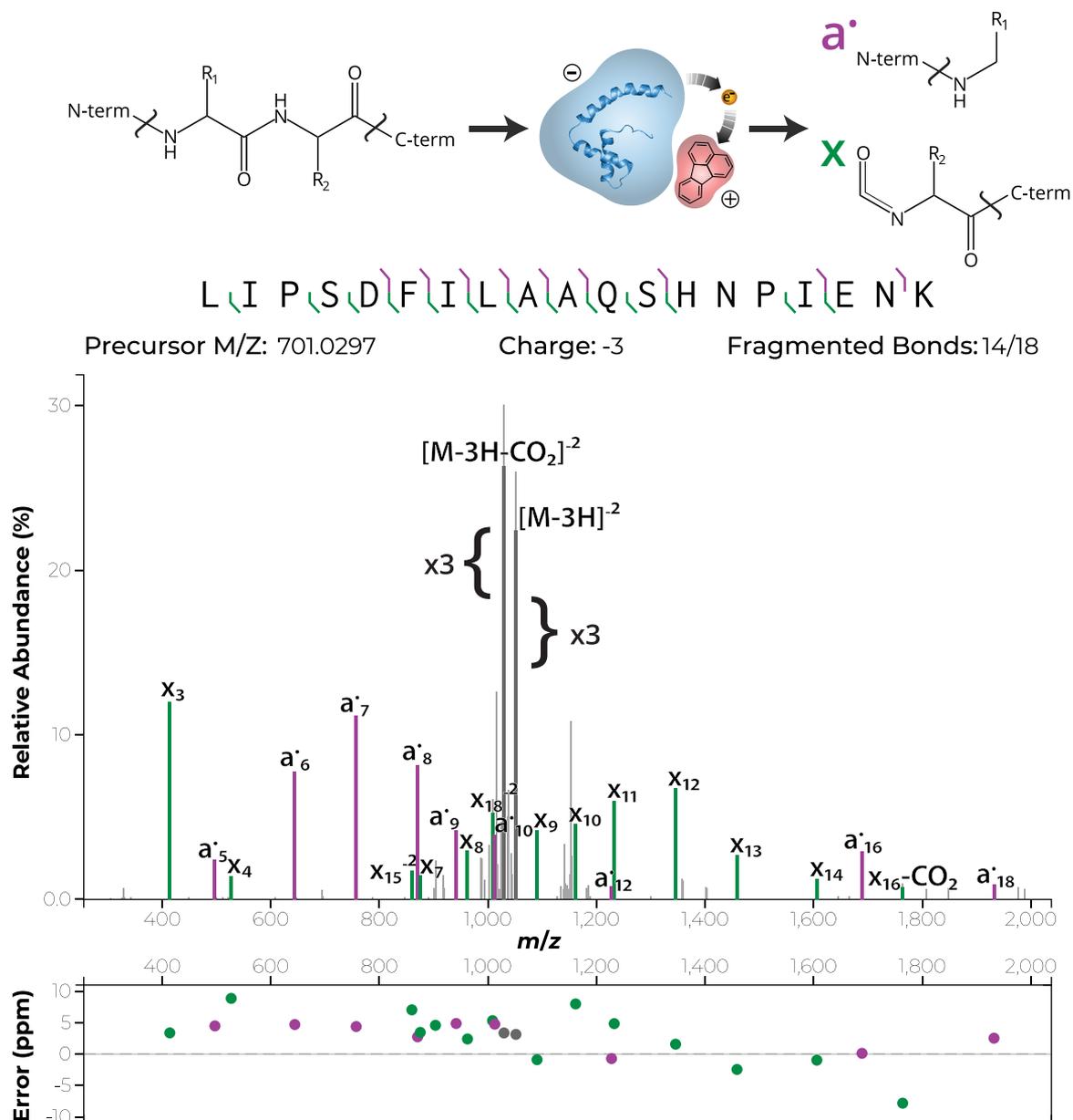


Fig. 3. **Negative-Mode Annotation.** The annotated AI-NETD spectrum of the triply deprotonated peptide LIPSDFILAAQSHNPIENK generated using IPISA. The charge-reduced precursor was downscaled by a factor of 3. Unreacted precursor was cleaned from this spectrum.

DISCUSSION

Modern MS-based proteomics techniques are widely used to identify and characterize tens of thousands of peptides and proteins originating from a variety of biological samples. The annotation of the tandem mass spectra used to identify these species is an arduous task requiring extensive expertise. Our web-based and open-source peptide spectral annotator, IPISA, provides a resource for generating and investigating annotated spectra for peptide identifications to a wide research community. IPISA can generate customizable annotated peptide spectra using a clean and intuitive user interface, allowing researchers to export customizable, publication-ready annotated spectra as vector graphics to aid in figure creation. It can

process MS/MS spectra from both anionic and cationic precursors, and it has built-in support to annotate fragment ions generated from a diverse assortment of dissociative techniques. Additionally, IPISA can extract fragment ion statistics from any number of peptide spectra and return results in a tabular format, giving researchers a deeper and more comprehensive view of their peptide analyses.

We chose to develop IPISA as an online platform to reach a wide audience of proteomics researchers: those with an Internet connection on a computer with a web browser. Web-based software also allowed us to use the flexibility of the well-established JavaScript visualization library D3.js while avoiding software compatibility issues and version control.

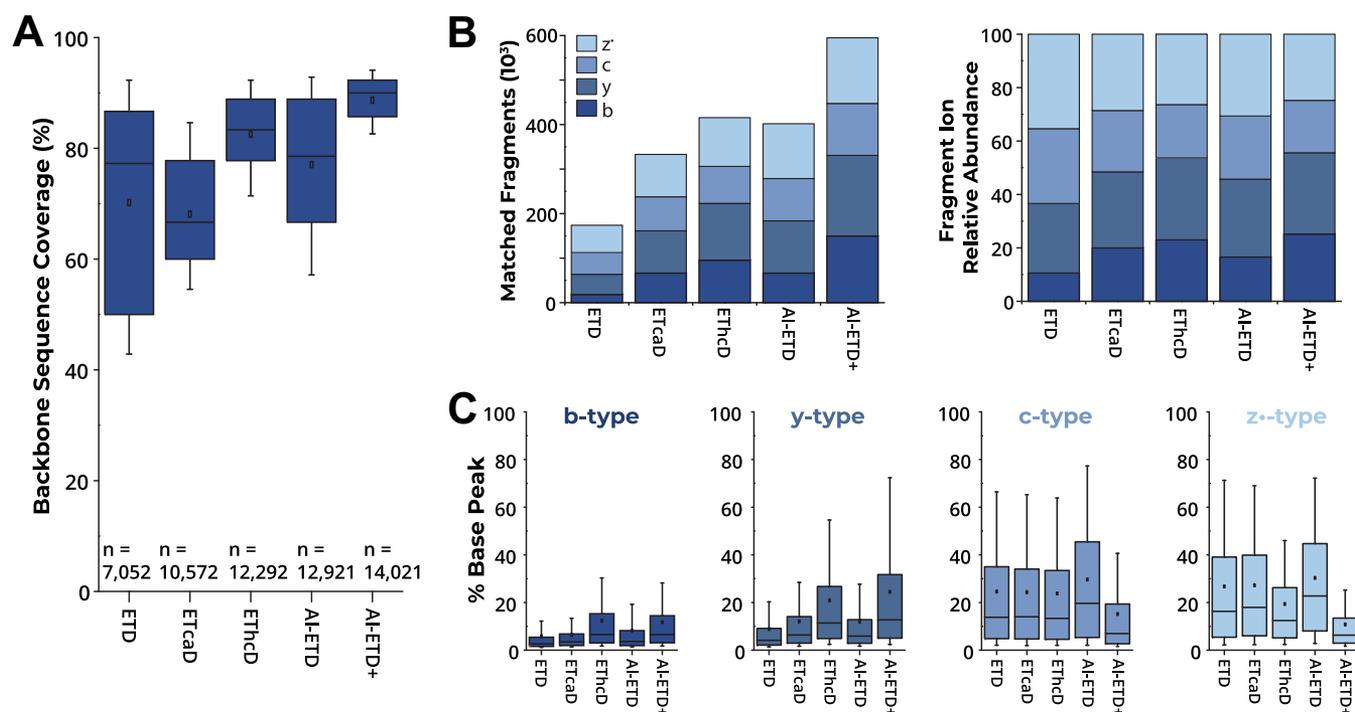


FIG. 4. Fragment Ion Statistics. Exported fragment statistics facilitate comparisons between different electron transfer dissociation techniques. IPSA matching parameters were set to a 10 ppm mass error window and a 1% base peak intensity cutoff. **A**, AI-ETD+ provides superior peptide backbone fragmentation compared with ETD, ETcAd, ETHcD, and AI-ETD. **B**, AI-ETD+ generates higher counts of fragment ions over an experiment compared with other techniques. **C**, Distributions of all matched fragment ions for each respective dissociative technique.

Through IPSA, we aim to increase the approachability of spectral annotation to proteomics novices and experts alike.

The IPSA source code is freely available for inspection and download at <https://github.com/coongroup/IPSA> alongside additional guides regarding software usage. We recommend using an updated web browser to access IPSA at <http://interactivepeptidespectralannotator.com> as outdated browsers may not provide support for critical functions. IPSA can be easily installed on a private desktop or server using a prebuilt Docker image and instructions at <https://hub.docker.com/r/dbrademan/ipsa>, or IPSA's project files can be manually configured to operate on private web servers with full functionality. Additionally, the JavaScript file used to render the interactive visualization, *IPSA.js*, is configured to be used as an AngularJS directive. This directive can be attached to custom annotation scripts in many website environments, allowing the use of our software beyond that of the platform we described here.

Acknowledgments—We thank Kevin Schauer for providing feedback during IPSA's design.

DATA AVAILABILITY

Raw spectral data, peptide identifications, and protein databases have been deposited to the ProteomeXchange Consortium via the PRIDE (45) partner repository with the dataset identifier PXD011695.

* This work was supported by NIH grants R35 GM118110, P41 GM108538, and DOE Great Lakes Bioenergy Research Center Office of Science BER Award Number DE-SC0018409. DRB was funded through an NHGRI training grant awarded to the Genomic Sciences Training Program (5T32HG002760). NMR was funded by an NIH Predoctoral to Postdoctoral Transition Award (F99 CA212454).

** To whom correspondence should be addressed. E-mail: jcoon@chem.wisc.edu.

Author contributions: D.R.B., N.M.R., N.W.K., and J.J.C. designed research; D.R.B. and N.M.R. performed research; D.R.B. contributed new reagents/analytic tools; D.R.B. and N.M.R. analyzed data; D.R.B., N.M.R., and J.J.C. wrote the paper.

REFERENCES

1. Hebert, A. S., Richards, A. L., Bailey, D. J., Ulbrich, A., Coughlin, E. E., Westphall, M. S., and Coon, J. J. (2014) The One Hour Yeast Proteome. *Mol. Cell Proteomics* **13**, 339–347
2. Richards, A. L., Hebert, A. S., Ulbrich, A., Bailey, D. J., Coughlin, E. E., Westphall, M. S., and Coon, J. J. (2015) One-hour proteome analysis in yeast. *Nat. Protoc.* **10**, 701–714
3. Senko, M. W., Remes, P. M., Canterbury, J. D., Mathur, R., Song, Q., Eliuk, S. M., Mullen, C., Earley, L., Hardman, M., Blethrow, J. D., Bui, H., Specht, A., Lange, O., Denisov, E., Makarov, A., Horning, S., and Zabrouskov, V. (2013) Novel parallelized quadrupole/linear ion trap/orbitrap hybrid mass spectrometer improving proteome coverage and peptide identification rates. *Anal. Chem.* **85**, 11710–11714
4. Richards, A. L., Merrill, A. E., and Coon, J. J. (2015) Proteome sequencing goes deep. *Curr. Opin. Chem. Biol.* **24**, 11–17
5. Hebert, A. S., Thöing, C., Riley, N. M., Kwiecien, N. W., Shiskova, E., Huguet, R., Cardasis, H. L., Kuehn, A., Eliuk, S., Zabrouskov, V., Westphall, M. S., McAlister, G. C., and Coon, J. J. (2018) Improved precursor characterization for data-dependent mass spectrometry. *Anal. Chem.* **90**, 2333–2340

6. Scheltema, R. A., Hauschild, J.-P., Lange, O., Hornburg, D., Denisov, E., Damoc, E., Kuehn, A., Makarov, A., and Mann, M. (2014) The Q exactive HF, a benchtop mass spectrometer with a pre-filter, high-performance quadrupole and an ultra-high-field orbitrap analyzer. *Mol. Cell Proteomics* **13**, 3698–3708
7. Kelstrup, C. D., Bekker-Jensen, D. B., Arrey, T. N., Högrefe, A., Harder, A., and Olsen, J. V. (2018) Performance evaluation of the Q exactive HF-X for shotgun proteomics. *J. Proteome Res.* **17**, 727–738
8. Bekker-Jensen, D. B., Kelstrup, C. D., Batth, T. S., Larsen, S. C., Haldrup, C., Bramsen, J. B., Sørensen, K. D., Høyer, S., Ørntoft, T. F., Andersen, C. L., Nielsen, M. L., and Olsen, J. V. (2017) An optimized shotgun strategy for the rapid generation of comprehensive human proteomes. *Cell Syst.* **4**, 587–599.e4
9. Shishkova, E., Hebert, A. S., and Coon, J. J. (2016) Now, more than ever, proteomics needs better chromatography. *Cell Syst.* **3**, 321–324
10. Eng, J. K., McCormack, A. L., and Yates, J. R. (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *Am. Soc. Mass Spectrom.* **5**, 976–989
11. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372
12. Perkins, D. N., Pappin, D. J. C. C., Creasy, D. M., and Cottrell, J. S. (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551–3567
13. Taylor, J. A., and Johnson, R. S. (1997) Sequence database searches via de Novo peptide sequencing by tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **11**, 1067–1075
14. Ma, B., Zhang, K., Hendrie, C., Liang, C., Li, M., Doherty-Kirby, A., and Lajoie, G. (2003) PEAKS: Powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **17**, 2337–2342
15. Chi, H., Sun, R. X., Yang, B., Song, C. Q., Wang, L. H., Liu, C., Fu, Y., Yuan, Z. F., Wang, H. P., He, S. M., and Dong, M. Q. (2010) PNovo: De novo peptide sequencing and identification using HCD spectra. *J. Proteome Res.* **9**, 2713–2724
16. Sinitcyn, P., Daniel Rudolph, J., Cox, J., Rudolph, J. D., and Cox, J. (2018) Computational methods for understanding mass spectrometry-based shotgun proteomics data. *Annu. Rev. Biomed. Data Sci.* **1**, annurev-biodatasci-080917-013516
17. Bradshaw, R. A., Burlingame, A. L., Carr, S., and Aebersold, R. (2006) Reporting protein identification data: the next generation of guidelines. *Mol. Cell Proteomics* **5**, 787–788
18. Jones, A. R., Eisenacher, M., Mayer, G., Kohlbacher, O., Siepen, J., Hubbard, S. J., Selley, J. N., Searle, B. C., Shofstahl, J., Seymour, S. L., Julian, R., Binz, P.-A., Deutsch, E. W., Hermjakob, H., Reisinger, F., Griss, J., Vizcaino, J. A., Chambers, M., Pizarro, A., and Creasy, D. (2012) The mzIdentML data standard for mass spectrometry-based proteomics results. *Mol. Cell Proteomics* **11**, M111.014381
19. Seymour, S. L., Farrah, T., Binz, P. A., Chalkley, R. J., Cottrell, J. S., Searle, B. C., Tabb, D. L., Vizcaino, J. A., Prieto, G., Uszkoreit, J., Eisenacher, M., Martínez-Bartolomé, S., Ghali, F., and Jones, A. R. (2014) A standardized framing for reporting protein identifications in mzIdentML 1.2. *Proteomics* **14**, 2389–2399
20. Burlingame, A., Carr, S. A., Bradshaw, R. A., and Chalkley, R. J. (2015) On credibility, clarity, and compliance. *Mol. Cell Proteomics* **14**, 1731–1733
21. Baker, P. R., and Chalkley, R. J. (2014) MS-viewer: a web-based spectral viewer for proteomics results. *Mol. Cell Proteomics* **13**, 1392–1396
22. Strohal, M., Hassman, M., Košata, B., and Kodíček, M. (2008) mMass data miner: An open source alternative for mass spectrometric data analysis. *Rapid Commun. Mass Spectrom.* **22**, 905–908
23. Colinge, J., Masselot, A., Carbonell, P., and Appel, R. D. (2006) InSilico-Spectro: An open-source proteomics library. *J. Proteome Res.* **5**, 619–624
24. Sharma, V., Eng, J. K., MacCoss, M. J., and Riffle, M. (2012) A mass spectrometry proteomics data management platform. *Mol. Cell Proteomics* **11**, 824–831
25. Riffle, M., Jaschob, D., Zelter, A., and Davis, T. N. (2016) ProXL (Protein Cross-Linking Database): A platform for analysis, visualization, and sharing of protein cross-linking mass spectrometry data. *J. Proteome Res.* **15**, 2863–2870
26. Perez-Riverol, Y., Alpi, E., Wang, R., Hermjakob, H., and Vizcaino, J. A. (2015) Making proteomics data accessible and reusable: Current state of proteomics databases and repositories. *Proteomics* **15**, 930–950
27. Riffle, M., Merrihew, G. E., Jaschob, D., Sharma, V., Davis, T. N., Noble, W. S., and MacCoss, M. J. (2015) Visualization and dissemination of multidimensional proteomics data comparing protein abundance during *Caenorhabditis elegans* development. *J. Am. Soc. Mass Spectrom.* **26**, 1827–1836
28. Ly, T., and Julian, R. R. (2009) Ultraviolet photodissociation: developments towards applications for mass-spectrometry-based proteomics. *Angew. Chemie - Int. Ed.* **48**, 7130–7137
29. Yu, Q., Wang, B., Chen, Z., Urabe, G., Glover, M. S., Shi, X., Guo, L. W., Kent, K. C., and Li, L. (2017) Electron-transfer/higher-energy collision dissociation (ETHC/D)-enabled intact glycopeptide/glycoproteome characterization. *J. Am. Soc. Mass Spectrom.* **28**, 1751–1764
30. Ledvina, A. R., Beauchene, N. A., McAlister, G. C., Syka, J. E. P., Schwartz, J. C., Griep-Raming, J., Westphall, M. S., and Coon, J. J. (2010) Activated-ion electron transfer dissociation improves the ability of electron transfer dissociation to identify peptides in a complex mixture. *Anal. Chem.* **82**, 10068–10074
31. Bostock, M., Ogievetsky, V., and Heer, J. (2011) D³ data-driven documents. *IEEE Trans. Vis. Comput. Graph.* **17**, 2301–2309
32. Geer, L. Y., Markey, S. P., Kowalak, J. A., Wagner, L., Xu, M., Maynard, D. M., Yang, X., Shi, W., and Bryant, S. H. (2004) Open mass spectrometry search algorithm. *J. Proteome Res.* **3**, 958–964
33. Wenger, C. D., Phanstiel, D. H., Lee, M. V., Bailey, D. J., and Coon, J. J. (2011) COMPASS: A suite of pre- and post-search proteomics software tools for OMSSA. *Proteomics* **11**, 1064–1074
34. Riley, N. M., Westphall, M. S., Hebert, A. S., and Coon, J. J. (2017) Implementation of activated ion electron transfer dissociation on a quadrupole-orbitrap-linear ion trap hybrid mass spectrometer. *Anal. Chem.* **89**, 6358–6366
35. Riley, N. M., Rush, M. J. P., Rose, C. M., Richards, A. L., Kwiecien, N. W., Bailey, D. J., Hebert, A. S., Westphall, M. S., and Coon, J. J. (2015) The negative mode proteome with activated ion negative electron transfer dissociation (AI-NETD). *Mol. Cell Proteomics* **14**, 2644–2660
36. Roepstorff, P., and Fohlman, J. (1984) Letter to the editors. *Biol. Mass Spectrom.* **11**, 601–601
37. Johnson, R. S., Martin, S. A., Biemann, K., Stults, J. T., and Watson, J. T. (1987) Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: differentiation of leucine and isoleucine. *Anal. Chem.* **59**, 2621–2625
38. Griss, J., Jones, A. R., Sachsenberg, T., Walzer, M., Gatto, L., Hartler, J., Thallinger, G. G., Salek, R. M., Steinbeck, C., Neuhauser, N., Cox, J., Neumann, S., Fan, J., Reisinger, F., Xu, Q.-W., Del Toro, N., Pérez-Riverol, Y., Ghali, F., Bandeira, N., Xenarios, I., Kohlbacher, O., Vizcaino, J. A., and Hermjakob, H. (2014) The mzTab data exchange format: communicating mass-spectrometry-based proteomics and metabolomics experimental results to a wider audience. *Mol. Cell Proteomics* **13**, 2765–2775
39. Martens, L., Chambers, M., Sturm, M., Kessner, D., Levander, F., Shofstahl, J., Tang, W. H., Römpf, A., Neumann, S., Pizarro, A. D., Montecchi-Palazzi, L., Tasman, N., Coleman, M., Reisinger, F., Souda, P., Hermjakob, H., Binz, P.-A., and Deutsch, E. W. (2011) mzML—a community standard for mass spectrometry data. *Mol. Cell Proteomics* **10**, R110.000133
40. Chambers, M. C., Maclean, B., Burke, R., Amode, D., Ruderman, D. L., Neumann, S., Gatto, L., Fischer, B., Pratt, B., Egerton, J., Hoff, K., Kessner, D., Tasman, N., Shulman, N., Frewen, B., Baker, T. A., Brusniak, M.-Y., Paulse, C., Creasy, D., Flashner, L., Kani, K., Moulding, C., Seymour, S. L., Nuwaysir, L. M., Lefebvre, B., Kuhlmann, F., Roark, J., Rainer, P., Detlev, S., Hemenway, T., Huhmer, A., Langridge, J., Connolly, B., Chadick, T., Holly, K., Eckels, J., Deutsch, E. W., Moritz, R. L., Katz, J. E., Agus, D. B., MacCoss, M., Tabb, D. L., and Mallick, P. (2012) A cross-platform toolkit for mass spectrometry and proteomics. *Nat. Biotechnol.* **30**
41. McAlister, G. C., Russell, J. D., Rumachik, N. G., Hebert, A. S., Syka, J. E. P., Geer, L. Y., Westphall, M. S., Pagliarini, D. J., and Coon, J. J.

- (2012) Analysis of the acidic proteome with negative electron-transfer dissociation mass spectrometry. *Anal. Chem.* **84**, 2875–2882
42. Rumachik, N. G., McAlister, G. C., Russell, J. D., Bailey, D. J., Wenger, C. D., and Coon, J. J. (2012) Characterizing peptide neutral losses induced by negative electron-transfer dissociation (NETD). *J. Am. Soc. Mass Spectrom.* **23**, 718–727
43. Madsen, J. A., Xu, H., Robinson, M. R., Horton, A. P., Shaw, J. B., Giles, D. K., Kaoud, T. S., Dalby, K. N., Trent, M. S., and Brodbelt, J. S. (2013) High-throughput database search and large-scale negative polarity liquid chromatography–tandem mass spectrometry with ultraviolet photodissociation for complex proteomic samples. *Mol. Cell Proteomics* **12**, 2604–2614
44. Ewing, N. P., and Cassady, C. J. (2001) Dissociation of multiply charged negative ions for hirudin (54–65), fibrinopeptide B, and insulin A (oxidized). *J. Am. Soc. Mass Spectrom.* **12**, 105–116
45. Vizcaino, J. A., Csordas, A., del-Toro, N., Dienes, J. A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Ternent, T., Xu, Q.-W., Wang, R., and Hermjakob, H. (2016) 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* **44**, D447–D456