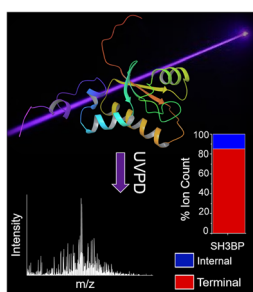


## RESEARCH ARTICLE

# The Ups and Downs of Repeated Cleavage and Internal Fragment Production in Top-Down Proteomics

Yana A. Lyon,<sup>1</sup> Dylan Riggs,<sup>1</sup> Luca Fornelli,<sup>2</sup> Philip D. Compton,<sup>2</sup> Ryan R. Julian<sup>1</sup><sup>1</sup>Department of Chemistry, University of California, Riverside, 501 Big Springs Road, Riverside, CA 92521, USA<sup>2</sup>Departments of Chemistry and Molecular Biosciences, and the Proteomics Center of Excellence, Northwestern University, 2145 N. Sheridan Road, Evanston, IL 60208, USA

**Abstract.** Analysis of whole proteins by mass spectrometry, or top-down proteomics, has several advantages over methods relying on proteolysis. For example, proteoforms can be unambiguously identified and examined. However, from a gas-phase ion-chemistry perspective, proteins are enormous molecules that present novel challenges relative to peptide analysis. Herein, the statistics of cleaving the peptide backbone multiple times are examined to evaluate the inherent propensity for generating internal versus terminal ions. The raw statistics reveal an inherent bias favoring production of terminal ions, which holds true regardless of protein size. Importantly, even if the full suite of internal ions is generated by statistical dissociation, terminal ions are predicted to account for at least 50% of the total ion current,

regardless of protein size, if there are three backbone dissociations or fewer. Top-down analysis should therefore be a viable approach for examining proteins of significant size. Comparison of the purely statistical analysis with actual top-down data derived from ultraviolet photodissociation (UVPD) and higher-energy collisional dissociation (HCD) reveals that terminal ions account for much of the total ion current in both experiments. Terminal ion production is more favored in UVPD relative to HCD, which is likely due to differences in the mechanisms controlling fragmentation. Importantly, internal ions are not found to dominate from either the theoretical or experimental point of view.

**Keywords:** UVPD, HCD, Statistical analysis, Internal ion

Received: 28 April 2017/Revised: 8 September 2017/Accepted: 23 September 2017/Published Online: 16 October 2017

## Introduction

Top-down analysis of whole proteins is a rapidly expanding area in mass spectrometry attributable to several unique capabilities [1–11]. Proteins synthesized from a single gene can exist in several unique molecular forms because of events such as genetic mutation, alternative splicing, and post-translational modifications [12]. Top-down mass spectrometry makes it possible to distinguish these proteoforms and their biological roles [13]. Additionally, information about three-dimensional structure can be obtained through fragmentation of whole proteins and protein complexes [14, 15]. The suite of widely available methods for top-down fragmentation includes electron capture dissociation, electron transfer dissociation (ECD, ETD)

[16, 17], and higher-energy collisional dissociation (HCD) [18], with ultraviolet photodissociation (UVPD) also rising in popularity [19]. Relative to peptides, whole proteins are much larger molecules with vastly more atoms and vibrational degrees of freedom. To fragment such large species requires either (1) significantly higher total energy input or (2) weakening of specific bonds [20]. HCD operates under the first principle whereas ECD/ETD utilize primarily the second, and UVPD likely functions via a combination of both. Regardless of which fragmentation paradigm is relevant for a given experiment, the potential for fragmenting the same protein multiple times must be considered. In other words, multiple photons or electrons can initiate independent fragmentation events, or for HCD the energy can be sufficiently high that product ions subsequently undergo secondary fragmentation.

Cleaving the peptide backbone in multiple places creates internal fragment ions, which have not received significant attention in the literature and are typically ignored during analysis in most experiments. Furthermore, there is a general misunderstanding within the mass spectrometry community of

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s13361-017-1823-8>) contains supplementary material, which is available to authorized users.

Correspondence to: Ryan Julian; e-mail: ryan.julian@ucr.edu

the statistics of internal ions in top-down analysis. Because the number of possible internal ions grows significantly with protein size, it is often assumed that internal ions will consume most of the ion intensity when large proteins are fragmented. Only a few reports have examined internal ion generation within the context of peptide dissociation [21, 22]. From a top-down perspective, Kelleher and co-workers previously reported that inclusion of internal fragment ions can increase sequence coverage for smaller proteins [23].

Herein we report the statistical analysis of internal fragment ions to reveal the inherent propensities for generating internal versus terminal fragment ions as a function of both protein size and the number of dissociation events. It is confirmed that as molecular size increases, the total number of unique internal ions that can be generated increases dramatically. However, the statistical probability for making *any given* internal fragment ion also decreases simultaneously at the same rate while the propensity for making terminal ions remains constant. Terminal ions are therefore predicted to be generated frequently, regardless of protein size. Although multiple dissociation events (i.e., more than two) can increase the proportion of internal ions generated from each precursor ion, terminal ions are still statistically predicted to dominate spectra, even with increasing protein size. Indeed, the fraction of precursor ion charge that ends up on terminal ions does not depend on protein size and can be easily estimated if the number of dissociation events is known. The statistical results were compared with actual HCD and UVPD data for proteins of varying sizes, revealing that a significant fraction of the ion intensity is attributed to terminal ions regardless of protein size. In general, UVPD results in a greater fraction of ion intensity being apportioned to terminal ions and more closely resembles statistical fragmentation.

## Experimental

### Simulations

A program was written in [VB.net](#) for statistical analysis of internal and terminal ions. With a given input sequence, the program cleaves the protein amide bond from one to four times at random locations to produce b and y ions and internal fragments. The masses of each fragment produced by a single run are recorded and the frequency of each product is then summed after multiple runs have been carried out. For each analysis, the number of runs was sufficient to sample all possible product ions repeatedly (typically hundreds of times). All cleavage points are treated equally, i.e., there is no provision for preferred cleavage at certain sites. It is assumed that disulfide bonds have all been reduced. All product ions are recorded, including single amino acids and short sequences without basic residues. Many of these products would be unlikely to be detected in real experiments, as is discussed in greater detail below.

### Mass Spectrometry

Human whole cell lysate was prepared as previously described [24]. Briefly, primary IMR90 human fibroblasts were lysed

using a buffer composed of 4% SDS (w/v), 10 mM Tris-HCl (pH 7.8), 1 mM dithiothreitol, 10 mM sodium butyrate, in the presence of a protease inhibitors cocktail (Thermo Scientific), and the protein content was cleaned by acetone precipitation and subsequently fractionated using a GELFREE 8100 Fractionation System (Expedeon, Harston, Cambridgeshire, UK) equipped with a 10% and 8% GELFREE 8100 cartridge. Selected eluted fractions in the mass range from 0 to 30 kDa were cleaned from SDS and other contaminants via MeOH/H<sub>2</sub>O/CHCl<sub>3</sub> extraction. Dried protein pellets were resuspended in 25  $\mu$ L of Solution A (composed of 5% acetonitrile and 0.2% formic acid in water). Protein reversed-phase chromatographic separation was performed using a Dionex Ultimate 3000 chromatographic system (Thermo Scientific, Sunnyvale, CA, USA), with the nanobore column packed with  $\sim$ 30 cm of PLRP-S stationary phase (75  $\mu$ m i.d., 5  $\mu$ m particle size, Agilent, Santa Clara, CA, USA) on-line coupled to a nanoelectrospray ionization source. The analytical gradient was set as follows: Solution B (4.8% water in acetonitrile with 0.2% formic acid) was raised from 5% to 15% in 2 min, then from 15% to 50% in 50 min (followed by a wash at 95% B and re-equilibration at 5% B). The column outlet was linked to a 15  $\mu$ m i.d. electrospray emitter (New Objective, Woburn, MA, USA), packed with  $\sim$ 0.5 mm of PLRP-S resin to avoid outgassing, through a high voltage union to which a  $\sim$ 2 kV potential was applied.

All mass spectrometry measurements were carried out on an Orbitrap Fusion Lumos mass spectrometer (Thermo Scientific, San Jose, CA, USA) operating in “protein mode,” under reduced (2 mTorr) N<sub>2</sub> pressure in the HCD cell combined with “extended trapping” of ions in the HCD cell. The transfer capillary temperature was set at 320  $^{\circ}$ C, the source rf was set at 30% amplitude, and a 15 V offset was applied to favor protein ion desolvation and adduct removal. Acquisition parameters were set as follows: broadband MS scans used a resolving power (r.p.) of 120,000 (at 200  $m/z$ ) and were acquired within a 500–2000  $m/z$  window averaging 4 microscans, with an AGC target 1e6 and a maximum injection time of 200 ms. Tandem MS (MS<sup>2</sup>) was performed in a data-dependent fashion by quadrupole-isolating the two most abundant species in the MS<sup>1</sup> spectrum using a 3  $m/z$ -wide isolation window. Ion activation/fragmentation used either HCD or UVPD @213 nm. For HCD, a normalized collision energy of 23% was used. UVPD was performed in the high pressure chamber of the linear ion trap using the fifth harmonic (corresponding to 213 nm photons,  $\sim$ 50  $\mu$ J/pulse) of a Nd:YAG solid-state laser (CryLas, Berlin, Germany). The number of laser pulses was varied depending on the average protein size of the analyzed GELFrEE fraction (typically 20–40 pulses were used). All MS<sup>2</sup> scans were recorded with a r.p. of 60,000 (at 200  $m/z$ ) over a 400–2000  $m/z$  window and four microscans, using an AGC target of 1e6 charges and a maximum injection time of 800 ms. Recorded .RAW files were analyzed using TDPportal (<https://portal.nrtdp.northwestern.edu/>) for protein identification, using a previously described workflow for human database search [25]. Selected

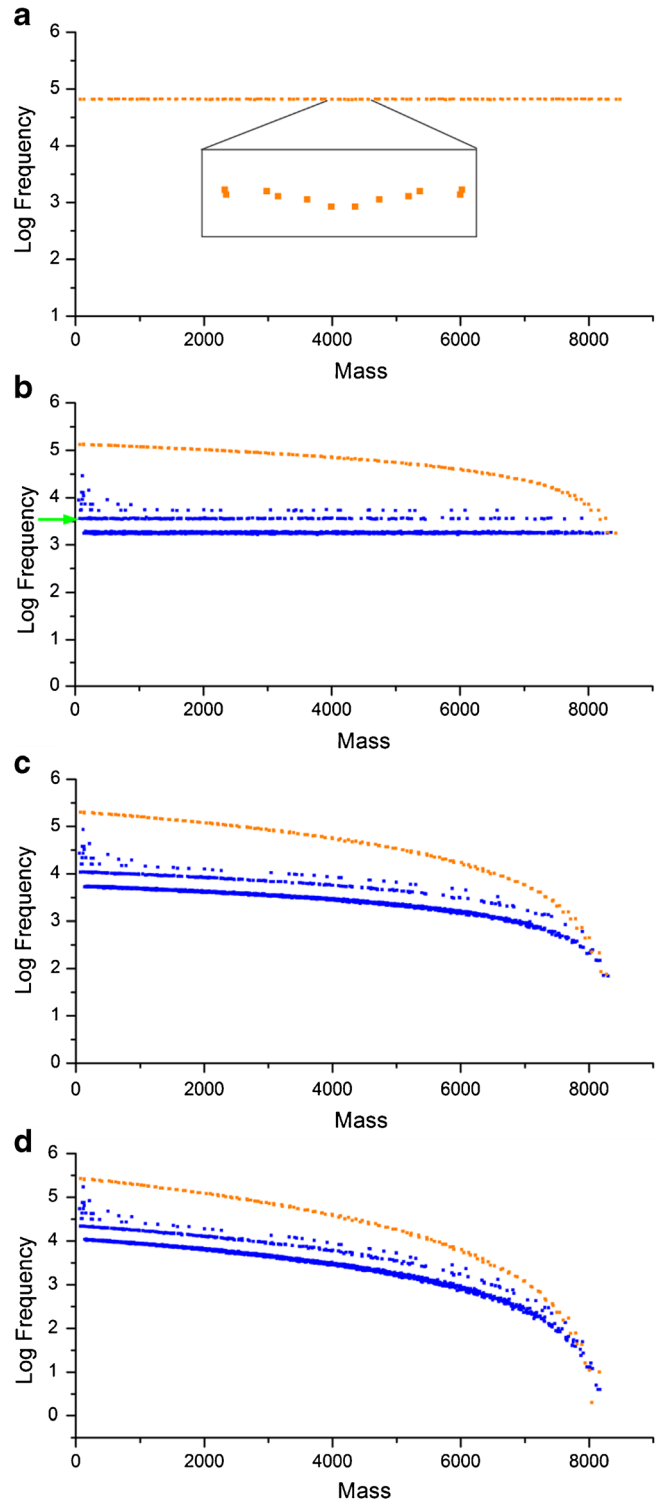
proteoforms, initially identified via TDPortal, were manually annotated, with spectral deconvolution performed using Xtract (Thermo Scientific).

## Results and Discussion

### Statistical Dissociation

To understand the statistical propensities for internal and terminal ion production in top-down analysis, stochastic fragmentation of various proteins was carried out *in silico* as detailed in the Experimental section. Results for ubiquitin are shown in Figure 1, where the number of times each fragment was observed after repeated trials is plotted as a function of mass. Fragmentation was simulated in a completely random fashion, with no preference for any cleavage site. The results for breaking the backbone a single time are shown in Figure 1a, which illustrates a completely flat abundance profile. Single cleavage of the backbone yields complementary N-terminal and C-terminal fragments, and close inspection of the results in Figure 1a reveals symmetry that results from this complementarity (see zoomed box in Figure 1a). As expected, all fragments are generated with essentially equal probability. Figure 1b shows results for the same approach where the backbone was cleaved twice, revealing distinct populations with significantly different abundances. The terminal fragments (orange) occur with much greater frequency than internal fragments (blue) for most of the mass range. The abundance of terminal fragments drops as mass increases because the peptide backbone is cleaved in two locations. In order to observe a large mass terminal ion, both fragmentation points must be located near the opposite terminus, which is statistically unlikely. Owing to this requirement, the longest terminal fragments are generated with frequency similar to the most abundant internal fragments.

Double cleavage yields two terminal fragments and one internal fragment per precursor ion, but this difference alone is insufficient to account for the observed disparity in abundance. A more compelling explanation is that any given internal fragment is generated (on average) less frequently due to dilution because the total number of internal fragments greatly exceeds the total number of terminal ions. For terminal fragments of ubiquitin, there are only 150 possibilities that can be populated, but there are 2401 unique mass internal fragments. The likelihood for generating a particular internal fragment is therefore significantly lower, although there is no inherent length bias for internal fragments in double cleavage, as illustrated by the flatness of the distribution. However, the density of internal fragments within a given mass window decreases with increasing mass (see Supporting Information for a histogram). There are also privileged internal fragments that occur with significantly higher frequency than most. Those at low mass correspond to single amino acids, with the highest frequency peak matching leucine/isoleucine, the most abundant amino acids in ubiquitin. The secondary line of internal fragments just above the primary distribution (indicated by the



**Figure 1.** Statistical fragmentation of ubiquitin for (a) single, (b) double, (c) triple, and (d) quadruple cleavages of the peptide backbone. Terminal fragments are displayed in orange, internal fragments, blue. Green arrow indicates single-frame-shifting internal fragments.

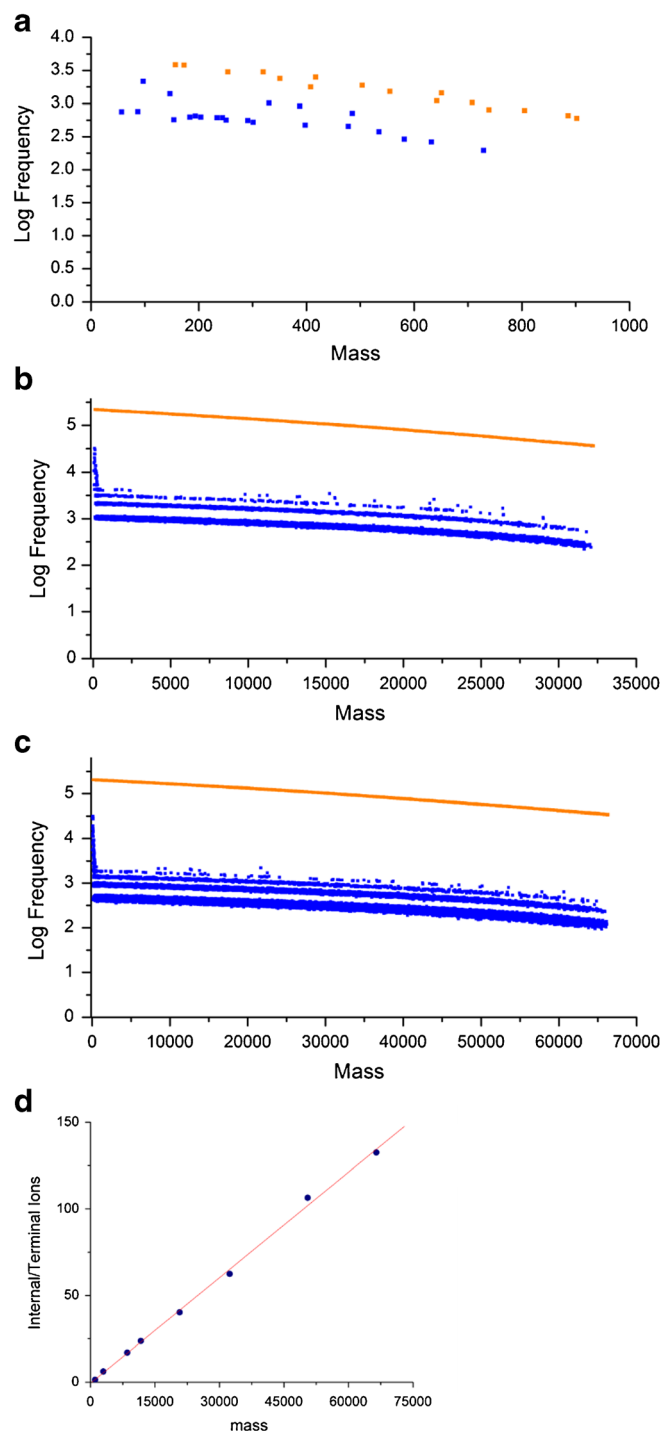
green arrow) correspond to frame-shifting fragments, i.e., fragments that are flanked by the same amino acid on either side, creating two mass equivalent sequences. For example, an

internal sequence flanked by glycine GXXXXG, where X could be any amino acid would create mass equivalent GXXXX and XXXXG sequences, increasing the probability for this mass relative to non-frame-shifting sequences. Double and higher order frame-shifting sequences can also be observed.

Cleaving the backbone in three locations yields the results shown Figure 1c. The trends are largely similar to those observed for double cleavage with the exception of greater bias against long fragments, either terminal or internal. Terminal fragments are still generated with greater frequency than internal fragments for most of the mass range, although the difference is less than the margin observed in Figure 1b, and the point at which long terminal fragments are observed with equal probability to internal fragments occurs at lower mass. Triple cleavage creates two internal fragments per precursor ion, which accounts for the narrowing between distributions. Quadruple cleavage is shown in Figure 1d. When the backbone is cleaved repeatedly, bias against longer terminal fragments becomes more extreme, with the longest fragments being generated orders of magnitude less often than shorter fragments. The gap between internal and terminal fragments also decreases relative to triple cleavage as does the crossover point where long terminal ions are equivalent to shorter internal fragments. It is expected that these same trends would continue for additional backbone fragmentation, with the gap between terminal and internal fragments decreasing incrementally with each additional cleavage.

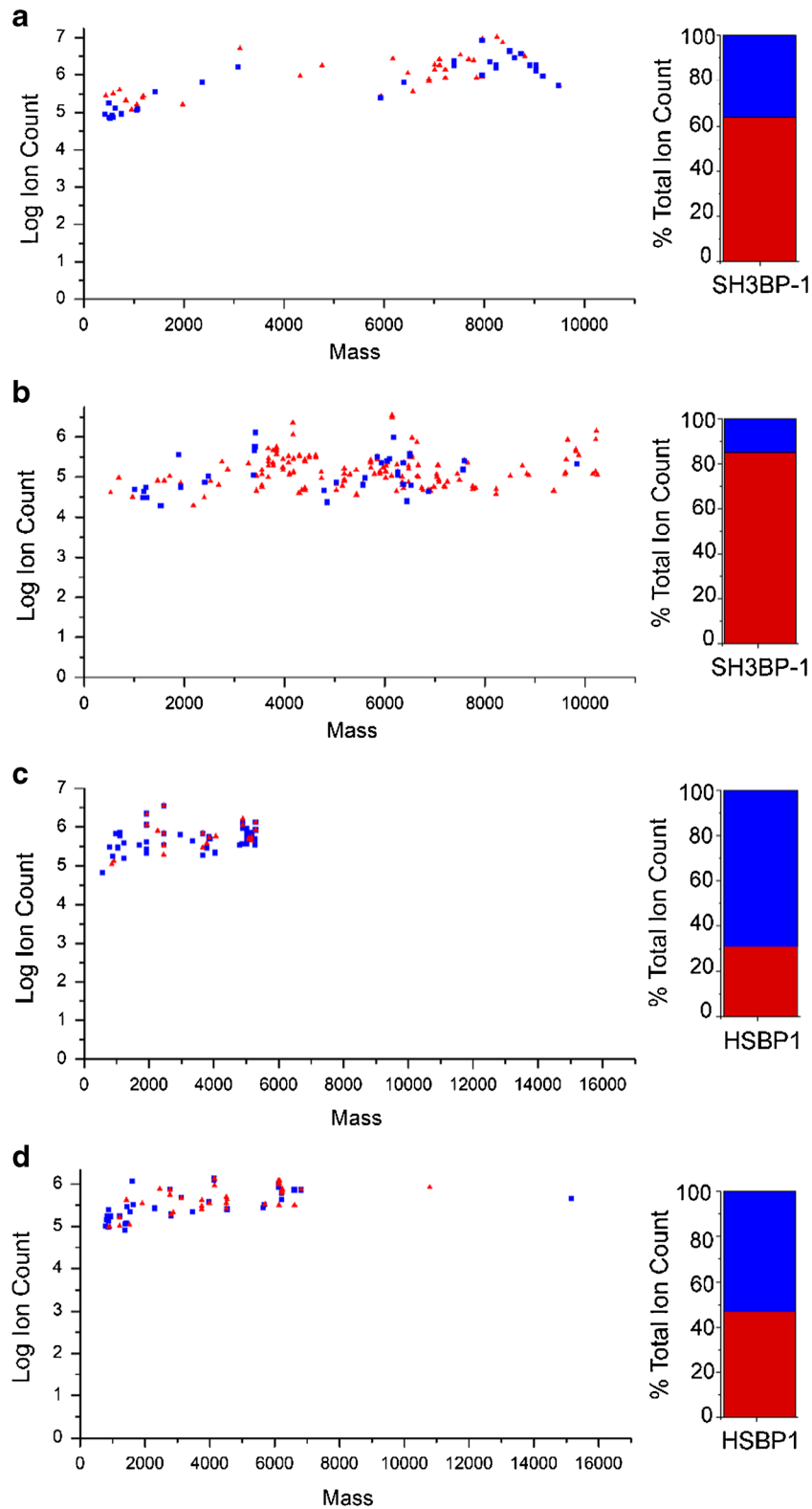
The results in Figure 1 allow examination of the statistics as a function of dissociation events, but true stochastic fragmentation would also lead to a random number of backbone cleavages. In other words, a fraction of molecules would be cleaved once, others twice, etc. Illustrative results for this type of process are shown for molecules of various sizes in Figure 2, where equivalent amounts of single, double, and triple cleavages were summed together. Results for bradykinin are shown in Figure 2a. For peptides of this length, the distinction between internal and terminal fragments is small. Short peptides, such as those routinely produced in proteolytic digests for proteomics experiments, have similar numbers of total internal and terminal fragments (16 terminal, 22 internal for bradykinin), mitigating the statistical dilution effect. In contrast, results for helicase illustrate clear differences between the propensity for creating internal and terminal fragments (see Figure 2b). Helicase is a ~32 kDa protein with 550 terminal and 34,356 internal unique mass fragments. The distinction between terminal and internal fragments becomes even larger for human serum albumin (HSA), as shown in Figure 2c. In fact, the ratio of the average abundance of terminal fragments to internal fragments of similar mass follows a linear trend as shown in Figure 2d. With larger molecular size, the probability for generating any particular internal fragment becomes less favorable because of the dilution effect.

In order to relate the information in Figures 1 and 2 to mass spectra, the charge must also be taken into consideration. If a



**Figure 2.** Linear sum of single, double, and triple statistical cleavages of (a) bradykinin, (b) helicase, and (c) human serum albumin. (d) Illustration of trend for ratio of internal/terminal ions as a function of protein size. Terminal fragments are displayed in orange, internal fragments, blue

fixed amount of charge is assigned to the total population of fragments according to mass, i.e., the amount of charge ending up on a fragment is equal to its proportion of mass, the amount of charge attributed to terminal and internal ions can be approximated. The results are summarized in Table 1 for



**Figure 3.** Fragment intensity versus mass for terminal (red triangles) and internal (blue squares) ions from SH3BP-1 **(a)** HCD, **(b)** UVPD and HSPB1, **(c)** HCD, and **(d)** UVPD. The percent of the total ion count for each type of fragment is shown in the bar graphs

double and triple cleavage of the peptide backbone for proteins of various sizes. Interestingly, the statistics predict that terminal ions account for the same fraction of total ion current regardless of protein size for a given number of

backbone fragmentations. For double cleavage, terminal ions account for  $\sim 2/3$  of the available charge. For triple cleavage, terminal ions retain  $\sim 1/2$  of the charge. These results may seem counterintuitive given that the number of



**Table 1.** Charge Distribution

Protein/#cleavages	Fraction charge terminal ions	Fraction charge internal ions	Average intensity internal ion	Number internal ions
Ubiquitin/double	0.667	0.333	137.9	2412
Gamma c/double	0.665	0.335	24.6	13267
Helicase/double	0.663	0.337	16.7	34359
Ubiquitin/triple	0.501	0.499	207.2	2409
Gamma c/triple	0.498	0.502	38.7	13264
Helicase/triple	0.495	0.505	16.7	34356

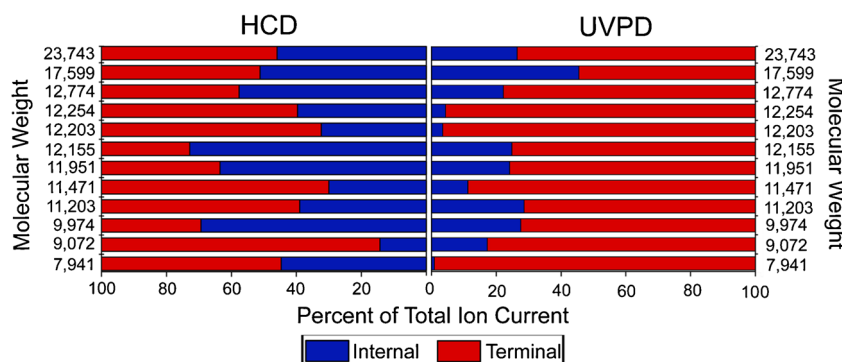
unique internal ions surges with increasing protein size. However, inspection of the results reveals that although the number of internal ions increases dramatically for larger proteins, the propensity for generating each internal ion decreases at essentially the same rate. These offsetting factors cause roughly the same amount of ion current to end up on internal ions, *regardless of protein size*.

These complex statistical results can also be easily predicted by simply considering the experiment from the molecular point of view. A single cleavage of the peptide backbone yields two terminal ions, and consequently 100% of the ion current goes to terminal ions. Cleavage of the backbone twice yields two terminal ions and one internal ion. When averaged out over all potential cleavage points, this will lead to 1/3 of the charge being apportioned to each fragment, leaving 2/3 of the charge on terminal ions. The number of potential internal ions is actually irrelevant in this single molecule picture. By extension, the charge will be split 50/50 after triple cleavage, which will yield two internal and two terminal ions. Importantly, these considerations imply that top-down analysis should not be overly hindered by internal ion production as protein size increases, as long as the number of multiple backbone fragmentations is not excessive. Fortunately, given that large molecules are typically difficult to fragment, excessive sequential fragmentation is not likely to be an issue.

### Comparison with Experiments

To gain additional insight, we analyzed representative results extracted from a high-throughput top-down experiment utilizing HCD and UVPD @213 nm. Detailed results are shown in Figure 3 for the +9 charge state of SH3

domain-binding glutamic acid-rich-like protein 3 (SH3BP), a ~10 kDa protein, and the +28 charge state of heat shock protein beta-1 (HSPB1), a ~23 kDa protein. Assignable fragments from the HCD spectrum for SH3BP are shown in Figure 3a using the same presentation scheme established in Figures 1 and 2. For experimental data, terminal and internal fragments do not separate into clearly distinct groupings as was observed in the statistical model. The primary reasons for this difference are: (1) not all fragments are detectable, and (2) dissociation is favored at some positions. More meaningful comparison between statistics and experiment can be extracted from the percentage of total ion current attributed to terminal and internal fragments, as shown in the bar graphs in Figure 3. For SH3BP, the majority of the ion current is attributed to terminal ions for both HCD and UVPD, although terminal ions are more dominant in UVPD. For the larger HSPB1 protein, internal ions account for more than half of the ion intensity in HCD and roughly half in UVPD. Based on the results in Table 1, to achieve the balance of terminal/internal ions observed in Figure 3c by statistical dissociation, the backbone would have to be fragmented five times on average (yielding two terminal and four internal ions, giving a 1:2 ratio). This degree of dissociation seems unlikely, suggesting that another factor influences the balance of ions detected. Indeed, inspection of the results reveals that many of the intense internal ions (e.g., residues 151-203, 154-203, 179-198, 2-42, 2-43, 3-42, 3-30) are generated by fragmentation near the termini. It is unlikely that the handful of residues comprising the terminal counterparts for these internal ions would be detected. HCD experiments are prone to favored



**Figure 4.** Total ion current distributions for 12 additional proteins sorted by mass

fragmentation pathways [20, 26, 27]. If dissociation is favored near a terminus, the balance of charge attributed to terminal ions will be negatively impacted. In contrast, analysis of the UVPD data reveals that only two backbone fragmentations are needed to account for completely statistical generation of the ion intensity taken up by internal ions in Figure 3d. Double cleavage of the backbone is more experimentally plausible, but it is also possible that some sequential dissociation may contribute to the overall intensity of internal ions in UVPD because recent reports have suggested direct dissociation of the backbone in UVPD is rare [28–30]. Inspection of the internal ions generated by UVPD does not reveal abundant dissociation near the termini. Overall, the results hint that backbone dissociation in UVPD is more statistical in nature relative to HCD.

The percentage of total ion current distributions for twelve additional proteins for HCD and UVPD are shown sorted by precursor mass in Figure 4. In general, UVPD leads to more ion current in terminal ions and less variability (min ~50%, max ~95%). In contrast, HCD yields more sporadic results, which may be connected to the locations of facile fragmentation points within any given protein sequence. However, for both UVPD and HCD, a significant portion of the total ion count is detected in terminal ions, and no obvious connection is noted between molecular weight and the distribution of total ion current.

## Conclusions

Statistical fragmentation of proteins significantly favors production of terminal ions over internal ions even if the protein backbone is fragmented repeatedly. This effect becomes more pronounced for larger proteins, with predicted abundances of terminal ions exceeding internals by orders of magnitude. Although the number of unique mass internal ions increases greatly with protein size, the statistical probability for observing any given internal ion decreases proportionately due to dilution. It is therefore incorrect to assume that the staggering number of internal ions that can be generated for large proteins will make detection of terminal ions unlikely. On the contrary, examination of the ion current in actual experiments reveals that terminal ions represent a significant fraction of the total, even for large proteins. The results further suggest that HCD is more prone for favored dissociation pathways while UVPD more closely reflects statistical dissociation.

## Acknowledgments

The authors thank Neil Kelleher for helpful discussions and Connor Julian for assistance with the coding. Funding from the National Science Foundation (CHE-1401737) and NIH (R01 GM107099) supported this research, which was carried out in collaboration with the National Resource for Translational and Developmental Proteomics under grant P41 GM108569 from the National Institute of General Medical Sciences, National Institutes of Health.

## References

- Catherman, A.D., Skinner, O.S., Kelleher, N.L.: Top Down proteomics: facts and perspectives. *Biochem. Biophys. Res. Commun.* **445**, 683–693 (2014)
- Toby, T.K., Fomelli, L., Kelleher, N.L.: Progress in top-down proteomics and the analysis of proteoforms. *Annu. Rev. Anal. Chem.* **9**, 499–519 (2016)
- Riley, N.M., Westphall, M.S., Coon, J.J.: Activated ion electron transfer dissociation enables comprehensive top-down protein fragmentation. *J. Proteome Res.* **16**, 2653–2659 (2017)
- Gregorich, Z.R., Cai, W., Lin, Z., Chen, A.J., Peng, Y., Kohmoto, T., Ge, Y.: Distinct sequences and post-translational modifications in cardiac atrial and ventricular myosin light chains revealed by top-down mass spectrometry. *J. Mol. Cell Cardiol.* **107**, 13–21 (2017)
- Cleland, T.P., Dehart, C.J., Fellers, R.T., Vannispén, A.J., Greer, J.B., LeDuc, R.D., Parker, W.R., Thomas, P.M., Kelleher, N.L., Brodbelt, J.S.: High-throughput analysis of intact human proteins using UVPD and HCD on an Orbitrap mass spectrometer. *J. Proteome Res.* **16**, 2072–2079 (2017)
- Cai, W., Tucholski, T., Chen, B., Alpert, A.J., McIlwain, S., Kohmoto, T., Jin, S., Ge, Y.: Top-down proteomics of large proteins up to 223 kDa enabled by serial size exclusion chromatography strategy. *Anal. Chem.* **89**, 5467–5475 (2017)
- Lee, A.E., Geis-Asteggiane, L., Dixon, E.K., Kim, Y., Kashyap, T.R., Wang, Y., Fushman, D., Fenselau, C.: Preparing to read the ubiquitin code: characterization of ubiquitin trimers by top-down mass spectrometry. *J. Mass Spectrom.* **51**, 315–321 (2016)
- Shen, Y., Tolić, N., Pichowski, P.D., Shukla, A.K., Kim, S., Zhao, R., Qu, Y., Robinson, E., Smith, R.D., Paša-Tolić, L.: High-resolution ultrahigh-pressure long column reversed-phase liquid chromatography for top-down proteomics. *J. Chromatogr. A.* **1498**, 99–110 (2017)
- Anderson, D.M.G., Van de Plas, R., Rose, K.L., Hill, S., Schey, K.L., Solga, A.C., Gutmann, D.H., Caprioli, R.M.: 3-D imaging mass spectrometry of protein distributions in mouse neurofibromatosis 1 (NF1)-associated optic glioma. *J. Proteom.* **149**, 77–84 (2016)
- Zhao, Y., Sun, L., Zhu, G., Dovichi, N.J.: Coupling capillary zone electrophoresis to a Q exactive HF mass spectrometer for top-down proteomics: 580 proteoform identifications from yeast. *J. Proteome Res.* **15**, 3679–3685 (2016)
- Song, Y., Laskay, Ü.A., Vilcins, I.M.E., Barbour, A.G., Wsocki, V.H.: Top-down-assisted bottom-up method for homologous protein sequencing: hemoglobin from 33 bird species. *J. Am. Soc. Mass Spectrom.* **26**, 1875–1884 (2015)
- Smith, L.M., Kelleher, N.L., Linial, M., Goodlett, D., Langridge-Smith, P., Ah Goo, Y., Safford, G., Bonilla, L., Kruppa, G., Zubarev, R., Rontree, J., Chamot-Rooke, J., Garavelli, J., Heck, A., Loo, J., Penque, D., Hornshaw, M., Hendrickson, C., Pasa-Tolić, L., Borchers, C., Chan, D., Young, N., Agar, J., Masselon, C., Gross, M., McLafferty, F., Tsybin, Y., Ge, Y., Sanders, I., Langridge, J., Whitelegge, J., Marshall, A.: Proteoform: a single term describing protein complexity. *Nat. Methods.* **10**, 186–187 (2013)
- Tran, J.C., Zamdborg, L., Ahlf, D.R., Lee, J.E., Catherman, A.D., Durbin, K.R., Tipton, J.D., Vellaichamy, A., Kellie, J.F., Li, M., Wu, C., Sweet, S.M.M., Early, B.P., Siuti, N., LeDuc, R.D., Compton, P.D., Thomas, P.M., Kelleher, N.L.: Mapping intact protein isoforms in discovery mode using top-down proteomics. *Nature.* **480**, 254–258 (2011)
- O'Brien, J.P., Li, W., Zhang, Y., Brodbelt, J.S.: Characterization of native protein complexes using ultraviolet photodissociation mass spectrometry. *J. Am. Chem. Soc.* **136**, 12920–12928 (2014)
- Skinner, O.S., McLafferty, F.W., Breuker, K.: How ubiquitin unfolds after transfer into the gas phase. *J. Am. Soc. Mass Spectrom.* **23**, 1011–1014 (2012)
- Tureček, F., Julian, R.R.: Peptide radicals and cation radicals in the gas phase. *Chem. Rev.* **113**, 6691–6733 (2013)
- Tsybin, Y.O., Fomelli, L., Stoerner, C., Luebeck, M., Parra, J., Nallet, S., Wurm, F.M., Hartmer, R.: Structural analysis of intact monoclonal antibodies by electron transfer dissociation mass spectrometry. *Anal. Chem.* **83**, 8919–8927 (2011)
- Schuhmann, K., Herzog, R., Schwudke, D., Metelmann-Strupat, W., Bornstein, S.R., Shevchenko, A.: Bottom-up shotgun lipidomics by higher energy collisional dissociation on LTQ orbitrap mass spectrometers. *Anal. Chem.* **83**, 5480–5487 (2011)

19. 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.* **135**, 12646–12651 (2013)
20. Marzluff, E.M., Beauchamp, J.L.: Collisional activation studies of large molecules. in large ions: their vaporization, detection, and structural analysis. Baer T, Ed. Wiley, New York (1996)
21. Ballard, K.D., Gaskell, S.J.: Sequential mass spectrometry applied to the study of the formation of “internal” fragment ions of protonated peptides. *Int. J. Mass Spectrom. Ion Processes.* **111**, 173–189 (1991)
22. Khatun, J., Ramkissoon, K., Giddings, M.C.: Fragmentation characteristics of collision-induced dissociation in MALDI TOF/TOF mass spectrometry. *Anal. Chem.* **79**, 3032–3040 (2007)
23. Durbin, K.R., Skinner, O.S., Fellers, R.T., Kelleher, N.L.: Analyzing internal fragmentation of electrosprayed ubiquitin ions during beam-type collisional dissociation. *J. Am. Soc. Mass Spectrom.* **26**, 782–787 (2015)
24. Fomelli, L., Durbin, K.R., Fellers, R.T., Early, B.P., Greer, J.B., LeDuc, R.D., Compton, P.D., Kelleher, N.L.: Advancing top-down analysis of the human proteome using a benchtop quadrupole-Orbitrap mass spectrometer. *J. Proteome Res.* **16**, 609–618 (2017)
25. Sham, D.P., Early, B.P., Fellers, R.T., Greer, J.B., Thomas, P.T., Fomelli, L., LeDuc, R.D., Shwab, D.J., Kelleher, N.L.: Accurate estimation of false discovery rates for protein and proteoform identification in top down proteomics. *Submitted*
26. Tsapralis, G., Nair, H., Somogyi, Á., Wysocki, V.H., Zhong, W., Futrell, J.H., Summerfield, S.G., Gaskell, S.J.: Influence of secondary structure on the fragmentation of protonated peptides. *J. Am. Chem. Soc.* **121**, 5142–5154 (1999)
27. Loo, J.A., Edmonds, C.G., Smith, R.D.: Tandem mass spectrometry of very large molecules. 2. Dissociation of multiply charged proline-containing proteins from electrospray ionization. *Anal. Chem.* **65**, 425–438 (1993)
28. Julian, R.R.: The mechanism behind top-down UVPD experiments: making sense of apparent contradictions. *J. Am. Soc. Mass Spectrom.* **28**, 1823–1826 (2017)
29. Hansen, K., Skinnerup Byskov, C., Nielsen, S.B.: Energy flow in peptides after UV photoexcitation of backbone linkages. *Phys. Chem., Chem. Phys.* **19**, 19640–19645 (2017)
30. Byskov, C.S., Jensen, F., Jørgensen, T.J.D., Nielsen, S.B.: On the photostability of peptides after selective photoexcitation of the backbone: prompt versus slow dissociation. *Phys. Chem., Chem. Phys.* **16**, 15831–15838 (2014)