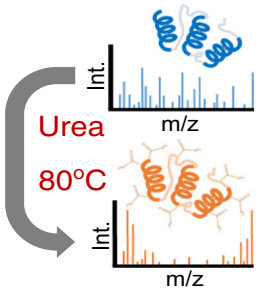
J. Am. Soc. Mass Spectrom. (2017) 28:1587–1599
DOI: 10.1007/s13361-017-1648-5

RESEARCH ARTICLE

Modulation of Protein Fragmentation Through Carbamylation of Primary Amines

Sylvester M. Greer,[1] Dustin D. Holden,[1] Ryan Fellers,[2] Neil L. Kelleher,[2,3] Jennifer S. Brodbelt[1]

[1]Department of Chemistry, University of Texas at Austin, Austin, TX 78712, USA
[2]National Resource for Translational and Developmental Proteomics, Northwestern University, Evanston, IL 60208, USA
[3]Departments of Chemistry, Molecular Biosciences, and the Feinberg School of Medicine, Northwestern University, Evanston, IL 60208, USA



Abstract. We evaluate the impact of carbamylation of the primary amines of the side-chains of Lys and the N-termini on the fragmentation of intact protein ions and the chromatographic properties of a mixture of *E. coli* ribosomal proteins. The fragmentation patterns of the six unmodified and carbamylated proteins obtained by higher energy collision dissociation (HCD) and ultraviolet photodissociation (UVPD) were compared. Carbamylation significantly reduced the total number of protons retained by the protein owing to the conversion of basic primary amines to non-basic carbamates. Carbamylation caused a significant negative impact on fragmentation of the protein by HCD (i.e., reduced sequence coverage and fewer diagnostic fragment ions) consistent with the mobile proton model, which correlates peptide fragmentation with charge distribution and the opportunity for charge-directed pathways. In addition, fragmentation was enhanced near the N- and C-termini upon HCD of carbamylated proteins. For LCMS/MS analysis of *E. coli* ribosomal proteins, the retention times increased by 16 min on average upon carbamylation, an outcome attributed to the increased hydrophobicity of the proteins after carbamylation. As noted for both the six model proteins and the ribosomal proteins, carbamylation had relatively little impact on the distribution or types of fragment ions product by UVPD, supporting the proposition that the mechanism of UVPD for intact proteins does not reflect the mobile proton model.

Keywords: Protein, Top down, MS/MS, Ultraviolet photodissociation, Carbamylation

Received: 30 January 2017/Revised: 28 February 2017/Accepted: 2 March 2017/Published Online: 3 April 2017

Introduction

Improved chromatographic methods coupled with high performance mass analyzers and increasingly sophisticated informatics have facilitated the efficient separation, analysis, and identification of intact proteins in the gas phase, thus inspiring great interest in top-down strategies for proteomics [1–3]. While measurement of the accurate mass of a protein is a crucial first step, complete characterization of a proteoform (i.e., a unique molecular form of a protein including its mutations and specific post-translational modifications) requires much more information about the sequence, as well as the identity, number, and position of modifications [4]. There are several established methods to activate and dissociate intact proteins; collisionally activated dissociation (CAD [5], beam-type higher energy collision dissociation (HCD [6, 7]), and electron-based methods (most commonly electron transfer dissociation (ETD [8, 9]), have been used for the most significant high throughput top-down studies. Ultraviolet photodissociation (UVPD) is the newest activation method that has been developed for the analysis of intact proteins [10–16]. The absorption of high energy photons (typically 6.4 eV per 193 nm photon) results in extensive backbone cleavages that result in formation of *a*, *b*, *c*, *x*, *y*, and *z* ions. UVPD affords high sequence coverage, the ability to map sites of post-translational modifications, and has shown promise for top-down LC-MS applications [16].

Electronic supplementary material The online version of this article (doi:10.1007/s13361-017-1648-5) contains supplementary material, which is available to authorized users.

Correspondence to: Jennifer Brodbelt; *e-mail:* jbrodbelt@cm.utexas.edu

Top-down methods have not reached the widespread adoption of bottom-up methods for high throughput proteomics, in part owing to less effective activation methods for intact proteins [1]. In the context of activation of proteins, performance metrics tend to decrease with increasing mass, and charge state plays a major role [17, 18]. Upon electrospray ionization, protein ions are generated in a wide array of charge states, thus making it especially important to more extensively evaluate and understand the impact of charge state on the fragmentation of proteins [5, 19]. There have been several systematic studies of protein dissociation using collisional activation, including ones that have examined the influence of charge state and other factors on fragmentation pathways [19–23]. CAD of intact proteins depends on proton mobility for fragmentation, as also well-recognized for peptide fragmentation induced by collisional activation [24]. Protons are typically sequestered at the more basic sites (Arg, Lys, His, N-terminus), but these protons can be mobilized via addition of energy to the ion [25]. For proteins in higher charge states, the additional protons associated with less basic sites along the backbone facilitate *b*/*y* fragmentation pathways [24, 26, 27]. As similarly noted for peptides, cleavages are preferentially enhanced at acidic residues for lower charge states (i.e., absence of mobile protons). McLuckey and co-workers have shown that CAD of intact proteins in low charge states (i.e., ones typified by low proton mobility) results in enhanced cleavage at Glu and Asp residues and a reduction in other diagnostic backbone *b*/*y* fragments compared with fragmentation of higher charge states having a greater number of mobile protons [20]. In another CAD study of intact proteins, Agar et al. reported enhanced cleavages adjacent to Gly, Lys, Gln, and N-terminal to Ser, Tyr, Ile, Leu, and Pro, none of which are prominent in typical CAD spectra of tryptic peptides [23].The irregular distribution of basic residues in proteins may give rise to these uncommon cleavage pathways. Not surprisingly, CAD spectra of tryptic peptides with missed cleavages (i.e., peptides having a basic Arg residue other than at the C-terminus) often display these same dissociation pathways [23].

The production of diagnostic *c*/*z* sequence ions upon ETD of intact proteins is also highly dependent on the charge density and charge state. ETD of proteins in low charge states result in far fewer fragments and reduced sequence coverage compared with ones in higher charge states owing to the propensity for non-dissociative charge reduction associated with electron attachment [28, 29]. The ability of ETD to map post-translational modifications remains a particularly compelling advantage, which balances the sub-par performance for proteins in lower charge states [30, 31]. In contrast to collisional and electron-based activation, UVPD has shown less dependence on charge state, and many of the fragmentation processes do not require mobile protons [32].

The number and locations of charge sites can be further modulated by addition of supercharging agents to the solutions or by derivatization to convert specific functional groups to more or less basic ones. Iavarone and colleagues studied the impact of supercharging on fragmentation of intact proteins. Supercharging was achieved via addition of *m*-nitrobenzyl alcohol to the solution prior to ESI [33]. Collisionally activated fragmentation of the supercharged states yielded a small number of highly abundant fragment ions clustered around narrow stretches of the protein backbone compared with the more widespread fragmentation observed upon CAD of intermediate charge states [34]. Recently, new supercharging reagents have been discovered that facilitate production of higher charge states even beyond those observed with traditional supercharging reagents [35]. As shown for supercharged ubiquitin, Zenaidee and Donald reported greater than 90% sequence coverage by ECD utilizing recently discovered supercharging reagents [36]. The Smith group manipulated the charge states of intact proteins via chemical derivatization of acidic sites and addition of basic moieties or ones with fixed charges [37]. Capping acidic sites with neutral moieties shifted the charge states very little during ESI, suggesting that carboxylic acid side-chains played a relatively minor role in determining the charge states of proteins upon ESI. Addition of basic and fixed charge moieties had a significant impact on the charge state of intact denatured proteins, suggesting that the number of basic sites modulated the range of charge states adopted by the proteins [37]. Guanidination increases the basicity of Lys residues and promotes proton sequestration (reducing proton mobility), and was used as a means to probe the influence of proton mobility on CAD of ubiquitin [38]. For the 10+ charge state of ubiquitin, the resulting fragmentation of the guanidinated protein occurred largely C-terminal to Asp in a charge remote fashion as predicted by the mobile proton theory. For the 10+ charge state of non-guanidinated ubiquitin, nonspecific amide bond cleavages to produce traditional *b*/*y* ions and enhanced cleavage N-terminal to Pro were observed. This contrast in fragmentation behavior that arose from guanidination demonstrated that reduction in proton mobility restricted the nonspecific fragmentation pathways [38].

In the present study, we directly evaluate the dependence of UVPD and HCD on proton mobility and charge state for several proteins. In order to affect both protein charge state and proton mobility, we employ a highly efficient carbamylation reaction that converts the basic primary amines of Lys side-chains and the N-terminus to less basic carbamates. Not only does this reduce the average charge state adopted by a given protein upon electrospray ionization, it removes sites of proton sequestration, which alters proton mobility. The impact of charge state and proton mobility on HCD and UVPD of six proteins and a mixture of ribosomal proteins were investigated. Comparisons of sequence coverage, distributions of sequence ions, and fragment ion type are reported for multiple charge states of unmodified and carbamylated proteins.

Experimental

Materials

Ubiquitin (bovine), cytochrome *c* (equine), myoglobin (bovine), superoxide dismutase (bovine), lysozyme (galline),

carbonic anhydrase (bovine), and urea were obtained from Sigma-Aldrich (St. Louis, MO, USA). *E. coli* 70S ribosome was obtained from New England Biolabs (Ipswich, MA, USA). Solvents were obtained from EMD Millipore (Billerica, MA, USA). Tris(2-carboxyethyl) phosphine (TCEP) was obtained from Thermo-Scientific (Rochford, IL).

Carbamylation

Carbamylation was performed as previously reported [39]. Briefly, each sample was split into two aliquots; one for derivatization and one as a control. Each was suspended in 200 mM Tris-HCl in the presence or absence of 8 M urea. Both samples were incubated at 80 °C for 4 h. Samples were desalted using Amicon Ultra 3 kDa MWCO spin columns (EMD Millipore, Billerica, Ma, USA), then evaporated to dryness and resuspended in solvent to match the LC starting conditions (2% acetonitrile/98% water/0.1% formic acid) or infusion conditions (50% methanol/50% water/1% formic acid).

Separation

Proteins were separated using a Dionex RSLC 3000 nano-liquid chromatograph (Thermo Fisher, San Jose, CA, USA). Approximately 1 μg of proteins were injected onto a 3 cm PLRP reverse phase trapping column (75 μm i.d.) packed with 5 μm particles (1000 Å pore size). Proteins were then eluted onto a 40 cm fritted 75 μm i.d. pulled tip analytical column (New Objective, Woburn, MA, USA) packed in-house with PLRP (5 μm particles, 1000 Å pore size) at a flow rate of 300 nL/min using a linear gradient of 2%–50% solvent B (acetonitrile/0.1% formic acid) over 120 min. Solvent A was water/0.1% formic acid.

Mass Spectrometry

Proteins for infusion were suspended in a solution of water, acetonitrile, and formic acid (49.5/49.5/1) at a final concentration of 10 μM. For proteins having known disulfide bonds, a 20X molar excess of TCEP was added to the solution prior to infusion. The proteins were either infused directly at 3 μL/min using a HESI II Source (Thermo Fisher Scientific, San Jose, CA, USA) or introduced by nano LC ESI into an Orbitrap Fusion Lumos mass spectrometer (Thermo Fisher Scientific, San Jose, CA, USA) customized for implementation of UVPD as described previously [40].

Spectra were analyzed in the Orbitrap mass analyzer at a resolving power of 120,000 at *m/z* 200, using Intact Protein Mode. Two hundred fifty scans were collected and averaged for infusion experiments. LC-MS data was collected in a top speed (7 s cycle) data-dependent manner where each MS1 consisted of 4 μs cans (AGC target of 1.0E + 05, max injection time of 100 ms), and MS2 consisted of 6 μs cans (AGC target of 5.5E + 05, max injection time of 250 ms). Precursor ions were filtered according to intact protein monoisotopic precursor selection, thus focusing on proteins with charge states greater than 5+. Spray voltage was set to 1.8 kV. MS2 isolation width was set to 5 *m/z* using the quadrupole for mass filtering. Precursors selected more than five times in 120 s were excluded from MS2 selection for 120 s. HCD normalized collision energy (NCE) was optimized (10–30 NCE) per charge state for infusion experiments and set to 20 NCE for LC-MS experiments. UVPD performed in the high pressure cell of the dual linear ion trap was achieved via a single 5 ns laser pulse from a Coherent ExciStar XS 500 (Santa Clara, CA, USA) 193 nm excimer laser. Laser power was set to 1.0 mJ for both infusion and LC-MS experiments.

Data Analysis

High resolution intact protein fragmentation spectra were deconvoluted using the Xtract algorithm enabled in Thermo Xcalibur Qual Browser (Thermo Fisher Scientific, San Jose, CA, USA). The deconvoluted data was further processed via Prosight Lite build 1.3.5744.1622 (http://prosightlite.northwestern.edu/) to generate sequence coverage maps and confirm the degree of carbamylation based on the presence of fragment ions within a 10 ppm tolerance [41]. Only fragment ions that contain the N-terminal or C-terminal residue of the sequence are searched and identified; internal ions are not identified. Protein backbone cleavage maps were generated using msProduct (http://prospector.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msproduct) outputs to assign cleavage position, fragment intensity, and ion type. These results were further processed and represented graphically using Microsoft Excel (Microsoft, Redmond, WA, USA)

LC-MS data was processed using ProSight PC 4.0. The protein sequence database was populated using the Uniprot reviewed *E. coli* K12 database (accessed October 2016). To enable analysis of the MS/MS spectra of carbamylated proteins, a custom sequence database was created, which treated all Lys residues as carbamylated and N-termini as carbamylated or acetylated. The results were filtered at the proteoform level using a P-Score cutoff of 1.0E-05.

Results and Discussion

This study focuses on evaluating the impact of modifying charge sites (Lys, N-termini) of intact proteins on the outcome and metrics of HCD and UVPD. Six model proteins and a ribosomal protein mixture were introduced using infusion or via nanoLC, respectively, then subjected to HCD and UVPD. In particular, this work aims to compare the fragmentation of carbamylated and unmodified proteins to evaluate the influence of mobile protons and charge state upon HCD and UVPD. Carbamylation of the primary amines of Lys and the N-terminal amine converts them to non-basic groups (Supplementary Scheme S1), leaving Arg residues as the most basic sites, followed by His. Carbamylation of proteins changes not only the number of protons retained by each protein during ESI but also the localization and mobility of protons, as evidenced by the sometimes significant variations in fragmentation observed in the MS/MS spectra generated upon HCD, as described in more detail below. Six model proteins were selected to span an array of molecular sizes and have a range of number, locations, and distributions of basic Lys/Arg sites. For

example, lysozyme, cytochrome *c*, and myoglobin have similar total numbers of highly basic sites (20, 18, and 22, respectively), but the number of Arg residues varies considerably (only two for cytochrome *c* and myoglobin, but 11 for lysozyme). Since fragmentation near the termini is dominant upon collisional activation of intact proteins, proteins were selected with Lys near the N-terminus (ubiquitin, cytochrome *c*, lysozyme, superoxide dismutase) or near the C-terminus (cytochrome *c*, superoxide dismutase, carbonic anhydrase) or with Arg in those segments. A list of the six protein sequences is provided in Supplementary Figure S1.

Effect of Carbamylation on Charge State Distribution of Intact Proteins

Prior to evaluating the variations in fragmentation patterns upon carbamylation, first the distributions of charge states of unmodified versus carbamylated proteins were examined. The charge states adopted by a protein upon ESI are influenced by several factors, including protein size, number of basic and acidic residues, solvent composition and solvent additives, among others. In this study, proteins were sprayed using conventional denaturing conditions prior to or after carbamylation of all Lys and N-termini. For example, ubiquitin (8.5 kDa) contains seven Lys residues and was modified a total of eight times upon carbamylation with nearly 100% efficiency, indicating highly efficient carbamylation of all seven primary amines of the Lys side-chains plus the N-terminus (Figure 1). Figure 1 shows the ESI mass spectra of unmodified and carbamylated ubiquitin, myoglobin, and carbonic anhydrase (and the same pairs of ESI mass spectra are shown in Supplementary Figure S2 for cytochrome *c*, lysozyme, and superoxide dismutase). For each of the six proteins, the shift in the charge state distributions after carbamylation is dramatic. The range of charge states for unmodified ubiquitin is +6 to +13, whereas it is +4 to +9 after carbamylation. Myoglobin (16.9 kDa) has 19 Lys and displays charge states from +10 to +24 prior to carbamylation and +7 to +16 after carbamylation. Carbonic anhydrase (29 kDa, 18 Lys) also shows a significant shift in charge state distribution upon carbamylation, ranging from +14 to +35 prior to carbamylation and from +10 to +28 after carbamylation. The same shift in charge states also occurs for cytochrome *c*, lysozyme, and superoxide dismutase (Supplementary Figure S2). The significantly lower basicity of carbamylated groups in comparison to primary amines accounts for the notable reduction in charge states of the proteins.

For the subsequent MS/MS experiments described in the next sections, several charge states were selected for HCD and UVPD. Typically one charge state higher than the median charge state and one charge state lower than the median were selected for MS/MS analysis, as well as one charge state that "overlapped" between each unmodified and carbamylated protein. Owing to the incredibly rich MS/MS spectra of intact proteins [as exemplified by the HCD and UVPD spectra of lysozyme (12+) shown in Supplementary Figure S3], often containing more than 100 fragment ions, displaying numerous annotated spectra is cumbersome and thus an alternative graphical representation was used for this study. In order to show the

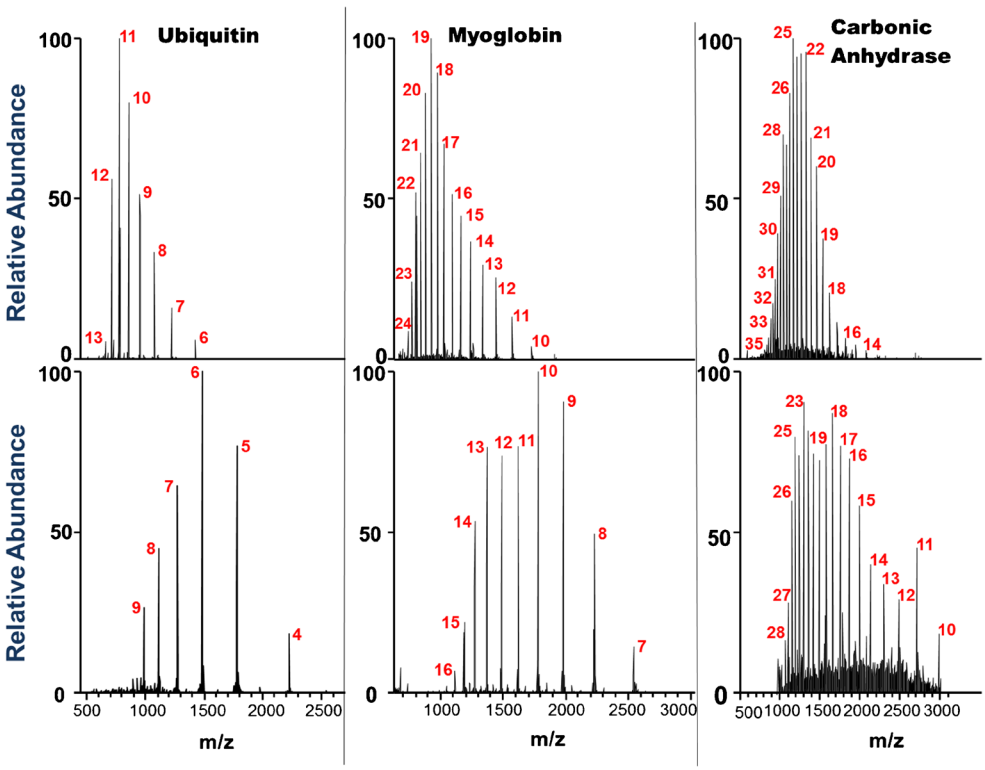


Figure 1. (**a**), (**b**), (**c**) ESI-MS of unmodified proteins; (**d**), (**e**), (**f**) ESI MS of the analogous carbamylated proteins

distribution of cleavages of the backbone, the relative fragment ion abundances originating from cleavages at each backbone position were plotted as histograms spanning the protein sequence. The abundances of all N-terminal (*a*, *b*, *c* ions) and C-terminal (*x*, *y*, *z* ions) corresponding to each inter-residue position were summed, and the two sums were stacked and placed at their appropriate inter-residue cleavage site along the protein backbone. This data format is shown in Figure 2 for lysozyme and Figure 3 for myoglobin, as well as in Supplementary Figures S4–S27 for all six proteins. In addition, sequence maps for several selected charge states of each protein (unmodified and carbamylated) are displayed in Supplementary Figures S28–33 based on HCD and UVPD data.

Ubiquitin

Collisional activation of the 10+ and 12+ charge states of ubiquitin (Supplementary Figure S4) resulted in remarkably similar fragmentation patterns, exhibiting significant cleavage C-terminal to three Glu residues (Glu16, Glu18, Glu18), as well as N-terminal to Pro (Pro19). This pattern shows that fragmentation of ubiquitin is dominated by preferential pathways (e.g., adjacent to acidic residues and Pro); similar behavior has been reported previously [15, 20, 23]. There were also many nonspecific cleavages across much of the backbone, yielding numerous low abundance *b* and *y* ions and resulting in similar total sequence coverages for both charge states (84% for 10+ and 89% for 12+). Collisional activation of the 10+ charge state of carbamylated ubiquitin (Supplementary Figure S5) resulted in a similar preference towards Pro-mediated cleavage; however, the dominant cleavage occurred N-terminal to a different Pro residue: Pro37 (IleI36/Pro37) instead of Pro19 and HCD resulted in somewhat lower sequence coverage (72%) compared with the 10+ charge state of the unmodified protein.

HCD of one representative low charge state (6+) of unmodified ubiquitin displayed extensive cleavage across the entire backbone (Supplementary Figure S4), along with enhanced cleavage N-terminal to Pro19 and N-terminal to Pro37 and a number of enhanced cleavages C-terminal to acidic residues (Asp32, Glu34, Asp39, Asp52, and Asp57). HCD of the corresponding 6+ charge state of carbamylated ubiquitin showed prominent fragmentation channels N-terminal to Pro19 and Pro37 as well as an array of nonspecific backbone cleavages spanning residues Ile3 to Glu18 (Supplementary Figure S5). The fragmentation pattern of carbamylated ubiquitin (6+) upon HCD most closely resembled the fragmentation of the 10+ charge state of unmodified ubiquitin. With respect to this similarity in fragmentation behavior, it appears that the lower charge state of carbamylated ubiquitin (6+) was in part compensated by the greater mobility of protons upon carbamylation of the Lys side-chains. Interestingly, upon HCD the 4+ charge

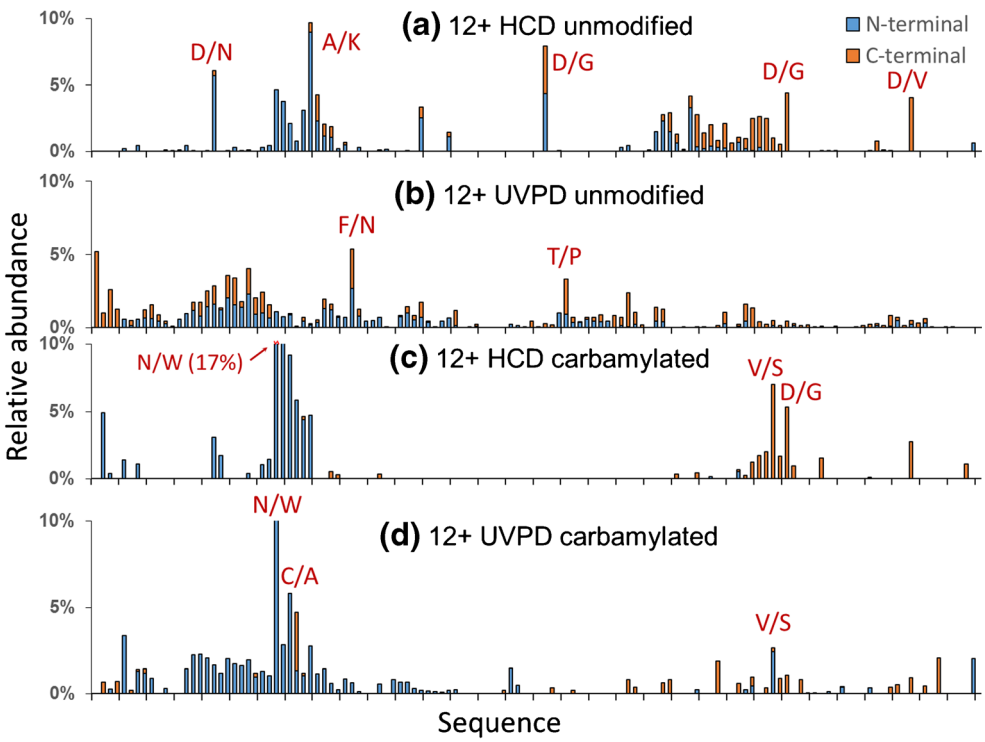


Figure 2. Backbone cleavage histograms of lysozyme (12+): (**a**) HCD of unmodified lysozyme, (**b**) UVPD of unmodified lysozyme, (**c**) HCD of carbamylated lysozyme, and (**d**) UVPD of carbamylated lysozyme. Some of the most enhanced backbone cleavage sites are labeled. The backbone was separated into four-residue stretches marked by ticks along the x-axis, going from the N-terminus (left) to the C-terminus (right). All N-terminal sequence ions are shown as blue bars; all C-terminal sequence ions are shown as orange bars

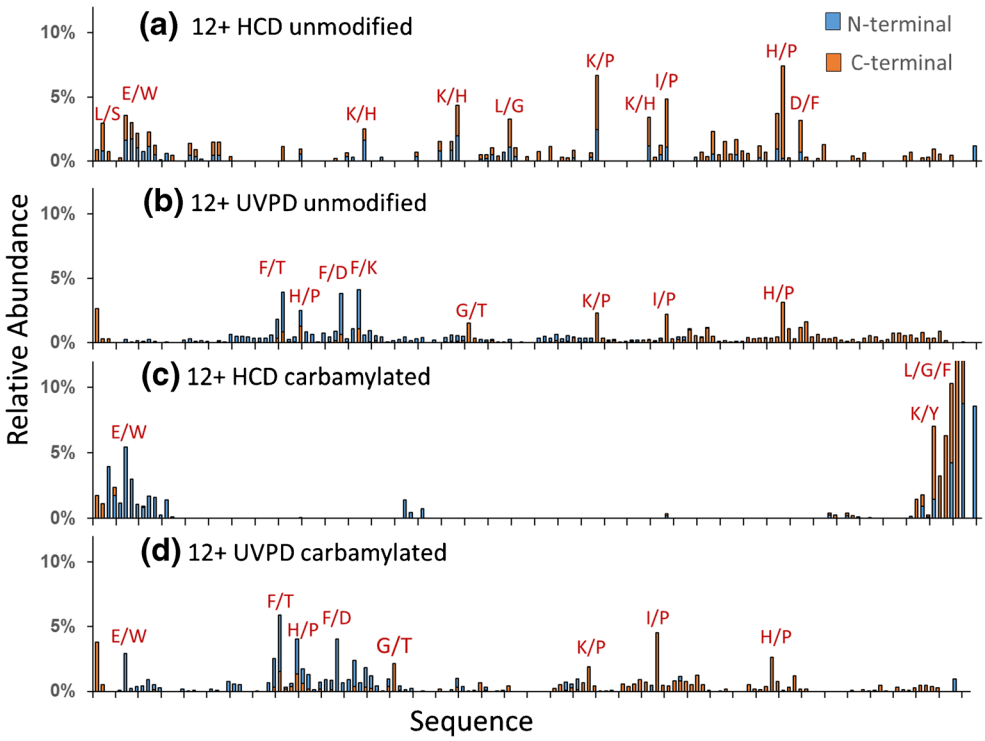


Figure 3. Backbone cleavage histograms of myoglobin (12+): (**a**) HCD of unmodified lysozyme, (**b**) UVPD of unmodified lysozyme, (**c**) HCD of carbamylated lysozyme, and (**d**) UVPD of carbamylated lysozyme. Some of the most enhanced backbone cleavage sites are labeled. The backbone was separated into four-residue stretches marked by ticks along the x-axis, going from the N-terminus (left) to the C-terminus (right). All N-terminal sequence ions are shown as blue bars; all C-terminal sequence ions are shown as orange bars

state of carbamylated ubiquitin (a charge state only accessible after carbamylation) did not exhibit the preferential Pro cleavage observed for all of the other charge states, and instead, cleavages C-terminal to acidic residues were exceptionally prominent (Asp21, Asp32, Asp39, Asp52, Asp58, Glu64). The exaggerated enhancement of cleavages C-terminal to acidic residues has been noted previously for low charge states of unmodified proteins [15, 42–44], and it is echoed for the very low charge state (4+) of carbamylated ubiquitin in the present study. Although this result may seem counterintuitive, the number of charges (4) of carbamylated ubiquitin is equal to the number of Arg residues (4 Arg), and thus the four charges could be sequestered at Arg sites, resulting in the persistence of charge remote fragmentation despite the carbamylation of Lys residues and the N-terminus. In general, the fragmentation patterns of each of the three representative charge states of carbamylated ubiquitin (4+, 6+, 10+) displayed significant differences in the locations and sites of preferential cleavages, and the total sequence coverages (56%–72%) were notably lower than the coverages obtained for the unmodified protein (83%–89%) upon HCD (see Table 1).

Unlike the variations in fragmentation patterns observed upon HCD of unmodified ubiquitin, the fragmentation patterns generated upon UVPD are nearly independent of charge state and only modest differences are noted among the backbone cleavage histograms of the 6+, 10+, and 12+ charge states (Supplementary Figure S6). Closer inspection of the histograms reveals subtle variations in the relative portions of C- and N-terminal product ions, but overall the fragmentation is considerably more uniform across the backbone upon UVPD compared with HCD. This trend is also reflected in the consistently high sequence coverage obtained from UVPD of ubiquitin irrespective of charge state (99%–100%, Table 1), an outcome that also holds true for carbamylated ubiquitin (87%–99%). For all charge states, UVPD of carbamylated ubiquitin showed suppressed backbone fragmentation in the stretch from Pro19 to Ile36 compared with the unmodified protein, along with significantly lower abundances of N-terminal fragment ions across the entire sequence (Supplementary Figure S7). However, in contrast to HCD of carbamylated ubiquitin, which favored preferential cleavages at Pro and acidic residues, UVPD of carbamylated ubiquitin largely exhibited nonspecific backbone fragmentation akin to the pattern observed upon UVPD of unmodified ubiquitin.

Cytochrome c

Upon ESI, cytochrome *c* retains the heme group (+616.191 Da) bound at C14 and C17, and this heme group is incorporated in assignment of all fragment ions that encompass those Cys residues. HCD of the 12+, 14+, and 16+ charge states of cytochrome *c* (Supplementary Figure S8) resulted in a dominant cleavage N-terminal to Pro (Pro77). Backbone cleavages adjacent to two Lys residues (K26 and K28) were also

Table 1. Sequence Coverages Obtained for Various Charge States of Unmodified and Carbamylated Proteins by HCD and UVPD

Ubiquitin	4+	6+	10+	12+		
Unmod. HCD		83%	84%	89%		
Carb. HCD	64%	56%	72%			
Unmod. UVPD		99%	100%	100%		
Carb. UVPD	87%	97%	99%			
Cytochrome *c*	**6+**	**8+**	**10+**	**12+**	**14+**	**16+**
Unmod. HCD				88%	83%	72%
Carb. HCD	63%	70%	50%			
Unmod. UVPD				89%	90%	87%
Carb. UVPD	84%	90%	91%			
Lysozyme	**8+**	**10+**	**12+**	**15+**	**18+**	
Unmod. HCD			52%	55%	54%	
Carb. HCD	28%	28%	27%			
Unmod. UVPD			85%	85%	85%	
Carb. UVPD	59%	67%	59%			
SOD	**12+**	**14+**	**16+**	**20+**		
Unmod. HCD	55%		54%	50%		
Carb. HCD	35%	39%	43%			
Unmod. UVPD	73%		77%	71%		
Carb. UVPD	77%	78%	75%			
Myoglobin	**8+**	**10+**	**12+**	**16+**	**20+**	
Unmod. HCD			52%	64%	45%	
Carb. HCD	18%	21%	24%			
Unmod. UVPD			90%	94%	91%	
Carb. UVPD	74%	83%	74%			
Carbonic anhydrase	**16+**	**22+**	**29+**	**34+**		
Unmod. HCD		43%	38%	26%		
Carb. HCD	24%	17%	19%			
Unmod. UVPD		65%	60%	66%		
Carb. UVPD	28%	34%	34%			

prominent, and HCD promoted a wide variety of nonspecific backbone cleavages to produce ample series of *b*/*y* ions. For each of these charge states, HCD resulted in high sequence coverage (72%–88%, Table 1) as evidenced by relatively broad fragmentation along the backbone, aside from a few stretches of little or no fragmentation (Cys14 to Thr19, Phe36 to Gln42). HCD of carbamylated cytochrome *c* (Supplementary Figure S9) also resulted in enhanced backbone cleavage N-terminal to Pro77; however, aside from consistent fragment ions near the C- and N-termini, the MS/MS spectra obtained for the 6+, 8+, and 10+ charge states exhibited large stretches lacking any fragmentation, and the overall sequence coverages were significantly lower for carbamylated cytochrome *c* (50%–70%).

In contrast to the fragmentation patterns observed upon HCD of cytochrome *c*, UVPD (Supplementary Figure S10) does not result in a dominant cleavage N-terminal to Pro77, and rather backbone cleavage N-terminal to Pro30 and Pro45 were more prominent pathways. Moreover, UVPD resulted in nonspecific fragmentation across nearly the entire backbone that yielded high sequence coverages for all charge states (87%–90%). The heme-binding domain covering the stretch from Cys14 though Cys17 remains resistant to fragmentation by both HCD and UVPD, confirming the stabilization of this region owing to the thioether bonds between the heme and two Cys residues.

UVPD of carbamylated cytochrome *c* resulted in extensive nonspecific fragmentation across the backbone (Supplementary Figure S11). For the lowest charge state (6+) carbamylated cytochrome *c* exhibited preferential cleavage N-terminal to Pro77; however, as charge state increased (8+ and 10+) this cleavage became less prominent and the cleavages N-terminal to Pro30 and Pro45 were enhanced. As noted earlier for the ubiquitin analysis, carbamylation had a much lower impact on the fragmentation of cytochrome *c*, regardless of charge state, compared with the far more striking impact on HCD in which backbone fragmentation throughout the protein was significantly curbed. The sequence coverage afforded by UVPD was 84%–91% for the various charge states of carbamylated cytochrome *c* (Table 1), nearly identical to the range obtained for the unmodified protein and well above that obtained upon HCD.

Lysozyme

Lysozyme, which naturally has four disulfide bonds, was reduced prior to MS/MS analysis to mitigate the well-known suppression of fragmentation caused by disulfide bonds in proteins. Upon ESI, lysozyme produced ions in charge states that spanned 9+ to 20+ (Supplementary Figure S2). HCD of lysozyme in the 12+, 15+, and 18+ charge states resulted in approximately 50% sequence coverage (Table 1), with significant gaps at the N- and C-termini as well as the mid-section of the protein. For the two lower charge states (12+ and 15+), some C-terminal and N-terminal ions were observed across the entire backbone, whereas only N-terminal *b* ions were observed exclusively for the first half of the sequence and only C-terminal *y* ions were observed for the last half of the

sequence (Figure 2 and Supplementary Figure S12). This trend for selective formation of N-terminal and C-terminal product ions was further enhanced for the 18+ charge state, and fragmentation was preferentially clustered around a few residues (e.g., Asn27/Trp28, Ile88/Thr89).

Lysozyme has six Lys residues and therefore was carbamylated at seven positions (N-terminus, Lys1, Lys13, Lys33, Lys96, Lys97, Lys116.) After carbamylation, the charge states ranged from 7+ to 14+. HCD of carbamylated lysozyme (Figure 2 and Supplementary Figure S13) resulted in markedly sparser fragmentation dominated by cleavages consolidated in the stretch from Asn26 to Phe33 (NWVCAAKF) and Lys96 to Asn103 (KKIVSDGN). These two regions are the longest stretches of the protein sequence that contain no basic Arg or His residues and only contain carbamylated Lys residues. The limited fragmentation observed for the carbamylated protein resulted in low sequence coverages (<30% Table 1) for all three charge states. Similar to the behavior observed upon HCD of the higher charge states of the unmodified protein, predominantly N-terminal *b* type ions were produced for the N-terminal half of the protein, and C-terminal *y* ions for the C-terminal half of the protein. Four of the most dominant products generated upon HCD of unmodified lysozyme arose from cleavage C-terminal to Asp. Only one of these cleavages remained prominent after carbamylation.

As illustrated in Figure 2 and Supplementary Figure S14, UVPD of lysozyme (12+, 15+, and 18+ charge states) yielded nonspecific cleavages across nearly the protein backbone independent of charge state and resulted in far greater sequence coverages (averaging 85%, Table 1) than observed for HCD (averaging 54%). Interestingly, the number of complementary pairs of N-terminal and C-terminal fragment ions decreased as charge state increased, and in the 18+ charge state the majority of C-terminal fragment ions were restricted to the C-terminal half of the protein just as the majority of N-terminal fragment ions were restricted to the N-terminal half of the protein. This result was similar to the segregation of C-terminal and N-terminal fragment ions noted above for HCD of lysozyme, albeit with a far greater total number of backbone cleavage sites observed upon UVPD.

The sequence coverage obtained upon UVPD of carbamylated lysozyme (Figure 2 and Supplementary S15 for 8+, 10+, 12+) decreased (averaging 62%) compared with the coverage obtained upon UVPD of unmodified lysozyme (averaging 85%), just like was noted upon HCD. However, there was still significantly greater fragmentation throughout the protein upon UVPD than observed upon HCD. The prominent fragmentation across the Asn27-Phe34 stretch was also observed upon UVPD, like HCD. The UVPD fragmentation trends showed less dependence on charge state for both the unmodified and carbamylated protein compared with HCD, again attesting to the reduced impact of mobile protons on modulation of fragmentation.

Superoxide Dismutase

Upon ESI, superoxide dismutase (SOD) produced ions in charge states ranging from 10+ to 23+ after reduction of disulfide bonds. HCD of SOD in the 12+, 16+, and 20+ charge states resulted in nonspecific cleavage (Supplementary Figure S16) of the backbone to yield approximately 50% sequence coverage (Table 1). The most extensive fragmentation was concentrated in the regions spanning Val27 to Gly35, and Gly91 to Tyr108. SOD exhibited C-terminal and N-terminal product ion segregation that increased with charge state such that only N-terminal *b* ions are dominant for the first half of the protein, and C-terminal *y* ions are dominant for the second half of the protein.

SOD underwent efficient carbamylation at all 10 Lys residues (the acetylated N-terminus was not modified) and the dominant charge state shifted from 20+ for the unmodified protein to 14+ after carbamylation. Similar to the trend noted for the other proteins, the sequence coverage obtained upon HCD of carbamylated SOD decreased significantly relative to that of unmodified SOD, even for the same charge state (e.g., 35% coverage for carbamylated SOD (12+) and 55% coverage for unmodified SOD (12+)) (Table 1). HCD of carbamylated SOD decreased as the charge state decreased (Supplementary Figure S17), and overall was sparse compared with HCD of the unmodified protein. For the 12+ charge state, fragmentation was prominent only near the N-terminus and the region spanning Gly92 to Asp99 with minor contributions near the C-terminus. For the 14+ and 16+ charge states, HCD resulted in some selective fragmentation in the middle region of the sequence, including stretches from Val27-Thr37 and Asn84-Pro100.

UVPD of both unmodified SOD (12+, 16+, and 20+) and carbamylated SOD (12+, 14+, 16+) resulted in numerous nonspecific cleavages (Supplementary Figures S18, S19) across the protein backbone and was virtually independent of charge state, yielding sequence coverages that averaged 75% (Table 1). Both N-terminal and C-terminal products were generated throughout the protein. Enhancement of cleavages adjacent to Pro residues (Pro13, Pro64, Pro72, Pro100, Pro121) was observed for both unmodified and carbamylated SOD. What is perhaps most remarkable about the fragmentation of SOD is the lack of overlap among the regions of enhanced fragmentation for HCD versus UVPD and for the carbamylated versus unmodified protein. For example, the region of greatest backbone fragmentation spanned residues Val92 to Gly106 for HCD of unmodified SOD (16+), residues Val92 to Pro100 for HCD of carbamylated SOD (16+), residues His19 to Thr86 for UVPD of unmodified SOD (16+), and residues Phe43 to Pro100 for UVPD of carbamylated SOD (12+). This comparison highlights the complementary nature of HCD and UVPD, as well as the impact of carbamylation on fragmentation.

Myoglobin

Myoglobin produced ions in charge states ranging from 9+ to 24+; after carbamylation the charge states ranged from 7+ to 16+ and the clean spectra confirmed that carbamylation

occurred with near 100% efficiency at 20 sites (19 Lys plus the N-terminus). HCD of myoglobin in the 12+, 16+, and 20+ charge states resulted in nonspecific cleavage across much of the backbone (Figure 3 and Supplementary Figure S20) and yielded an average of 50% sequence coverage (Table 1). Several prominent preferential cleavages were observed, such as N-terminal to Pro120, and C-terminal to both Leu2 and Glu6. The trend of significant *b* and *y* ion segregation (i.e., *b* ions preferentially observed for the N-terminal half of the protein and *y* ions dominating for the C-terminal half) was again noted, especially for the higher charge states. After carbamylation, there was a significant decrease in fragmentation of myoglobin upon HCD as evidenced by sequence coverages that averaged only 20% (Table 1). The few sequence ions that were observed were clustered near the first 20 residues of the N-terminus (*b* ions) or C-terminus (mostly *y* ions), leaving the mid-section of the protein unsequenced (Supplementary Figure S21). Myoglobin contains a large number of Lys residues in the region of the protein devoid of fragments (15 Lys out of 19 total Lys), suggesting that heavy modification of the internal portion of the protein disrupts the formation of fragment ions by HCD. Myoglobin contains only two highly basic Arg residues (Arg31, Arg139) in its entire sequence, meaning that the 11 His residues serve as the other basic sites most likely to be protonated during ESI. Interestingly, the regions that do yield fragment ions do not contain His residues, suggesting that His may sequester protons or otherwise hinder fragmentation of carbamylated proteins.

UVPD of myoglobin resulted in mainly nonspecific cleavages that resulted in high sequence coverage (up to 94% for the 16+ charge state and averaging 92% for all of the charge states examined, Table 1). Cleavage N-terminal to Pro120 was observed upon UVPD, just as it was prominent upon HCD, but cleavage C-terminal to Phe33, Phe46, and Phe48 were of similar relative abundance to that of the Pro120 cleavage (Supplementary Figure S22), suggesting that fragmentation may be enhanced adjacent to aromatic residues, which are known to have high UV photo-absorption cross-sections. UVPD of carbamylated myoglobin results in numerous nonspecific cleavage along the backbone (Figure 3 and Supplementary Figure S23), yielding sequence coverages from 74% to 83% for the 8+, 10+, and 12+ charge states (Table 1). Although this level of sequence coverage is lower than the average 92% coverage observed for unmodified myoglobin upon UVPD, it is three to four times greater than observed upon HCD. Several prominent backbone cleavages were noted C-terminal to Phe33 and N-terminal to several Pro residues, which were identical to those observed upon UVPD of unmodified myoglobin. In general, myoglobin showed the greatest difference in overall fragmentation between HCD of the carbamylated and unmodified protein and most similarity for UVPD of the unmodified and modified protein. This finding suggests that for proteins that are Arg poor, modification of Lys residues severely hinders production of *b* and *y* type ions upon collisional activation but has little impact on the performance or outcome of UVPD.

Carbonic Anhydrase

Upon ESI, carbonic anhydrase produced ions in charge states ranging from 14+ to 35+ (Figure 1). HCD resulted in rather sparse fragmentation of all charge states (Supplementary Figure S24), yielding sequence coverages averaging only 35%. Several dominant cleavages occurred at Pro residues, and the majority of fragments entailed backbone cleavages within 75 residues of either the C- or N-termini. The highest charge state examined by HCD (34+) displayed a shift towards fragmentation at only the C-terminus region. Carbamylation of carbonic anhydrase was efficient, resulting in modification of all 18 Lys residues plus the C-terminus and shifting the charge states to 10+ to 28+ upon ESI (Figure 1). HCD of carbamylated carbonic anhydrase resulted in preferential cleavages similar to that observed for unmodified carbonic anhydrase, albeit with even lower sequence coverage (averaging 20%) and with virtually no fragmentation along the N-terminal half of the protein (Supplementary Figure S25).

UVPD of carbonic anhydrase resulted in a larger degree of nonspecific cleavages (Supplementary Figure S26), yielding sequence coverages that averaged 63% (Table 1). Similar to HCD, several cleavages adjacent to Pro residues were moderately enhanced. UVPD of carbamylated carbonic resulted in an average sequence coverage of 32% (Table 1). Similar to what was observed for HCD, a large stretch of the protein remained unsequenced upon UVPD of the carbamylated protein (Supplementary Figure S27), although the suppression of fragmentation was less dramatic than noted for HCD. The UVPD fragmentation patterns of carbamylated carbonic anhydrase (16+, 22+, 29+) most closely resembled the UVPD fragmentation pattern of unmodified carbonic anhydrase in the 34+ charge state, supporting the observation that fragmentation patterns of carbamylated proteins are most similar to the ones obtained for the highest charge states of the unmodified counterparts.

Effect of Carbamylation on Sequence Coverage

Sequence coverage for all six proteins studied was lowest for carbamylated proteins activated by HCD and highest for unmodified proteins activated by UVPD (Table 1 with sequence maps shown in Supplementary Figures S28–33 for each of the six proteins). In general, carbamylation caused reduced sequence coverage regardless of activation method, and the reduction in sequence coverage for the carbamylated proteins was notably more precipitous for HCD than UVPD. UVPD routinely achieved higher coverage and consistently outperformed HCD for both carbamylated and unmodified proteins. Myoglobin, the most Lys-rich protein, displayed the largest disparity between HCD and UVPD of the carbamylated species (77% coverage for UVPD and only 22% for HCD over three charge states), an outcome also mirrored for the unmodified proteins (92% coverage for UVPD and only 54% for HCD on average). The relatively modest reduction in sequence coverage for UVPD of carbamylated proteins relative to unmodified proteins recapitulates the lack of significant

dependence on proton mobility for the UVPD process relative to HCD. Based on analysis of the series of fragment ion maps, HCD of carbamylated proteins enhances terminal-mediated fragmentation (i.e., enhancement of smaller *a*, *b*, *c* ions near the N-terminal portion of the protein, *x*, *y*, *z* ions near the C-terminal portion, and sparse fragmentation in the mid-section), which could explain the reduced sequence coverage as protein size increases.

Effect of Carbamylation on Cleavage Preferences

Under certain circumstances, collisional activation of proteins generates dominant fragment ions resulting from preferential cleavages, typically ones directly related to specific amino acids [15, 23]. In general, this phenomenon is most prominent upon activation of the highest and lowest charge states and is most frequently manifested in preferential cleavages N-terminal to Pro and C-terminal to Asp and Glu residues [45–47]. To assess the overall impact of carbamylation on these preferential cleavages, the distributions of key categories of fragment ions of ubiquitin (Supplementary Figure S34), cytochrome *c* (Supplementary Figure S35), and lysozyme (Figure 4) were compiled owing to the relatively extensive sequence coverage of these three proteins regardless of carbamylation or activation method.

HCD of unmodified ubiquitin showed an increasing degree of preferential N-terminal Pro cleavage with increasing charge state and a significant portion of preferential cleavages C-terminal to acidic residues (Glu and Asp) (Supplementary Figure S34), similar to previous observations by Reid et al. [20]. Upon carbamylation of ubiquitin, N-terminal Pro cleavage was significantly suppressed for the +4 charge state, whereas cleavage C-terminal to Asp and Glu were greatly enhanced, suggesting that charge remote fragmentation dominates at this low charge state generated upon carbamylation (Supplementary Figure S34). The 10+ charge state of carbamylated ubiquitin displayed a degree of N-terminal Pro cleavage upon HCD similar to that observed for the 12+ of unmodified ubiquitin, and a reduced percentage of C-terminal Asp and Glu cleavages. UVPD of both unmodified and carbamylated ubiquitin resulted in lower portions of preferential cleavages and significantly more contributions from nonspecific cleavages of the entire backbone.

HCD resulted in similar portions of preferential N-terminal Pro and C-terminal Glu and Asp cleavages of cytochrome *c* regardless of charge state, and these distributions changed only modestly upon carbamylation (Supplementary Figure S35). In particular, Pro-specific fragmentation was enhanced after carbamylation for the 8+ charge state of cytochrome *c*. The contribution of preferential cleavages diminished for HCD of the 10+ charge state of carbamylated cytochrome *c*, but this distribution was likely skewed owing to the notable enhancement of nonspecific cleavages at the C-terminus end of the protein. UVPD of cytochrome *c* displayed more dominant fragmentation C-terminal to Phe than HCD, reiterating that the aromatic chromophore played a role in directing site-specific fragmentation by UVPD. Interestingly, UVPD of both unmodified and carbamylated cytochrome *c* resulted in enhanced N-terminal Pro cleavage and somewhat suppressed C-terminal Asp and Glu cleavages compared with HCD. In addition, the sites of the preferential cleavages varied for HCD relative to UVPD. Cleavage adjacent to Pro44 and Pro76 were favored for HCD, whereas UVPD displayed enhanced N-terminal cleavage at two additional Pro residues: Pro30 and Pro71. Thus, UVPD did not show any particular discrimination of Pro residues, whereas the preferential Pro cleavage was selective upon HCD, suggesting that specific sequence motifs or charge site locations influenced the Pro cleavages of cytochrome *c* upon HCD.

Lysozyme contains only two Pro residues with both located in the middle section of the protein, a region where fragmentation is typically suppressed for top-down MS/MS methods. Upon HCD, both unmodified and carbamylated lysozyme displayed little Pro-selective cleavage, and instead preferential cleavages adjacent to Glu and Asp were more prominent (Figure 4). Upon carbamylation, a slight increase in C-terminal Asp and Glu cleavages was observed for the 8+ charge state upon HCD. The more extensive and non-selective fragmentation across the backbone by UVPD generated higher sequence coverage and resulted in ample N-terminal Pro and C-terminal Glu and Asp cleavages. Upon carbamylation, the degree of Pro cleavage decreased significantly upon UVPD, as well as a decrease in cleavage C-terminal to Phe.

LC-MS of Carbamylated E. coli Ribosomal Proteins

Carbamylation not only causes a significant change in the charge states of proteins and their fragmentation patterns but also alters the chromatographic properties of proteins. To examine the potential impact of carbamylation on a top-down LCMS/MS workflow, the *E. coli* ribosome containing 56 proteins was used as a benchmark mixture. Supplementary Figure S36 shows the base peak LCMS traces for a mixture of ribosomal proteins prior to and after carbamylation. For the mixture subjected to carbamylation, the composition of eluting proteins was checked throughout the LC run, and there was no evidence for non-carbamylated forms. In essence, carbamylation proceeded with near 100% efficiency. As specific examples of eluting proteins, Supplementary Figures S37 and S38 show extracted ion chromatograms and corresponding MS1 spectra (Figure 5) of ribosomal protein L24 from the 50S subunit prior to and after carbamylation. The significant reduction in charge state upon carbamylation is evident in Figure 5 for which the most abundant charge state shifts from 16+ (unmodified) to 11+ (carbamylated) for 50S L24. The retention time changes by over 36 min upon carbamylation, thus reflecting the increase in hydrophobicity after modification of the 16 Lys and N-terminus of this 11.2 kDa (103 residues) protein. In general, carbamylation increases the retention times of the ribosomal proteins by an average of 16 ± 7 min, an increase that is modulated by the number of Lys residues per protein (e.g., the hydrophobicity and concomitant change in

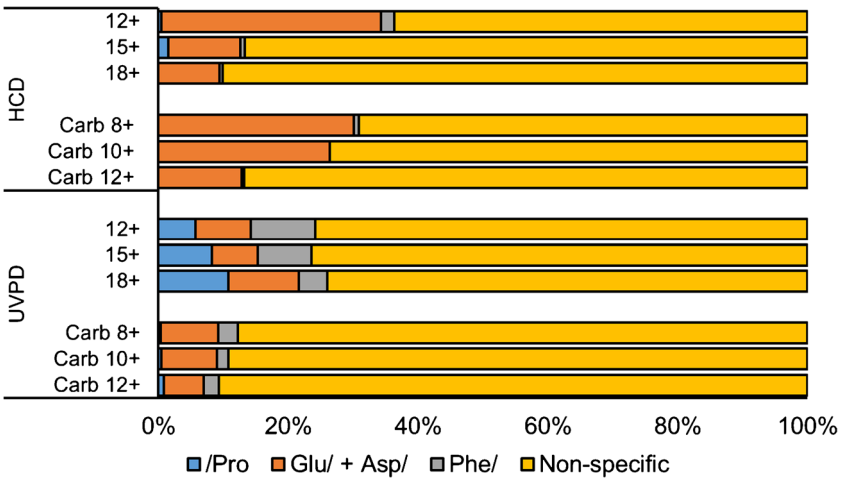


Figure 4. Distribution of fragment ions generated by HCD and UVPD categorized as preferential cleavages (N-terminal to Pro, C-terminal to Glu and Asp, C-terminal to Phe) and all nonspecific pathways (other N-terminal and C-terminal cleavages) for unmodified and carbamylated lysozyme

retention time scales with the number of carbamylated groups). The sequence maps are shown in the lower half of Figure 5, and the following sequence coverages were obtained: 62% for HCD and 80% for UVPD of the L24 protein (16+), and 27% for HCD and 71% for UVPD of the carbamylated L24 protein (10+), again exhibiting a significant decrease in coverage upon HCD of the carbamylated protein and a much smaller decrease in coverage upon UVPD.

Table 2 summarizes LC-MS results comparing HCD and UVPD for the collective set of ribosomal proteins, with specific performance results (including molecular weight, retention time, and sequence coverage) for a subset of six individual

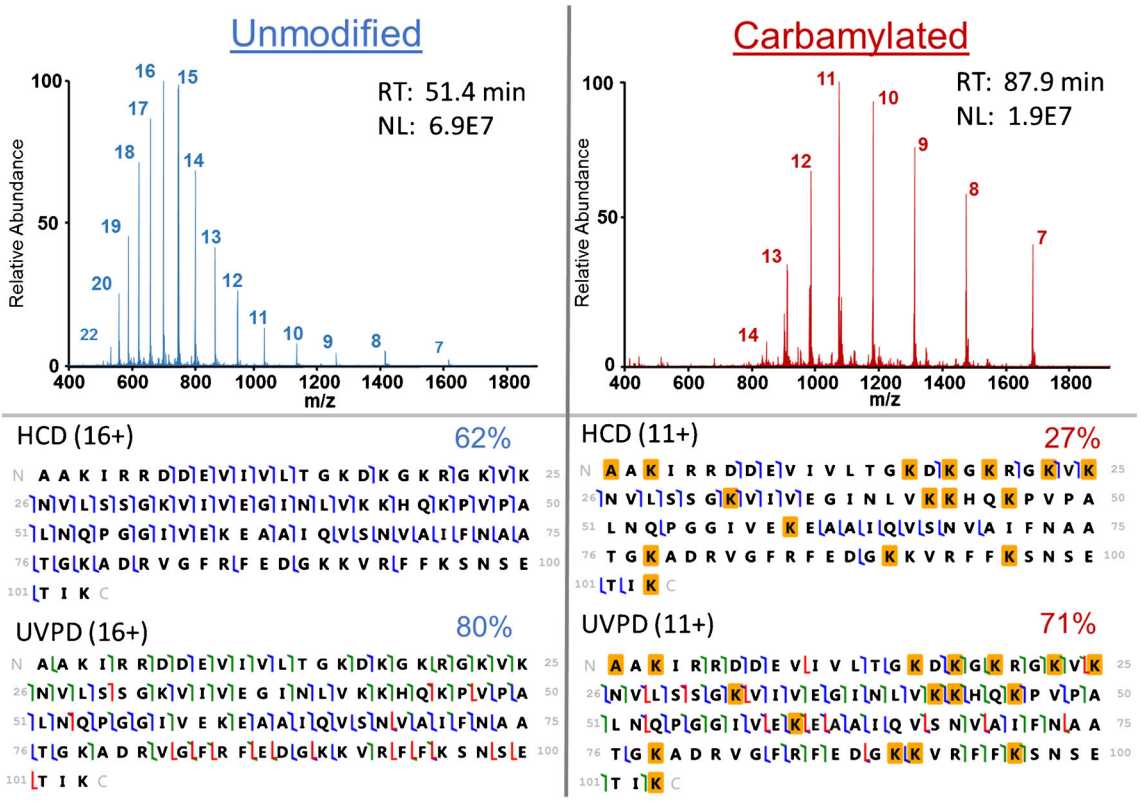


Figure 5. **Top**: ESI mass spectra of unmodified and carbamylated ribosomal protein 50S L24 from *E. coli*. **Bottom**: Sequence maps annotated for HCD fragmentation (upper sequence maps) and UVPD (lower sequence maps). The types of fragment ions generated from backbone cleavages are color coded as follows: a/x green; b/y blue; c/z red. Gold-highlighted boxes denote sites of carbamylation

Table 2. LC-MS Metrics for UVPD and HCD of a Mixture of *E. coli* Ribosomal Proteins

	HCD		UVPD	
	Unmodified	Carbamylated	Unmodified	Carbamylated
Proteins*	53 ± 1	24 ± 1	46 ± 2**	48 ± 1**
Proteoforms	180 ± 2	75 ± 3	134 ± 1	134 ± 11
Ribosomal proteins	42 ± 1	22 ± 0	40 ± 2	32 ± 4
Number of matching fragments	23 ± 12	13 ± 5	45 ± 25	38 ± 18

* Includes proteins identified in sample not associated with *E. coli* ribosomal subunits
** Not significantly different ($p = 0.19$)

proteins shown in Supplementary Table S1. Carbamylation significantly reduced the number of ribosomal proteins and matched fragment ions identified by HCD but had a far more modest impact on the UVPD results. In fact, carbamylation resulted in a greater than 50% reduction in total protein identifications for HCD (from 53 to 24 proteins), with a similar trend for the number of proteoforms characterized (180 to 75 proteoforms), number of ribosomal proteins identified (42 to 22 proteins), and average number of fragments matched (23 to 13 matching fragments per protein). For the subset of proteins analyzed in detail in Supplementary Table S1, the sequence coverage averaged 46% for the six unmodified proteins but plunged to 22% for the same carbamylated proteins. UVPD generated nearly 50% more matched fragments on average compared with HCD, regardless of carbamylation, indicating better protein characterization and sequence coverage (Table 2). Carbamylation caused little change in the number of proteins (46 versus 48) and proteoforms (134 versus 134) identified by UVPD. For the same subset of proteins reported in Supplementary Table S1, the sequence coverage averaged 70% for the six unmodified proteins by UVPD and decreased to 60% for the carbamylated proteins.

Conclusion

Carbamylation offers a highly efficient means to modify the Lys side-chains and N-terminus of proteins, occurring with nearly 100% efficiency. Carbamylation of intact protein molecules causes (1) reduction of the observed charge states upon ESI, (2) modulation of HCD and UVPD fragmentation, and (3) increases in LC retention times owing to the greater hydrophobicity after conversion of primary amines to carbamylated groups. Overall carbamylation significantly decreased the sequence coverage produced by HCD and resulted in a much more modest impact on UVPD. Fragmentation was preferentially enhanced near the N- and C-termini for HCD of carbamylated proteins (typically resulting in smaller fragment ions) with a concomitant reduction in fragmentation in the mid-sections (larger fragment ions). Although this apparent decrease in fragmentation in the mid-section of the protein could suggest that the larger fragments were converted to smaller N- and C-terminal fragment ions by secondary fragmentation pathways, there was no evidence for this premise based on examination of the fragmentation patterns, as the HCD collision energy was varied. Carbamylation had relatively little influence on the distribution or types of fragment ions generated upon UVPD, recapitulating the premise that the mechanism of UVPD is not highly dependent on mobile protons. MS/MS analysis of intact proteins remains a formidable challenge, and carbamylation offers a convenient way to modulate charge states and resolve overlapping charge state distributions as has been previously shown for gas-phase proton transfer reactions [15].

Acknowledgements

The authors acknowledge the following funding sources: NSF (Grant CHE1402753) and the Welch Foundation (Grant F-1155). S.M.G. acknowledges a graduate fellowship from the American Chemical Society Division of Analytical Sciences. Funding from the UT System for support of the UT System Proteomics Core Facility Network is gratefully acknowledged. This work was supported in part by the National Institute of General Medical Sciences P41GM108569 for the National Resource for Translational and Developmental Proteomics (NRTDP) based at Northwestern University (to N.L.K.).

References

1. Toby, T.K., Fornelli, L., Kelleher, N.L.: Progress in top-down proteomics and the analysis of proteoforms. Annu. Rev. Anal. Chem. **9**, 499–519 (2016)
2. Fornelli, L., Ayoub, D., Aizikov, K., Beck, A., Tsybin, Y.O.: Middle-down analysis of monoclonal antibodies with electron transfer dissociation Orbitrap Fourier transform mass spectrometry. Anal. Chem. **86**, 3005–3012 (2014)
3. Catherman, A.D., Skinner, O.S., Kelleher, N.L.: Top down proteomics: facts and perspectives. Biochem. Biophys. Res. Commun. **445**, 683–693 (2014)
4. 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-Tolic, 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)
5. Senko, M.W., Speir, J.P., McLafferty, F.W.: Collisional activation of large multiply charged ions using Fourier transform mass spectrometry. Anal. Chem. **66**, 2801–2808 (1994)
6. Zhou, M., Paša-Tolić, L., Stenoien, D.L.: Profiling of histone post-translational modifications in mouse brain with high-resolution top-down mass spectrometry. J. Proteome. Res. **16**(2), 599–608 (2017)
7. Fornelli, 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**(2), 609–618 (2017).
8. Syka, J.E.P., Coon, J.J., Schroeder, M.J., Shabanowitz, J., Hunt, D.F.: Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. Proc. Natl. Acad. Sci. U.S.A. **101**, 9528–9533 (2004)
9. Anderson, L.C., Karch, K.R., Ugrin, S.A., Coradin, M., English, A.M., Sidoli, S., Shabanowitz, J., Garcia, B.A., Hunt, D.F.: Analyses of histone proteoforms using front-end electron transfer dissociation-enabled Orbitrap instruments. Mol. Cell. Proteom. **15**, 975–988 (2016)
10. Cammarata, M.B., Brodbelt, J.S.: Characterization of intra- and intermolecular protein cross-linking by top down ultraviolet photodissociation mass spectrometry. ChemistrySelect. **1**, 590–593 (2016)
11. Cannon, J.R., Kluwe, C., Ellington, A., Brodbelt, J.S.: Characterization of green fluorescent proteins by 193 nm ultraviolet photodissociation mass spectrometry. Proteomics **14**, 1165–1173 (2014)
12. Cannon, J.R., Martinez-Fonts, K., Robotham, S.A., Matouschek, A., Brodbelt, J.S.: Top-down 193-nm ultraviolet photodissociation mass spectrometry for simultaneous determination of polyubiquitin chain length and topology. Anal. Chem. **87**, 1812–1820 (2015)
13. Cammarata, M., Lin, K.-Y., Pruet, J., Liu, H., Brodbelt, J.: Probing the unfolding of myoglobin and domain C of PARP-1 with covalent labeling and top-down ultraviolet photodissociation mass spectrometry. Anal. Chem. **86**, 2534–2542 (2014)
14. 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)
15. Holden, D.D., McGee, W.M., Brodbelt, J.S.: Integration of ultraviolet photodissociation with proton transfer reactions and ion parking for analysis of intact proteins. Anal. Chem. **88**, 1008–1016 (2016)
16. Cannon, J.R., Cammarata, M.B., Robotham, S.A., Cotham, V.C., Shaw, J.B., Fellers, R.T., Early, B.P., Thomas, P.M., Kelleher, N.L., Brodbelt, J.S.: Ultraviolet photodissociation for characterization of whole proteins on a chromatographic time scale. Anal. Chem. **86**, 2185–2192 (2014)
17. LeDuc, R.D., Fellers, R.T., Early, B.P., Greer, J.B., Thomas, P.M., Kelleher, N.L.: The C-Score: a Bayesian framework to sharply improve proteoform scoring in high-throughput top down proteomics. J. Proteome Res. **13**, 3231–3240 (2014)
18. Armirotti, A., Damonte, G.: Achievements and perspectives of top-down proteomics. Proteomics **10**, 3566–3576 (2010)
19. Chanthamontri, C., Liu, J., McLuckey, S.A.: Charge state dependent fragmentation of gaseous α-synuclein cations via ion trap and beam-type collisional activation. Int. J. Mass Spectrom. **283**, 9–16 (2009)
20. Reid, G.E., Wu, J., Chrisman, P.A., Wells, J.M., McLuckey, S.A.: Charge-state-dependent sequence analysis of protonated ubiquitin ions via ion trap tandem mass spectrometry. Anal. Chem. **73**, 3274–3281 (2001)
21. Amunugama, R., Hogan, J.M., Newton, K.A., McLuckey, S.A.: Whole protein dissociation in a quadrupole ion trap: identification of an a priori unknown modified protein. Anal. Chem. **76**, 720–727 (2004)
22. Mitchell Wells, J., McLuckey, S.A.: Collision-induced dissociation (CID) of peptides and proteins. In: Methods in enzymology; pp. 148–185. Elsevier: (2005)
23. Cobb, J.S., Easterling, M.L., Agar, J.N.: Structural characterization of intact proteins is enhanced by prevalent fragmentation pathways rarely observed for peptides. J. Am. Soc. Mass Spectrom. **21**, 949–959 (2010)
24. Wysocki, V.H., Tsaprailis, G., Smith, L.L., Breci, L.A.: Mobile and localized protons: a framework for understanding peptide dissociation. J. Mass Spectrom. **35**, 1399–1406 (2000)
25. Bythell, B.J., Suhai, S., Somogyi, Á., Paizs, B.: Proton-driven amide bond-cleavage pathways of gas-phase peptide ions lacking mobile protons. J. Am. Chem. Soc. **131**, 14057–14065 (2009)
26. Tabb, D.L., Huang, Y., Wysocki, V.H., Yates, J.R.: Influence of basic residue content on fragment ion peak intensities in low-energy collision-induced dissociation spectra of peptides. Anal. Chem. **76**, 1243–1248 (2004)
27. Dongré, A.R., Jones, J.L., Somogyi, Á., Wysocki, V.H.: Influence of peptide composition, gas-phase basicity, and chemical modification on fragmentation efficiency: evidence for the mobile proton model. J. Am. Chem. Soc. **118**, 8365–8374 (1996)
28. Riley, N.M., Westphall, M.S., Coon, J.J.: Activated ion electron transfer dissociation for improved fragmentation of intact proteins. Anal. Chem. **87**, 7109–7116 (2015)
29. Zhurov, K.O., Fornelli, L., Wodrich, M.D., Laskay, Ü.A., Tsybin, Y.O.: Principles of electron capture and transfer dissociation mass spectrometry applied to peptide and protein structure analysis. Chem. Soc. Rev. **42**, 5014 (2013)
30. Pesavento, J.J., Kim, Y.-B., Taylor, G.K., Kelleher, N.L.: Shotgun annotation of histone modifications: a new approach for streamlined characterization of proteins by top down mass spectrometry. J. Am. Chem. Soc. **126**, 3386–3387 (2004)
31. Karch, K.R., DeNizio, J.E., Black, B.E., Garcia, B.A.: Identification and interrogation of combinatorial histone modifications. Front. Genet. **4** (2013)
32. Morrison, L.J., Rosenberg, J.A., Singleton, J.P., Brodbelt, J.S.: Statistical examination of the a and a + 1 fragment ions from 193 nm ultraviolet photodissociation reveals local hydrogen bonding interactions. J. Am. Soc. Mass Spectrom. **27**, 1443–1453 (2016)
33. Iavarone, A.T., Jurchen, J.C., Williams, E.R.: Supercharged protein and peptide ions formed by electrospray ionization. Anal. Chem. **73**, 1455–1460 (2001)
34. Iavarone, A.T., Williams, E.R.: Collisionally activated dissociation of supercharged proteins formed by electrospray ionization. Anal. Chem. **75**, 4525–4533 (2003)
35. Teo, C.A., Donald, W.A.: Solution additives for supercharging proteins beyond the theoretical maximum proton-transfer limit in electrospray ionization mass spectrometry. Anal. Chem. **86**, 4455–4462 (2014)
36. Zenaidee, M.A., Donald, W.A.: Extremely supercharged proteins in mass spectrometry: profiling the pH of electrospray generated droplets, narrowing charge state distributions, and increasing ion fragmentation. Analyst **140**, 1894–1905 (2015)
37. Krusemark, C.J., Frey, B.L., Belshaw, P.J., Smith, L.M.: Modifying the charge state distribution of proteins in electrospray ionization mass spectrometry by chemical derivatization. J. Am. Soc. Mass Spectrom. **20**, 1617–1625 (2009)
38. Pitteri, S.J., Reid, G.E., McLuckey, S.A.: Affecting proton mobility in activated peptide and whole protein ions via lysine guanidination. J. Proteome Res. **3**, 46–54 (2004)
39. Greer, S.M., Cannon, J.R., Brodbelt, J.S.: Improvement of shotgun proteomics in the negative mode by carbamylation of peptides and ultraviolet photodissociation mass spectrometry. Anal. Chem. **86**, 12285–12290 (2014)
40. Klein, D.R., Holden, D.D., Brodbelt, J.S.: Shotgun Analysis of Rough-Type Lipopolysaccharides using ultraviolet photodissociation mass spectrometry. Anal. Chem. **88**, 1044–1051 (2016)
41. Fellers, R.T., Greer, J.B., Early, B.P., Yu, X., LeDuc, R.D., Kelleher, N.L., Thomas, P.M.: ProSight Lite: graphical software to analyze top-down mass spectrometry data. Proteomics **15**, 1235–1238 (2015)
42. Tsaprailis, 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)
43. Qin, J., Chait, B.T.: Preferential fragmentation of protonated gas-phase peptide ions adjacent to acidic amino acid residues. J. Am. Chem. Soc. **117**, 5411–5412 (1995)
44. Jockusch, R.A., Schnier, P.D., Price, W.D., Strittmatter, E.F., Demirev, P.A., Williams, E.R.: Effects of charge state on fragmentation pathways, dynamics, and activation energies of ubiquitin ions measured by blackbody infrared radiative dissociation. Anal. Chem. **69**, 1119–1126 (1997)
45. Breci, L.A., Tabb, D.L., Yates, J.R., Wysocki, V.H.: Cleavage N-terminal to proline: analysis of a database of peptide tandem mass spectra. Anal. Chem. **75**, 1963–1971 (2003)
46. Newton, K.A., Pitteri, S.J., Laskowski, M., McLuckey, S.A.: Effects of single amino acid substitution on the collision-induced dissociation of intact protein ions: Turkey ovomucoid third domain. J. Proteome Res. **3**, 1033–1041 (2004)
47. Paizs, B., Suhai, S.: Fragmentation pathways of protonated peptides. Mass Spectrom. Rev. **24**, 508–548 (2005)