HHS Public Access

Author manuscript

J Am Soc Mass Spectrom. Author manuscript; available in PMC 2016 September 01.

Published in final edited form as:

J Am Soc Mass Spectrom. 2015 September ; 26(9): 1580–1587. doi:10.1007/s13361-015-1168-0.

ETD Outperforms CID and HCD in the Analysis of the Ubiquitylated Proteome

Tanya R. Porras-Yakushi, Michael J. Sweredoski, and Sonja Hess

California Institute of Technology, Beckman Institute, Pasadena, CA 91125, USA

Abstract

Comprehensive analysis of the ubiquitylome is a prerequisite to fully understand the regulatory role of ubiquitylation. However, the impact of key mass spectrometry parameters on ubiquitylome analyses has not been fully explored. In this study, we show that using electron transfer dissociation (ETD) fragmentation, either exclusively or as part of a decision tree method, leads to ca. 2-fold increase in ubiquitylation site identifications in K- ϵ -GG peptide-enriched samples over traditional collisional-induced dissociation (CID) or higher-energy collision dissociation (HCD) methods. Precursor ions were predominantly observed as 3+ charged species or higher and in a mass range 300–1200 m/z . *N*-ethylmaleimide was used as an alkylating agent to reduce false positive identifications resulting from overalkylation with halo-acetamides. These results demonstrate that the application of ETD fragmentation, in addition to narrowing the mass range and using *N*-ethylmaleimide yields more high-confidence ubiquitylation site identification than conventional CID and HCD analysis.

Graphical abstract

Correspondence to: Sonja Hess; shess@caltech.edu.

Electronic supplementary material The online version of this article (doi:10.1007/s13361-015-1168-0) contains supplementary material, which is available to authorized users.

Ubiquitylated Peptides

Keywords

Ubiquitin; Ubiquitylation; Mass spectrometry; Proteomics; K- ϵ -GG antibody; Electron transfer dissociation; Collisional-induced dissociation

Introduction

Proteomic studies have sought to understand the role of ubiquitin conjugation by identifying proteome-wide ubiquitin modified proteins. Recent advances in the use of mass spectrometry have greatly enhanced our current understanding of the repertoire of proteins that are ubiquitylated [1–6]. The combined use of the di-glycine (GG) remnant antibody [7, 8] and advances in orthogonal peptide prefractionation, have led to the identification of more than 10,000 putative ubiquitylation sites in human cell lines [6, 9–13]. Advances in mass spectrometers have been instrumental in the characterization of the ubiquitylome; however, optimization of specific parameters needs further exploration. Currently, the highest number of ubiquitylation sites reported in *Saccharomyces cerevisiae*, Baker's yeast, is 2299, but this is believed to be a fraction of the total sites that can occur in the whole proteome, especially when one considers that most proteins are eventually targeted for degradation by the ubiquitin proteasome system (UPS) [12, 14].

Ubiquitin is a regulatory protein that is conjugated to target proteins via a cascade of E1, E2, and E3 enzymes that determine the site and complexity of ubiquitylation. Conjugation of ubiquitin to a specific lysine residue in the target protein results in the formation of an isopeptide bond between the ϵ -amino group of lysine and the C-terminus of ubiquitin. Fortunately, the C-terminus of ubiquitin ends in RGG, allowing ubiquitylated proteins to yield tryptic peptides that have a GG remnant covalently attached to the modified lysine,

giving rise to K-ε-GG peptides. Therefore, the GG moiety can be used to identify specific lysine residues in the target protein where ubiquitin was linked, based on a mass modification of +114.0429.

Many initial proteomic ubiquitylome studies employed low resolution collisional-induced dissociation (CID) to fragment K-ε-GG peptides [4, 15, 16]. With the advent of higher-energy collision dissociation (HCD), it became possible to acquire high-resolution MS/MS spectra with an acceptable cycle time leading to similar or even increased identification over CID [9, 14, 17]. Studies have shown that K-ε-GG peptides, hereafter referred to as GG peptides, are predominantly present as higher charged species [11]. The higher charge results from the additional N-terminus that remains at the end of the GG adduct after trypsin cleavage and the additional unmodified Lys or Arg following the modified Lys. Note that trypsin typically does not cleave after the modified Lys, which results in a missed cleavage. As shown in Figure 1a, one missed cleavage resulting from a conjugated GG adduct gives rise to a triply charged species (N-terminus, Lys or Arg side chain at the C-terminus, N-terminus of GG remnant attached via isopeptide bond, thus creating a branched peptide). The presence of a histidine residue in a GG peptide can add yet an additional charge to the peptide, resulting in higher charged species.

Promising studies using electron capture dissociation (ECD) on a Fourier transform ion cyclotron mass spectrometer (FT-ICR) was able to use ECD fragmentation to fully characterize the ubiquitylation sites on a tryptically digested purified protein and additionally characterize the autoubiquitylation sites on ubiquitin demonstrating that a softer fragmentation method could better preserve the GG moiety at the site of modification [18]. Later fundamental studies, using electron transfer dissociation (ETD), suggested that a higher yield of ubiquitylation identifications could be achieved using ETD instead of CID and HCD because it is more amenable to fragmenting higher charged species and preserving PTMs than CID [19–21]. The use of ETD techniques for the preferential fragmentation of GG modified peptides has been tested on single ubiquitylated proteins and has been shown to result in greater site coverage than CID [20, 22]. However, the use of ETD on a complex mixture of digested proteins enriched in ubiquitylated peptides has so far not been tested. Although it has previously been shown that individual standard GG peptides have benefited from ETD analysis, the proteomics field has been hesitant to adopt ETD fragmentation on global ubiquitylation site analyses, possibly because the increased length of the duty cycles is expected to negatively affect the outcome. Additionally, ETD requires specialized equipment and optimization, which may have discouraged adoption. In this study, we demonstrate that ETD outperforms both CID and HCD in a global ubiquitylation analysis and explore methodological and mass spectrometric factors that will encourage the use of ETD for global ubiquitylation studies.

Overalkylation is an added problem that must be considered while performing ubiquitylome analysis [23, 24]. Nielsen et al. demonstrated that overalkylation of proteins with iodoacetamide (IAA) can result in the addition of two carbamidomethyl groups at lysine residues, equal in molecular composition ($C_8H_{12}N_4O_4$) and mass (+114.04292 Da) to a GG remnant [24]. The authors suggested instead using chloroacetamide (CIAA) because it is less reactive [24]. However, the underlying problem can remain since overalkylation with IAA

and CIAA yield the same product leading to potential false PTM assignments. *N*-ethylmaleimide (NEM) is an alternate alkylating agent that can be used to prevent disulfide bridges from reforming, but most importantly the use of NEM adds a mass tag of +125.0477 Da. By using NEM to alkylate the side chains of cysteine residues, we circumvent the potential problem since the mass modification by NEM would be different than that of the GG remnant.

In this study, we explored GG peptide identification by CID, HCD, ETD, and data-dependent decision tree (DT) fragmentation. We demonstrate the vast improvement made by the use of ETD fragmentation for the proteome-wide identification of ubiquitylation supported by the prevalence of these peptides in higher charge states. In our study, ETD fragmentation resulted in the identification of a higher number of unique GG peptides in a complex proteomic sample than traditional CID or HCD analysis.

Experimental Procedures

Synthetic Peptide Sample Processing

Custom synthetic peptides purchased from JPT Innovative Peptide Solutions (Berlin, Germany) were synthesized with the following amino acid sequence AMLK(GG)SEQNR, where (GG) represents a di-glycine unit connected to the peptide via the ϵ -amino group of lysine. Dried peptides were resuspended in LC-MS water to a concentration of 1 nmol/ μ L and stored at -20°C . Immediately before analysis, peptide solutions were diluted to a concentration of 500 fmol/ μ L and a total of 1 pmol was analyzed by nanoLC-MS/MS.

Yeast strains and Growth Conditions

Wild type *S. cerevisiae* strain, RJD360 (W303 background) was used in this study. Yeast cells were grown at 30°C in YPD media using standard methods and conditions. A culture of 800 mL was inoculated at a starting OD_{600nm} of 0.1 and grown to an optical density of 0.6–1.0. Cells were then harvested by centrifugation at $5000 \times g$ for 10 min at 4°C , washed once with 20 mL of ice-cold water, and harvested at $5000 \times g$ for 5 min at 4°C . The cell pellet was then frozen in liquid nitrogen and stored at -80°C until lysis.

Digestion and Desalting

Cell lysis, digestion and peptide desalting procedures were adapted from the PTMScan Ubiquitin Remnant Motif (K- ϵ -GG) Kit #5562 Cell Signaling Technology product manual. Briefly, yeast cells were lysed in 5 mL of lysis buffer (20 mM HEPES (pH 8.0), 9 M urea, $1 \times$ protease inhibitor cocktail (Promega, Madison (WI), USA), 1 mM PMSF) and 4 mL of glass beads by vortexing 1 min followed by a 1 min incubation on ice, seven times. The lysate was collected and centrifuged at $16,000 \times g$ for 15 min. Protein concentration of the lysate was then determined by Bradford. Cleared lysate containing 10 mg of protein was reduced for 45 min by adding 1/278th (v/v) of 1.25 M DTT. Alkylation of cysteines was performed by treating the lysate with 250 mM NEM dissolved in H₂O (25 $\times$ stock) to achieve a final concentration of 10 mM NEM, for 30 min at room temperature in the dark. For trypsin digestion, lysate was diluted to 2 M urea by adding 100 mM Tris (pH 8.0). Proteins were digested by trypsin using a ratio of 1:100. Digestion was carried out overnight

(15 h) at room temperature in the dark. The following morning the reaction was quenched by the addition of formic acid to a final concentration of 0.2%.

Digested peptides were centrifuged at $16,000 \times g$ for 15 min to remove insoluble material. Cleared peptides were desalted by SepPak using a 500 mg capacity column. Briefly, resin was hydrated using 7 column volumes of acetonitrile (21 mL), followed by equilibration with 7 column volumes of Buffer A (0.2% TFA in H_2O) (21 mL). Peptides were loaded onto the resin by gravity flow. After binding, the resin was washed with 7 column volumes of Buffer A and 3 column volumes of wash buffer (0.2% TFA, 5% acetonitrile in H_2O). Desalted peptides were recovered using 2 column volumes of elution buffer (0.2% TFA, 40% acetonitrile in H_2O) and lyophilized to dryness.

K- ϵ -GG Antibody Cross-Linking and Immunoprecipitation

In short, K- ϵ -GG peptide-specific antibody (PTMScan Ubiquitin Remnant Motif (K- ϵ -GG) Kit #5562, Limited Use License, Cell Signaling Technology) was washed with 3×1 mL aliquots of 100 mM sodium borate (pH 9.0). Antibody bound beads were pelleted after each wash by centrifugation at $2000 \times g$ for 30 s and kept on ice whenever possible. After washing, the beads were incubated for 30 min in 1 mL of DMP cross-linking solution (100 mM sodium borate, pH 8.0, 20 mM dimethyl pimelimidate, DMP) for 30 min at room temperature with gentle rotation. The cross-linking reaction was quenched by first washing the beads with 3×1 mL aliquots of 200 mM ethanolamine blocking buffer (pH 8.0) then incubating with 1 mL of ethanolamine blocking buffer for 2 h at 4°C. After blocking the antibody-bound beads were washed with 3×1 mL aliquots of 1X IAP buffer (50 mM MOPS, pH 7.2, 10 mM sodium phosphate, and 50 mM NaCl), then incubated with the desalted peptide sample for 1 h at 4°C. Before incubating with cross-linked antibody, the desalted peptide sample was first resuspended in 1.0 mL of 1X IAP buffer, the pH was measured (should be $pH \approx 7$), and cleared by spinning at maximum speed for 5 min. After incubating the beads with the peptide sample, the beads were pelleted by centrifugation at $2000 \times g$ for 1 min, resuspended in 500 μ L of 1X IAP, and transferred to a 0.67-mL tube and washed three times with 500 μ L of 1X IAP buffer. Following the IAP washes, the beads were washed twice with 1X PBS and once with mass spectrometry grade water (Fluka, Seelze, Germany). Finally, the bound K- ϵ -GG peptides were eluted with 2×150 μ L aliquots of 0.15% TFA, each time incubating the beads with elution buffer for 10 min at room temperature with constant mixing. The eluents were combined, dried, desalted by HPLC using a Michrom Bioresources, (Auburn (CA), USA) C18 macrotrap, (Buffer A: 0.2% formic acid in H_2O ; Buffer B: 0.2% formic acid in acetonitrile) and concentrated in vacuo.

NanoLC-MS/MS Analysis

Dried peptide samples were acidified by resuspending in Buffer A (0.2% formic acid, 2% acetonitrile), and subjected to proteomic analysis using an EASY II nano-UPLC (Thermo Scientific, Waltham (MA), USA) connected on-line to an Orbitrap Elite hybrid mass spectrometer with a nanoelectrospray ion source (Thermo Scientific) with settings and instrument arrangements similar to those described previously [25]. Peptide separation was performed on a 15 cm silica analytical column with a 75 μ m i.d. packed in-house with reversed phase ReproSil-Pur C18AQ 3 μ m resin (Dr. Maisch GmbH, Amerbuch-Entringen,

Germany). Flow rate was set to 350 nL/min, using a gradient of 2%–32% B (Buffer A: 0.2% formic acid, 2% acetonitrile, 97.8% nanoLC grade H₂O; Buffer B: 0.2% formic acid, 80% acetonitrile, 19.8% nanoLC grade H₂O). Whole cell digest immuno-precipitates and mass spectrometry-detectable samples were analyzed on 159 min gradients, while synthetic peptide samples were analyzed by 30 min gradients. The mass spectrometer was programmed to collect data in a data-dependent mode, switching automatically between full-scan MS and tandem MS acquisition. Survey full scan MS spectra were acquired in the Orbitrap using an AGC target of 1,000,000, between an m/z range of 300 to 1200 and a resolution of 120,000. The 15 most intense precursor ions were isolated and after the accumulation of 5000 ions were fragmented in the linear ion trap by ETD using a normalized collisional energy of 35% and an isolation window of 2.0 Da, while the 20 most intense ions were selected for CID. Unlike CID and ETD, HCD fragment ions were analyzed in the Orbitrap using a normalized collisional energy of 30, an activation time of 100 ms, and selection of the top 15 parent ions for fragmentation. Supplemental activation was enabled while analyzing by ETD, in both ETD only and DT mode. For all analyses, precursor ion charge state screening was enabled, singly charged and unassigned charge state species were rejected, and the dynamic exclusion window was set to 60 s maximum retention time. During synthetic peptide analysis, dynamic exclusion was not enabled. Samples were analyzed using various fragmentation methods including only ETD, only CID, only HCD, and DT (i.e., CID or ETD depending on peptide charge). DT settings [26] were set as follows: all 2+ charged species with a m/z value of ≤ 1200 were targeted for fragmentation by CID, whereas 3+ and higher charge states with an m/z value of ≤ 1200 were set to trigger ETD fragmentation. The parameters used for ETD and DT fragmentation are summarized in Table 1.

Data Analysis

For the ubiquitylome analysis, a GG peptide-enriched sample immunoprecipitated from yeast whole cell lysate was analyzed by five technical replicates of CID, HCD, ETD, and DT. A total of 20 raw files were processed together using MaxQuant (ver. 1.5.1.2) [27–29]. Spectra were searched against the yeast proteome from SGD (5898 entries, download 1/15/2010) and a contaminant database (245 entries) as well as a decoy database of equal size. Protein, peptide, and site false discovery rates were less than 1% across the entire dataset and were estimated using a target-decoy approach [30]. Trypsin was specified as the digestion enzyme with up to two allowed missed cleavages. Precursor mass tolerance was 6 ppm and fragment ion tolerance was 0.5 Da for CID and ETD analyses, whereas a fragment mass tolerance of 20 ppm was used for HCD. At the spectrum level, a posterior error probability of less than 0.05 gave a spectrum FDR of less than 0.001 across the entire dataset. *N*-ethylmaleimide modification of cysteine (+125.0477) was specified as a fixed modification. Variable modifications included protein N-terminal acetylation (+42.0106), methionine oxidation (+15.9949), and the GG remnant (+114.0429). Losses of 57.0215 and 114.0429 Da were used to account for fragmentation in the GG remnant. Peptides modified at the C-terminus by a GG adduct were excluded from the MaxQuant analysis and consequently from the final results. Please note, by “unique site” we are referring to a unique position in a protein that is modified by ubiquitin, whereas “unique peptide” refers to a specific combination of residues and modifications in a peptide. Figures 2 and 4 were

plotted to show the differences between the mean performances for the various fragmentation methods as recommended by Kryzwywinski and Altman [31].

Results and Discussion

Comparing Fragmentation Strategies

Our initial study began by trying to understand the nature of GG peptides and, consequently, use this information to select the optimal fragmentation strategy for their analysis. A synthetic peptide containing a GG remnant covalently attached to the ϵ -amino group of the single internal lysine residue in the peptide was analyzed by conventional CID and HCD fragmentation and compared with ETD fragmentation. The sequence of the synthetic peptide is depicted in Figure 1a and illustrates the prevalence of GG peptides in higher charge states attributable to the presence of an additional N-terminus. The abundance of the triply charged species was compared with that of the doubly charged species in Skyline [32] and found to be considerably higher (Figure 1b). Additional synthetic GG containing peptides were evaluated and found to display a similar trend (data not shown). MS/MS spectrum produced during ETD fragmentation of the synthetic peptide, AMLK(GG)SEQNR, is illustrated in Figure 1c. Complete sequence coverage was observed when the AMLK(GG)SEQNR peptide was fragmented by ETD. We were able to reliably observe fragment ions c_2 – c_8 along with z_2 – z_8 in the nine-residue peptide. CID fragmentation of the synthetic peptide was also performed and we observed similar fragmentation coverage of the synthetic peptide sample (Supplementary Figure 1).

Another consideration in the identification of ubiquitin modification sites in global analyses is the possibility of observing false positive identifications resulting from overalkylation. During bottom-up ubiquitylome analyses, before digestion with trypsin, proteins are reduced, then alkylated to prevent the reforming of disulfide bridges between cysteine residues. Unfortunately, as with all proteomic analyses, the more treatments performed, the greater the likelihood of observing a larger mixture of contaminant peptides. In the case of using the alkylation reagent iodoacetamide (IAA), the possibility of observing false positive identification of GG-containing peptides increases, resulting from the equivalent mass and atomic composition of a carbamidomethyl group and a glycine [22, 23]. If two carbamidomethyl groups become attached to the ϵ -amino group of lysine, it is practically indistinguishable from the addition of a GG remnant and may lead to a false identification. This observation has been reported before and the authors concluded that the use of chloroacetamide (CIAA) is better suited for ubiquitylome analyses because it is less reactive than IAA and should prevent the addition of carbamidomethyl groups at lysine residues [23]. However, the problem remains that the mass tag of the reagent is still equivalent to the mass tag of a glycine residue and, therefore, indistinguishable from a true GG addition, if overalkylation occurs. We chose to preclude any false positive identification of ubiquitylation sites due to overalkylation by using NEM to alkylate reduced cysteine residues in our peptide samples. The conditions for treating a complex mixture with NEM were determined (data not shown) and found to be optimal at a concentration of 10 mM for 30 min at room temperature in the dark. At this concentration, we achieved adequate modification of cysteine residues, while minimizing NEM adduct addition to non-cysteine

residues. For the remainder of the ubiquitylation analyses described, 10 mM NEM was used for the alkylation of cysteine residues.

It was previously suggested that ETD fragmentation would be well suited for the analysis of ubiquitylation sites because of its ability to fragment higher charged species and species with labile post-translational modifications (PTMs) better than CID, resulting in increased sequence coverage [18–20]. Previous studies had demonstrated the efficiency of ETD fragmentation on trypsin-digested purified ubiquitylated proteins, (diubiquitin, histones, and DNA polymerase B1) [20, 22], yet ETD has never been used in a global bottom-up analysis. We therefore chose to evaluate the performance of ETD on the identification of GG-containing peptides in a complex proteomic sample. In short, GG peptides were immunoprecipitated from whole cell yeast lysate using the GG-specific antibody and analyzed by five technical replicates each of CID, HCD, ETD, and decision tree (DT) fragmentation for comparison (Figure 2). DT parameters were set to allow all 2+ species to be fragmented by CID fragmentation irrespective of m/z value, whereas all species 3+ and higher irrespective of m/z were set to fragment by ETD. When we simply counted the total number of unique sites identified by each method after five technical replicates, we observed that the use of ETD fragmentation consistently outperformed CID/HCD fragmentation, while DT fragmentation with the ETD option, performed similarly to all ion ETD (Figure 2a). When we compared the difference between the average number of GG sites identified by each fragmentation method to traditional CID fragmentation, we observed that using ETD or DT led to significantly higher numbers of GG modification sites with P -values of 1.89×10^{-3} and 1.04×10^{-3} , respectively (Figure 2b). ETD and DT were also compared and not found to be statistically different in their performance; the difference in the average number of unique GG site identification was 4.0 with a P -value of 0.725 (Figure 2b). When the uniqueness of the sites identified in each fragmentation strategy was compared, we found ETD and DT fragmentation identified more unique sites not identified in traditional CID and HCD analyses (Figure 2c). ETD identified 100 and 126 ubiquitylation sites, while DT identified 96 and 135 sites, not found by CID or HCD, respectively. This clearly demonstrates that the use of ETD or DT greatly increases ubiquitin site identification by ca. 2 fold.

In order to demonstrate that GG peptides have unique characteristics that make them better suited for analysis by ETD fragmentation, we evaluated the charge state distribution of GG peptides and compared them with the charge distribution of peptides observed in a non-enriched sample. The non-enriched sample was a tryptically digested yeast lysate. We observed that GG peptides did not follow the normal distribution observed for traditional peptide mixtures. In the non-enriched sample, approximately 60% of the peptides were observed in the 2+ charge state, approximately 30% were observed in the 3+ charge state, while ~10% were in a charge state of 4+ and higher (Figure 3). This observation was consistent with previous findings where roughly 70% of identifiable peptides from a whole cell tryptic peptide mixture were found as 2+ species [33]. In the GG peptide-enriched sample, the distribution shifted and displayed prevalence for higher charged states. We observed approximately 18% of the peptides in a 2+ charge state, ~62% of the peptides were in a 3+ charge state, and approximately 20% of the total GG peptides observed were in a

charge state of 4+ and higher (Figure 3). The GG peptide charge state population was consistent with what we had observed with the synthetic GG peptides.

When considering all peptides identified in the immunoprecipitate, GG modified peptides plus unmodified peptides, we observed that ETD and CID performed similarly, while DT identified more peptides than strictly CID or ETD (Figure 4). The average number of peptides identified in ETD and CID replicates differed by only 166 peptides (*P*-value: 0.654). This is in sharp contrast to the comparison of DT to CID and DT to ETD of all peptides, where 1236.6 (*P*-value: 0.011) and 1070.2 (*P*-value: 0.00362) more peptides were observed, respectively. Thus, DT would prove advantageous over ETD alone, when the identification of GG peptides and non-GG peptides are important in the analysis.

In addition to type of fragmentation, we wanted to determine if there were other parameters that could be optimized for the analysis of GG peptides. We compared the number of unique GG peptides identified binned by precursor ion *m/z* values (Figure 5). In preliminary studies, we observed very few GG peptides above a mass range of 1200 *m/z*. Additionally, sites were observed as low as 300 *m/z*; smaller peptides were not considered since they would not yield reliable peptide spectrum matches. GG peptides were preferentially present between a mass range 300–1200 *m/z* across all 20 replicates, independent of fragmentation method used. We therefore propose adjusting the mass range to 300–1200 *m/z* (from commonly used standard settings of 300–1700 *m/z*) [25] for future GG peptide analyses.

Analysis of Ubiquitylation Identifications

To increase confidence in the ubiquitylation sites presented, we performed a thorough analysis of the sites identified. Potential false positives were excluded from consideration in MaxQuant by restricting GG remnant modifications to internal or peptide N-terminal lysines. This is because trypsin is unlikely to cleave on the carboxyl side of a modified lysine [34]. Additionally, a localization probability of 0.5 or higher was used to exclude false identification of sites. If a site was identified as having a localization probability equal to 0.5 at two different lysine residues in the same peptide, they were combined into a single identification. A complete list of ubiquitylation sites identified in this study is presented in Supplementary Table 1. Sites identified are listed along with their corresponding protein open reading frame, position in the protein, modified peptide sequence, precursor charge states, and whether the site was previously reported in SCUD, SGD, or other recent studies [12, 14]. Additionally, the fragmentation methods in which each site was observed are included in the column labeled “Fragmentation method site observed” and a column indicating if the site was only observed while using ETD or DT fragmentation has also been included. The specific site of modification for each entry listed in the “Modified peptide sequences” column is represented by a K(gl) highlighted in bold. Interestingly, of the 235 unique ubiquitylation sites identified in this study, 106 sites (45%) were only found by ETD and/or DT, and not by CID and HCD; and 38 of these sites had not been previously reported in the literature [12, 14]. Another interesting aspect of this study is that the yeast samples used were not treated with proteasome inhibitors or peptide pre-fractionated prior to performing the GG-peptide immunoprecipitation. Future studies should demonstrate the usefulness of ETD fragmentation to more highly enriched GG-peptide samples.

Peptide spectrum matches of all ubiquitylation sites identified in this study are included in the Supplementary Data (SiteSpectra.pdf). Sequence information on the GG peptides identified is listed in Supplementary Table 1, whereas sequence coverage information on all proteins identified in this analysis is included in Supplementary Table 2. A total of 1141 proteins were identified in this study and 216 unique proteins were found to be ubiquitylated, of which are included protein isoforms and paralogs. Contaminants and decoys were removed from the tables for brevity. For completeness, we also determined the duty cycle values for all fragmentation methods used. The duty cycle time for CID was ~1.4 s, HCD was ~7 s, ETD was ~4.8 s, and DT was ~4.2 s. Although the duty cycle times were considerably longer for the ETD and DT methods compared with CID fragmentation, we were still able to identify a higher number of ubiquitylation sites using ETD fragmentation.

Conclusion

In this study, we explored several parameters for increasing the number of high-confidence ubiquitylation site identifications. We used the novel approach of applying ETD and DT fragmentation to the proteomic analysis of ubiquitylated peptides. Additionally, we precluded false identification of ubiquitylation sites by using the alkylation agent NEM. Finally, we extensively curated our list of identifications by excluding false positive GG peptide identifications resulting from unlikely tryptic cleavages. We conclusively demonstrate that ETD fragmentation in either an all ion or DT method greatly improves the comprehensive analysis of ubiquitylation proteome-wide. In the future, we anticipate ETD being applied to other global ubiquitylation studies and aiding in better cataloging sites of ubiquitylation in future biological work.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Dr. Raymond J. Deshaies for helpful suggestions and critical discussions of the work described. In addition, they thank Dr. Deshaies for kindly providing them with the yeast strain RJD 360 (W303 background). Lastly, they thank members of the Proteome Exploration Laboratory, housed in the Beckman Institute at the California Institute of Technology, for helpful discussions during the course of this work. This work was supported by the Gordon and Betty Moore Foundation, through grant GBMF775, the Beckman Institute and the NIH through grant 1S10RR029594-01A1.

References

1. Porras-Yakushi TR, Hess S. Recent advances in defining the ubiquitylome. *Expert Rev Proteomics*. 2014; 11:477–490. [PubMed: 24961939]
2. Low TY, Magliozzi R, Guardavaccaro D, Heck AJR. Unraveling the ubiquitin-regulated signaling networks by mass spectrometry-based proteomics. *Proteomics*. 2013; 13:526–537. [PubMed: 23019148]
3. Sylvestersen KB, Young C, Nielsen ML. Advances in characterizing ubiquitylation sites by mass spectrometry. *Curr Opin Chem Biol*. 2013; 17:49–58. [PubMed: 23298953]
4. Anania VG, Pham VC, Huang X, Masselot A, Lill JR, Kirkpatrick DS. Peptide level immunoaffinity enrichment enhances ubiquitination site identification on individual proteins. *Mol Cell Probes*. 2014; 13:145–156.

5. Polge C, Uttenweiler-Joseph S, Leulmi R, Heng AE, Burlet-Schiltz O, Attaix D, Taillandier D. Deciphering the ubiquitin proteome: limits and advantages of high throughput global affinity purification-mass spectrometry approaches. *Int J Biochem Cell Biol.* 2013; 45:2136–2146. [PubMed: 23764619]
6. Xu G, Jaffrey SR. Comprehensive profiling of protein ubiquitination for drug discovery. *Curr Pharm Des.* 2013; 19:3315–3328. [PubMed: 23151132]
7. Warren MRE, Parker CE, Mocanu V, Klapper D, Borchers CH. Electrospray ionization tandem mass spectrometry of model peptides reveals diagnostic fragment ions for protein ubiquitination. *Rapid Commun Mass Spectrom.* 2005; 19:429–437. [PubMed: 15655800]
8. Peng J, Schwartz D, Elias JE, Thoreen CC, Cheng D, Marsischky G, Roelofs J, Finley D, Gygi SP. A proteomics approach to understanding protein ubiquitination. *Nat Biotechnol.* 2003; 21:921–926. [PubMed: 12872131]
9. Udeshi ND, Svinkina T, Mertins P, Kuhn E, Mani DR, Qiao JW, Carr SA. Refined preparation and use of anti-diglycine remnant (K-ε-GG) antibody enables routine quantification of 10,000 s of ubiquitination sites in single proteomics experiments. *Mol Cell Proteomics.* 2013; 12:825–831. [PubMed: 23266961]
10. Udeshi ND, Mertins P, Svinkina T, Carr SA. Large-scale identification of ubiquitination sites by mass spectrometry. *Nat Protoc.* 2013; 8:1950–1960. [PubMed: 24051958]
11. Udeshi ND, Mani DR, Eisenhaure T, Mertins P, Jaffe JD, Clauser KR, Hacohen N, Carr SA. Methods for quantification of in vivo changes in protein ubiquitination following proteasome and deubiquitinase inhibition. *Mol Cell Proteomics.* 2012; 11:148–159. [PubMed: 22505724]
12. Starita LM, Lo RS, Eng JK, von Haller PD, Fields S. Sites of ubiquitin attachment in *Saccharomyces cerevisiae*. *Proteomics.* 2012; 12:236–240. [PubMed: 22106047]
13. Bustos D, Bakalarski CE, Yang Y, Peng J, Kirkpatrick DS. Characterizing ubiquitination sites by peptide-based immunoaffinity enrichment. *Mol Cell Proteomics.* 2012; 11:1529–1540. [PubMed: 22729469]
14. Iesmantavicius V, Weinert BT, Choudhary C. Convergence of ubiquitylation and phosphorylation signaling in rapamycin-treated yeast cells. *Mol Cell Proteomics.* 2014; 13:1979–1992. [PubMed: 24961812]
15. Kim DY, Scalf M, Smith LM, Vierstra RD. Advanced proteomic analyses yield a deep catalog of ubiquitylation targets in Arabidopsis. *Plant Cell.* 2013; 25:1523–1540. [PubMed: 23667124]
16. Cannon J, Nakasone M, Fushman D, Fenselau C. Proteomic identification and analysis of K63-linked ubiquitin conjugates. *Anal Chem.* 2012; 84:10121–10128. [PubMed: 23101881]
17. Stes E, Laga M, Walton A, Samyn N, Timmerman E, De Smet I, Goormachtig S, Gevaert K. A COFRADIC protocol to study protein ubiquitination. *J Proteome Res.* 2014; 13:3107–3113. [PubMed: 24816145]
18. Cooper HJ, Heath JK, Jaffray E, Hay RT, Lam TT, Marshall AG. Identification of sites of ubiquitination in proteins: a Fourier transform ion cyclotron resonance mass spectrometry approach. *Anal Chem.* 2004; 76:6982–6988. [PubMed: 15571350]
19. Mikesch LM, Ueberheide B, Chi A, Coon JJ, Syka JEP, Shabanowitz J, Hunt DF. The utility of ETD mass spectrometry in proteomic analysis. *Biochim Biophys Acta.* 2006; 1764:1811–1822. [PubMed: 17118725]
20. Sobott F, Watt SJ, Smith J, Edelman MJ, Kramer HB, Kessler BM. Comparison of CID versus ETD based MS/MS fragmentation for the analysis of protein ubiquitination. *J Am Soc Mass Spectrom.* 2009; 20:1652–1659. [PubMed: 19523847]
21. Sarbu M, Ghiulai RM, Zamfir AD. Recent developments and applications of electron transfer dissociation mass spectrometry in proteomics. *Amino Acids.* 2014; 46:1625–1634. [PubMed: 24687149]
22. Jufvas A, Strålfors P, Vener AV. Histone variants and their post-translational modifications in primary human fat cells. *PLoS One.* 2011; 6:e15960. [PubMed: 21249133]
23. Boja ES, Fales HM. Overalkylation of a protein digest with iodoacetamide. *Anal Chem.* 2001; 73:3576–3582. [PubMed: 11510821]

24. Nielsen ML, Vermeulen M, Bonaldi T, Cox J, Moroder L, Mann M. Iodoacetamide-induced artifact mimics ubiquitination in mass spectrometry. *Nat Methods*. 2008; 5:459–460. [PubMed: 18511913]
25. Kalli A, Smith GT, Sweredoski MJ, Hess S. Evaluation and optimization of mass spectrometric settings during data-dependent acquisition mode: focus on LTQ-Orbitrap mass analyzers. *J Proteome Res*. 2013; 12:3071–3086. [PubMed: 23642296]
26. Swaney DL, McAlister GC, Coon JJ. Decision tree-driven tandem mass spectrometry for shotgun proteomics. *Nat Methods*. 2008; 5:959–964. [PubMed: 18931669]
27. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized ppb-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol*. 2008; 26:1367–1372. [PubMed: 19029910]
28. Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res*. 2011; 10:1794–1805. [PubMed: 21254760]
29. Neuhauser N, Michalski A, Cox J, Mann M. Expert system for computer-assisted annotation of MS/MS spectra. *Mol Cell Proteomics*. 2012; 11:1500–1509. [PubMed: 22888147]
30. Elias JE, Gygi SP. Target-decoy search strategy for mass spectrometry-based proteomics. *Methods Mol Biol*. 2010; 604:55–71. [PubMed: 20013364]
31. Krzywinski M, Altman N. Points of view: designing comparative experiments. *Nat Methods*. 2014; 11:597–598. [PubMed: 25019145]
32. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, Kern R, Tabb DL, Liebler DC, MacCoss MJ. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics*. 2010; 26:966–968. [PubMed: 20147306]
33. Sweredoski MJ, Smith GT, Kalli A, Graham RLJ, Hess S. LogViewer: A software tool to visualize quality control parameters to optimize proteomics experiments using Orbitrap and LTQ-FT mass spectrometers. *J Biomol Tech*. 2011; 22:122–126. [PubMed: 22131886]
34. Kirkpatrick DS, Denison C, Gygi SP. Weighing in on ubiquitin: the expanding role of mass-spectrometry-based proteomics. *Nat Cell Biol*. 2005; 7:750–757. [PubMed: 16056266]

Figure 1.

ETD analysis of a synthetic GG peptide (a) structure of synthetic GG peptide purchased from JPT. The sequence of the peptide is AMLK(GG)SEQNR, where (GG) represents the di-glycine moiety covalently linked via the carboxy terminus to the ϵ -amino group of lysine. Ubiquitin conjugation to target lysine residues prevents trypsin cleavage at the modified lysine. As a result, a tryptic fragment containing a ubiquitin remnant would yield a peptide with charges contributed by the true N-terminus, a second N-terminus at the end of the GG isopeptide, and the R-group charge of the C-terminal lysine or arginine. (b) Skyline analysis of the charge state distribution observed for the synthetic peptide. The 3+ precursor is depicted in red, whereas the 2+ precursor is depicted in blue. (c) Example ETD fragmentation spectrum observed for the synthetic GG peptide AMLK(gl)SEQNR. *Please note gl denotes the GG moiety in all MS/MS spectra

Figure 2.

ETD fragmentation improves the identification of GG sites in a ubiquitylome analysis. (a) Comparison of CID, HCD, all ion ETD, and DT methods in the identification of unique GG sites. (b) Difference between the average number of sites observed in ETD, DT, and HCD in comparison to CID and difference measured between DT and ETD, along with respective *P*-values. Five technical replicates of each method were used in this analysis [31]. (c) Venn diagrams demonstrating the overlap in unique GG site identifications. Conventional CID and HCD were directly compared with ETD or DT. Additionally, ETD and DT methods were compared with each other. A unique GG site was included as long as it was observed in any of the five replicates. There were no additional criteria beyond the 1% false discovery rate threshold for proteins, peptides, and modified sites for selection

Figure 3.

Comparison of charge state distribution between an enriched and unenriched yeast sample. GG peptides enriched from a yeast whole cell lysate using the K-ε-GG peptide-specific antibody were analyzed by five technical replicates each of CID, HCD, ETD, and DT. The charge state distribution of the GG peptides identified were plotted based on the charge states observed for the parent ions and compared with the charge states of all peptides observed in the analysis of an aliquot of the yeast lysate used to perform the enrichment. Values are expressed as percentage of total peptides identified

Figure 4.

All peptide comparison. The average number of all peptides, GG peptide, plus non-GG peptides, identified by each fragmentation was compared. The difference between the median values is plotted and the *P*-value is provided. ETD was compared with CID, DT was compared with CID and ETD, respectively [31]

Figure 5. Optimization of the mass range for GG peptide analysis. GG peptides identified in all analyses were binned based on the m/z value of the parent ion. Bin widths were 100 m/z . We did not detect GG peptides at a mass range higher than 1200 and propose narrowing the mass range in the analysis of GG peptide enriched samples to 300–1200 m/z .

Table 1

Mass Spectrometry Settings for ETD and DT Analysis

Parameters	ETD analysis values	DT analysis values
Activation type	ETD	CID/ETD
MS mass range	300–1200 <i>m/z</i>	300–1200 <i>m/z</i>
Minimum signal required for MS/MS	500	500
Number of microscans	1	1
Number of data-dependent MS/MS	15	20
FT master scan preview	Not enabled	Not enabled
Automatic gain control (AGC) Target Value for MS	1×10^6	1×10^6
AGC Target value for MS/MS	5×10^3	5×10^3
Maximum ion injection time for MS/MS	50 ms	50 ms
Supplemental activation	Enabled	Enabled
Dynamic exclusion duration	60 s	60 s
Activation time	100 ms	10 ms
ETD reaction time	–	100 ms
Use decision tree or other parameters	Not enabled	Enabled
Default charge state parameter	2+	CID = 2+ (300–1200 <i>m/z</i>)/brETD 3+ (300–1200 <i>m/z</i>)