Unexpected Trypsin Cleavage at Ubiquitinated Lysines

Meghan C. Burke,† Yan Wang,‡ Amanda E. Lee,† Emma Kimm Dixon,† Carlos A. Castaneda,† David Fushman,† and Catherine Fenselau*†,‡

†Department of Chemistry & Biochemistry and ‡Proteomics Core Facility, University of Maryland, College Park, Maryland 20742, United States

Supporting Information

ABSTRACT: Unexpected tryptic cleavage has been characterized at modified K48 residues in polyubiquitins. In particular, the tryptic products of all seven of the lysine-linked dimers of ubiquitin and of three trimers—linear Ub−48Ub−49Ub, linear Ub−63Ub−64Ub, and the branched trimer Ub3−6,48Ub—have been analyzed. In addition to the peptide products expected under commonly used tryptic conditions, we observe that peptides are formed with an unexpected ε-glycinylglycinyl-Lys carboxyl terminus when the site of linkage is Lys48. Trypsin from three different commercial sources exhibited this aberration. Initial cleavage at R74 is proposed in a distal ubiquitin ε-glycinylglycinyl-Lys tag on lysine at the site of conjugation.

Proteolysis by trypsin is commonly used in proteomic workflows and other biochemical investigations. It is solidly established that the enzyme cleaves proteins and peptides on the carboxyl side of lysine and arginine residues1,2 employing a well-understood specific binding mechanism that recognizes the basic side chains of those amino acids.3 It is generally observed that derivatized lysine and arginine residues are not substrates for the enzyme,1,2,6 and this behavior has been used to good advantage to recognize and localize modifications on Lys and Arg in histones and other proteins.

Ubiquitin is enzymatically conjugated with substrate proteins and other ubiquitin moieties through isopeptide bonds between its terminal Gly76 and ε-amino groups of lysine and arginine residues (Figure 1). The extent and nature of such conjugations have profound effects on protein function and disposition in the cell. Trypsin can cleave all the potential sites in these polypeptides in an exhaustive incubation; however, it cleaves ubiquitin at Arg74 very rapidly,7 leaving Gly75 and Gly76 residues on the modified lysine of the substrate peptide.8−11 Likewise, in the case of polyubiquitins, rapid tryptic cleavage of a distal ubiquitin at Arg74 leaves an ε-glycinylglycinyl tag on lysine at the site of conjugation in the proximal ubiquitin.12−13 This glycinylglycine tag is susceptible to analysis by HPLC, tandem mass spectrometry, and bioinformatics, and its formation and localization have formed the basis for much of our understanding of the biochemistry of ubiquitination.14−16 Neither ubiquitinated lysine nor ε-glycinylglycinyl-Lys is expected to be cleaved by trypsin. However, in a recent study of ubiquitinated conjugates in exosomes shed by myeloid-derived suppressor cells, 15 of 65 ε-glycinylglycinyl-peptide products had been formed by tryptic cleavage directly at the derivatized lysine residues.17 These included products from polyubiquitins as well as protein conjugates. Other laboratories have also recently reported peptides identified from tryptic digests of ubiquitin polymers and conjugates that are terminated in ε-glycinylglycinyl-Lys.10,14,15 The experiments reported here were designed to confirm rigorously or not that this irregular tryptic cleavage occurs in polyubiquitins and, by implication, in proteins conjugated with ubiquitins. Unexpectedly, we have observed that the irregular cleavage occurs in polyubiquitins specifically at linkages located at Lys48.

EXPERIMENTAL SECTION

Polyubiquitins. All seven of the lysine-linked ubiquitin dimers were available for this study, as well as two unbranched trimers, Ub−48Ub−49Ub and Ub−63Ub−64Ub, and the branched trimer Ub3−6,48Ub. Ubiquitin dimers were assembled either enzymatically (K48, K63) or chemically (K6, K11, K27, K29, K33) using recombinant ubiquitin monomers, as detailed in ref 16. Ubiquitin trimers were assembled enzymatically using linkage-specific ubiquitin-conjugating E2 enzymes.

Tryptic Digestion. All tryptic digestions were carried out under commonly used conditions: 16 h at 37 °C, in 50 mM ammonium bicarbonate, pH 7.8, using a 1:50 (ubiquitin dimers) or 1:20 (ubiquitin trimers) ratio of enzyme to substrate. Briefer incubation times were implemented in kinetic studies. Trypsin Gold, mass spectrometry grade porcine trypsin from Promega, was used in the majority of the 16 h digestions and in kinetic studies. However, TPCK (tosylphenylalanyl chloromethyl ketone)-immobilized bovine pancreatic trypsin

Received: February 9, 2015
Accepted: July 16, 2015
Published: July 16, 2015
Results and Discussion

Examination of tryptic products from more than 30 experiments indicates that digestion is exhaustive under the conditions used for the 16 h proteolysis with each of the three samples of trypsin. The tryptic peptide LIFAGKGG-

Figure 1. Sequence of the ubiquitin dimer Ub−64Ub. Top: The peptide product that results from the expected formation of the glycylglycylglycinyl tag from the distal chain and conventional cleavages at Arg42 and Arg54 in the proximal chain is highlighted. Bottom: The peptide product that results from formation of the glycylglycylglycinyl tag from the distal chain and cleavage at Arg42 and the unexpected cleavage at Lys48 in the proximal chain is highlighted. (In a ubiquitin—ubiquitin conjugate, distal refers to the ubiquitin unit that is conjugated through its C-terminal Gly76 to a lysine on another ubiquitin, referred to as proximal. In a branched triubiquitin such as [Ub]6−64Ub, two distal ubiquitins are attached to two different lysines, in this case Lys6 and Lys48, on the same proximal ubiquitin.)
identified in which the carboxyl terminus carries the glycineglycineyl tag (Figure 2B). The third trimer offers a Lys48 modification and also one at Lys6. Digestion with Trypsin Gold produced the expected peptides with missed cleavages at glycineglycineyl-Lys6 and -Lys48. In addition, a third product was identified, on the basis of accurate mass determination and tandem mass spectra, which terminates in glycineglycineyl-Lys48 (Figure 2C). Overall, the digestion products detected from these trimers further support the conclusion that proteolysis occurs uniquely at modified Lys48, and not at Lys residues modified at other positions in polyubiquitins.

These observations introduce several questions: What is the substrate for trypsin that produces the unexpected cleavage? Why does the aberrant cleavage take place selectively at substituted K48 in polyubiquitins?

A kinetic study presented in Supplemental Figure 1 (Supporting Information) shows rapid tryptic cleavage at R74 through the first 90 min of digestion of Ub−48Ub and Ub−63Ub dimers and the absence of detectable cleavage at K48
and K63, respectively. This is consistent with earlier reports.\textsuperscript{7,12,13} We propose that peptides carrying glycinylglycine-lysine are bound by trypsin as substrates for subsequent cleavage to produce the unexpected product (Figure 1). This binding would occur in competition with normal substrates. Because the glycinylglycine fragment carries a positive charge at its N-terminus and contains no bulky groups, it is bound in the recognition pocket by which trypsin normally recognizes Lys and Arg side chains. This surrogate binding may not position the polypeptide chain correctly to allow cleavage to occur at the normal rate. The binding pocket of trypsin is 10–12 Å deep, on the basis of crystal structures.\textsuperscript{18,19} On the basis of a crystal structure available for the K48-linked ubiquitin dimer,\textsuperscript{20} the length of the extended lysine side chain modified by glycinylglycine is 10.9 Å. It is relevant that S-(β-1-aminoethyl)cysteine, a side chain that also presents a positive terminal charge, was recognized as a substrate for trypsin some years ago.\textsuperscript{21}

The second question is more challenging. If the substrate for the unexpected cleavage is not ubiquitin-modified Lys\textsubscript{48}, but rather glycinylglycinyl-Lys\textsubscript{48}-Ub, then we need to ask why trypsin binds and cleaves glycinylglycinyl-Lys at position 48 of diubiquitin, but does not bind and cleave glycinylglycinyl-Lys at other positions. In an initial consideration of the primary

![Figure 3. Extracted ion chromatograms of \(m/z\) 762.5, the molecular ion of the peptide formed by tryptic cleavage at glycinylglycine-modified K48, and \(m/z\) 1460.8, the molecular ion of the peptide formed by tryptic cleavage at K54 with glycinylglycine attached at K48.](image)

### Table 1. Glycinylglycinyl-Peptides Observed\textsuperscript{a} from Proteolysis of Ubiquitin Dimers Linked at Lys\textsubscript{63}, Lys\textsubscript{48}, Lys\textsubscript{33}, Lys\textsubscript{29}, Lys\textsubscript{27}, Lys\textsubscript{11}, and Lys\textsubscript{6}

| Dimer Linkage | Peptide Sequence | Precursor Mass (Da) | Obsd | Peptide Sequence | Precursor Mass (Da) | Obsd |
|---------------|------------------|---------------------|------|------------------|---------------------|------|
| K63 Ub       | TLSDYNIQKGGESTLHLVLR | 2244.19             | yes  | TLSDYNIQKGG      | 1195.59             | no   |
| K48 Ub       | LIFAGKGGQLEDGR    | 1460.78             | yes  | LIFAGKGG         | 762.45              | yes  |
| K33 Ub       | IQDKGGEGIPPDDQQR  | 1637.82             | yes  | IQDKGG           | 617.32              | no   |
| K29 Ub       | AKGGQLQDK         | 816.46              | yes  | TITLVEPSDTIENVKGG | 2101.10             | no   |
| K27 Ub       | TITLVEPSDTIENVKGGAK | 2101.10           | yes  | TITLVEPSDTIENVKGG | 1901.97             | no   |
| K11 Ub       | TLTVKGGAKTLTVEPSDTIENVK | 2402.27        | yes  | TLTVKGG          | 633.36              | no   |
| K6 Ub        | MQIFVKGGTLTVEPSDTIENVK | 1379.77        | yes  | MQIFVKGG         | 879.48              | no   |

\(\text{"Yes" indicates that the peptide was observed; "no" indicates that the peptide was not observed.}

### Table 2. Glycinylglycinyl-Peptides Observed\textsuperscript{a} from Proteolysis of Ubiquitin Trimers Ub\textsubscript{−63}Ub−48Ub, Ub−63Ub−48Ub, and [Ub\textsubscript{−6,48}]\textsubscript{−6,48}Ub

| Trimer Linkage | Peptide Sequence | Precursor Mass (Da) | Obsd | Peptide Sequence | Precursor Mass (Da) | Obsd |
|----------------|------------------|---------------------|------|------------------|---------------------|------|
| Ub−63Ub−48Ub  | TLSDYNIQKGGESTLHLVLR | 2244.19             | yes  | TLSDYNIQKGG      | 1195.59             | no   |
| Ub−63Ub−48Ub  | LIFAGKGGQLEDGR    | 1460.78             | yes  | LIFAGKGG         | 762.45              | yes  |
| [Ub\textsubscript{−6,48}]−6,48Ub | MQIFVKGGTLTVEPSDTIENVK | 1379.77       | yes  | MQIFVKGG         | 879.48              | no   |

\(\text{"Yes" indicates that the peptide was observed; "no" indicates that the peptide was not observed.}
sequence, we used the MEROPS tool\(^\text{22}\) to correlate cleavage frequencies experimentally associated with the three residues N-terminal and four residues C-terminal to the cleavage sites in the various ubiquitin dimers. Although the sequence around K48 scored well, considerations of primary sequence were not definitive and suggest that tertiary structure and topology should be considered. The compact and dynamic structures of all the linkage types of diubiquitin have been modeled recently and found to be distinctive.\(^\text{23,24}\) We speculate that trypsin has a specific interaction with the surface of ubiquitin around Lys48 and that this interaction promotes binding and/or cleavage of glycine-glycinyl-Lys48 as a substrate. We are currently exploring the possibility of such an interaction using NMR.

In summary, we have confirmed that trypsin does cleave some polyubiquitins to produce peptides that carry glycine-glycinyl-lysine at the carboxyl terminus. We propose that rapid cleavage of a distal ubiquitin chain occurs first in these cases, and that the resulting glycine-glycinyl-lysine side chain is bound as a surrogate to arginine (or lysine) in trypsin’s active site. Subsequent proteolysis may occur if the polypeptide chain is suitably positioned on the enzyme. This hypothesis is extended to proteins conjugated with ubiquitin, consistent with reports of unexpected formation of peptides with glycine-glycinyl-lysine carboxyl termini.\(^\text{10,11,14,15}\) These observations suggest that qualitative and quantitative studies that rely on analysis only of peptides carrying internal glycine-glycinyl tags may require reinterpretation.

**ASSOCIATED CONTENT**

Supporting Information
Table listing the peptide identifications from the Ub\(^{-}\text{48}\)Ub dimer (Excel) and figure showing the peak area as a function of time for the digestion of Ub\(^{-}\text{48}\)Ub and Ub\(^{-}\text{53}\)Ub. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.analchem.5b01960.

**AUTHOR INFORMATION**

Corresponding Author
*E-mail: Fenselau@umd.edu. Phone: 301-405-8616.

Notes
The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

This research was supported by grants from the National Institutes of Health (GM021248 and GM065334).

**REFERENCES**

1. Olsen, J. V.; Ong, S.-E.; Mann, M. *Mol. Cell. Proteomics* 2004, 3, 608–614.
2. Rodriguez, J.; Gupta, N.; Smith, R. D.; Pevzner, P. A. *J. Prot. Res.* 2008, 7, 300–305.
3. Sprang, S.; Standing, T.; Fletterick, R. J.; Stroud, R. M.; Finer-Moore, J.; Xuong, N.-H.; Hamlin, R.; Rutter, W. J.; Craik, C. S. *Science* 1987, 237, 905–909.
4. Craik, C. S.; Rocznik, S.; Largman, C.; Rutter, W. J. *Science* 1987, 237, 909–913.
5. Corey, D. R.; Craik, C. S. *J. Am. Chem. Soc.* 1992, 114, 1784–1790.
6. Ong, S.-E.; Mittler, G.; Mann, M. *Nat. Methods* 2004, 1, 119–126.
7. Wilkinson, K. D.; Audhya, T. *K. J. Biol. Chem.* 1981, 256, 9235–9241.
8. Van Nocker, S.; Vierstra, R. D. *Proc. Natl. Acad. Sci. U. S. A.* 1991, 88, 10297–10301.
9. Peng, J.; Schwartz, D.; Elias, J. E.; Thoreen, C. C.; Cheng, D.; Marsischky, G.; Roelofs, J.; Finley, D.; Gygi, S. P. *Nat. Biotechnol.* 2003, 21, 921–926.
10. Denis, N. J.; Vasilescu, J.; Lambert, J. P.; Smith, J. C.; Figeys, D. *Proteomics* 2007, 7, 868–874.
11. Burke, M. C.; Oei, M. S.; Edwards, N. J.; Ostrand-Rosenberg, S.; Fenselau, C. *J. Prot. Res.* 2014, 13, 5965–5972.
12. Xu, P.; Peng, J. *Anal. Chem.* 2008, 80, 3438–3444.
13. Valkevich, E. M.; Sanchez, N.; Ge, Y.; Strieter, E. R. *Biochemistry 2014*, 53, 4979–4989.
14. Danielsen, J. R.; Sylvestersen, K. B.; Bekker-Jensen, S.; Szklarczyk, D.; Poulsen, J. W.; Horn, H.; Jensen, L. J.; Mailand, N.; Nielsen, M. *L. Mol. Cell. Proteomics* 2011, 10, M110.003590.
15. Xu, G.; Paige, J. S.; Jaffrey, S. R. *Nat. Biotechnol.* 2010, 28, 868–873.
16. Lee, A. E.; Castaneda, C.; Wang, Y.; Fushman, D.; Fenselau, C. *J. Mass Spectrom.* 2014, 49, 1272–1278.
17. Walmsley, S. J.; Rudnick, P. A.; Liang, Y.; Dong, Q.; Stein, S. E.; Nesvizhskii, A. I. *J. Prot. Res.* 2013, 12, 5666–5680.
18. Steitz, T. A.; Hendekson, R.; Blow, D. M. *J. Mol. Biol.* 1969, 46, 337–348.
19. Liebschner, D.; Dauter, M.; Brzuszkiewicz, A.; Dauter, Z. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 2013, 69, 1447–1462.
20. Lai, M.-Y.; Zhang, N.; LaRonde-LeBlanc, N.; Fushman, D. *Biochim. Biophys. Acta, Mol. Cell Res.* 2012, 1823, 2046–2056.
21. Lindley, H. *Nature* 1956, 176, 647–648.
22. Rawlings, N. D.; Waller, M.; Barrett, A. J.; Bateman. *Nucleic Acids Res.* 2014, 42, D503–D509.
23. Fushman, D.; Walker, O. *J. Mol. Biol.* 2010, 395, 803–814.
24. Wang, Y.; Tang, C.; Wang, E.; Wang, J. *PLoS Comput. Biol.* 2014, 10, e1003691.