Middle-Down Mass Spectrometry Enables Characterization of Branched Ubiquitin Chains

Ellen M. Valkevich,† Nicholas A. Sanchez,† Ying Ge,‡ and Eric R. Strieter*†

†Department of Chemistry, University of Wisconsin—Madison, 1101 University Avenue, Madison, Wisconsin 53706, United States
‡Department of Cell and Regenerative Biology, School of Medicine and Public Health, University of Wisconsin—Madison, 1300 University Avenue, Madison, Wisconsin 53706, United States

ABSTRACT: Protein ubiquitylation, one of the most prevalent post-translational modifications in eukaryotes, is involved in regulating nearly every cellular signaling pathway. The vast functional range of ubiquitylation has largely been attributed to the formation of a diverse array of polymeric ubiquitin (polyUb) chains. Methods that enable the characterization of these diverse chains are necessary to fully understand how differences in structure relate to function. Here, we describe a method for the detection of enzymatically derived branched polyUb conjugates in which a single Ub subunit is modified by two Ub molecules at distinct lysine residues. Using a middle-down mass spectrometry approach in which restricted trypsin-mediated digestion is coupled with mass spectrometric analysis, we characterize the polyUb chains produced by bacterial effector E3 ligases NleL (non-Lee-encoded effector ligase from enterohemorrhagic Escherichia coli O157:H7) and IpaH9.8 (from Shigella flexneri). Because Ub is largely intact after minimal trypsinolysis, multiple modifications on a single Ub moiety can be detected. Analysis of NleL- and IpaH9.8-derived polyUb chains reveals branch points are present in approximately 10% of the overall chain population. When unanchored, well-defined polyUb chains are added to reaction mixtures containing NleL, longer chains are more likely to be modified internally, forming branch points rather than extending from the end of the chain. These results suggest that middle-down mass spectrometry can be used to assess the extent to which branched polyUb chains are formed by various enzymatic systems and potentially evaluate the presence of these atypical conjugates in cell and tissue extracts.

Through covalent attachment to intracellular proteins, a process termed protein ubiquitylation, the small protein ubiquitin (Ub) regulates a wide range of biological processes. Ub is tethered to proteins via the action of three proteins: E1 Ub activating enzymes, E2 Ub conjugating enzymes, and E3 Ub ligating enzymes (Figure 1A).7,8 This enzymatic cascade results in the formation of an isopeptide linkage between the C-terminus of Ub and an ε-amino group of a substrate lysine residue. Analogous to protein glycosylation, once Ub is anchored to a substrate protein additional rounds of conjugation afford polymeric Ub (polyUb) chains. A variety of chains can be assembled due to the presence of seven Ub lysine residues (K6, K11, K27, K29, K33, K48, and K63) and an amino terminus (M1) (Figure 1B). The prevailing view is that distinct polyUb chains govern specific biological pathways. This supposition is based, in large part, on the disparate activities of single-linkage (homotypic) K48- and K63-linked polyUb chains. K48-linked Ub chains with a minimum of four subunits endow cells with the ability to remove intracellular proteins by acting as the principal signal for proteasomal degradation.9−11 By contrast, K63-linked polyUb chains serve as nondegradative signals during the DNA damage response and cytokine signaling.12,13 These functional discrepancies have motivated efforts to uncover the roles of other polyUb chains, particularly because proteomics studies have revealed all eight Ub−Ub linkages are present in eukaryotic cells.14 While it has become evident that the abundance of non-K48 and K63 linkages increases under certain cellular conditions (e.g., M1, K6, and K11),15−17 the precise function of these conjugates is unclear. Moreover, the exact topology of chains (e.g., whether a chain is linear or branched) that govern a particular cellular pathway has been a mystery.

Connecting chain topology to a particular biological function presents a significant analytical challenge. The three most common methods for analyzing polyUb chains include expression of Ub lysine-to-arginine (K-to-R) variants in different cell lines, linkage-specific antibodies, and mass spectrometry (MS).18 Use of Ub K-to-R variants blocks chain extension through specific sites and therefore prevents downstream events associated with a particular polyUb chain. The implementation of this approach has led to seminal discoveries related to K48 and K63 linkages.10,12 However, Ub...
K-to-R variants are overexpressed in living cells, and the presence of wild-type Ub often complicates analysis. Furthermore, definitive conclusions cannot be drawn with regard to the formation of linear or branched heterotypic chains, and particular K-to-R substitutions may affect the surface of Ub and bias polyUb assembly. Although linkage-specific antibodies (four of which are available, M1, K11, K48, and K63) provide an opportunity to explore the impact of chain linkage on the fate of endogenous polyUb-modified proteins, it is difficult to differentiate between homotypic and heterotypic chains.19,20

MS approaches have also been instrumental in characterizing polyUb chains. The most common strategy is to use a bottom-up approach in which Ub conjugates are first digested extensively by trypsin and then analyzed by MS. Modified lysine residues of a particular protein, including Ub itself, are then identified by the presence of a diglycine (GG) motif appended to the ε-amino group (trypsin cleaves between R74 and G75 of Ub, leaving behind the GG motif). While bottom-up MS enables characterization of the linkages between two Ub molecules, it is difficult to assess chain length and topology because the connectivity of Ub fragments harboring branch points is destroyed during digestion. To address this, there has been interest in using other approaches that examine intact proteins such as middle-down MS. Middle-down MS combines the benefits of both bottom-up and top-down approaches by exploiting minimal protease digestion of protein samples and the ability to detect multiple post-translational modifications on a single polypeptide chain. However, this approach has not been applied to the analysis of Ub chain topology.

Here we demonstrate the utility of middle-down MS in the characterization of branched polyUb chains produced by bacterial E3 ligases. Several pathogenic bacteria deliver proteins, termed effectors, into host cells to undermine the defense response. Because the ubiquitylation network plays a major role in the immune response, it is one of the primary targets of these effector proteins. Many effectors functionally mimic eukaryotic E3 ligases and catalyze the assembly of polyUb chains on a distinct subset of host proteins. One particular example is the E3 ligase NleL (non-Lee-encoded effector ligase) from enterohemorrhagic Escherichia coli (EHEC) O157:H7. NleL is important for modulating the actin cytoskeleton of the host cell and has recently been shown to build heterotypic polyUb chains bearing K6 and K48 linkages in vitro. Because of the challenges associated with the assessment of chain topology, however, the extent to which NleL constructs polyUb chains bearing branch points is unknown. Therefore, we sought to employ middle-down MS to evaluate the formation of this underexplored chain topology.

**MATERIALS AND METHODS**

**Protein Expression and Purification.** Wild-type ubiquitin (Ub) and lysine-to-cysteine (K-to-C) Ub variants were
expressed in *E. coli* RosettaTM 2(DE3)pLysS cells (Novagen) and purified by perchloric acid precipitation, following a procedure adapted from ref 35. DNA encoding the human E1 Ub-activating enzyme was amplified from the HeLa cell cDNA library and cloned into pET24a(+). The UbcH5c(Ube2D3)-pET14a DNA construct was purchased from Addgene. The catalytic domain of IpaH9.8−545 was cloned into pET28a(+). Human E1, UBE2D3, and IpaH9.8 (*S. flexneri*) with six-histidine (His) tags were expressed and purified as previously described.36 A bacterial expression vector encoding glutathione S-transferase (GST)-tagged NleL170−782 was a gift from D. Y. W. Lin and J. Chen. Briefly, GST-tagged NleL was expressed in *E. coli* BL-21 cells grown in LB medium (OD600 of 0.06), induced with IPTG (0.1 mM), and grown at 16 °C (16 h). GST-NleL was then purified by glutathione sepharose affinity chromatography. The GST tag was cleaved from the eluted protein with TEV protease (4 °C for 16 h; minimal N-terminal perturbation is imperative for chain synthesis activity with NleL), and the protein was further purified by cation exchange chromatography.

**Thiol−Ene Ub Chain Synthesis.** Homotypic Ub chains linked via non-native isopeptide bonds at position 6 or 48 were synthesized using thiol−ene coupling (TEC) chemistry as previously described.37,38

**Native PolyUb Chain Synthesis.** To solutions containing reaction buffer A [50 mM Tris-HCl (pH 7.4), 50 mM NaCl, 5 mM MgCl2, and 0.1 mM DTT] were added Ub (50 μM), E1 (150 nM), E2 (1 μM UBE2D3 or UBE2L3), and E3 (0.05−5 μM NleL or IpaH9.8). Reactions were then initiated using ATP (2 mM) and allowed to proceed at 37 °C. PolyUb chain formation was analyzed by sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE) and MS (see below).

**Chain Elongation Using Thiol−Ene-Derived Ub Substrates.** To each reaction mixture were added Ub oligomers

Figure 2. Minimal trypsinolysis of substrates modified with polyUb chains of different topologies. What differentiates the formation of linear and branched chains is the presence of a Ub1−74 derivative harboring a single GG motif (linear; GG-Ub1−74) or two GG motifs (branched; 2xGG-Ub1−74).

4981
derived from thiol–ene chemistry (50 μM), Ub (25 μM), E1 (150 nM), UBE2D3 (1 μM), and NleL (5 μM) in buffer A. ATP (2 mM) was then added, and polymerization was allowed to occur for 3 h at 37 °C. It is important to note the concentrations of Ub dimers and tetramers are based on the molecular weight of each chain. By contrast, the concentrations of heterogeneous mixtures of Ub oligomers were measured on the basis of the molecular weight of a single Ub molecule.

Minimal Trypsin Digestion of Ub Chains. After chain synthesis, reaction mixtures were concentrated and exchanged into water using Amicon spin filters [0.5 mL with a 3.5 kDa molecular weight cutoff (MWCO)]. The enzyme/chain mixture (30 μL or half of the total reaction mixture) was digested with trypsin (0.5 μg; Cal Biochem MS grade) in ammonium bicarbonate buffer at 37 °C for 6 h. Trypsin was deactivated with 10% acetic acid, and the resulting mixtures were dialyzed into water (Slide-A-lyzer MINI dialysis units, 3.5 kDa MWCO) to remove small peptides arising from conjugating enzymes.

Middle-Down Mass Spectrometry Analysis. Samples were dissolved in a water/acetonitrile/acetic acid (45:45:10) solution and injected into a 7T linear ion trap/Fourier transform ion cyclotron resonance (LTQ/FT-ICR) hybrid mass spectrometer (Thermo Scientific Inc., Bremen, Germany) equipped with an automated chip-based nanoESI source (Triversa NanoMate, Advion BioSciences, Ithaca, NY) as described previously. The resolving power of the FT-ICR mass analyzer was set at 10000. All FT-ICR spectra were processed with in-house software (MASH Suite) using a signal-to-noise threshold of 3 and a fit factor of 60% and then validated manually.

Electron Capture Dissociation (ECD) Analysis of Ub Chain Linkages. For tandem mass spectrometry (MS/MS) experiments using ECD, individual charge states of protein molecular ions were first isolated. Then, the ions were dissociated by ECD using 3–4% “electron energy” and a 70 ms duration with a 65 μs delay. All FT-ICR spectra were processed with the MASH software suite using a signal-to-noise threshold of 3 and a fit factor of 60% and then validated manually. The resulting mass lists were further assigned on the basis of the protein sequence of Ub with or without the diglycine (GG) modification at each lysine residue using tolerances of 10 and 20 ppm for precursor and fragment ions, respectively. All reported calculated (calc) and experimental (exptl) values correspond to the most abundant molecular weights.

RESULTS

Middle-Down Analysis of Branched PolyUb Chains. Under conditions in which Ub is completely folded, trypsin cleaves only the peptide bond between arginine 74 and glycine 75, thereby releasing the C-terminal diglycine (GG) motif. By exploiting the minimal trypsinolysis of Ub, we reasoned middle-down MS could be employed to unambiguously characterize the presence of branching within polyUb chains. The general concept is that minimal trypsinolysis furnishes Ub1−74 monomers still harboring a GG motif at any lysine previously engaged in the formation of a chain. Because chain branching involves the addition of multiple Ub molecules to a single Ub subunit, minimal trypsinolysis should then afford a Ub1−74 derivative harboring two or more GG motifs tethered to specific lysine residues. If there is chain branching, we would expect to observe at least three distinct species by MS: Ub1−74 arising from chain caps and/or unpolymerized Ub (calc, 8450.57 Da); Ub1−74 with a single GG motif (GG-Ub1−74; calc, 8564.62 Da), resulting from the linear portion of the chain; and Ub1−74 with two GG motifs (2xGG-Ub1−74; calc, 8678.66 Da), originating from the branch point of the chain (Figure 2).

To determine whether middle-down MS could be used to detect branched conjugates, we examined polyUb chain formation by the E3 ligase NleL from EHEC O157:H7. NleL is composed of an N-terminal unstructured region (residues 1–169), a pentapeptide repeat domain (residues 170–370), and a catalytic HECT-like domain (HECT; homologous to the C-terminus of E6AP) with a conserved cysteine residue (C753) at the C-terminus. Previous studies with NleL have shown that residues 170–782 comprise a catalytically active E3 ligase capable of assembling chains on itself through a process termed autoubiquitylation as well as building unanchored/free chains. In accordance, the addition of NleL170−782 to

![Figure 3. Middle-down MS analysis of NleL-catalyzed reactions. (A) Generation of unanchored and NleL-anchored polyUb chains using two different E2s: UBE2D3 and UBE2L3. (B) FT-ICR MS analysis of NleL-catalyzed reactions after minimal trypsinolysis with UBE2D3 as the E2. The spectra correspond to the Ub1−74 charge state. The top spectrum shows all three Ub1−74 species after restricted trypsin digestion. The middle spectrum displays the isolated GG-Ub1−74 M+ parent ion, and the bottom spectrum shows the isolated 2xGG-Ub1−74 M+ parent ion.](image-url)
chains of varying length that functionally mimic native Ub. FT-ICR MS shows that each Ub$_{1-74}$ species can be detected at all three NleL concentrations. (B) FT-ICR MS analysis of minimal trypsin digests of wild-type Ub (top) and NleL-catalyzed polyUb chain formation (bottom). GG-modified Ub$_{1-74}$ is only present in reaction mixtures with NleL. (C) Homotypic 6- and 48-linked polyUb chains of varying length were assembled using thiol–ene coupling and subsequently digested with trypsin. FT-ICR MS shows the presence of Ub$_{1-74}$ and GG-Ub$_{1-74}$, but not 2xGG-Ub$_{1-74}$, in the spectrum for both 6-linked (top) and 48-linked (bottom) polyUb chains. The red box indicates where the signal for 2xGG-Ub$_{1-74}$ should be.

Figure 4. Formation of branch points under different conditions. (A) Generation of 2xGG-Ub$_{1-74}$ and GG-Ub$_{1-74}$ as a function of NleL concentration. Reactions were performed with three different concentrations of NleL. FT-ICR MS shows that each Ub$_{1-74}$ species can be detected at all three NleL concentrations. (B) FT-ICR MS analysis of minimal trypsin digests of wild-type Ub (top) and NleL-catalyzed polyUb chain formation (bottom). GG-modified Ub$_{1-74}$ is only present in reaction mixtures with NleL. (C) Homotypic 6- and 48-linked polyUb chains of varying length were assembled using thiol–ene coupling and subsequently digested with trypsin. FT-ICR MS shows the presence of Ub$_{1-74}$ and GG-Ub$_{1-74}$, but not 2xGG-Ub$_{1-74}$, in the spectrum for both 6-linked (top) and 48-linked (bottom) polyUb chains. The red box indicates where the signal for 2xGG-Ub$_{1-74}$ should be.

Reaction mixtures containing human E1, the E2 UBE2D3 or UBE2L3, ATP, and Ub resulted in the formation of polyUb chains (Figure 3A). To characterize the topology of these chains, minimal trypsinolysis was performed in an ammonium bicarbonate buffer to maintain Ub in a native, folded state. Initial experiments focused on determining the amount of trypsin required to completely disassemble chains into Ub$_{1-74}$ fragments. Using a trypsin concentration of 12.5 μg/mL and an enzyme-to-substrate ratio of approximately 1:1, chains were completely digested within 5 h as judged by SDS–PAGE (Figure S1 of the Supporting Information). The resulting Ub$_{1-74}$ derivatives were then analyzed using FT-ICR MS. Although Ub species were detected, low-molecular weight peptides from other proteins in the reaction mixture dominated the spectra. This was not entirely unexpected considering that, unlike peptides, the concentration of intact proteins is typically diluted by isotopic distributions and the presence of multiple charge states. To eliminate signals arising from peptides, a postdigest dialysis step was incorporated into the workflow (Figure S2 of the Supporting Information). This led to the detection of all three Ub$_{1-74}$ species (Figure 3B).

Next, we sought to confirm that 2xGG-Ub$_{1-74}$ was generated from a branch point. We first tested whether there was a correlation between the concentration of NleL and the levels of 2xGG-Ub$_{1-74}$. Changing the concentration of NleL from 50 nM to 5 μM was commensurate with an increase in the peak intensity of both 2xGG-Ub$_{1-74}$ and GG-Ub$_{1-74}$ (Figure 4A), indicating the presence of these Ub species is directly related to the catalytic activity of NleL. Analysis of minimally digested wild-type Ub supported this conclusion, as signals corresponding to 2xGG-Ub$_{1-74}$ and GG-Ub$_{1-74}$ were absent in the MS spectra (Figure 4B). Because the 2xGG-Ub$_{1-74}$ signal could also reflect a Ub derivative in which the GG motif is not removed from the C-terminal, i.e., proximal, subunit of an unanchored polyUb chain but is cleaved from the adjoining Ub, we wanted to evaluate minimal digestion of different linear chains. Previously, our lab has shown that a bifunctional Ub derivative harboring both a cysteine in lieu of a specific lysine residue and a C-terminal allylamine adduct could be polymerized using thiol–ene coupling chemistry to furnish single-linkage polyUb chains of varying length that functionally mimic native oligomers.37,38 Thus, we assembled homotypic 6- and 48-linked polyUb chains using thiol–ene chemistry and subjected these polymers to restricted digestion conditions. FT-ICR analysis of the resulting Ub$_{1-74}$ derivatives showed that 2xGG-Ub$_{1-74}$ is absent (Figure 4C). These data indicate that within a complex mixture of homotypic polyUb chains the GG motif on the proximal subunit is always removed by trypsin. Collectively, these results demonstrate that middle-down MS can be used to characterize the presence of branched polyUb chains.

Identification of Modified Lysine Residues by ECD. To identify residues of 2xGG-Ub$_{1-74}$ modified with Ub GG motifs, we used ECD45 to induce protein fragmentation, resulting in the cleavage of N–Cα bonds and the formation of c- and z*-type ions. Analysis of the fragments then identifies the exact residues modified with GG. In the case of the 2xGG-Ub$_{1-74}$ species, the M$^8$ charge state was isolated and fragmented by ECD, resulting in 19 c ions and 25 z* ions with a total of 35 of 73 bond cleavages (Figure 5B). The fragmentation patterns around K6 and K48 unambiguously confirmed the presence of isopeptide linkages at these positions (Figure 5A and Figure 5S).
of the Supporting Information). ECD analysis was also performed on the $M^+$ charge state of GG-UB$_{1-74}$ and the fragmentation pattern showed the GG modification largely resides on K6 (Figure S4 of the Supporting Information). These results are consistent with previous reports in that NleL assembles K6-linked polymers with an efficiency greater than that with which it generates K48 linkages.

**Branching Depends on Chain Length.** With an established characterization method, we wanted to gain more insight into the formation of branched polyUb chains. We reasoned that because the Ub sequence remains largely the same among Ub$_{1-74}$ GG-UB$_{1-74}$ and 2xGG-UB$_{1-74}$, the effect of the GG motif on ionization efficiency should be negligible. As a result, FT-ICR MS could be used to measure the relative ratios of Ub$_{1-74}$ GG-UB$_{1-74}$ and 2xGG-UB$_{1-74}$ and monitor the formation of branch points over time. Indeed, a similar approach has been used to calculate the ratio of Ub$_{1-74}$ to GG-UB$_{1-74}$ to determine the length of homotypic polyUb chains. 

To ensure this strategy would be suitable for relative quantitation, signal ratios were measured over a range of charge states (Figure S6 of the Supporting Information) from Ub$^{+7}$ to Ub$^{+12}$; in each case, the ratios remained invariant.

Using ratios of different Ub$_{1-74}$ signals, the extent of chain branching was then examined. E3 ligases catalyze polyUb formation via a step-growth mechanism in which individual Ub subunits are sequentially added to the growing chain. According to this mechanism, the product distribution should be composed of low-molecular weight conjugates such as dimers, trimers, etc., at early time points. As the reaction proceeds, the distribution should shift to high-molecular weight species, increasing the number of lysine residues in each chain that can serve as acceptors for another Ub. Accordingly, the probability of building a branched chain depends on the length of the polyUb chain because the internal acceptor sites outnumber the sites at the end of the chain as the chain gets longer. To test this, the population of branch points was measured at different times during chain formation. At each time point, signal intensities were measured and divided by the total population to arrive at the amount of individual Ub$_{1-74}$ derivatives. With UBE2D3 as the E2 partner, approximately 3% of the total chain population (GG-UB$_{1-74}$ and 2xGG-UB$_{1-74}$) contains a branch point at the onset of the reaction (Figure 6A). As the reaction progresses, the amount of branching doubles, reaching 7% of the total chain population by 4 h (Figure 6A). Additional branch points were not observed after the reaction had been allowed to proceed overnight. SDS–PAGE analysis showed that the polyUb chain population is largely composed of dimers, trimers, and tetramers at 0.5 h (Figure 3A). However, by 4 h, a significant proportion of the chains exist as high-molecular weight species. Using UBE2L3 as the E2, the formation of polyUb chains appeared to be faster during the initial phases of the reaction than polymerizations conducted with UBE2D3 (Figure 6B). For instance, at 0.5 h, ~30% of the Ub$_{1-74}$ derivatives are composed of GG-UB$_{1-74}$ and 2xGG-UB$_{1-74}$ compared to only 20% with UBE2D3 (Figure 6C). There is also a slight increase in the total amount of branch points with UBE2L3, as 2xGG-UB$_{1-74}$ represents 10% of the chains by 4 h. These results suggest a combination of UBE2L3-NleL affords a chain assembly complex more active than that of UBE2D3-NleL. Consistent with this notion, SDS–PAGE analysis showed that with UBE2L3 dimers and trimers are almost completely consumed after 4 h, whereas with UBE2D3 short oligomers persist at longer incubation times (Figure 3A).

To provide additional support for a model in which branching depends on chain length, we assessed the ability of NleL to extend unanchored polyUb chains. Because NleL is related to the HECT family of human E3 ligases and members of this family of enzymes are capable of binding and elongating free polyUb chains, we surmised NleL might also interact with and modify well-defined 6- and 48-linked oligomers. To evaluate this possibility, we again employed homotypic polyUb chains derived from thioldiene coupling chemistry (Figure 7A). These chains contain cysteine residues in the distal subunits, which block extension in the form of homotypic chains and allow us to assess the ability of NleL to exclusively catalyze the formation of linear or branched heterotypic chains as a function of chain length. The thioldiene-derived chains are also composed of a proximal subunit bearing a C-terminal allylamine adduct, which prevents activation by E1 and subsequent transfer to E2 and NleL. Thus, thioldiene-derived chains can be modified only if (i) Ub is also present in the reaction mixture and NleL catalyzes the transfer of Ub to the...
Encouraged by the results with 48-linked dimers, we sought to evaluate 6-linked oligomers as substrates of NleL. First, we wanted to investigate whether the linkage influences the efficiency of chain extension. Comparing reactions with 48- and 6-linked trimers, we measured the ratio of branched product (II) to unmodified substrate (I) (Figure 8A). From these data, we observed that II is formed to a much greater extent with 6-linked chains than those bearing 48-linkages, suggesting NleL prefers the former as substrates. Next, we wanted to evaluate the extent of branching as a function of chain length. Starting with 6-linked dimers, we observed two new product peaks: one corresponding to a linear heterotypic chain (calcd, 8539.83 Da) and another commensurate with a branched chain (Figure 8B). With 6-linked trimers and tetramers, however, only 2xGG-Ub1−74 could be detected, suggesting linear chains are not formed and branching occurs by placing Ub on one or more of the internal subunits (Figure 8A,B). From these data, we conclude that (i) branching depends on the linkage of the synthetic chains or (ii) the allylamine moiety has been removed from the C-terminus using the yeast Ub C-terminal hydrolase Yuh1.51 As evidenced by a new peak indicative of a branch point (calcd, 8710.83 Da), NleL catalyzes the modification of allylamine-capped, 48-linked Ub dimers only in the presence of free Ub (Figure 7B). With allylamine removed, the addition of free Ub is not required for heterotypic chain formation, indicating that 48-linked dimers are shuttled from E1 to E2 and finally to the active site cysteine of NleL.

**Figure 7.** Extension of preformed 48-linked Ub dimers using NleL. (A) Reaction scheme depicting the modification of 48-linked dimers along with the subsequent trypsin digestion. (B) FT-ICR MS analysis of NleL-catalyzed extension reactions. The resulting branched chains are colored red. Preformed dimers retained the allylamine moiety from the thiol–ene reaction, or the allylamine was removed prior to chain extension using the C-terminal hydrolase Yuh1.
linear chain initially assembled by the E3 ligase and (ii) as the chain length increases, it is more favorable to install a branch point rather than extend the chain in a linear fashion.

**Formation of Branched Chains with Other Bacterial E3 Ligases.** Recently, another class of bacterial effector proteins that catalyze polyUb chain formation via a thioester intermediate was discovered. Unlike NleL, these proteins, termed NELs (novel E3 ligases), bear no sequence or structural similarity to eukaryotic HECT or RING (really interesting new gene) ligases. NELs contain an N-terminal leucine-rich repeat and a C-terminal $\alpha$-helical catalytic domain. Despite differences between bacterial HECT-like ligases and the NELs, we sought to determine whether a representative member of the NEL family, IpaH9.8, assembles branched polyUb chains. IpaH9.8 is produced by the pathogen *S. flexneri* to dampen the host inflammatory response during infection. Consistent with previous reports, when IpaH9.8 was added to a reaction mixture containing UBE2D3 and Ub, polyUb chain formation was observed by SDS–PAGE (Figure 9B). Chain formation, however, was negligible with UBE2L3 as the E2. Minimal trypsin digest of the resulting Ub conjugates then led to the detection of Ub$_{1-74}$, GG-Ub$_{1-74}$, and 2xGG-Ub$_{1-74}$ by FT-ICR MS (Figure 9A), and ECD analysis of GG-Ub$_{1-74}$ revealed that K48 is the predominant linkage (Figure S11 of the Supporting Information). Compared to reactions with NleL, branch points are formed to a much lesser extent, making it difficult to characterize the linkages in 2xGG-Ub$_{1-74}$ using ECD. To overcome this problem, we used a bottom-up approach. Complete trypsinolysis of the Ub conjugates followed by MS analysis using an Orbitrap instrument showed that both K6 and K48 linkages are present (Figure S12 of the Supporting Information), suggesting IpaH9.8 catalyzes the formation of the same branch points as NleL.

Because the efficiency with which IpaH9.8 assembles branched chains is lower than that of NleL, we sought to further investigate the relationship between chain extension and branching. The population of branch points (2xGG-Ub$_{1-74}$) was assessed at different time points and compared to the amount of linear polyUb (GG-Ub$_{1-74}$) (Figure 9C). With NleL, branch points are detected within the first 20% conversion of Ub into polyUb chains. By contrast, IpaH9.8 does not generate branch points until $\sim$50% of Ub has been transformed into polyUb chains. Analysis of IpaH9.8-catalyzed reactions by SDS–PAGE indicated that while chains form rapidly, they have molecular weights lower than the molecular weights of those produced by NleL. These results provide additional support for the conclusion that branching is directly related to chain length. Alternatively, because IpaH9.8 primarily builds chains bearing K48 linkages and our results with NleL show that it is difficult to generate branch points starting from preformed K48-linked chains, the low abundance of 2xGG-
Ub$_{1-74}$ could be a consequence of the linkage type within the linear chains.

DISCUSSION

Bottom-up MS methods have been instrumental in shaping our understanding of the repertoire of polyUb chain linkages in eukaryotic cells.\textsuperscript{14,21,56} Trypsinolysis of polyUb chains furnishes signature peptides bearing a lysine residue with a GG motif from the C-terminus of Ub. The seven tryptic peptide fragments generated from this process then inform on the sites of polyUb chain formation. Because the signature GG motifs reside in distinct Ub peptides, it remains a formidable challenge to assess chain length and topology, i.e., the degree to which branched chains are formed unless modifications occur on adjacent lysines, e.g., K6+K11, K27+K29, and K29+K33.\textsuperscript{22} Characterizing length and topology is important because both factors control the dynamics of biochemical pathways.\textsuperscript{4,5} For instance, chains with a minimum of four subunits provide an effective signal for proteasomal degradation,\textsuperscript{11} and differences in the intracellular trafficking of major histocompatibility complex class II (MHC II) in professional antigen-presenting cells are a consequence of differences in chain length.\textsuperscript{57} With regard to topology, chain branching through K11 and K63 has also been implicated in controlling the rate at which MHC I is internalized by endocytosis,\textsuperscript{58} and just recently, the anaphase-promoting complex (APC/C) was found to assemble branched chains containing K11 linkages to promote efficient degradation by the proteasome during prometaphase of the cell cycle.\textsuperscript{59}

In the work presented here, we exploited middle-down MS to investigate the formation of branched polyUb chains containing isopeptide linkages at nonadjacent lysines. Our studies focused on the bacterial E3 ligase NleL, as it has the ability not only to tether heterotypic chains to itself, but also to form free, unanchored polyUb chains. The results of these studies demonstrate that NleL constructs branched chains containing isopeptide linkages at K6 and K48 when polyUb formation is completely unrestricted by the presence of K-to-R substitutions. Experiments with well-defined polyUb chains as substrates for NleL showed that it is equally probable to extend a dimer in the form of a linear or branched heterotypic chain. However, when longer chains, such as tetramers, are used as substrates, branching becomes a more prevalent modification. The observation that another bacterial E3 ligase, unrelated to NleL, along with the metazoan APC/C\textsuperscript{59} also assembles branched conjugates suggests these atypical chains could be more widespread than currently appreciated.

In light of our results, we speculate that mixtures of linkages, which are often present in high-molecular weight Ub conjugates, could represent branched chains in both enzymatic assays and cellular extracts. For example, high-molecular weight K48-linked polyUb chains immunoprecipitated from mammalian cells using the K48 linkage-specific antibody typically contain K6, K11, and K63 linkages.\textsuperscript{60,61} Our results could also apply to the action of E4 Ub ligases as these enzymes collaborate with E1, E2, and E3 enzymes to catalyze the extension of polyUb chains.\textsuperscript{60} The prototype of this activity is the yeast enzyme Ufd2, which elongates existing K29-linked chains by forming K48 linkages.\textsuperscript{61} The initial Ufd2-dependent elongation step may occur by extending the chain from the end to form a linear heterotypic chain. Alternatively, Ufd2 could first build a branch point before continuing to catalyze the sequential addition of Ub through K48.

The distinct Ub topologies created by different chain linkages could also play an important role in the ability to construct branched conjugates. Although NleL generates both K6 and K48 linkages, we found that by using unanchored polyUb chains as substrates K6-linked chains could be elongated in the form of a branched conjugate, whereas K48-linked chains were modified to a much lesser extent. These results, along with observations that long (more than two subunits) K6-linked chains are constructed by NleL at a rate much faster than the rates of those bearing K48 linkages,\textsuperscript{33} suggest the structural ensemble of K6-linked chains is well-suited for NleL to maintain a persistent and productive interaction. By contrast, the conformations adopted by K48-linked chains may not promote high-affinity interactions with NleL, and as a result, processive chain formation does not occur. Although additional binding studies are required to test this hypothesis, we surmise that as K6-linked chains become longer and the number of available K48 residues increases within each chain, entropy becomes an important factor in targeting the placement of a Ub unit at an internal subunit rather than at the chain terminus. This model could pertain to other ligases. For example, the yeast HECT ligase Rsp5 prefers to assemble K63 linkages, but because chains can also be extended through K11, K33, and K48, there could be a preponderance of arborization within these oligomers.\textsuperscript{62} Middle-down MS will be instrumental in testing this model and further elucidating the biochemical details of E3 ligases, one of the most predominant classes of enzymes encoded by the human genome. Moreover, middle-down MS can ultimately be combined with affinity chromatography steps to uncover the extent to which branched chains form under different cellular conditions, thus providing unprecedented insight into how Ub chain topology influences function.

ASSOCIATED CONTENT

Supporting Information

Supplemental Figures S1–S12. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

*E-mail: strieter@chem.wisc.edu.

Funding

Financial support is provided by the University of Wisconsin—Madison, the Greater Milwaukee Shaw Scientist Program (awarded to E.R.S.), Susan G. Komen for the Cure (Grant KG110081 to E.R.S.), the Wisconsin Partnership Fund for the establishment of the University of Wisconsin Human Proteomics Program (Y.G.), and National Institutes of Health Grant R01HL096971 to Y.G.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Hershko, A., and Ciechanover, A. (1998) The ubiquitin system. Annu. Rev. Biochem. 67, 425−479.
(2) Mukhopadhyay, D., and Riezman, H. (2007) Proteasome-independent functions of ubiquitin in endocytosis and signaling. Science 315, 201−205.
(3) Vucic, D., Dixit, V. M., and Wertz, I. E. (2011) Ubiquitylation in apoptosis: A post-translational modification at the edge of life and death. Nat. Rev. Mol. Cell Biol. 12, 439−452.
The unexplored world of polyubiquitin beyond Lys48 and Lys63 linkages.

Biochemistry 13, 508–523.

(19) Newton, K., Matsumoto, M. L., Wertz, I. E., Kirkpatrick, D. S., Lill, J. R., Tan, J., Dugger, D., Gordon, N., Sidhu, S. S., Fellouse, F. A., Komueves, L., French, D. M., Ferrando, R. E., Lam, C., Compaan, D., Yu, C., Bosanac, I., Hymowitz, S. G., Kelley, R. F., and Dixit, V. M. (2008) Ubiquitin chain editing revealed by polyubiquitin linkage-specific antibodies. Cell 134, 668–678.

(20) Phu, L., Izrael-Thomas, A., Matsumoto, M. L., Bostos, D., Dynel, J. N., Fedorova, A. V., Bakalaris, C. E., Armott, D., Deshayes, K., Dixit, V. M., Kelley, R. F., Vucic, D., and Kirkpatrick, D. S. (2011) Improved quantitative mass spectrometry methods for characterizing complex ubiquitin signals. Mol. Cell Proteomics 10, M110.003756.

(21) Xu, P., and Peng, J. (2006) Diisecting the ubiquitin pathway by mass spectrometry. Biochim. Biophys. Acta 1764, 1940–1947.

(22) Kim, H. T., Kim, K. P., Lledias, F., Kisselev, A. F., Scaglione, K. M., Skowrya, D., Gygi, S. P., and Goldberg, A. L. (2007) Certain pairs of ubiquitin-conjugating enzymes (E2s) and ubiquitin-protein ligases (E3s) synthesize nondegradable forked ubiquitin chains containing all possible isopeptide linkages. J. Biol. Chem. 282, 17375–17386.

(23) Xu, P., and Peng, J. (2008) Characterization of polyubiquitin chain structure by middle-down mass spectrometry. Anal. Chem. 80, 3438–3444.

(24) Strachan, J., Roach, L., Sokratois, K., Tooth, D., Long, J., Garner, T. P., Searle, M. S., Oldham, N. J., and Layfield, R. (2012) Insights into the molecular composition of endogenous unanchored polyubiquitin chains. J. Proteome Res. 11, 1969–1980.

(25) Cannon, J. R., Edwards, N. J., and Fenselau, C. (2013) Mass-biased partitioning to enhance middle down proteomics analysis. J. Mass Spectrom. 48, 340–343.

(26) Chait, B. T. (2006) Chemistry. Mass spectrometry: Bottom-up or top-down? Science 314, 65–66.

(27) Forbes, A. J., Mazur, M. T., Patel, H. M., Walsh, C. T., and Kelleher, N. L. (2001) Toward efficient analysis of >70 kDa proteins with 100% sequence coverage. Proteomics 1, 927–933.

(28) Catherman, A. D., Skinner, O. S., and Kelleher, N. L. (2014) Top down proteomics: Facts and perspectives. Biochem. Biophys. Res. Commun. 445, 683–693.

(29) Hicks, S. W., and Galan, J. E. (2013) Exploitation of eukaryotic subcellular targeting mechanisms by bacterial effectors. Nat. Rev. Microbiol. 11, 316–326.

(30) Hicks, S. W., and Galan, J. E. (2010) Hijacking the host ubiquitin pathway: Structural strategies of bacterial E3 ubiquitin ligases. Curr. Opin. Microbiol. 13, 41–46.

(31) Lin, D. Y., Diao, J., Zhou, D., and Chen, J. (2011) Biochemical and structural studies of a HECT-like ubiquitin ligase from Escherichia coli O157:H7. J. Biol. Chem. 286, 441–449.

(32) Piscatelli, H., Kotkar, S. A., McBeef, M. E., Muthupalani, S., Schauer, D. B., Mandrell, R. E., Leong, J. M., and Zhou, D. (2011) The EHEC type III effector NleL is an E3 ubiquitin ligase that modulates pedestal formation. PLoS One 6, e19331.

(33) Hospenthal, M. K., Freund, S. M., and Komander, D. (2013) Assembly, analysis and architecture of atypical ubiquitin chains. Nat. Struct. Mol. Biol. 20, 555–565.

(34) Ashida, H., Kim, M., Schmidt-Supprian, M., Ma, A., Ogawa, M., and Sasakawa, C. (2010) A bacterial E3 ubiquitin ligase IpaH9.8 targets NEMO/IKKγ to dampen the host NF-κB-mediated inflammatory response. J. Biol. Cell 12, 66–73 (supplementary pages 61–69).

(35) Pickart, C. M., and Rassu, S. (2005) Controlled synthesis of polyubiquitin chains. Methods Enzymol. 399, 21–36.

(36) Shekhawat, S., Pham, G. H., Prabakaran, J., and Strierer, E. R. (2014) unpublished results.

(37) Trang, V. H., Valkevich, E. M., Minami, S., Chen, Y. C., Ge, Y., and Strierer, E. R. (2012) Nonenzymatic polymerization of ubiquitin: Single-step synthesis and isolation of discrete ubiquitin oligomers. Angew. Chem., Int. Ed. 51, 13085–13088.

(38) Valkevich, E. M., Genge, R. G., Sanchez, N. A., Chen, Y. C., Ge, Y., and Strierer, E. R. (2012) Forging isopeptide bonds using thiol-ene chemistry: Site-specific coupling of ubiquitin molecules for studying the activity of isopeptidases. J. Am. Chem. Soc. 134, 6916–6919.

(39) Ayaz-Guner, S., Zhang, J., Li, L., Walker, J. W., and Ge, Y. (2009) In vivo phosphorylation site mapping in mouse cardiac troponin I by high resolution top-down electron capture dissociation mass spectrometry: Ser22/23 are the only sites basally phosphorylated. J. Biol. Chem. 284, 38654–38664.

(40) Ge, Y., Rybakova, I. N., Xu, Q., and Moss, R. L. (2009) Top-down high-resolution mass spectrometry of cardiac myosin binding protein C revealed that truncation alters protein phosphorylation state. Proc. Natl. Acad. Sci. U.S.A. 106, 12658–12663.

(41) Zhang, J., Guy, M. J., Norman, H. S., Chen, Y. C., Xu, Q., Dong, X., Guner, H., Wang, S., Komoto, T., Young, K. H., Moss, R. L., and Ge, Y. (2011) Top-down quantitative proteomics identified phosphorylation of cardiac troponin I as a candidate biomarker for chronic heart failure. J. Proteome Res. 10, 4054–4065.

(42) Guner, H., Close, P. L., Cai, W., Zhang, H., Peng, Y., Gregorich, Z. R., and Ge, Y. (2014) MASH Suite: A user-friendly and versatile software interface for high-resolution mass spectrometry data.
interpretation and visualization. J. Am. Soc. Mass Spectrom. 25, 464–470.

(43) Wilkinson, K. D., and Audhya, T. K. (1981) Stimulation of ATP-dependent proteolysis requires ubiquitin with the COOH-terminal sequence Arg-Gly-Gly. J. Biol. Chem. 256, 9235–9241.

(44) Compton, P. D., Zamdberg, L., Thomas, P. M., and Kelleher, N. L. (2011) On the scalability and requirements of whole protein mass spectrometry. Anal. Chem. 83, 6868–6874.

(45) Zubarev, R. A., Kelleher, N. L., and McLafferty, F. W. (1998) Electron capture dissociation of multiply charged protein cations. A nonergodic process. J. Am. Chem. Soc. 120, 3265–3266.

(46) Pierce, N. W., Kleiger, G., Shan, S. O., and Deshaies, R. J. (2009) Detection of sequential polyubiquitylation on a millisecond timescale. Nature 462, 615–619.

(47) Kamadurai, H. B., Qiu, Y., Deng, A., Harrison, J. S., Macdonald, C., Actis, M., Rodrigues, P., Miller, D. J., Souphron, J., Lewis, S. M., Kurinov, I., Fuji, N., Hammel, M., Piper, R., Kuhlman, B., and Schulman, B. A. (2013) Mechanism of ubiquitin ligation and lysine prioritization by a HECT E3. eLife 2, e00828.

(48) Wang, M., Cheng, D., Peng, J., and Pickart, C. M. (2006) Molecular determinants of polyubiquitin linkage selection by an HECT ubiquitin ligase. EMBO J. 25, 1710–1719.

(49) Ronchi, V. P., Klein, J. M., and Haas, A. L. (2013) E6AP/UBE3A ubiquitin ligase harbors two E2∼ubiquitin binding sites. J. Biol. Chem. 288, 10349–10360.

(50) Johnston, S. C., Riddle, S. M., Cohen, R. E., and Hill, C. P. (1999) Structural basis for the specificity of ubiquitin C-terminal hydrolases. EMBO J. 18, 3877–3887.

(51) Li, W., Tu, D., Brunger, A. T., and Ye, Y. (2007) A ubiquitin ligase transfers preformed polyubiquitin chains from a conjugating enzyme to a substrate. Nature 446, 333–337.

(52) Rohde, J. R., Breitkreutz, A., Chenal, A., Sansonetti, P. J., and Parsot, C. (2007) Type III secretion effectors of the IpAH family are E3 ubiquitin ligases. Cell Host Microbe 1, 77–83.

(53) Singer, A. U., Rohde, J. R., Lam, R., Skarina, T., Kagan, O., Dileo, R., Chirgadze, N. Y., Cuff, M. E., Joachimiak, A., Tyers, M., Sansonetti, P. J., Parsot, C., and Savchenko, A. (2008) Structure of the Shigella T3SS effector IpaH defines a new class of E3 ubiquitin ligases. Nat. Struct. Mol. Biol. 15, 1293–1301.

(54) Zhu, Y., Li, H., Hu, L., Wang, J., Zhou, Y., Pang, Z., Liu, L., and Shao, F. (2008) Structure of a Shigella effector reveals a new class of ubiquitin ligases. Nat. Struct. Mol. Biol. 15, 1302–1308.

(55) Quezada, C. M., Hicks, S. W., Galan, J. E., and Stebbins, C. E. (2009) A family of Salmonella virulence factors functions as a distinct class of autoregulated E3 ubiquitin ligases. Proc. Natl. Acad. Sci. U.S.A. 106, 4864–4869.

(56) Chen, P. C., Na, C. H., and Peng, J. (2012) Quantitative proteomics to decipher ubiquitin signaling. Amino Acids 43, 1049–1060.

(57) Ma, J. K., Platt, M. Y., Eastham-Anderson, J., Shin, J. S., and Mellman, I. (2012) MHC class II distribution in dendritic cells and B cells is determined by ubiquitin chain length. Proc. Natl. Acad. Sci. U.S.A. 109, 8820–8827.

(58) Boname, J. M., Thomas, M., Stagg, H. R., Xu, P., Peng, J., and Lehner, P. J. (2010) Efficient internalization of MHC I requires lysine-11 and lysine-63 mixed linkage polyubiquitin chains. Traffic 11, 210–220.

(59) Meyer, H. J., and Rape, M. (2014) Enhanced protein degradation by branched ubiquitin chains. Cell 157, 910–921.

(60) Hoppe, T. (2005) Multiubiquitylation by E4 enzymes: ‘One size’ doesn’t fit all. Trends Biochem. Sci. 30, 183–187.

(61) Koegl, M., Hoppe, T., Schlenker, S., Ulrich, H. D., Mayer, T. U., and Jentsch, S. (1999) A novel ubiquitination factor, E4, is involved in multiubiquitin chain assembly. Cell 96, 635–644.

(62) Kim, H. C., and Hubbregtse, J. M. (2009) Polyubiquitination by HECT E3s and the determinants of chain type specificity. Mol. Cell. Biol. 29, 3307–3318.