



Published in final edited form as:

J Am Chem Soc. 2011 November 9; 133(44): 17855–17868. doi:10.1021/ja207220g.

Nonenzymatic assembly of natural polyubiquitin chains of any linkage composition and isotopic labeling scheme

Carlos Castañeda^a, Jia Liu^b, Apurva Chaturvedi^a, Urszula Nowicka^a, T. Ashton Cropp^b, and David Fushman^a

T. Ashton Cropp: tacropp@vcu.edu; David Fushman: fushman@umd.edu

^aDepartment of Chemistry and Biochemistry, Center for Biomolecular Structure and Organization, University of Maryland, College Park, Maryland, 20742, USA

^bDepartment of Chemistry, Virginia Commonwealth University, Richmond, VA 23284

Abstract

Polymeric chains made of a small protein ubiquitin act as molecular signals regulating a variety of cellular processes controlling essentially all aspects of eukaryotic biology. Uncovering the mechanisms that allow differently linked polyubiquitin chains to serve as distinct molecular signals requires the ability to make these chains with the native connectivity, defined length, linkage composition, and in sufficient quantities. This however has been a major impediment in the ubiquitin field. Here we present a robust, efficient, and widely accessible method for controlled iterative non-enzymatic assembly of polyubiquitin chains using recombinant ubiquitin monomers as the primary building blocks. This method uses silver-mediated condensation reaction between the C-terminal thioester of one ubiquitin and the ϵ -amine of a specific lysine on the other ubiquitin. We augment the non-enzymatic approaches developed recently by using removable orthogonal amine-protecting groups, Alloc and Boc. The use of bacterially expressed ubiquitins allows cost-effective isotopic enrichment of any individual monomer in the chain. We demonstrate that our method yields completely natural polyubiquitin chains (free of mutations and linked through native isopeptide bonds) of essentially any desired length, linkage composition, and isotopic labeling scheme, and in milligram quantities. Specifically, we successfully made Lys11-linked di-, tri-, and tetra-ubiquitins, Lys33-linked di-ubiquitin, and a mixed-linkage Lys33,Lys11-linked tri-ubiquitin. We also demonstrate the ability to obtain, by high-resolution NMR, residue-specific information on ubiquitin units at any desired position in such chains. This method opens up essentially endless possibilities for rigorous structural and functional studies of polyubiquitin signals.

Introduction

Polymeric chains composed of a small protein ubiquitin (Ub) function as molecular signals in numerous cellular processes in eukaryotes, including protein turnover, progression through the cell cycle, transcriptional activation, antigen processing, and vesicular trafficking of proteins (reviewed in ^{1–7}). Ub monomers in these chains are usually linked via an isopeptide bond between the C-terminal G76 of one Ub (referred to here as the distal Ub) and the ϵ -amino group of one of seven lysines of the other Ub (proximal). This post-translational modification is performed and tightly controlled by a series of enzymes, and the regiospecificity of the isopeptide linkage is defined by various Ub-conjugating enzymes, E2s.^{8,9} The same chemistry is involved in the attachment of a monomeric Ub (monoUb) or a

polyubiquitin (polyUb) chain (via its proximal Ub unit) to a lysine residue of the target protein.

The functional outcome of polyubiquitination of a target protein depends on the length of the polyUb tag and the lysine residue involved in the Ub-Ub linkage. For example, long ($n \geq 4$) polyUb chains linked via K48 serve as the principal signal targeting proteins for degradation by the 26S proteasome^{2,10,11}, whereas K63-linked chains act as regulatory rather than proteolytic signals in a variety of non-degradative processes.^{12–15} The biological roles and the recognition features of the polyUb chains linked through the other five lysines (K6, K11, K27, K29, and K33) or head-to-tail, as well as mixed-linkage or branched chains are poorly understood and are currently the focus of extensive research. A central question in ubiquitin biology concerns how the broad functional range of Ub signaling is achieved. Despite the wealth of information on the various processes controlled or regulated by polyubiquitination, the molecular mechanisms underlying the ability of different polyUb chains to act as distinct molecular signals for diverse cellular events remain poorly understood. It is believed that the specificity of the recognition signal carried by a particular polyUb chain is determined by the unique conformations that a particular chain can adopt, which in turn are dictated by the linkage type.¹⁶ For example, the conformational differences between K48-linked and K63-linked chains^{17,18} are responsible for the ability of K48-linked polyUb to bind K48-selective receptors in a “sandwich”-like mode¹⁹, which contrasts the avid binding of K63-linked chains to their specific receptors.²⁰

Studies of the relationship between the linkage, structure, and function of the polyUb signals require the ability to generate these chains with the native connectivity, controlled length, defined linkage and composition, and in sufficient (mg-scale) quantities. This, however, is a significant challenge, because lysine-specific E2 enzymes are not available for all linkages. The problem is further exacerbated for *in vitro* ubiquitination of target proteins, because of the need for substrate-specific Ub-ligases (E3) which are often not known. Until recently, controlled enzymatic assembly of polyUb chains of defined length required the introduction of chain-terminating mutations that inevitably resulted in polyUb chains having surrogate linkages or some lysines permanently replaced with other residues (e.g. Arg or Cys).^{21–23} We have demonstrated²⁴ that this problem can be addressed by using lysines with removable protecting groups (incorporated into Ub as genetically encoded unnatural amino acids), thus allowing controlled assembly of native polyUb chains of any desired length. Nevertheless, the lack of linkage-specific E2s and substrate-specific E3s remains a significant bottleneck.

These limitations have motivated several research groups to develop non-enzymatic methods to form the isopeptide bond.^{25–35} In particular, it has been demonstrated that by using total chemical synthesis combined with native isopeptide chemical ligation (through mercaptolysine residues) it is possible to make Ub₂ chains^{28,29} and even Ub₄.³³ While this is a remarkable accomplishment that opens new opportunities to build and characterize polyUb chains, these methods are limited in their broad implementation by the need to use total chemical synthesis, which is not readily available in every biochemical laboratory. Also, the need to desulfurize the final product, in order to obtain a native Gly-Lys isopeptide linkage, limits the applicability of various mercaptolysine-based approaches in cases of substrate proteins containing sulfur-bearing side chains, e.g. cysteines.

Even given access to such polyUb chains, another major challenge remains, namely the need to study these chains by high resolution structural methods. The weak noncovalent intra-chain interactions in polyUb complicate the use of crystallographic methods: the chains either evade crystallization or the resulting crystal structures might not represent the physiologically relevant conformations.^{16,17,36,37} This highlights the necessity to study

polyUb chains in solution, and thus makes NMR the method of choice for such studies. Unfortunately, the homopolymeric nature of polyUb, combined with the chemical and spectroscopic similarity of Ub monomers, makes it almost impossible to resolve the NMR signals from the individual Ub units within the chain, required for monomer-specific characterization, unless the individual Ubs are differentially isotopically labeled (^{23,38}, also Figure S1). The latter task, however, is essentially impractical for chemically synthesized chains due to the high cost of incorporating isotopic labeling into chemical synthesis. The need to circumvent the formidable cost of total peptide synthesis of isotope-labeled proteins motivated us to seek alternative ways of achieving non-enzymatic ubiquitination, which would (1) be affordable and accessible, (2) yield fully natural polyUb chains of any desired length and linkage, and (3) allow isotopic labeling of any selected Ub moiety in the polyUb chain. We believe that this can be achieved by taking advantage of accessible and cost-effective production of isotope-enriched proteins using bacterial expression.

Here we present a robust and cost-effective method for iterative non-enzymatic assembly of bacterially expressed Ub monomers into an *all-natural* polyUb chain (i.e. containing no mutations and linked through native isopeptide bonds) of any desired length and composition. This method uses a silver-mediated condensation reaction³⁹ between a thioester on the C-terminus of one Ub and the ε-amine of a specific lysine in the other Ub (or another target protein). Our strategy is based on incorporation (as an unnatural amino acid) of a lysine derivative bearing a removable protecting group (e.g., Boc-lysine (Lys(Boc))) at the desired position in the protein (for the isopeptide linkage) and blocking all other primary amines with an orthogonal protecting group, allyloxycarbonyl (Alloc), followed by removal of one or the other protecting group as necessary (Scheme 1). We were inspired by the recent work by Virdee et al.³⁰ that utilized an analogous strategy to make K6- and K29-linked Ub₂s. The novelty and advantages of the method presented here are (1) the use of the Alloc group which, unlike the carbobenzyloxy (Cbz) group used in³⁰, is entirely orthogonal to the Boc group and requires less harsh conditions for removal, and that our method allows (2) synthesis of polyUb chains of any length/composition and (3) isotopic labeling of any desired Ub monomer in the chain. Using this method, we successfully assembled K11-linked Ub₂, Ub₃, and Ub₄, K33-linked Ub₂, and mixed-linkage, K33,K11-linked Ub₃ chains containing all natural isopeptide linkages. Moreover, by ¹⁵N enrichment of a specific Ub within the Ub₂ and Ub₃ chains, we were able, for the first time, to characterize, using high-resolution NMR, Ub units at any position in such chains.

Materials and Methods

Assembly of Ub chains

Below we describe in detail all steps and protocols involved in the non-enzymatic assembly of a Ub chain, using Ub₂ as an example (Scheme 1). Iterative assembly of longer chains (Schemes 2, 3) from preassembled Ub₂ or longer constructs involves similar procedures.

Preparation of Lys(Boc)-containing Ub—Plasmids pTXB1, containing the *E. coli* codon-optimized Ub gene with the TAG mutation at residue position 11 or 33, and pSUP-PylT-PylS, containing cellular machinery to incorporate Lys(Boc) as a genetically encoded unnatural amino acid at the TAG codon, were double transformed into *E. coli* BL21(DE3) chemically competent cells as detailed elsewhere.²⁴ ¹⁵N-labeled wild type (WT) Ub and ¹⁵N-labeled Ub containing Lys(Boc) were expressed in *E. coli* using auto-inducing minimal media with ¹⁵NH₄Cl as the sole source of nitrogen⁴⁰ and purified as described.²⁴

Generation of Ub-SR—Ub variants with the thioester group (Ub-SR) at the C-terminal G76 were generated from respective Ub monomers (unlabeled or ¹⁵N labeled) using the

following procedure. A 1 mL solution of 5 mg/mL Ub was prepared in 20 mM sodium phosphate buffer (pH 8.0) with 10 mM ATP, 10 mM MgCl₂, 100 mM sodium 2-mercaptoethane-sulphonate (MESNA), and 250 nM of Ub-activating enzyme E1.³¹ The mixture was incubated at 37°C for 6 hours and then stored at 4°C. The E1 enzyme was precipitated by adding a drop of glacial acetic acid to the solution. The protein solution was buffer-exchanged into dH₂O containing 0.4% TFA and subsequently lyophilized. Ub-SR generation was confirmed by ESI-MS (electrospray ionization mass spectrometry).

Alloc protection of Ub monomers—The Alloc protection reaction took place by initially dissolving 5 mg of each Ub monomer in 450 μL DMSO for 15–20 minutes on a platform rocker. To this solution, 17 μL of diisopropylethylamine (DIEA) and 75 μL of freshly made 40 mg/mL Alloc-OSu (from TCI America) solution in DMSO were added and allowed to react for at least 1 hour at room temperature on the platform rocker. Complete Alloc protection was monitored by performing ESI-MS on the sample. For each 100 μL of protein solution, the protein was precipitated into a small white pellet by mixing ~2 mL of ice-cold ether, vortexing for 15 seconds, and centrifuging for 10 minutes at 4°C. The top organic layer was removed after centrifugation and two more rounds of ether precipitation were conducted. The pellet was then air-dried for 15–20 minutes.

Boc deprotection of Ub containing Lys(Boc) at desired position for isopeptide linkage—The pellet (5 mg) from the previous step was dissolved in 500 μL of ice-cold 3:2 TFA/dH₂O solution and left to react for at least 2–3 hours at 4°C.³⁰ Complete dissolution occurred after 15 minutes; and can be aided by gently pipetting the solution. Removal of Boc group (molecular weight loss of 100 Da) was monitored by ESI-MS. For every 100 μL of solution, the protein was precipitated with three rounds of cold ether as described above. Typically the protein precipitated instantly after initial ether addition, and the resulting white pellet was air-dried for 15–20 minutes.

Ligation of Ub monomers—Approximately 2 mg of each of the two protected Ub monomers were dissolved in DMSO and then added together to a total volume of 90 μL as originally described.³⁰ DIEA (4 μL), hydroxysuccinimide (H-OSu) (1 μL of fresh 390 mg/mL solution dissolved in DMSO), and AgNO₃ (1 μL of fresh 57 mg/mL solution dissolved in DMSO) were added to the Ub monomers, mixed, and incubated at room temperature in the dark for at least 20–30 hours. Formation of ligated Ub₂ species was confirmed by SDS-PAGE. The solution was subjected to three rounds of ether precipitation as described above and then air-dried. The pellet typically was slightly yellowish in color.

Global Alloc deprotection—Ub was deprotected using 50 mol % chloro-pentamethylcyclopentadienyl-cyclooctadiene-ruthenium(II) ([Cp**Ru*(cod)Cl], from Sigma) and 50 equivalents of thiophenol, relative to the number of moles of protected amines in the solution (counting all lysine and histidine residues and the N-terminal amine). For example, for 4 mg of total Ub monomer and assuring slight excess over nine amine protection sites on each Ub molecule, the reaction pellet was dissolved in 420 μL of DMSO to which 240 μL of H₂O, 116.2 μL of 20 mM [Cp**Ru*(cod)Cl] (freshly dissolved in DMSO), and 23.8 μL of neat thiophenol were added. The resulting dark-brown/black solution was divided into 200 μL aliquots in PCR tubes. The reactions were incubated in a thermal cycler at 50°C for 2 hours. After this the tubes were allowed to cool for 10–15 minutes; the resulting solution was typically a dark orange cloudy solution with black precipitate.

Approximately 100 μL of each reaction tube were aliquoted into 2 mL eppendorf tubes for cold ether precipitation as described above. Typically, at least 10 rounds of ether precipitation were necessary to form a small dark-brown pellet in each of these eppendorf tubes. After each round of centrifugation, the top organic layer was very carefully removed,

so as to not disturb the aqueous bottom layer. When cold ether was added, vigorous vortexing for 15 seconds was performed to ensure proper mixing of the solution. After pellet formation, the pellet was allowed to air dry for 10–15 minutes.

Renaturation and purification of the ligation product—The above protein solution from Alloc deprotection was initially dissolved in a total of 500 μL – 1 mL of filtered 6 M GdnHCl, 20 mM sodium phosphate (pH 6.8) solution for 10–15 minutes. The solution was centrifuged for 5 minutes to remove the undissolved precipitate. To this solution, 200 μL of 20 mM sodium phosphate (pH 6.8) buffer containing 130 mM NaCl were added step-wise for 10–15 minutes at a time, to a final volume of 10–15 mL. The clear solution was then transferred to 3K MWCO dialysis tubing and dialyzed overnight against 2 L of the same buffer at 4°C.

The next morning, the dialysate was concentrated to < 1 mL using Amicon Ultra-15 3K MWCO units and then purified on a Superdex 75 120 mL column on an AKTA FPLC system. The fractions pertaining to the desired Ub species were collected, concentrated, and exchanged into NMR buffer (20 mM sodium phosphate, pH 6.8). Purity was assessed to be > 99% based on SDS-PAGE gel analysis. Typically from an input of 5 mg of Ub monomers, 0.7 mg – 1 mg of Ub₂ was obtained. To ensure that complete Alloc deprotection took place, a small protein aliquot (3–4 μL of a 100 μM sample) was analyzed by ESI-MS.

Ub₃ assembly from Ub₂ and Ub—To assemble K11-linked Ub₃ with the middle Ub ¹⁵N-labeled, approximately 0.75 mg of K11-linked Ub₂ (¹⁵N-labeled on the distal Ub containing K11Boc) was reacted with 1.25 mg of thioesterified WT Ub. Both proteins were protected with Alloc groups, and K11-linked Ub₂ was treated with TFA to expose the ϵ -amine of K11 for the chemical condensation reaction. Immediately prior to the reaction, each protein was dissolved in 45 μL of DMSO. After mixing the proteins together, the remaining ingredients (DIEA, AgNO₃, and H-OSu) were added. Alloc deprotection was performed assuming 3 mg of Ub monomer with ten Alloc-protected amines per Ub. This overestimation of the amount of Ub ensured complete Alloc deprotection. The reaction pellet was dissolved in 315 μL of DMSO to which 180 μL of H₂O, 87.2 μL of 20 mM [Cp*¹⁵N-Ru(cod)Cl] (freshly dissolved in DMSO), and 17.9 μL of neat thiophenol were added. Protein was precipitated with ether, renatured, and purified as described above. A small aliquot (5 μL of 40 μM protein) was analyzed by ESI-MS to confirm that all Alloc groups on K11-linked Ub₃ were removed (Supporting Information). Total amount of purified Ub₃ was determined to be 0.3 mg (15% yield).

To assemble K33,K11-linked Ub₃ with the middle Ub ¹⁵N-labeled, 0.75 mg of K33-linked Ub₂ (¹⁵N-labeled on the proximal Ub) was reacted with 3 mg (an unintentional excess amount) of K11Boc Ub. The C-terminus of K33-Ub₂ was thioesterified using MESNA and E1 enzyme as described above; complete thioesterification was verified by ESI-MS. The protein was protected with Alloc groups as described above, assuming 2 mg of Ub, and checked by ESI-MS. K11Boc Ub was protected with Alloc groups and treated with TFA as described above. For the condensation reaction, each protein was dissolved in 45 μL of DMSO before mixing both proteins together. DIEA, AgNO₃, and H-OSu were added in the same amounts as described above. Subsequent Alloc deprotection took place assuming a total of 5 mg of Ub monomer with ten Alloc-protected amines per Ub. The reaction pellet was dissolved in 525 μL of DMSO to which 300 μL of H₂O, 145.3 μL of 20 mM [Cp*¹⁵N-Ru(cod)Cl] (freshly dissolved in DMSO), and 29.8 μL of neat thiophenol were added. Protein was renatured and purified as detailed above. Total amount of K33,K11-linked Ub₃ was 0.35 mg (9% yield). The apparent lower yield reflects the excess amount of unreacted K11Boc Ub present in the reaction.

NMR experiments

All NMR measurements were performed on Avance III 600 MHz spectrometer (Bruker Biospin) equipped with TXI cryoprobe using standard or in-house pulse sequences; the sample temperature was set to 23°C. The data were processed using NMRPipe⁴¹ and analyzed using Sparky⁴².

Mass spectrometry

For Ub monomers and some Ub₂, high resolution mass spectra of m/z 250–2500 were acquired with JEOL AccuTOF-CS mass spectrometer in electrospray positive mode using flow injection. To determine the molecular weight, spectra were deconvoluted using MagTran software with the maximum charge set to 30. For dilute Ub₂ and Ub₃ samples, high resolution mass spectra of m/z 400–2000 (or 600–4000 for protein with alloc protection group) were acquired on Thermo Scientific LTQ Orbitrap XL mass spectrometer using flow injection. The resolution was 60,000 at m/z 400. Deconvolution was carried out using either the Xtract program in Xcalibur software or MagTran with a maximum charge set to 30.

Results

We devised a strategy for iterative non-enzymatic assembly of Ub chains of any desired length and linkage composition using bacterially expressed recombinant Ub monomers as primary building blocks. At the heart of this approach is the formation of an isopeptide bond through chemical condensation reaction³⁹ between the thioesterified C-terminus of one Ub and a specific lysine residue of the other Ub (Scheme 1). The same reaction can be used for site-specific non-enzymatic (poly)ubiquitination of any other protein. To control the reaction and direct the linkage to a specific lysine, the latter is initially protected with the Boc group (introduced as a genetically encoded unnatural amino acid, Lys(Boc)²⁴), while all other amines on both reactant proteins are protected with the Alloc groups. Subsequent removal of the Boc group from Lys(Boc) makes this lysine the only available amine to interact with the C-terminal thioester (SR) of another Ub to condense into an isopeptide bond. This step is conceptually similar to that in ref³⁰, with the principal difference being the use of the Alloc protecting group instead of the Cbz group. The complete orthogonality of the Alloc and Boc groups (see below) allows their removal entirely independently from one another. This enables iterative Ub chain assembly (Schemes 2, 3) to construct *all-natural* Ub chains of *any* desired length and linkage, comprised of either homogeneous or mixed linkages (Scheme 3), and to have *any* Ub in the chain isotopically enriched at will (Schemes 1, 2). The basic steps in our iterative ubiquitination/chain-assembly procedure are as follows (Schemes 1–3):

1. activation (by thioesterification) of the C-terminus of Ub (or a Ub chain) for the ligation with a target protein (e.g., another Ub or some other protein);
2. incorporation of a lysine side chain bearing a removable protecting group (e.g., Lys(Boc)) as a genetically encoded unnatural amino acid at the desired position in the target protein;
3. protection of all available amines (lysine and histidine side chains and the N-terminal amine) on both reactant proteins with a different, orthogonal protecting group (Alloc);
4. removal of the Boc protecting group to expose the specific Lys side chain as the sole ligation site;
5. ligation of the two proteins via chemical condensation reaction;
6. complete removal of all Alloc groups;
7. renaturation and purification of the desired product.

Optimization of the Alloc deprotection reaction

Central to the success of this method is complete Alloc deprotection. Recently it has been shown⁴³ that Alloc protecting groups can be efficiently removed using relatively mild conditions in neutral aqueous solution using chloro-pentamethylcyclopentadienyl-cyclooctadiene-ruthenium(II) ([Cp**Ru*(cod)Cl], *Ru* catalyst) and thiophenol. However, application of these exact conditions to Ub protein, which can have up to nine Alloc groups attached to it simultaneously, failed to completely remove all Alloc groups. To determine the correct conditions for Alloc deprotection, we tested multiple experimental conditions on a ¹⁵N- labeled sample of WT Ub (Figure S2). After many rounds of optimization, we discovered the correct amounts of *Ru* catalyst, thiophenol, and water to use in the Alloc deprotection reaction. The most critical features to the success of the reaction are (i) near-stoichiometric amounts of *Ru* catalyst (50% mol) relative to the moles of amine present, (ii) at least 30% H₂O (v/v), and (iii) temperature raised to at least 50°C. When any of these conditions were not met, ESI-MS analysis revealed that Ub species typically still contained a variable number of Alloc protecting groups covalently attached to Ub (Figure S2). For example, lowering the amount of ruthenium metal in the reaction did not yield full deprotection. Once the Alloc deprotection reaction began, typically the reaction reached completion after a maximum of two hours. Allowing reactions to proceed overnight did not change the ESI-MS results.

Thioesterification of the C-terminus of ubiquitin

To prepare the monomers for the ligation reaction (Scheme 1) the C-terminus of the (distal) Ub must be first activated with a thioester functional group. Traditionally, Ub-SR is made via MESNA-induced cleavage reaction of a Ub-intein-CBD (chitin binding domain) fusion construct. Unfortunately, this approach is slow (each cleavage reaction takes 24–40 hours), necessitates HPLC purification (removal of an unwanted byproduct of unactivated Ub), and produces Ub-SR in low yield (only 10 mg of pure Ub-SR from expression in 2 L of culture). A promising alternative, first introduced by Oualid et al.³¹, generates Ub-SR by reacting wild type (WT) Ub with Ub-activating enzyme, E1, and MESNA. We found this approach very simple and efficient at generating Ub-SR. Complete thioesterification of Ub's C-terminus occurs after 6 hours of treatment with the E1 enzyme and MESNA, resulting in addition of 125 Da (MES) to the molecular weight of Ub, as detected by ESI-MS (Figure 1a). Subsequent treatment with glacial acetic acid precipitates E1 and lowers the pH to maintain Ub-SR stability. Finally, buffer exchange removes other reaction components to yield a pure Ub-SR product. This method produces Ub-SR at high yield, since it is easy to express and purify WT Ub in large amounts (50–100 mg from 1 L culture) and react with E1 to make Ub-SR. Importantly, the same procedure works nicely for activating Lys(Boc) Ub as well as all preassembled polyUb chains that we tested (see below).

Assembly of natural K11-linked Ub₂

As a proof of principle, we first assembled natural K11-linked Ub₂ ¹⁵N-labeled on the proximal Ub. We chose this di-Ub because we can compare its properties to the enzymatically synthesized K11-linked Ub₂ construct made using the K11-specific E2 Ube2s. K11 linkages could be as abundant as K48 and K63 linkages,⁴⁴ and K11-linked polyUb chains appear to act both as regulatory and proteolytic signals.^{7,45} To assemble the K11-linked Ub₂ non-enzymatically, we prepared two proteins, WT Ub and ¹⁵N K11Boc Ub. The ligation via chemical condensation reaction was carried out as outlined above and indicated in Scheme 1.

All of the unprotected amines (lysines, N-terminus, and the histidine³⁰) on both the distal Ub-SR and the proximal ¹⁵N K11Boc Ub were protected with the Alloc group (molecular weight 84 Da). ESI-MS results confirmed that a total of nine Alloc groups were covalently

attached to Ub-SR (Figure 1a). On ^{15}N K11Boc Ub, generally only eight Alloc groups were attached, consistent with the fact that K11 still contains a Boc group (Figure 1b). The Alloc-protected ^{15}N K11Boc Ub was then Boc-deprotected at residue 11 via TFA treatment (see Materials and Methods). ESI-MS confirmed a loss of 100 Da, attributed to the molecular weight of a single Boc group (Figure 1b). Importantly, the Boc group removal occurred without affecting the Alloc groups. Both Alloc-protected Ub-SR and Alloc-protected ^{15}N Ub (bearing deprotected K11) were reacted together with AgNO_3 and H-OSu to form an Alloc-protected K11-linked Ub_2 . Typically the reaction yields between 30–50% formation of Ub_2 as monitored by SDS-PAGE (Figure 2a).

After Alloc deprotection followed by renaturation, the desired Ub_2 product was purified from the unreacted monomers using size exclusion chromatography (Figure 2b), concentrated, and buffer exchanged into the appropriate NMR buffer. The complete removal of all Alloc protecting groups from the purified Ub_2 was confirmed by ESI-MS (Figure 1c). As expected, the observed molecular weight of the all-natural Ub_2 with uniform ^{15}N labeling on the proximal Ub was 17210 Da.

Using Scheme 1 it is also straightforward to assemble a K11-linked Ub_2 ^{15}N -labeled on the distal Ub. This can be accomplished by either (i) ^{15}N labeling the (distal) WT Ub rather than K11Boc Ub in Scheme 1, or (ii) using ^{15}N K11Boc Ub for the distal unit and unlabeled K11Boc Ub for the proximal. Note that using K11Boc Ub as the distal unit is essential for assembling longer chains by elongation on the distal end (see below). Therefore, as proof of principle, we prepared K11-linked Ub_2 ^{15}N -labeled on the distal Ub in this manner. All of the steps of Scheme 1 were followed as is, and verified by ESI-MS (Figure 3). Most importantly, Alloc deprotection did *not* remove the remaining Boc group on the distal Ub (Figure 3c, also Figure S3). Together with the results in Figure 1b, this demonstrates the complete orthogonality of the Boc and Alloc protecting groups. Only after purification of the K11-linked Ub_2 , the protein was treated with 3% TFA for 4–6 hours to remove the remaining Boc group on the distal Ub. A significant advantage of this step is that it can be completed without the need to fully denature the product (Ub_2).

Unit-specific characterization of natural K11-linked Ub_2

NMR is exquisitely sensitive to nuances in the chemical environment of the nucleus under observation that are not detectable by other methods. In the ^1H - ^{15}N TROSY-HSQC⁴⁶ and SOFAST-HMQC⁴⁷ experiments used here, primarily signals from the backbone ^1H - ^{15}N amide bonds are recorded, thus providing a characteristic spectral ‘fingerprint’ of a protein. To examine the resulting Ub_2 by NMR, we recorded ^1H - ^{15}N TROSY spectra of the K11-linked Ub_2 chains ^{15}N -labeled on the distal or the proximal Ub, assembled as detailed above, and overlaid them separately on ^1H - ^{15}N TROSY spectra of WT monoUb (Figure 4a,b). The NMR spectra of both Ub units in the chemically ligated K11-linked Ub_2 revealed an excellent spectral dispersion of NMR resonances indicative of a fully renatured Ub_2 . No minor peaks were observed, indicating that the chemically ligated K11-linked Ub_2 is free of contaminating species. On the distal Ub, the largest signal shifts compared to monoUb were observed for the C-terminal residues G75 and G76, which is consistent with the formation of an isopeptide linkage involving the C-terminus of the distal Ub. The K11 backbone amide signal is not present in the distal-Ub spectrum because Lys(Boc) at residue 11 in this Ub was introduced as an unlabeled (naturally abundant, ^{14}N) amino acid, which is “invisible” in the ^{15}N NMR spectra. Likewise, the K11 backbone amide signal is absent also in the spectrum of the ^{15}N -labeled proximal Ub (Figure 4b). Most importantly, the G76-K11 isopeptide signal is also absent (Figure S4), and there are no unaccounted for signals in the spectrum. Taken together, these observations indicate that the resulting Ub_2 species is comprised of a single, K11-specific isopeptide linkage. Moreover, the overall similarity of the NMR spectra of both Ub units with the spectra of monoUb and the relatively small

magnitudes of the signal shifts (except for the C-terminus of the distal Ub) strongly suggest that the structure of each Ub unit in the chemically assembled K11-linked Ub₂ is intact.

To validate that the chemically ligated natural K11-linked Ub₂ had the same structural/conformational properties as the enzymatically assembled chain, we compared the ¹H-¹⁵N NMR spectra of this K11-linked Ub₂ with those of the K11-linked Ub₂ made using K11-specific E2 enzyme, Ube2s. Unfortunately, the enzymatically assembled Ub₂ contained chain-terminating mutations (K11R and K63R on the distal Ub and K63R and D77 on the proximal Ub), complicating direct comparison of the spectra of the two Ub₂ chains. Therefore, we compared NMR spectra for each Ub unit in these chains with its respective unconjugated Ub monomer. Residue-specific chemical shift perturbations (CSPs) were quantified and plotted in Figures 4c–f. The remarkable similarity between the CSPs in the chemically ligated (Figure 4c,d) and enzymatically assembled (Figure 4e,f) K11-linked Ub₂s, serves as a strong indicator that the two chains have identical structural and conformational properties.

Notably, the two Ub units in K11-linked Ub₂ show strikingly different CSP patterns, namely, significant CSPs are spread throughout the proximal Ub but almost absent in the distal Ub (except for the ligated C-terminus). This is in contrast with the CSP patterns observed in K48-linked¹⁷ and K63-linked¹⁸ Ub₂s, raising the question about the nature of the CSPs detected in the proximal Ub: do these reflect some noncovalent interdomain contacts in K11-linked Ub₂ or merely are caused by the chemical (isopeptide) modification of K11? A close inspection shows that the residues with significant CSPs in the proximal Ub cluster around K11 in the 3D structure of Ub, thus suggesting that these perturbations could be due to changes in the microenvironment of K11 upon the isopeptide linkage formation. To test this hypothesis, we measured the CSPs in K11Boc Ub vs. WT Ub (Figure S5a), where the only difference between the two proteins is the presence of the Boc protecting group on the side chain of K11. Indeed, the results show a remarkable overall similarity between the CSP patterns in K11Boc Ub (vs. WT Ub, Figure S5a) and in the proximal Ub of K11-linked Ub₂ (Figure 4d,f). This strongly suggests that the spectral perturbations observed in the latter are primarily caused by the chemical modification of the side chain of K11 resulting from the isopeptide linkage. Further structural studies by solution NMR (currently underway) are required to fully address this issue.

Assembly and characterization of K33-linked Ub₂

To demonstrate that our method is not exclusive to K11 linkages, we also assembled, for the first time, all-natural K33-linked Ub₂ ¹⁵N-labeled on the proximal Ub. K33-linked polyUb chains have been found to play a non-proteolytic, regulatory role in T-cell signaling.⁴⁸ We made this chain using the same approach as in Scheme 1, but substituting K11Boc Ub with K33Boc Ub. The ESI-MS data (Figure S6) confirm that the purified K33-linked Ub₂ has the expected molecular weight of 17210 Da, assuming one Ub is ¹⁵N-labeled and the other is not.

NMR characterization of the proximal Ub in K33-linked Ub₂ revealed an excellent spectral dispersion in both ¹⁵N and ¹H dimensions (Figure 5), strongly indicating that this Ub unit is well folded. Overall, the ¹H-¹⁵N spectrum of the proximal Ub unit is very similar to that of monoUb (Figure 5a,c), and the relatively small magnitude of the observed CSPs (Figure 5c) suggests that structurally this Ub unit is essentially intact. As expected, the ¹H-¹⁵N signal of K33 is absent, because this residue was introduced as an unlabeled Boc-containing amino acid. Note that the CSP pattern for the proximal Ub in this chain is distinct from that in the K11-linked Ub₂ (Figure 4d). Interestingly, in K33-linked Ub₂, the CSPs are quite small for the hydrophobic patch residues L8, I44, and V70, and instead, cluster around residues 12–17 and 26–35 located close in space or adjacent to K33. This suggests that the majority of the

spectral perturbations observed in this Ub unit are a consequence of the chemical modification of K33 by the isopeptide linkage. To further support this conclusion, we overlaid the spectra of K33Boc Ub and WT Ub (Figure 4b) and calculated the corresponding CSPs (Figure 5d). Indeed, the pattern of spectral perturbations for K33Boc Ub is overall very similar to that for the proximal Ub in K33-linked Ub₂, with the exception of few residues around V70 in K33Boc Ub.

Assembly of longer Ub chains

The power of our approach is in its ability to easily extend from producing natural di-Ubs to natural polyUb chains of any desired length, either homogeneously linked or comprised of mixed linkages. In Scheme 2, we devised an iterative approach that allowed us to assemble, for the first time, an all-natural Ub₃ chain comprised of only K11 linkages. Moreover, the middle Ub unit in this chain was enriched with ¹⁵N to enable its characterization by heteronuclear NMR. Furthermore, to demonstrate the ability to make even longer chains, we also assembled a K11-linked Ub₄ chain (see below).

As the preparatory step for the construction of K11-linked Ub₃, we assembled and purified K11-linked Ub₂ whose distal Ub both was ¹⁵N-labeled and still contained the Boc group on K11 (Scheme 2). Figure S5b demonstrates the ability to successfully renature and examine by NMR the K11-linked Ub₂, with the Boc group attached to the distal Ub. Treating an aliquot of this Ub₂ with 3% TFA removed the Boc group without denaturing the protein and yielding a completely natural K11-linked Ub₂ ¹⁵N labeled on the distal Ub (Figure 4a). This illustrates that our iterative chain assembly procedure allows characterization of any intermediate species in the process of forming longer chains. For example, as shown in Figure S5, NMR can be used to detect the presence of a Boc group on ¹⁵N-labeled Ub species. This provides a useful tool for monitoring and understanding the effects of Ub chain elongation on the structure and intra-chain interactions of individual Ub units and on the conformational properties of the polyUb chain.

As illustrated in Scheme 2, the assembly of K11-linked Ub₃ involved the use of two proteins, Ub-SR and K11-linked Ub₂ containing ¹⁵N-labeled K11Boc Ub as the distal unit. First, Ub-SR was prepared from WT Ub using E1 and MESNA as described above. Second, Alloc protection was performed on both Ub-SR and K11-linked Ub₂. Subsequent TFA treatment of the Ub₂ unmasked K11 on the distal Ub, enabling the formation of an isopeptide linkage with the C-terminus of Ub-SR. Ligation of the two proteins was performed as described above, and after 16–20 hours formation of Ub₃ was confirmed using SDS-PAGE (Figure 2c). Following Alloc deprotection, the Ub₃ was renatured and purified. The size exclusion chromatogram in Figure 2d revealed that the desired Ub₃ species can be isolated from the unreacted monoUb and Ub₂ components. Finally, ESI-MS of the purified Ub₃ product confirmed that it indeed has the correct molecular weight (25758 Da) expected for all-natural Ub₃ with one Ub ¹⁵N-labeled (Figure S7). It should be mentioned here, for completeness, that the Ub₃ chain could also be assembled by reacting K11-linked Ub₂-SR with monoUb, as exemplified in Scheme 3c for a mixed-linkage chain.

Unit-specific characterization of K11-linked Ub₃

Isotopic labeling of the middle Ub in K11-linked Ub₃ allowed us to selectively examine by NMR this specific Ub unit in the chain (Figure 6). The excellent spectral dispersion of both ¹H and ¹⁵N resonances of the backbone amides clearly indicates that the middle Ub is fully folded despite being tethered to two other Ub units in the chain. The ¹H transverse relaxation time (T₂) of amide protons in K11-linked Ub₃ was 19 ms, consistent with the ¹H T₂ expected of a protein of this molecular weight, and between the T₂ values of 25 ms measured for its precursor, K11-linked Ub₂, as well as K48-linked Ub₂¹⁷ and 13 ms for

K48-linked Ub₄.¹⁷ Also the average ¹⁵N T₁ of 870 ms measured for the middle Ub is consistent with the size of the resulting chain.¹⁹

The observed spectral differences between the middle Ub in K11-linked Ub₃ and monoUb are a composite of the perturbations observed above for both distal and proximal Ubs upon formation of K11-linked Ub₂ (Figure 6d). This is not unexpected given that the middle Ub unit is both proximal and distal with regard to the two Ub units flanking it. To deconvolute the observed spectral perturbations in the middle Ub, we calculated the corresponding CSPs versus the distal and proximal Ubs of natural K11-linked Ub₂ (Figures 6e and 6f, respectively). The CSPs versus the distal Ub of K11-linked Ub₂ (Figure 6e) are remarkably similar to the CSPs in the proximal Ub (versus monoUb) upon formation of the K11-linked Ub₂, shown in Figure 4d. Also the directions of the signal shifts (Figure 6b) are identical to those in Figure 4b. The remarkable similarity of the CSP patterns in Figures 4d and 6e strongly indicates that (i) the K11 isopeptide linkage is formed between the middle Ub and the distal Ub in K11-linked Ub₃, and (ii) no additional Ub/Ub interface (compared to that in K11-linked Ub₂) exists between the two Ub units in this chain.

Likewise, the CSPs in the middle Ub versus the proximal Ub of K11-linked Ub₂ (Figure 6f) are similar to those observed for the distal Ub (versus monoUb) upon its incorporation into K11-linked Ub₂ (Figure 4c). The strongest CSPs are observed for the C-terminal residues G75 and G76. Such perturbations are typically the hallmark of G76's involvement in an isopeptide bond.^{17,18,23} Thus, this observation verifies that the C-terminus of the middle Ub in K11-linked Ub₃ is connected via an isopeptide bond to the proximal Ub. Interestingly, the magnitude of the CSPs for the rest of the residues in the middle Ub is somewhat higher than in the distal Ub of K11-linked Ub₂ (Figure 4c). Here the elevated CSPs cluster around residues 7–15, 40–50, and 67–70, which make up part of the hydrophobic surface patch in Ub. It is conceivable that there is an increase in Ub/Ub interfacial contacts between the middle Ub and the proximal Ub in K11-linked Ub₃ compared to K11-linked Ub₂. This might indicate an onset of additional inter-Ub contacts as the chain gets longer. However, it is important to bear in mind that the magnitudes of these CSPs are quite small overall, especially compared to those observed in K48-linked Ub₂¹⁷, suggesting that the Ub/Ub interfaces that form in K11-linked Ub₃ are transient.

Assembly and characterization of mixed-linkage Ub₃

To demonstrate the versatility of our method, we also assembled, for the first time, a mixed-linkage Ub₃ chain comprised of K33- and K11-linkages, as outlined in Scheme 3c. For this we used K33-linked Ub₂ (assembled and characterized above, with the proximal Ub ¹⁵N labeled), and K11Boc Ub. The Ub₂ was treated with E1 and MESNA to activate (thioesterify) its C-terminus and then subjected to Alloc protection of all available seventeen amines (see an example in Fig. S11). K11Boc Ub underwent Alloc protection of all available amines followed by the removal of the Boc group to expose K11 as the sole lysine side chain for the ligation reaction. Ligation of the two proteins was performed as described above, and after 16–20 hours the formation of Ub₃ was confirmed by SDS-PAGE (Figure 2e). Following Alloc deprotection, the Ub₃ was renatured and purified (Figure 2f). The correct size (number of Ub units) of the assembled Ub₃ chain was independently confirmed by NMR relaxation measurements, which revealed a decrease in the ¹H T₂ value from 25 ms for the precursor K33-linked Ub₂ to 18.5 ms for the final Ub₃ product, similar to that for K11-linked Ub₃ (see above).

The ¹H-¹⁵N NMR spectra of the middle Ub in the mixed-linkage K33,K11 Ub₃ chain are shown in Figure 7. As with the K11-linked Ub₃ (see above), the spread of the NMR signals and the strong similarity with the spectra of monoUb indicate that the structure of this Ub unit is essentially intact. Similar to the middle Ub in K11-linked Ub₃, the spectral

perturbations in the middle Ub of K33,K11-linked Ub₃ are a composition of the perturbations in the respective Ub₂ constructs, reflecting the role of the middle Ub as both the distal and the proximal Ub with respect to its neighbors in the chain.

Assembly of K11-linked tetra-Ub

As proof of principle, we also assembled K11-linked Ub₄ using as reaction components the above-characterized homogeneously K11-linked Ub₃ and K11Boc Ub. Following Scheme 3b, we first treated K11-linked Ub₃ with E1 enzyme to generate a thioester at its C-terminus, and subsequently Alloc-protected all the amines. As detailed above, we used K11Boc Ub as the proximal Ub and protected all the amines with Alloc groups prior to TFA treatment that exposed the ε-amine of K11 for the ligation reaction. The two proteins were mixed and ligated for 20 hours, and the formation of a Ub₄ product was confirmed by SDS-PAGE (Figure 8). Combined with the data shown above, this result demonstrates the ability to assemble Ub chain of any desired length using the method proposed here.

Discussion

Deciphering the “ubiquitin code” requires understanding of the linkage-structure-function relationship for various polyUb signals. Structural and functional studies of polyUb chains necessitate the ability to generate natural Ub chains of controlled length, linkage composition, and isotopic labeling (for studies in their native milieu). With the methods devised in this paper, we have demonstrated, for the first time, the ability to make completely *natural* Ub chains of *any* desired length and linkage composition, and at the same time to isotopically label *any* Ub in the resulting chain. The latter renders such chemically assembled chains amenable for unit-specific studies by NMR and any other biophysical technique where isotopic labeling is critical, e.g., for contrast variation in small angle neutron scattering (SANS). Our method can also be used to incorporate modified Ub variants at any specific location in the chain, to study the effect of monomer/site-directed mutations on inter-subunit interactions, chain structure, and receptor recognition, and to enable monomer-selective labeling with various physical probes (e.g. fluorophores, spin-labels, paramagnetic ions) for structural and binding studies by fluorescence/FRET, ESR, NMR etc.

Our approach builds upon several recent significant breakthroughs in enzyme-free ubiquitination. First, our method was inspired by the successful demonstration by Virdee et al.³⁰ that natural di-Ub chains can be synthesized via the GOPAL (genetically encoded orthogonal protection and activated ligation) approach. We use the same basic concept, but replace the Cbz protecting group with the Alloc group.

Second, our method was facilitated by the recent finding by several research groups^{43,49} of a relatively mild set of conditions for removal of the Alloc protecting group from a target lysine on a protein. As we showed here (Figure S3), the use of Ru catalyst together with thiophenol in a partially aqueous pH-neutral solution allows complete Alloc deprotection of multiple amines in a protein without affecting the Boc protecting group. This is in contrast to the Cbz group used in ref³⁰, whose removal with a strong acid cocktail (TFMSA, TFA) also removes the Boc group. Moreover, Boc deprotection can be achieved without affecting the Alloc protecting groups (Figure 3b). This demonstrates a complete, bi-directional orthogonality of the Boc and Alloc groups, i.e. each group can be removed independently of the other. This feature is absolutely essential for the iterative assembly of Ub chains longer than Ub₂. Also critical is the fact that Alloc protection does not interfere with the thioester-activated C-terminus of Ub (Figures 1a, 3a). Together with the ability to activate the C-terminus of an existing chain, this feature is essential for chain elongation on the proximal end. It is worth mentioning here that the milder, Alloc removal conditions proved harmless

to Ub chains and have potentially important utility for applying the same strategy to ubiquitinate other proteins which could be less stable than Ub in acidic conditions.

Third, the recent use of the E1 enzyme combined with MESNA by Ouilad et al.³¹ to activate the C-terminus of Ub with a thioester functional group had profound implications for our iterative assembly of longer Ub chains. Indeed, as demonstrated above, the same method can be used to efficiently activate the C-terminal G76 of various Ub chains studied here: K11-linked Ub₂ and Ub₃ and K33-linked Ub₂, and is likely to work with essentially any polyUb chain. By ESI-MS we verified (see an example in Fig. S11) that we can activate the C-terminus of a Ub₂ with E1 and MESNA with high yield and with no need to purify the product, as it is completely converted into thioester-activated Ub₂.

As demonstrated here on multiple examples, our method of non-enzymatic chain assembly allows isotopic labeling of any Ub unit at will and in any type of Ub chain. Consider, for example, a tri-Ub chain. Here we successfully achieved isotopic labeling of the middle Ub in two different Ub₃ chains, homogeneously K11-linked and mixed-linkage, K33,K11-Ub₃. In order to label the distal Ub in Ub₃, one can use an isotope-labeled Ub-SR in Scheme 2 and an unlabeled K11-linked Ub₂. Similarly, to obtain a K11-linked Ub₃ isotope-labeled on the proximal Ub, one can simply substitute the K11-linked Ub₂ in Scheme 2 with a similar Ub₂ isotope-labeled on the proximal Ub. Moreover, this method is not limited to labeling a single Ub unit in the chain: the use of recombinant Ub monomers enables a “mix-and-match” approach, where several (or all) Ub units could be isotopically labeled, each using a different labeling scheme.

The ability to isotopically label individual Ub units in the chain is absolutely critical for unit-specific structural and functional studies of polyUb in solution (see Fig. S1) Thus, selective isotopic labeling allowed us to obtain, for the first time, residue-specific information on the middle Ub unit in a Ub₃ chain. An important observation here is that the spectral perturbations in the middle Ub are a composition of the spectral perturbations in both the distal and proximal Ubs upon formation of the corresponding Ub₂ fragments. Moreover, a detailed analysis revealed that the non-covalent Ub/Ub contacts in K11-linked Ub₃ and in K33,K11-linked Ub₃ are relatively weak or transient, such that the observed spectral perturbations in the middle Ub largely reflect alterations in the electronic environment caused by the formation of isopeptide linkages.

Our methodology can be easily extended to iteratively assemble completely natural polyUb chains of any length and linkage composition (as exemplified in Scheme 3) and, if required, with any desired composition of isotopic enrichments of the individual Ub units. For example, a Ub₄ comprised only of K11 linkages can be built by either reacting two K11-linked Ub₂ chains or a K11-linked Ub₃ with monoUb (as demonstrated in Figure 8).

Overall, the proposed method of non-enzymatic Ub chain assembly using thioester-based, silver-mediated chemical condensation combined with Alloc/Boc-based protection/deprotection described here is extremely robust: (1) the method is relatively simple and straightforward, with the most time-intensive step being the ether precipitation of the proteins from the Alloc deprotection reaction, (2) all Ub chains assembled so far (including Ub₃s) have been denatured and refolded successfully, and (3) the K11-linked Ub₂ chains assembled using our method possess the same spectral and structural properties as the Ub₂s produced enzymatically. The latter is an important control as we begin to make chains consisting of non-canonical linkages (K6, K27, K29, K33) for which no linkage-specific E2 enzymes are known. In summary, this makes our method particularly well suited for synthesis of *any* Ub chain containing homogeneous or mixed linkages, and potentially for E2- and E3-free ubiquitination of other (target) proteins with these chains.

Ultimately, the goal is to be able to attach Ub chains of any length and linkage composition to any target protein (with no need for specific E2 and E3 enzymes) to begin to understand the biological impact of ubiquitination. Such studies have been hindered by the unavailability of substrate-specific Ub ligases (E3). Our non-enzymatic ubiquitination method can be extended to enable *in vitro* construction of a fully natural ubiquitinated substrate. For example, the reaction shown in Scheme 1 (or Scheme 3b,c) can be used to attach a Ub-SR (or a pre-assembled polyUb-SR chain) to a recombinant target protein bearing Lys(Boc) at the desired ubiquitination site. The latter can be bacterially expressed and purified using the same protocol as for Lys(Boc)-Ub^{24,50}, see also Materials and Methods). Alternatively, polyUb-SR chains pre-assembled using our method can be attached through native chemical ligation to a mercaptolysine on a substrate of interest obtained using recently developed techniques for site-specific incorporation of a mercaptolysine into a target protein using chemical synthesis^{28,29} or bacterial expression³⁴. This provides essentially endless possibilities for non-enzymatic ubiquitination of virtually any substrate protein that can be refolded or can withstand TFA or desulfurization treatment, which will open new, previously unavailable opportunities for structural and functional studies of the outcome of ubiquitination.

Conclusions

The dearth of native polyUb chains of any defined length, linkage composition, and isotopic labeling has been a major hindrance to the ubiquitin field. To address this challenge, we developed an affordable and widely accessible method for controlled, iterative non-enzymatic assembly of completely natural polyUb chains using recombinant monomers as the primary building blocks. Moreover, the use of bacterially expressed Ub monomers allows cost-effective isotopic enrichment of any individual Ub unit in the chain using various isotopic labeling schemes (uniform or residue/group selective) already developed for biomolecular NMR applications (e.g.,⁵¹). This opens up broad opportunities to study polyUb chains by high-resolution NMR, SANS, and other biophysical and biochemical methods, to obtain unit-specific atomic-resolution information on the structure, intra-chain interactions, and receptor-recognition properties of the polyUb signal. Using this method, we synthesized homogeneously K11-linked Ub₂, Ub₃, and Ub₄, and K33-linked Ub₂, as well as a mixed-linkage K33,K11-linked Ub₃. These Ub₃ and Ub₄ chains have never been assembled before. This allowed us to characterize by NMR, for the first time, the middle Ub unit in homogeneously K11-linked Ub₃ and in K33,K11-linked Ub₃, as well as the proximal Ub in K33-linked Ub₂. With this chain-assembly method in hand, it is now possible to generate and study essentially any polyUb chain, both homogeneously linked and with mixed linkages, in order to uncover the structure and receptor recognition of the polyUb signal and its processing by linkage-specific deubiquitinases. Importantly, the assembled chains contain natural isopeptide linkages and can be made from wild type Ub monomers, with no need for permanent mutations, thus yielding *completely* natural Ub chains of essentially any length and in milligram (or larger) quantities. Our method can be extended to (poly)ubiquitinate a target protein, with no need for specific E2 and E3 enzymes, and could also be used to form an isopeptide bond between virtually any two proteins.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Supported by the National Institutes of Health (GM084396 to T.A.C and GM065334 to D.F.) and the National Science Foundation (CHE-0848398 to T.A.C and NSF Postdoctoral Fellowship DBI-0905967 to C.A.C.). We thank

Yan Wang (U. Maryland Proteomics facility) for help with MS studies, and Mike McCrane and Steve Rokita for HPLC advice and usage. We also thank Michael Rape (UC Berkeley) for the Ube2s plasmid.

References

1. Hershko A, Ciechanover A. *Annu Rev Biochem.* 1998; 67:425–480. [PubMed: 9759494]
2. Glickman MH, Ciechanover A. *Physiol Rev.* 2002; 82:373–428. [PubMed: 11917093]
3. Muratani M, Tansey WP. *Nat Rev Mol Cell Biol.* 2003; 4:192–201. [PubMed: 12612638]
4. Aguilar RC, Wendland B. *Curr Opin Cell Biol.* 2003; 15:184–90. [PubMed: 12648674]
5. Osley MA. *Biochim Biophys Acta.* 2004; 1677:74–8. [PubMed: 15020048]
6. Ikeda F, Dikic I. *EMBO Rep.* 2008; 9:536–42. [PubMed: 18516089]
7. Wickliffe K, Williamson A, Jin L, Rape M. *Chem Rev.* 2009; 109:1537–48. [PubMed: 19146381]
8. Ye Y, Rape M. *Nat Rev Mol Cell Biol.* 2009; 10:755–64. [PubMed: 19851334]
9. David Y, Ziv T, Admon A, Navon A. *J Biol Chem.* 2010; 285:8595–604. [PubMed: 20061386]
10. Chau V, Tobias JW, Bachmair A, Marriott D, Ecker DJ, Gonda DK, Varshavsky A. *Science.* 1989; 243:1576–83. [PubMed: 2538923]
11. Thrower JS, Hoffman L, Rechtenstein M, Pickart CM. *EMBO J.* 2000; 19:94–102. [PubMed: 10619848]
12. Hicke L, Dunn R. *Annu Rev Cell Dev Biol.* 2003; 19:141–72. [PubMed: 14570567]
13. Sun L, Chen ZJ. *Curr Opin Cell Biol.* 2004; 16:119–26. [PubMed: 15196553]
14. Spence J, Sadis S, Haas A, Finley D. *Mol Cell Biol.* 1995; 15:1265–73. [PubMed: 7862120]
15. Spence J, Gali RR, Dittmar G, Sherman F, Karin M, Finley D. *Cell.* 2000; 102:67–76. [PubMed: 10929714]
16. Pickart CM, Fushman D. *Curr Opin Chem Biol.* 2004; 8:610–6. [PubMed: 15556404]
17. Varadan R, Walker O, Pickart C, Fushman D. *J Mol Biol.* 2002; 324:637–47. [PubMed: 12460567]
18. Varadan R, Assfalg M, Haririnia A, Raasi S, Pickart C, Fushman D. *J Biol Chem.* 2004; 279:7055–7063. [PubMed: 14645257]
19. Varadan R, Assfalg M, Raasi S, Pickart C, Fushman D. *Mol Cell.* 2005; 18:687–98. [PubMed: 15949443]
20. Sims JJ, Cohen RE. *Mol Cell.* 2009; 33:775–83. [PubMed: 19328070]
21. Piotrowski J, Beal R, Hoffmann L, Wilkinson KD, Cohen RE, Pickart CM. *J Biol Chem.* 1997; 272:23712–21. [PubMed: 9295315]
22. Pickart CM, Raasi S. *Methods Enzymol.* 2005; 399:21–36. [PubMed: 16338346]
23. Varadan, R.; Assfalg, M.; Fushman, D. *Ubiquitin and Protein Degradation, Methods in Enzymology, Vol.399 part B.* Deshaies, RJ., editor. Vol. 399. 2005. p. 177-192.
24. Castaneda CA, Liu J, Kashyap TR, Singh RK, Fushman D, Cropp TA. *Chem Commun (Camb).* 2011; 47:2026–8. [PubMed: 21212884]
25. Chatterjee C, McGinty RK, Pellois JP, Muir TW. *Angew Chem Int Ed Engl.* 2007; 46:2814–8. [PubMed: 17366504]
26. Yang R, Pasunooti KK, Li F, Liu XW, Liu CF. *J Am Chem Soc.* 2009; 131:13592–3. [PubMed: 19728708]
27. Ajish Kumar KS, Haj-Yahya M, Olschewski D, Lashuel HA, Brik A. *Angew Chem Int Ed Engl.* 2009; 48:8090–4. [PubMed: 19780082]
28. Kumar KS, Spasser L, Erlich LA, Bavikar SN, Brik A. *Angew Chem Int Ed Engl.* 2010; 49:9126–31. [PubMed: 20815002]
29. Yang R, Pasunooti KK, Li F, Liu XW, Liu CF. *Chem Commun (Camb).* 2010; 46:7199–201. [PubMed: 20737108]
30. Virdee S, Ye Y, Nguyen DP, Komander D, Chin JW. *Nat Chem Biol.* 2010; 6:750–7. [PubMed: 20802491]
31. El Oualid F, Merckx R, Ekkebus R, Hameed DS, Smit JJ, de Jong A, Hilkmann H, Sixma TK, Ovaa H. *Angew Chem Int Ed Engl.* 2010; 49:10149–53. [PubMed: 21117055]

32. Kumar KS, Spasser L, Ohayon S, Erlich LA, Brik A. *Bioconjug Chem.* 2011; 22:137–43. [PubMed: 21235224]
33. Kumar KS, Bavikar SN, Spasser L, Moyal T, Ohayon S, Brik A. *Angew Chem.* 2011; 50:6137–6141. [PubMed: 21591043]
34. Virdee S, Kapadnis PB, Elliott T, Lang K, Madrzak J, Nguyen DP, Riechmann L, Chin JW. *J Am Chem Soc.* 2011; 133:10708–11. [PubMed: 21710965]
35. For a recent example of using non-natural amino acids for making an isopeptide mimic see Eger S, Scheffner M, Marx A, Rubini M. *J Am Chem Soc.* 2010; 132:16337–9. [PubMed: 21033666]
36. Eddins MJ, Varadan R, Fushman D, Pickart CM, Wolberger C. *J Mol Biol.* 2007; 367:204–211. [PubMed: 17240395]
37. Datta AB, Hura GL, Wolberger C. *J Mol Biol.* 2009; 392:1117–24. [PubMed: 19664638]
38. For a first demonstration of using chemical ligation of recombinant proteins for selective segmental isotopic labeling of protein domains see Xu R, Ayers B, Cowburn D, Muir TW. *Proc Natl Acad Sci USA.* 1999; 96:388–393. [PubMed: 9892643]
39. Aimoto S. *Biopolymers.* 1999; 51:247–65. [PubMed: 10618594]
40. Studier FW. *Protein Expr Purif.* 2005; 41:207–34. [PubMed: 15915565]
41. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. *J Biomol NMR.* 1995; 6:277–93. [PubMed: 8520220]
42. Goddard, TD.; Kneller, DG. SPARKY3. University of California; San Francisco:
43. Streu C, Meggers E. *Angew Chem Int Ed Engl.* 2006; 45:5645–8. [PubMed: 16856188]
44. Xu P, Duong DM, Seyfried NT, Cheng D, Xie Y, Robert J, Rush J, Hochstrasser M, Finley D, Peng J. *Cell.* 2009; 137:133–45. [PubMed: 19345192]
45. Williamson A, Wickliffe KE, Mellone BG, Song L, Karpen GH, Rape M. *Proc Natl Acad Sci U S A.* 2009; 106:18213–8. [PubMed: 19822757]
46. Pervushin K, Riek R, Wider G, Wuthrich K. *Proc Natl Acad Sci USA.* 1997; 94:12366–71. [PubMed: 9356455]
47. Schanda P, Kupce E, Brutscher B. *J Biomol NMR.* 2005; 33:199–211. [PubMed: 16341750]
48. Huang H, Jeon MS, Liao L, Yang C, Elly C, Yates JR 3rd, Liu YC. *Immunity.* 2010; 33:60–70. [PubMed: 20637659]
49. Ai HW, Lee JW, Schultz PG. *Chem Commun (Camb).* 2010; 46:5506–8. [PubMed: 20571694]
50. Liu J, Castaneda CA, Wilkins BJ, Fushman D, Cropp TA. *Bioorg Med Chem Lett.* 2010; 20:5613–6. [PubMed: 20805030]
51. Tugarinov V, Hwang PM, Kay LE. *Annu Rev Biochem.* 2004; 73:107–46. [PubMed: 15189138]

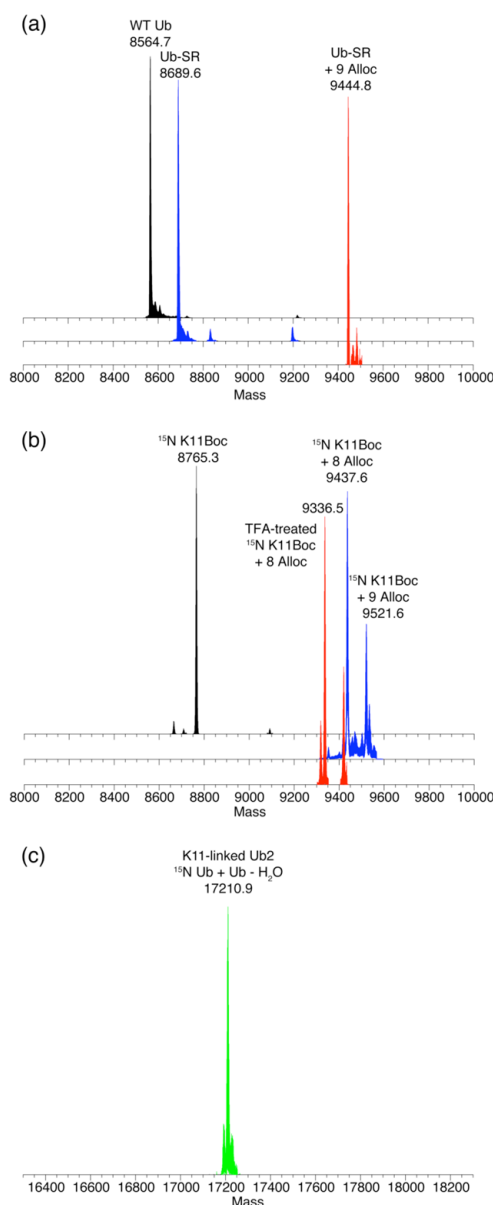


Figure 1.

ESI-MS spectra of the various steps in the non-enzymatic assembly of K11-linked Ub₂. (a) Preparation of Ub for incorporation as the distal unit in K11-linked Ub₂ ¹⁵N- labeled on the proximal Ub. The molecular weight of wild-type Ub is 8564 Da (black). Reaction with E1 and MESNA adds a C-terminal thioester functional group to Ub, increasing its molecular weight by 125 Da to 8689 Da (blue), and fully converts all Ubs into Ub-SR. Alloc protection of Ub-SR (red) adds a total of nine Alloc groups (each Alloc protecting group is 84 Da). (b) Preparation of ¹⁵N K11Boc Ub for incorporation as the proximal Ub. The molecular weight of ¹⁵N K11Boc Ub relative to WT Ub is increased by 200 Da to 8765 Da (black) as a result of ¹⁵N isotopic enrichment (100 Da) and addition of the Boc protecting group (100 Da). Alloc protection (blue) adds eight Alloc groups to the protein, one less than for Ub-SR because of the Boc protection on residue K11. TFA treatment (red) removes only the Boc group on K11, reducing the molecular weight by 100 Da to 9336 Da. (c) After chemical

condensation and complete Alloc deprotection, the expected molecular weight of purified K11-linked Ub₂ is 17210 Da as shown, resulting from the sum of one unlabeled Ub (8564 Da) and one ¹⁵N-labeled Ub (8664 Da) and the loss of one water molecule from the K11 isopeptide linkage. See also Fig. S10.

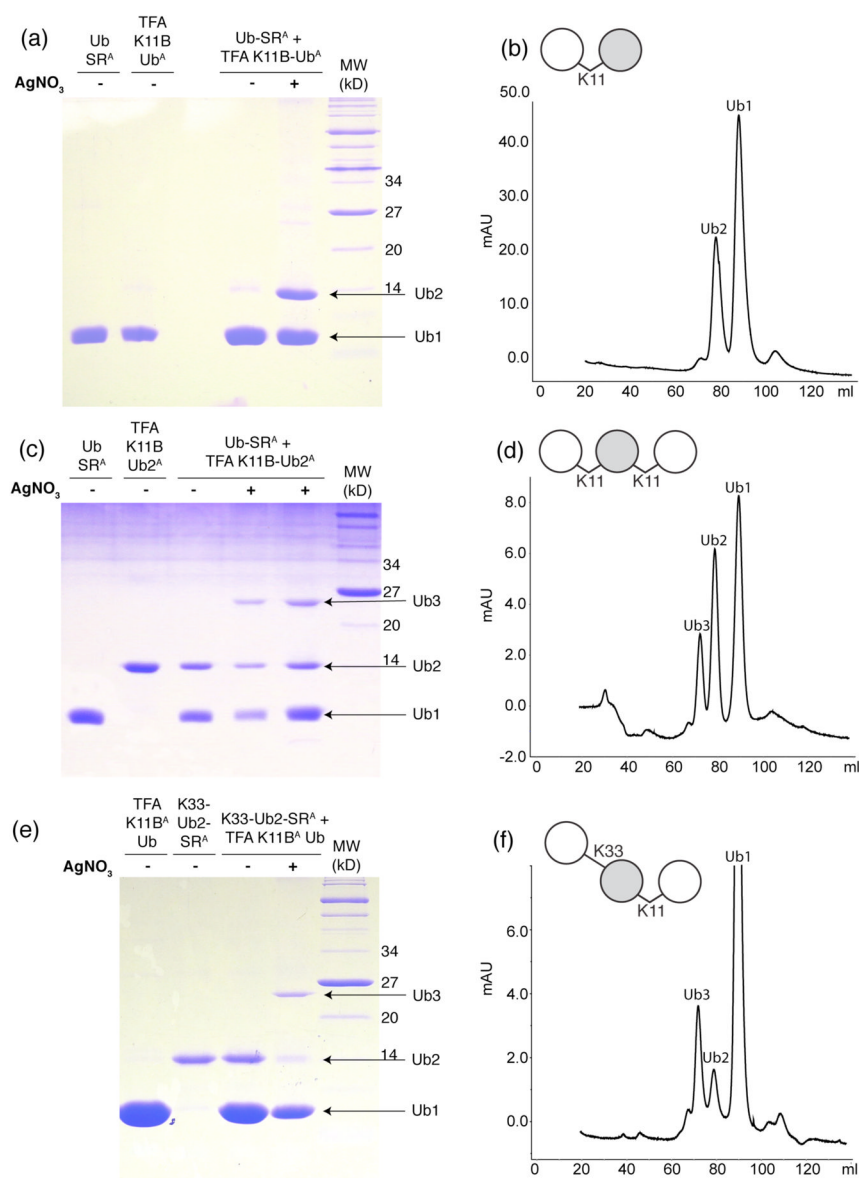


Figure 2. Chemical assembly of K11-linked di- and tri-Ub chains. (a,c,e) Coomassie-stained 15% SDS PAGE gels of the chemical condensation reactions of (a) K11-linked Ub₂ ¹⁵N-labeled on the proximal Ub, (c) K11-linked Ub₃ with the middle Ub unit ¹⁵N-labeled, and (e) K33,K11-linked Ub₃ with the middle Ub unit ¹⁵N-labeled. K11-linked Ub₂ (panel a) was assembled from Alloc-protected monomers of Ub-SR (Ub-SR^A) and TFA-treated ¹⁵N-labeled K11Boc Ub (TFA K11B Ub^A). K11-linked Ub₃ (panel c) was assembled from Alloc-protected Ub-SR and distal-¹⁵N-labeled K11-linked Ub₂ (TFA K11B Ub₂^A) whose distal-K11 side chain was made available for ligation by TFA treatment. K33,K11-linked Ub₃ (panel e) was assembled from Alloc-protected proximal-¹⁵N-labeled K33-linked Ub₂ (K33-Ub₂-SR) and TFA-treated K11Boc Ub (TFA K11B Ub^A). Formation of Ub₂ or Ub₃ is observed after 20 hours of reacting the proteins with AgNO₃, H-OSu, and DIEA. 0.3 – 0.4 μL are loaded from the reaction directly onto gel. Notations used here: “TFA” indicates that the corresponding protein was treated with TFA to remove the Boc group in order to make

the corresponding Lys available for the ligation reaction, while the superscript “A” indicates that the protein is Alloc-protected.

(b,d,f) Size-exclusion chromatograms of (b) K11-linked Ub₂, (d) homogeneously K11-linked Ub₃, and (f) mixed-linkage K33,K11-linked Ub₃ performed after Alloc deprotection and protein renaturation to purify the desired product from the unreacted components. The ability to separate Ub₂ or Ub₃ products from unreacted species is illustrated in Fig. S8.

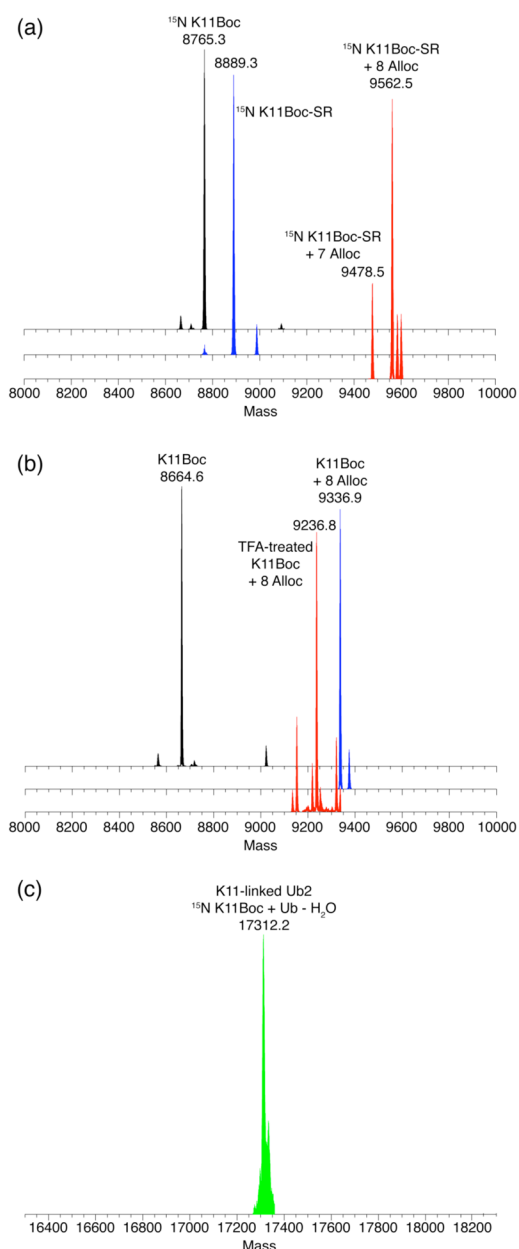


Figure 3.

ESI-MS spectra of various Ub components in the assembly of K11-linked Ub₂ with the distal Ub unit both ^{15}N -labeled and K11Boc-protected. (a) Preparation of ^{15}N K11Boc Ub for incorporation as the distal unit in distal- ^{15}N -labeled K11-linked Ub₂. The molecular weight of ^{15}N K11Boc is 8764 Da (black). Reaction with E1 and MESNA adds a C-terminal thioester functional group to the Ub protein, increasing its molecular weight by 125 Da to 8889 Da (blue), and fully converts all Ubs into ^{15}N -K11Boc-SR. Alloc protection of ^{15}N -K11Boc-Ub-SR (red) adds a total of eight Alloc groups (each Alloc protecting group is 84 Da). (b) Preparation of K11Boc Ub for incorporation as the proximal Ub. The molecular weight of K11Boc Ub is 8665 Da (black) as a result of addition of the Boc protecting group (100 Da). Alloc protection (blue) adds eight Alloc groups to the protein, similar to the distal ^{15}N -K11Boc-SR Ub protein. TFA treatment (red) removes only the Boc group on

K11, reducing the molecular weight by 100 Da to 9236 Da. (c) After chemical condensation and complete Alloc deprotection, the expected molecular weight of the purified distal-¹⁵N-labeled and distal-K11Boc K11-linked Ub₂ is 17310 Da as shown, resulting from the sum of one Ub (8564 Da) and one ¹⁵N-labeled K11Boc Ub (8764 Da) and the loss of one water molecule from the K11 isopeptide linkage.

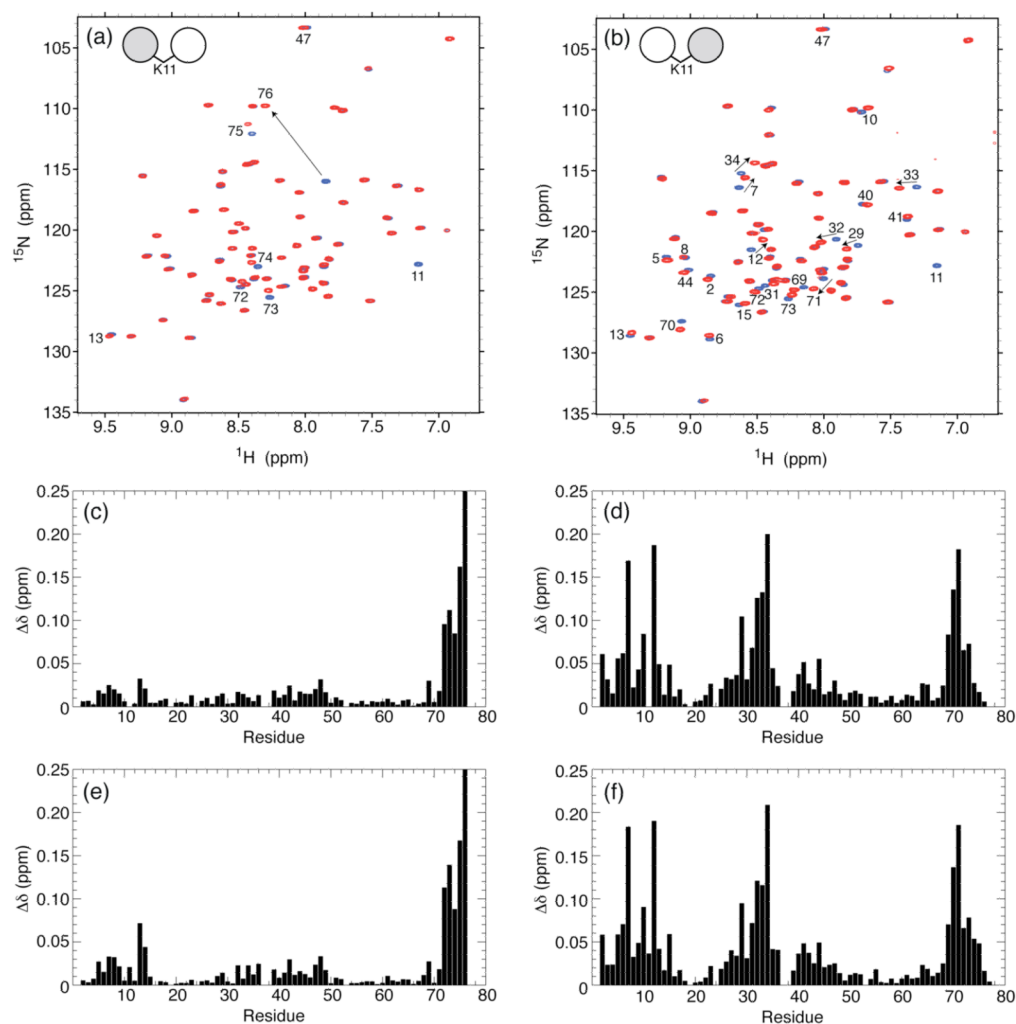
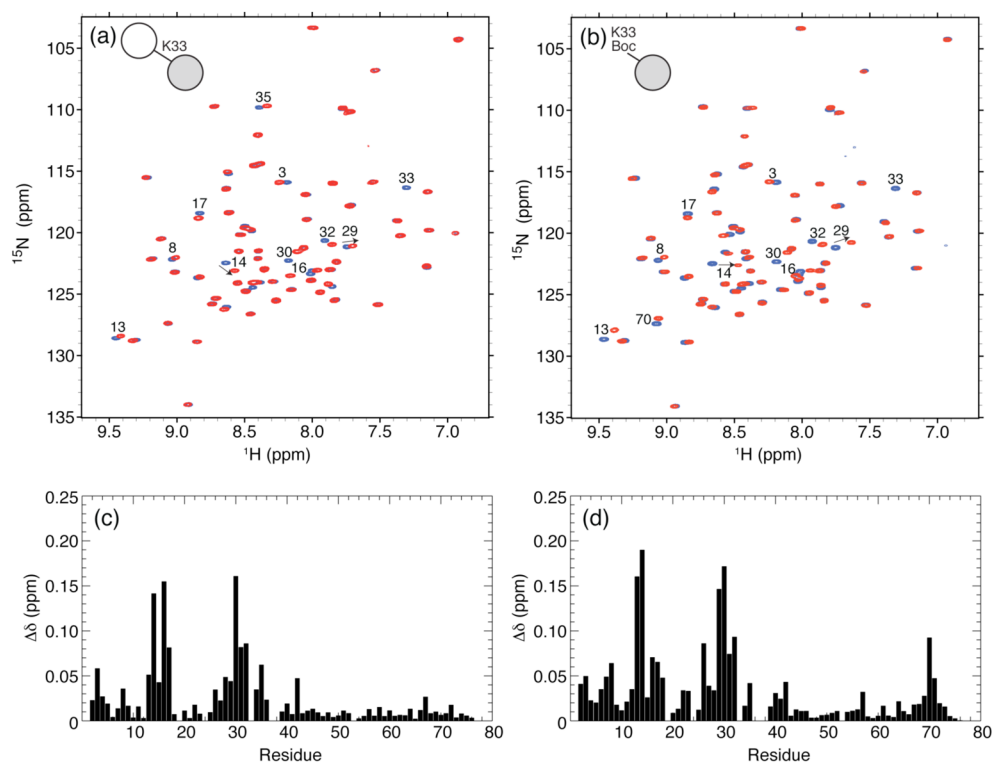


Figure 4. Overlay of ^1H - ^{15}N TROSY-HSQC spectra of (a) the distal Ub and (b) the proximal Ub in all-natural K11-linked Ub₂ (red) and of WT Ub (blue). The spectral differences between K11-linked di-Ub and mono-Ub, quantified as amide chemical shift perturbations (CSPs), are plotted as a function of the residue number for (c,e) distal Ub and (d,f) proximal Ub in (c,d) all-natural K11-linked Ub₂ and (e,f) enzymatically-synthesized K11-linked Ub₂. The CSPs were calculated as $\Delta\delta = [(\Delta\delta_{\text{H}})^2 + (\Delta\delta_{\text{N}}/5)^2]^{1/2}$, where $\Delta\delta_{\text{H}}$ and $\Delta\delta_{\text{N}}$ are chemical shift differences for ^1H and ^{15}N , respectively. Residues with significant CSPs are indicated on the spectra in panels a and b. Note that the absence of the K11 backbone amide signal in the spectra of K11-linked Ub₂ (panels a and b) serves as a direct confirmation of the incorporation of the unlabeled Lys at residue 11.

**Figure 5.**

(a) Overlay of ^1H - ^{15}N TROSY spectra of the proximal Ub in K33-linked Ub_2 (red) and WT monoUb (blue). (b) Overlay of ^1H - ^{15}N TROSY spectra of K33Boc Ub (red) and WT Ub (blue). (c,d) The spectral differences between residues in the (c) proximal Ub of K33-linked Ub_2 and WT Ub and in (d) K33Boc Ub and WT Ub are quantified as amide chemical shift perturbations (CSPs). Residues with significant CSPs are indicated on the spectra in panels a and b. The absence of the K33 backbone amide signal in the spectra of K33-linked Ub_2 and K33Boc Ub (panels a and b, respectively) serves as a direct confirmation of the incorporation of the unlabeled Lys at residue 33.

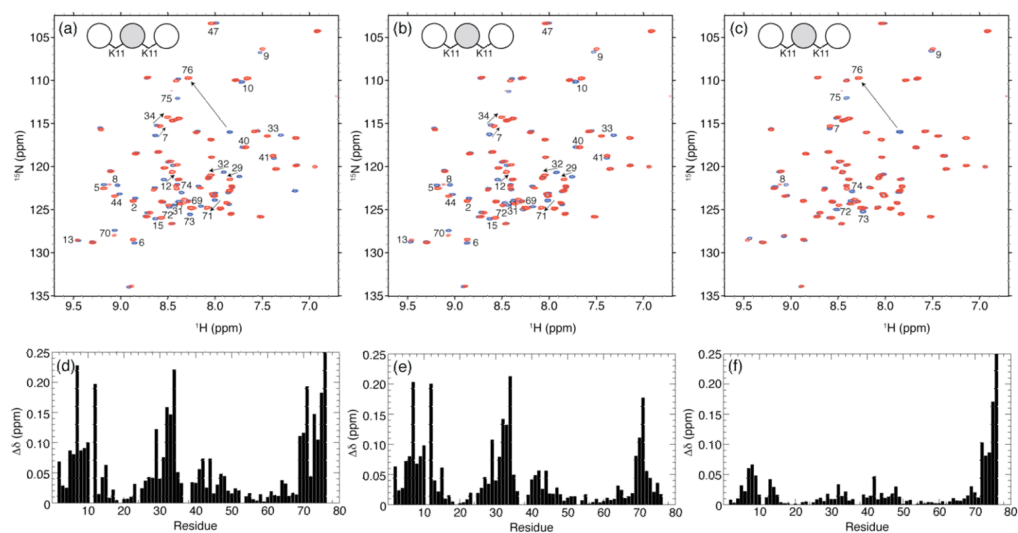


Figure 6. Overlay of ^1H - ^{15}N TROSY spectra of the middle Ub in all-natural homogeneously K11-linked Ub_3 (red) and in blue, (a) WT Ub, (b) the distal Ub of K11-linked Ub_2 , and (c) the proximal Ub of K11-linked Ub_2 . The spectral differences between residues in the middle Ub unit of K11-linked Ub_3 and (d) WT Ub, (e) distal Ub of K11-linked Ub_2 , and (f) proximal Ub of K11-linked Ub_2 are quantified as amide chemical shift perturbations (CSPs). Residues with significant CSPs are indicated on the spectra in panels a–c.

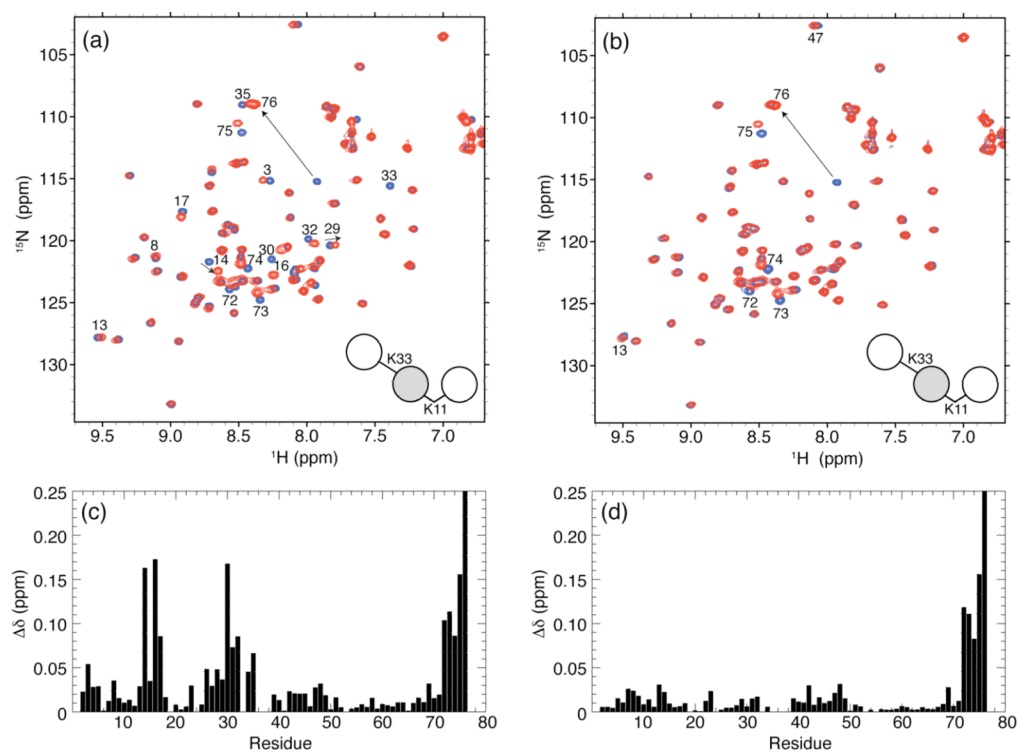


Figure 7. Overlay of ^1H - ^{15}N SOFAST spectra of the middle Ub in all-natural mixed K33,K11-linked Ub₃ (red) and in blue, (a) WT Ub, (b) the proximal Ub of K33-linked Ub₂. The spectral differences between residues in the middle Ub unit of K33,K11-linked Ub₃ and (c) WT monoUb and (d) proximal Ub of K33-linked Ub₂ are quantified as amide chemical shift perturbations (CSPs). Residues with significant CSPs are indicated on the spectra in panels a and b.

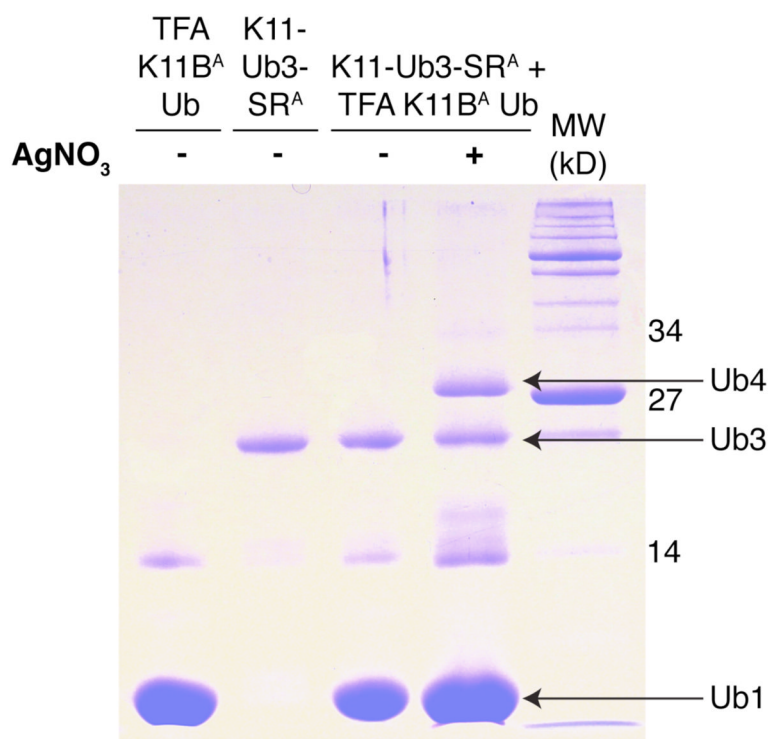
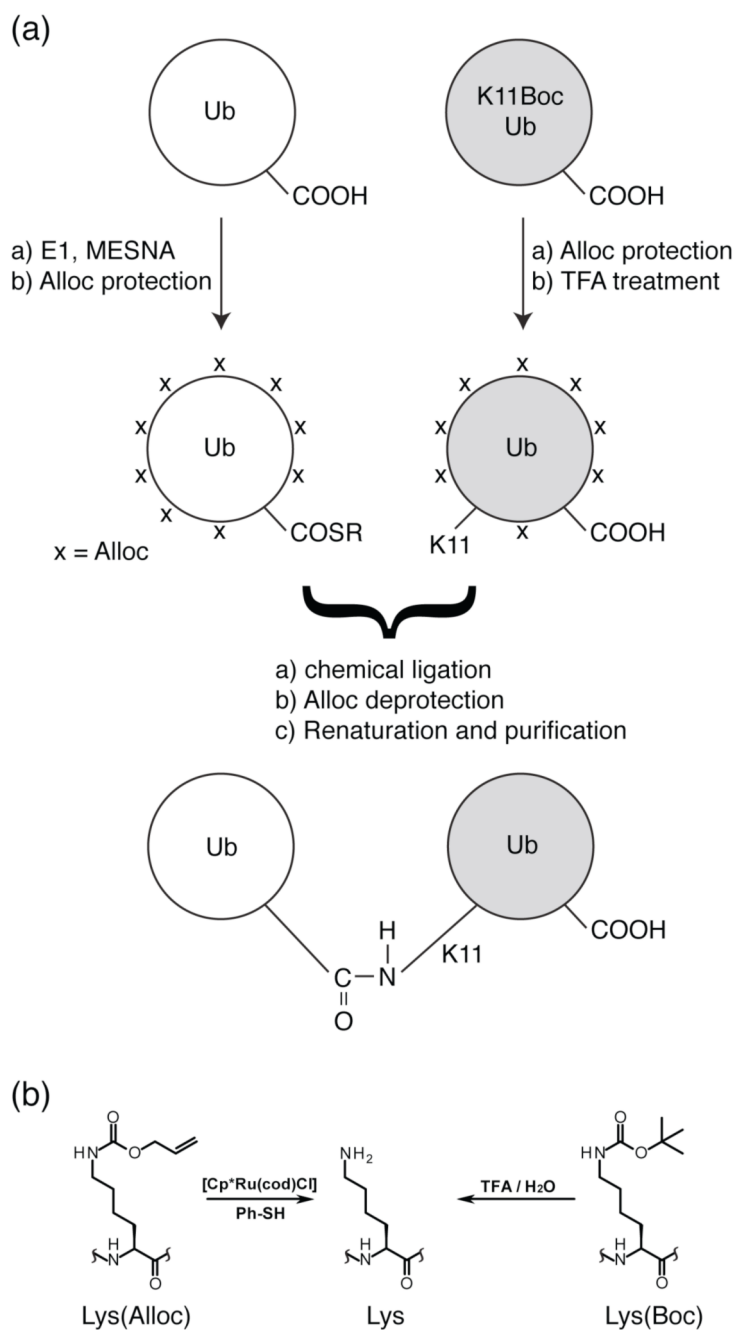
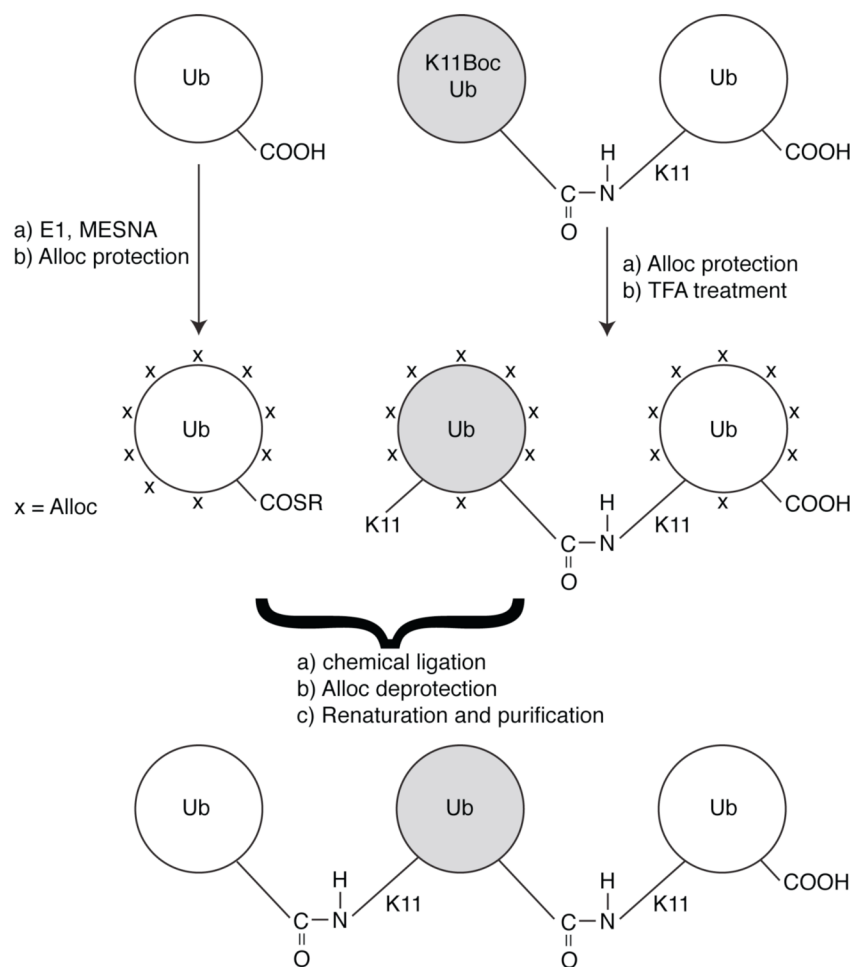


Figure 8.

Coomassie-stained urea SDS-PAGE gel demonstrating the formation of homogeneously K11-linked Ub₄ via a 20-hour incubation of Alloc-protected middle-¹⁵N-labeled K11-linked Ub₃-SR and TFA-treated, Alloc-protected K11Boc Ub monomer (TFA K11Boc Ub^A). The contaminating Ub₂ band from an impure sample of K11Boc is inert as it is visible both before and after the ligation reaction. The sample after ligation was loaded at higher concentration to emphasize the formation of Ub₄.

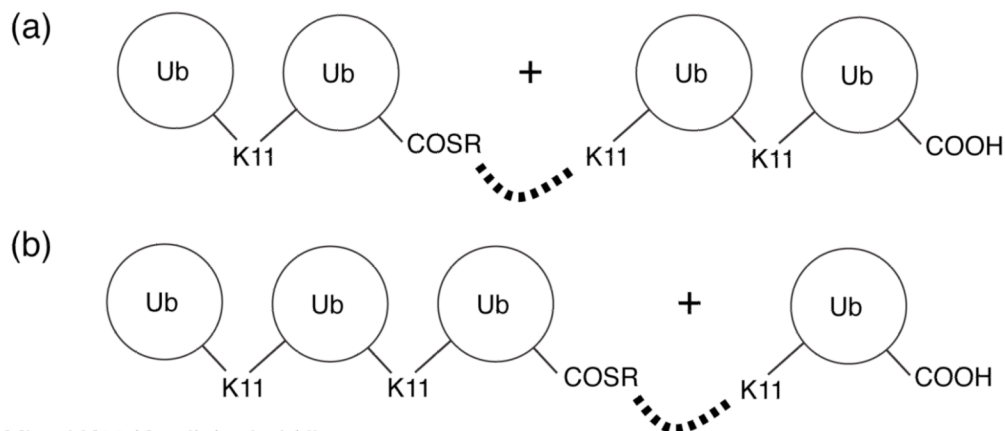
**Scheme 1.**

(a) Assembly of K11-linked Ub₂ from recombinant Ub monomers. The two Ubs are shaded differently to illustrate the concept of unit-specific isotopic labeling. In particular, the Ub₂ chain with an isotope-labeled proximal Ub (shaded gray) is assembled from the unlabeled Ub monomer and isotope-labeled K11Boc Ub. Similarly, Ub₂ with a labeled distal Ub is made from an isotope-labeled Ub and unlabeled K11Boc Ub. The same concept applies to making Ub₂ chains linked via other lysines (by incorporating Lys(Boc) at other Lys positions) and/or with both Ub units labeled but with different isotopes (²H, ¹⁵N, ¹³C or combination thereof) uniformly or using different residue- or group-specific labeling schemes. (b) Structure of Alloc and Boc protecting groups on a lysine residue.

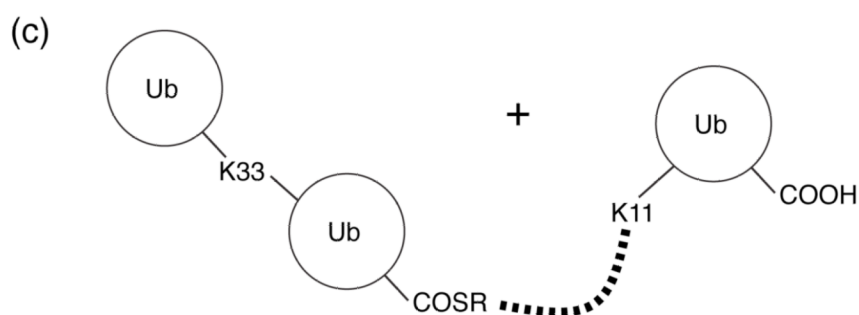
**Scheme 2.**

Assembly of homogeneously K11-linked Ub₃ from Ub monomer and K11-linked Ub₂. To illustrate the concept of unit-specific isotope labeling, the middle Ub unit (highlighted in gray) in this chain is isotope-labeled, as in the current study.

Homogeneous K11-linked tetra-Ub



Mixed K33,K11-linked tri-Ub

**Scheme 3.**

Schematics on the assembly of long Ub chains with either homogeneous or mixed linkages. Shown in (a,b) are two methods to form a homogeneously K11-linked Ub₄ from either (a) two K11-linked dimers, or (b) a combination of a homogeneously K11-linked Ub₃ and K11Boc Ub monomer. (c) The ability to make Ub chains with mixed linkages can be easily exploited with our present method. Here we illustrate the formation of a mixed K33,K11-linked Ub₃ from a K33-linked Ub₂ and K11Boc Ub monomer. The dotted line indicates the isopeptide bond that would be formed during the ligation reaction. Longer Ub chains can be assembled iteratively by using the same concepts as illustrated here.