universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Tailor-made ubiquitin chains“

verfasst von / submitted by

Marina Kujundzic BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2021 / Vienna 2021

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 066 862
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Chemie
Betreut von / Supervisor:	Univ.-Prof. Dr. Christian Friedrich Wilhelm Becker

Acknowledgement

First and foremost, I would like to thank Professor Christian Becker for not only giving me the opportunity to work in his group, but also for his guidance and support during the experimental work and writing of this thesis.

Furthermore, I would like to thank Mag. Gerhard Niederacher for teaching me the molecular biology needed in this thesis, his advice, and the friendly working atmosphere. Also, I am very thankful to Somnath Mukherjee, PhD and Maximilian Schrems, MSc for their supervision and advice during the organic synthesis work. I would also like to give special thanks to Alanca Schmid, MSc for always helping me out in the lab, even with little things. Of course, I also would like to thank the rest of the working group for their companionship and support.

Finally, I would like to thank my parents, for enabling my studies and their support during hardship, my brother for always helping me with graphics and layout issues and my partner for his meals and care during this master thesis.

Table of contents

List of abbreviations

AA	Allylamide
ACN	Acetonitrile
APS	Ammonium persulfate
Boc	Tert-butyloxycarbonyl
CBD	Chitin binding domain
CuAAC	Cu(I)-catalysed azide-alkyne cycloaddition
DTT	Dithiothreitol
DI	Direct injection
DUB	Deubiquitinase
Fmoc	Fluorenylmethyloxycarbonyl
GOPAL	Genetically encoded orthogonal protection and activated ligation
IPTG	Isopropyl β-D-1-thiogalactopyranoside
KxC	Lysine to cysteine mutation at position x
LAP	Lithium phenyl-2,4,6 trimethylbenzoylphosphinate
LB	Lysogeny broth
MesNa	2-Mercaptoethane sulfonate
MPAA	4-Mercaptophenylacetic acid
MSP	Monosodium phosphate
NCL	Native chemical ligation
PA	Propargylamide
PBS	Phosphate-buffered saline
PTM	Post-translational modification
SPPS	Solid phase peptide synthesis
TCEP	Tris(2-carboxyethyl)phosphin
TAE	Tris-acetate-EDTA
TEMED	Tetramethylethylenediamine
TFA	Trifluoroacetic acid
TBS	Tris-buffered saline
Tris	Tris(hydroxymethyl)aminomethane
Ub	Ubiquitin
UBD	Ubiquitin-binding domain

Abstract

Ubiquitination is a post-translational modification (PTM) involved in the regulation of a wide range of cellular functions. This versatility results from the many possibilities of ubiquitin arrangement, like for example polyubiquitination of target proteins. To understand the impact of these differences in such arrangement, many synthetic and semi-synthetic strategies to access all kinds of ubiquitin variants and conjugates have been developed. This master thesis aimed to create ubiquitin-building blocks to facilitate the assembly of ubiquitin chains using different synthesis strategies. This was achieved by recombinant expression of a ubiquitin-intein-fusion construct, intein-mediated cleavage, and further conversion of the ubiquitin C-terminus. Moreover, point mutations of lysine to cysteine in ubiquitin were performed, to enable thiol-ene coupling of ubiquitins. The organic synthesis of a δ-thiolysine was also attempted to generate a building block for SPPS of ubiquitin. In summary, ten different ubiquitin building blocks were generated, which can be used for Cu(I)-catalysed azide-alkyne cycloaddition reaction or thiol-ene coupling, to potentially generate ubiquitin dimers and chains. During this master thesis thiol-ene chemistry was already used to generate a di-ubiquitin from ubiquitin-allylamide and ubiquitin(K48C)-hydrazide. Despite this achievement, further improvement of the thiol-ene reaction is needed to increase product yield.

Zusammenfassung

Ubiquitylierung ist eine posttranslationale Modifikation (PTM), die an der Regulation einer Vielzahl von zellulären Funktionen beteiligt ist. Diese Vielseitigkeit resultiert aus den vielen Möglichkeiten der Anordnung von Ubiquitin, wie zum Beispiel der Polyubiquitylierung von Zielproteinen. Um die Auswirkungen dieser unterschiedlichen Anordnungen zu verstehen, wurden viele synthetische und semi-synthetische Strategien entwickelt, um alle möglichen Ubiquitin-Varianten und -Konjugate zu erzeugen. Ziel dieser Masterarbeit war es, Ubiquitin-Bausteine zu kreieren, um den Aufbau von Ubiquitin-Ketten durch verschiedene Synthesestrategien zu ermöglichen. Dies wurde durch rekombinante Expression eines Ubiquitin-Intein-Fusionskonstrukts, Intein-mediated cleavage und weitere Umsetzung des Ubiquitin-C-Terminus erreicht. Außerdem wurden Punktmutationen von Lysin zu Cystein im Ubiquitin durchgeführt, um Thiol-En-Kopplungen von Ubiquitinen zu ermöglichen. Weiters wurde die organische Synthese eines δ-Thiolysins angestrebt, um einen Baustein für SPPS von Ubiquitin zu generieren. Zusammenfassend wurden zehn verschiedene Ubiquitin-Bausteine generiert, die für Cu(I)-katalysierte Azid-Alkin-Cycloadditionsreaktionen oder Thiol-En-Kupplungen verwendet werden können, um Ubiquitin-Dimere und -Ketten zu erzeugen. Im Rahmen dieser Masterarbeit wurde Thiol-En-Chemie bereits zur Erzeugung eines Ubiquitin-Dimers aus Ubiquitin-Allylamid und Ubiquitin(K48C)-Hydrazid eingesetzt. Trotz dieses Erfolges ist eine weitere Verbesserung der Thiol-En-Reaktion erforderlich, um die Produktausbeute zu erhöhen.

1. Introduction

1.1. Ubiquitin

During their life span, most proteins undergo various changes, as they are modified by post-translational modifications (PTMs), which impacts their structure, function and cellular localization (1). One of the most significant PTMs is called ubiquitination. Through ubiquitination, ubiquitin (Ub), a 76aminoacidlong protein, or a polyubiquitin chain, consisting of linked ubiquitins, is covalently attached to a target protein. Depending on the resulting architecture of the attached ubiquitin(s) the biological fate of the target protein is changed in various ways.

Ubiquitination is mediated by a cascade of three enzymes: E1-ubiquitin-activating enzymes, E2-ubiquitin-conjugating enzymes and E3-ubiquitin ligases (2). Firstly, ubiquitin is activated at its C-terminal glycine residue by E1 under ATP-consumption, resulting in a E1-Ub-thioester at the catalytic cysteine of E1. Ubiquitin is then transferred to E2 *via* trans-thioesterification, resulting in a E2-Ub-thioester. Finally, E3 facilitates the transfer of ubiquitin to the protein's ε-amino group of the target lysine residue, thereby forming an isopeptide bond. *Via* this procedure ubiquitin chains are assembled, and mono-ubiquitin or ubiquitin chains are attached to proteins (3). Importantly, which kind of ubiquitin chain is generated, and which protein is ubiquitinated, depends on the combination of E2 and E3 enzymes. Currently, two E1, around 40 E2 and over 600 E3 enzymes have been identified in humans (2, 4).

The first clearly assigned function of ubiquitination was the Ub-mediated proteasomal degradation of proteins (5). However, nowadays a wide range of events are known to be regulated by ubiquitination, such as intracellular trafficking, DNA repair, cell division and inflammation (3). This versatility is due to the high complexity resulting from the variable architecture of ubiquitinated proteins.

The first layer of complexity begins at the ubiquitin level since di-ubiquitin formation can occur at any of eight primary amines: On the N-terminal amino group of M1 or the ε-amino group of K6, K11, K27, K29, K33, K48, and K63 (1, 3). All types of ubiquitin linkages have been found to co-exist in cells (6). Of these eight linkage types K48-linked and K63-linked are the most common, with K48-linked chains regulating the proteasomal degradation of proteins and K63-linked chains regulating non-degradative signals like intercellular trafficking (3). Chains resulting from linkages at the same lysine position are called homotypic (2). Ubiquitin chains are recognized by ubiquitin-binding domains (UBDs), with some UBDs exhibiting linkage specificity. Linkage-specific UBDs have been identified for five out of the eight linkage types (3).

The complexity is further increased by heterotypic chains. These chains have a mixed or branched architecture, meaning that either ubiquitins are attached to each other at different lysine residues or one ubiquitin is ubiquitinated at multiple different lysine residues, respectively (1). A mixed architecture seems to regulate several non-degradative processes, while branched architectures have been linked to proteasomal degradation as their structures demonstrate high affinity towards the proteasome. Additionally, target proteins can have multiple anchoring sites for ubiquitination, leading to constructs with mono-ubiquitination, polyubiquitination and multiple mono-ubiquitination (7). Recently, another layer of complexity in ubiquitin signaling has been uncovered, as ubiquitin itself can undergo further PTMs apart from ubiquitination, such as phosphorylation and acetylation (1). What kind of physiological

roles they fulfill is mostly unknown and currently a hot research topic. Ultimately, ubiquitination is a reversible PTM. Enzymes named deubiquitinases (DUBs) can cleave ubiquitin chains and ubiquitins of proteins. About 100 DUBS have been identified so far in humans (8), with some DUBs demonstrating high linkage specificity (3).

Research efforts to decipher the complexity of ubiquitination have been steadily increasing as new insights into the topic further uncover the significance of ubiquitin in a multitude of diseases such as cancer and neurodegeneration (1). One of the first methods to create ubiquitin chains for studying the effects of different ubiquitin architectures was by enzymatic means *via* linkage-specific pairings of E2 and E3 (8). This strategy was, however, limited as many E2-E3 pairs were not known, especially in the case of atypical linkage types, and the resulting products often lack homogeneity (1). Consequently, non-enzymatic methods like chemical synthesis and semi-synthesis have been employed to expand the toolbox for ubiquitin-conjugate creation.

1.2. Synthetic and semi-synthetic strategies to access Ub

1.2.1. Native isopeptide bond

1.2.1.1. Native chemical ligation

Native chemical ligation (NCL), developed by Dawson *et al.*, is a synthesis strategy that has vastly facilitated chemical protein synthesis, as it enabled the total synthesis and semi-synthesis of hundreds of proteins beyond the limits of classic solid phase peptide synthesis (SPPS) (9). Unsurprisingly, NCL has also helped the understanding of ubiquitination as it has been employed in the creation of ubiquitin, ubiquitin chains and ubiquitin conjugates (8). In the originally reported NCL reaction an N-terminal cysteine residue and a C-terminal thioester react *via* trans-thioesterification and subsequent S-to-N-acyl shift, resulting in an amide bond (10, 11). This synthesis method has been expanded upon to tend to different synthesis problems over the years (10). For the creation of ubiquitin conjugates and chains, one of the most popular NCL strategies is based on the incorporation of a γ-thiolysine or δ-thiolysine moiety at the target lysine residue which can then reacts with a thioester moiety of the target protein (8). The syntheses of γ- and δ-thiolysine have been established independently by the Liu group and Brik group, respectively (12, 13). Thiolysine is incorporated into ubiquitin *via* SPPS, with the ubiquitin synthesized either by linear synthesis or assembled from two fragments (8). When synthesizing a ubiquitin dimer, a thioester needs to be introduced at its C-terminus to facilitate NCL with a thiolysine-containing ubiquitin. This thioester moiety can be introduced *via* different methods, for example by E1-mediated enzymatic conversion with MesNa, during Fmoc-based SPPS or *via* cleavage of an intein-fusion construct with MesNa (14). Finally, NCL is carried out and afterward a desulfurization step is employed, resulting in a ubiquitin conjugate linked through a native isopeptide bond (8). One of the drawbacks of the thiolysine strategy is the lengthy synthesis of thiolysine (1). Recently, the Liu group reported an alternative NCL strategy which neither needs γ- or δ-thiolysine (15). The assembly of atypical ubiquitin chains is carried out in a modular manner facilitated by premade isopeptide linked Ub isomer, bearing an N-terminal cysteine and a C-terminal hydrazide. Another recent development was the establishing of γ-thionorleucine, a thiolated amino acid which is incorporated at the M1 position of ubiquitin, enabling the incorporation of N-terminal ubiquitin (16).

1.2.1.2. Genetic code expansion

The utilization of expressed proteins is a very convenient strategy to generate large amounts of proteins such as ubiquitin (8). Genetic code expansion with unnatural amino acids has proven itself to be a useful tool in introducing novel reactive groups for ligation of recombinant proteins. For example, GOPAL, genetically encoded orthogonal protection and activated ligation, developed by the Komander and Chin groups, is the first reported method to combine genetic code expansion with chemoselective ligation to enable site-selective isopeptide bond formation in ubiquitin conjugates (2, 17). With this strategy K6- and K29-linked di-ubiquitins were synthesized. However, the GOPAL strategy involves multiple protection and deprotection steps that can result in poor yields (2). Subsequently to GOPAL the Chin group has developed a method to incorporate δ-thiolysine site-specifically into recombinant proteins by using newly engineered tRNA-synthetase/tRNA pair (2, 18). With this method, recombinant ubiquitin containing δ-thiolysine was ligated *via* NCL to recombinantly expressed Ub-thioester to generate K6-linked di-ubiquitin.

1.2.1.3. Auxiliary-mediated chemical ubiquitination

Using auxiliary-mediated ligation the first ubiquitinated peptide was generated by the Muir group (14). With this method a recombinant Ub(1-75)-α-thioester was ligated to a 11-mer peptide from the human H2B histone, resulting in a native isopeptide bond once the auxiliary was removed by UV light (14). Based on auxiliary-mediated ligation two other ubiquitination methods have been developed (1). Firstly, Weller *et al.* have reported a method applying 2-(aminooxy)ethanethiol as an auxiliary, which is cleaved by zinc under acidic conditions (19). With this method a ubiquitin conjugate was created allowing the generation of a native or non-native isopeptide bond, depending on whether the auxiliary is cleaved. Such a type of linkage is useful for biophysical and biochemical studies. Another method developed by Liu *et al.* utilizes the TFA-labile glycyl auxiliary 1-(2,4- dimethoxyphenyl)-2-mercaptoethyl (20). With this auxiliary K27-linked di- and tri-Ub and ubiquitinated H2B histone at K120 (21) have been synthesized. Furthermore, the combination of the glycyl auxiliary with GOPAL was used to create K48-linked and K6-linked tri-Ub and mixed-linkage K6, K48-linked tri-Ub (22).

1.2.2. Non-native isopeptide bond

Non-native isopeptide bonds are of interest as some information cannot be obtained with native isopeptide bonds, for example in cases where susceptibility to DUBs is counterproductive (1). Such cases include studies of DUBs interaction with Ub conjugates and DUBs structure or studying Ub conjugate interactions with specific proteins without the interference of degradation by DUBs in cells. Additionally, the synthesis of Ub conjugates linked *via* non-native bonds is often easier compared to native linkage and some hydrolysable non-native linkages have been developed, which mimic the native bond well enough that linkage selective DUBs can cleave them (23).

1.2.2.1. Oxime-linked ubiquitin conjugates

Oxime linkages have been used to create non-hydrolysable ubiquitin conjugates (24, 25). They occur as the result of condensation reactions between aldehyde and aminooxy moieties and are sufficiently stable under physiological conditions. When used for linking ubiquitin and a target protein the oxime linkage isosterically mimics a native isopeptide bond and is stable towards DUBs (24). The required reactive groups can be installed interchangeable at the ubiquitin or target protein and do not interfere with other types of ligation methods as the reactive groups of the oxime reaction are chemoselective (24, 25). With this synthesis strategy UbK48 and UbK63 isopeptide isosteres, ubiquitin conjugates with

parts of PCNA, H2B, H2A and FANCD2 have been synthesized. For the study of DUBs and proteasomal degradation, native and non-native bonds in di- to tetra-Ub-α-globin and bis(diUb)β-globin have been synthesized with the oxime linkage installed proximal to the target protein (24).

1.2.2.2. CuAAC chemistry

Another useful strategy to generate non-hydrolysable ubiquitin conjugates is the Cu(I)-catalysed azide-alkyne cycloaddition (CuAAC) reaction (26, 27). For this type of reaction, an azide and alkyne moiety incorporation can be achieved in various ways: Ubiquitin and other proteins can be mutated through the expansion of the genetic code to incorporate amino acids containing the azide or alkyne moiety (28). For mono-ubiquitination the azide-containing group is placed at the C-terminus of ubiquitin and the alkyne-containing amino acid at the target site of the target protein. Further, both azide and alkyne-containing amino acids can be incorporated into ubiquitin simultaneously, enabling the creation of ubiquitin chains and poly-ubiquitinated proteins. Another approach to generate linkages *via* click reaction was reported by Mootz *et al.*, where *via* genetic code expansion an azide containing amino acid was incorporated into the target protein (29). The alkyne moiety was induced in ubiquitin *via* the expression of a ubiquitin-intein construct, cleaved to generate an Ub-thioester which was reacted with propargylamine. Alternatively, E1-mediated enzymatic conversion can be employed to install the azide or alkyne at the C-terminus of ubiquitin (30).

1.2.2.3. Cysteine chemistry

Cysteine chemistry has proven itself to be one of the most versatile ways to generate site-specific non-native isopeptide bonds for ubiquitin conjugation. Naturally ubiquitin has no cysteine in its amino acid sequence, which makes cysteine chemistry a suitable ligation strategy when cysteine is introduced into ubiquitin. This method is especially useful for the creation of polyubiquitins. For the conjugation of ubiquitin with other proteins, the occurrence of cysteine in the target protein can limit the options of cysteine chemistry. Notably, depending on the ligation method non-native isopeptide bonds with hydrolysable or non-hydrolysable linkages can be created. In the case of non-hydrolysable linkages, dichloroacetone, bromoacetamide, and maleimide have been used to generate ubiquitin conjugates with target proteins (31).

An approach to prepare hydrolysable ubiquitin-conjugates is to generate disulfide bonds replacing the native isopeptide bond (1). Although the disulfide strategy is a highly efficient method to generate a non-native isopeptide linkage, the use of disulfides bears its own limitations. Firstly, the disulfide bond can be easily reduced in biological environments, which limits the variety of biological experiments possible. Secondly, this strategy is limited to proteins that contain only one available cysteine.

Another method to prepare hydrolysable ubiquitin-conjugates is using thiol-ene chemistry for the synthesis of ubiquitin chains (23). For this method, the target lysine for ubiquitination is replaced with cysteine and coupled with ubiquitin in which allylamine was introduced at the C-terminus *via* the UCH YUH1 enzyme. Wang *et al.* have also shown that the allylamine functionality can be introduced *via* E1-mediated enzymatic conversion (30). The thiol-ene reaction is induced by a radical starter, lithium acyl phosphinate (LAP) and UV-irradiation, leading to the formation of a non-native isopeptide linkage with an additional sulfur atom. This type of linkage has been shown to be processed by DUBs and a multitude of ubiquitin chains have been created with this method, including all lysine-linked ubiquitin dimers (23), some higher-order chains and even branched chains (23, 32).

1.3. Objective

This master thesis elaborates on the creation of ubiquitin-building blocks to facilitate the assembly of ubiquitin chains using different synthesis strategies. The building blocks were created for the assembly of ubiquitin chains with linkages between K48 and/or K63, with a maximal chain length of four ubiquitins. To attain a variety of building blocks, molecular biology and chemical methods were employed. Ubiquitin was chosen to be generated *via* recombinant expression of an intein fusion construct, enabling the easy introduction of different C-terminal modifications *via* intein cleavage. The resulting ubiquitin variations can be further chemically modified at the C-terminus to enable a wider range of assembly possibilities. The point mutation of lysine to cysteine in ubiquitin was also performed to enable thiol-ene coupling. Lastly, organic synthesis of a δ-thiolysine was attempted to generate a building block for SPPS of ubiquitin.

2. Materials and methods

2.1. Chemicals

All chemicals and solvents were obtained from commercial sources and used without further purification. Restriction enzymes were obtained from New England Biolabs, components of the expression media were produced by PanReac AppliChem. ACN, TFA, petroleum spirit and diethyl ether were purchased from VWR. Ethyl acetate was produced by Acros Organics. MesNa was obtained from Sigma Aldrich and Fluorochem, hydrazine monohydrate and DMAP were purchased from Fluka and L-glutamic acid was produced by Roth. Allylamine, propargylamine, TCEP, MPAA and all other chemicals for organic synthesis were obtained from Sigma Aldrich. Deionized Water (Milli-Q Reference A+) was used for all buffer preparations and experiments.

2.2. Molecular biology

2.2.1. General protocols

2.2.1.1. Agarose gel electrophoresis

For the preparation of a 1% agarose gel, 0.5 g agarose was dissolved in 50 ml TAE buffer (40 mM Tris, 20 mM acetate, 1 mM EDTA, pH 8.4) by heating in the microwave. After a short cooling phase to room temperature, the gel was poured. For migration tracking during electrophoresis, the samples were combined with DNA Gel Loading Dye Purple (6X) and 12 µl of dyed sample was loaded per slot. A DNA marker (GeneRuler 1kb DNA Ladder, Thermo Scientific) was loaded to determine the fragment sizes. The gel was run at 80 V for 120 minutes. After electrophoresis, the gel was incubated with ethidium bromide solution (0.5 µg/ml) for 30 minutes and washed with TAE buffer. Stained gels were documented using a gel imaging system (ChemiDocTM, BioRad).

2.2.1.2. SDS-PAGE

SDS-PAGE was performed after Laemmli *et al.* (33). First, the separating gel was prepared, poured and levelled with 70% isopropanol. After the separating gel has polymerized, isopropanol was removed, the stacking gel was prepared, poured on top of the separating gel and combs were inserted. After hardening, combs were removed, and the gels were stored in moist towels at 4°C until further use. All chemicals used for gel preparation and their proportions are listed in Table 1. The composition of the separating gel buffer and stacking gel buffer can be found in Table 2, next to all other buffers and solutions further used in sample preparation and electrophoresis.

Three types of samples were analysed by SDS-PAGE: proteins bound to chitin beads, (cell) pellets and liquid samples. For chitin beads, the samples were washed with TBS buffer (150 mM NaCl, 50 mM Tris, pH 7.5), centrifuged for 1 minute (14000 rpm) and the supernatant discarded. Washed chitin beads were then resuspended in ddH_2O and SDS sample buffer (1:1). Cell pellets were directly resuspended in ddH_2O and SDS sample buffer (1:1). For all liquid samples an equal amount of SDS sample buffer (1:1) was added. All three types of samples were vortexed and heated at 95°C for 5 minutes to ensure complete solubilization and denaturation. Afterwards, the samples were vortexed again and centrifuged for 1 minute (14000 rpm). 10 µl of the supernatant were then loaded per slot. The gels were run at 250 V for 38 minutes.

After electrophoresis, the gels were stained with Coomassie-staining solution for 20 minutes, washed with ddH_2O and destained in ddH_2O overnight. Finished gels were documented using a gel imaging system (ChemiDocTM, BioRad).

Table 1. SDS gel composition (for 12 gels).

	Separating gel	Stacking gel
ddH2O	12.5 ml	12.5 ml
Separating gel buffer	13.5 ml	-
Stacking gel buffer	-	2.5 ml
30% Acrylamide	26 ml	3.5 ml
10% SDS	550 µl	185 µl
10% APS	550 µl	185 µl
TEMED	18 µl	18 µl

Table 2. List of buffer and solution composition for SDS-PAGE.

Separating gel buffer	1.5 M TrisHCl, 0.4% (w/v) SDS, pH 8.8
Stacking gel buffer	0.5 M TrisHCl, 0.4% (w/v) SDS, pH 6.8
2x Sample buffer (loading buffer)	500 mM Tris, 6% (w/v) SDS, pH 6.8, 35% (v/v) glycine, 3.55% β-mercaptoethanol, 0.05% (w/v) bromphenolblue
10x Laemmli buffer	250 mM Tris, 2 M glycine, 1% (w/v) SDS
Coomassie-staining solution	0.1% (w/v) Coomassie R250, 10% (v/v) Acetic acid, 45% (v/v) methanol in water

2.2.2. Cloning of ubiquitin-intein fusion constructs

2.2.2.1. Site-directed mutagenesis of ubiquitin-intein fusion construct

Site-directed mutagenesis was used to perform two different point mutations: 1. from Ub to Ub(K48C) and 2. from Ub to Ub(K63C). For this the corresponding codon (AAA) of the desired lysine positions had to be changed to a codon for cysteine (TGC). With this change, back-to-back primers for the substitutions and a matching PCR protocol were designed using NEBaseChanger. The template used for both substitutions was pTXB1-Ub-Mxe-H7-CBD (plasmid map in supplementary data) in varying concentrations: 65.2 ng/µl (1), 6.52 ng/µl (1:10) and 0.652 ng/µl (1:100). A 7x master mix of Q5 Hot Start High-Fidelity (Table 3) was prepared and used for PCR according to the protocol created by NEBaseChanger (Table 4 and 5). The PCR was performed on a Bio-Rad T100 Thermocycler. The success of the substitutions was investigated by agarose gel electrophoresis as described in section 2.2.1.1.

Table 3. 7X Master Mix of Q5 Hot Start High-Fidelity.

Reaction Buffer (5x)	35.0 µl
10 mM dNTPs	3.5 µl
Q5 high GC Enhancer (5x)	35.0 µl
Nuclease-Free Water	75.2 µl
Q5 High-Fidelity DNA Polymerase	1.75 µl

Table 4. Reaction mix for PCR.

Q5 Hot Start High-Fidelity Master Mix	12.5 µl
10 µM Forward Primer	1.25 µl
10 µM Reverse Primer	1.25 µl
Template DNA	1.00 µl
Nuclease-free water	9.00 µl

Table 5. PCR protocol used in both site-directed mutagenesis substitutions.

1x	90°C	30 s
30x	98°C	10 s
	60°C	30 s
	72°C	4.5 minutes
1x	72°C	2 minutes

2.2.2.2. Purification and ligation of PCR product

The PCR product of the undiluted template was chosen for further experiments. The PCR product was purified following the step-by-step instructions of the PCR Purification Kit (GeneJet, Thermo Scientific). The purified PCR product was then ligated with the KLD Enzyme Mix (NEB). Therefore, the KLD Enzyme Mix Reaction Protocol (Table 6) was followed and the sample was incubated at room temperature for 5 minutes. The finished plasmid was stored at -20°C.

Table 6. Components of KLD Enzyme Mix Reaction.

Purified PCR Product	1 µl
KLD Reaction Buffer (2x)	5 µl
KLD Enzyme Mix (10x)	1 µl
Nuclease-free Water	3 µl

2.2.2.3. Transformation of mutated ubiquitin-intein fusion construct

The ligated plasmids were transformed into *E. coli* XL-1 competent cells using the heat shock method. 10 µl of the desired plasmid was added to 100 µl competent XL-1 cells and incubated on ice for 5 minutes. Then, the cells were incubated for 90 s at 42°C and afterwards cooled with ice. 1 ml LB-medium (10 g/l tryptone, 5 g/l yeast extract, 10 g/l NaCl) was added and the cells were incubated at 37°C for 45 minutes. Afterwards, the cells were centrifuged shortly, and the LB-medium was removed. The cells were then resuspended in 200 µl fresh LB-medium, plated onto selective LB-agar plates (15 mg/l agar, 10 g/l tryptone, 5 g/l yeast extract, 10 g/l NaCl and 100 µg/ml of ampicillin) and incubated at 37°C overnight. The plates were stored at 4°C.

2.2.2.4. Sequencing of mutated ubiquitin-intein fusion construct

From each of the two different XL-1 containing plates three colonies were picked and incubated in 3 ml of selective LB-medium (LB-medium plus 100 µg/ml ampicillin) at 37°C and 180 rpm, overnight. The overnight cultures were then used for plasmid extraction (GeneJET Plasmid Miniprep Kit, Thermo Scientific). The isolated plasmid was then tested for its DNA concentration using a Nanodrop 2000c

system (Thermo Scientific). The two extracts of each plasmid with the highest concentrations were sent for sequencing (Microsynth). Plasmid extracts were stored at -20°C.

2.2.2.5. *Creation of glycerol stocks*

Plasmid extracts of pTXB1-Ub(K48C)-Mxe-H7-CBD and pTXB1-Ub(K63C)-Mxe-H7-CBD with the correct DNA sequence were transformed into *E. coli* BL21 Gold (DE3) and plated as described in section 2.2.2.3. From each of the plates one colony was picked and incubated in 3 ml of selective LB-medium, at 37°C and 180 rpm, overnight. These overnight cultures were used to prepare glycerol stocks by diluting 1:1 with 40% glycerol in water. Glycerol stocks were labelled and stored at -80°C.

2.2.3. Expression of original and mutated ubiquitin-intein fusion constructs

2.2.3.1. *Expression of ubiquitin-intein fusion construct*

The *E. coli* BL21(DE3) pTXB1-Ub-Mxe-H7-CBD expression strain was already available and expression was optimized previously by Gerhard Niederacher. Therefore, it was immediately possible to perform a large-scale 6 l expression. For this, three flasks of 200 ml 2YT-medium (3.2 g/l tryptone, 2 g/l yeast extract, 1 g/l NaCl) were prepared and inoculated with *E. coli* BL21(DE3) containing the pTXB1-Ub-Mxe-H7-CBD plasmid at 37°C, 160 rpm, overnight. These overnight cultures were then used to inoculate three flasks of 2 l 2YT-medium containing 100 µg/ml ampicillin. The bacteria were grown at 37°C and 160 rpm until reaching an OD600 between 0.6 and 0.8. The OD600 was measured on a Nanodrop 2000c system (Thermo Scientific). Then, overexpression was induced by addition of 1mM IPTG and incubated for 2 h (37°C, 160 rpm) until the cells were harvested. At one-hour time intervals samples for SDS-PAGE analysis were taken. For this the OD600 was measured and used to calculate the sample volume needed, which was then taken from the expression medium, centrifuged (1 minute at 14000 rpm) and the supernatant discarded to create a cell pellet. The cell pellet was stored at -20°C until SDS-PAGE analysis. The first sample was taken shortly before induction of overexpression. For cell harvesting the expression medium was centrifuged (Beckman Coulter Avanti J26XP) at 6000 rpm, 4°C for 20 min. The resulting supernatant was discarded, and the cell pellet was stored at -80°C.

2.2.3.2. *Expression of mutated ubiquitin-intein fusion constructs*

The Ub(KxC)-Mxe-H7-CBD plasmids were not only transformed into *E. coli* BL21 Gold (DE3), but also into *E. coli* BL21(DE3) and *E. coli* Rosetta 2 (as described in 2.2.2.3.) and have been used for glycerol stock preparation (as described in 2.2.2.5.). As the plasmids have been newly created and, therefore, not been used in expression before, small scale test expressions (200 ml of 2YT-medium, 37°C, 200 rpm) were performed with all three strains for both Ub(K48C)-Mxe-H7-CBD and Ub(K63C)-Mxe-H7-CBD to identify suitable expression strains. After the test expressions, *E. coli* Rosetta 2 was chosen for Ub(K48C)-Mxe-H7-CBD, while for Ub(K63C)-Mxe-H7-CBD *E. coli* BL21Gold (DE3) was determined to work best. The expressions were then scaled up (1-6 l of 2YT-medium) and performed in a similar fashion as described in section 2.2.3.1, with the exceptions that the expression of Ub(K48C)-Mxe-H7-CBD contained 30 µg/ml chloramphenicol next to 100 µg/ml ampicillin, as needed for *E. coli* Rosetta 2 and both proteins, Ub(K48C)-Mxe-H7-CBD and Ub(K63C)-Mxe-H7-CBD, being overexpressed for 4h instead of 2h. As with Ub-Mxe-H7-CBD, the cell pellets were stored at -80°C until further use.

2.2.3.3. Lysis of cell pellets

For cell lysis high pressure cell disruption was used. The cell pellets from a 1 l expression were resuspended in 100 ml TBS buffer (150 mM NaCl, 50 mM Tris, pH 7.5) using an Ultra Turrax (IKA T18 basic). Then, the cell suspension was lysed by running it two times through a cell disruptor (Constant Systems). The lysed suspension was centrifuged (18000 rpm, 4°C, 30 minutes) and the supernatant was decanted and stored at 4°C until further processing. Samples of supernatant were taken for SDS-PAGE analysis.

2.2.4. Intein-mediated cleavage of Ub and Ub(KxC)

Further purification and cleavage of the expressed proteins was facilitated by using a chitin affinity matrix. The protein was first loaded on the chitin beads and afterwards an intein-mediated cleavage was induced by shaking the chitin beads in a cleavage solution.

2.2.4.1. Loading of chitin beads

About 10 ml of chitin beads were equilibrated for 15 minutes with 40 ml TBS buffer in a 50 ml tube, centrifuged and the supernatant discarded. The supernatant from cell lysis was poured onto the chitin beads and shaken for 1 h. After loading, a small sample of the chitin beads was taken and prepared for SDS-PAGE as described in 2.2.1.2. The loaded chitin beads were centrifuged (4000 rpm for 2 minutes) and the supernatant was discarded. The loaded chitin beads were washed either with PBS buffer (150 mM NaCl, 100 mM monosodium phosphate, pH 5.0), when cleaving with MesNa, or washed with TBS buffer, when cleaving with hydrazine. After centrifugation the buffer was discarded, and the protein cleaved either by thiol intein-mediated cleavage (MesNa) (section 2.2.4.2) or hydrazine intein-mediated cleavage (section 2.2.4.3).

2.2.4.2. Thiol-mediated cleavage

Loaded chitin beads washed with PBS buffer were suspended in PBS buffer containing 500 mM MesNa, resulting in a pH of 5.5 of the cleaving solution. The chitin beads were then shaken for three days in the cleaving solution, while every 24h a sample of the chitin beads was taken and prepared for SDS-PAGE analysis as described in 2.2.1.2. After 72 h the chitin beads were centrifuged and the supernatant (cleaving solution) was transferred and filtered. The solution was then stored at 4°C or directly purified by HPLC. The chitin beads were regenerated as recommended by the manufacturer and stored in ethanol at 4°C.

2.2.4.3. Hydrazine-mediated cleavage

Loaded chitin beads washed with TBS buffer were suspended in TBS buffer containing 5% hydrazine and 50 mM DTT, resulting in a pH of 9.8 of the cleaving solution. The chitin beads were then shaken for 48 h in the cleaving solution, while every 24h a sample of the chitin beads was taken and prepared for SDS-PAGE analysis as described in 2.2.1.2. The chitin beads were then centrifuged and the supernatant (cleaving solution) was transferred and filtered. The solution was then stored at 4°C or the pH lower to 7.0 immediately before purification. The chitin beads were regenerated as recommended by the manufacturer and stored in ethanol at 4°C.

2.2.5. Purification of modified Ub and Ub(KxC)

The solutions from section 2.2.4.2 and 2.2.4.3 were purified on a Waters Prep 150 System. Depending on the amount of solution, the cleavage solution was either loaded on a C4 semiprep, or a C4 prep column, preequilibrated with 5% buffer B (ACN + 0.08% TFA) and 95% buffer A (ddH2O + 0.1% TFA). The protein was eluted with a gradient, going from 5 – 50% buffer B in 45 minutes with a prior 20 minutes desalting at 5% B. The elution fractions, containing the protein, were analysed by ESI-MS, combined and lyophilized. The purity of the obtained protein was assessed by performing analytical HPLC analyses on a Dionex Ultimate 3000 instrument and ESI-MS analyses on Waters Auto Purification HPLC/MS system. The protein was stored at -20°C.

2.3. Conversion of C-terminal modifications

2.3.1. C-terminal hydrazide to allylamide conversion via azide intermediate

For this reaction a protocol from Zheng et *al.* was adapted (34). Ub(K48C)-$NHNH_2$ was dissolved in a filtered and degassed reaction buffer, containing 0.2 M Na_2HPO_4, 6 M Gua-HCl, pH 3.0, resulting in a protein concentration of 2 mg/ml. The solution was placed in a -15°C ice-salt bath, 20 mM $NaNO_2$ was added to the solution, stirred for 15 minutes and a sample for LC-MS was taken. Then, either 500 mM allylamine or 200 mM MPAA was added to the reaction and the pH was adjusted to 7.0 with NaOH. The reaction was stirred at room temperature. After 1 h, samples were taken. To the reaction containing MPAA, 500 mM allylamine was added and stirred for 1 h, after which a sample was taken. All samples were immediately analysed by LC-MS and the reactions were stored at -80°C.

2.4. Thiol–ene coupling

The desired thiol-ene reaction has been reported by Valkevich et *al.* (23) and their protocol has been adapted for this master thesis.

2.4.1. Coupling Ub-allylamide with glutathione

For the test reaction of thiol-ene coupling, it was chosen to couple Ub-Allylamine with Glutathione. For this 1 mM Ub-Allylamine, 1 mM Glutathione, 0.5 mM LAP and 0.5 mM TCEP were dissolved in 6 M GUA pH 4.4, sonicated if needed, and placed in an ice bath for radiation. The sample was radiated for 1 h with a UV-lamp at wavelength 365 nm (CS2010 - UV Curing LED System, Thorlabs). The reaction was monitored by LC-MS on Waters Auto Purification HPLC/MS system. The product was stored at -20°C.

2.4.2. Coupling Ub-allylamide with Ub(KXC)-hydrazide

For thiol-ene coupling of Ub proteins, Ub-allylamide was coupled to Ub(KxC)-hydrazine. For this 1 mM Ub-allylamine, 1 mM Ub(KxC)-hydrazine, 0.5 mM LAP and 0.5 mM TCEP were dissolved in 6 M GUA pH 4.4, sonicated if needed, and placed in an ice bath for radiation. The sample was radiated for 1 h with a UV-lamp at wavelength 365 nm (CS2010 - UV Curing LED System, Thorlabs). The reaction was monitored by LC-MS on Waters Auto Purification HPLC/MS system. The product was stored at -20°C.

3. Results

3.1. Cloning of ubiquitin-intein fusion constructs

The goal was to introduce cysteine mutations at specific lysine positions in ubiquitin, to install the needed thiol group for thiol-ene coupling. Two different ubiquitin variations were desired: ubiquitin with a cysteine at position 48 and a version with cysteine at position 63. The *E. coli* vector pTXB1-Ub-Mxe-H7-CBD was already available and therefore site-directed mutagenesis was chosen to generate the desired cysteine mutations. As described in section 2.2.2.1, back-to-back primers, including the codon change from AAA for lysine to TGC for cysteine at the desired positions 48 and 63, were designed and purchased (primer design in supplementary data). For the site-directed mutagenesis, PCRs were performed using 1 µl of three different template concentrations (undiluted: 65.2 ng/µL, 1:10, and 1:100 dilutions) for each ubiquitin variant.

The PCR reactions were analysed *via* agarose gel electrophoresis to visualize the obtained DNA fragments. As can be seen in Figure 1 a strong band was detected in all samples between 8000 and 6000 bp, with the desired PCR product have a size of 6883 bp, thereby indicating that the PCRs were successful.

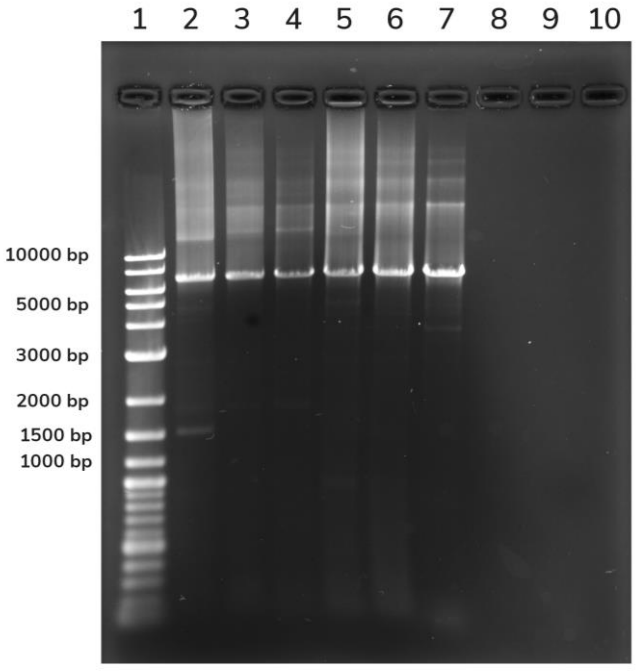


Figure 1. Analytical agarose gel of PCR products from site-directed mutagenesis. Lane 1 contains DNA ladder, lane 2-4 contain PCR products of K48C mutation, with undiluted, 1:10 and 1:100 diluted template concentration, respectively. Lane 5-7 contain PCR products of K63C mutation, with undiluted, 1:10 and 1:100 diluted template concentration, respectively. Lane 8-10 are empty.

The PCR products obtained from the reaction with the undiluted plasmid template were purified and ligated using the site-directed mutagenesis kit (see section 2.2.2.2). The ligated plasmids were transformed into *E. coli* XL-1 competent cells and three colonies of both ubiquitin variations were inoculated and the overnight cultures were used for plasmid extraction. The purified plasmids were

sequenced to confirm successful mutation. The results of the sequencing verified the correct substitution of lysine with cysteine at the desired positions and no changes in the proximate sequence. With this information, the plasmids were used for the expression of Ub(KxC)-Mxe-H7-CBD (with x = 48 or 63). The original DNA sequence of pTXB1-Ub-Mxe-H7-CBD and the results of the sequencing for both ubiquitin variants can be found in the supplementary data.

3.2. Expression of ubiquitin-intein fusion constructs

3.2.1. Expression of Ub-Mxe-H7-CBD

E. coli BL21 (DE3) transformed with pTXB1-Ub-Mxe-H7-CBD was already available for expression and optimized previously by Gerhard Niederacher. Therefore, a large-scale expression of 6 l (three times 2 l) was performed. The cells were harvested by centrifugation two hours after induction at 37°C and lysed in a cell disruptor. After another centrifugation step, the Ub-Mxe-H7-CBD protein was found in the soluble fraction of the lysate. For SDS-PAGE analysis, samples were taken at the following times points: before induction (0 h), 1h and 2 h after induction. Further, samples of the supernatant after cell lysis (crude), the supernatant after loading of the chitin beads and a loaded chitin bead sample were taken and examined by SDS-PAGE (Figure 2).

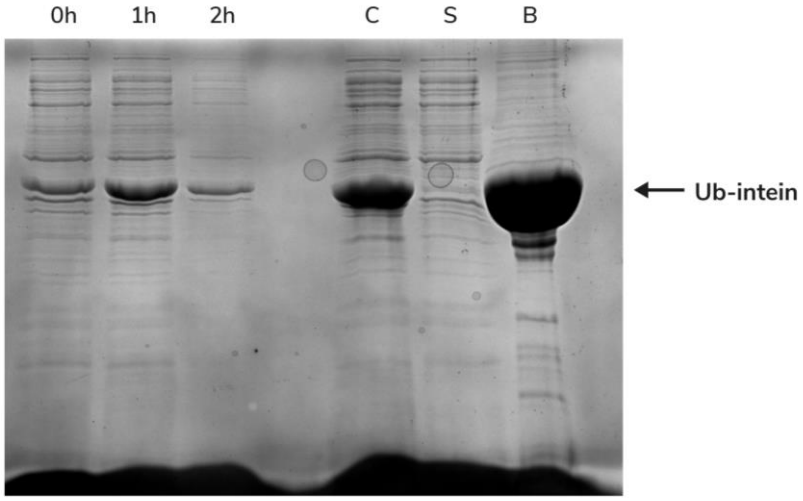


Figure 2. SDS-PAGE of Ub-Mxe-H7-CBD expression samples, with samples at time point 0 h, 1 h and 2 h after induction, crude product (C) after cell lysis, supernatant (S) and beads (B) after loading on chitin-beads.

Unfortunately, no marker was loaded on the gel, however, the strong bands at 0 h, 1 h, and 2 h indicate the overexpression and based on available comparison we concluded that the Ub-intein was expressed. The SDS-gel also revealed that the protein has already been expressed before induction.

The band occurring in lane C (crude) showed that the protein resided in the soluble fraction after cell lysis. Lane S (supernatant), which showed no band for the protein, and lane B (loaded chitin beads), which showed a large band, indicating that binding of ubiquitin to the chitin beads was successful. With the successfully loaded chitin beads, further experiments were carried out (see section 3.3).

3.2.2. Selection of expression strains for Ub(KxC)-Mxe-H7-CBD

Both ubiquitin mutants were transformed into *E. coli* BL21 Gold (DE3). A 1 l test expression was performed to examine differences between expressions of native ubiquitin and cysteine mutants. For test expressions, the same conditions were used as for the wild-type ubiquitin expression. As before, samples of the expression (time points 0 h, 1 h, 2 h), the crude after cell lysis, and the chitin beads after loading were taken and analysed by SDS-PAGE (Figure 3).

On the SDS-gel of the ubiquitin cysteine mutants, the samples taken from the culture do not show the strong overexpression, which has been observed for the wild type protein. However, some protein was expressed in both cases, which can be seen by the band between 45 and 30 kDa found in the lanes of the loaded chitin bead sample (lane B, Figure 3). When comparing the wild type expressed in *E. coli* BL21(DE3) with the expression of the ubiquitin variations in *E. coli* BL21 Gold (DE3), a weaker expression is observed.

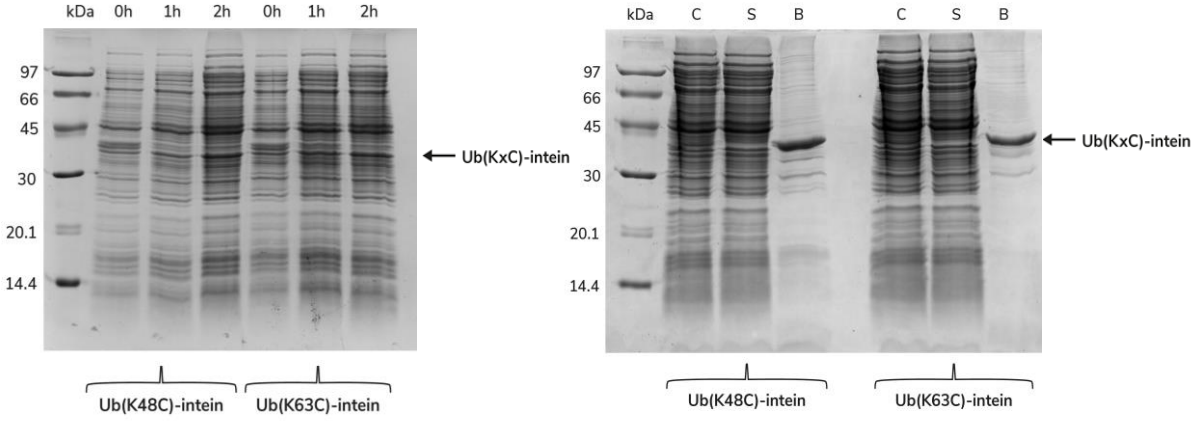


Figure 3. SDS-PAGE of Ub(KxC)-Mxe-H7-CBD expressions, with samples at time point 0 h, 1 h and 2 h after induction, crude product (C) after cell lysis, supernatant (S) and beads (B) after loading on chitin-beads.

From these findings, experiments were performed to find more efficient expression strains and better expression conditions. The expression strains *E. coli* BL21 (DE3) and *E. coli* Rosetta 2 were chosen for testing, with the pTXB1-Ub(KxC)-Mxe-H7-CBD plasmid being transformed into these strains. The test expressions were done in 1 l cultures, with expression time increased to four instead of two hours at 37°C. Samples for SDS-PAGE analysis were taken from the expression media of all strains before induction (0 h), as well as 1 h, 2 h, 3 h, and 4 h after induction. The SDS-gels of the test expressions can be found in the supplementary data.

To compare expression results, cells were harvested, and cell lysate was loaded on chitin beads. Samples of the loaded chitin beads were analysed by SDS-PAGE to compare the amount of expressed protein (Figure 4). Beforehand, samples from both the supernatant and the pellet after cell lysis were analysed *via* SDS-PAGE to ensure that the protein was in the soluble fraction after disruption (data not shown). In all expressions, the protein was in the soluble fraction of the cell lysate.

The SDS-gel showed that the Ub(K48C) variant was expressed in all three tested *E. coli* strains, as indicated by the band between 45 and 30 kDa. The most intense band was observed in *E. coli* Rosetta 2. Meanwhile, for the Ub(K63C) variant, all three tested *E. coli* strains, as indicated by the band between

45 and 30 kDa, expressed Ub(K63C). However, the bands for *E. coli* BL21(DE3) and *E. coli* Rosetta 2 are very faint in comparison to *E. coli* BL21 Gold (DE3) strain (see Figure 4). Consequently, *E. coli* Rosetta 2 was selected for Ub(K48C), while *E. coli* BL21 Gold (DE3) was chosen for Ub(K63C) to perform larger-scale expressions.

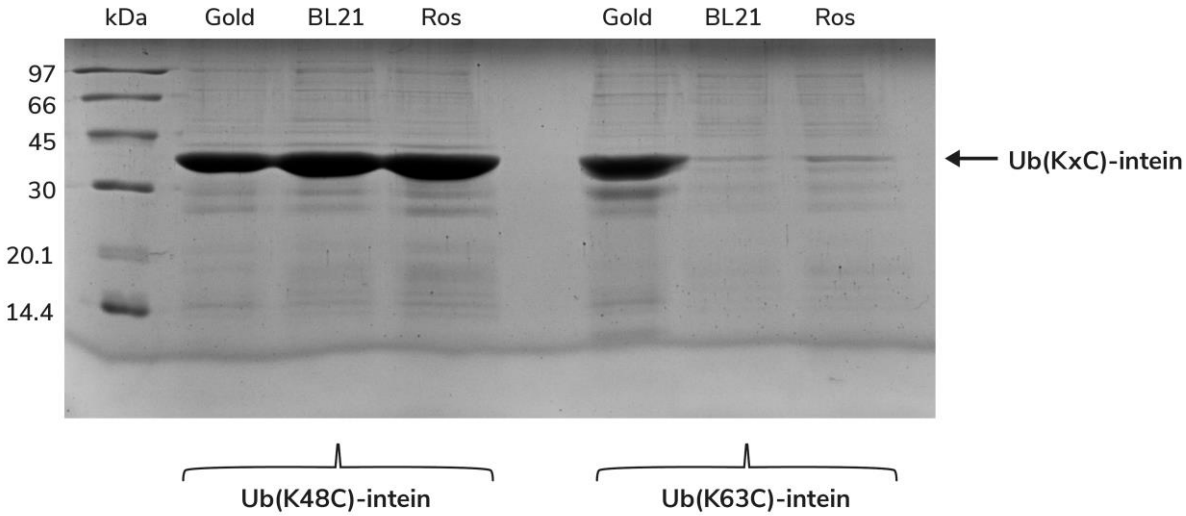


Figure 4. SDS-PAGE comparing the loaded chitin bead samples of different expression strains of Ub(KxC)-intein..

3.2.3. Expression of Ub(KxC)-Mxe-H7-CBD

Expressions of 6 l (three times 2 l) were performed, with cells being harvested 4 hours after induction at 37°C. The cells were lysed using a cell disruptor, with the desired proteins residing in the soluble fraction of the cell lysate. The lysates were incubated with chitin beads, resulting in the transfer of the protein to the chitin beads. The loaded chitin beads were used in further experiments (see section 3.3).

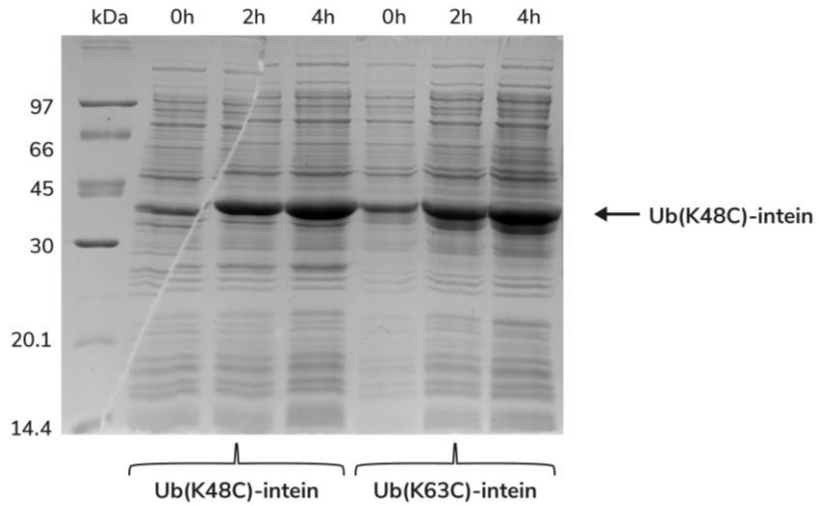


Figure 5. SDS-PAGE of large-scale expressions of Ub(K48C)-Mxe-H7-CBD and Ub(K63C)-Mxe-H7-CBD monitored over 4 h.

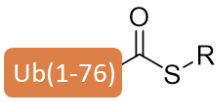
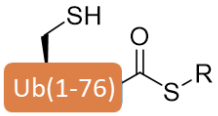
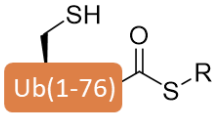
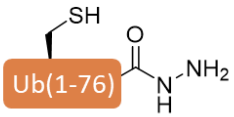
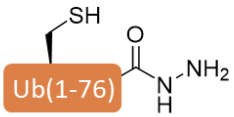
Samples of the expression were taken before induction (0 h), as well as 2 h and 4 h after induction (Figure 5). Further, samples of the supernatant after cell lysis (crude) and the supernatant and chitin beads after loading were taken and examined by SDS-PAGE (data not shown).

As indicated by the SDS-gel in Figure 5, the expression of Ub(K48C)-Mxe-H7-CBD and Ub(K63C)-Mxe-H7-CBD was successful. The expressed proteins were successfully loaded on chitin beads (data not shown), which were used in subsequent experiments (see section 3.3).

3.3. Intein-mediated cleavage of Ub- and Ub(KxC)-intein fusion constructs

In this chapter intein-mediated cleavage of Ub- and Ub(KxC)-intein constructs is discussed. An overview of the Ub-variants generate by intein-mediated cleavage can be seen in Table 7.

Table 7. Overview of all generated Ub-variation described in section 3.3.

	Ub(wildtype)	Ub(K48C)	Ub(K63C)
Ub-thioester (R: $CH_2CH_2SO_3H$)			
Ub-hydrazide			

3.3.1. Thiol-mediated cleavage

Thiol-mediated intein cleavage was used to install a thioester at the C-terminus of Ub or Ub(KxC), as depicted in Figure 6.

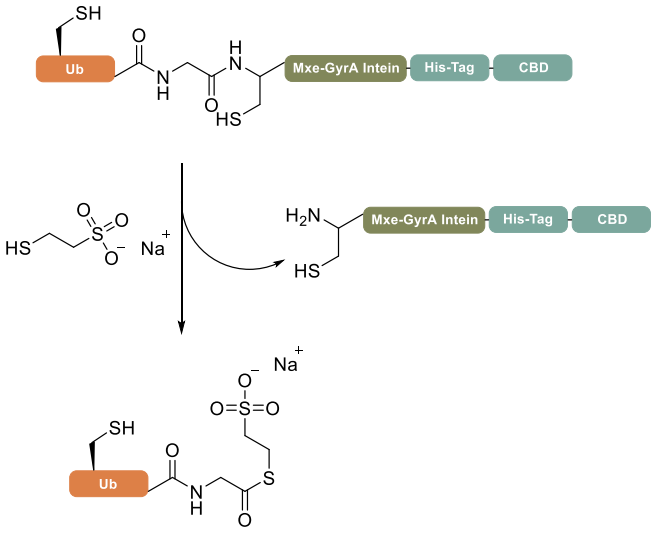


Figure 6. Overview of MesNa-based intein-mediated cleavage of Ub(KxC)-Mxe-H7-CBD.

For this, loaded chitin beads were washed with TBS buffer (pH 7.5), the buffer was discarded after centrifugation and the washed beads were suspended in cleavage solution consisting of 500 mM MesNa in TBS buffer. The resulting pH of the chitin beads suspension was 8.0. The chitin beads were shaken in the cleavage solution for 72 h, with bead samples for SDS-PAGE taken every 24 h. Afterward, the suspension was centrifuged, the supernatant was transferred, filtered and the pH was lowered to 7.0 with HCl before purification with an HPLC preparative system. The resulting fractions were screened for the desired protein by ESI-MS, pooled, and lyophilized. The final product was analysed by ESI-MS and analytical HPLC.

The successful cleavage was verified by SDS-PAGE (Figure 7), as indicated by a second band occurring beneath 30 kDa, caused by the cleaved Mxe-H7-CBD construct. No band indicating the Ub(KxC)-SR can be observed on the gel, as the mass of the protein is too low to be detected on a 15% acrylamide gel. All other bands seen in the SDS-PAGE, which are not labelled in Figure 7, were impurities carried over from the crude product. Comparing the band intensities of the uncleaved ubiquitin before and after cleavage indicated that approximately 95% of the fusion protein was cleaved after 72 hours. No further cleavage was observed after 72 h of incubation (data not shown).

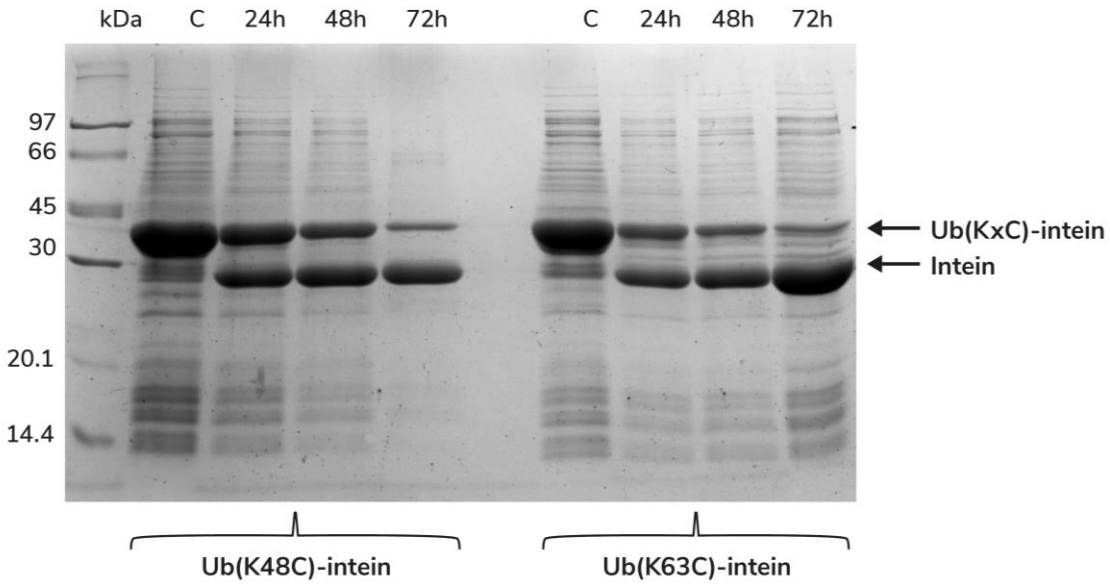


Figure 7. SDS-PAGE of the MesNa cleavage of Ub(K48C)-Mxe-H7-CBD and Ub(K63C)-Mxe-H7-CBD monitored over 72 h. C: before MesNa, 24 h/48 h/72 h: after MesNa

After 72h cleavage, the samples were purified by HPLC and the fractions containing the desired protein were identified, pooled, and lyophilized. Afterward, analytical HPLC and ESI-MS were performed to characterize the final product (Figure 8 A). Both measurements indicate that a certain part of the sample was contaminated with hydrolysed ubiquitin, which was most likely caused by a basic pH (8.0) during cleavage. The chromatogram showed the main product eluting at 21.7 minutes, with some impurity eluting slightly before, while in the mass spectrum of the same sample both the desired mass of 8689 Da and the mass of the hydrolysed product Ub-COOH, with a mass of 8565 Da, were observed.

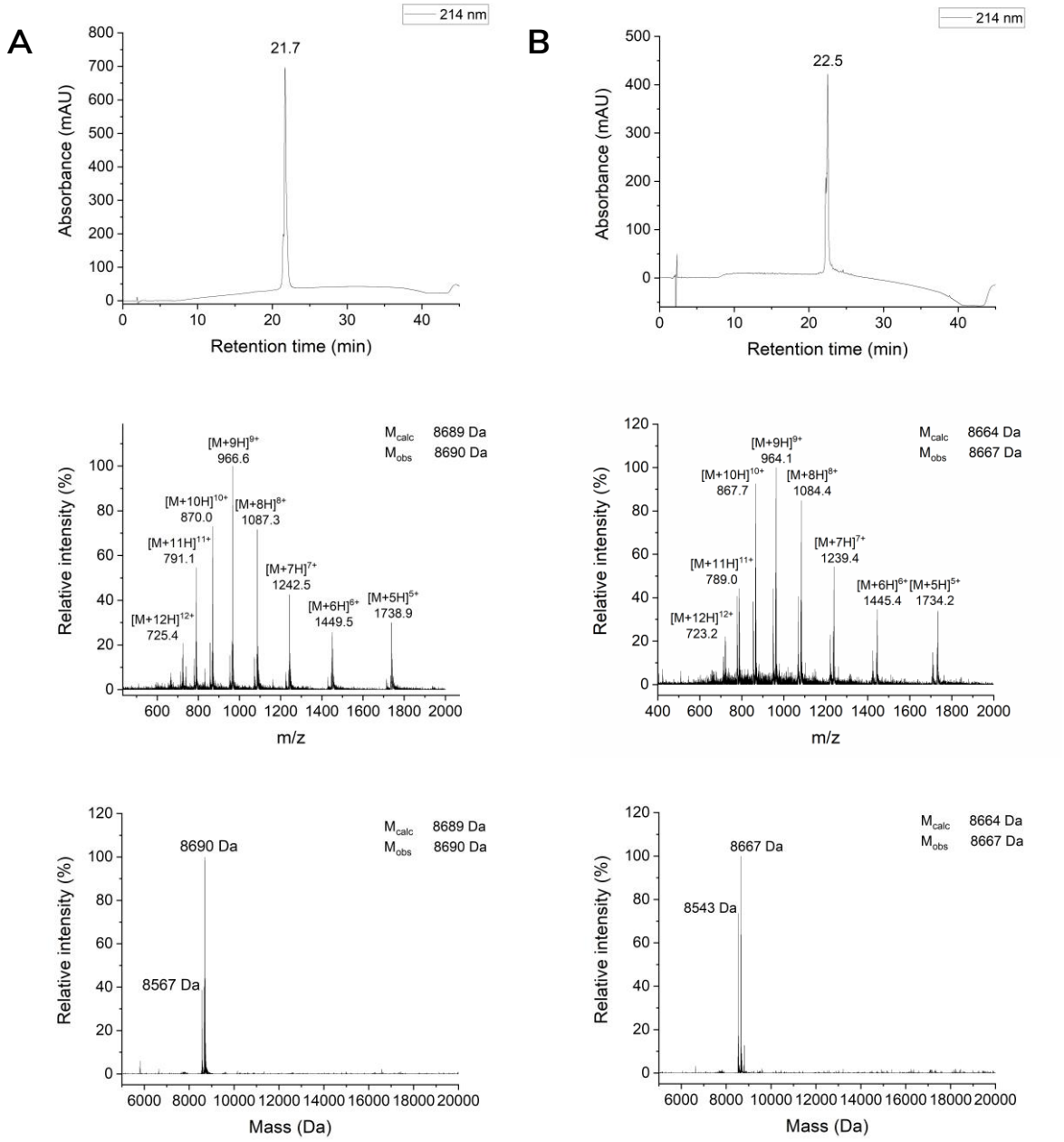


Figure 8. A shows the final analysis data of Ub-SR, while B shows the final analysis data of Ub(K48)-SR. The chromatograms of analytical HPLC are at the top, the original MS-data in the middle and the deconvoluted MS-data on the bottom. For Ub-SR the expected mass of the desired product (8690 Da) has been detected, with some hydrolysed product (8567 Da). For Ub(K48)-SR the expected mass of the desired product (8667 Da) has been detected, with some hydrolysed product (8543 Da).

In the case of Ub(K48C)-SR, which was cleaved using the same protocol, an even higher amount of the hydrolysed product was observed (Figure 8 B). Therefore, it was attempted to optimize the MesNa cleavage to minimize hydrolysis. Previous experiments performed by Gerhard Niederacher found that lowing the pH for the MesNa cleavage of ubiquitin does not affect the efficiency of the cleavage and may improve hydrolysis. This is the case for a solution above pH 5.0. Below pH 5.0, the efficiency of the cleavage was compromised.

With this information, a test cleavage at pH 5.5 was performed. To enable a consistent pH of 5.5, the TBS buffer pH 7.5 for washing and cleavage solution preparation was exchanged with a PBS buffer pH 5.0. The loaded chitin beads were, as before, shaken in the test cleavage solution for three days, with samples taken every 24 hours. The SDS-PAGE indicated no difference in cleavage efficiency, still showing cleavage at 95% with band sizes staying the same (comparison of SDS-PAGE in supplementary data). After 72 h, the cleavage solution underwent the same treatment as in the previous protocol, with the exemption of pH adjustment. When analysing the resulting fraction after HPLC purification, MS-data showed less hydrolysis and some fractions being entirely free of hydrolysis (data not shown). The final analysis of the protein with the improved protocol showed less hydrolysis compared to the old protocol (compare Figure 8 B with Figure 9A).

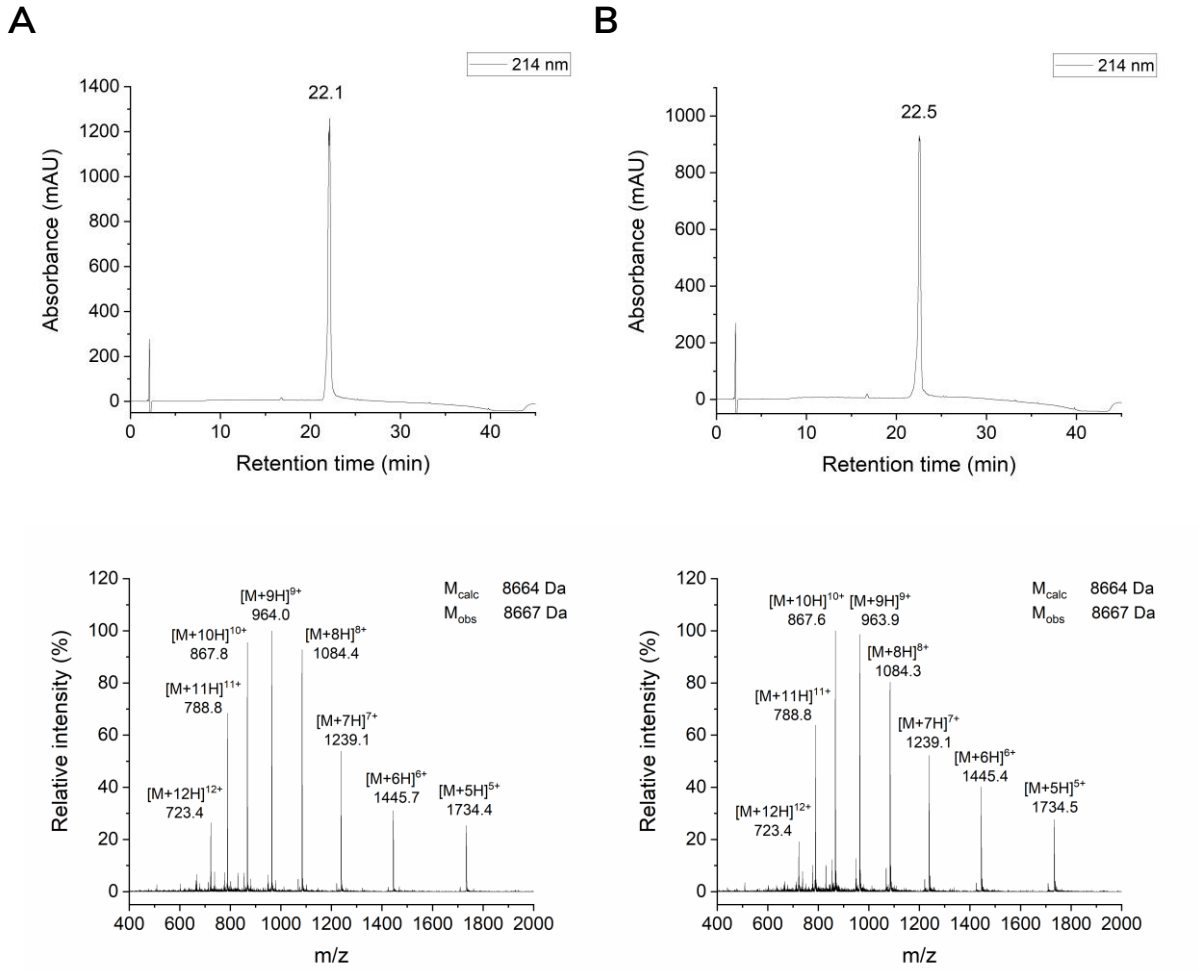


Figure 9. A shows the final analysis data of the improved Ub(K48C)-SR, while B shows the final analysis data of Ub(K63C)-SR. The chromatograms of analytical HPLC are at the top and the original MS-data om the bottom. For Ub(K48C)-SR and Ub(K63C)-SR the expected masses of the desired products (8667 Da) have been detected. Small amounts of hydrolysed product were observed in both Ub-variants (8543 Da).

The test cleavage was performed with Ub(K48C) protein and was successfully applied for Ub(K63C) protein (final analysis data in Figure 9). With the new protocol, yields of 12 mg/l for Ub(K48C)-SR and 10 mg/l for Ub(K63)-SR of expression medium were achieved. Ub-SR has been produced before protocol optimization and a yield of 3 mg/l was achieved.

3.3.2. Hydrazine-mediated cleavage

Hydrazine-mediated intein cleavage was used to install a hydrazide at the C-terminus of Ub or Ub(KxC), as depicted in Figure 10.

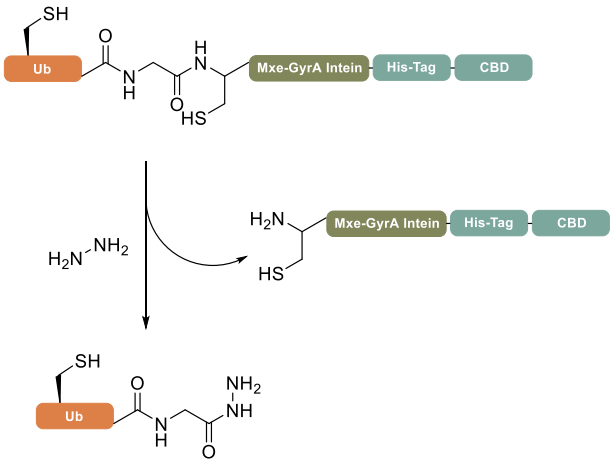


Figure 10. Overview of hydrazine-based intein-mediated cleavage of Ub(KxC)-Mxe-H7-CBD.

For the intein-mediated cleavage with hydrazine, an already established procedure was available (Gerhard Niederacher personal communication). Loaded chitin beads were washed with TBS buffer (pH 7.5), the buffer was discarded after centrifugation and the washed beads were suspended in a cleavage solution composed of TBS buffer with 5% hydrazine and 50 mM DTT. The resulting pH of the chitin beads suspension was 9.8. The chitin beads were shaken in cleavage solution for two days, with bead samples taken every 24 h. Afterward, the cleavage solution was centrifuged, transferred, filtered and the pH lowered to 7.0 before purification on an HPLC prep system. The resulting fractions were screened for the desired protein with ESI-MS, pooled, and lyophilized. The final product was analysed via ESI-MS and analytical HPLC.

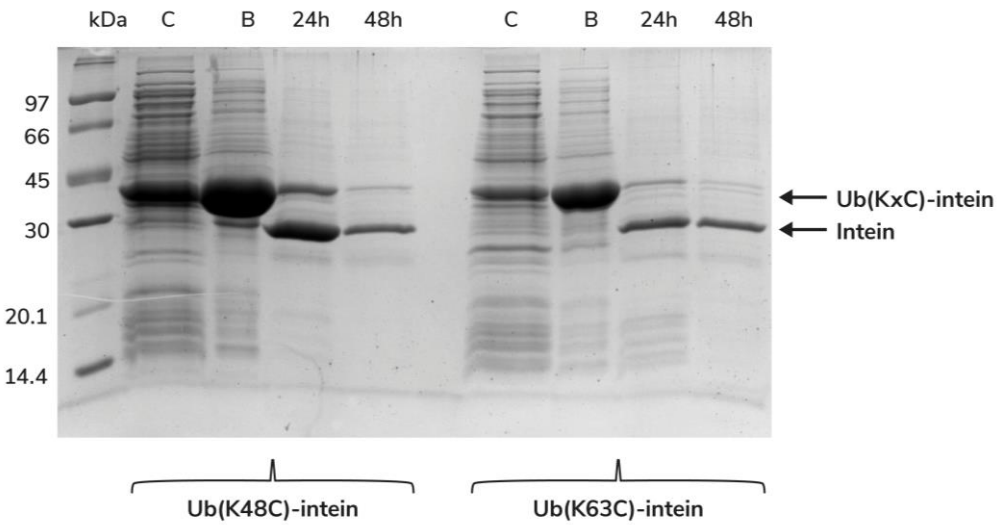


Figure 11. SDS-PAGE of the hydrazine cleavage of Ub(K48C)-Mxe-H7-CBD and Ub(K63C)-Mxe-H7-CBD monitored over 48 h. C: before MesNa, B: protein on chitin beads, 24 h/48 h: after MesNa.

The reaction was monitored by SDS-PAGE (Figure 11). Cleavage of the protein was indicated by a second band occurring slightly below 30 kDa, originating from the cleaved Mxe-H7-CBD construct. The desired protein Ub(KxC)-$NHNH_2$ cannot be observed on the gel, as the mass of the protein is too low to be detected on a 15% acrylamide gel. All other bands seen in the SDS-PAGE, which are not labelled in Figure 11, were impurities carried over from the crude product. The samples after 48 h showed that after this time point there was almost no uncleaved Ub(K48C)-Mxe-H7-CBD left.

Final analysis with ESI-MS and HPLC showed that in comparison to the MesNa cleavage, nearly no protein was hydrolysed, although the cleavage was performed at pH 8.0. The HPLC-chromatogram showed the final product eluting after 22.2 minutes, with no shoulder peaks. The MS-spectrum of the same sample displayed a mass of 8557 Da, which is close to the expected mass of 8553 Da, with little to no traces of hydrolysed product. This was observed for both Ub(K48C)-$NHNH_2$ and Ub(K63C)-$NHNH_2$ (Figure 12). With this protocol, yields of 13 mg/l for Ub(K48C)-$NHNH_2$ and 5 mg/l Ub(K63C)-$NHNH_2$ were achieved.

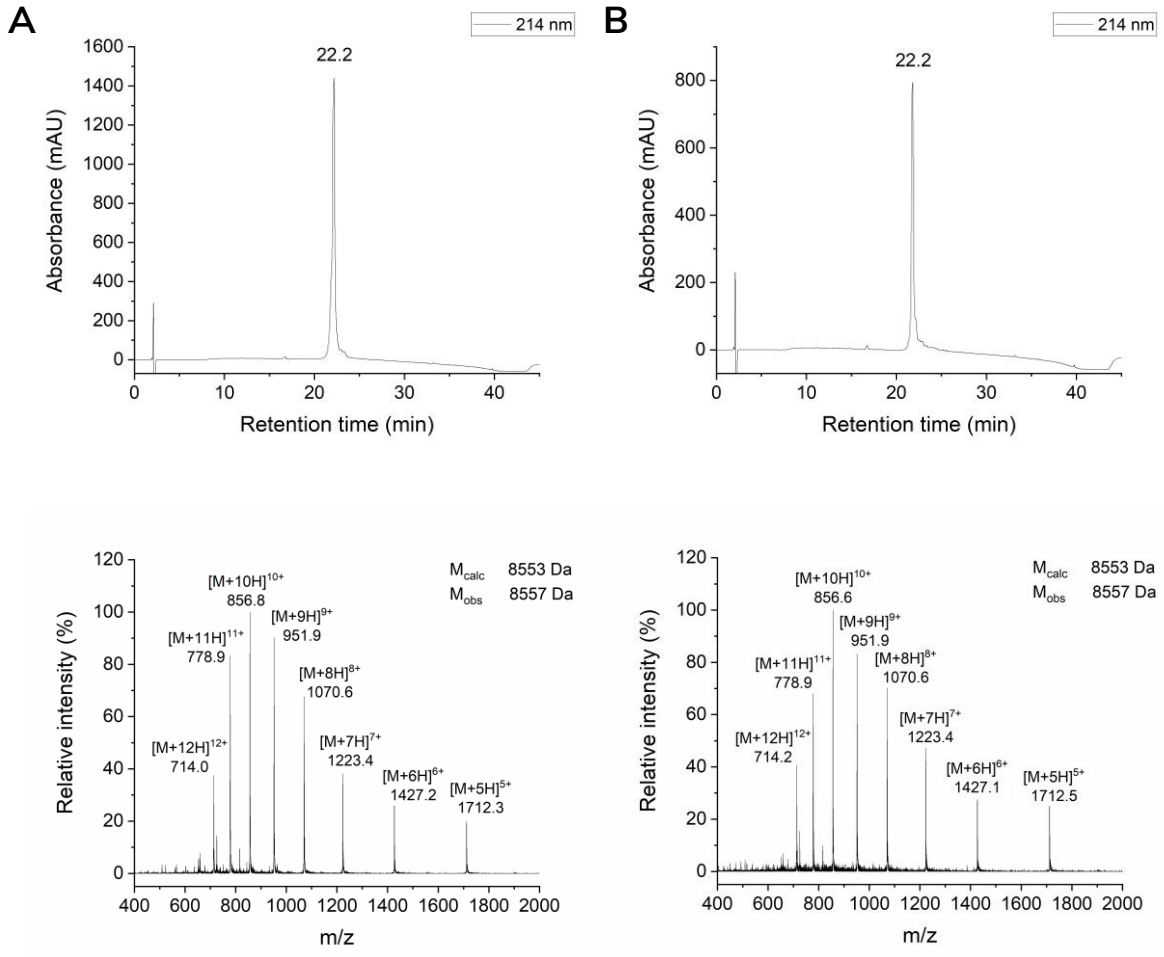


Figure 12. A shows the final analysis data of Ub(K48C)-$NHNH_2$, while B shows the final analysis data of Ub(K63C)-$NHNH_2$. The chromatograms of analytical HPLC are shown at the top and the original MS-data on the bottom. For Ub(K48C)-$NHNH_2$ and Ub(K48C)-$NHNH_2$ the expected masses of the desired products (8557 Da) have been detected. Trace amounts of hydrolysed product were observed in both Ub-variants

3.3.3. Allylamine-mediated cleavage

3.3.3.1 Cleavage solution containing only allylamine

The induction of intein-mediated cleavage *via* allylamine was investigated. As with the thiol or hydrazine mediated cleavage, chitin beads were loaded with protein, the beads washed with buffer, in this case, TBS buffer pH 7.5, and then suspended in cleavage solution containing TBS buffer with 500 mM allylamine. The beads were incubated for two days, with samples taken every 24 h. The cleavage samples were loaded next to MesNa cleavage samples, for a conclusive comparison of cleavage efficiency.

The SDS-gel, seen in Figure 13, indicated that no cleavage was induced by allylamine, as no Mxe-H7-CBD band beneath 30 kDa was found after 24 h and 48 h of incubation. In contrast, the control cleavages with MesNa displayed this band. Therefore, the cleavage with an allylamine-containing cleavage solution was deemed not possible under the tested conditions (48h at rt and pH above 12), and no further purification after incubation was performed. All other bands seen in the SDS-PAGE, which are not labelled in Figure 13, were impurities carried over from the crude product.

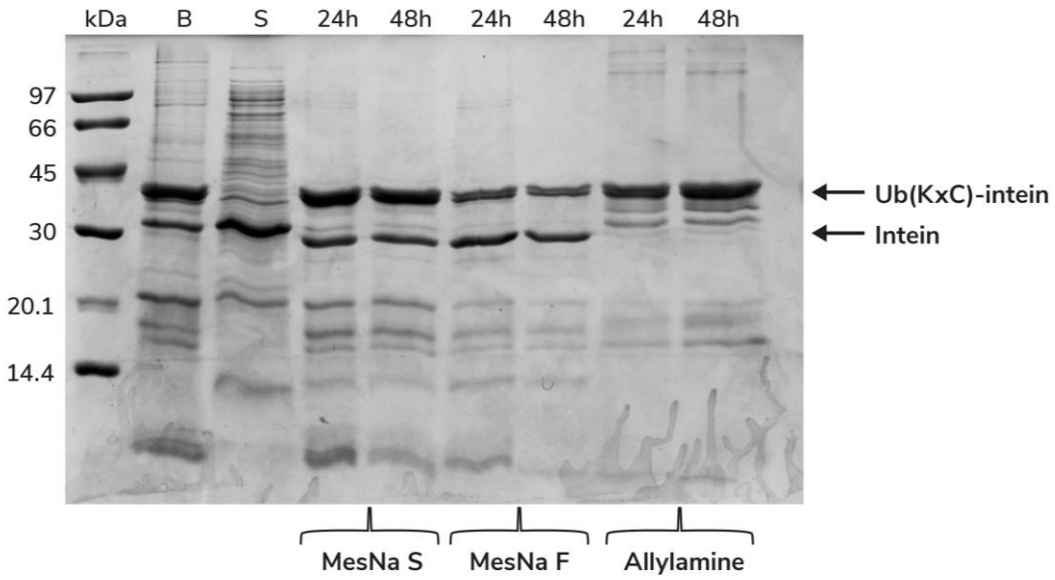


Figure 13. SDS-PAGE of cleavage with MesNa S (form Aldrich Sigma), MesNa F (from Fluorchem) and allylamine. monitored over two days. Lane B contains the loaded bead sample, lane S the supernatant after loading. The lanes marked with MesNa S and MesNa F contain samples of MesNa cleavage, while the lanes labled allylamine contain samples of allylamine cleavage.

3.3.3.2. Addition of allylamine to thiol cleavage solution

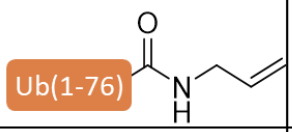
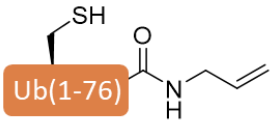
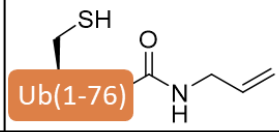
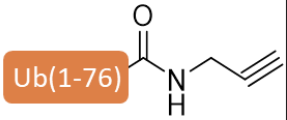
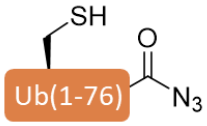
A different approach for generating proteins with a C-terminal allylamine moiety *via* intein cleavage was conceived while studying literature. Either an intein-mediated cleavage with MesNa ca be performed followed by aminolysis with propargylamine (35) or a cleavage with the cleavage solution containing both MesNa (250 mM) and propargylamine (250 mM) can be employed (36) to install a propargylamine group at the C-terminus of ubiquitin-like proteins. As allylamine is similar to propargylamine, an attempt was made to use the first protocol (MesNa then aminolysis) with allylamine replacing propargylamine. However, the protocol was changed slightly, using 500 mM allylamine and

MesNa instead of 250 mM, as all other experiments performed during this master thesis used 500 mM concentrations of allylamine and MesNa. With allylamine added the beads were incubated one more day. Afterward, LC-MS was performed to see whether the solution contained Ub-SR or Ub-allylamine. No information was obtained from the LC-MS, as the cleavage solution contained too many impurities to identify any components and no SDS-PAGE was performed. The attempt to purify the cleaving solution *via* HPLC yielded no fractions with desired product. Therefore, further testing was abandoned.

3.4. Conversion of C-terminal modifications

In this chapter the conversion of the previously installed C-terminal modifications of Ub and Ub(KxC) is discussed. An overview of the Ub-variants generate by different reactions can be seen in Table 8.

Table 8. Overview of all generated Ub-variation described in section 3.4

	Ub(wildtype)	UbK48C	UbK63C
Ub-allylamide			
Ub - propargylamide			
Ub-azide			

3.4.1. Conversion of C-terminal thioester to C-terminal allylamide

For the conversion of Ub(KxC)-SR to Ub(KxC)-allylamide (Ub(KxC)-AA) (see Figure 14), the starting material was dissolved in ddH_2O with 500 mM allylamine, resulting in a protein concentration of 1 mg/mL. The pH of the reaction solution was 12.0. This solution was then incubated at room temperature for up to 2 h, monitoring the reaction with ESI-MS every 30 minutes. Once ESI-MS indicated full conversion, the pH was adjusted to 7.0 with HCl, and the solution was purified *via* HPLC. The resulting fractions were screened for the desired protein by ESI-MS, pooled, and lyophilized. The final product was analysed *via* ESI-MS and analytical HPLC.

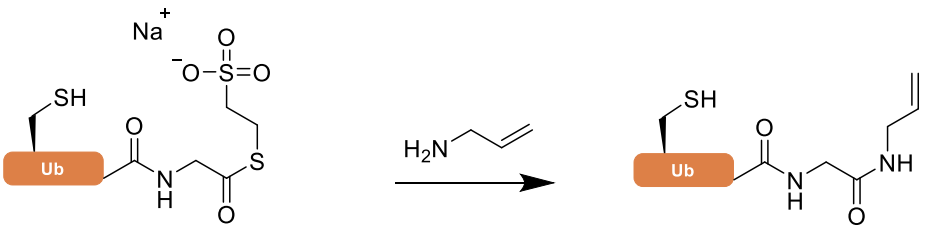


Figure 14. Overview of conversion from Ub(KxC)-SR to Ub(KxC)-AA.

The protocol for this conversion was established *via* experiments with Ub(K48C)-SR. Unfortunately, while experimenting, previously created Ub(K48C)-SR ran out before proper purification was established. Therefore, the conversion to Ub(K48C)-AA was only confirmed by ESI-MS of the reaction solution (Figure 15) and no yield can be provided.

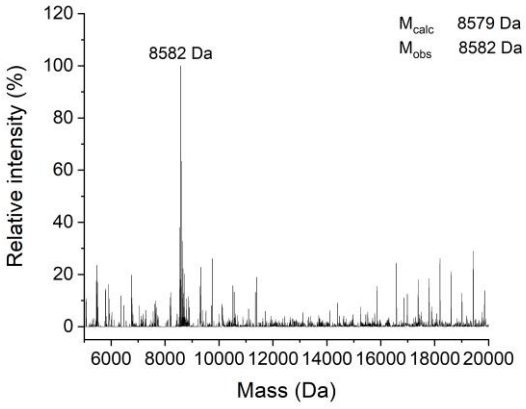


Figure 15. Deconvoluted MS-data of the reaction solution from Ub(K48C)-SR to Ub(K48C)-AA conversion. Expected mass: 8579 Da, observed mass: 8582 Da.

Nevertheless, the established conversion protocol was directly transferable to Ub(K63C)-SR, and Ub(K63C)-AA was subsequently purified and obtained as lyophilized protein. The final analysis data of Ub(K63C)-AA can be seen in Figure 16. The chromatogram showed two products, one eluting at 22.0, the other at 22.4 minutes, indicating impurities or side-products. The MS-data of the same sample displayed both the desired mass of 8579 Da and a mass of an unidentified by-product, with a mass around 8715 Da, which corresponds to the second peak seen in the chromatogram. The identification of the by-product was not possible. The yield was at 16%, however, the high amount of by-product lowers the actual yield of Ub(K63C)-AA significantly.

For the conversion of wild-type Ub-SR to Ub-AA, the same procedure was used. However, the conversion was more successful compared to Ub(K63C). The final analysis data of Ub-AA (Figure 16) indicated a pure product, with the chromatogram displaying one peak at 21.6 minutes and the MS-data showing the desired mass of 8604 Da. Small amounts of hydrolysed product were found in the MS-data. The yield was 52%.

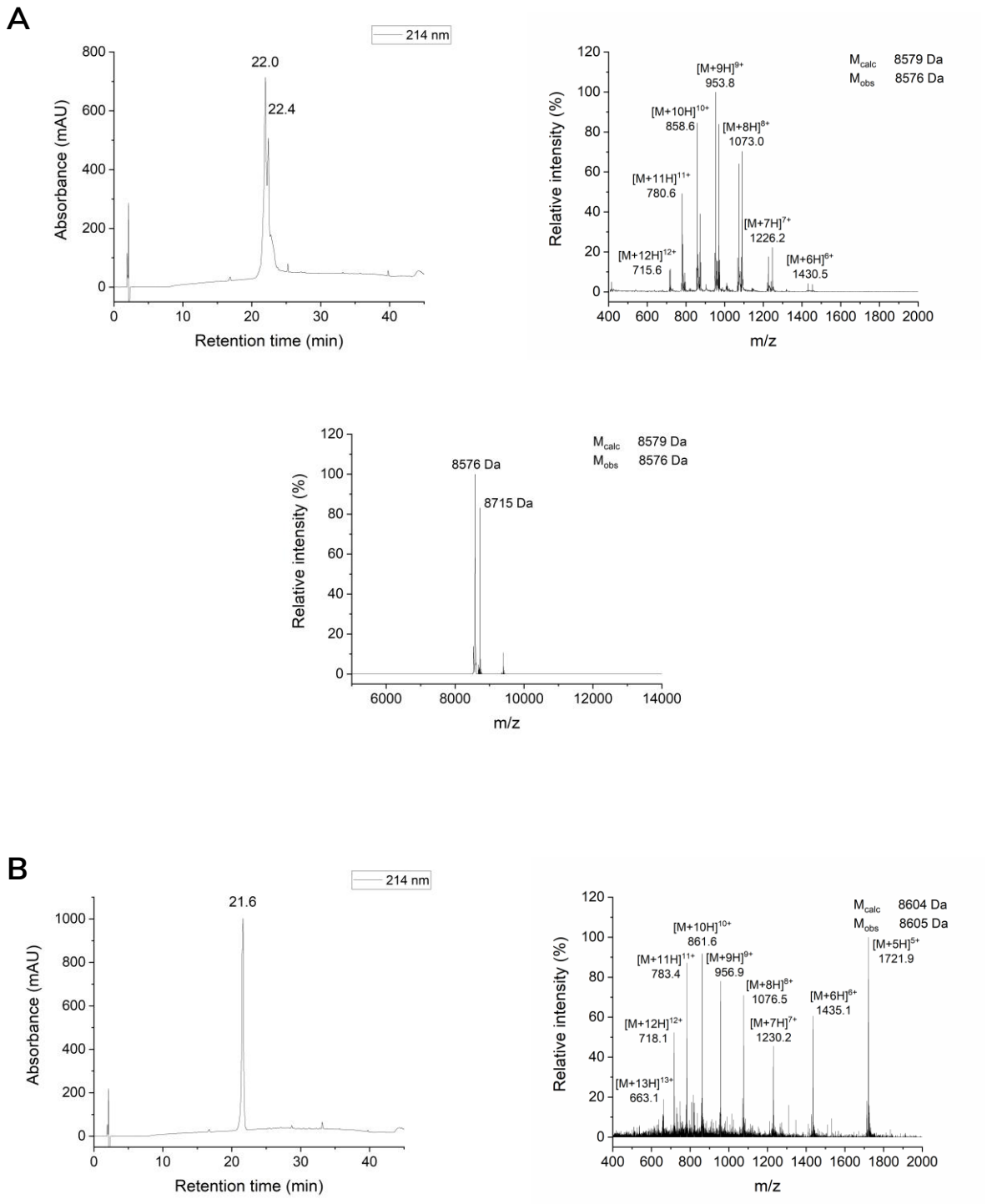


Figure 16. A shows the final analysis data of Ub(K63C)-AA. On the top left, the chromatogram of analytical HPLC is shown, with two product eluting at 22.0 and 22.4 minutes. On the top right, the MS-data of the same sample is shown, with the expected and observed mass of the protein. Another series indicates the presence of a by-product. Deconvoluted MS shows both the mass of the desired product (8576 Da) and the unidentified by-product (8715 Da). B shows the final analysis data of Ub-AA. On the left, the chromatogram of analytical HPLC is shown, with the product eluting at 21.6 minutes. On the right, the MS-data of the same sample is shown, with the expected and observed mass of the protein. Small amounts of hydrolysed product were found (8565 Da).

3.4.2. Conversion of C-terminal hydrazide to C-terminal allylamide

3.4.2.1. Thioester to allylamide conversion protocol

Ub(K48C)-$NHNH_2$ has also been examined to see whether a direct conversion to Ub(K48C)-AA, as in case of Ub(KxC)-SR, was possible. To this end, the same conditions found to convert Ub(KxC)-SR to Ub(KxC)-AA were used with Ub(K48C)-$NHNH_2$. The protein was dissolved in ddH_2O with 500 mM allylamine, resulting in a protein concentration of 1mg/ml and a pH of 12.0. The reaction was observed over 24 hours. Samples were taken at 0 h, 1 h, 2 h, and 24 h and analysed by LC-MS. No changes could be observed after 0 h, 1 h, and 2 h (Figure 17). Also, after 24h, only the mass of Ub(K48C)-$NHNH_2$ and some hydrolysed protein could be observed (data not shown). An attempt to purify the reaction solution by HPLC resulted in no desired product.

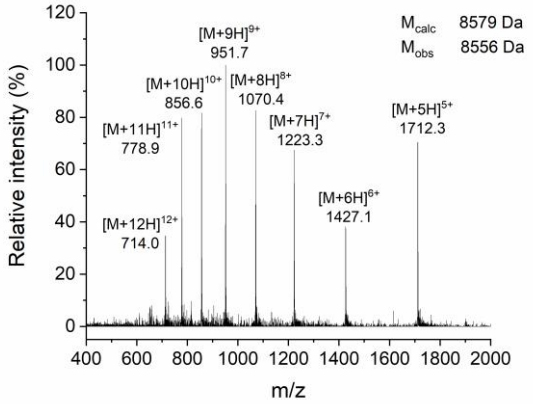


Figure 17. MS-spectrum of Ub(K48C)-$NHNH_2$ conversion to Ub(K48C)-allylamine after 2 h of reaction time. The starting material did not react with allylamine, as seen by the unchanged mass (8553 Da).

3.4.2.2. C-terminal hydrazide to allylamide conversion via azide intermediate

As reacting Ub(K48C)-$NHNH_2$ directly with allylamine, was not successful, an established protocol to exchange the $NHNH_2$ with an amine group was adapted (34). The original protocol was established for NCL in protein synthesis. In this transformation, Ub(K48C)-$NHNH_2$ was converted to Ub(K48C)-N_3, which was then either reacted directly with allylamine or reacted with MPAA after which allylamine was added.

For both reaction variants, Ub(K48C)-$NHNH_2$ was dissolved in a filtered and degassed reaction buffer, containing 0.2 M Na_2HPO_4, 6 M Gua-HCl, pH 3.0, resulting in a protein concentration of 2 mg/mL. The solution was placed in a -15°C ice-salt bath, 20 mM $NaNO_2$ was added to the solution, stirred for 15 minutes and a sample for LC-MS was taken. Then, either 500 mM allylamine or 200 mM MPAA was added to the reaction and the pH was adjusted to 7.0 with NaOH. The reaction was stirred at room temperature. After 1 h, samples were taken. To the reaction containing MPAA, 500 mM allylamine was added and stirred for 1 h, after which a sample was taken. All samples were immediately analysed by LC-MS and the reactions were stored at -80°C.

The LC-MS analysis of the first sample (Figure 18), taken 15 minutes after $NaNO_2$ addition, showed that most of the starting material reacted to Ub(K48C)-N_3, while a small part of the starting material hydrolysed to Ub-K48C-COOH.

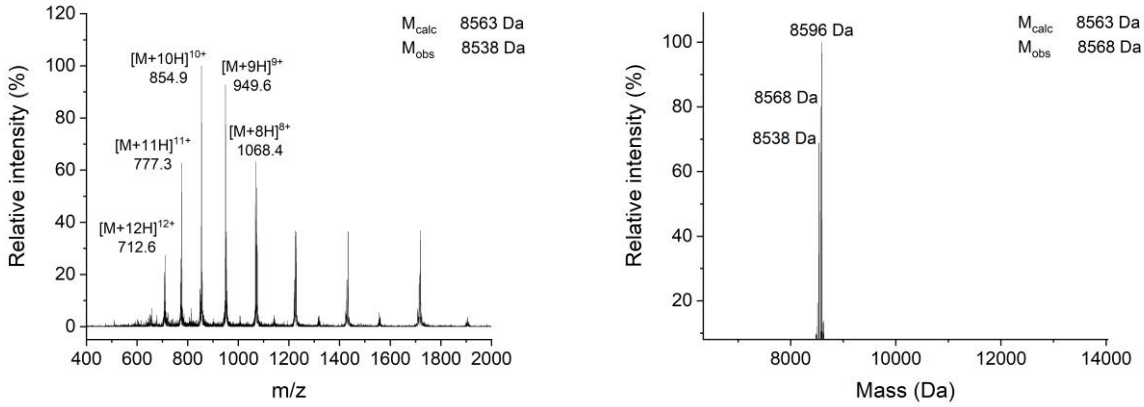


Figure 18. MS-data of Ub(K48C)-$NHNH_2$ to Ub(K48C)-N_3. In the original mass spectrum, the observed main series, is of the hydrolysis product (8538 Da). However, there are also two more series. After deconvolution, the mass of the desired product and its sodium adduct (8568 Da and 8596 Da, respectively) are also detected.

Once the azide intermediate was generate a reaction with either MPAA, as suggested by the original protocol, and then allylamine was performed or allylamine was directly added to the azide intermediate, as in theory a direct exchange of the azide group with the amine group of allylamine was plausible.

Following the original protocol and adding MPAA, one hour after the addition, LC-MS of the reaction was performed (Figure 19). The LC-MS showed that the azide group had been transformed into the MPAA group. As Ub(K48C)-$NHNH_2$ degraded at the azide stage, some hydrolysis product was also detected *via* LC-MS. Further, two additional masses were found. One mass indicating the addition of a second MPAA to the protein (mass 8859 Da) and the other mass of 17049 Da was not possible to identify.

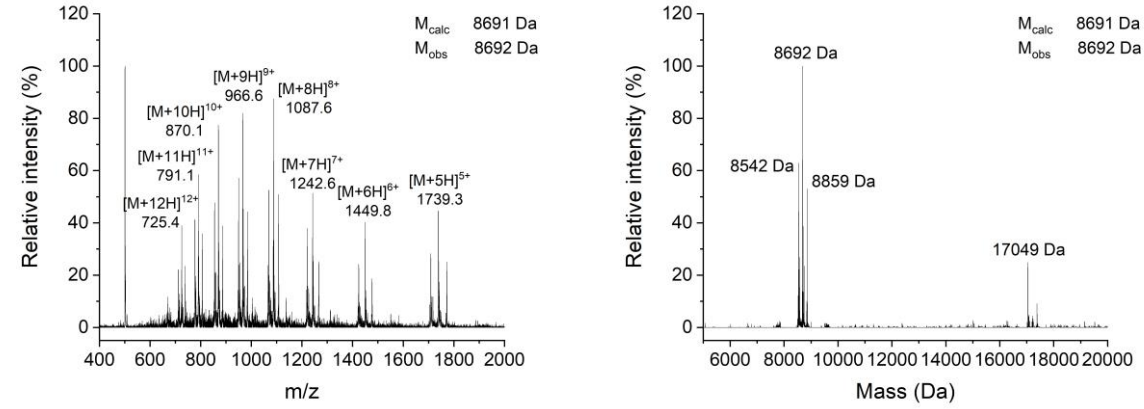


Figure 19. MS-data of Ub(K48C)-N_3 to Ub(K48C)-MPAA. In the original mass spectrum, the observed main series, is of the desired product (8692 Da). However, there are also two more series. After deconvolution, the mass of the hydrolysis product (8542 Da) and two additional masses (8859 Da and 17049 Da) were detected. The mass of 8859 Da is suspected to be the addition of a second MPAA, while 17049 Da were not identified.

After the sucessful conversion with MPAA, allylamine was added and 1 h after addition a sample was taken and LC-MS analysis performed (Figure 20). The MS-data indicated that the MPAA group was

exchanged by the allylamine, however, this was not the favored reaction. Instead, allylamine reacted with the MPAA group to form a new product (mass 8748 Da) (suggested by-product structure in supplementary data). Further, an mass at 8709 Da was detected, which was not possible to identify.

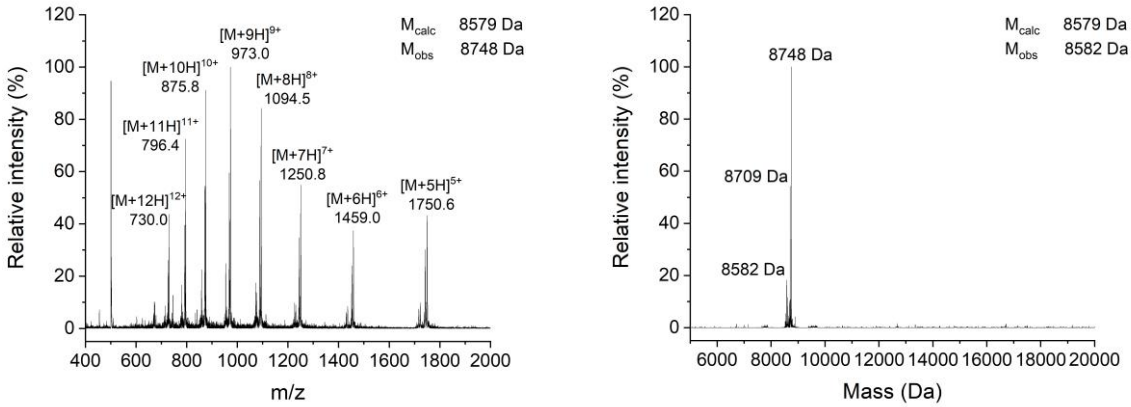


Figure 20. MS-data of Ub(K48C)-MPAA to Ub(K48C)-AA. In the original mass spectrum, the observed main series, is of a by-product (8748 Da). However, there are also two more series. After deconvolution, the mass of the desired product and an unidentifiable by-product (8582 Da and 8709 Da, respectively) are also detected.

The direct addition of allylamine to the azide intermediate was also examine. One hour after the direct addition LC-MS analysis of the reaction was performed (Figure 21). The mass spectrum showed that the azide group had been fully converted into the allylamide. As some Ub(K48C)-$NHNH_2$ degraded at the azide intermediate, some hydrolysis product was also detected *via* LC-MS.

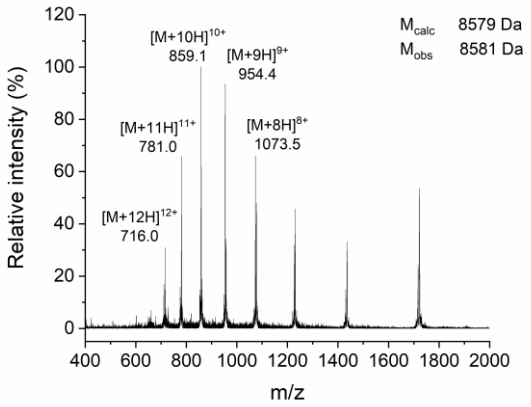


Figure 21. MS-data of Ub(K48C)-N_3 to Ub-K48C-AA after 1 h reaction time. The expected mass of the desired protein (8579 Da) was detected.

From these experiments, it was concluded that the one-pot reaction from Ub(K48C)-$NHNH_2$ over an azide intermediate to Ub(K48C)-allylamine, without the addition of MPAA, was the best method to generate Ub(K48C)-allylamide. However, the resulting product was not isolated and analysed, therefore, no yield was provided.

3.4.3. Conversion of C-terminal thioester to C-terminal propargylamide

As propargylamine is similar to allylamine and can also be used in the generation of ubiquitin chains, it was tested whether propargylamine can be added to the C-terminus as previously shown for allylamine. Therefore, Ub-SR was dissolved in ddH_2O containing 500 mM propargylamine. After 1 h the desired protein was detected via LC-MS (Figure 22). Hence, the conversion of Ub-SR to Ub-propargylamine (Ub-PA) can be performed in the same manner as Ub-SR to Ub-AA. As the resulting product was not isolated and analysed no yield has been evaluated.

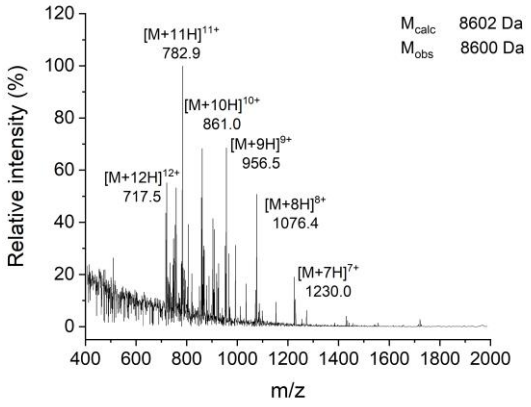


Figure 22. MS-spectrum of Ub-SR to Ub-PA after 1 h reaction time. The expected mass of the desired protein (8600 Da) was detected.

3.5. Thiol-ene coupling

Thiol-ene chemistry is based on a thiol and an alkene reacting to form a thioether. As seen in Figure 23, the thiol-ene reaction used for the creation of di-Ub is initiated by a radical starter, LAP.

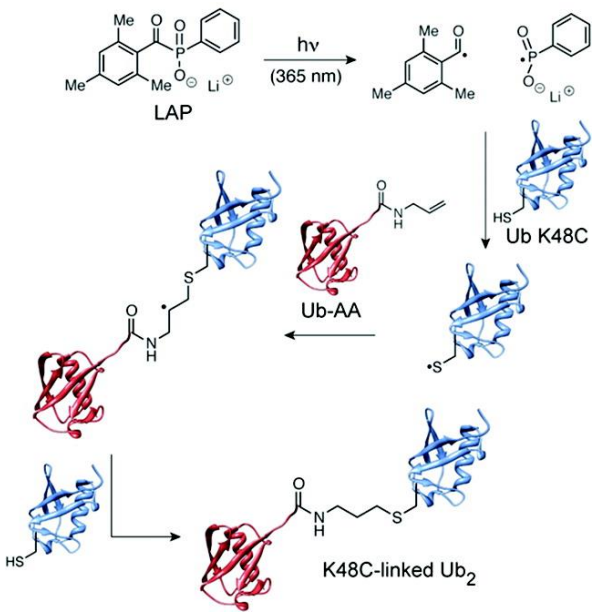


Figure 23. Overview of the thiol-ene coupling of Ub-AA and Ub(K48C) using the UV-light activated LAP radical starter, to generate K48C-linked di-Ub (23).

3.5.1. Finding suitable reaction conditions

The desired thiol-ene reaction of Ub-AA with Ub(KxC) has been previously reported (23). The original protocol called for Ub(KxC)-COOH and Ub-AA to be dissolved in 50 µl of 250 mM NaOAc buffer pH 5.1 containing 0.5 mM LAP, resulting in 1 mM concentrations of each protein. The reaction solution was placed on ice and irradiated at 365 nm for 30 minutes. With these reaction condition Valkevich *et al.* were able to generate six out of seven di-Ub(KxC).

With these promising results, the original protocol was tested for this master thesis. However, the above described conditions were not appliable for the Ub-AA and Ub(KxC) generated during this master thesis, as the proteins were not soluble under the Valkevich *et al* conditions. Therefore, the original protocol had to be changed to dissolve the proteins and generate the desired product. The procedure established during this master thesis consisted of dissolving the proteins in Gua-HCl pH 4.4., 0.5 mM LAP and 0.5 mM TCEP, resulting in protein concentrations of 1 mM. The reaction was performed in an Eppendorf tube and the reaction volume ranged from 70 - 120 µL. The solution was then placed on ice and irradiated for 1 h, resulting in the coupling of Ub-AA with a glutathion and Ub-AA with Ub(K48C)-$NHNH_2$. The coupling of Ub(KxC)-$NHNH_2$ with allylamine did not succeed.

3.5.2. Thiol-ene coupling of Ub(KxC)-hydrazide with allylamine

For future experiments, a test reaction for thiol-ene coupling needed to be established. The first option was to react Ub(KxC)-$NHNH_2$ with allylamine. However, neither the condition of the reported thiol-ene coupling nor the improved conditions resulted in the desired product. LC-MS showed that most of the starting material did not react (data not shown).

3.5.3. Thiol-ene coupling of Ub-allylamide with a thiol-group

The other option for testing thiol-ene coupling was to react Ub-AA with a chemical containing a thiol group. For this, glutathione was chosen. The reaction was performed with the newly established conditions and the desired mass of the conjugate (8912 Da) was detected. However, the MS-data (Figure 24) revealed that most of the starting material did not react and a by-product (8747 Da) was detected. The same observations have been made by Valkevich *et al.* (23) and are further discussed in section 4.5. Further, some of the staring material hydrolysed. After 1 h no further changes in the reaction were observed.

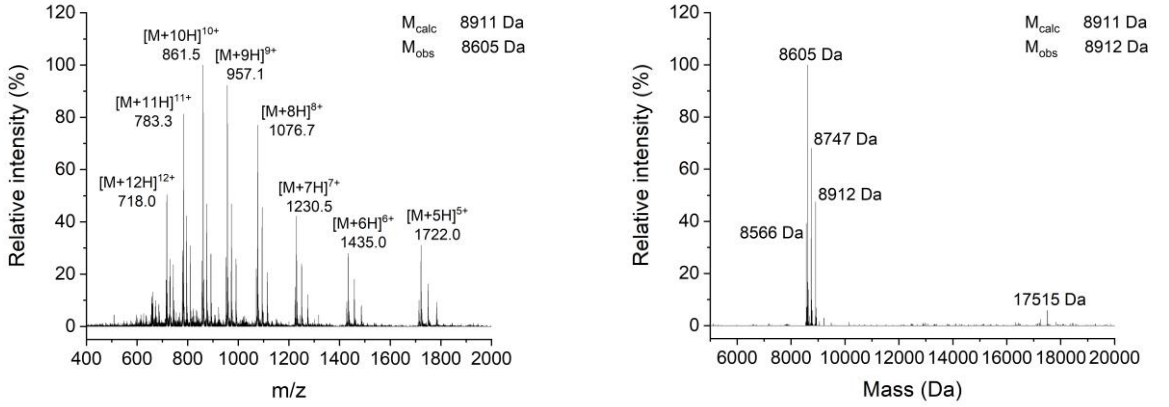


Figure 24. MS-data of thiol-ene reaction of Ub-AA and glutathione. In the original mass spectrum, the observed main series, is of the Ub-AA (8605 Da). However, there are more series. After deconvolution, the mass of the desired product (8912 Da), hydrolysed protein (8566 Da) and a by-product (8747 Da) were detected.

3.5.4. Thiol-ene coupling of Ub-allylamide with Ub(K48C)-hydrazide

After the test thiol-ene reaction of Ub-AA with glutathion succeeded, the thiol-ene coupling of Ub-AA with Ub(K48C)-$NHNH_2$ under newly established conditions was attempted. The reaction did succeed as the desired mass (17159 Da) was detected, however, the MS-data revealed that, as observed in the test reaction, most of the starting material did not react and again the same by-product was identified (Figure 25). No further changes were observed after 1 h of reaction time. Similar results were also observed by Valkevich et *al.* (23) and are discussed in section 4.5.

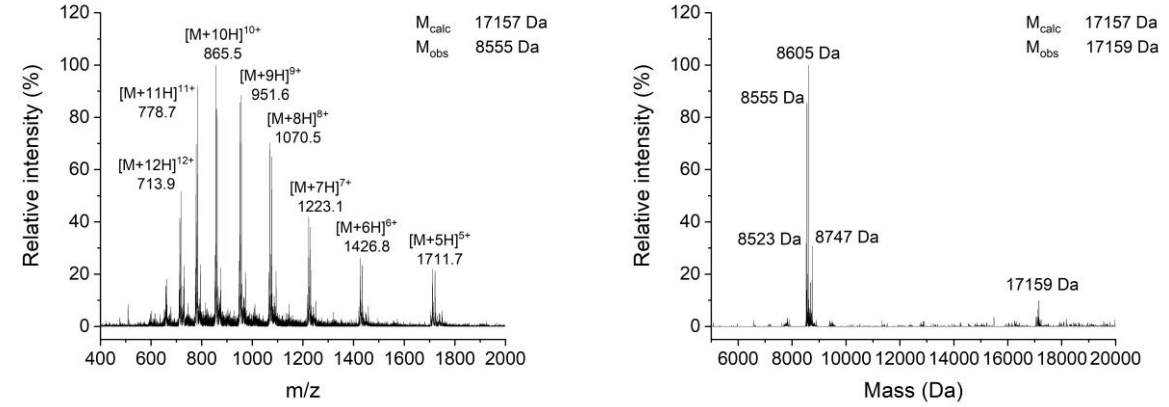


Figure 25. MS-data of thiol-ene reaction of Ub-AA and Ub(K48C)-$NHNH_2$. In the original mass spectrum, the observed main series, is of the Ub(K48C)-$NHNH_2$ (8555 Da). However, there are more series. After deconvolution, the mass of the desired product (17159 Da), starting material (8605 Da and 8555 Da), hydrolysed protein (8523 Da) and of a by-product (8747 Da) were detected.

3.6. Synthesis of thiazolidine protected δ-mercapto-L-Fmoc-lysine

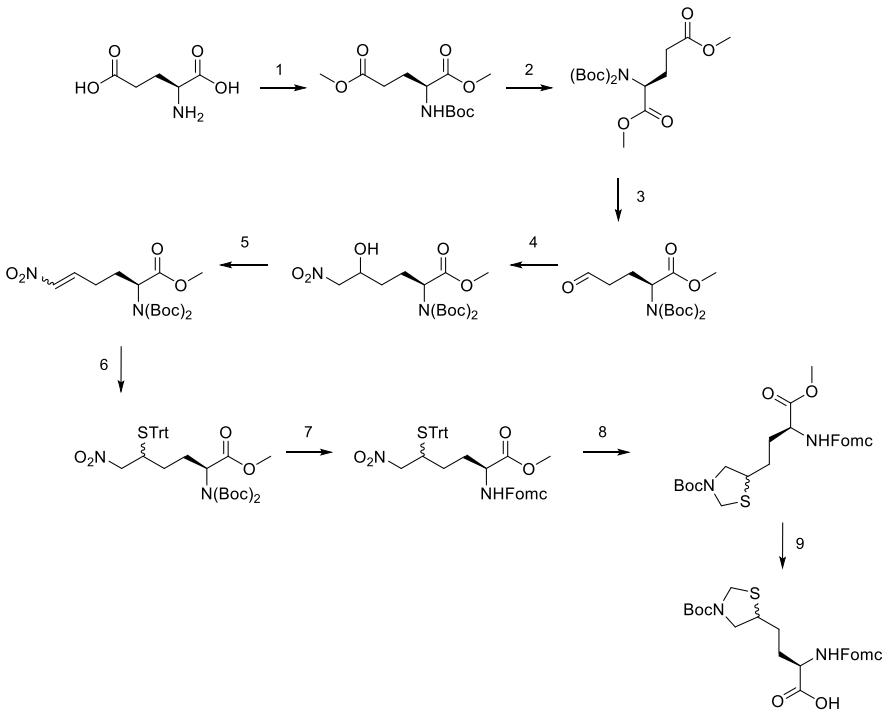


Figure 26. Overwie of full synthesis route for thiazolidine protected δ-mercapto-L-Fmoc-lysine based on (13, 37, 38).

The amino acid δ-mercapto-lysine is commonly used in the isopeptide chemical ligation (ICL) of ubiquitin. Therefore, its synthesis was reproduced. For this several literatures were consulted to generate a synthesis route, which resulted in a nine-step synthesis as seen in Figure 26. In this master thesis, the synthesis was not completed. The detailed synthesis instructions, NMR, and MS-data can be found in the supplementary data.

The first three steps were replicated from Padrón *et al.* (37). Most of the synthesis was performed as described in the experimental section, however, some small changes were made, and a suitable purification protocol was established as no conditions for silica gel column chromatography were described. For step 1, after the addition of Et_3N, the reaction was stirred overnight, instead of regularly monitoring the reaction by TLC, as after 4 hours of stirring no complete protection was observed. For purification, a flash column chromatography (silica gel, ethyl acetate/petroleum ether (1/1)) was performed, which resulted in a yield of 55%. Step 2 was modified in respect to $(Boc)_2O$ being added as 1.6 equiv. right away, instead of two separate portions and the reaction stirring overnight. For purification, a flash column chromatography (silica gel, ethyl acetate/petroleum ether (1/4)) was performed, resulted in a yield of 83%. For step 3, after the DIBAL addition at -78°C, the reaction was stirred for 30 minutes at -78°C. After quenching the reaction, no purification was performed, resulting in a yield of 85%. Step 4 was performed as a test reaction following the protocol described by Kumar *et al.* (13). NMR of the test reaction product confirmed the success of the reaction. Step 5 would have further followed Kumar *et al.* (13), while steps 6 to 9 would have been carried out as described by Haj-Yahya *et al.* (38).

4. Discussion

This thesis aims to generate different types of ubiquitin-building blocks to facilitate the assembly of ubiquitin chains, with a focus on creating ubiquitin-dimers *via* thiol-ene coupling. In the pursuit of this goal, ten different ubiquitin building blocks were generated from which seven were isolated. Two of these ubiquitin-building blocks were coupled *via* thiol-ene chemistry. However, further experimentation is necessary to improve the yield and purity of the examined synthesis of ubiquitin building blocks and their coupling reactions.

4.1. Cloning of ubiquitin-intein fusion constructs

As cysteine does not occur naturally in ubiquitin, a cysteine mutation is a crucial step, in enabling the thiol-ene coupling between ubiquitins. Naturally, the target sites for such a mutation are the lysine positions in ubiquitin, as they enable the linkage of ubiquitins in nature as well. Site-directed mutagenesis of the ubiquitin sequence has therefore become a common practice to generate lysine to cysteine mutations, for polyubiquitin creation. The site-directed mutagenesis of the *E. coli* vector pTXB1-Ub-Mxe-H7-CBD to generate the cysteine mutations Ub(K48C) and Ub(K63C) was achieved, with relative ease, as the mutation was installed at the first attempt. Further, the mutation was installed at all used templet concentrations (undiluted: 65.2 ng/µL, 1:10, and 1:100 dilutions). Next to the K48C and K63C mutations, which were desired in this master thesis, all other lysine positions in ubiquitin have been mutated before, enabling all types of ubiquitin linkages (23). This information is very value if the mutation of other lysine positions is pursued in the future.

4.2. Expression of ubiquitin-intein fusion constructs

The recombinant expression of ubiquitin was chosen as this method enables the generation of large amounts of ubiquitin in a short amount of time. The expressed protein itself was designed as a C-terminal intein-fusion construct followed by seven histidines (H7) and a CBD (Chitin Binding Domain) for purification. The intein Mxe GyrA enables the easy modification of the C-terminus of ubiquitin *via* intein cleavage with different cleavage solutions (further discussion in section 4.3).

As the expressed protein accumulates in the soluble fraction of the cell lysate, the CBD could be used for the main purification of the fusion-protein via chitin beads, while the H7 was implemented as a failsafe in case the CBD purification would have failed or the expressed protein formed inclusion bodies.

The expression of Ub-Mxe-H7-CBD in *E. coli* BL21 (DE3) has previously been established by Gerhard Niederarcher, while the expression of the mutated variants had to be established. The main difference between wild-type and mutants was the amount of overexpression, with the mutants exhibiting lower protein expression in comparison. Therefore, the expression time was prolonged and more efficient strains for the mutants were employed. The Ub(K48C) variant was expressed in all tested *E. coli* strains, with *E. coli* Rosetta 2 exhibiting the most intense protein band. The Ub(K63C) variant was also expressed in all tested *E. coli* strains, however *E. coli* BL21 Gold (DE3) showed an intense band, while the other strain showed a very faint band in comparison. The reason for such a difference in overexpression has not been identified. It has been reported that the *E. coli* strain Rosetta 2(DE3)pLysS can express all Ub(KxC) variants (23), therefore, a switch to this strain may be considered, especially if other Ub(KxC) variants next to the already established K48C and K63C are needed in future experiments.

4.3. Intein-mediated cleavage of Ub- and Ub(KxC)-intein fusion constructs

Intein-mediated cleavage has been chosen, as different C-terminal modifications can be generated *via* this method. As carried out in this thesis, one large batch of 6 l expression can be easily split and treated with different cleaving solutions to afford different C-terminal modifications, which can be used for a wide range of experiments. The desired C-terminal modifications in this master thesis were thioesters and amides, specifically generated with MesNa, hydrazine and allylamine, respectively.

The intein used was Mxe GyrA, a natural non-self-cleaving intein, which undergoes an N-S acyl shift, thereby creating a thioester susceptible to thiols. However, other chemicals can also induce cleavage as seen in the case of hydrazine. The optimal cleavage conditions as recommended by the manufacturer are pH 8.0-8.5 at 4°C. However, the cleavage can be induced at a pH and temperature range of 6.0-9.0 and 4-23°C, respectively.

4.3.1. Thiol-mediated cleavage

As discussed before, the recommended cleavage conditions for the Mxe GyrA intein are at pH 8.0, which has been used in the previous cleavage protocol. However, this protocol resulted in hydrolysis of the generated Ub-SR. Such hydrolysis has also been observed in another study on ubiquitin-like proteins where cleavage of the intein fusion protein with MesNa at pH 8.0 resulted in a 1:4 ratio of hydrolysed product and thioester-protein (35). This is caused by the increased rate of hydrolysis at a basic pH such as 8.0. As hydrolysed Ub-SR is unreactive in further experiments and no way has been found to separate Ub-SR and its hydrolysed counterpart, a procedure to minimize hydrolysis was searched for, to increase the yield of useable protein.

The issue with hydrolysis was solved by cleaving with MesNa at pH 5.5 instead of pH 8.0. At this pH, the thioester is sufficiently below pH 8.0 therefore decreasing hydrolysis effectively, however, the low pH of 5.5 may cause misfolding of the intein, which would compromise the N-S shift and subsequently slow the cleaving and decreased cleavage efficiency. However, no decrease in efficiency was observed *via* SDS-PAGE (see comparison in supplementary section).

4.3.2. Hydrazine-mediated cleavage

The hydrazine cleavage was performed at pH 9.8 and executed for two days, resulting in complete cleavage. Although the hydrazine cleavage was performed at higher pH nearly no hydrolysis of Ub-$NHNH_2$ was observed. This is explained by the stability of amides. As thioesters are stable in mildly acidic conditions and unstable at basic conditions, amides exhibit the opposite properties, remaining stable in basic conditions and hydrolysis occurring under acidic conditions. Therefore, the basic conditions did not affect the stability of Ub-$NHNH_2$. Further, a complete cleavage was observed within one to two days. This may be due to the basic conditions, which are preferable for the N-S shift in Mxe GyrA, or due to hydrazine acting as a stronger nucleophile in comparison to MesNa.

Interestingly, the difference in yield of Ub(K48C)-$NHNH_2$ and Ub(K63C)-$NHNH_2$ is exceptionally high and cannot be explained by the discrepancy in expression strength, as the difference is significantly larger than compared to the yields of Ub(K48C)-SR and Ub(K63C)-SR. Therefore, further investigation of the hydrazine cleavage has to be performed, to check whether this difference is an artifact or has a particular cause.

4.3.3. Allylamine-mediated cleavage

Intein-mediated cleavage directly resulting in a C-terminally attached allylamine would be desirable, as it would reduce time spent on multiple purifications and conversions. To get to this result, two different methods were examined. Either, the cleavage solution consisted solely of allylamine or the cleavage was performed via MesNa and then aminolysis was induced *via* the addition of allylamine.

The first method has not been cited in literature yet, and therefore no established protocol was followed. The reaction conditions were therefore replicated from the MesNa and hydrazine protocols. From the SDS-PAGE it was evident that no cleavage occurred at all. Unfortunately, at the time of testing, it was not considered that the TBS buffer 7.5 used in previous experiments would not buffer 500 mM allylamine sufficiently, thereby resulting in a very basic solution. The resulting pH at the time of testing has not been evaluated, however from subsequent experiments with 500 mM allylamine solutions, a pH of 11 or 12 was very likely. This pH is far above the pH working range of the intein. Therefore, misfolding or aggregation of the intein is likely, which compromises the cleavage, as an intact folding of the intein is needed from cleavage. Hence, a re-examination of allylamine as cleavage solution should be performed, either with a lowered concentration or a stronger buffer system, resulting in a pH of 8.0-9.0 of the cleavage solution.

The second method employed has previously been documented for propargylamine, which is similar to allylamine, with two variations: First, cleaving with MesNa and then performing aminolysis with propargylamine (35), or using a mix of MesNa and propargylamine as cleavage solution (36). In this thesis, the first variation has been examined, where MesNa was used for cleavage, allylamine was added to the solution. However, this resulted in neither Ub-SR nor Ub-AA. A potential cause may be due to the increase in pH as the buffer system was not able to buffer 500 mM allylamine, resulting in rapid hydrolysis of Ub-SR at pH 12.0. Therefore, this method should be re-examined with the addition of a suitable buffering system, and additionally, the second cleavage variation should also be examined.

4.4. Conversion of C-terminal modifications

4.4.1. Conversion of C-terminal thioester to C-terminal allylamide

Previous studies have reported on the generation of Ub-AA, as it has been used in thiol-ene coupling before. In these studies, allylamine has been attached via enzymatic means, either by using the enzyme YUH1 (23) or E1 enzyme (30). These reactions yielded 30 % to close to quantitative conversion, respectively. In this thesis, the conversion of Ub-SR to Ub-AA has been examined.

The conversion of Ub-SR to Ub-AA, of wild-type and cysteine mutants was performed by dissolved purified Ub-SR in ddH_2O containing 500 mM allylamine, resulting in a conversion within minutes. However, depending on the Ub-variation significantly differing yields were observed. Ub-AA yielded good results, with only a small amount of hydrolysed product and a yield of 52 %. In comparison, Ub(K63C)-AA gave non satisfying results, with an unidentified by-product occurring and overall yield of only 16%. No statement on the yield of Ub(K48C)-AA can be made, as the product was not isolated. The difference in the results of Ub-AA and Ub(K63C)-AA cannot be explain with identifying the generate by-product during Ub(K63C)-AA creation. Therefore, further analysis, for example an MS-MS measurement should be performed to identify the by-product. Moreover, further experiments concerned with the purification of Ub(K48C)-AA should be carried out to examine whether both Ub-mutations show (the same) by-products and low yields, to potentially change the synthesis strategy.

4.4.2. Conversion of C-terminal hydrazide to C-terminal allylamide

4.4.2.1. Application of thioester to allylamide conversion protocol

The conversion of Ub(K48C)-$NHNH_2$ to Ub(K48C)-AA was performed in the same manner as Ub(KxC)-SR to Ub(KxC)-AA, however, no reaction was observed over a period of 24 h. This negative result is probably due to $NHNH_2$ being a worse leaving group in comparison to the thiol, as hydrazide are more stable compounds. Therefore, another synthesis route with Ub(K48C)-$NHNH_2$ as starting material has been examined.

4.4.2.2. C-terminal hydrazide to allylamide conversion over azide stage

As the protocol for directly transforming Ub(K48C)-$NHNH_2$ to Ub(K48C)-AA failed, an established protocol, used for native chemical ligation, was tested (34). In this one-pot-reaction, Ub(K48C)-$NHNH_2$ is reacted to Ub(K48C)-N_3, with the generated azide group acting as a better leaving group than the hydrazine group. MPAA, a thiol, is added to convert the azide into a thioester which can subsequently be converted to an allylamide. In theory, allylamine can be added directly added at the azide stage, as the azide group is such a good leaving group. Therefore, both reactions were tested.

The generation of the azide intermediate proved to work well with standard reaction conditions (see 3.4.2.2), as the majority of the starting material reacted to Ub(K48C)-N_3, with some hydrolysis occurring. When comparing the original process of MPAA addition or the direct addition of allylamine, the direct addition of allylamine prove more favourable, as the direct method was not only faster but also without any by-products, compared to the method using MPAA. As these reactions were only performed in small test reactions, the product was not isolated and therefore no yield was obtained. However, if in further experiments the yields prove to be better compared to the yields of the Ub(KxC)-SR to Ub(KxC)-AA, this method may be employed as the main strategy to generate Ub(KxC)-AA. Further experiments incorporating the other Ub-variations should be performed.

4.4.3. Conversion of C-terminal thioester to C-terminal propargylamide

As propargylamine is the alkyne variation of allylamine and Ub-PA has been used in the creation of polyubiquitin, it was examined if propargylamine can be attached to Ub via the same conversion as Ub-SR to Ub-AA. The established protocol was translatable to propargylamine in a test reaction. The isolation of Ub-PA was, however, not performed and therefore no yield was established. Isolating Ub-PA may be of interest, as Ub-PA can be used to generate uncleavable linkages between ubiquitins *via* Cu(I)-catalysed azide-alkyne cycloaddition reaction. For this Ub-PA needs to be reacted with a Ub-azide, which are often generated *via* unnatural amino acid mutagenesis with an azide carrying amino acid (28, 29). These Ub-azide variations differ from the Ub-azide generate during this thesis, as the azide functionality is located in the incorporated amino acid side chain, while in our experiments the azide group is positioned directly at the carboxy-group of the C-terminus. Therefore, the in this study created Ub-azide is not suitable for CuAAC chemistry. Another option to generate a Ub-azide is the C-terminal installation of 2-azidoethylamine as demonstrated by Wang *et al.* (30). Potentially, 2-azidoethylamine could be installed C-terminally in Ub similar to Ub-AA or Ub-PA, as examined in this thesis by the conversion of Ub-SR.

4.5. Thiol-ene coupling

As mention in the results section 3.5.1 the protocol for thiol-ene coupling has been adapted from Valkevich *et al.* (23). In the original protocol, ubiquitin is dissolved in 250 mM NaOAc buffer, however, the ubiquitin-variants generated during this thesis were poorly soluble in NaOAc, which therefore prompted the use of Gua-HCl, a strong denaturant aiding protein solubility. Usually, ubiquitin is well soluble in H_2O, a potential cause for the observed poor solubility may be based on misfolding of ubiquitin, occurring during lyophilisation. However, no CD has been measured to confirm this hypothesis. Once the solubility issues of ubiquitin were solved, the thiol-ene coupling proceeded as described by Valkevich *et al*.

Thiol-ene coupling experiments were performed with three different reactions set consisting of different reagents. Two of these reaction sets were considered test reactions as they were not intended to yield di-Ub, but rather test the thiol-ene rection itself. These two reactions were Ub(KxC)-$NHNH_2$ with allylamine and Ub-AA with a thiol-containing molecule, in this case, glutathione. The Ub(KxC)-$NHNH_2$ with allylamine reaction proved to be unsuccessful, as no thiol-ene coupling was detected via LC-MS. In contrast, the Ub-AA with glutathione coupling was successful and similar to the thiol-ene reaction observed by Valkevich *et al.* Therefore, the thiol-ene reaction of Ub-AA with Ub(K48C)-hydrazine was performed, with MS-data indicating a low yield of di-UB.

The results showed that next to the coupling product, which made up a relatively small portion of the reaction, a large amount of unreacted starting material remained. Further, a portion of Ub-AA reacted with LAP to create an irreversible phosphinate side-product (Figure 27). This side-product formation explains the large amounts of unreacted starting material, as the radical starter was consumed during coupling. Valkevich *et al.* have performed experiments to find a well-working ratio of Ub(KxC) to Ub-AA and to find a reasonable ratio between the UB-AA+LAP and Ub-Ub, however, the reagent amounts used in this thesis were already the optimized conditions (1:1). Further experiments concerning the amount of LAP used and whether subsequent addition of LAP improves yields still need to be examined. Moreover, it could be tested whether Ub(K48C)-hydrazide and LAP should be irradiated first and then Ub-AA added after a while, to ensure that the activated LAP has enough time to react with the thiol of Ub(K48C)-hydrazide, before Ub-AA can interfere. However, this approach may also backfire as more activate LAP may be available to react with Ub-AA once added. These experiments concerning LAP may prove to be vital to improve di-Ub yields.

A further issue of the described protocol is the use of small reaction volumes, consisting of 50 µl overall reaction volume. Valkevich *et al.* needed to combine 15 reaction solutions to generate enough volume for purification. In the experiments during the thesis, the reaction volume was increased to a maximum of 120 µl, which seemed to result in no changes in the efficiency of the overall reaction. However, the purification of this small amount resulted in no viable fraction, resulting in no evaluated yield.

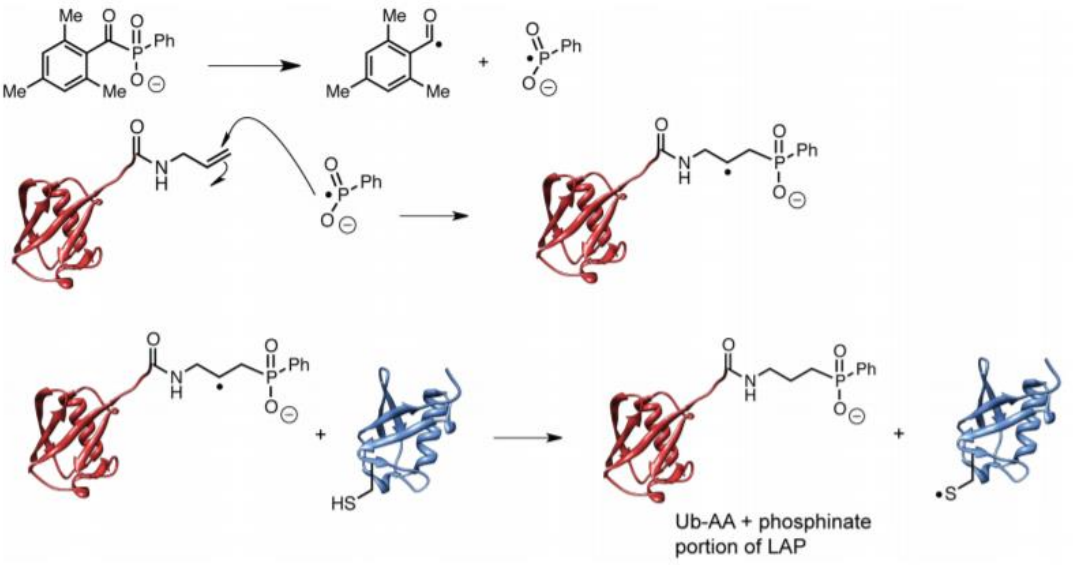


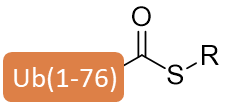
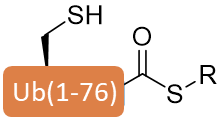
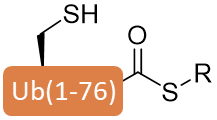
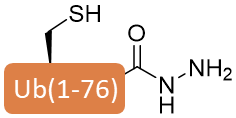
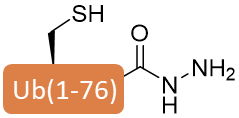
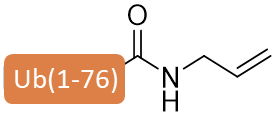
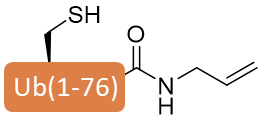
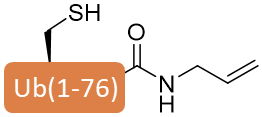
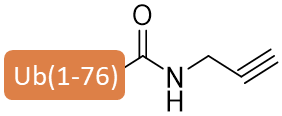
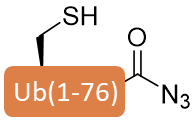
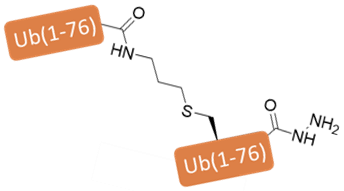
Figure 27. Proposed mechanism for the generation of Ub-AA/phosphinate by-product during thiol-ene coupling by Valkavich et *al.* (23).

To summarise, the thiol-ene reaction protocol adapted from Valkevich et *al*. can be used to generate Ub-dimers. However, further experimentation to improve yield is necessary, as the radical starter is consumed during the reaction, leading to unreacted starting material. A potential solution could lie in changing LAP amounts or changing the radical starter. Another issue that needs to be addressed is the small reaction volume, which is probably due to the high concentrations needed for coupling and the limited power of UV-radiation sources, leading to increased experimentation time as multiple reactions need to be combined for each purification. This issue will become smaller, as the efficiency of the coupling reaction is improved, however, additional experimentation may show that the reaction volume can be upscaled which would, compensate or further improve the coupling reaction.

4.6. Conclusion

During this thesis, ten different ubiquitin building blocks were generated (Table 9), facilitating two different coupling reactions of ubiquitin dimer creation: Cu(I)-catalysed azide-alkyne cycloaddition reaction and thiol-ene chemistry. By recombinant expression of a ubiquitin-intein-fusion construct large amounts of C-terminally modified ubiquitin have been produced, with the resulting ubiquitin building blocks Ub-SR and Ub-$NHNH_2$ forming the base structure for the generation of further modifications like Ub-AA. Moreover, through the site-directed mutagenesis of lysine to cysteine the creation of K48-ubiquitin-dimer *via* thiol-ene coupling was enabled. Although the employed protocol successfully allowed for the creation of a K48-dimer (Table 9), further optimization is needed to ensure greater conversion efficiency and increase yields. In conclusion, the creation of a ubiquitin building block set with various modifications could assist the assembly of different types of homo- and heterotypic ubiquitin chains required for further investigating ubiquitination and its associated complex signalling pathways.

Table 9. Overview of all generated Ub-variations during this master thesis.

	Ub(wildtype)	Ub(K48C)	Ub(K63C)
Ub-thioester (R: $CH_2CH_2SO_3H$)			
Ub-hydrazide			
Ub-allylamide			
Ub-propargylamide			
Ub-azide			
K48 di-Ub via thiol-ene chemistry			

5. References

1. *Mali SM, Singh SK, Eid E, Brik A. Ubiquitin signaling: chemistry comes to the rescue. Journal of the American Chemical Society. 2017;139(14):4971-86.*

2. *Abeywardana T, Pratt MR. Using chemistry to investigate the molecular consequences of protein ubiquitylation. ChemBioChem. 2014;15(11):1547-54.*

3. *Michel MA, Swatek KN, Hospenthal MK, Komander D. Ubiquitin linkage-specific affimers reveal insights into K6-linked ubiquitin signaling. Molecular cell. 2017;68(1):233-46. e5.*

4. *Spasser L, Brik A. Chemistry and biology of the ubiquitin signal. Angewandte Chemie International Edition. 2012;51(28):6840-62.*

5. *Hershko A, Ciechanover A. The ubiquitin system. Annual review of biochemistry. 1998;67(1):425-79.*

6. *Peng J, Schwartz D, Elias JE, Thoreen CC, Cheng D, Marsischky G, et al. A proteomics approach to understanding protein ubiquitination. Nature biotechnology. 2003;21(8):921-6.*

7. *Gopinath P, Ohayon S, Nawatha M, Brik A. Chemical and semisynthetic approaches to study and target deubiquitinases. Chemical Society Reviews. 2016;45(15):4171-98.*

8. *van Tilburg GB, Elhebieshy AF, Ovaa H. Synthetic and semi-synthetic strategies to study ubiquitin signaling. Current opinion in structural biology. 2016;38:92-101.*

9. *Dawson PE, Muir TW, Clark-Lewis I, Kent S. Synthesis of proteins by native chemical ligation. Science. 1994;266(5186):776-9.*

10. *Conibear AC, Watson EE, Payne RJ, Becker CF. Native chemical ligation in protein synthesis and semi-synthesis. Chemical Society Reviews. 2018;47(24):9046-68.*

11. *Mulder MP, Witting KF, Ovaa H. Cracking the ubiquitin Code: the ubiquitin toolbox. Current issues in molecular biology. 2020;37:1-20.*

12. *Yang R, Pasunooti KK, Li F, Liu X-W, Liu C-F. Dual native chemical ligation at lysine. Journal of the American Chemical Society. 2009;131(38):13592-3.*

13. *Ajish Kumar K, Haj-Yahya M, Olschewski D, Lashuel HA, Brik A. Highly efficient and chemoselective peptide ubiquitylation. Angewandte Chemie. 2009;121(43):8234-8.*

14. *Chatterjee C, McGinty RK, Pellois JP, Muir TW. Auxiliary-mediated site-specific peptide ubiquitylation. Angewandte Chemie. 2007;119(16):2872-6.*

15. *Tang S, Liang LJ, Si YY, Gao S, Wang JX, Liang J, et al. Practical chemical synthesis of atypical ubiquitin chains by using an isopeptide-linked Ub isomer. Angewandte Chemie. 2017;129(43):13518-22.*

16. Xin B-T, van Tol BD, Ovaa H, Geurink PP. Native chemical ligation at methionine bioisostere norleucine allows for N-terminal chemical protein ligation. Organic & biomolecular chemistry. 2018;16(34):6306-15.

17. Virdee S, Ye Y, Nguyen DP, Komander D, Chin JW. Engineered diubiquitin synthesis reveals Lys29-isopeptide specificity of an OTU deubiquitinase. Nature chemical biology. 2010;6(10):750-7.

18. Virdee S, Kapadnis PB, Elliott T, Lang K, Madrzak J, Nguyen DP, et al. Traceless and site-specific ubiquitination of recombinant proteins. Journal of the American Chemical Society. 2011;133(28):10708-11.

19. Weller CE, Huang W, Chatterjee C. Facile Synthesis of Native and Protease-Resistant Ubiquitylated Peptides. ChemBioChem. 2014;15(9):1263-7.

20. Pan M, Gao S, Zheng Y, Tan X, Lan H, Tan X, et al. Quasi-racemic X-ray structures of K27-linked ubiquitin chains prepared by total chemical synthesis. Journal of the American Chemical Society. 2016;138(23):7429-35.

21. Qi Y-K, He Q-Q, Ai H-S, Li J-B, Zheng J-S. Convergent total synthesis of histone H2B protein with site-specific ubiquitination at Lys120. Synlett. 2017;28(15):1907-12.

22. Xu L, Huang J-F, Chen C-C, Qu Q, Shi J, Pan M, et al. Chemical synthesis of natural polyubiquitin chains through auxiliary-mediated ligation of an expressed ubiquitin isomer. Organic letters. 2018;20(2):329-32.

23. Valkevich EM, Guenette RG, Sanchez NA, Chen Y-c, Ge Y, Strieter ER. Forging isopeptide bonds using thiol–ene chemistry: site-specific coupling of ubiquitin molecules for studying the activity of isopeptidases. Journal of the American Chemical Society. 2012;134(16):6916-9.

24. Shanmugham A, Fish A, Luna-Vargas MP, Faesen AC, Oualid FE, Sixma TK, et al. Nonhydrolyzable Ubiquitin– Isopeptide Isosteres as Deubiquitinating Enzyme Probes. Journal of the American Chemical Society. 2010;132(26):8834-5.

25. Singh SK, Sahu I, Mali SM, Hemantha HP, Kleifeld O, Glickman MH, et al. Synthetic uncleavable ubiquitinated proteins dissect proteasome deubiquitination and degradation, and highlight distinctive fate of tetraubiquitin. Journal of the American Chemical Society. 2016;138(49):16004-15.

26. Tornøe CW, Christensen C, Meldal M. Peptidotriazoles on solid phase:[1, 2, 3]-triazoles by regiospecific copper (I)-catalyzed 1, 3-dipolar cycloadditions of terminal alkynes to azides. The Journal of organic chemistry. 2002;67(9):3057-64.

27. Rostovtsev VV, Green LG, Fokin VV, Sharpless KB. A stepwise huisgen cycloaddition process: copper (I)-catalyzed regioselective "ligation" of azides and terminal alkynes. Angewandte Chemie. 2002;114(14):2708-11.

28. Rösner D, Schneider T, Schneider D, Scheffner M, Marx A. Click chemistry for targeted protein ubiquitylation and ubiquitin chain formation. Nature protocols. 2015;10(10):1594.

29. Weikart ND, Sommer S, Mootz HD. Click synthesis of ubiquitin dimer analogs to interrogate linkage-specific UBA domain binding. Chemical Communications. 2012;48(2):296-8.

30. Wang XA, Kurra Y, Huang Y, Lee YJ, Liu WR. E1-Catalyzed Ubiquitin C-Terminal Amidation for the Facile Synthesis of Deubiquitinase Substrates. ChemBioChem. 2014;15(1):37-41.

31. Pham GH, Strieter ER. Peeling away the layers of ubiquitin signaling complexities with synthetic ubiquitin–protein conjugates. Current opinion in chemical biology. 2015;28:57-65.

32. Valkevich EM, Sanchez NA, Ge Y, Strieter ER. Middle-down mass spectrometry enables characterization of branched ubiquitin chains. Biochemistry. 2014;53(30):4979-89.

33. Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. nature. 1970;227(5259):680-5.

34. Zheng J-S, Tang S, Qi Y-K, Wang Z-P, Liu L. Chemical synthesis of proteins using peptide hydrazides as thioester surrogates. Nature protocols. 2013;8(12):2483-95.

35. Weikart ND, Mootz HD. Generation of Site-Specific and Enzymatically Stable Conjugates of Recombinant Proteins with Ubiquitin-Like Modifiers by the CuI-Catalyzed Azide–Alkyne Cycloaddition. ChemBioChem. 2010;11(6):774-7.

36. Sommer S, Weikart ND, Brockmeyer A, Janning P, Mootz HD. Expanded click conjugation of recombinant proteins with ubiquitin-like modifiers reveals altered substrate preference of SUMO2-modified Ubc9. Angewandte Chemie International Edition. 2011;50(42):9888-92.

37. Padrón JM, Kokotos G, Martín T, Markidis T, Gibbons WA, Martín VcS. Enantiospecific synthesis of α-amino acid semialdehydes: a key step for the synthesis of unnatural unsaturated and saturated α-amino acids. Tetrahedron: Asymmetry. 1998;9(19):3381-94.

38. Haj-Yahya M, Ajish Kumar K, Erlich LA, Brik A. Protecting group variations of δ-mercaptolysine useful in chemical ubiquitylation. Peptide Science. 2010;94(4):504-10.

6. Supplementary data

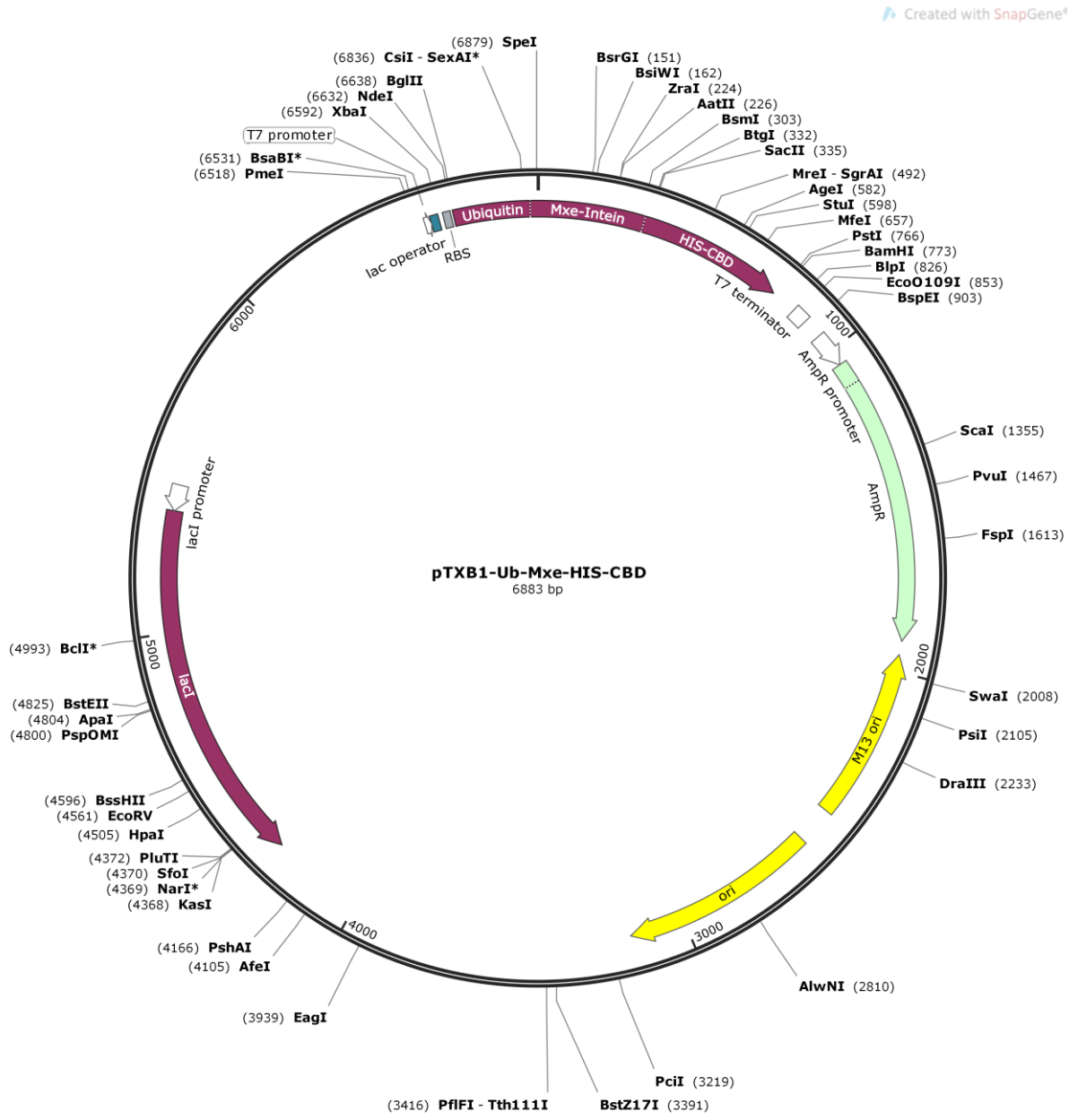
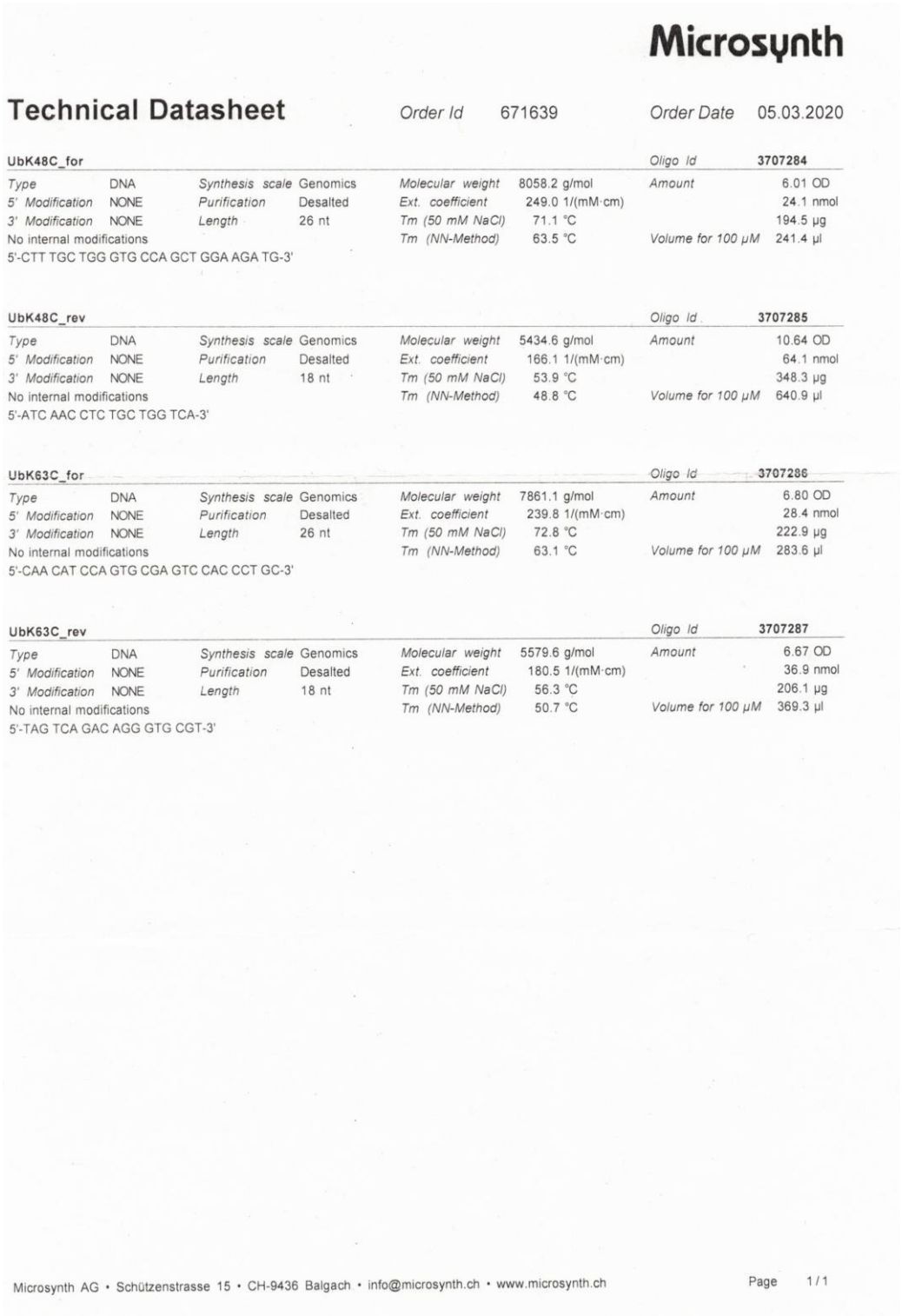


Figure S1. Plasmid map of pTXB1-Ub-Mxe-H7-CBD.



Microsynth

Technical Datasheet Order Id 671639 Order Date 05.03.2020

UbK48C_for Oligo Id **3707284**

Type	DNA	*Synthesis scale*	Genomics	*Molecular weight*	8058.2 g/mol	*Amount*	6.01 OD
5' Modification	NONE	*Purification*	Desalted	*Ext. coefficient*	249.0 1/(mM·cm)		24.1 nmol
3' Modification	NONE	*Length*	26 nt	*Tm (50 mM NaCl)*	71.1 °C		194.5 µg
No internal modifications				*Tm (NN-Method)*	63.5 °C	*Volume for 100 µM*	241.4 µl

5'-CTT TGC TGG GTG CCA GCT GGA AGA TG-3'

UbK48C_rev Oligo Id **3707285**

Type	DNA	*Synthesis scale*	Genomics	*Molecular weight*	5434.6 g/mol	*Amount*	10.64 OD
5' Modification	NONE	*Purification*	Desalted	*Ext. coefficient*	166.1 1/(mM·cm)		64.1 nmol
3' Modification	NONE	*Length*	18 nt	*Tm (50 mM NaCl)*	53.9 °C		348.3 µg
No internal modifications				*Tm (NN-Method)*	48.8 °C	*Volume for 100 µM*	640.9 µl

5'-ATC AAC CTC TGC TGG TCA-3'

UbK63C_for Oligo Id **3707286**

Type	DNA	*Synthesis scale*	Genomics	*Molecular weight*	7861.1 g/mol	*Amount*	6.80 OD
5' Modification	NONE	*Purification*	Desalted	*Ext. coefficient*	239.8 1/(mM·cm)		28.4 nmol
3' Modification	NONE	*Length*	26 nt	*Tm (50 mM NaCl)*	72.8 °C		222.9 µg
No internal modifications				*Tm (NN-Method)*	63.1 °C	*Volume for 100 µM*	283.6 µl

5'-CAA CAT CCA GTG CGA GTC CAC CCT GC-3'

UbK63C_rev Oligo Id **3707287**

Type	DNA	*Synthesis scale*	Genomics	*Molecular weight*	5579.6 g/mol	*Amount*	6.67 OD
5' Modification	NONE	*Purification*	Desalted	*Ext. coefficient*	180.5 1/(mM·cm)		36.9 nmol
3' Modification	NONE	*Length*	18 nt	*Tm (50 mM NaCl)*	56.3 °C		206.1 µg
No internal modifications				*Tm (NN-Method)*	50.7 °C	*Volume for 100 µM*	369.3 µl

5'-TAG TCA GAC AGG GTG CGT-3'

Microsynth AG • Schützenstrasse 15 • CH-9436 Balgach • info@microsynth.ch • www.microsynth.ch Page 1/1

Figure S 2. Ordering list of back-to-back primers designed for the codon change from AAA for lysine to TGC for cysteine at the positions 48 and 63. Purchased from Microsynth.

TCTAGAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGCAGATCTTCGTGAAGACTCTG
ACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG
ATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG
AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTACTCCG
TCTCAGAGGTGGTTGCATCACGGGAGATGCACTAGTTGCCCTACCCGAGGGCGAGTCGGTACGC
ATCGCCGACATCGTGCCGGGTGCGCGGCCCAACAGTGACAACGCCATCGACCTGAAAGTCCTTG
ACCGGCATGGCAATCCCGTGCTCGCCGACCGGCTGTTCCACTCCGGCGAGCATCCGGTGTACAC
GGTGCGTACGGTCGAAGGTCTGCGTGTGACGGGCACCGCGAACCACCCGTTGTTGTGTTTGGTC
GACGTCGCCGGGGTGCCGACCCTGCTGTGGAAGCTGATCGACGAAATCAAGCCGGGCGATTAC
GCGGTGATTCAACGCAGCGCATTCAGCGTCGACTGTGCAGGTTTTGCCCGCGGGAAACCCGAAT
TTGCGCCCACAACCTACACAGTCGGCGTCCCTGGACTGGTGCGTTTCTTGGAAGCACACCACCG
AGACCCGGACGCCCAAGCTATCGCCGACGAGCTGACCGACGGGCGGTTCTACTACGCGAAAGT
CGCCAGTGTCACCGACGCCGGCGTGCAGCCGGTGTATAGCCTTCGTGTCGACACGGCAGACCAC
GCGTTTATCACGAACGGGTTCGTCAGCCACGCTACTGGCCTCACCGGAATTCACCACCACCCAC
CACCACTCCGGTCTGAACTCAGGCCTCACGACAAATCCTGGTGTATCCGCTTGGCAGGTCAACA
CAGCTTATACTGCGGGACAATTGGTCACATATAACGGCAAGACGTATAAATGTTTGCAGCCCCA
CACCTCCTTGGCAGGATGGGAACCATCCAACGTTCCTGCCTTGTGGCAGCTTCAATGACTGCAG
GAAGGGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAA
CAATAACTAGCATAACCCCTTGGGGCCYYTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAG
GAACTW

Figure S3. DNA sequence of Ub-Mxe-H7-CBD. The blue underlined codon shows the beginning of Ub, while green shows the codon for position 48 and yellow the codon for position 63. In both cases the codon is AAA for lysine.

TCTAGAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGCAGATCTTCGTGAAGACTCTG
ACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG
ATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGTGCCAGCTGG
AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTACTCCG
TCTCAGAGGTGGTTGCATCACGGGAGATGCACTAGTTGCCCTACCCGAGGGCGAGTCGGTACGC
ATCGCCGACATCGTGCCGGGTGCGCGGCCCAACAGTGACAACGCCATCGACCTGAAAGTCCTTG
ACCGGCATGGCAATCCCGTGCTCGCCGACCGGCTGTTCCACTCCGGCGAGCATCCGGTGTACAC
GGTGCGTACGGTCGAAGGTCTGCGTGTGACGGGCACCGCGAACCACCCGTTGTTGTGTTTGGTC
GACGTCGCCGGGGTGCCGACCCTGCTGTGGAAGCTGATCGACGAAATCAAGCCGGGCGATTAC
GCGGTGATTCAACGCAGCGCATTCAGCGTCGACTGTGCAGGTTTTGCCCGCGGGAAACCCGAAT
TTGCGCCCACAACCTACACAGTCGGCGTCCCTGGACTGGTGCGTTTCTTGGAAGCACACCACCG
AGACCCGGACGCCCAAGCTATCGCCGACGAGCTGACCGACGGGCGGTTCTACTACGCGAAAGT
CGCCAGTGTCACCGACGCCGGCGTGCAGCCGGTGTATAGCCTTCGTGTCGACACGGCAGACCAC
GCGTTTATCACGAACGGGTTCGTCAGCCACGCTACTGGCCTCACCGGAATTCACCACCACCACC
ACCACCACTCCGGTCTGAACTCAGGCCTCACGACAAATCCTGGTGTATCCGCTTGGCAGGTCAA
CACAGCTTATACTGCGGGACAATTGGTCACATATAACGGCAAGACGTATAAATGTTTGCAGCCC
CACACCTCCTTGGCAGGATGGGAACCATCCAACGTTCCTGCCTTGTGGCAGCTTCAATGACTGCA
GGAAGGGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGA
ACAATAACTAGCATAACCCCTTGGGGCCYYTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGA
GGAACTW

Figure S4. DNA sequence of Ub(K48C)-Mxe-H7-CBD. The blue underlined codon shows the beginning of Ub, while green shows the codon for position 48. For Ub(K48C) the green codon has been switched to TGC for cysteine.

TCTAGAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGCAGATCTTCGTGAAGACTCTG
ACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG
ATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG
AAGATGGACGCACCCTGTCTGACTACAACATCCAGTGCGAGTCCACCCTGCACCTGGTACTCCG
TCTCAGAGGTGGTTGCATCACGGGAGATGCACTAGTTGCCCTACCCGAGGGCGAGTCGGTACGC
ATCGCCGACATCGTGCCGGGTGCGCGGCCCAACAGTGACAACGCCATCGACCTGAAAGTCCTTG
ACCGGCATGGCAATCCCGTGCTCGCCGACCGGCTGTTCCACTCCGGCGAGCATCCGGTGTACAC
GGTGCGTACGGTCGAAGGTCTGCGTGTGACGGGCACCGCGAACCACCCGTTGTTGTGTTTGGTC
GACGTCGCCGGGGTGCCGACCCTGCTGTGGAAGCTGATCGACGAAATCAAGCCGGGCGATTAC
GCGGTGATTCAACGCAGCGCATTCAGCGTCGACTGTGCAGGTTTTGCCCGCGGGAAACCCGAAT
TTGCGCCCACAACCTACACAGTCGGCGTCCCTGGACTGGTGCGTTTCTTGGAAGCACACCACCG
AGACCCGGACGCCCAAGCTATCGCCGACGAGCTGACCGACGGGCGGTTCTACTACGCGAAAGT
CGCCAGTGTCACCGACGCCGGCGTGCAGCCGGTGTATAGCCTTCGTGTCGACACGGCAGACCAC
GCGTTTATCACGAACGGGTTCGTCAGCCACGCTACTGGCCTCACCGGAATTCACCACCACCACC
ACCACCACTCCGGTCTGAACTCAGGCCTCACGACAAATCCTGGTGTATCCGCTTGGCAGGTCAA
CACAGCTTATACTGCGGGACAATTGGTCACATATAACGGCAAGACGTATAAATGTTTGCAGCCC
CACACCTCCTTGGCAGGATGGGAACCATCCAACGTTCCTGCCTTGTGGCAGCTTCAATGACTGCA
GGAAGGGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGA
ACAATAACTAGCATAACCCCTTGGGGCCYYTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGA
GGAACTW

Figure S5. DNA sequence of Ub(K63C)-Mxe-H7-CBD. The blue underlined codon shows the beginning of Ub, while yellow shows the codon for position 63. For Ub(K63C) the yellow codon has been switched to TGC for cysteine.

Microsynth AG | Sequencing Report | 13.05.2020

Economy Run (Tube)
30504572 90256454
Ub K48C_2 T7
NNNNATTCCCTCTAGAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGCAGATCTTCGTGAAGACTCTGACTGGT
AAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCC
CTCCTGACCAGCAGAGGTTGATCTTTGCTGGGTGCCAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA
AGAGTCCACCCTGCACCTGGTACTCCGTCTCAGAGGTGGTTGCATCACGGGAGATGCACTAGTTGCCCTACCCGAGGG
CGAGTCGGTACGCATCGCCGACATCGTGCCGGGTGCGCGGCCCAACAGTGACAACGCCATCGACCTGAAAGTCCTTGA
CCGGCATGGCAATCCCGTGCTCGCCGACCGGCTGTTCCACTCCGGCGAGCATCCGGTGTACACGGTGCGTACGGTCG
AAGGTCTGCGTGTGACGGGCACCGCGAACCACCCGTTGTTGTGTTTGGTCGACGTCGCCGGGGTGCCGACCCTGCTG
TGGAAGCTGATCGACGAAATCAAGCCGGGCGATTACGCGGTGATTCAACGCAGCGCATTCAGCGTCGACTGTGCAGGT
TTTGCCCGCGGGAAACCCGAATTTGCGCCCACAACCTACACAGTCGGCGTCCCTGGACTGGTGCGTTTCTTGGAAGCA
CACCACCGAGACCCGGACGCCCAAGCTATCGCCGACGAGCTGACCGACGGGCGGTTCTACTACGCGAAAGTCGCCAG
TGTCACCGACGCCGGCGTGCAGCCGGTGTATAGCCTTCGTGTCGACACGGCAGACCACGCGTTTATCACGAACGGGTT
CGTCAGCCACGCTACTGGCCTCACCGGAATTCACCACCACCACCACCACCACTCCGGTCTGAACTCAGGCCTCACGAC
AAATCCTGGTGTATCCGCTTGGCAGGTCAACACAGCTTATACTGCGGGACAATTGGTCACATATAACGGCAAGACGTAT
AAATGTTTGCAGCCCCACACCTCCTTGGCAGGATGGGAACCATCCAACGTTCCTGCCTTGTGGCAGCTTCAATGACTGC
AGGAAGGGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGA

Quality control passed.
Status: Finished

Economy Run (Tube)
30504573 90256455
Ub K48C_2 T7term
NANCTTCTTTCGGGCTTTGTTAGCAGCCGGATCCCCTTCCTGCAGTCATTGAAGCTGCCACAAGGCAGGAACGTTGGAT
GGTTCCCATCCTGCCAAGGAGGTGTGGGGCTGCAAACATTTATACGTCTTGCCGTTATATGTGACCAATTGTCCCGCAG
TATAAGCTGTGTTGACCTGCCAAGCGGATACACCAGGATTTGTCGTGAGGCCTGAGTTCAGACCGGAGTGGTGGTGGT
GGTGGTGGTGAATTCCGGTGAGGCCAGTAGCGTGGCTGACGAACCCGTTCGTGATAAACGCGTGGTCTGCCGTGTCGA
CACGAAGGCTATACACCGGCTGCACGCCGGCGTCGGTGACACTGGCGACTTTCGCGTAGTAGAACCGCCCGTCGGTC
AGCTCGTCGGCGATAGCTTGGGCGTCCGGGTCTCGGTGGTGTGCTTCCAAGAAACGCACCAGTCCAGGGACGCCGAC
TGTGTAGGTTGTGGGCGCAAATTCGGGTTTCCCGCGGGCAAAACCTGCACAGTCGACGCTGAATGCGCTGCGTTGAAT
CACCGCGTAATCGCCCGGCTTGATTTCGTCGATCAGCTTCCACAGCAGGGTCGGCACCCCGGCGACGTCGACCAAACA
CAACAACGGGTGGTTCGCGGTGCCCGTCACACGCAGACCTTCGACCGTACGCACCGTGTACACCGGATGCTCGCCGG
AGTGGAACAGCCGGTCGGCGAGCACGGGATTGCCATGCCGGTCAAGGACTTTCAGGTCGATGGCGTTGTCACTGTTGG
GCCGCGCACCCGGCACGATGTCGGCGATGCGTACCGACTCGCCCTCGGGTAGGGCAACTAGTGCATCTCCCGTGATG
CAACCACCTCTGAGACGGAGTACCAGGTGCAGGGTGGACTCTTTCTGGATGTTGTAGTCAGACAGGGTGCGTCCATCTT
CCAGCTGGCACCCAGCAAAGATCAACCTCTGCTGGTCAGGAGGGATGCCTTCCTTGTCTTGGATCTTTGCCTTGACATT
CTCAATGGTGTCACTCGGCTCCACTTCGAGARTGATGGTCTTACCAGTCAGAGTCTTCA

Quality control passed.
Status: Finished

Method
Sanger Cycle Sequencing / Capillary Electrophoresis

Test Laboratory
Microsynth Austria GmbH
Leberstrasse 20
A-1110 Wien
Austria

Figure S6. Sequencing report for Ub(K48C)-intein.

Microsynth AG | Sequencing Report | 13.05.2020

Economy Run (Tube)
30504574 90256456
Ub K63C_1 T7
AATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACTCTC
GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAG
AGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGTGCGAGTCCACCCTGC
ACCTGGTACTCCGTCTCAGAGGTGGTTGCATCACGGGAGATGCACTAGTTGCCCTACCCGAGGGCGAGTCGGTACGCA
TCGCCGACATCGTGCCGGGTGCGCGGCCCAACAGTGACAACGCCATCGACCTGAAAGTCCTTGACCGGCATGGCAATC
CCGTGCTCGCCGACCGGCTGTTCCACTCCGGCGAGCATCCGGTGTACACGGTGCGTACGGTCGAAGGTCTGCGTGTG
ACGGGCACCGCGAACCACCCGTTGTTGTGTTTGGTCGACGTCGCCGGGGTGCCGACCCTGCTGTGGAAGCTGATCGA
CGAAATCAAGCCGGGCGATTACGCGGTGATTCAACGCAGCGCATTCAGCGTCGACTGTGCAGGTTTTGCCCGCGGGAA
ACCCGAATTTGCGCCCACAACCTACACAGTCGGCGTCCCTGGACTGGTGCGTTTCTTGGAAGCACACCACCGAGACCC
GGACGCCCAAGCTATCGCCGACGAGCTGACCGACGGGCGGTTCTACTACGCGAAAGTCGCCAGTGTCACCGACGCCG
GCGTGCAGCCGGTGTATAGCCTTCGTGTCGACACGGCAGACCACGCGTTTATCACGAACGGGTTCGTCAGCCACGCTA
CTGGCCTCACCGGAATTCACCACCACCACCACCACCACTCCGGTCTGAACTCAGGCCTCACGACAAATCCTGGTGTATC
CGCTTGGCAGGTCAACACAGCTTATACTGCGGGACAATTGGTCACATATAACGGCAAGACGTATAAATGTTTGCAGCCC
CACACCTCCTTGGCAGGATGGGAACCATCCAACGTTCCTGCCTTGTGGCAGCTTCAATGACTGCAGGAAGGGGATCCG
GCTGCTAACAAAC

Quality control passed.
Status: Finished

Economy Run (Tube)
30504575 90256457
Ub K63C_1 T7term
CTTTCGGGCTTTGTTAGCAGCCGGATCCCCTTCCTGCAGTCATTGAAGCTGCCACAAGGCAGGAACGTTGGATGGTTCC
CATCCTGCCAAGGAGGTGTGGGGCTGCAAACATTTATACGTCTTGCCGTTATATGTGACCAATTGTCCCGCAGTATAAG
CTGTGTTGACCTGCCAAGCGGATACACCAGGATTTGTCGTGAGGCCTGAGTTCAGACCGGAGTGGTGGTGGTGGTGGT
GGTGAATTCCGGTGAGGCCAGTAGCGTGGCTGACGAACCCGTTCGTGATAAACGCGTGGTCTGCCGTGTCGACACGAA
GGCTATACACCGGCTGCACGCCGGCGTCGGTGACACTGGCGACTTTCGCGTAGTAGAACCGCCCGTCGGTCAGCTCG
TCGGCGATAGCTTGGGCGTCCGGGTCTCGGTGGTGTGCTTCCAAGAAACGCACCAGTCCAGGGACGCCGACTGTGTA
GGTTGTGGGCGCAAATTCGGGTTTCCCGCGGGCAAAACCTGCACAGTCGACGCTGAATGCGCTGCGTTGAATCACCGC
GTAATCGCCCGGCTTGATTTCGTCGATCAGCTTCCACAGCAGGGTCGGCACCCCGGCGACGTCGACCAAACACAACAA
CGGGTGGTTCGCGGTGCCCGTCACACGCAGACCTTCGACCGTACGCACCGTGTACACCGGATGCTCGCCGGAGTGGA
ACAGCCGGTCGGCGAGCACGGGATTGCCATGCCGGTCAAGGACTTTCAGGTCGATGGCGTTGTCACTGTTGGGCCGC
GCACCCGGCACGATGTCGGCGATGCGTACCGACTCGCCCTCGGGTAGGGCAACTAGTGCATCTCCCGTGATGCAACCA
CCTCTGAGACGGAGTACCAGGTGCAGGGTGGACTCGCACTGGATGTTGTAGTCAGACAGGGTGCGTCCATCTTCCAGC
TGTTTCCCAGCAAAGATCAACCTCTGCTGGTCAGGAGGGATGCCTTCCTTGTCTTGGATCTTTGCCTTGACATTCTCAAT
GGTGTCACTCGGCTCCACTTCGAGAGTGATGGTCTTACCAGTCAGAGTCTTCMCGAAGATCTGCATATGTATWTCTCCT
TCTTAAAGTTAAACAAAATTATTTCTAGAGGGGAATTGTTATCCGCTCMCAATTCCCCTAWAGTGAGTCGTATTAATTTCC

Quality control passed.
Status: Finished

Method
Sanger Cycle Sequencing / Capillary Electrophoresis

Test Laboratory
Microsynth Austria GmbH
Leberstrasse 20
A-1110 Wien
Austria

Figure S7. Sequencing report for Ub(K63C)-intein.

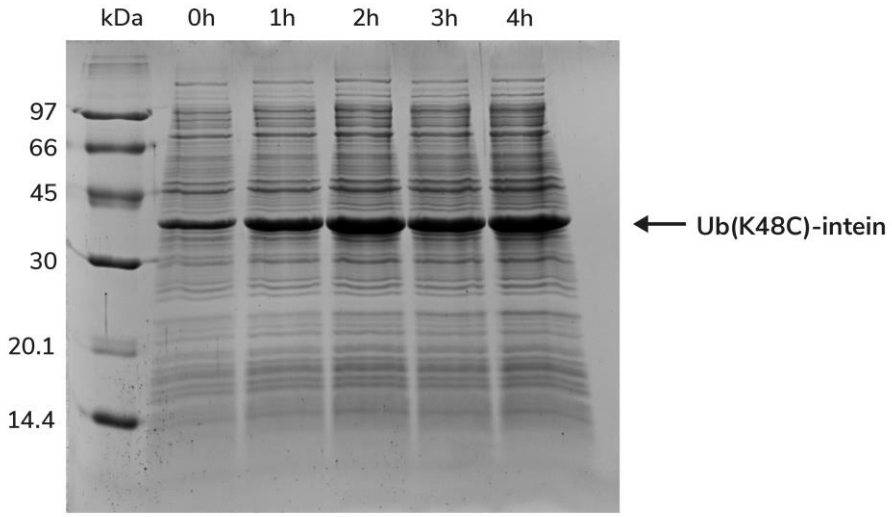


Figure S8. SDS-PAGE of Ub(K48C)-Mxe-H7-CBD expressions in *E. coli* BL21, with samples at time point 0 h, 1 h, 2 h, 3h and 4 h after induction.

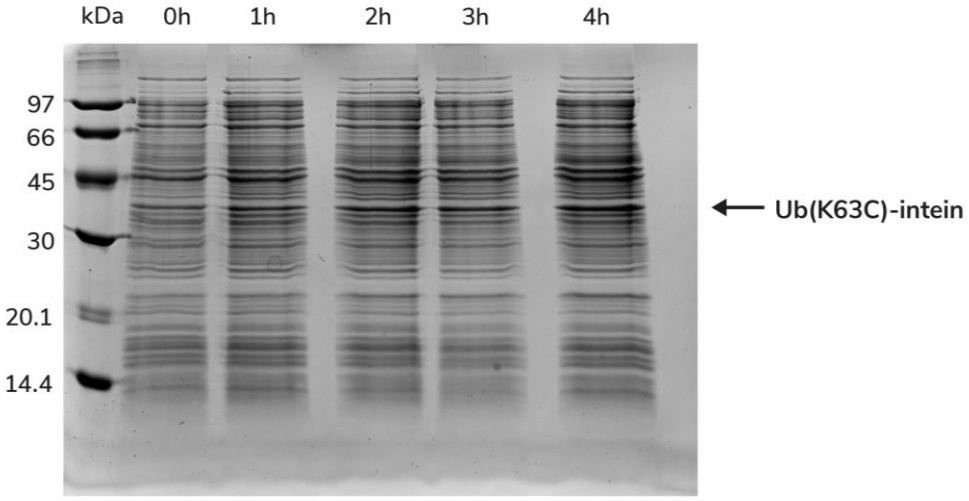


Figure S9. SDS-PAGE of Ub(K63C)-Mxe-H7-CBD expressions in *E. coli* BL21, with samples at time point 0 h, 1 h, 2 h, 3h and 4 h after induction.

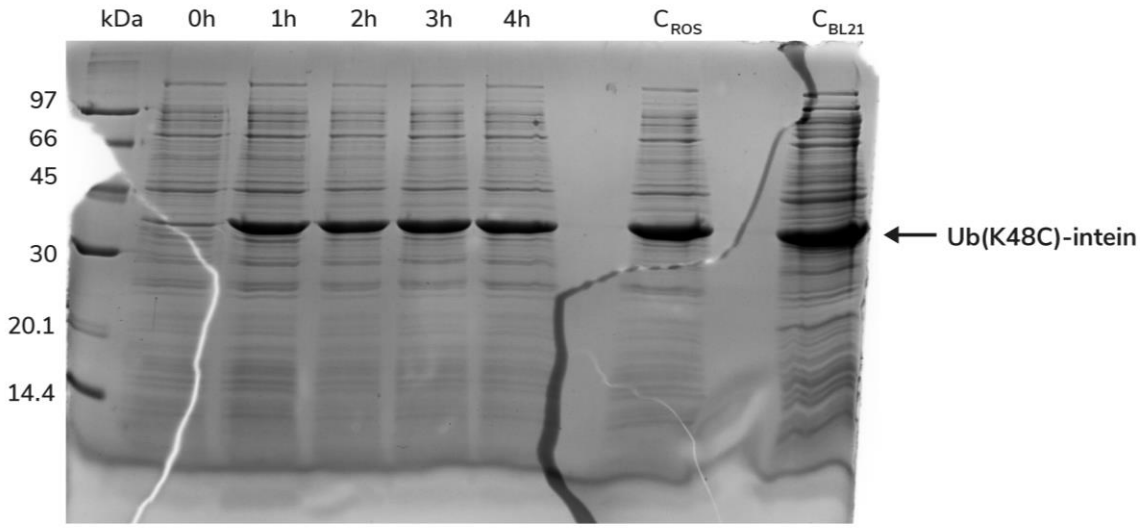


Figure S10. SDS-PAGE of Ub(K48C)-Mxe-H7-CBD expressions in *E. coli* Rosetta 2, with samples at time point 0 h, 1 h, 2 h, 3h and 4 h after induction. Further the crude of Ub(K48C)-Mxe-H7-CBD from the *E. coli* Rosetta 2 (C_{ROS}) and *E. coli* BL21 (C_{BL21}) expression are shown.

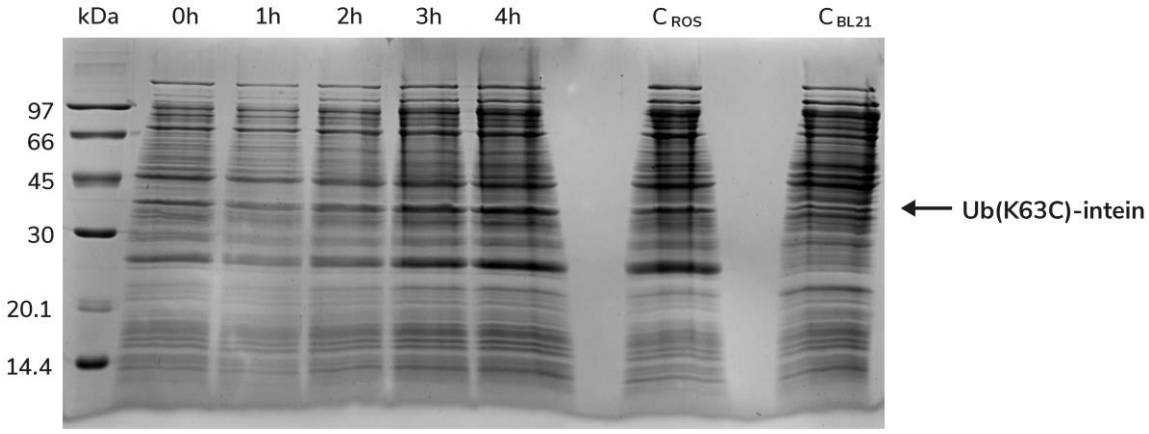


Figure S11. SDS-PAGE of Ub(K63C)-Mxe-H7-CBD expressions in *E. coli* Rosetta 2, with samples at time point 0 h, 1 h, 2 h, 3h and 4 h after induction. Further the crude of Ub(K63C)-Mxe-H7-CBD from the *E. coli* Rosetta 2 (C_{ROS}) and *E. coli* BL21 (C_{BL21}) expression are shown.

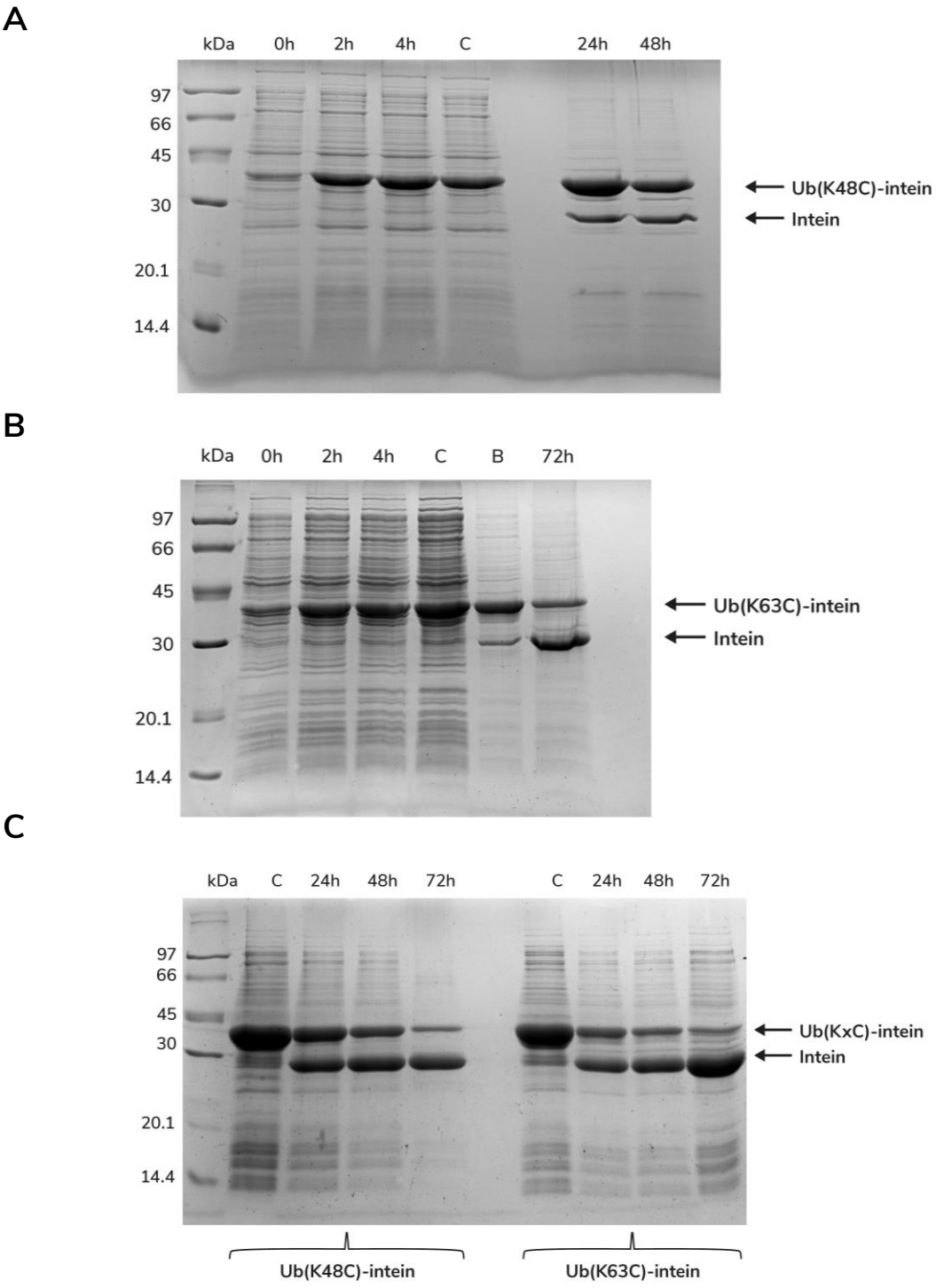


Figure S 12. The SDS-PAGE of MesNa-mediated cleavage for comparison of cleavage efficiency before and after pH change from 8.0 to 5.5. A shows the SDS-PAGE of Ub(K48C)-intein cleaved at pH 8.0 with MesNa at the time points 24 h and 48 h. B shows the SDS-PAGE of Ub(K63C)-intein cleaved at pH 8.0 with MesNa at the time point 0 h and 72 h. C shows the SDS-PAGE of Ub(K48C)-intein and Ub(k63C)-intein cleaved at pH 5.5 with MesNa at the time points 24 h, 48 h and 72 h.

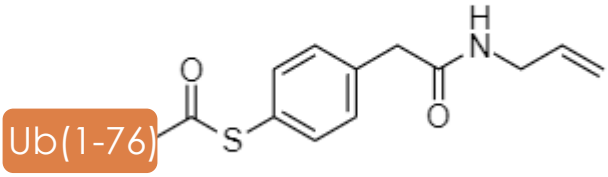
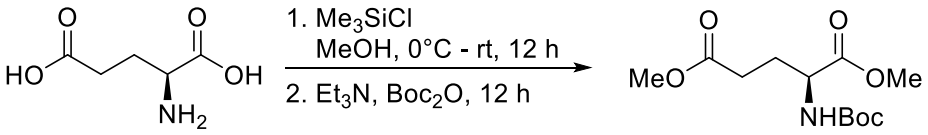


Figure S 13. Suggested by-product (mass 8748 Da) structure from allylamine reacting with the MPAA group of Ub-MPAA (from section 3.4.2.2. Figure 21.)

Dimethyl (S)-2-tertbutoxycarbonylamino-pentanodioate



1

L-glutamic acid

2, 55%

L-glutamic acid (1.47 g, 10 mmol) was suspended in dry MeOH (33 ml, 0.3 M) under Ar. Then Me_3SiCl (5.6 ml, 44 mmol, 4.4 equiv.) was added slowly with the flask in an ice-cold bath. After the addition was completed the ice-cold bath was removed and the reaction was stirred overnight until TLC showed complete conversion. Then, Et_3N (9 ml, 65 mmol, 6.5 equiv.) and $(Boc)_2O$ (2.4 g, 11 mmol, 1.1 equiv.) were added one after the other. The reaction mixture was stirred overnight. The solvent was removed under reduced pressure and the residue thoroughly washed with Et_2O (3×150 ml, 15 ml/mmol), which was filtered using a pad of Celite. The combined organic layers were evaporated, and the resulting crude was purified by silica gel chromatography, ethyl acetate/petroleum ether (1/1), resulting in a yield of 55%: 1H NMR ($CDCl_3$) δ 1.44 (s, 9H), 1.91 (m, 1H), 2.20 (m, 1H), 2.38 (m, 2H), 3.68 (s, 3H), 3.74 (s, 3H), 4.33 (br s, 1H), 5.10 (br s, 1H).

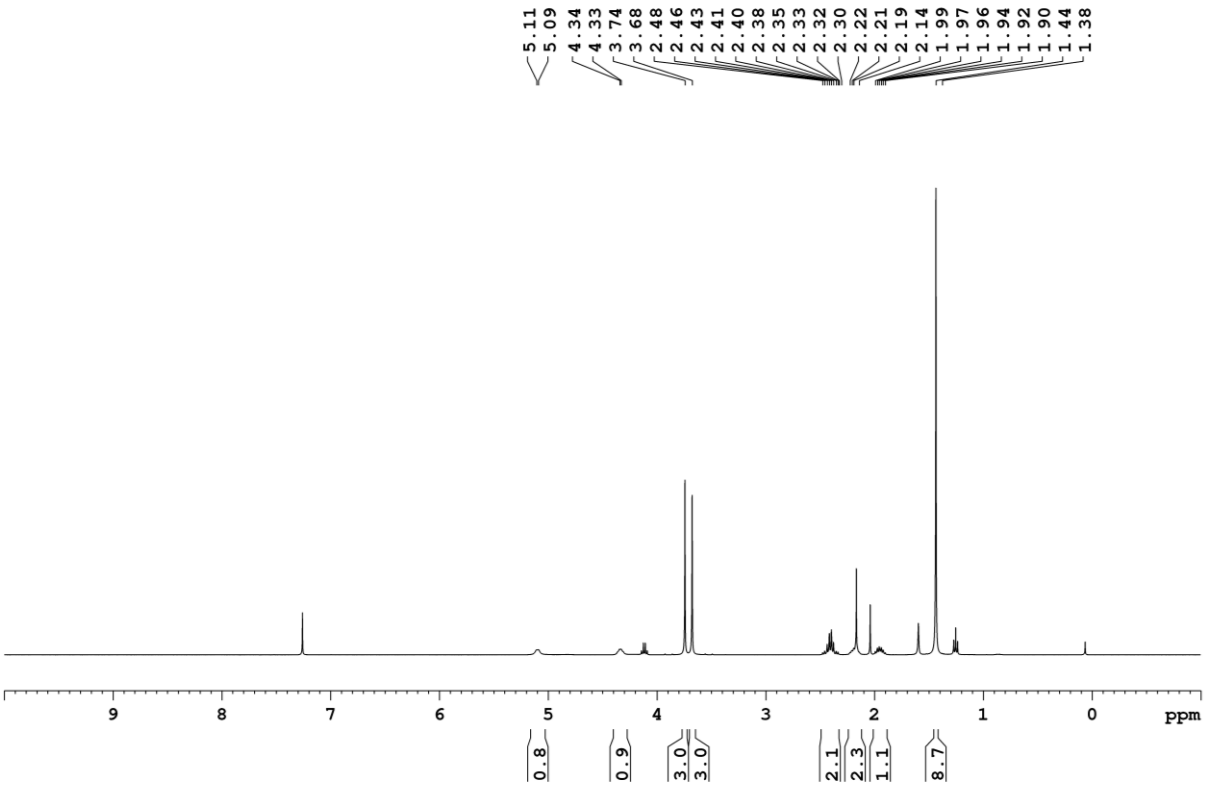
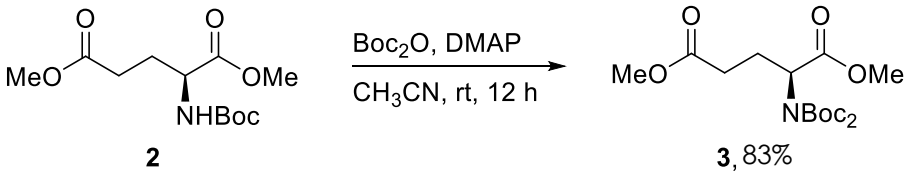


Figure S 14. 1-NMR of substance 2 in $CDCl_3$ with traces of ethyl acetate from purification.

1-(3-fluoro-4-methoxyphenyl)-2-(2-(phenylthio)phenyl)ethan-1-one



2 **3**, 83%

N-Boc amino ester **2** (0.90 g, 3.3 mmol) and DMAP (79 mg, 0.6 mmol, 0.2 equiv.) were stirred in dry CH_3CN (11 ml) under Ar, before $(Boc)_2O$ (1.1 g, 5.2 mmol, 1.6 equiv.) was added at rt. The reaction became slightly red with gas evolution. The mixture was stirred overnight. The solvent was evaporated, and the crude purified by silica gel chromatography, ethyl acetate/petroleum ether (1/4), resulting in a yield of 83% yield: 1H NMR ($CDCl_3$) δ 1.49 (s, 18H), 2.20 (m, 1H), 2.42 (m, 2H), 2.46 (m, 1H), 3.67 (s, 3H), 3.71 (s, 3H), 4.93 (dd, 1H).

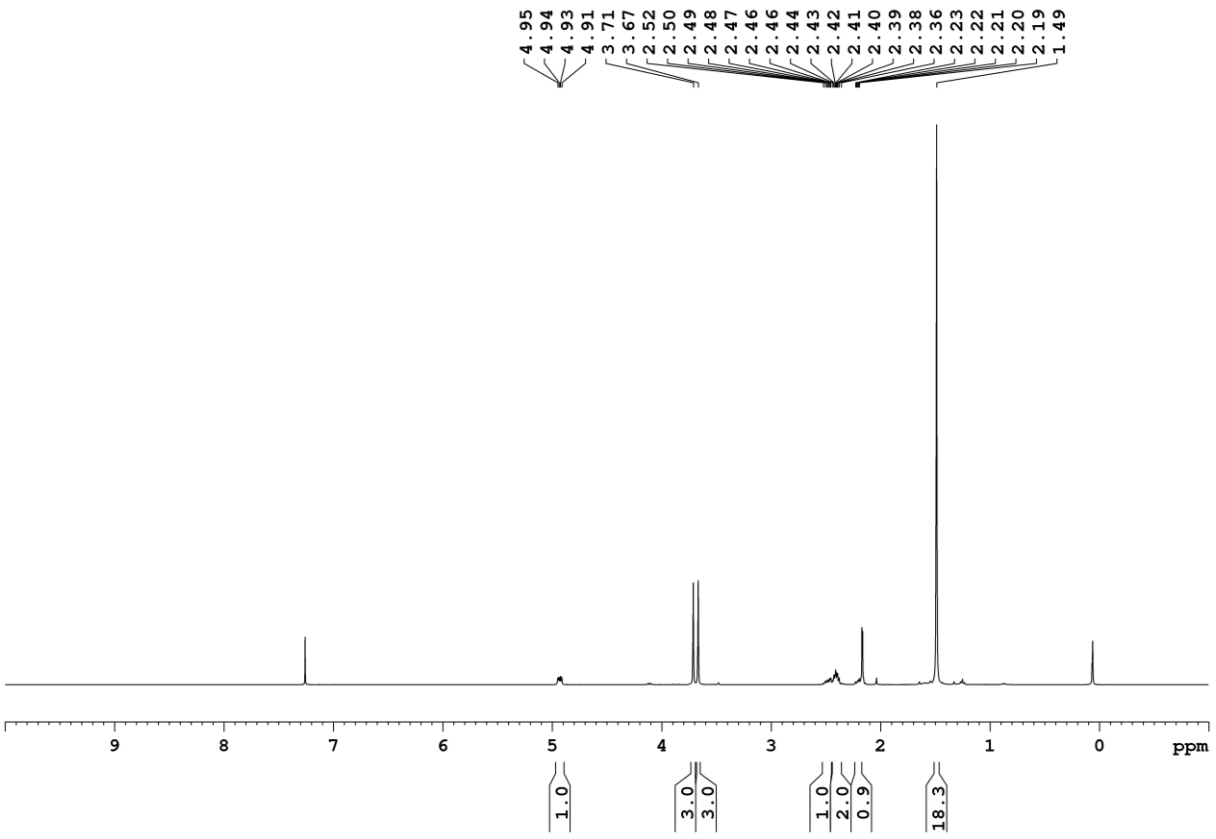
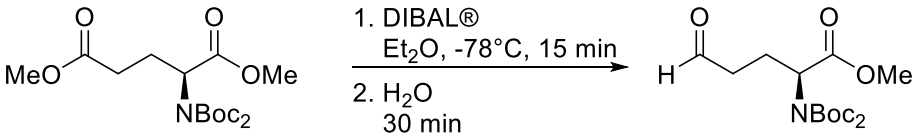


Figure S 15. 1-NMR of substance 3 in $CDCl_3$ with traces of ethyl acetate from purification.

Methyl (2S)-2-{(tert-butoxy)-N-[(tert-butyl)oxycarbonyl]-carbonylamino}-5-oxopentanoate



3 **4**, 85%

The dimethyl ester **3** (0.67 g, 1.8 mmol) was suspended in dry Et_2O (18 ml) under Ar, before dropwise addition of DIBAL® (2 ml, 1.0 M in hexane, 2 mmol, 1.1 equiv.) at −78°C. The reaction was stirred for 30 min and then quenched with H_2O (0.27 ml, ≈7 equiv.). The mixture was stirred for another 30 min, dried over $MgSO_4$, and filtered through a pad of Celite. The solvent was evaporated, and resulted in a yield of 8% yield: 1H NMR ($CDCl_3$) δ 1.49 (s, 18H), 2.18 (m, 1H), 2.50 (m, 2H), 2.60 (m, 1H), 3.71 (s, 3H), 4.89 (dd, 1H), 9.76 (s, 1H).

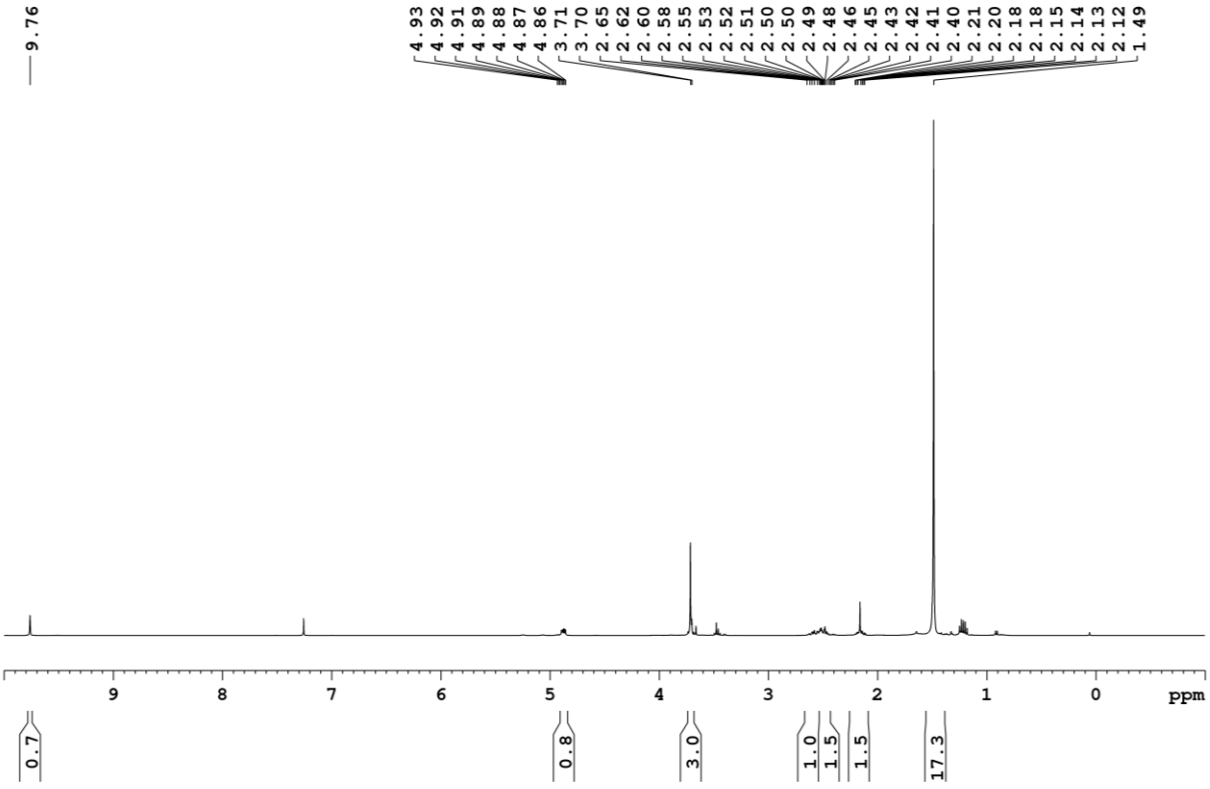
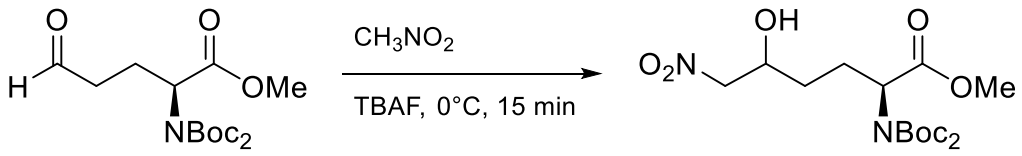


Figure S 16. 1-NMR of substance 4 in $CDCl_3$ with traces of ethyl acetate from reaction solvent.

Methyl (5R,S)-hydroxy-6-nitro-(2S)-2-{(tert-butoxy)-N-[tert-butyloxy carbonyl] carbonylamino}-hexanoate



4 **5**

The aldehyde **4** (0.137 g, 0.6 mmol) was suspended in dry nitromethane (0.44 ml) at -5 °C under Ar, before TBAF (1M solution in THF, 0.20 ml, 0.2 mmol, 0.3 equiv.) was added and the reaction mixture was stirred at 0 °C. After 15 min the reaction mixture was analysed with NMR: 1 H NMR ($CDCl_3$): δ 1.44 (bs, 18H), 1.45 (m, 2H), 1.90-2.07 (m, 1H), 2.21-2.41 (m, 1H), 2.96 (br dd, 1H), 3.66 (s, 3H, OCH3), 4.39 (m, 3H), 4.85 (m, 1H).

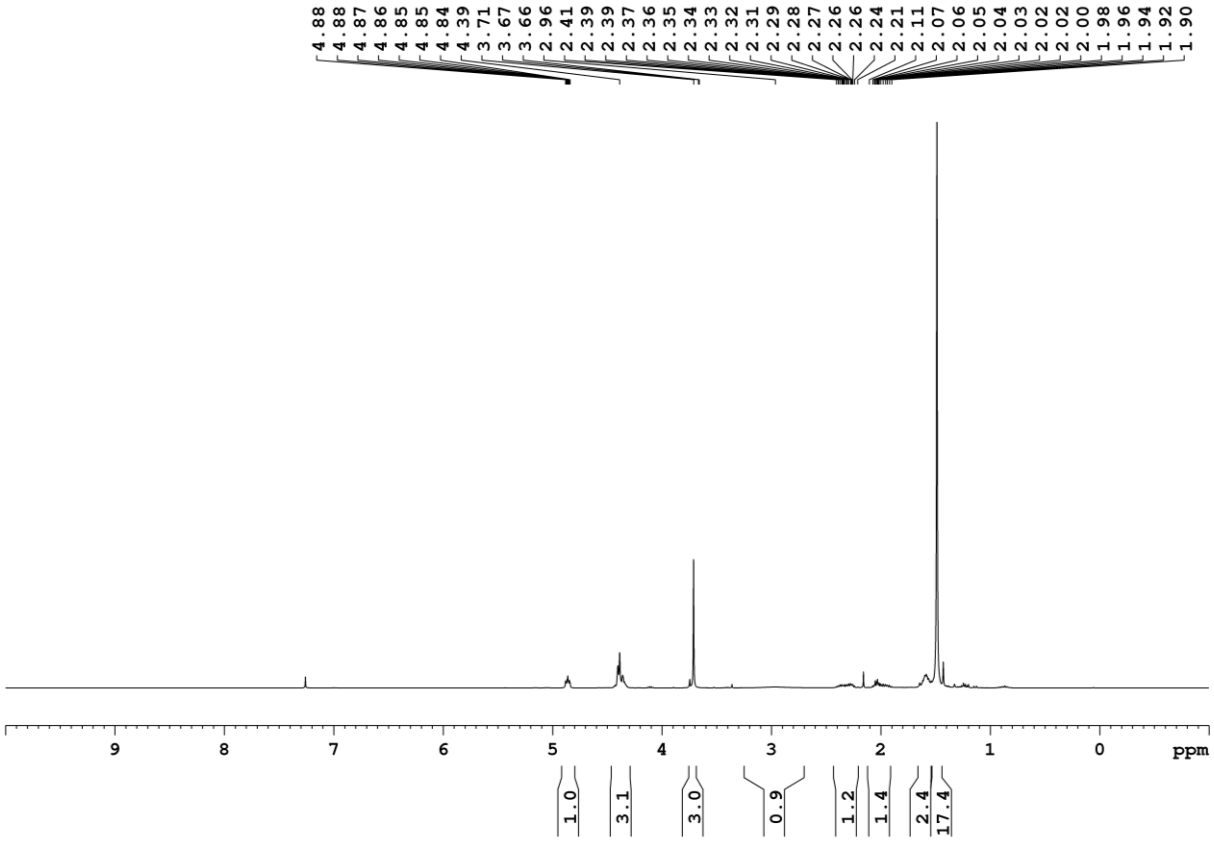


Figure S 17. 1-NMR of substance 5 in $CDCl_3$ with traces of ethyl acetate from aldehyde educt.