Characterization of conformational dynamics of Lys48-linked ubiquitin chains as design frameworks for creating allosterically controllable multidomain proteins

Methanee HIRANYAKORN

Department of Functional Molecular Science
School of Physical Sciences
The Graduate University for Advanced Studies, SOKENDAI

2021 September

A dissertation submitted for the degree of Doctor of Philosophy

List of published thesis

The chapter 2 of this doctoral thesis is mainly a reprint of the material as it appears in "NMR Characterization of conformational interconversions of Lys48-linked ubiquitin chains" as listed below. The co-author listed in this publication directed and supervised the research that forms the basis for the doctoral thesis.

Hiranyakorn, M., Yanaka, S., Satoh, T., Wilasri, T., Jityuti, B., Yagi-Utsumi, M. and Kato, K. NMR Characterization of conformational interconversions of Lys48-linked ubiquitin chains *Int. J. Mol. Sci.* **21**, 5351 (2020).

Abbreviations:

Ub	: Ubiquitin
ATP	: Adenosine trihosphate
DTT	: Dithiothreitol
TCEP	: Tris (2-carboxyethyl) phosphine
IPTG	: Isopropyl β-D-1-thiogalactopyranoside
DEAE	: Diethylaminoethyl
DNA	: Deoxyribonucleic acid
cDNA	: complementary Deoxyribonucleic acid
UBD	: Ubiquitin binding domain
GST	: glutathione S-transferase-tag
TEV	: Tobacco Etch Virus
YUH1	: Yeast ubiquitin hydrase1
NMR	: Nuclear magnetic resonance
DUB	: Deubiquitinating enzyme
UIM	: ubiquitin interacting motif
FRET	: Förster resonance energy transfer
SDS-PAGE	: Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis
H/D	: Hydrogen/deuterated
HSQC	: Heteronuclear Single Quantum Correlation
Ni-NTA	: Nickel-Nitrilotriacetic Acid
OTUB1	: Otubain 1
PDI	: Protein Disulfide Isomerase

Thesis contents:

Chapter 1: General introduction

1. Introduction

1.1 Multidomain proteins

Sophisticated protein functions are, in many cases, mediated through the cooperative interplay between two or more domains [1]. Proteins with a modular architecture of multiple domains connected by linkers often exhibit diversity in relative positions of the individual domains organized through weak and even transient inter-domain interactions [1, 2]. Thus, the plasticity in dynamical spatial arrangements of these domains is a key feature of multidomain proteins endowed with sophisticated molecular mechanisms typified by allosteric regulation [3]. Moreover, it is suggested that motion, orientation, and interaction of these domains could be regulated and affected by various environmental factors in the cell, including pH, temperature, oxidative stress, and molecular crowding. Therefore, in order to extend our understanding of the working mechanisms of multidomain proteins in living system, the quantitative characterization of their conformational interchanges in solution is necessary. In this context, I have chosen Lys48-linked polyubiquitin chains as excellent models for structural study.

1.2 Ubiquitin chains

Ubiquitin (Ub) is a small protein composed of 76 amino acid residues [4]. Polymeric Ub chains, in which several Ub proteins are connected through specific isopeptide bonds, are known to play regulatory roles in various cellular processes, including cell cycle progression, DNA repair, transcriptional regulation, and apoptosis [4-6]. These regulatory functions of polymeric Ub chain are expressed by its modification of target proteins through the formation of isopeptide linkages at the C terminus. The C-terminal group of Ub can also be linked to another Ub (termed the distal and proximal moieties, respectively) through all seven lysine residues at positions 6, 11, 27, 29, 33, 48, and 63 as well as the N terminal amino group [4].

Ubiquitylation requires the sequential action of three enzymes; ubiquitin-activating enzyme (E1s), ubiquitin-conjugating enzymes (E2s), and ubiquitin ligase enzymes (E3s) [7-9]. First, an E1 uses ATP to link the C-terminal Gly of Ub through a thioester bond to a Cys in the E1 active site. The activated Ub intermediate is next transferred to the catalytic Cys of an E2. Ub is finally linked by its C-terminus to ε-amine of Lys residues of a substrate protein through an isopeptide bond. Target specificity and independent regulation of different linkages in ubiquitination are determined by the E2 and E3 enzymes (Figure 1.2.1).

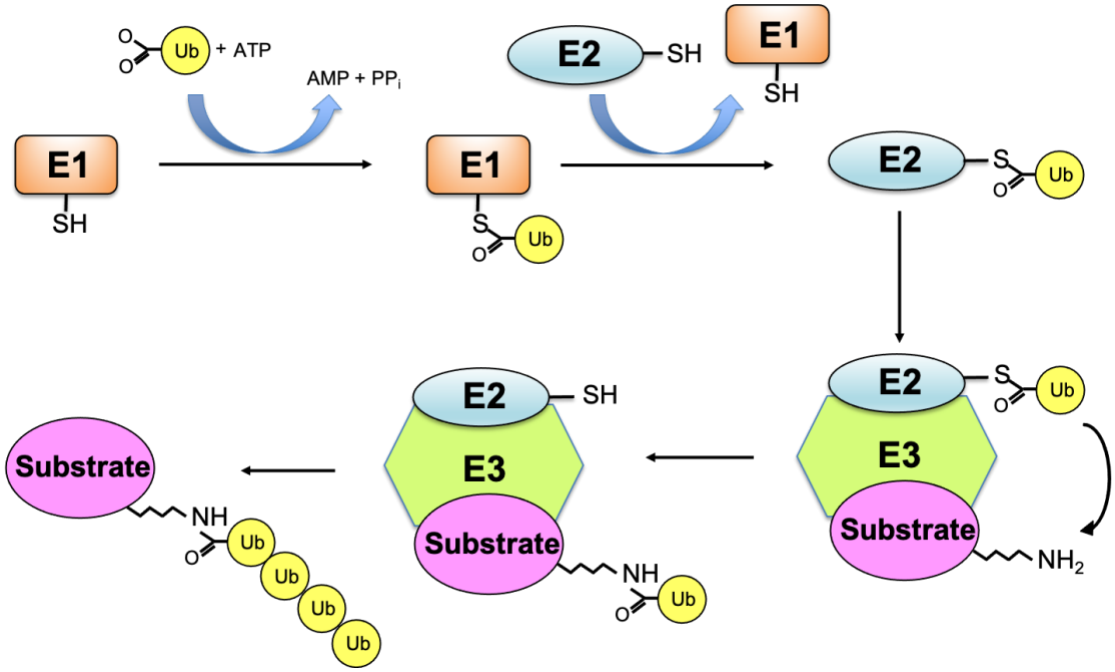


Figure 1.2.1 Ubiquitination system

The Ub chains conjugated by different linkages carry distinct biological information in the form of a “Ub code” that is read out by specific Ub-interacting proteins [4, 10]. These proteins generally recognize the hydrophobic surfaces displayed on the Ub chains. The Lys48-linked Ub chain serves as a tag for protein degradation by the 26S proteasome [11] (Fig. 1.2.2)., whereas the Lys63-linked Ub chain is involved in nonproteolytic functions, such as signaling in DNA repair and transcriptional regulation [12, 13].

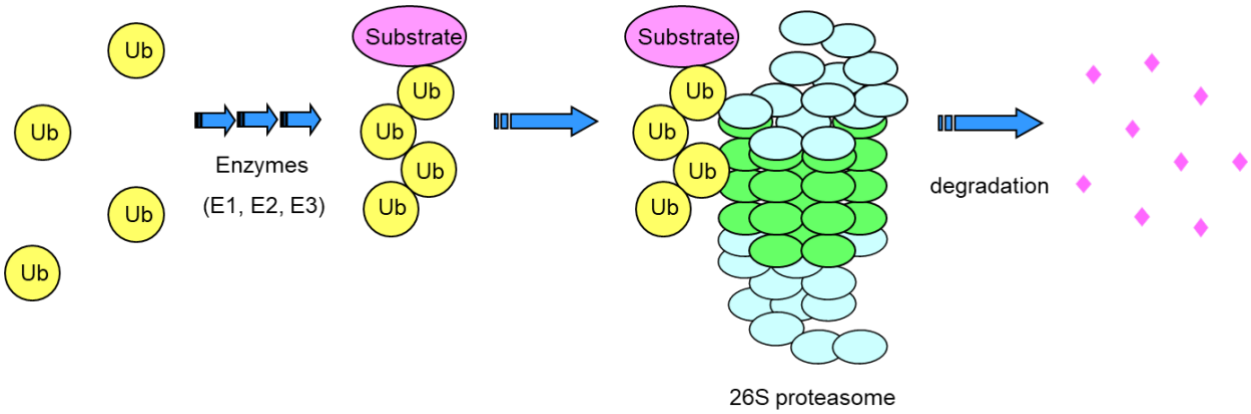


Figure 1.2.2 Ubiquitin-proteasome system

The proteins that recognize the specific Ub chains interact with a hydrophobic surface, which includes Leu8, Ile44, and Val70 of Lys48-linked Ub [10, 14] (Figure 1.2.3). The Ile44 patch is recognized by the proteasome and most ubiquitin-binding domains (UBDs). Another hydrophobic surface is centered on Ile36 and involves Leu71 and Leu73 of the ubiquitin tail (Fig. 1.2.3). The Ile36 patch can mediate interactions between ubiquitin molecules in chains, and it is recognized by UBDs and deubiquitinating enzymes (DUBs).

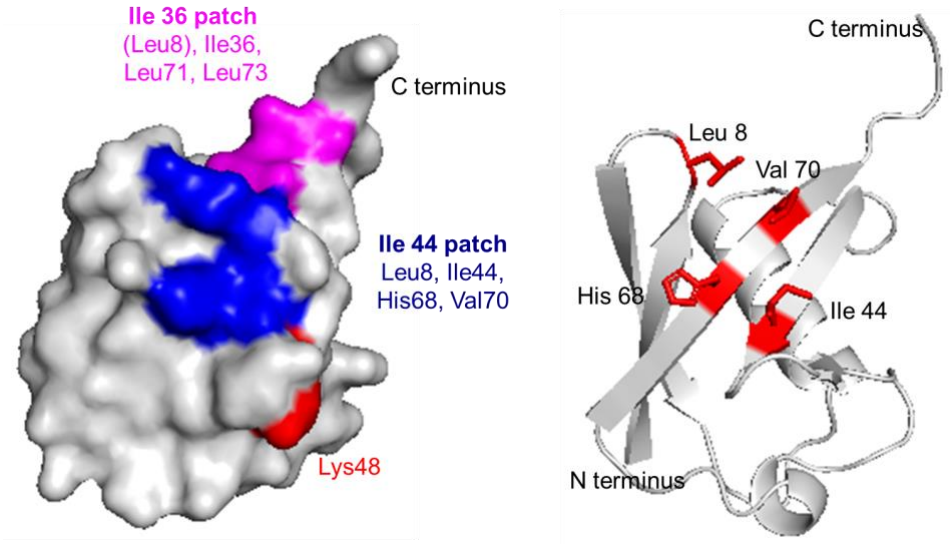


Figure 1.2.3 Structural features of Ub. The Ub surface is shown with Ile44 (blue) and Ile36 (magenta) highlighted (left). Structure of Ub, indicating the C-terminal tail and residues of the Ile44 hydrophobic surface (right). PDB code: 1UBQ.

1.3 Structural studies of polyUb chains

Since the first Lys48-linked diUb structure appeared in 1992 [15], several crystallographic and NMR structures of polymeric Ub chains have been solved, revealing that different linkages result in distinct conformations of the ubiquitin chains [10]. These conformational varieties of Ub chains mediated by specific linkage types are thought to create distinct molecular recognition mechanisms with different signals. The Lys63-linked diUb and tetraUb chains exhibit extended conformations in which the hydrophobic surfaces are exposed to a solvent, as revealed by crystallographic and NMR studies [16-18]. The crystal structure of Lys29-linked diUb also adopts an extend conformation, distinct from the conformation observed in Lys63-linked Ub chains [19].

In addition, the overall conformation of the Met1-linked chain is virtually equivalent to the Lys63-linked chain and consists of structures in which individual Ub units are rotationally unrestrained and highly flexible [17]. In contrast, the crystal structures of Lys48-linked Ub chains often exhibit closed conformations, in which the hydrophobic patches are shielded due to Ub-Ub interactions in the chains [15, 20]. In addition, their dynamic domain rearrangements have been observed in solution using NMR spectroscopy [21-26] and molecular simulation [27-29]. For example, our previous NMR study enabled the quantitative characterization of the conformational interchange of the native form of Lys48-linked diUb (n-diUb) between the open and closed conformations, based on conventional chemical shift data [25]. Furthermore, Lys48-linked tetraUb chains form a pseudo-tetragonal structure in which pairs of hydrophobic patches interact, suggesting that diUb is the minimal structural unit [22, 24]. As for diUb chains, tetraUb chains are thought to visit closed and open conformations in fast equilibrium and use this dynamical conformation state for partner recognition [26]. Similar to Lys48 linkages, Lys6- and Lys11-linked diUb chains adopt asymmetric compact conformations [30, 31]. Lys33-linked diUb also adopts a compact structure where the two monomeric subunits interact extensively through two hydrophobic surfaces, while the crystal structure of Lys33-linked triUb reveals an open conformation [19, 32].

1.4 Scope of this study

Previous NMR studies of structural dynamics of polymeric Ub chains in aqueous solutions have focused exclusively on diUb, including our characterization of the conformational interchange of the native form of Lys48-linked diUb between the open and closed conformations, based on conventional chemical shift data [25]. I thus extended the previous work by characterizing conformational interconversions of the native forms of Lys48-linked triUb and tetraUb chains in solution in my Ph.D. study.

These conformational revelations could not only provide insights into the relationship between domain motions and molecular functions mediated by Ub chains but also offer new strategies for probing the conformational dynamics of multidomain proteins. I also attempted to expand the structural features of Ub chains to unique design frameworks for creating allosterically controllable multidomain proteins.

Chapter 2: NMR characterization of conformational interconversions of Lys48-linked ubiquitin chains

2.1 Introduction

The Ub chains conjugated by different linkages carry distinct biological information in the form of a "Ub code" that is read out by specific Ub-interacting proteins [4, 10]. The Lys48-linked Ub chain serves as a tag for protein degradation by the 26S proteasome and interacts with the related proteasomal proteins through a hydrophobic surface [11]. According to the crystal structures, Lys48-linked Ub chains often exhibit closed conformations, in which the hydrophobic patches are shielded due to Ub-Ub interactions in chains [15, 20]. On the other hand, the previous NMR study in our group enabled the characterization of the conformational interchange of the native form of Lys48-linked diUb (n-diUb) between the open and closed conformations, based on conventional chemical shift data [25].

In my research, I have thus extended the previous work by characterizing conformational interconversions of the native forms of Lys48-linked triUb (n-triUb) and tetraUb (n-tetraUb) chains in solution. Using enzymatically synthesized Ub chains in cyclic and native, non-cyclic forms, conformational equilibria of Ub chains in terms of the exposure of the hydrophobic surfaces are quantitatively delineated based on the NMR spectroscopic data.

2.2 Materials and methods

2.2.1 Expression and purification of human Ub and derivative

Human Ub protein was expressed from pGEX6p-1 plasmid in BL21(DE3) CodonPlus cells and the cells were grown in M9 minimal media containing [^{15}N]NH_4Cl (1 g/l) to produce the isotopically labeled protein. After sonication, the supernatant was purified using DEAE column in buffer 50 mM Tris-HCl (pH 8.0). The pH of the collected flow-through was adjusted to 4.5 by using 10% acetic acid. After centrifugation, the supernatant was purified by SP-TOYOPERAL column (TOSOH) in 20 mM acetate pH 4.5 buffer. Finally, the buffer was exchanged to 50 mM Tris-HCl (pH 8.0).

C-terminally hexahistidine-tagged Ub (Ub-His) protein was expressed from the pET15C plasmid in BL21(DE3) CodonPlus cells and purified using DEAE and Ni^{2+} - nitrilotriacetic acid (Ni^{2+}-NTA) agarose (Qiagen). The cells were grown in M9 minimal media containing [^{15}N]NH_4Cl (1 g/l) to produce the isotopically labeled protein.

K48S and K48C Ub mutant were generated using site-directed mutagenesis techniques. The mutants were expressed from pGEX6p-1 plasmid in *Escherichia coli* BL21(DE3) CodonPlus cells, which were grown in M9 minimal media containing [^{15}N]NH_4Cl (1 g/L) in order to produce the isotopically labeled protein. After sonication, the supernatant was purified using a DEAE column in a buffer of 50mMTris-HCl (pH8.0). The pH of the collected flow-through was adjusted to 4.5 by using 10% acetic acid. After centrifugation, the supernatant was purified in a 20 mM sodium acetate (pH 4.5) using a SP-TOYOPEARL column (TOSOH, Japan). Finally, the buffer was exchanged with 50 mMTris-HCl (pH 8.0).

2.2.2 Expression and purification of Ub-related enzymes

2.2.2.1 E1

Hi5 insect cells were cultured in Grace's medium at 27ºC. The cells were infected with baculovirus containing hexahistidine-tagged-E1 (His_6-E1) cDNA for 60 h. The cells were harvested and then suspended into buffer A (50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 10% glycerol, 1% triton-X100, 0.01 mM dithiothreitol (DTT), His-protease inhibitor cocktail) incubated 30 min on ice. After centrifugation, His_6–E1 was purified by a Ni^{2+}-NTA affinity column (GE Healthcare). I also express His-E1 using *Escherichia coli* system. The plasmid vector encoding E1 was constructed and cloned as a fusion protein with His_6-E1 using the pF1KB4151 vector subsequently transformed into *E. coli* strain BL21-CodonPlus (Stratagene). Transformed bacteria were grown at 37ºC in LB media containing 50 μg/ml of ampicillin.

Protein expression was induced by adding 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG) when the absorbance reached 0.8 at 600 nm and harvested the cells. Protein was purified using the same protocol with His-E1 from insect cells.

2.2.2.2 Cdc-34

The plasmid vector encoding Cdc-34 was constructed and cloned as a fusion protein with glutathione S-transferase-tagged Cdc-34 (GST-Cdc-34) using the pCold vector. Subsequently, the vector was transformed into *Escherichia coli* strain BL21-CodonPlus (Stratagene). The transformed bacteria were grown at 37^{o}C in LB media containing 50 μg/mL ampicillin, until the OD = 0.65 at 600 nm, and then the cultures were rapidly cooled in an ice bath for 10–15 min. The GST-Cdc-34 protein was induced with 0.5 mM IPTG and the cultures were incubated at 15^{o}C overnight. The cell pellets were dissolved in 20 mM Tris-HCl (pH 8.0), 150 mM NaCl, and were disrupted by sonication.

GST-Cdc-34 was purified by a glutathione affinity column (GE Healthcare, Chicago, IL, USA). The Cdc-34 protein was enzymatically cleaved from GST-Cdc-34 by incubation with TEV protease. The cleaved Cdc-34 was purified by glutathione column, again, in order to remove GST-tag, and was then concentrated up to 2 mg/mL in the presence of 0.5 mM dithiothreitol (DTT), and stored at -80^{o}C.

2.2.2.3 E2-25K

The plasmid vector encoding E2-25K was constructed and cloned as a fusion protein with His_6-Ub using the pET28a(+) vector (Novagene), and subsequently transformed into *Escherichia coli* strain BL21-CodonPlus (Stratagene). Transformed bacteria were grown at 37^{o}C in LB media containing 50 μg/ml of ampicillin. Protein expression was induced by adding 0.5 mM IPTG when the absorbance reached 0.8 at 600 nm.

After 3 h, cells were harvested and then suspended into buffer A (50 mM Tris-HCl (pH 8.0), 100 mM NaCl, 10% glycerol, 1% triton-X100, 1 mM DTT, protease inhibitor cocktail) incubated 1 h on ice. After centrifugation, His_6–E2-25K was purified by a Ni^{2+} NTA affinity column (GE Healthcare).

2.2.2.4 YUH-1

Yeast ubiquitin hydrase 1 (YUH1) cDNA was inserted into the pGEX vector to encode GST-YUH1. The GST-YUH1 proteins were expressed in *Escherichia coli* BL21(DE3) CodonPlus and purified using a glutathione-Sepharose column (GE Healthcare) in a buffer (50 mM Tris-HCl (pH 8.0) and 0.1 M NaCl).

2.2.3 Enzymatic synthesis of cyclic Lys48-linked polyUb chains

The cyclic Lys48-linked polyUb chains were prepared by *in vitro* enzymatic reaction using E2-25K. [^{15}N]Ub in 50 mM Tris-HCl (pH 8.0) was incubated at 37^{o}C for 16 h in the presence of 0.6 μM E1, 20 μM E2–25K, 1 mM DTT, 5 mM $MgCl_2$, 10 mM ATP, 0.6 units/ml creatine phosphokinase, and 10 mM creatine phosphate. After the reaction, cyclic Ub chains were separated from longer Ub chains using Mono S (GE Healthcare) cation exchange chromatography (Figure 2.2.1). Protein purity was verified by SDS-PAGE.

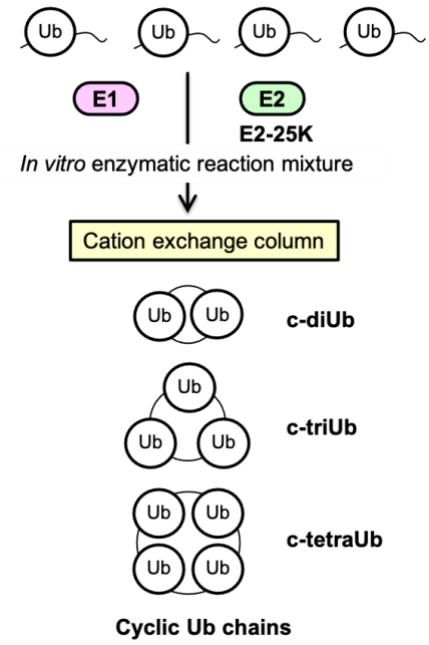


Figure 2.2.1 Schematic diagram of the preparation of cyclic Lys48-linked Ub chains

2.2.4 Enzymatic synthesis of native Lys48-linked polyUb chains

The native Lys48-linked polyUb chains were prepared by *in vitro* enzymatic reaction using Cdc-34. [^{15}N]Ub-His and [^{15}N]Ub were mixed at a molar ratio of 2:1 in 50 mM Tris-HCl (pH 8.0) and incubated at 37^{o}C for 16 h in the presence of 0.6 μM E1, 20 μM Cdc-34, 1 mM DTT, 5 mM $MgCl_2$, 10 mM ATP, 0.6 units/ml creatine phosphokinase, and 10 mM creatine phosphate. After the reaction, native Ub chains with the hexahistidine tag were separated from the reaction mixture including cyclic Ub chains by Mono S (GE Healthcare) cation exchange chromatography. Finally, native di, tri, and tetra Lys48-linked Ub chains with the hexahistidine tag were separated using Superdex 75 column (GE Healthcare) in buffer 50 mM Tris-HCl pH 8.0 (Figure 2.2.2). Protein purity was verified by SDS-PAGE. K48S and K48C mutant polyUb chains were prepared using enzymatic reaction with [^{15}N]Ub-His and [^{15}N]Ub mutant at a molar ratio of 2:1. After the reaction, using the same protocol with native Lys48-linked polyUb chains.

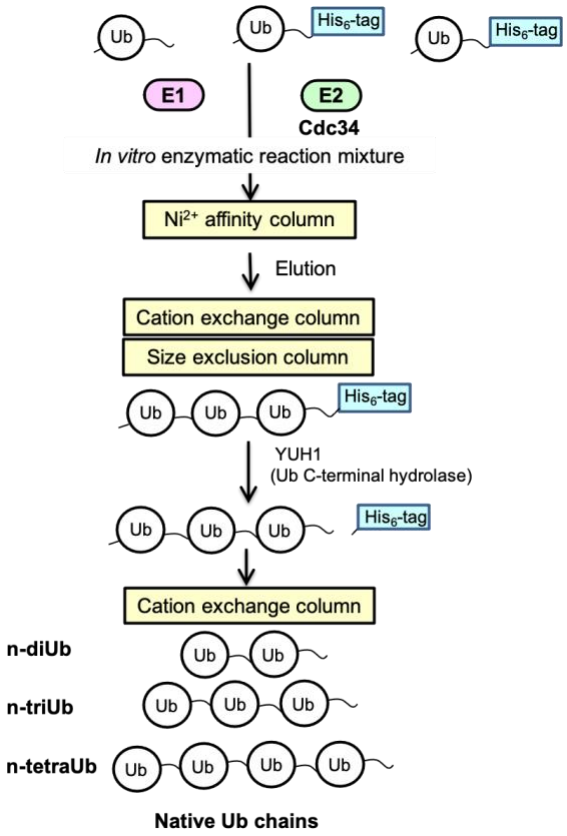


Figure 2.2.2 Schematic diagram of the preparation of native Lys48-linked Ub chains

2.2.5 Preparation of K48C-triUb with carboxymethyl

K48C-triUb 1 mg/ml were incubated with 10 mM DTT in 100 mM Tris-HCl (pH 8.5) and incubated at 25^{o}C for 4 h. After that, adding 25 mM iodoacetic acid and incubate for 30 min in the dark. After the reaction, 20 mM DTT was added for quenching the iodoacetic acid and K48C-triUb with carboxymethyl were purified by Mono S (GE Healthcare) cation exchange chromatography.

2.2.6 NMR measurements

All of the NMR samples were prepared in a 10mM sodium phosphate buffer at pH 7.0 with 95% H_2O/5% D_2O (*v/v*). NMR spectra were recorded at 42°C using a Bruker AVANCE-III HD 500 and AVANCE 800US spectrometers equipped with a 5-mm cryogenic triple-resonance probe. For temperature dependency experiment, NMR spectra were recorded at 42°C, 37°C, 30°C, 25°C, 20°C, 15°C and 10°C.The data were processed using NMR Topspin and analyzed using Sparky [33].

2.2.7 H/D exchange measurement

The buffer of uniformly ^{15}N-labeled n-triUb was replaced with 10mM sodium phosphate buffer at pH 7.0 with 99.9% (*v/v*) D_2O by centrifugation using a spin desalting column (Zeba spin desalting column, Thermo Fisher scientific, Waltham, MA). NMR spectra were immediately recorded at 42°C using a Bruker AVANCE 800US spectrometer. The data were processed using NMR Topspin and analyzed using Sparky [33].

2.2.8 Crystallization, X-ray data collection, and structure determination

For crystallization, purified c-triUb was dissolved in a protein concentration of 8.0 mg/mL in 20 mM Tris-HCl (pH 7.5) and 150 mM NaCl. Protein crystals were obtained in a buffer containing 20% PEG3350 and 200 mM zinc acetate with incubation at 20°C, and cryoprotected with a crystallization buffer supplemented with 20% glycerol. The diffraction data were integrated and scaled using HKL2000 [34]. The crystals of c-triUb belonged to space group C2 and were diffracted up to a resolution of 1.33 Å. The crystal structure of c-triUb was solved by molecular replacement method using the program MOLREP [35] with a human ubiquitin (Protein Data Bank code 1UBQ) as a search model.

Model fitting to the electron density maps and the subsequent refinement were conducted using COOT [36] and REFMAC5 [37], respectively. The stereochemical quality of the final model was validated using MolProbity [38]. The data collection and refinement statistics of c-triUb are summarized in Table 1. The molecular graphics were prepared using PyMOL (Schrödinger, New York, NY, USA).

Table 1. Data collection and refinements statistics for the crystal structure of c-triUb.

Crystallographic data	
Space group	*C*2
Unit cell *a*/*b*/*c* (Å)	88.4/54.8/61.0
β (°)	122.0
Data processing statistics	
Beam line	Photon Factory AR-NE3A
Wavelength (Å)	1.0000
Resolution (Å)	50–1.33 (1.35–1.33)
Total/unique reflections	285,708/56,837
Completeness (%)	99.9 (100.0)
R_{merge} (%)	7.5 (78.4)
I/σ (I)	25.2 (1.5)
Refinement statistics	
Resolution (Å)	20.0–1.33
R_{work}/R_{free} (%)	13.5/17.4
RMS deviations from ideal	
Bond lengths (Å)	0.014
Bond angles (°)	1.74
Ramachandran plot (%)	
Favored	100
Allowed	0
Outliers	0

2.3 Results and discussion

2.3.1. Spectral comparison among cyclic forms of Lys48-linked diUb, triUb, and tetraUb

I first compared the ^{1}H-^{15}N HSQC spectra of the cyclic forms of Lys48-linked diUb (c-diUb), triUb (c-triUb), and tetraUb (c-tetraUb) (Figure 2.3.1a–c and Figure 2.3.2a). In the cyclic forms, the C-terminus of one Ub unit is conjugated to the Lys48 of its most distal Ub unit. Each of these cyclic forms gave a single set of chemical shifts, indicating their symmetrical structures.

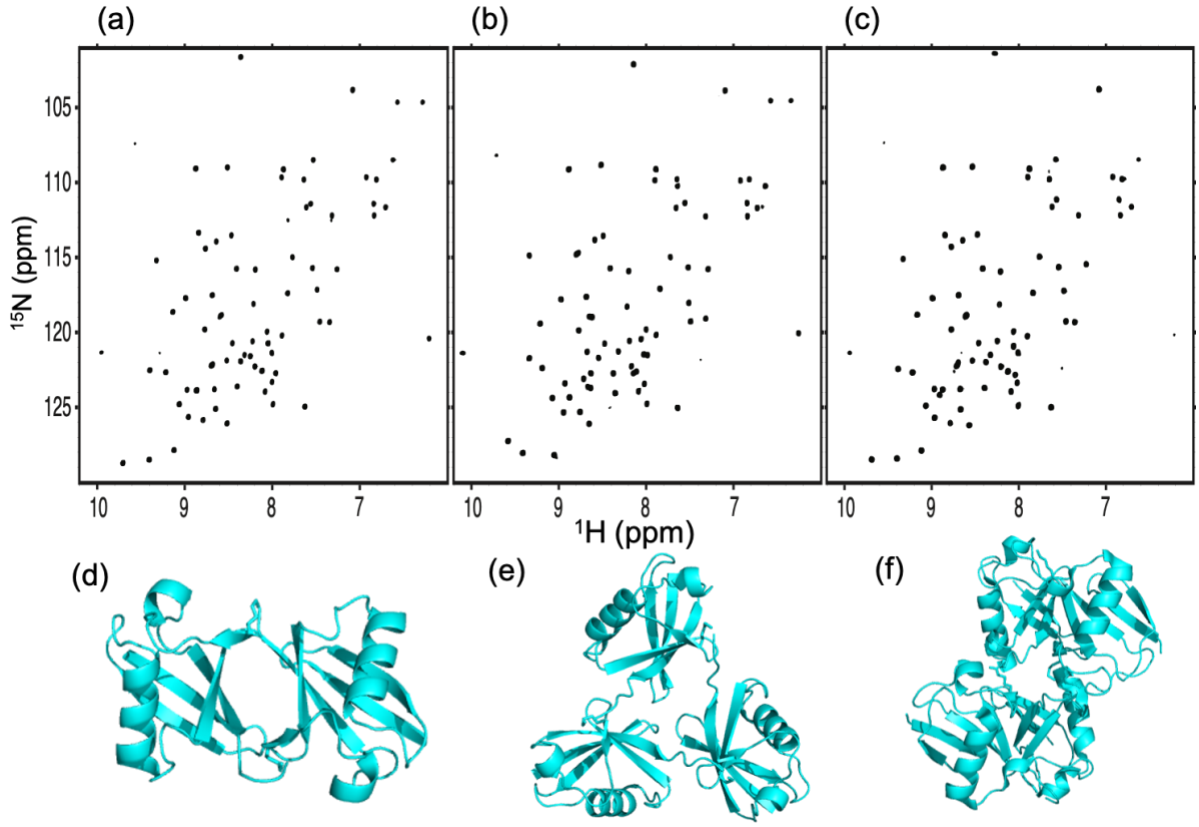


Figure 2.3.1 ^{1}H-^{15}N HSQC spectra of uniformly ^{15}N-labeled (a) c-diUb, (b) c-triUb, and (c) c-tetraUb. 3D structural models of (d) c-diUb [20], (e) c-triUb (solved in this study, PDB: 7CAP), and (f) c-tetraUb (PDB: 3ALB) [27].

The ^{1}H-^{15}N HSQC spectra of the c-tetraUb and c-diUb were well-superimposed, indicating their similarity with respect to the Ub–Ub interaction (Figure 2.3.2b). This is consistent with our previously reported structures of c-diUb and c-tetraUb, which share almost identical closed conformations [25, 39] (Figure 2.3.1d, f). In contrast, the spectrum of c-triUb was significantly different from that of c-diUb in terms of the chemical shifts of the peaks originating from the hydrophobic surface and the segment linking the Ub units (Figure 2.3.2a, c).

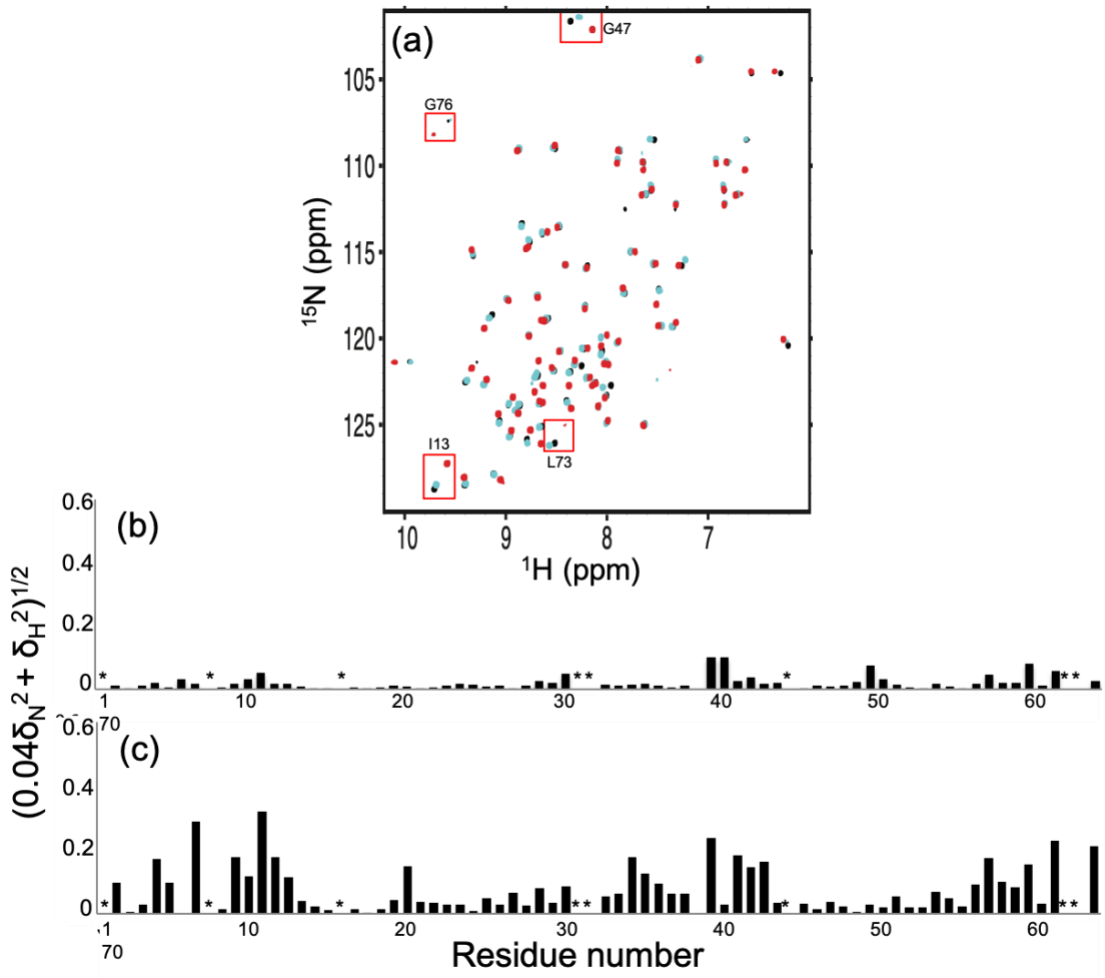


Figure 2.3.2 Spectral comparison of the cyclic Ub chains. (a) Superposition of the ^{1}H-^{15}N HSQC spectra of c-diUb (black), c-triUb (red), and c-tetraUb (cyan). Chemical shift differences (b) between c-tetraUb and c-diUb, and (c) between c-triUb and c-diUb. Data are shown according to the equation $(0.04\delta_N^2 + \delta_H^2)^{1/2}$, where δ_N and δ_H represent the difference in nitrogen and proton chemical shifts, respectively. The proline residues and the residues whose ^{1}H-^{15}N HSQC peak could not be used as a probe because of broadening are shown by asterisks.

We determined a crystal structure of c-triUb at a resolution of 1.33 Å (Figure 2.3.1e and Figure 2.3.3). As expected from the NMR spectra, c-triUb adopts a three-fold symmetric structure in which the hydrophobic surfaces, including Leu8, Ile44, and Val70, are partially exposed, distinct from the closed conformation observed for c-diUb and c-tetraUb.

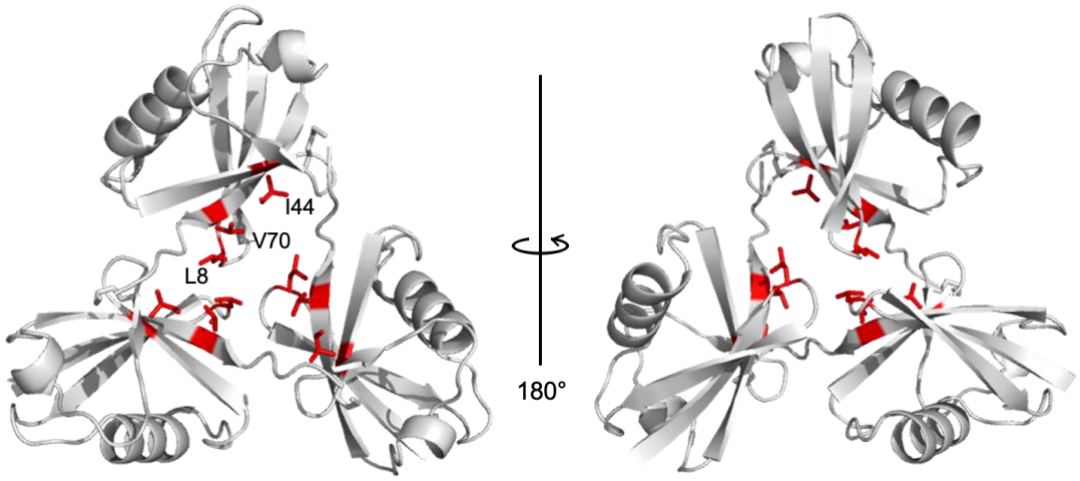


Figure 2.3.3 Crystal structure of c-triUb (solved in this study; PDB: 7CAP) highlighting Leu8, Ile44, and Val70 on the hydrophobic surface.

2.3.2. Spectral comparison among native forms of Lys48-linked diUb, triUb, and tetraUb

In the following, I compared the ^{1}H-^{15}N HSQC spectra of n-diUb, n-triUb, and n-tetraUb (Figure 2.3.4a–c). As our group reported previously [25], n-diUb has a single set of chemical shifts, except for the peaks originating from the Lys48-Gly76 linkage site, indicating that the two Ub units in n-diUb are structurally equivalent. In contrast, n-triUb and n-tetraUb exhibited multiple peaks for many residues, suggesting differences in the local environment among the Ub units. The amino acid residues showing different chemical shifts among different Ub units were located in the hydrophobic surfaces, as exemplified by Val70, in addition to those located in the Lys48-Gly76 linkage site (Figure 2.3.4d).

These observations suggest that the hydrophobic surface of the Ub units in the n-triUb and n-tetraUb experience different environments in time average.

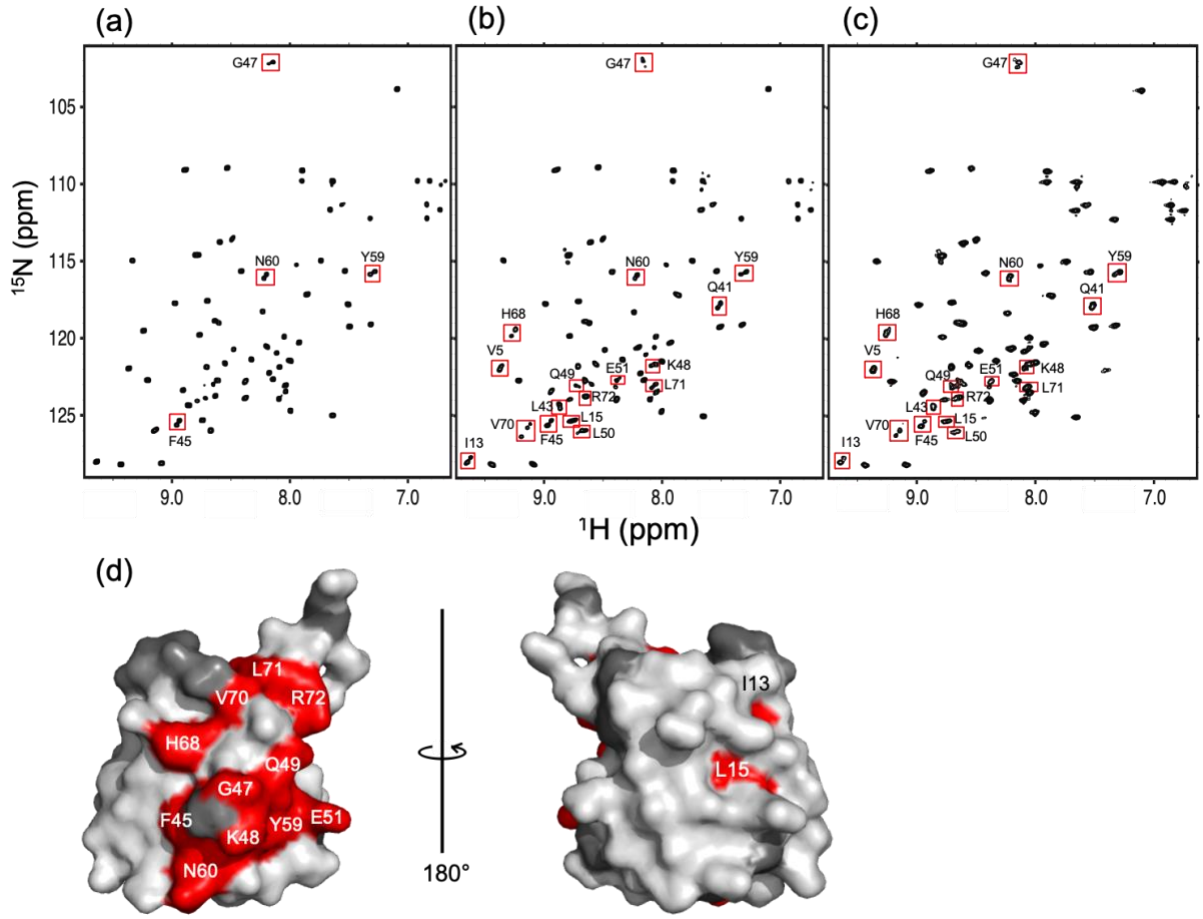


Figure 2.3.4 ^{1}H-^{15}N HSQC spectra of uniformly ^{15}N-labeled (a) n-diUb, (b) n-triUb, and (c) n-tetraUb. Multiple peaks (boxed) are displayed. (d) Mapping of the 3D structure of monomeric Ub (PDB: 1UBQ), with residues showing the multiple peaks. The proline residues and the residues whose ^{1}H-^{15}N HSQC peaks could not be observed as probes because of broadening and/or overlapping are shown in gray.

2.3.3. Conformational equilibrium of the native forms of Lys48-linked triUb and tetraUb

Our previous NMR study revealed that n-diUb undergoes a conformational transition in the fast exchange regime between the open and closed states, which is mimicked by monomeric Ub and c-diUb, respectively, in terms of the exposure of the hydrophobic surfaces to the solvent [25]. Namely, c-diUb exhibits a closed conformation, in which two hydrophobic surfaces are highly packed against each other, whereas the hydrophobic surface of the monomeric Ub is exposed to the solvent. Under this circumstance, each amino acid residue located on the hydrophobic surfaces of the Ub units gave an HSQC peak between the peaks originating from the corresponding residues in monomeric Ub and c-diUb in the same straight line, as exemplified by the Val70 of n-diUb (Figure 2.3.5a). This provided an opportunity to estimate the population of each conformation based on the internal division ratios by the n-diUb's peak of the chemical shift difference between monomeric Ub and the c-diUb, with the assumption that they represent fully open and closed states, respectively.

Intriguingly, this approach was applicable for estimating the conformer populations of n-triUb and n-tetraUb, because the multiple peaks originating from the hydrophobic surfaces of the Ub units in these Ub chains were also aligned in the same straight line between the corresponding peaks from c-diUb and monomeric Ub, exhibiting varying degrees of line broadening (Figure 2.3.5b–e).

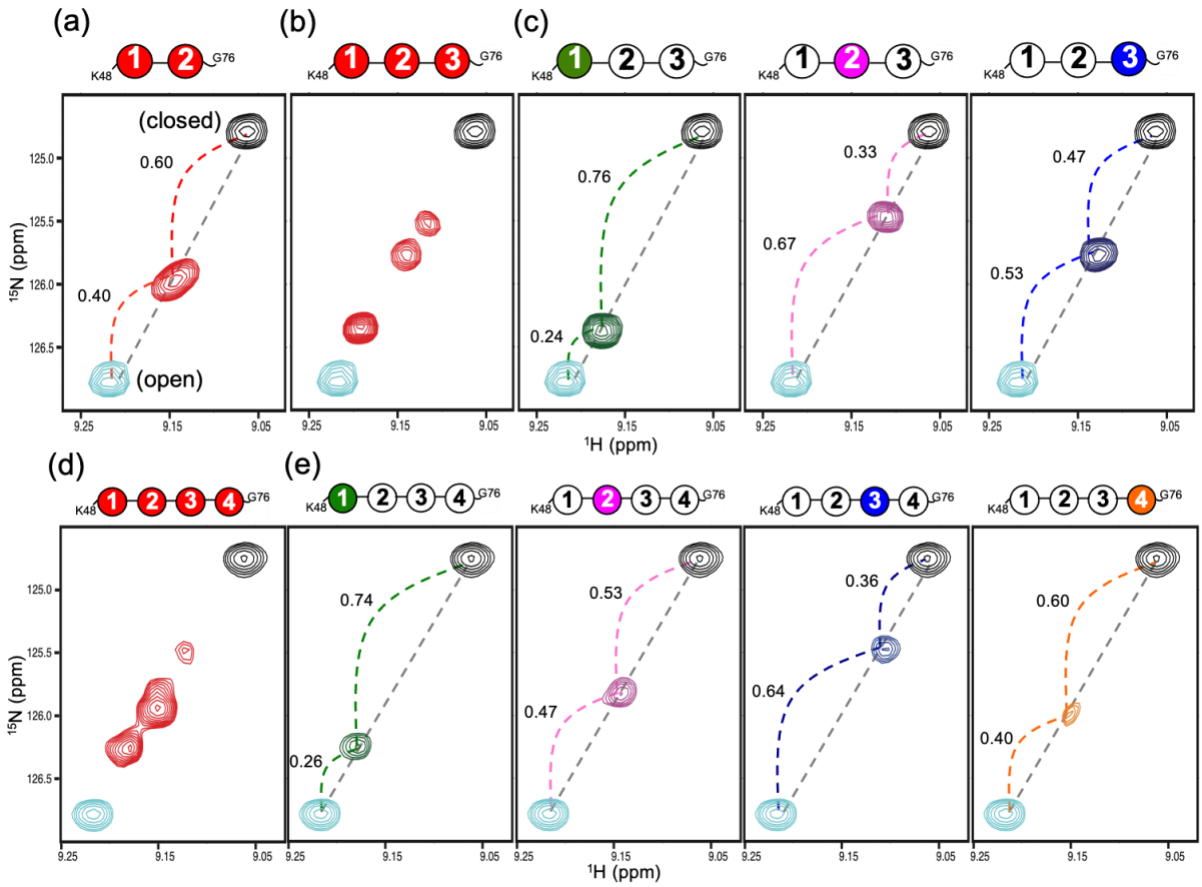


Figure 2.3.5 ^{1}H-^{15}N HSQC peaks originating from Val70 of (a) uniformly ^{15}N-labeled n-diUb (red), (b) uniformly^{15}N-labeled n-triUb (red), (c) unit-selectively ^{15}N-labeled n-triUb chains at the distal Ub1 (green), the middle Ub2 (magenta) and the proximal Ub3 (blue), (d) uniformly ^{15}N-labeled n-tetraUb (red), and (e) unit-selectively ^{15}N-labeled n-tetraUb at the distal Ub1 (green), the middle Ub2 (magenta), the middle Ub3 (blue), and the proximal Ub4 (orange). The peaks from the monomeric Ub (open form) and c-diUb (closed form) are plotted in cyan and black, respectively. The dividing ratios of the chemical shift differences of n-diUb, n-triUb, and n-tetraUb are indicated between monomeric Ub and c-diUb.

It should be noted that c-triUb was not suitable as a model of the closed state, because of its exposed hydrophobic surfaces (Figure 2.3.1e), which resulted in a significant chemical shift deviation from the straight line, as exemplified by the Val70 peak (Figure 2.3.6).

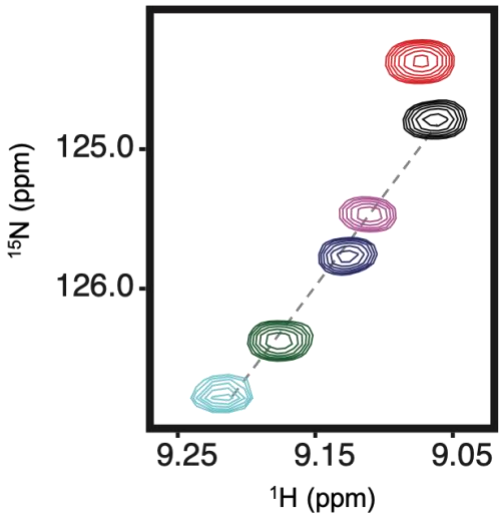


Figure 2.3.6 ^{1}H-^{15}N HSQC peaks originating from Val70 of monomeric Ub (cyan), c-diUb (black), c-triUb (red), and unit-selectively ^{15}N-labeled n-triUb chains at the distal Ub1 (green), the middle Ub2 (magenta) and the proximal Ub3 (blue).

These NMR data indicate that each Ub unit of n-triUb and n-tetraUb experienced a dynamic transition in the moderately fast exchange between the open and closed states on the relevant NMR timescale (approximately 100 Hz). Under this condition, the population of each conformer can be estimated on the basis of dividing the ratios of the chemical shift differences, using the monomeric Ub and c-diUb as references.

In n-triUb, the dividing ratio between the monomeric and cyclic forms in the Val70 peak indicated that the populations of open and closed conformers made up 76% and 24% of the total conformers in the distal unit (Ub1), 33% and 67% in the middle unit (Ub2), and 47% and 53% in the proximal unit (Ub3), respectively (Figure 2.3.5c and Figure 2.3.7).

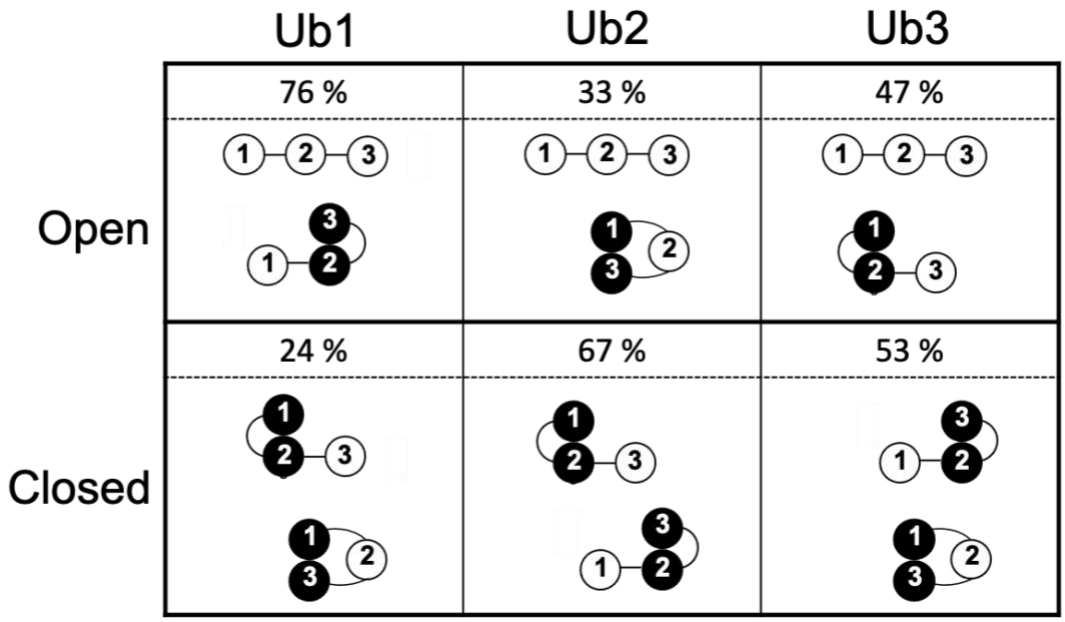


Figure 2.3.7 Conformer populations n-triUb estimated from the NMR spectral data. A cartoon model of the possible conformers in each state is shown. A pair of Ub units whose hydrophobic surfaces are shielded from each other are shown in black.

These data indicate that the hydrophobic surface of the distal unit (Ub1) is mostly exposed to the solvent, while that of the middle unit (Ub2) is mostly shielded. To consider the overall open–closed conformational equilibrium state of n-triUb, we formed a hypothetical model with four different conformational states (Figure 2.3.8), and estimated their populations based on the three different dividing ratios of the chemical shift differences of the Val70 peak, according to the following equations:

$$\text{Ub1 open : closed} = P_A + P_B : P_C + P_D = 0.76 : 0.24 \quad (1)$$

$$\text{Ub2 open : closed} = P_A + P_D : P_B + P_C = 0.33 : 0.67 \quad (2)$$

$$\text{Ub3 open : closed} = P_A + P_C : P_B + P_D = 0.47 : 0.53 \quad (3)$$

$$P_A + P_B + P_C + P_D = 1 \quad (4)$$

where P_X is the population of state x.

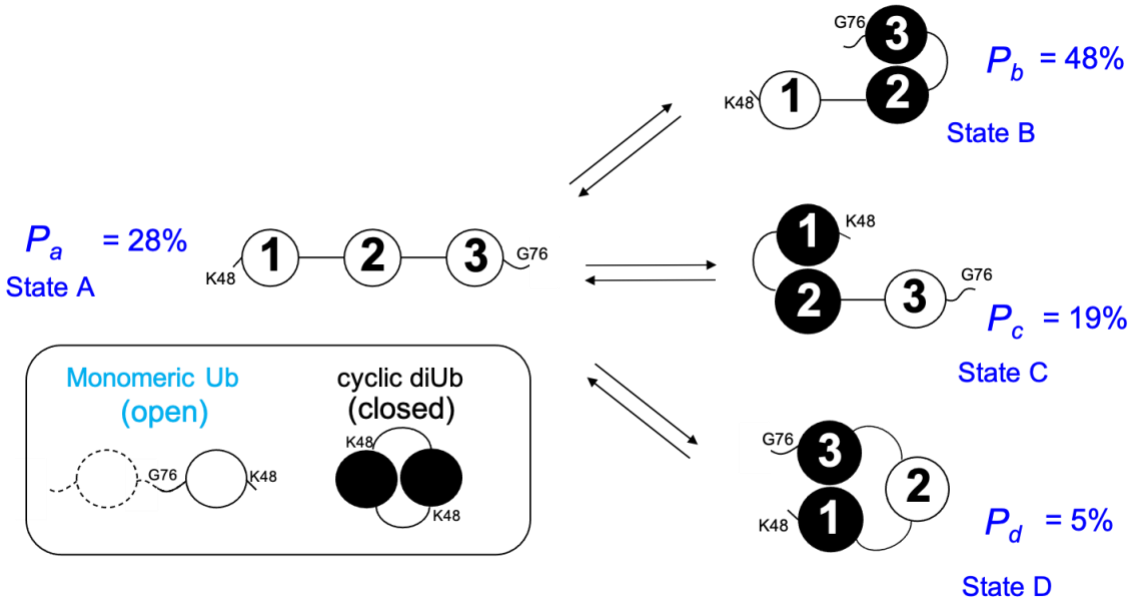


Figure 2.3.8 Cartoon model of the conformational equilibrium of n-triUb. The populations of states A, B, C, and D of n-triUb are denoted as P_A, P_B, P_C, and P_D, respectively.

According to the calculation results, the population of state A (fully open state) was 28% and those of state B, state C, and state D (three different closed states) were 48%, 19%, and 5%, respectively (Figure 2.3.8). This indicates that the "end-to-end" Ub interaction leaving the middle Ub2 unit open (state D) seldom occurs, presumably due to steric restriction, rendering Ub2 mostly involved in the interaction with either Ub1 or Ub3. It should be noted that the population of state B is significantly higher than that of state C, suggesting that Ub2 has a higher affinity for Ub3 than for Ub1.

In order to compare solvent accessibility among Ub1, Ub2, and Ub3 units of n-triUb, I performed a hydrogen/deuterated (H/D)-exchange NMR experiment using the V70 peak as a probe. In H_2O buffer, the relative intensity of the peak derived from Ub1 was the highest, followed by Ub3 and Ub2 (Figure 2.3.9a).

When the solution conditions of n-triUb were changed from H_2O to D_2O, the peak intensity changes with time were observed. As a result, after a certain time, the intensity of the V70 peak derived from Ub2 was the highest, while the intensities of the peaks derived from Ub1 and Ub3 were lower (Figure 2.3.9b).

These data indicate that the Val70 peak on Ub2 unit is the most protected from the solvent, supporting the fact that the hydrophobic surface of Ub2 unit is mostly shielded through the interaction with either Ub1 or Ub3.

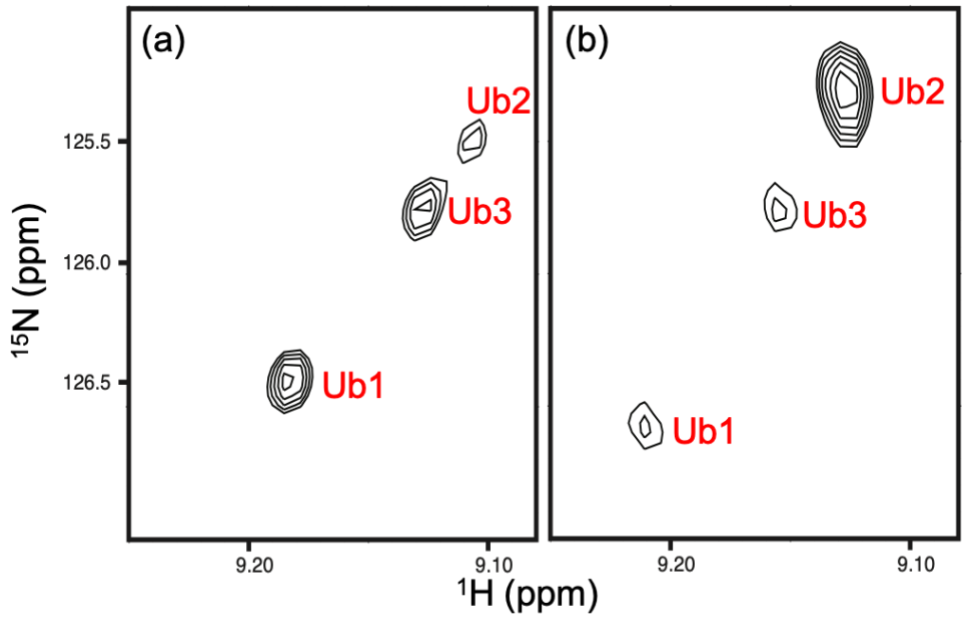


Figure 2.3.9 ^{1}H-^{15}N HSQC peaks originating from Val70 of n-triUb in 10mM sodium phosphate buffer at pH 7.0 with (a) 95%(v/v) H_2O/5% D_2O and (b) 6 h after replacement with 99.9%(v/v) D_2O.

In a similar way, I estimated the populations of the open and closed conformers of n-tetraUb based on dividing ratios with respect to the Val70 peak. The populations of the open state of Ub1, Ub2, Ub3, and Ub4 were estimated as 74%, 53%, 36%, and 60%, respectively (Figure 2.3.5e and Figure 2.3.10). Namely, the distal Ub1 unit has a more highly solvent-exposed hydrophobic surface than the remaining Ub units, as in the case of n-triUb.

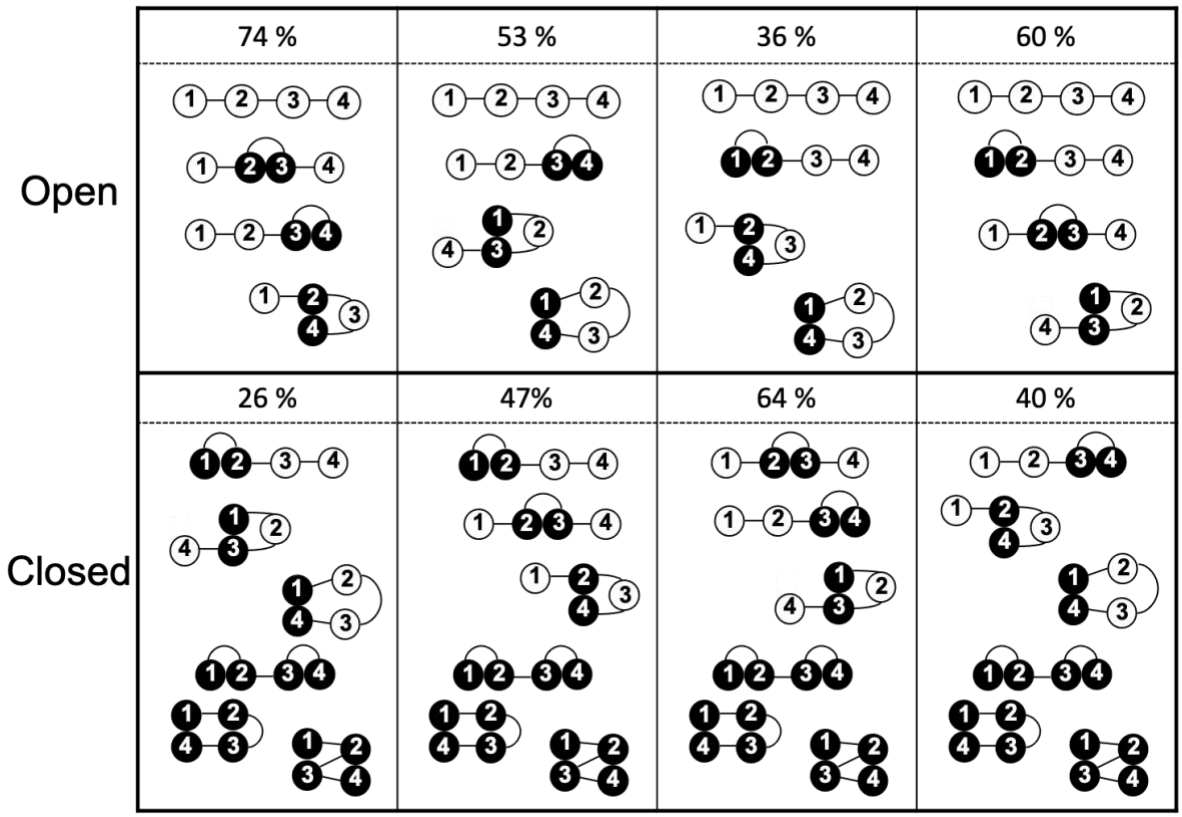
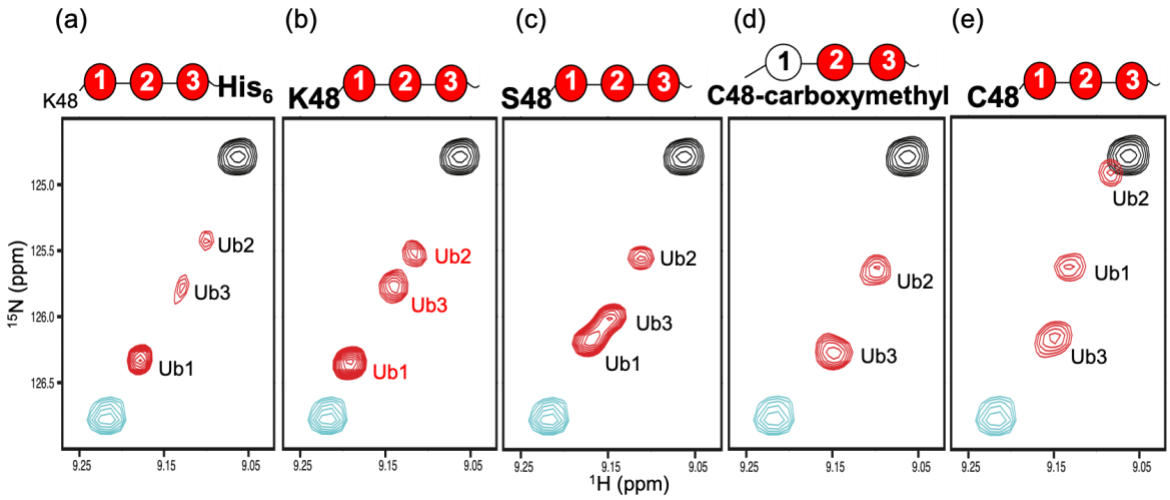


Figure 2.3.10 Conformer populations of n-tetraUb estimated from the NMR spectral data. A cartoon model of the possible conformers in each state is shown. A pair of Ub units whose hydrophobic surfaces are shielded from each other are shown in black.

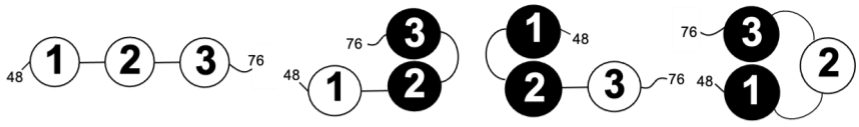
2.3.4 Inter-subunit interactions of Lys48-linked triUb

To explain the higher open-state propensity of the distal Ub units in n-triUb and n-tetraUb, I considered the possible end effects attributed to the distal and proximal end of the Ub chain, i.e., the Lys48 amino group of the distal Ub and the C-terminal carboxyl group of proximal Ub. The hydrophobic patch of Ub is surrounded by basic amino acid residues, i.e., Lys6, Arg42, Lys48, His68, and Arg72, which are likely to destabilize the closed conformation due to electrostatic repulsion. Indeed, previous study in our group demonstrated that the substitution of His68 with valine resulted in a significant increase in the population of the closed state in n-diUb [25]. The previously reported simulation data also showed that electrostatic interactions made a negative contribution to the formation of compact Lys48-linked diUb [27].

Firstly, I attempted to examine a possible C-terminal effect by using an n-triUb analog with a C-terminal extension with hexahistidine in the proximal Ub (n-triUb-His_6), which was also subjected to the NMR-based conformer population quantification. The multiple peaks from the Val70 of n-triUb-His_6 were distributed in the same straight line (Figure 2.3.11a). According to the calculated population, no significant difference was observed in each of the conformer populations between n-triUb and n-triUb-His_6 (Figure 2.3.11f).



(f)



	State A	State B	State C	State D
	P_A	P_B	P_C	P_D
wild-type n-triUb	28%	48%	19%	5%
n-triUb-His_6	27%	49%	21%	3%
K48S-triUb	33%	37%	26%	4%
K48C-carboxymethyl-triUb	35%	30%	33%	2%
K48C-triUb	7%	31%	62%	0%

Figure 2.3.11 ^{1}H-^{15}N HSQC peaks originating from Val70 of uniformly ^{15}N-labeled (a) n-triUb-His_6, (b) wild-type n-triUb, (c) K48S-triUb, (d) K48C-carboxymethyl-triUb, and (e) K48C-triUb. The peaks from monomeric Ub and c-diUb are shown in cyan and black, respectively. (f) Cartoon model of the conformational equilibrium of n-triUb. The populations of states A, B, C, and D of triUb are denoted as P_A, P_B, P_C, and P_D, respectively. The calculated P_A, P_B, P_C, and P_D values of n-triUb-His_6, wild-type n-triUb, K48S-triUb, K48C-carboxymethyl -triUb, and K48C-triUb are indicated.

Next, I examined the possible effect of the substitution of Lys48 with serine in the distal Ub of n-triUb (K48S-triUb) on its conformational states. The multiple peaks from the Val70 of the K48S-triUb exhibited clearly different chemical shifts but were still aligned in the same straight line (Figure 2.3.11c). The calculated populations of states A, B, C, and D were 33%, 37%, 26%, and 4%, respectively (Figure 2.3.11f). As compared to the wild-type n-triUb, the population of state C of K48S-triUb was increased by the mutation (from 19% to 26%), with a concomitant decrease of population B (from 48% to 37%), indicating that the positive charge of Lys48 of the distal Ub negatively contributes to its closed state formation. These data demonstrate that the positively charged Lys48 in the distal Ub affected the conformer population distribution of n-triUb, while the proximal C-terminal group had little impact on it.

In order to further investigate the effect of negative charge on the K48 position, K48 in the distal Ub of n-triUb was replaced with cysteine (K48C-triUb), and then C48 was carboxymethylated (K48C-carboxymethyl-triUb). As compared to the K48S-triUb, the population of state C of K48C-carboxymethyl-triUb was slightly increased (from 26% to 33%), with a concomitant decrease of population B (from 37% to 30%), confirming that the negative charge at Lys48 position of the distal Ub positively contributes to its closed state formation. Intriguingly, I should note that the calculated population of state C of K48C-triUb was dramatically increased (from 19% to 62%) as compared to the wild-type n-triUb. As cysteine is a neutral residue, such a large increase of the State C population cannot be explained solely by the charge effect at the K48 position. This result is probably due to the hydrophobic property of cysteine residue [40].

To further clarify the possible driving force behind the conformational interconversion among the Ub units, I examined the temperature dependence of the conformational equilibrium of n-triUb by conducting NMR measurements at various temperatures. The chemical shift of Val70 shifted with temperature but was still aligned in the same straight line between c-diUb and monomeric Ub (Figure 2.3.12).

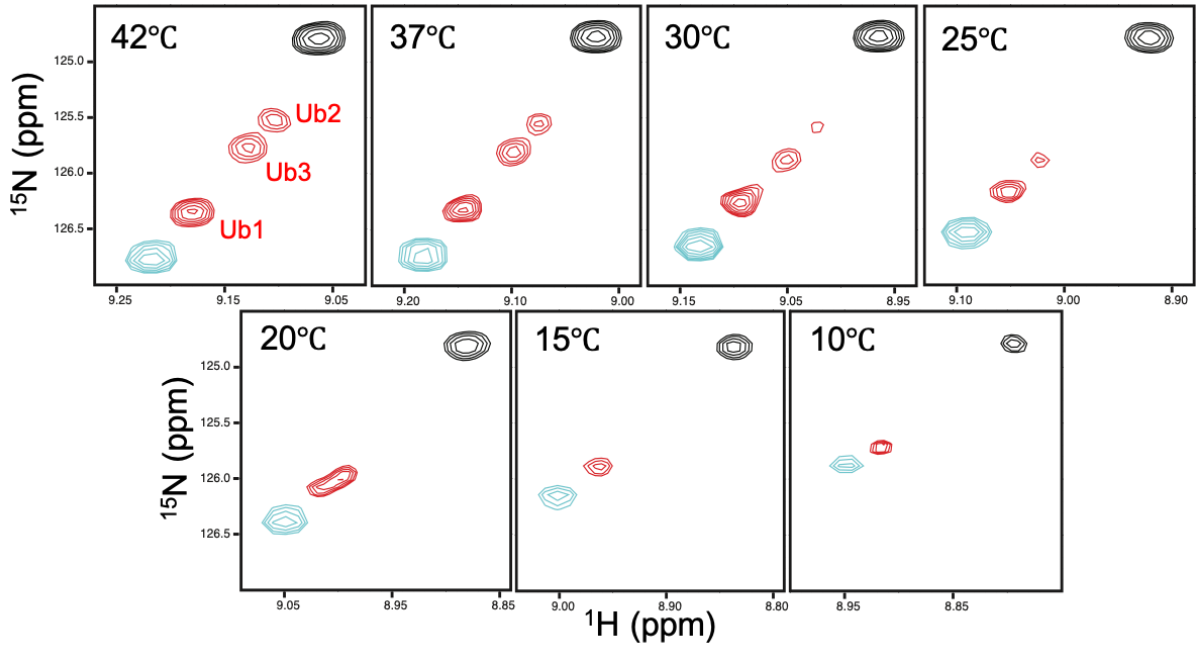


Figure 2.3.12 NMR characterization of conformational dynamics of Lys48-linked triUb in solution. ^{1}H-^{15}N HSQC peaks originating from Val70 of monomeric Ub (cyan), unit-selectively ^{15}N-labeled n-triUb chains (red), and cyclic Lys48-linked diUb (black) at the following different temperature conditions: 42°C, 37°C, 30°C, 25°C, 20°C, 15°C, and 10°C.

This provided a quantitative estimate of the temperature-dependent populational shift of each conformer of n-triUb. I have summarized the temperature dependence of the conformer populations of wild-type n-triUb, as well as K48S-triUb, K48C-triUb, and K48C-carboxymethyl-triUb (Figure 2.3.13).

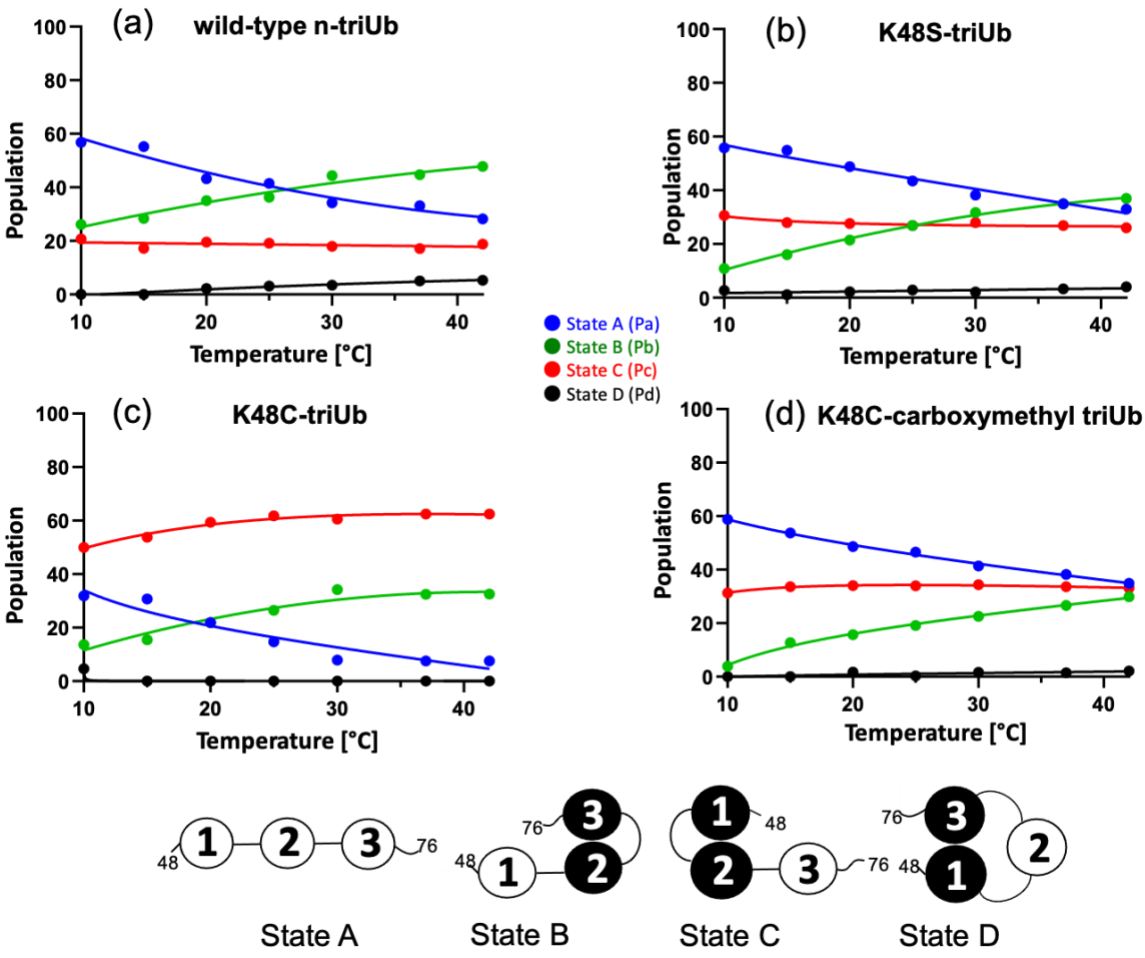


Figure 2.3.13 Temperature dependence of the open and closed conformational population of (a) wild-type n-triUb, (b) K48S-triUb, (c) K48C-triUb and (d) K48C-carboxymethyl triUb. Population of state A, B, C and D are shown in blue, green, red and black respectively. Population of each temperature estimated on the basis of the dividing ratio of chemical shift differences of Val70.

In all triUb derivatives, the population of State A and State B showed a temperature dependency, with the population of state B increasing and that of A concomitantly decreasing at higher temperatures. Since the hydrophobic interaction is generally enhanced at high temperatures, the formation of the closed structure between Ub2 and Ub3 in State B is mainly induced through the hydrophobic interaction between Ub2 and Ub3 units. On the other hand, the population of State C was found to be constant with temperature, although its initial population varied among the mutants. This result suggests that the formation of State C, i.e. the closed structure between Ub1 and Ub2, is mainly determined by the property of the K48 position in Ub1. By inspection of all these data, it is considered the proportion of multi-conformers corresponding to State A, State B, and State C is determined by the competition between the "Ub1-Ub2 interaction defined by the property of the K48 position of Ub1 unit" and the "Ub2-Ub3 interaction defined by the hydrophobic interaction".

These present data demonstrate that the amino acid substitutions at position 48 in the distal Ub of the Ub chain remotely affected the solvent exposure of the hydrophobic surfaces of the other Ub units through the competitive sharing of the hydrophobic surfaces among the Ub units. Namely, the modification of Ub1 unit can remotely modulate the solvent exposure of Ub3 unit. This allostery depends on the competitive sharing of Ub2 between Ub1 and Ub3 in the formation of closed states. Thus, the mutational effect at the most distal Ub is allosterically transmitted to the remaining Ub units in a chain-reaction manner.

Chapter 3: Exploration of Ub-based protein engineering

3.1 Introduction

In Chapter 2, I found that the amino acid substitutions at position 48 in the distal Ub of the Ub chains remotely affected the solvent exposure of the hydrophobic surfaces of the other Ub units through the competitive sharing of the hydrophobic surfaces among the Ub units. Thus, the mutational effect at the most distal Ub is allosterically transmitted to the remaining Ub units in a chain-reaction manner. These results suggested that the Lys48-linked Ub chains may offer unique design frameworks for Ub-based protein engineering including the creation of biosensing probes and allosterically controllable multidomain proteins. For proof of this concept, I attempted to design artificial multidomain proteins based on the Lys48-linked diUb.

Förster resonance energy transfer (FRET) can be used to study molecular events such as conformational changes, molecular interactions, and protein cleavages in an effective manner thus can provide a platform for designing sensor molecules to detect changes in the solution environments including pH, temperature, oxidative stress, and molecular crowding. I could construct environmentally responsive biosensing probes, in which FRET is enhanced in the closed state of Lys48-linked diUb.

Furthermore, I designed artificial hetero-multidomain proteins in which Ub is conjugated with another multidomain protein for controlling their domain rearrangements in environment-responsive manners. Protein disulfide isomerase (PDI) is a multidomain protein that consists of four tandem thioredoxin domains, *a, b, b'*, and *a'*, plus C-terminal extension. One unique property of PDI is that it undergoes conformational rearrangement of the *b'-a'* domains depending on the redox states of the *a′* active site [41, 42]. These two domains exhibit a closed conformation in the reduced form and are converted into an open conformation with the exposure of the hydrophobic surface upon oxidation of the *a′* domain. This conformational transition is supposed to be associated with the redox-dependent substrate binding of PDI.

In this study, I attempted to create an artificial hetero-multidomain protein by linking diUb and PDI$_{b'-a'}$. I expect that this would lead to the creation of a novel multidomain protein that exhibits the temperature/pH dependency of diUb and the sensitivity to redox states of PDI as well as regulates both properties in an allosteric manner.

3.2 Materials and methods

3.2.1 FRET biosensing probe

3.2.1.1 Preparation of diUb with fluorescent probes

Alexa647-maleimide (Cy5) and Cy3 Bis NTA-Ni were purchased from Thermo Fisher scientific (Waltham, USA) and AAT Bioquest (Sunnyvale, USA), respectively. K48C-diUb with a hexahistidine tag (K48C-diUb-His, 80 μM) in 50 mM Tris-HCl, pH 7.2, was prepared as described in 2.2.4 in chapter 2. Labeling reactions were carried out overnight at 4^{o}C a Cy5 dye: protein of 10:1 molar ratio in the presence of 10 mM TCEP. The Excess amounts of unincorporated dyes and unconjugated proteins were removed by a MonoS (GE Healthcare) cation exchange chromatography. Protein purity was verified by SDS-PAGE. The Cy3 NTA-Ni dye was incubated with Cy5-conjugated K48C-diUb-His at 1:1 molar ratio in 10mM sodium phosphate buffer at pH 7.0 (Cy5-diUb-Cy3).

3.2.1.2 Cy5-diUb-Cy3 with E2-25K binding

E2-25K was prepared as described in 2.2.2.3 in chapter 2. E2-25K solution was exchanged into 10mM sodium phosphate buffer at pH 7.0. The Cy5-diUb-Cy3 was incubated with E2-25K at 1:1, 1:2, and 1:5 molar ratio at 25^{o}C for 30 min.

3.2.1.3 Cy5-diUb-Cy3 with DUB

OTUB1, a K48-linkage specific DUB, was purchased from Life sensors. For cleavage

reaction, 100 µM of Cy5-K48C-diUb-His was incubated at a DUB: protein of 1:5 molar ratio at 37ºC for 1 h in the presence of 2 mM DTT. After the reaction, the sample was diluted to 0.8 µM for decreasing DTT concentration and then Cy3 NTA-Ni dye was added at a molar ratio of 1:1 in 10mM sodium phosphate buffer at pH 7.0.

3.2.1.4 Cy5-diUb-Cy3 with cytoplasmic extracts

The cytoplasmic extracts were collected from HEK293 cells. 0.8 µM Cy5-K48C-diUb-His was incubated with 0.25 mg/ml of cytoplasmic extracts at 25ºC, then Cy3 NTA-Ni dye was added at a molar ratio of 1:1.

3.2.1.5 Cy5-diUb (ethylamine linker)-Cy3

Synthesized ethylamine was provided by Dr. Momiyama and Dr. Ohtsuka. K48C Ub mutants with and without hexahistidine tag (K48C-Ub/ K48C-Ub-His) were prepared as described in 2.2.1 in chapter 2. Cy5-conjugated K48C-Ub was prepared as described in 3.2.1.1 in chapter 3. Firstly, K48C-Ub-His was incubated with ethylamine at 37ºC for 4 h for the creation of K48C-Ub-His conjugated with ethylamine as a linker. Cy5-K48C-Ub and monomeric K48C-Ub-His conjugated with ethylamine were mixed at molar ratio of 2:1 in 50 mM Tris-HCl (pH 8.0), and then subjected to *in vitro* enzymatic reaction using Cdc-34 as described in 2.2.4 in chapter 2 for creation of Cy5-diUb (ethylamine linker)-His. After the reaction, artificially-linked diUb, Cy5-diUb (ethylamine linker)-His, were separated from the reaction mixture by MonoS (GE Healthcare) cation exchange chromatography. Cy3 NTA-Ni dye was added at a molar ratio of 1:1 with Cy5-diUb (ethylamine linker)-His to create Cy5-diUb (ethylamine linker)-Cy3.

3.2.1.6 FRET measurement and analyses

Fluorescence emission spectra were recorded with Cy5-diUb-Cy3 (0.8 μM) in 10mM sodium phosphate buffer at pH 7.0 using a fluorescence spectrometer (Hitachi F2700) at 25ºC in the presence and absence of E2-25K. Emission spectra were recorded at 400-800 nm with excitation wavelengths of 550 nm. FRET peak was observed at 670 nm with excitation wavelengths of 550 nm. Percent of FRET efficiency was calculated using Percent of FRET efficiency = (Total FRET spectral - spectral bleed-through - cross-excitation) / emission of donor (Cy3). Monomeric Ub-His-Cy3 and monomeric K48C-Cy5 were used for measuring spectral bleed-through (Ex: 550 nm, Em: 570 nm) and for cross-excitation (Ex: 650 nm, Em: 670 nm), respectively. Fluorescence emission spectra of Cy5-diUb-Cy3 and Cy5-diUb (ethylamine linker)-Cy3 with cytoplasmic extracts or OTUB1 were recorded using the same protocol.

3.2.2 Artificial hetero-multidomain proteins

3.2.2.1 Protein expression and purification of linear Ub- PDI*b'-a'*-Ub and Ub- PDI*b'-a'*

Ub-PDI*b'-a'*-Ub and Ub- PDI*b'-a'* protein was constructed from pET21-a plasmid and subsequently transformed into *E.coli* strain BL21-CodonPlus (Stratagene). Transformed bacteria were grown at 37ºC in M9 minimal media containing 50 μg/ml of ampicillin with[^{15}N]NH_4Cl (1 g/l) to produce the isotopically labeled protein. Protein expression was induced by adding 0.5 mM IPTG when the absorbance reached 0.8 at 600 nm. After 3 h, cells were harvested and then suspended into 50 mM Tris-HCl (pH 8.0). After sonication, the supernatant was applied onto a Superdex 75 column in a buffer of 50 mM Tris-HCl (pH 8.0) for purification.

3.2.2.2 Protein expression and purification of PDI$_{b'-a'}$-Ub

PDI$_{b'-a'}$-Ub protein was expressed from pET28-a plasmid in BL21-CodonPlus (Stratagene) cells and the cells were grown in M9 minimal media containing 50 μg/ml of kanamycin with[^{15}N]NH_4Cl (1 g/l) to produce the isotopically labeled protein. After sonication, the supernatant was purified using Ni^{2+}-nitrilotriacetic acid (Ni^{2+}-NTA) agarose (Qiagen) in a buffer of 50 mM Tris-HCl (pH8.0). The N-terminus hexahistidine-tag of PDI$_{b'-a'}$-Ub protein was enzymatically cleaved by incubation with Thrombin. Then, PDI$_{b'-a'}$-Ub was applied onto a Superdex 75 column in a buffer of 50 mM Tris-HCl (pH 8.0).

3.2.2.3 Preparation of cyclic diUb-PDI$_{b'-a'}$

The cyclic diUb-PDI$_{b'-a'}$ protein was prepared by *in vitro* enzymatic reaction using E2-25K. [^{15}N] linear Ub-PDI$_{b'-a'}$-Ub in 50 mM Tris-HCl (pH 8.0) was incubated at 37^{o}C for 16 h in the presence of 0.6 μM E1, 20 μM E2–25K, 1 mM DTT, 5 mM $MgCl_2$, 10 mM ATP, 0.6 units/ml creatine phosphokinase, and 10 mM creatine phosphate. After the reaction, cyclic diUb-PDI$_{b'-a'}$ were separated using Superdex 75 column in a buffer of 50 mM Tris-HCl (pH 8.0). Protein purity was verified by SDS-PAGE.

3.2.2.4 Mass spectrometry

The sample solutions were buffer-exchanged into 150 mM ammonium acetate at pH 6.8 using centrifugal spin columns (Micro BioSpin-6 column, Bio-Rad, Hercules, California, USA) at 4°C. In order to prepare the denatured samples, formic acid was added to the buffer-exchanged samples at a volume fraction of 30%. The samples were loaded into gold-coated glass capillaries made in house (approximately 2-5 μl sample loaded per analysis) and analyzed by nanoflow electrospray ionization mass spectrometry. Spectra were recorded on a SYNAPT G2-*Si* HDMS mass spectrometer (Waters, Massachusetts, Milford, USA) in positive ionization

mode at 1.33 kV capillary voltage with 37°C source temperature, 150 V sampling cone, and source offset voltages, 0 V trap and transfer collision energies, and 2 ml/min trap gas flow. The spectra were calibrated using 2 mg/ml cesium iodide and analyzed with MassLynx software 4.1 (Waters).

3.2.2.5 NMR measurements

All of the NMR samples were prepared in 10 mM sodium phosphate buffer at pH 7.0 with 95% H_2O/5% D_2O (*v/v*). NMR spectra were recorded at 30°C using a Bruker AVANCE 800US spectrometers equipped with a 5-mm cryogenic triple-resonance probe with three-axis gradient coils. The data were processed using NMR Topspin and analyzed using Sparky.

3.3 Result and discussion

3.3.1 FRET biosensing probe

It has been found that the spatial arrangement of the Lys48 residue on the distal Ub unit and the C-terminus on the proximal Ub unit are close to each other in the closed form of n-diUb. I already confirmed that K48C-mutated diUb (K48C-diUb) was also under the open-closed equilibrium in solution. Therefore, the open-closed conformational changes of diUb can provide a platform for designing sensor molecules to detect changes in the solution environments. I used the K48C-diUb with a hexahistidine tag to introduce fluorescent dyes. A donor (Cy3 Bis NTA-Ni complex) was attached to the hexahistidine tag *via* a coordination bond, and an acceptor (Alexa Fluor 647 maleimide, or Cy5) was attached to the cysteine residue *via* a disulfide bond (Figure 3.3.1).

As a result of fluorescence measurement, FRET was observed in this K48C-diUb-based probe, Cy5-diUb-Cy3 (Figure 3.3.2a). Upon addition of E2-25K as a binding protein, it was expected that the open structure of diUb would increase due to the binding of E2-25K to the

hydrophobic surface of Ub. Indeed, the FRET peak showed intensity decrease upon addition of E2-25K to diUb, demonstrating a successful construction of FRET biosensor based on the conformational interconversion of diUb chain (Figure 3.3.2b).

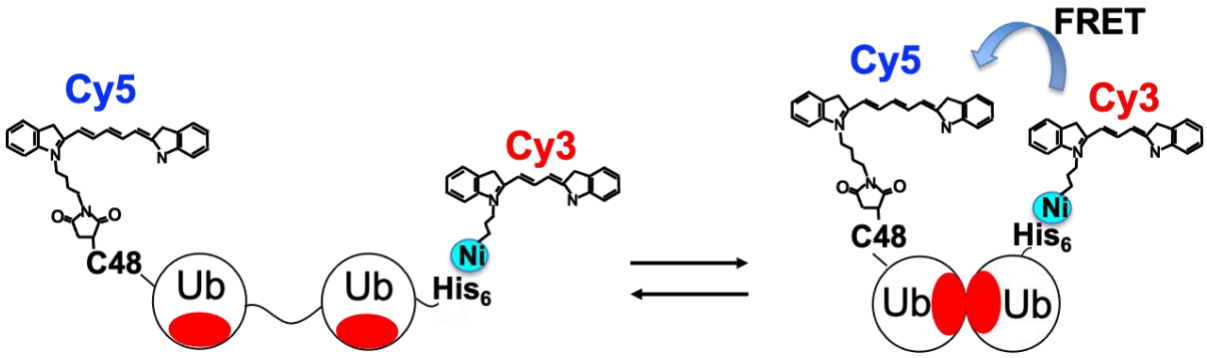


Figure 3.3.1 Strategy of FRET measurement of K48C-diUb-His conjugated with Cy3 and Cy5.

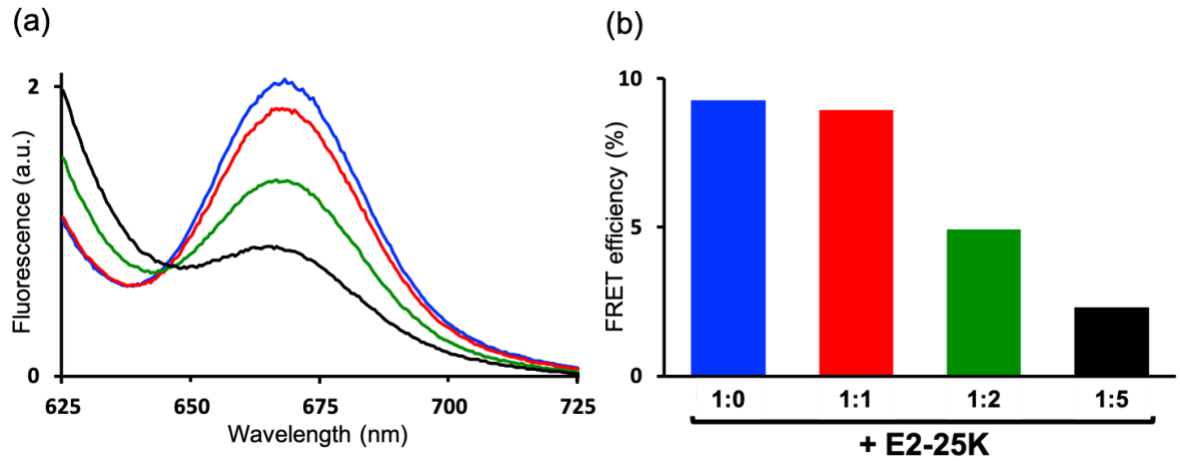


Figure 3.3.2 (a) The emission spectra corresponding to FRET of Cy5-diUb-Cy3 without E2-25K (blue), with E2-25K at a molar ration of 1:1 (red), 1:2 (green) and 1:5 (black). (b) Percent of FRET efficiency of Cy5-diUb-Cy3 calculated from the corresponding FRET intensities. The colors correspond to (a).

In order to consider the use of this diUb-based FRET biosensor probing cellular environments, I performed the FRET experiment using cytoplasmic extracts. The result showed that FRET efficiency was dramatically decreased in the cytoplasmic condition (Figure 3.3.3c). This decrease may suggest the existing of Ub binding proteins in cytoplasmic extracts. Moreover, there is another possibility that the isopeptide linker between Ub units was cleaved by deubiquitinating enzyme (DUB) in the cytosol. To access these possibilities, the diUb probe was incubated with DUB. The FRET efficiency was decreased upon addition of DUB (Figure 3.3.3c). These data suggested that the linker of Ub units was cleaved by DUB and the Cy5 and Cy3 were isolated from each other, resulting in a decrease in the FRET efficiency.

Therefore, further improvement is necessary, such as linking Ub units with an ethylamine that is expected to be non-cleavable by cytosolic DUBs in order to use this FRET biosensor in the cellular environments. The FRET results clearly indicated that the artificial ethylamine linker was not cleaved by DUBs in the cytosol in comparison with the native linker. However, in the case of Cy5-diUb (ethylamine linker)-Cy3, I found that the small but significant differences of FRET efficiency between the control and the samples with cytoplasmic extracts as well as DUB, which may be ascribed to binding to some cytosolic proteins (Figure 3.3.3d).

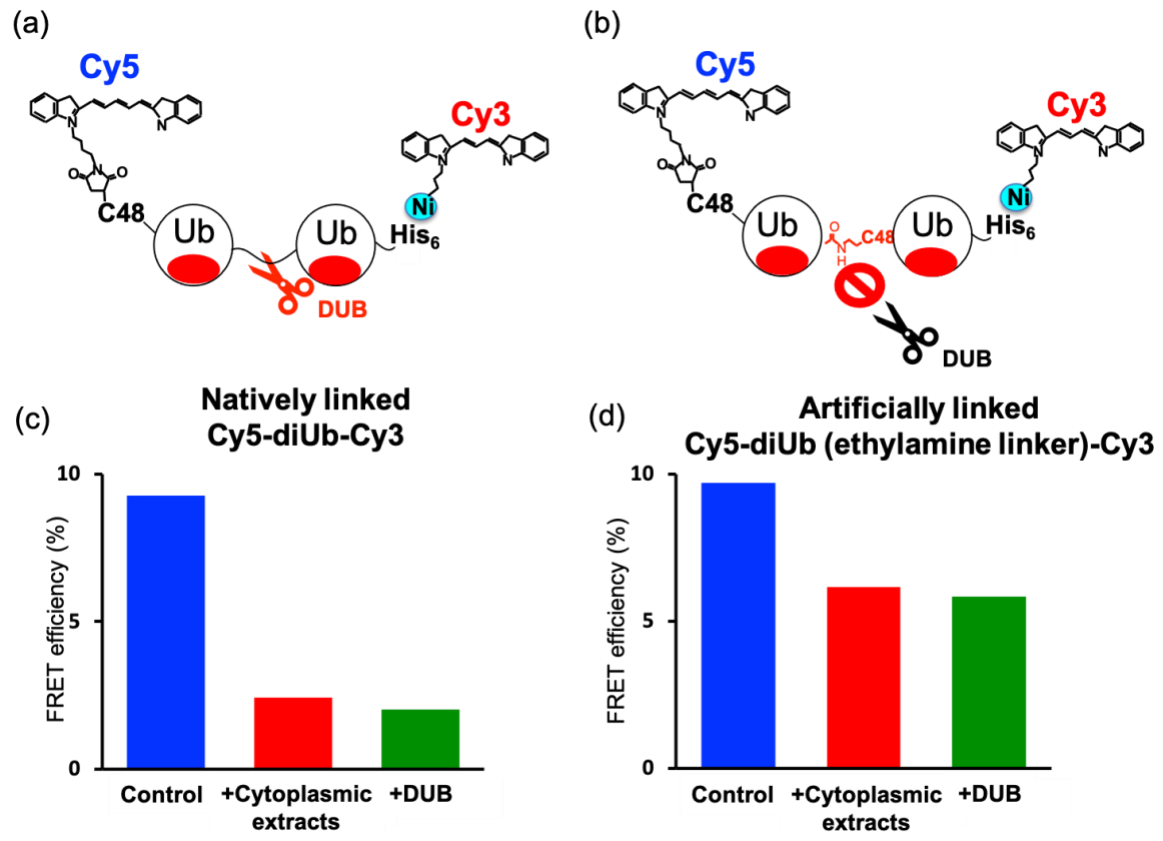


Figure 3.3.3 (a) cartoon model of natively linked Cy5-diUb-Cy3 (b) cartoon model of artificially linked Cy5-diUb (ethylamine linker)-Cy3. (c) Percent of FRET efficiency of natively linked Cy5-diUb-Cy3 calculated from the corresponding FRET intensities. FRET efficiency of Cy5-diUb-Cy3 in buffer (blue), in cytoplasmic extracts (red) and buffer with DUB (green). (d) Percent of FRET efficiency of artificially linked Cy5-diUb (ethylamine linker)-Cy3.

3.3.2 Artificial hetero-multidomain proteins

To realize my concept, I created novel hetero-multidomain protein that contains PDI$_{b'-a'}$ flanked by Ub units. In other words, One Ub was connected to the N-terminus or C-terminus of PDI$_{b'-a'}$ and two Ubs were attached to the N-terminus and C-terminus of PDI$_{b'-a'}$. I

expected that these Ub conjugations may affect the active site or the domain-domain interface of PDI$_{b'-a'}$. Furthermore, one of the most sensitive spectroscopic probes was the indole $^{1}H–^{15}N$ HSQC peak of Trp364 which is located in the vicinity of the active site of *a'* domain as well as on the domain-domain interface of PDI$_{b'-a'}$ (Figure 3.3.4a). The result showed that conjugation of Ub to the N-terminus of PDI$_{b'-a'}$ altered the chemical shift of the Trp peak of PDI$_{b'-a'}$ (Figure 3.3.4b). The chemical shift of Trp peak was also observed when conjugation of Ub to the C-terminus of PDI$_{b'-a'}$ (Figure 3.3.4c). Moreover, when two Ubs were conjugated to the N-terminus and C-terminus of PDI$_{b'-a'}$, the Trp peak was further shifted (Figure 3.3.4d and 3.3.5a). These data suggest that the conjugation of Ub unit to PDI$_{b'-a'}$ can cause micro-environmental change of active site as well as domain interface. I then performed an enzymatic reaction to link Lys48 on the N-terminal Ub to Gly76 of the C-terminal Ub *via* the isopeptide bond, resulting in the creation of a cyclized Ub- PDI$_{b'-a'}$ protein consisting of diUb and PDI$_{b'-a'}$ (cyclic diUb-PDI$_{b'-a'}$). The chemical shift of the Trp peak was further changed couple with cyclization of Ub- PDI$_{b'-a'}$ protein (Figure 3.3.4e and 3.3.5b). All these results demonstrated the possibility that Ub can regulate the dynamic structure of PDI$_{b'-a'}$.

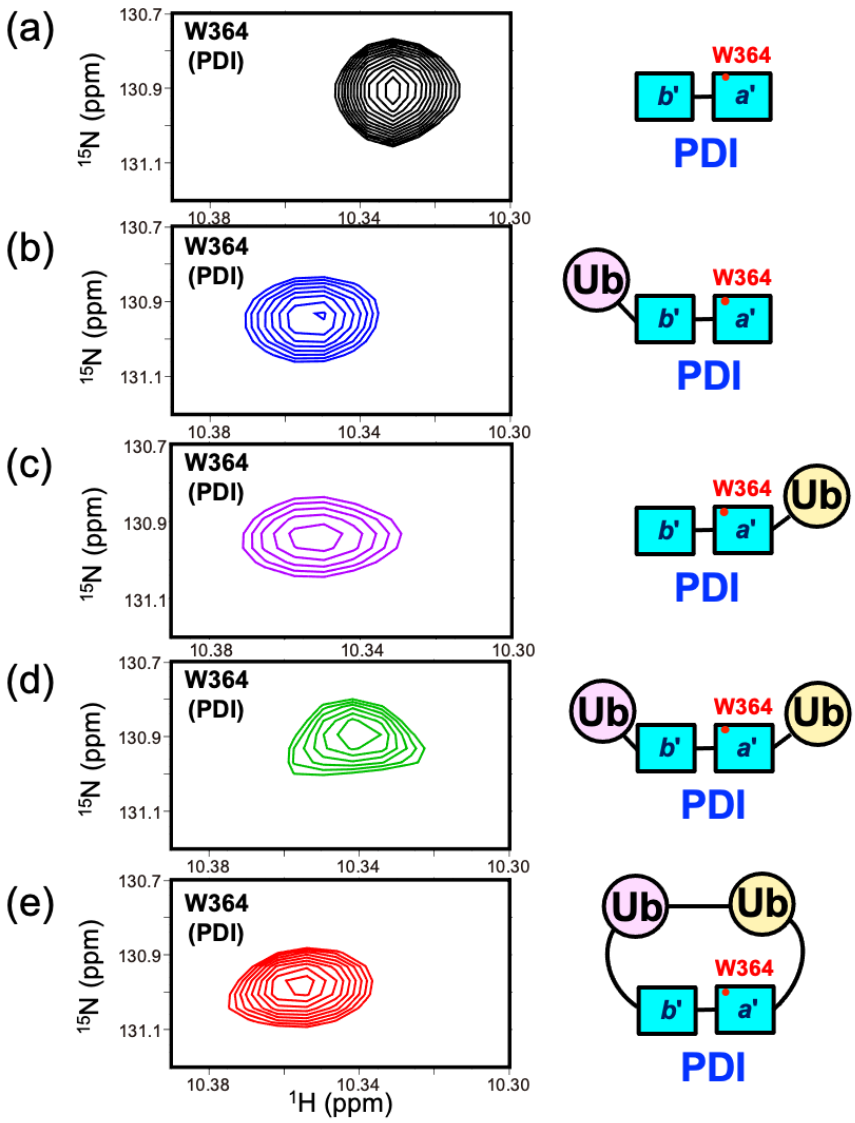


Figure 3.3.4 ^{1}H-^{15}N HSQC peaks originating from Trp364 on *a'* domain of PDI$_{b'-a'}$ with cartoon model. (a) PDI$_{b'-a'}$ (black) (b) Ub-PDI$_{b'-a'}$ (blue) (c) PDI$_{b'-a'}$-Ub (Magenta) (d) linear Ub-PDI$_{b'-a'}$-Ub (green) and (e) cyclic diUb-PDI$_{b'-a'}$ (red).

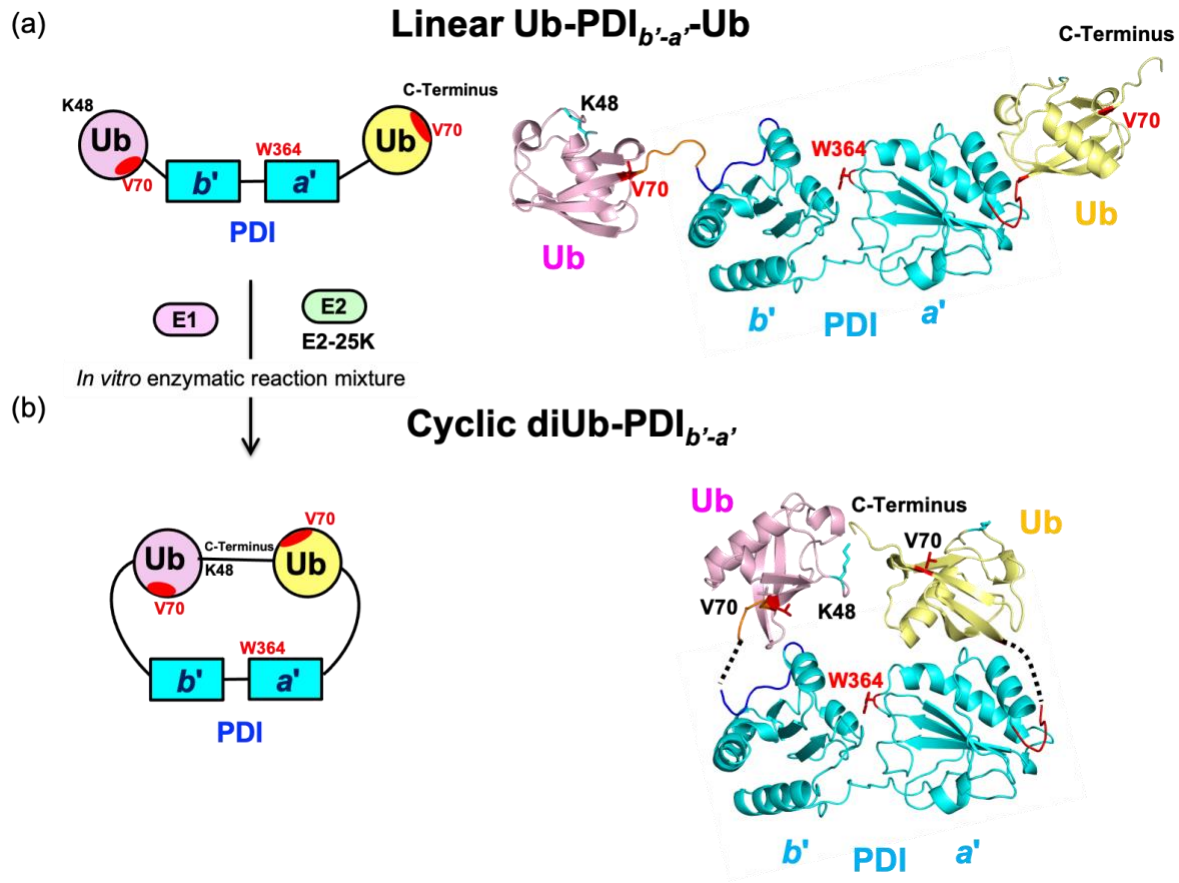


Figure 3.3.5 Structure of artificial hetero-multidomain protein. (a) linear Ub-PDI$_{b'-a'}$-Ub and (b) cyclic diUb-PDI$_{b'-a'}$ highlighting V70 on Ub and W364 on *a'* domain of PDI.

To confirm the formation of the isopeptide bond between Ub units, I performed native MS analysis. Native MS data indicated that the molecular mass determined for the linear Ub-PDI$_{b'-a'}$-Ub and the cyclic diUb-PDI$_{b'-a'}$ was 44093.62±2.39 Da and 44076.52±9.77 Da, respectively, with a difference of about 18 (Figure 3.3.6). Thus, I confirmed that the cyclic diUb-PDI$_{b'-a'}$ was successfully formed.

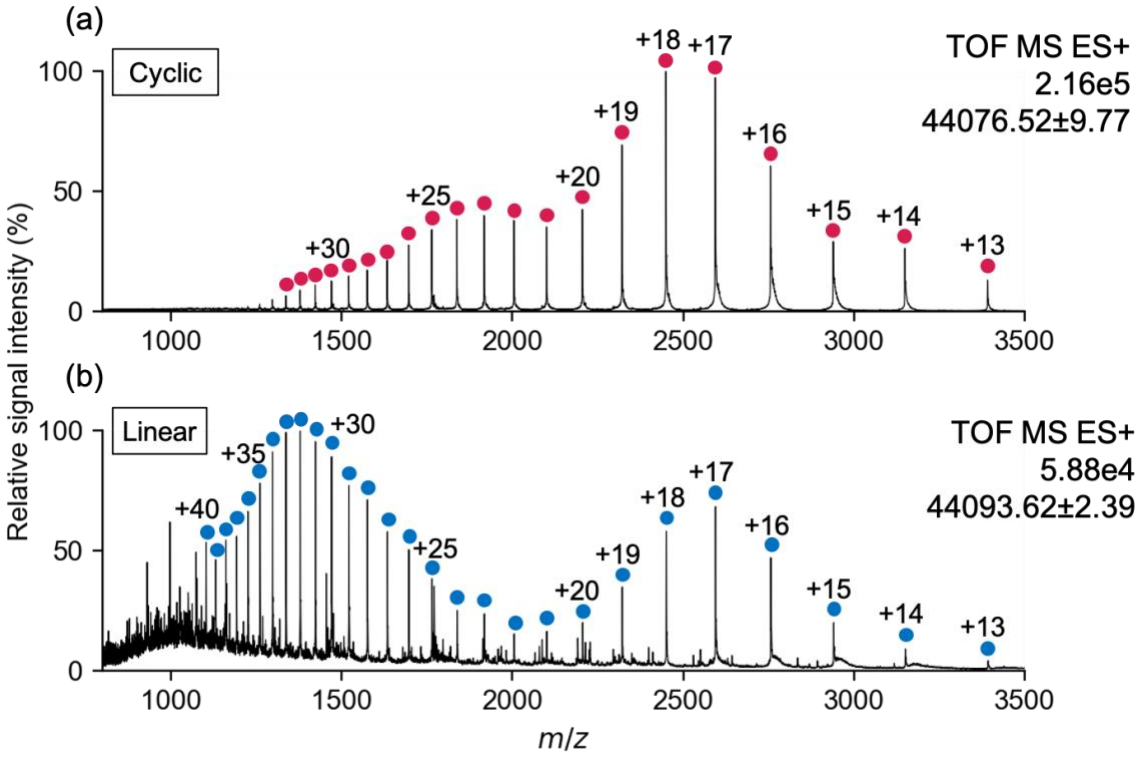


Figure 3.3.6 Native mass spectra of (a) cyclic diUb-PDI$_{b'\text{-}a'}$ showing the molecular weight was 44076.53 and (b) linear Ub-PDI$_{b'\text{-}a'}$-Ub showing the molecular weight was 44093.62.

It is intriguing to examine whether the redox-dependent domain interaction of PDI$_{b'\text{-}a'}$ affects the open-closed equilibria of diUb. Using the NMR peaks as spectroscopic probes, I firstly investigated the difference of the linear Ub-PDI$_{b'\text{-}a'}$ and the cyclic diUb-PDI$_{b'\text{-}a'}$. One of the most sensitive spectroscopic probes was the indole ^{1}H–^{15}N HSQC peak of Trp364 located in the vicinity of the active site of *a'* domain, which exhibited virtually no perturbation under reducing condition but showed a significant chemical shift change in oxidized condition upon cyclization. These results suggest that the conformation of the oxidized form of PDI$_{b'\text{-}a'}$ is more affected by cyclization than that of the reduced form (Figure 3.3.7). On the other hand, using the Val70 peak of diUb as a probe, I examined whether the redox-dependent domain interaction

of PDI affects the open-closed equilibrium of linear Ub-PDI$_{b'\text{-}a'}$ and cyclic diUb-PDI$_{b'\text{-}a'}$. In the linear form, the two Val70 peaks of Ub units were observed in the reduced state of PDI *a'* domain, whereas only one peak was in the oxidized state. In contrast, in the cyclic form, the chemical shifts of Val70 peaks of Ub units were quite different between in the reduced and in the oxidized states of PDI, although the two Val70 peaks of Ub units were observed in both the reduced and oxidized states of PDI (Figure 3.3.8). These results successfully demonstrated that the domain rearrangement of diUb was allosterically regulated according to the redox state of PDI *a'* domain. This proof of concept would be expected to lead to the creation of novel multidomain proteins for controlling their domain rearrangements in environment-responsive manners.

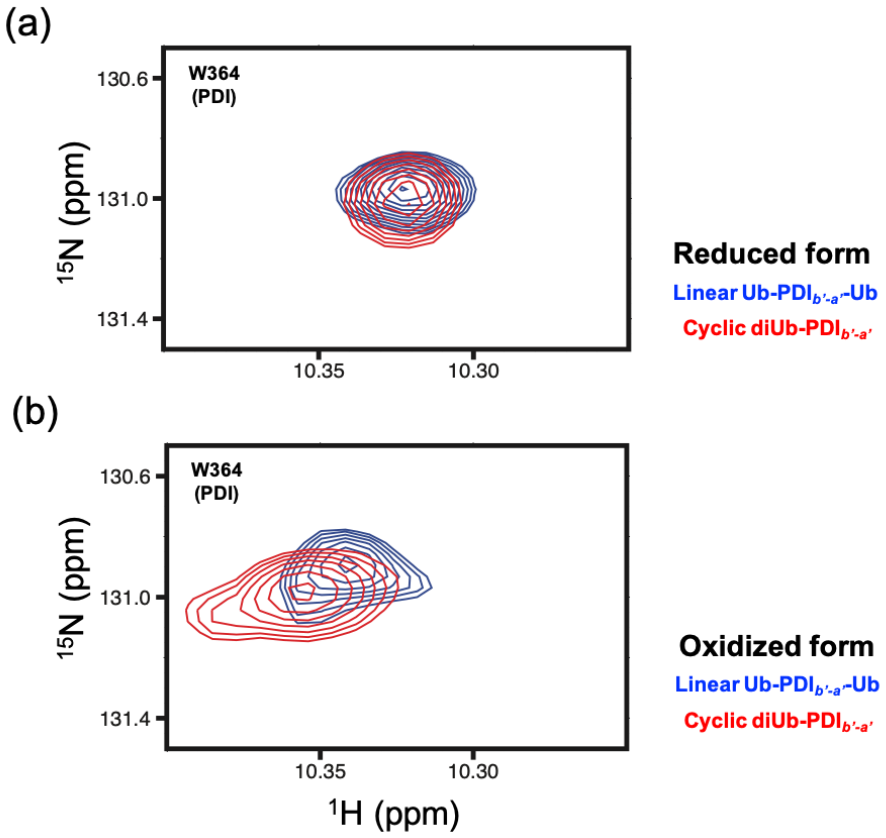


Figure 3.3.7 The indole ^{1}H-^{15}N HSQC peaks originating from W364 in PDI of the linear Ub-PDI$_{b'\text{-}a'}$-Ub (blue) and the cyclic diUb-PDI$_{b'\text{-}a'}$ (red) in (a) reduced and (b) oxidized state.

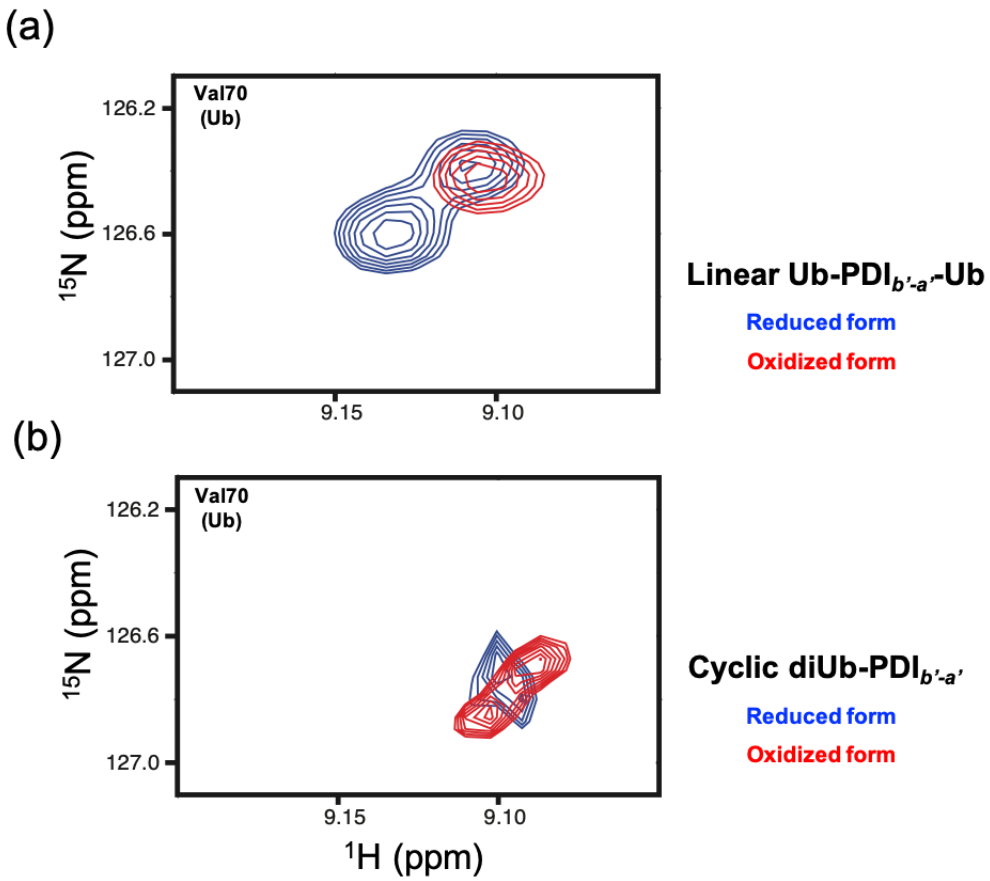


Figure 3.3.8 ^{1}H-^{15}N HSQC peaks originating from V70 of (a) the linear Ub-PDI$_{b'-a'}$-Ub and (b) the cyclic diUb-PDI$_{b'-a'}$ in reduced (blue) and oxidized state (red).

Chapter 4: Summary and perspective

Summary and perspective

In this study, I performed NMR analyses for characterizing conformational interconversions of the native forms of Lys48-linked triUb and tetraUb chains in solution. I successfully optimized the protocol of ubiquitylation reaction *in vitro* and prepared a series of Lys48-linked Ub chains with uniform isotope labeling in native and cyclic forms, carried out NMR studies, and found that the Lys48-linked triUb and tetraUb chains exhibited multiple peaks from the residues located on the hydrophobic surfaces, suggesting differences in the local environment among the Ub units. My NMR data indicated that the most distal Ub unit in the Ub chains is the most apt to expose its hydrophobic surface, suggesting its preferential involvement in interactions with the Ub-recognizing proteins. Furthermore, I found that the amino acid substitutions at position 48 in the distal Ub of the Ub chain remotely affected the solvent exposure of the hydrophobic surfaces of the other Ub units through the competitive sharing of the hydrophobic surfaces among the Ub units. This study underscores that the mutational effect at the most distal Ub is allosterically transmitted to the remaining Ub units in a chain-reaction manner.

Based on these results, I hypothesized that the Lys48-linked Ub chains may offer unique design frameworks for creating biosensing probes as well as allosterically controllable multidomain proteins. For proof of this concept, I attempted to design artificial multidomain proteins based on the Lys48-linked diUb. I constructed FRET biosensing probes, in which FRET is enhanced in the closed state of Lys48-linked diUb. Furthermore, as a prototype of artificial hetero-multidomain proteins for controlling their domain rearrangements in environment-responsive manners, I designed a chimeric protein in which diUb is cyclically conjugated with another multidomain protein, $PDI_{b'-a'}$, and successfully demonstrated that the domain rearrangement of diUb was allosterically regulated according to the redox state of this enzyme.

In summary, this study provided a quantitative view of conformational interconversions of the Lys48-linked Ub chains in solution, offering new strategies for probing and manipulating the conformational dynamics of multidomain proteins.

References

1. Vogel, C.; Bashton, M.; Kerrison, N. D.; Chothia, C.; Teichmann, S. A., Structure, function and evolution of multidomain proteins. *Curr Opin Struct Biol* **2004,** 14, (2), 208-16.
2. Göbl, C.; Madl, T.; Simon, B.; Sattler, M., NMR approaches for structural analysis of multidomain proteins and complexes in solution. *Prog Nucl Magn Reson Spectrosc* **2014,** 80, 26-63.
3. Boehr, D. D.; Nussinov, R.; Wright, P. E., The role of dynamic conformational ensembles in biomolecular recognition. *Nat Chem Biol* **2009,** 5, (11), 789-96.
4. Komander, D.; Rape, M., The ubiquitin code. *Annu Rev Biochem* **2012,** 81, 203-29.
5. Hershko, A.; Ciechanover, A., The ubiquitin system. *Annu Rev Biochem* **1998,** 67, 425-79.
6. Pickart, C. M., Mechanisms underlying ubiquitination. *Annu Rev Biochem* **2001,** 70, 503-33.
7. Deshaies, R. J.; Joazeiro, C. A., RING domain E3 ubiquitin ligases. *Annu Rev Biochem* **2009,** 78, 399-434.
8. Schulman, B. A.; Harper, J. W., Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways. *Nat Rev Mol Cell Biol* **2009,** 10, (5), 319-31.
9. Ye, Y.; Rape, M., Building ubiquitin chains: E2 enzymes at work. *Nat Rev Mol Cell Biol* **2009,** 10, (11), 755-64.
10. Alfano, C.; Faggiano, S.; Pastore, A., The Ball and Chain of Polyubiquitin Structures. *Trends Biochem Sci* **2016,** 41, (4), 371-385.
11. Finley, D., Recognition and processing of ubiquitin-protein conjugates by the proteasome. *Annu Rev Biochem* **2009,** 78, 477-513.

12. Hoege, C.; Pfander, B.; Moldovan, G. L.; Pyrowolakis, G.; Jentsch, S., RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. *Nature* **2002,** 419, (6903), 135-41.

13. Pickart, C. M.; Eddins, M. J., Ubiquitin: structures, functions, mechanisms. *Biochim Biophys Acta* **2004,** 1695, (1-3), 55-72.

14. Varadan, R.; Assfalg, M.; Raasi, S.; Pickart, C.; Fushman, D., Structural determinants for selective recognition of a Lys48-linked polyubiquitin chain by a UBA domain. *Mol Cell* **2005,** 18, (6), 687-98.

15. Cook, W. J.; Jeffrey, L. C.; Carson, M.; Chen, Z.; Pickart, C. M., Structure of a diubiquitin conjugate and a model for interaction with ubiquitin conjugating enzyme (E2). *J Biol Chem* **1992,** 267, (23), 16467-71.

16. Varadan, R.; Assfalg, M.; Haririnia, A.; Raasi, S.; Pickart, C.; Fushman, D., Solution conformation of Lys63-linked di-ubiquitin chain provides clues to functional diversity of polyubiquitin signaling. *J Biol Chem* **2004,** 279, (8), 7055-63.

17. Komander, D.; Reyes-Turcu, F.; Licchesi, J. D.; Odenwaelder, P.; Wilkinson, K. D.; Barford, D., Molecular discrimination of structurally equivalent Lys 63-linked and linear polyubiquitin chains. *EMBO Rep* **2009,** 10, (5), 466-73.

18. Liu, Z.; Gong, Z.; Jiang, W. X.; Yang, J.; Zhu, W. K.; Guo, D. C.; Zhang, W. P.; Liu, M. L.; Tang, C., Lys63-linked ubiquitin chain adopts multiple conformational states for specific target recognition. *Elife* **2015,** 4.

19. Michel, M. A.; Elliott, P. R.; Swatek, K. N.; Simicek, M.; Pruneda, J. N.; Wagstaff, J. L.; Freund, S. M.; Komander, D., Assembly and specific recognition of k29- and k33-linked polyubiquitin. *Mol Cell* **2015,** 58, (1), 95-109.

20. Cook, W. J.; Jeffrey, L. C.; Kasperek, E.; Pickart, C. M., Structure of tetraubiquitin shows how multiubiquitin chains can be formed. *J Mol Biol* **1994,** 236, (2), 601-9.

21. Varadan, R.; Walker, O.; Pickart, C.; Fushman, D., Structural properties of polyubiquitin chains in solution. *J Mol Biol* **2002,** 324, (4), 637-47.

22. Tenno, T.; Fujiwara, K.; Tochio, H.; Iwai, K.; Morita, E. H.; Hayashi, H.; Murata, S.; Hiroaki, H.; Sato, M.; Tanaka, K.; Shirakawa, M., Structural basis for distinct roles of Lys63- and Lys48-linked polyubiquitin chains. *Genes Cells* **2004,** 9, (10), 865-75.

23. Ryabov, Y.; Fushman, D., Interdomain mobility in di-ubiquitin revealed by NMR. *Proteins* **2006,** 63, (4), 787-96.

24. Eddins, M. J.; Varadan, R.; Fushman, D.; Pickart, C. M.; Wolberger, C., Crystal structure and solution NMR studies of Lys48-linked tetraubiquitin at neutral pH. *J Mol Biol* **2007,** 367, (1), 204-11.

25. Hirano, T.; Serve, O.; Yagi-Utsumi, M.; Takemoto, E.; Hiromoto, T.; Satoh, T.; Mizushima, T.; Kato, K., Conformational dynamics of wild-type Lys-48-linked diubiquitin in solution. *J Biol Chem* **2011,** 286, (43), 37496-502.

26. Kniss, A.; Schuetz, D.; Kazemi, S.; Pluska, L.; Spindler, P. E.; Rogov, V. V.; Husnjak, K.; Dikic, I.; Guntert, P.; Sommer, T.; Prisner, T. F.; Dotsch, V., Chain Assembly and Disassembly Processes Differently Affect the Conformational Space of Ubiquitin Chains. *Structure* **2018,** 26, (2), 249-258 e4.

27. Wang, Y.; Tang, C.; Wang, E.; Wang, J., PolyUbiquitin chain linkage topology selects the functions from the underlying binding landscape. *PLoS Comput Biol* **2014,** 10, (7), e1003691.

28. Bowerman, S.; Rana, A.; Rice, A.; Pham, G. H.; Strieter, E. R.; Wereszczynski, J., Determining Atomistic SAXS Models of Tri-Ubiquitin Chains from Bayesian Analysis of Accelerated Molecular Dynamics Simulations. *J Chem Theory Comput* **2017,** 13, (6), 2418-2429.

29. Berg, A.; Kukharenko, O.; Scheffner, M.; Peter, C., Towards a molecular basis of ubiquitin signaling: A dual-scale simulation study of ubiquitin dimers. *PLoS Comput Biol* **2018,** 14, (11), e1006589.

30. Virdee, S.; Ye, Y.; Nguyen, D. P.; Komander, D.; Chin, J. W., Engineered diubiquitin synthesis reveals Lys29-isopeptide specificity of an OTU deubiquitinase. *Nat Chem Biol* **2010,** 6, (10), 750-7.

31. Castaneda, C. A.; Kashyap, T. R.; Nakasone, M. A.; Krueger, S.; Fushman, D., Unique structural, dynamical, and functional properties of k11-linked polyubiquitin chains. *Structure* **2013,** 21, (7), 1168-81.

32. Kristariyanto, Y. A.; Choi, S. Y.; Rehman, S. A.; Ritorto, M. S.; Campbell, D. G.; Morrice, N. A.; Toth, R.; Kulathu, Y., Assembly and structure of Lys33-linked polyubiquitin reveals distinct conformations. *Biochem J* **2015,** 467, (2), 345-52.

33. Goddard, T. D.; Koeller, D. G., *Sparky*, Version 3.0, University of California: San Francisco, CA, USA. **1993**.

34. Otwinowski, Z.; Minor, W., Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **1997,** 276, 307-26.

35. Vagin, A.; Teplyakov, A., MOLREP: an Automated Program for Molecular Replacement. *J Appl Cryst* **1997,** 30, 1022-1025.

36. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **2010,** 66, (Pt 4), 486-501.

37. Murshudov, G. N.; Vagin, A. A.; Dodson, E. J., Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **1997,** 53, (Pt 3), 240-55.

38. Chen, V. B.; Arendall, W. B., 3rd; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C., MolProbity: all-

atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **2010,** 66, (Pt 1), 12-21.

39. Satoh, T.; Sakata, E.; Yamamoto, S.; Yamaguchi, Y.; Sumiyoshi, A.; Wakatsuki, S.; Kato, K., Crystal structure of cyclic Lys48-linked tetraubiquitin. *Biochem Biophys Res Commun* **2010,** 400, (3), 329-33.

40. Marino, S. M.; Gladyshev, V. N., Cysteine function governs its conservation and degeneration and restricts its utilization on protein surfaces. *J Mol Biol* **2010,** 404, (5), 902-16.

41. Serve, O.; Kamiya, Y.; Maeno, A.; Nakano, M.; Murakami, C.; Sasakawa, H.; Yamaguchi, Y.; Harada, T.; Kurimoto, E.; Yagi-Utsumi, M.; Iguchi, T.; Inaba, K.; Kikuchi, J.; Asami, O.; Kajino, T.; Oka, T.; Nakasako, M.; Kato, K., Redox-dependent domain rearrangement of protein disulfide isomerase coupled with exposure of its substrate-binding hydrophobic surface. *J Mol Biol* **2010,** 396, (2), 361-74.

42. Nakasako, M.; Maeno, A.; Kurimoto, E.; Harada, T.; Yamaguchi, Y.; Oka, T.; Takayama, Y.; Iwata, A.; Kato, K., Redox-dependent domain rearrangement of protein disulfide isomerase from a thermophilic fungus. *Biochemistry* **2010,** 49, (32), 6953-62.

Acknowledgement

First and foremost, I would like to express my sincere heartfelt thanks to my supervisor Prof. Koichi Kato for giving me the opportunity to join the wonderful group and for his professional guidance, constant support, and patience in supervision. His scientific proficiency helped me at various stages of my research especially during the concept defining stage.

Next, I would like to extend my sincere gratitude to Dr. Maho Yagi-Utsumi from Institute for Molecular Science (IMS) for her brilliant idea and immense help for teaching me the experimental skills and discussing for solving the problems. I have learned many experimental techniques from her. I am truly thankful for all her time and efforts her devoted for my project.

I would like to acknowledge Dr. Saeko Yanaka from IMS for her guidance, valuable comments, and experimental help. I would also like to extend my gratitude to Dr. Tadashi Satoh and Dr. Hirokazu Yagi from Nagoya City university for their kind help and valuable comments.

In regard to the collaborators, I sincerely thank Dr. Norie Momiyama and Dr. Naoya Ohtsuka for synthesis of the ethylamine linker. I thank Prof. Masaaki Sugiyama and Dr. Rintaro Inoue from Kyoto University Research Reactor Institute and Dr. Hiroshi Nakagawa from Japan Atomic Energy Agency for their guidance during SAXS and SANS measurement; Dr. Atsuji Kodama for his help in native mass measurement.; Prof. Takayuki Uchihashi, Dr. Shigetaka Nishiguchi, and Dr. Christian Ganser for their help in HS-AFM measurement.

Here I would like to thanks to my lab members including Dr. Sae Tanaka, Dr. Tatsuya Suzuki, Dr. Arunima Sikdar, Dr. Tong Zhu, Dr. Gengwei Yan, Ms. Rena Honda, Dr. Mesayamas Kongsema, Dr. Benjawan Jityuti, Mr. Eanwai Goh, Ms. Thunchanok Wilasri, Dr. Rina Yogo and Mr. Taichiro Sekiguchi. I would like to appreciate all the help provided by technician and secretaries including Ms. Kei Tanaka, Ms. Yukiyo Fukutomi Ms. Tomo Okada,

and Ms. Yukiko Isono.

I would like to thanks to all of my international friends in Okazaki include internship students every year for their love, friendship and support. I have enjoyed a quality time with them. Furthermore, a great thanks to Mr. Ryoma Seiki, Dr. Akasit Visootsat (IMS), Dr. Tianchai Chooppawa (IMS), Dr. Chanantida Jongwohan (IMS), Ms. Monique Honsa (IMS), Ms. Nawarat Rattanajearakul (NIPS), and Mr. Ruttapol Malatong (IMS) who have been supporting and listen to me all the time.

Last but not least, sincerely appreciate my mother, for her standing by me all the time and supporting all my pursuits, as well as her endless comfort and encouragement whenever I feel bad and down. It is her love made me keep the faith so that I could survive the most though period of time in my Ph.D. pursuit and stick to the end. Many thanks to my big family members for their love and encouragement.