Regulation and Disruption of Molecular Assembly in Intrinsically Disordered Protein by Molecular Chaperon

分子シャペロンによる天然変性タンパク質の分子集合制御メカニズムとその破綻

Honoka Kawamukai

川向　ほの香

Graduate School of Chemical Sciences and Engineering

Hokkaido University

北海道大学　大学院総合化学院

2024

ACKNOWLEDGEMENTS

I would like to express sincere gratitude to Professor Koichiro Ishimori. He always gave me sound advice. Especially, he encouraged me to study abroad so that I could have many good experiences. I am also pleased thank Secretary Maki Tanaka for accepting the troublesome office procedure, and the member of Structural Chemistry Laboratory for helps and assistance.

From February 2021, I have been conducting research at Tokushima University as a special student. I sincerely appreciate Professor Tomohide Saio. His attitude toward his research has inspired me. I also thank the previous and current member of Molecular Life Science Laboratory for helps and assistance, especially Dr. Motonori Matsusaki, Dr, Soichiro Kawagoe, Dr. Munehiro Kumashiro, Dr. Asuka Mukai, Dr. Yoshikazu Hattori, Ms. Eri Sakamoto, Ms, Miki Sanada, Ms. Miyako Sogawa, Ms. Azusa Hirano, and Ms Kotona Kato for contributing to this work.

It should be emphasized that the studies in this thesis have required the cooperation with many investigators. I wish to express sincere appreciation to Dr. Hiroyuki Kumeta (Hokkaido University) for the assistance with NMR measurement. I would like to express my special gratitude to Prof. Nobutaka Shimizu (High Energy Accelerator Research Organization) for the assistance with SAXS measurement.

At the review of this work, Professor Kazuyasu Sakaguchi (Laboratory of Biological Chemistry), Professor Kazuki Sada (Material Chemistry Laboratory), Professor Kei Murakoshi (Physical Chemistry Laboratory), and Professor Kenichiro Matsumoto (Laboratory of Biosynthetic Chemistry) gave me the valuable suggestion and guidance.

I have received a generous support from Japan Society for the Promotion of Science (JSPS) Research Fellowship DC1 since April 2021. This work was financially supported by a Grants-Aid for Scientific Research.

I express my deep appreciation to my family, who are Mr. Hitoshi Kawamukai, Ms. Fumie Kawamukai, Ms. Himeno Kawamukai with my whole heart. I am grateful that my pet Seven, Mamedaifuku, Bun, and Mikke has healed me.

March 2023

Graduate School of Chemical Science and Engineering

Hokkaido University

Honoka Kawamukai

LIST OF PUBLICATIONS

CHAPTER Ⅱ

"*C9orf72*-derived arginine-rich poly-dipeptides impede phase modifiers"

Hitoki Nanaura, Honoka Kawamukai , Ayano Fujiwara, Takeru Uehara, Yuichiro Aiba, Mari Nakanishi, Tomo Shiota, Masaki Hibino, Pattama Wiriyasermkul, Sotaro Kikuchi, Riko Nagata, Masaya Matsubayashi, Yoichi Shinkai, Tatsuya Niwa, Taro Mannen, Naritaka Morikawa, Naohiko Iguchi, Takao Kiriyama, Ken Morishima, Rintaro Inoue, Masaaki Sugiyama, Takashi Oda, Noriyuki Kodera, Sachiko Toma-Fukai, Mamoru Sato, Hideki Taguchi, Shushi Nagamori, Osami Shoji, Koichiro Ishimori, Hiroyoshi Matsumura, Kazuma Sugie, Tomohide Saio, Takuya Yoshizawa, Eiichiro Mori

Nat Commun **12**, 5301 (2021).

CHAPTER Ⅲ

"*Conserved loop of a phase modifier endows condensates with fluidity*"

Honoka Kawamukai, Motonori Matsusaki, Yoichi Shinkai4, Shunsuke Tomita, Ken Morishima, Mai Watabe, Taro Mannen, Tatsuya Niwa, Takuya Mabuchi, Yuichiro Aiba, Kotona Kato, Hiroyuki Kumeta, Hitoki Nanaura, Shingo Kanemura, Masaaki Sugiyama, Masaki Okumura, Takuya Yoshizawa, Eiichiro Mori, and Tomohide Saio

Manuscript in preparation

CHAPTER Ⅳ

"*Conformational Distribution of a Multi-Domain Protein Measured by Single-Pair Small Angle X-Ray Scattering*"

Honoka Kawamukai, Shumpei Takishita, Kazumi Shimizu, Daisuke Kohda,

Koichiro Ishimori, and Tomohide Saio

Phys. Chem. Let., **15**, 744-750 (2024).

LIST OF PRESENTATIONS

Oral Presentations

1. Honoka Kawamukai, Shumpei Takisita, Tomohide Saio, Koichiro Ishimori
   "Evaluation of Protein Conformation by Measuring the Distance between Lanthanide Ions"
   Chemistry Society Hokkaido Branch 2020 (Sapporo, Japan)
   Jan 28-29 (2020)

2. Honoka Kawamukai, Shumpei Takisita, Tomohide Saio, Koichiro Ishimori
   "Development of a new method for protein conformational analysis using small-angle X-ray scattering of heavy-atom ions."
   Annual Biophysical Society Hokkaido Branch 2108 (Sapporo, Japan)
   Mar 14 (2018)

3. Honoka Kawamukai, Koichiro Ishimori, Tomohide Saio,
   "Pro-Arg polydipeptide inhibits nuclear transport factor Kapβ2 from regulating liquid-liquid phase separation of FUS"
   The 22ed Annual Meeting of Protein Science Society of Japan (Nagoya, Japan)
   Jul 14 (2022)

Poster Presenation

1. Honoka Kawamukai, Shumpei Takisita, Tomohide Saio, Koichiro Ishimori
   "Measurement of interdomain distance of a multidomian protein"
   CSE International Student Symposium 2019 (Sapporo, Japan)
   Jul 18-21(2019)

2. Honoka Kawamukai, Shumpei Takisita, Tomohide Saio, Koichiro Ishimori
   "Measurement of interdomain distances in multidomain proteins using heavy metal ions"
   The 92th Annual Meeting of Japanese Biochemical Society (Yokohama, Japan)
   Sep 18-20 (2019)

3. Honoka Kawamukai, Koichiro Ishimori, Tomohide Saio
"Droplet formation mechanism by Fused in sarcoma (FUS) and amyloidation due to its disruption"
10th CSJ Chemistry Festa (online)
Oct 20 (2020)

4. Honoka Kawamukai, Koichiro Ishimori, Tomohide Saio
"Regulatory mechanism of liquid-liquid phase separation in Fused in sarcoma (FUS), the protein responsible for amyotrophic lateral sclerosis"
The 43th Annual Meeting of the Molecular Biology Society of Japan (online)
Dec 3 (2020)

5. Honoka Kawamukai, Koichiro Ishimori, Tomohide Saio
"Toxic PR poly-dipeptides inhibit the chaperone activity of Kapβ2"
ISMAR-APNMR2021 (online)
Aug 24 (2021)

6. Honoka Kawamukai, Koichiro Ishimori, Tomohide Saio
"PRn inhibits the binding of KapB2 to FUS"
The 16th International Symposium of the Institute Network for Biomedical Sciences KEY FORUM 2021 International Symposium (online)
Nov 12 (2021)

7. Honoka Kawamukai, Tomohide Saio, Koichiro Ishimori
"Mechanistic insight into regulation and dysregulation of liquid-liquid phase separation of FUS"
The 2021 International Chemical Congress of Pacific Basin Societies (online)
Dec 19 (2021)

8. Honoka Kawamukai, Koichiro Ishimori, Tomohide Saio
   “Pro-Arg polydipeptide inhibits nuclear transport factor Kapβ2 from regulating liquid-liquid phase separation of FUS”
   The 22nd Annual Meeting of Protein Science Society of Japan (Tsukuba, Japan)
   Jun 7-9 (2022)

9. Honoka Kawamukai, Koichiro Ishimori, Tomohide Saio
   “Mechanism of functional inhibition in impotin Kapβ2 by Pro-Arg polydipeptide”
   The 23rd Annual Meeting of Protein Science Society of Japan (Nagoya, Japan)
   Jul 5-7 (2023)

Contents

Ⅰ. General Introduction

1.1 Molecular assembly in intrinsically disordered protein

In general, proteins function by folding into a structure determined by their amino acid sequence. On the other hand, there are proteins that have intrinsically disordered region (IDR) that do not readily adopt a defined conformation, and these are known as intrinsically disordered protein (IDP). IDP form droplets by multivalent interactions between IDR to exclude surrounding molecules and function as membrane-less organelles (Figure 1.1). This phenomenon is also known as “biological phase separation”. Many membrane-less organelles contain RNA molecules, and the IDR of the RNA-binding proteins that bind to them have the activity to induce biological phase separation. The functions of membrane-less organelles include the incorporation of substrates and enzymes for specific biochemical reactions to increase reaction efficiency and the retention of specific factors in the droplet to inhibit their actions[1]. In the nucleus, membrane-less organelles exist in the gaps between chromatin regions and influence the arrangement and activity of specific chromatin sequences by interacting with chromatin[1]. In addition to these findings on molecular mechanisms, biological phase separation is widely implicated in physiological phenomena at various cellular and individual levels. In addition to gene expression involving RNA and DNA, many factors involved in various intracellular phenomena such as proteolysis, cytoskeletal polymerization, signal transduction, and autophagy have IDR. Biological phase separation through them controls the specificity and efficiency of biological phenomena[2].

Recently, biological phase separation has been reported to be associated with intractable diseases. A molecular mechanism has been proposed in which droplets formed through IDR are stabilized over time, transformed into aggregates, and accumulate within the cell. Thus, it has been suggested that disruption of the droplet regulation mechanism of IDP can lead to the development of disease.

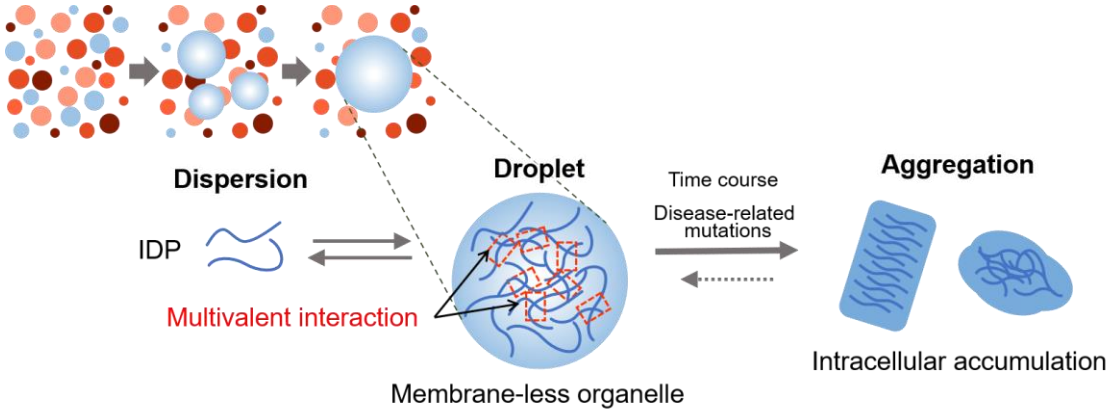


Figure 1.1 Molecular assembly in intrinsically disordered protein

IDP exclude surrounding molecules and form droplets.

1.2 Biological phase separation related to disease.

One disease in which IDP aggregates accumulate is the neurodegenerative disease amyotrophic lateral sclerosis (ALS). ALS is a disease characterized by selective and progressive degeneration and loss of upper and lower motor neurons. The disease progresses relatively rapidly and there are currently few effective treatments.

One of the proteins identified as a cause of ALS is Fused in sarcoma (FUS), an RNA-binding protein. FUS is a representative protein of biological phase separation and has been particularly well studied. FUS is a protein involved in transcription, splicing, DNA repair, and axonogenesis, and assembles a group of related molecules. It was reported in 2009 that FUS is the responsible gene for familial Amyotrophic Lateral Sclerosis (ALS) linked to chromosome 16[3]. The FUS is 526 amino acids long, of which the RNA recognition motif (RRM) of about 100 amino acids and the Zn finger (ZnF) of about 30 amino acids are tertiary structures, and about 70% of the rest is IDR (Figure 1.2). The N-terminal side of FUS is composed of a low-complexity (LC) domain whose amino acid composition is rich in four amino acids, serine, tyrosine, glycine, and glutamine (SYGQ-rich), followed by three RGG repeat regions with arginine-glycine-glycine sequence repeated between the RRM and ZnF. There are three RGG repeat regions with repeated arginine-glycine-glycine sequences after that. The C-terminal nuclear transfer signal (PY-NLS) is recognized by the nuclear transport receptor karyopherin-β2 (Kapβ2), and FUS is transported into the nucleus by Kapβ2.

The N-terminal LC domain is known to form an amyloid-like cross β-polymer, which is thought to be responsible for self-recognition in phase separation. The structure of the cross β-polymer

has been analyzed by various techniques including solid-state NMR, X-ray crystallography, electron crystallography, and cryo-EM[4,5]. The RGG region by itself has almost no phase separation property, but when combined with the LC domain, it promotes phase separation. Cation-π interactions between the tyrosine of the LC domain and the arginine residues of the RGG repeats have been shown to be important for this[6] (Figure 1.3). In droplets resulting from phase separation of FUS, the aforementioned interactions and others are considered to exist in a disorderly manner. However, there are many unknowns about the droplet formation process prior to hydrogel formation and the structure inside the droplet.

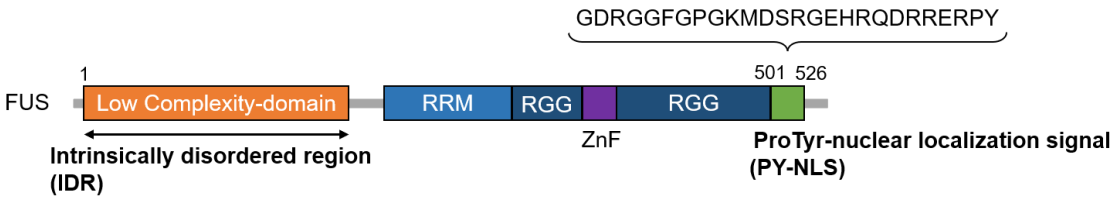


Figure 1.2 Domain structure of FUS

Low Complexity-domain is on the N-terminal side and Nuclear localization signal is on the C-terminal side.

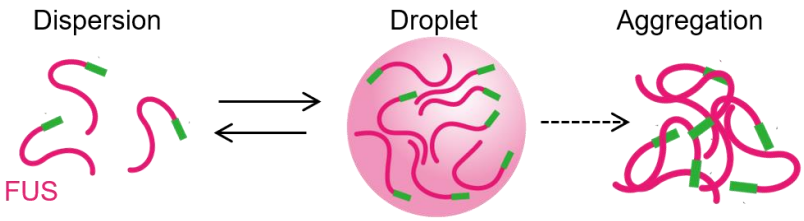


Figure 1.3 Overview of FUS molecular assembly

FUS is in equilibrium between the dispersed state and the droplet, but when this equilibrium is disrupted, aggregation occurs.

1.3 Chaperones inhibit FUS droplet formation (ChapterII, III)

The transporter that transports FUS into the nucleus is Kapβ2 (Figure 1.4). Kapβ2 is a member of the importin β family of nuclear transport receptors and transports PY-NLS-containing proteins into the nucleus. In humans, there are approximately 20 different members of the importin β family, each of which recognizes a specific NLS and transports a substrate into the nucleus[7]. Genetic mutations associated with ALS-associated amino acid mutations in FUS are concentrated in the PY-

NLS region. The pathogenesis of FUS deposition in the cytoplasm suggests a link between the disruption of nuclear transport of FUS by Kapβ2 and the disease. In addition, when Kapβ2 was added after FUS droplet formation, the FUS droplets disappeared quickly. This suggests that Kapβ2 also functions as a phase separation chaperone[8].Inhibition of Kapβ2 function may prevent Kapβ2 from recognizing FUS, but the molecular mechanism of this inhibition remains unclear. In Chapters 2 and 3, I aim to elucidate the mechanism for dysregulation of FUS phase separation caused by inhibition of Kapβ2 activity.

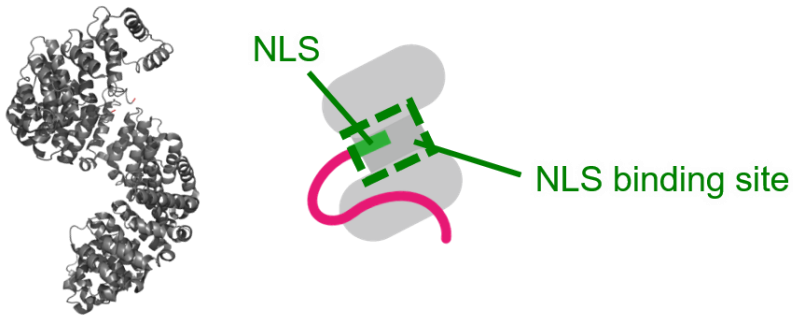


Figure 1.4 Structure of Kapβ2

Kapβ2 interacts with the NLS of FUS at the NLS binding site.

1.4 Toxic expression of proline-arginine poly-dipeptide (Chapter II,III)

This section discusses the causes of the inhibition of Kapβ2 function (Figure 1.5). A hexanucleotide repeat expansion in *C9orf72* is the most prevalent form of familial ALS and frontotemporal dementia (C9-ALS/FTD)[9]. Similar to other forms of ALS, mislocalization and aggregation of RBPs are observed in C9-ALS/FTD[10]. The main mechanisms of C9-ALS/FTD have been proposed as follows: *C9orf72* haploinsufficiency, toxic gain-of-function from repeat RNA or poly-dipeptides, or some combination of the above[11]. Repeated expansion in *C9orf72* produces five different poly-dipeptides, proline:arginine (PR), glycine:arginine (GR), proline:alanine (PA), glycine:proline (GP), and glycine:alanine (GA), through repeat-associated non-AUG (RAN) translation[12,13]. Arginine-rich, PR and GR, poly-dipeptides show similarly high toxic effect on membraneless organelles, proteins with LC-domains, and nucleocytoplasmic transport (NCT)[9,14,15]. Arginine-rich poly-dipeptides bind proteins with LC-domains and stabilize polymers formed by protein–protein interaction through LC-domains[16,17]. Existing genetic studies reveal that the repeat

expansion in *C9orf72* gene disrupts NCT[17,18].

In Chapter II, I explain the mechanism by which PR inhibits Kapβ2 function. First, to confirm the interaction between Kapβ2 and PR, I performed size-exclusion chromatography with multi-angle light scattering (SEC-MALS) and fluorescence-spectroscopy. In addition, NMR was used to determine the site of interaction of PR with Kapβ2. I discuss the mechanism by which PR inhibits the interaction between Kapβ2 and FUS. In Chapter III, I examine the mechanism by which excess amounts of PRn inhibit Kapβ2 function. There may be excess amounts of PRn relative to Kapβ2 in the cell because PRn is persistent. First, I used NMR to discover another site of interaction. Furthermore, I found that excess amounts of PRn reduce the diffusion rate of FUS in droplets.

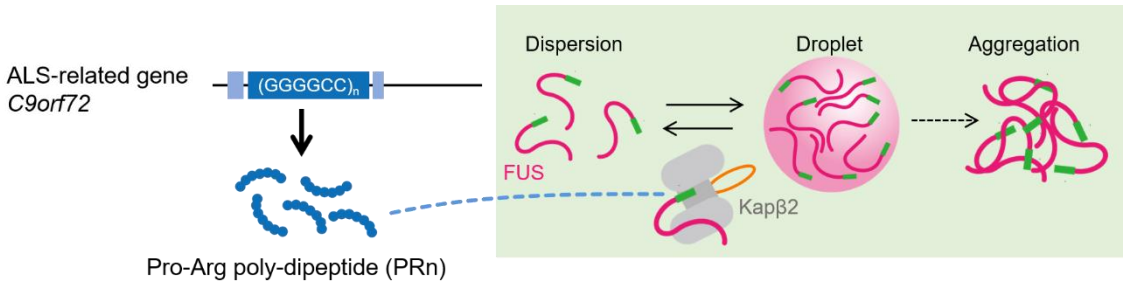


Figure 1.5 Relationship between PRn and Kapβ2
PRn inhibits the interaction between Kapβ2 and FUS.

1.5 Measurement of protein conformation distribution using Single-Pair Small Angle X-Ray Scattering (SAXS) (Chapter IV)

In Chapter Ⅳ, I discuss the development of a new methodology and its application to examine what kind of structural changes are induced in the molecules inside the droplet to induce their aggregation. When the droplet control function of FUS fails, the molecular structure inside the FUS droplet changes and aggregation occurs. Understanding the structure inside the droplet is necessary for the FUS aggregation mechanism. Because FUS is a IDP that can take on various conformations, and because multiple conformations exist in equilibrium, it is difficult to trace conformational changes in the protein by crystallography or NMR. Therefore, I focused on the fact that conformational changes of a protein can also be traced as changes in the distance between two specific amino acid residues in

its sequence or in the distribution of such distances. Since the distance between two points can be calculated from scattering interferences in small-angle X-ray scattering (SAXS), I fixed metal ions with high electron density near two specific amino acid residues and tried to extract only the scattering interferences originating from the metal ions from the protein-derived scattering interferences. To establish this approach, I utilized MurD, which is involved in peptidoglycan biosynthesis, where the steric structure has been reported and the electron-dense lanthanide metal ion can be anchored to a specific site.

References

1. Shin, Y. & Brangwynne, C. P. Liquid phase condensation in cell physiology and disease. *Science (80-.).* **357**, (2017).
2. Alberti, S. & Hyman, A. A. Biomolecular condensates at the nexus of cellular stress, protein aggregation disease and ageing. *Nat. Rev. Mol. Cell Biol.* **22**, 196–213 (2021).
3. Kwiatkowski, T. J. *et al.* Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science (80-.).* **323**, 1205–1208 (2009).
4. Kato, M. *et al.* Cell-free Formation of RNA Granules: Low Complexity Sequence Domains Form Dynamic Fibers within Hydrogels. *Cell* **149**, 753–767 (2012).
5. Murray, D. T. *et al.* Structure of FUS Protein Fibrils and Its Relevance to Self-Assembly and Phase Separation of Low-Complexity Domains. *Cell* **171**, 615-627.e16 (2017).
6. Qamar, S. *et al.* FUS Phase Separation Is Modulated by a Molecular Chaperone and Methylation of Arginine Cation-π Interactions. *Cell* **173**, 720-734.e15 (2018).
7. Xu, D., Farmer, A. & Chook, Y. M. Recognition of nuclear targeting signals by Karyopherin-β proteins. *Current Opinion in Structural Biology* vol. 20 782–790 at https://doi.org/10.1016/j.sbi.2010.09.008 (2010).
8. Yoshizawa, T. *et al.* Nuclear Import Receptor Inhibits Phase Separation of FUS through Binding to Multiple Sites. *Cell* **173**, 693-705.e22 (2018).
9. Odeh, H. M. & Shorter, J. Arginine-rich dipeptide-repeat proteins as phase disruptors in C9-ALS/FTD. *Emerg. Top. Life Sci.* **4**, 293–305 (2020).
10. Keller, B. A. *et al.* Co-aggregation of RNA binding proteins in ALS spinal motor neurons: evidence of a common pathogenic mechanism. *Acta Neuropathol.* **124**, 733–747 (2012).
11. Balendra, R. & Isaacs, A. M. C9orf72-mediated ALS and FTD: multiple pathways to disease. *Nat. Rev. Neurol.* **14**, 544–558 (2018).
12. Mori, K. *et al.* Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. *Acta Neuropathol.* **126**, 881–893 (2013).
13. Zu, T. *et al.* RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. *Proc. Natl. Acad. Sci. U. S. A.* **110**, (2013).
14. Kwon, I. *et al.* Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science (80-.).* **345**, 1139–1145 (2014).
15. Mizielinska, S. *et al.* C9orf72 repeat expansions cause neurodegeneration in Drosophila through

arginine-rich proteins. *Science (80-.).* **345**, 1192–1194 (2014).

16. Lin, Y. *et al.* Toxic PR Poly-Dipeptides Encoded by the C9orf72 Repeat Expansion Target LC Domain Polymers. *Cell* **167**, 789-802.e12 (2016).
17. Lee, K.-H. *et al.* C9orf72 Dipeptide Repeats Impair the Assembly, Dynamics, and Function of Membrane-Less Organelles. *Cell* **167**, 774-788.e17 (2016).
18. Freibaum, B. D. *et al.* GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. *Nature* **525**, 129–133 (2015).

Ⅱ. *C9orf72*-derived arginine-rich poly-dipeptides impede phase modifiers

Abstract

It has been reported that ALS disease produces Pro-Arg poly-dipeptide (PRn), an abnormally elongated repeat sequence in ALS-related genes. In Chapter 2, I attempt to identify the interaction site between Kapβ2 and PRn, assuming that this PRn interacts with Kapβ2 to inhibit the droplet regulatory function of FUS. However, since Kapβ2 has a molecular weight of 101 kDa, NMR analysis of Kapβ2 is difficult even when uniformly labeled with stable isotopes, so I used an amino acid side-chain-selective ^{13}C label to identify the interaction sites of hydrophobic amino acid residues on the side chains of Kapβ2. Therefore, I measured ^{1}H-^{13}C HSQC spectra of hydrophobic amino acid residue side chains using amino acid side chain-selective 13C labeling. I observed that the addition of PRn altered the chemical shifts of certain NMR signals, some of which were attributed to the nuclear translocation signal (NLS) binding site in Kapβ2. It has been reported that FUS also binds to the NLS binding site of Kapβ2. Furthermore, the dissociation constants of PRn and FUS for Kapβ2 are comparable. It was suggested that PRn competitively binds to the NLS binding site of FUS, thereby inhibiting the binding of Kapβ2 to FUS, resulting in excessive self-aggregation of FUS.

2.1 Introduction

FUS is a intrinsically disordered protein and one of the causative proteins of ALS. FUS has a low-complexity sequence, a flexible region in which up to several amino acids occur repeatedly and do not assume a specific structure. Through intermolecular π-π, cation-π, hydrophobic, and even electrostatic interactions via this low-complexity region, FUS self-assembles, resulting in the formation of droplets that exclude solvent, thereby performing RNA metabolism such as RNA transcription and splicing (Figure 2.1). Since excessive self-assembly of FUS results in the formation of FUS aggregates observed in ALS disease, the molecular chaperone Karyopherinβ2 (Kapβ2) is thought to interact with FUS and inhibit its self-assembly, thereby suppressing droplet formation. Although such functional inhibition of Kapβ2 has been postulated as a possible cause of ALS, the structural chemistry underlying the molecular mechanism has not been fully elucidated.

Hexanucleotide repeat elongation in the ALS-associated gene *C9orf72* is observed in ALS, where excessive FUS self-association progresses. The main mechanisms of C9-ALS/FTD have been proposed as follows: *C9orf72* haploinsufficiency, toxic gain-of-function from repeat RNA or poly-dipeptides, or some combination of the above. Repeat expansion in *C9orf72* produces five different poly-dipeptides, proline:arginine (PR), glycine:arginine (GR), proline:alanine (PA), glycine:proline (GP), and glycine:alanine (GA), through repeat-associated non-AUG (RAN) translation[1,2]. Arginine-rich, PR and GR, poly-dipeptides show similarly high toxic effect on membraneless organelles, proteins with LC-domains, and nucleocytoplasmic transport [3,4]. Unaffected individuals typically have between two and 23 GGGGCC repeats, whereas people with *C9orf72*-related ALS have hundreds or even thousands of repeats.

In Chapter 2, I attempted to identify the site of interaction between Kapβ2 and PRn, assuming that this PRn interacts with Kapβ2 to inhibit the droplet regulatory function of FUS. The approaches I utilized in this study include the following: (i) fluorescence spectroscopy (ii) a detailed solution nuclear magnetic resonance (NMR) analysis

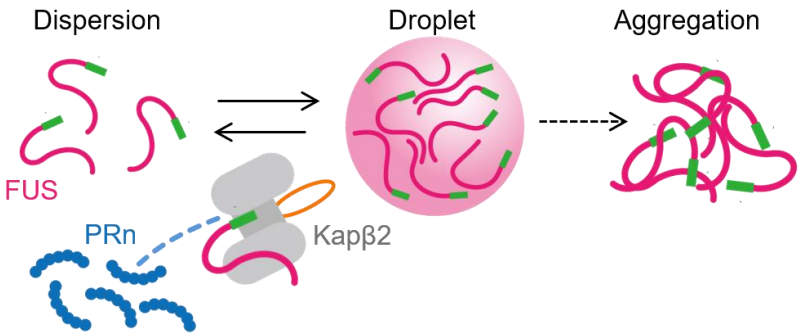


Figure 2.1 The specific binding between PR20 and Kapβ2.

Kapβ2 binds to FUS at the NLS binding site and returns FUS from the droplet to the dispersed state.

2.2 Material and Methods

2.2.1 Constructs, protein expression, and purification

Kapβ2 and M9M were expressed from GST fusion constructs using pGEX6P-1 and pGEX-TEV vectors, respectively. FUS and PRn proteins were expressed from MBP-fusion constructs using the pMAL-TEV vector. All recombinant proteins were expressed individually in BL21(DE3). *Escherichia coli* cells were induced with 0.5 mM ispropyl-β-D-1-thiogalactopyranoside (IPTG) for 12 h at 20 °C. Bacteria expressing Kapβ2 was lysed with a sonicator in buffer containing 50 mM Tris pH 7.5, 200 mM NaCl, 20% (v/v) glycerol, and 2 mM dithiothreitol (DTT). Kapβ2 was purified using GSH Sepharose beads (GS4B, GE Healthcare), cleaved with HRV3C protease, anion exchange chromatography (HiTrap Q HP, GE Healthcare), and gel filtration chromatography (Superdex200 16/60, GE Healthcare) in a buffer containing 20 mM HEPES pH 7.4, 150 mM NaCl, 2 mM DTT, 2 mM $Mg(OAc)_2$, and 10% glycerol. GST:M9M was purified using GSH Sepharose beads (GS4B, GE Healthcare) and gel filtration chromatography (Superdex200 16/60, GE Healthcare). To assemble the Kapβ2–M9M complex, purified Kapβ2 and GST:M9M were mixed, and GST tag was cleaved with Tobacco Etch Virus (TEV) protease. Kapβ2-M9M complex was purified by gel filtration chromatography (Superdex200 16/60, GE Healthcare) and the remaining GST was removed by GSH Sepharose beads (GS4B, GE Healthcare). MBP-FUS and MBP-PRn proteins were lysed in 50 mM Tris pH 7.5, 1.5 M NaCl, 10% glycerol, and 2 mM DTT. MBP-fusion proteins were purified by affinity chromatography, using amylose resin eluted with buffer containing 50 mM Tris pH 7.5, 150 mM NaCl, 10% glycerol, 2 mM DTT, and 20 mM Maltose. It was then further purified for MBP-PRn by cation-exchange chromatography (HiTrap SP HP, GE Healthcare) and gel filtration chromatography (Superdex200 16/60, GE Healthcare). Twenty repeats of PR/GR/PA/GP/GA with an HA epitope tag at the C terminus (PR/GR/PA/GP/GA20-HA) peptides were synthesized by SCRUM, Inc. (Japan).

2.2.2 Size-exclusion chromatography with multi-angle light scattering

SEC-MALS was measured using DAWN HELEOS8+ (Wyatt Technology Corporation) downstream of a Shimadzu liquid chromatography system connected to Superdex200 10/300 GL (GE Healthcare) gel filtration column. The differential refractive index (Shimadzu Corporation) downstream of MALS was used to obtain protein concentration. The column was equilibrated with a running buffer containing 20 mM HEPES pH 7.4, 150 mM NaCl, 2 mM $MgCl_2$, 10% glycerol, and

2 mM β-mercaptoethanol. Flow rate was set to 0.5 mL min^{-1} and 100 µL of the sample was injected. Kapβ2 at a concentration of 28 µM was injected in the absence and presence of 55 µM of MBP-PR18. The data were analyzed with ASTRA version 7.0.1 (Wyatt Technology Corporation).

2.2.3 Expression and purification of isotopically labeled Kapβ2

The Kapβ2 expression plasmid was transformed into BL21(DE3) cells. The protein samples with ^{1}H,^{13}C-labeled methyl groups in deuterium background were prepared as described in a previous study[5]. The cells were grown in medium with $^{15}NH_4Cl$ (2 gL^{-1}, CIL) and $^{2}H_7$-glucose (2 g L^{-1}, CIL) in 99.9% $^{2}H_2O$ (Isotec). The precursors for methyl groups of Ile, Val, and Leu, α-ketobutyric acid (50 mg L^{-1}) and α-ketoisovaleric acid (80 mg L^{-1}), and [$^{13}CH_3$] methionine (50 mg L^{-1}) were added to the culture 1 h before the addition of IPTG. Protein expression was induced by the addition of 0.5 mM IPTG at OD_{600} ~ 0.6, followed by ~16 h of incubation at 25 °C. Cells were collected and re-suspended in the lysis buffer containing 50 mM HEPES pH 7.4, 150 mM NaCl, 20% glycerol, 2 mM DTT, and 2 mM EDTA. Cells were disrupted by a sonicator and centrifuged at 38,900 × *g* for 45 min. Proteins were purified using GS4B resin (GE Healthcare) and eluted with the buffer containing 20 mM HEPES pH 7.4, 20 mM NaCl, 2 mM EDTA, 10% glycerol, and 30 mM GSH. The GST tag was removed by HRV3C protease at 4 °C for ~16 h. The protein was further purified by anion exchange using HiTrap Q FF (GE Healthcare) with the buffer containing 20 mM Imidazole pH 6.5, 20–1000 mM NaCl, 2 mM EDTA, 2 mM DTT, and 20% glycerol, followed by gel filtration using Superdex200 16/60 (GE Healthcare) with the buffer containing 20 mM HEPES pH 7.4, 150 mM NaCl, 2 mM $MgCl_2$, 2 mM DTT, and 2% glycerol. Protein concentration was determined spectrophotometrically at 280 nm using a corresponding extinction coefficient.

2.2.4 NMR experiments

The isotopically labeled Kapβ2 was prepared in the NMR buffer containing 20 mM ^{2}H-Tris pH 7.4, 150 mM NaCl, 2 mM $MgCl_2$, 2 mM DTT, and 2% glycerol, concentrated to 20 µM. NMR experiments were performed on Bruker 600 MHz and 800 MHz NMR operated with Topspin software at 20 °C. Perturbation of the side chain methyl resonances were each monitored using ^{1}H-^{13}C heteronuclear multiple quantum coherence. Spectra were processed using the NMRPipe software[6]. Data analyses were performed on Olivia software (https://github.com/yokochi47/Olivia). The perturbations of the resonances were evaluated based on chemical shift change or intensity change.

Chemical shift perturbations for the methyl groups were calculated based on the following equation:

$$\Delta\delta = \sqrt{\left(\frac{\Delta\delta_{\mathrm{H}}}{\alpha}\right)^2 + \left(\frac{\Delta\delta_{\mathrm{C}}}{\beta}\right)^2} \quad (1)$$

where $\Delta\delta_{\mathrm{H}}$ and $\Delta\delta_{\mathrm{C}}$ are the respective chemical shift changes of ^{1}H and ^{13}C by the addition of the ligand, and α and β are the respective chemical shift distributions of ^{1}H and ^{13}C of methyl groups as reported in the Biological Resonance Data Bank (http://www.bmrb.wisc.edu)[7]. Errors for the intensity ratio were estimated based on the peak intensity and noise level.

In order to investigate the interaction between Kapβ2 and PR poly-dipeptide by NMR, the NMR spectra of the isotopically labeled Kapβ2 in the absence and presence of the synthetic peptide of PR20-HA were measured. To investigate the interaction between Kapβ2 and M9M, the Kapβ2–M9M complex was prepared in the following manner. Purified GST:M9M protein was added to the isotopically labeled Kapβ2, followed by the removal of the GST tag by TEV protease at room temperature for 2 h and then at 4 °C for ~16 h before it was buffer-exchanged into the NMR buffer using Amicon Ultra-4 50k (Merck). To investigate the interaction between Kapβ2 and FUS-ΔNLS by NMR, the NMR spectra of isotopically labeled Kapβ2 in complex with M9M were measured in the absence and presence of FUS-ΔNLS.

2.3 Results

2.3.1 PR poly-dipeptides directly bind Kapβ2.

To investigate the interaction between Kapβ2 and PRn, I used fluorescence spectroscopy. The environment-sensitive fluorophores 5- (dimethylamino) naphthalene-1-sulfonyl (Dnc) was added to the PR20 (Dnc-PR) and the fluorescence of the Dnc group was measured. The affinity of Kapβ2 to Dnc-PR was estimated as $K_D = 1.9\times10^2$ nM (Figure 2.1 A, B). The affinity between Kapβ2 and FUS full length (1-526) determined by isothermal titration calorimetry (ITC) is $K_D = 1.6\times10^2$ nM[8], indicating that PRn binds to Kapβ2 with similar strength as FUS. To analyze the interaction quantitatively, I used size-exclusion chromatography with multi-angle light scattering (SEC-MALS). Maltose binding protein (MBP) was added to PR18 to enable detection by SEC-MALS. The theoretical molecular weight of Kapβ2 is 101 kDa and that of MBP-PR18 is 49 kDa. SEC-MALS showed that MBP-PR18 bound Kapβ2 (Figure 2.1 C). These results indicate that repeat extended PRn directly interact with Kapβ2.

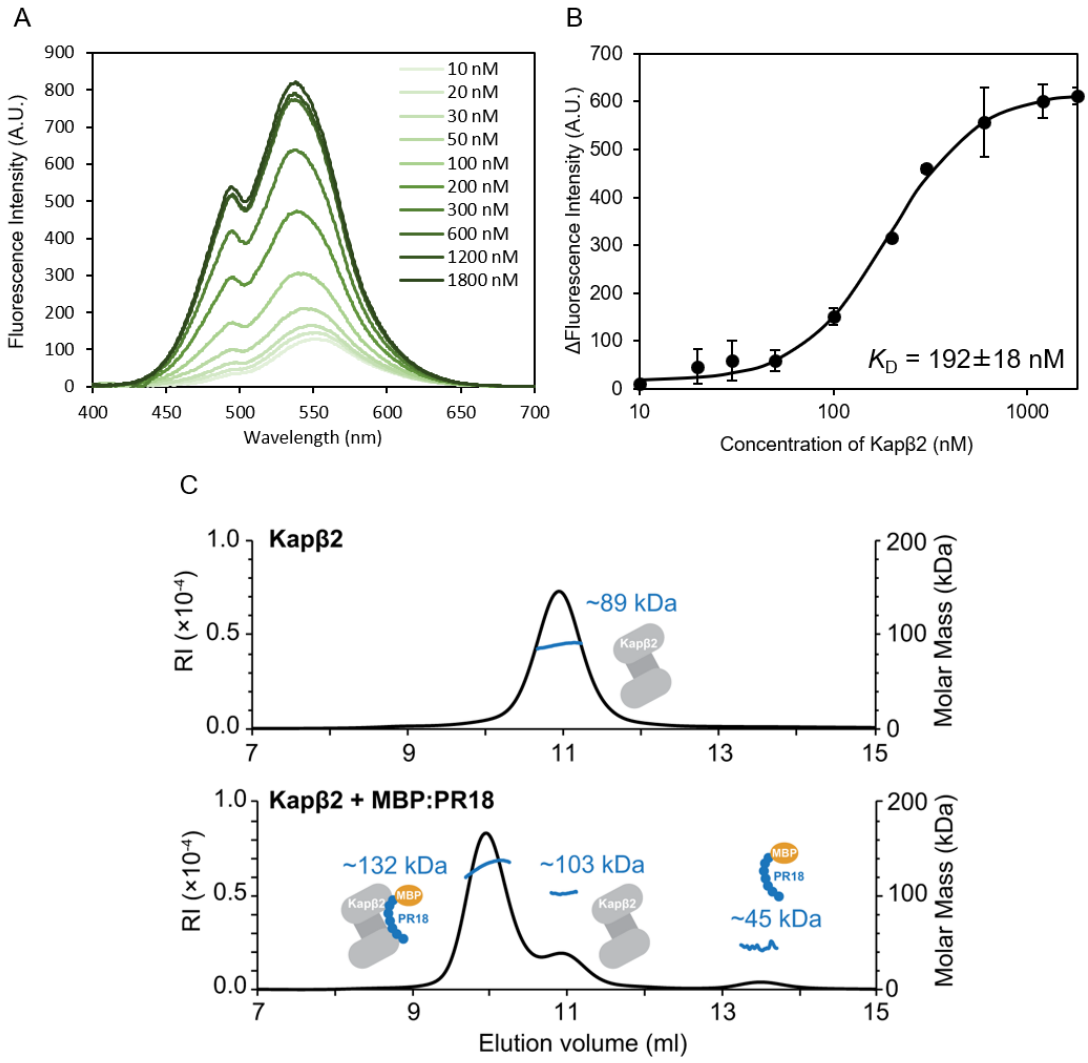


Figure 2.2 The specific binding between PR20 and Kapβ2.

(A) SEC-MALS of Kapβ2 in the absence and presence of MBP-PR18, showing a 1 : 1 complex formation between Kapβ2 and MBP-PR18.

(B) Fluorescence emission spectra of Dnc-PR20 in the presence of Kapβ2. Dnc-PR20 (300 nM) was mixed with 10–1800 nM Kapβ2 and Dnc group was excited with a 340 nm light.

(C) The *in vitro* interaction analysis between Kapβ2 and Dnc-PR peptide. Dnc-PR20 was titrated with Kapβ2 and the change in fluorescence intensity at 550 nm was measured. The concentration with half maximum fluorescence intensity (K_D) was determined from the fitting curve and is indicated by the green triangle. Data are indicated as mean ± SD from three independent experiments.

2.3.2 PR poly-dipeptides target the NLS-binding site of Kapβ2.

I investigated the binding sites of PRn on Kapβ2 using solution NMR. Despite the large size of Kapβ2 (100 kDa)[9], the use of advanced NMR techniques, including methyl-selective isotope labeling (Figure 2.2A) and methyl-transverse relaxation-optimized spectroscopy[5], achieved high-quality NMR spectra (Figure 2.3B). NMR spectra of the isotopically labeled Kapβ2 were acquired in the absence and presence of PR20 (Figure 2.3C). The addition of PR20 induced significant perturbations to several resonances of Kapβ2 (Figure 2.3C), indicating that PRn bind to specific regions of Kapβ2. In order to obtain information about the binding site of PRn, I performed reference NMR experiments. The perturbations induced by PR20 were compared with those induced by M9M, as the latter has previously been shown to bind to the NLS-binding site of Kapβ2[10]. As represented by peaks #63 and #69, several resonances show significant perturbations induced only by PR20, suggesting that these resonances are derived from regions dedicated for the binding of PR20 (Figure 2.4). Likewise, I found several other resonances that were perturbed only by M9M (e.g., Peak #2 and #19), which can be derived from the regions dedicated to the binding of NLS (Figure 2.4 and 2.5). Interestingly, several of the perturbed resonances (e.g., Peak #40 and #150) induced by the addition of PR20 were also perturbed by the binding of M9M, which suggest that these resonances were derived from regions responsible for the binding of both PR20 and NLS (Figure 2.4 and 2.5). Although the perturbations could have been caused by allosteric effect, the shared perturbations indicate that PR20 partially occupy the NLS-binding site of Kapβ2. The additional reference experiment using FUS (1-500, FUS-ΔNLS) show that perturbed resonances induced by FUS-ΔNLS (e.g., Peak #37 and #173) were not perturbed by PR20, indicating that the binding site for FUS-ΔNLS does not overlap with that for PR20 (Figure 2.4 and 2.6). Taken together, the NMR data suggest that the binding site of PRn on Kapβ2 partially overlaps with the Kapβ2 region responsible for the recognition of NLS. As seen in the crystal structure[9], NLS is recognized by the cavity of Kapβ2 in which several methyl-baring residues, including isoleucine residues (I457, I540, I642, I722, I773, I804), leucine residues (L419, L539, L767), valine residues (V643, V724), and methionine residue (M308), are located. Given that M9M binds to the NLS-binding site of Kapβ2, these residues can be attributed to methyl resonances that were perturbed by the binding of M9M (Figure 2.4 and 2.5). Among these residues, L539, I540, I642, and V643 were found to be located at the negatively charged cavity of Kapβ2 (Figure 2.4). Considering the positive charge of PRn, these four residues located at the negatively charged cavity of Kapβ2 may be important for its interaction with PRn.

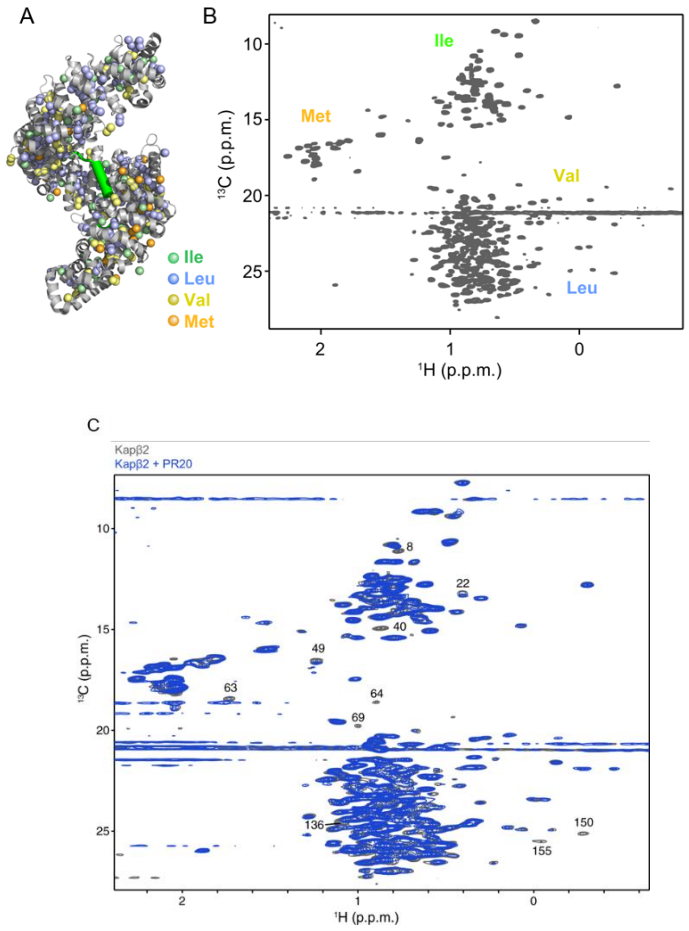
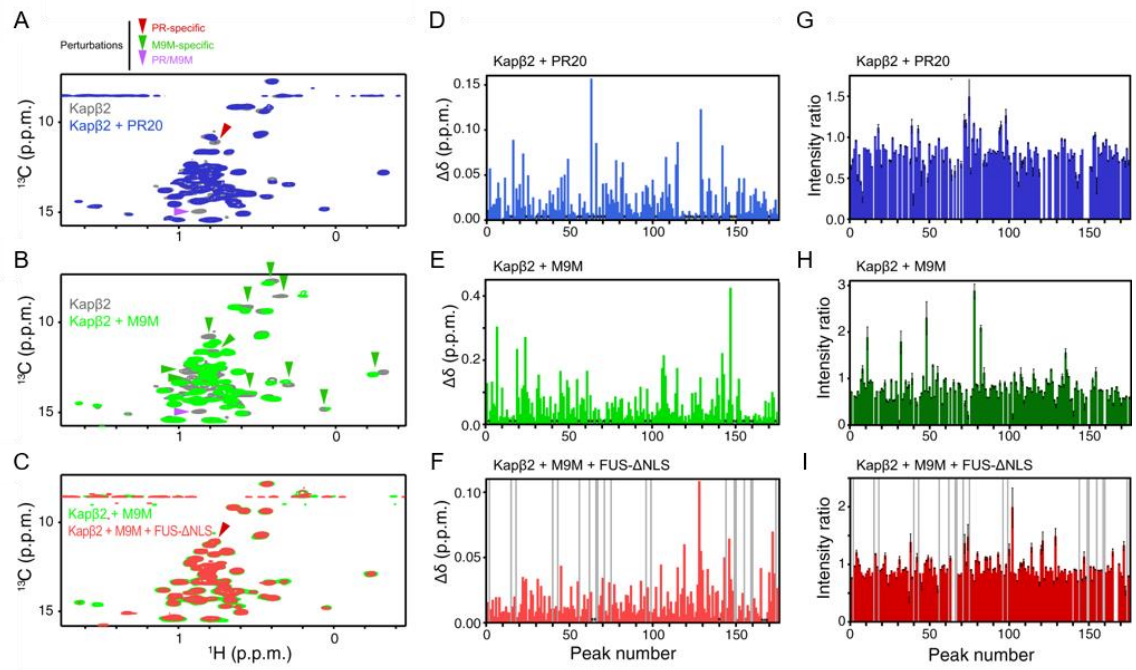


Figure 2.3 Solution NMR of Kapβ2 and interaction between Kapβ2 and PR20.

(A) The crystal structure of Kapβ2 in complex with NLS of FUS. NLS of FUS is presented as a green cylinder and Kapβ2 is shown as a gray ribbon, with spheres representing methyl groups of Ala (pink), Ile (green), Leu (blue), Val (yellow), and Met (orange). Kapβ2 is enriched in methyl-baring residues, indicating that sufficient information can be collected from the ^{1}H-^{13}C-correlated methyl NMR spectra.

(B) ^{1}H-^{13}C-correlated methyl NMR spectra of [U-^{2}H; Ile-δ1-$^{13}CH_3$; Leu, Val-$^{13}CH_3$/$^{12}CH_3$]-labeled Kapβ2. The spectral regions for the resonances of Ile, Met, Val, and Leu methyl groups are labeled.

(C) Interaction between Kapβ2 and PR poly-dipeptide investigated by solution NMR. ^{1}H-^{13}C-correlated methyl NMR spectra of [U-^{2}H; Ile-δ1-$^{13}CH_3$; Leu, Val-$^{13}CH_3$/$^{12}CH_3$]-labeled Kapβ2 in the absence (gray) and presence (blue) of PR20. Significant perturbations were observed for several resonances, indicating that PR poly-dipeptides bind to specific region(s) of Kapβ2. The perturbed resonances are indicated by peak numbers. The peak numbering corresponds to those in Figure 2.5.


A
Perturbations
PR-specific
M9M-specific
PR/M9M
Kapβ2
Kapβ2 + PR20
^{13}C (p.p.m.)
B
Kapβ2
Kapβ2 + M9M
C
Kapβ2 + M9M
Kapβ2 + M9M + FUS-ΔNLS
^{1}H (p.p.m.)
D
Kapβ2 + PR20
Δδ (p.p.m.)
E
Kapβ2 + M9M
F
Kapβ2 + M9M + FUS-ΔNLS
Peak number
G
Kapβ2 + PR20
Intensity ratio
H
Kapβ2 + M9M
I
Kapβ2 + M9M + FUS-ΔNLS
Peak number

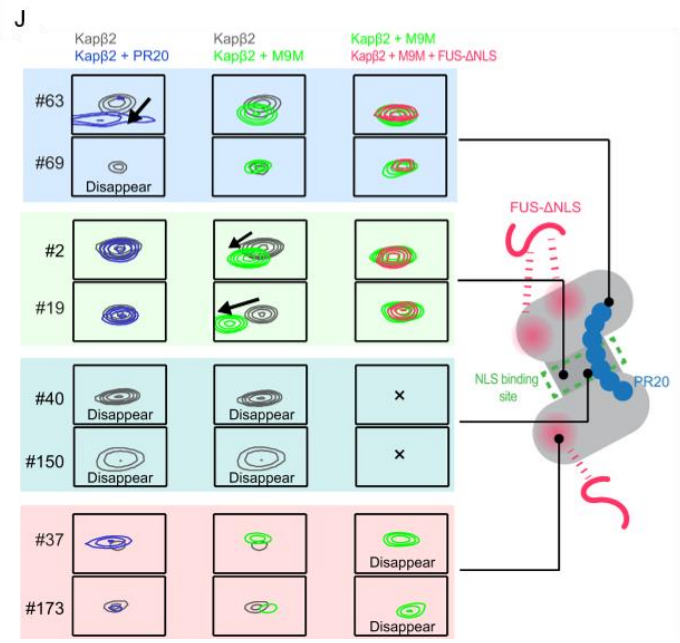

J
Kapβ2
Kapβ2 + PR20
Kapβ2
Kapβ2 + M9M
Kapβ2 + M9M
Kapβ2 + M9M + FUS-ΔNLS
#63
#69
Disappear
#2
#19
#40
#150
Disappear
#37
#173
FUS-ΔNLS
NLS binding site
PR20

Figure 2.4 Solution NMR of Kapβ2 and interaction between Kapβ2 and PR20.

^{1}H-^{13}C-correlated methyl NMR spectra of [U-^{2}H; Ile-δ1-$^{13}CH_3$; Leu, Val-$^{13}CH_3/C^{2}H_3$]-labeled Kapβ2 in the absence (gray) and presence (blue) of PR20 (A), in the absence (gray) and presence (green) of M9M (B), or in the presence of M9M (green) and the presence of M9M and FUS-ΔNLS (magenta) (C). Significant representative perturbations are indicated by arrow heads. Perturbations only seen for PR20 are indicated by red arrow heads. Perturbations only seen for M9M are indicated by green arrow heads. Perturbations common to PR20 and M9M are indicated by purple arrow heads. For clarity, only the region of the Ile methyl resonances is shown. The full-range spectra are shown in Figure 2.5 (for A), 2.5, 2.6 (for B and C). Chemical shift differences (D–F) and intensity ratios (g–i) of the methyl resonances of Kapβ2 by the interaction with PR20 (D, G), M9M (E, H), and FUS-ΔNLS (F, I) are shown. In D–F, resonances that disappeared with the addition of a binding partner are indicated by asterisks. In F and I, resonances that disappeared with the complex formation with M9M are indicated by gray bars. In order to compare the perturbations, peaks on the spectra are labeled in peak numbers as shown in Supplementary Figure 2.5.

(J) Selected views of the interaction between Kapβ2 and PR20/M9M/FUS-ΔNLS. Selected views of the representative resonance from ^{1}H-^{13}C-correlated methyl NMR spectra of [U-^{2}H; Ile-δ1-$^{13}CH_3$; Leu, Val-$^{13}CH_3/C^{2}H_3$]-labeled Kapβ2 in the absence (gray) and presence of PR20 (blue), M9M (green), and M9M and FUS-ΔNLS (magenta). Graphical representation corresponds to the interaction between Kapβ2 and PR20/M9M/FUS-ΔNLS.

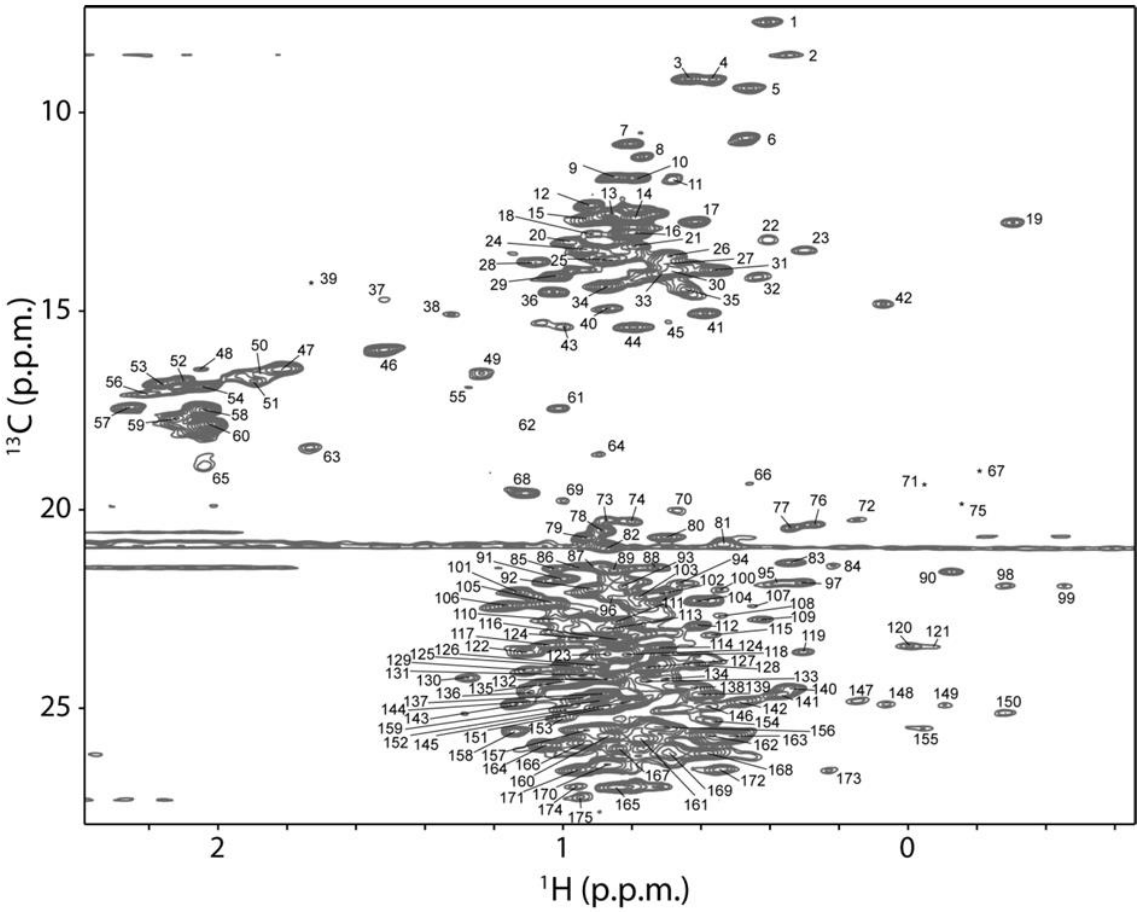


Figure 2.5. Solution NMR of Kapβ2.

^{1}H-^{13}C-correlated methyl NMR spectra of [U^{2}H; Ile-1-^{13}CH3; Leu, Val-$^{13}CH_3$/$^{12}CH_3$]-labeled Kapβ2. The peaks are labeled with peak numbers. Asterisks indicate the positions of the small peaks that were absent in the threshold setting of the figure.

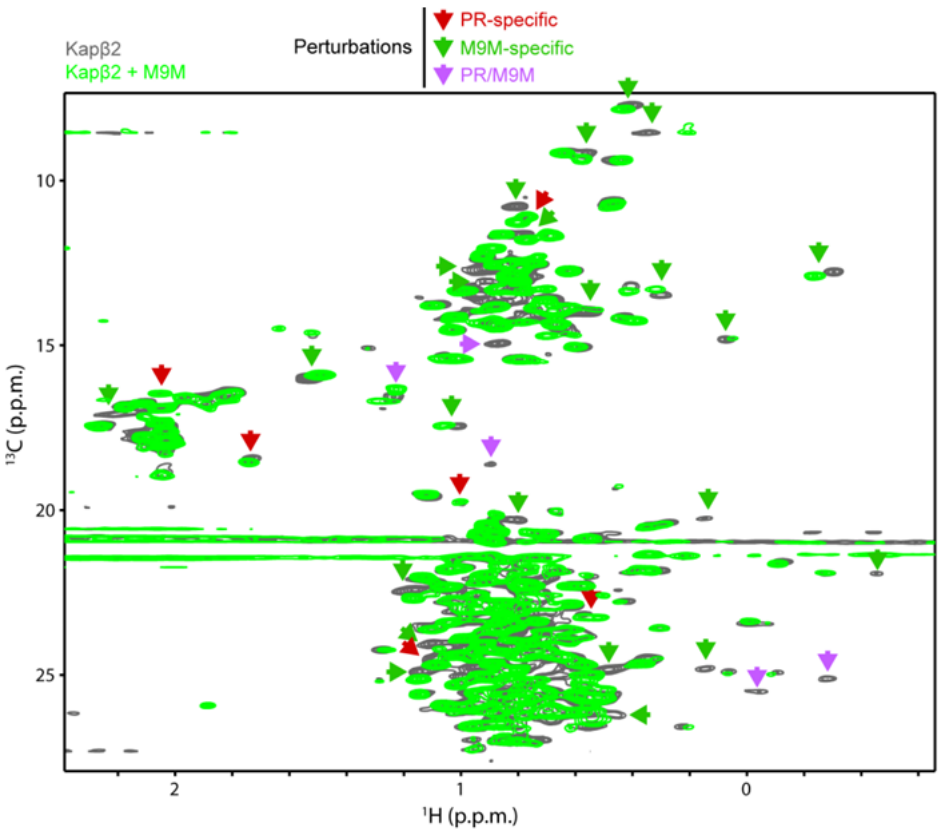


Figure 2.6 Interaction between Kapβ2 and M9M investigated by solution NMR.

^{1}H-^{13}C-correlated methyl NMR spectra of [U^{2}H; Ile-1-^{13}CH3; Leu, Val-$^{13}CH_3$/$^{12}CH_3$]-labeled Kapβ2 (gray) and that in complex with M9M (green). The significant perturbations were observed, which is consistent with the high affinity between Kapβ2 and M9M as reported in a previous study.

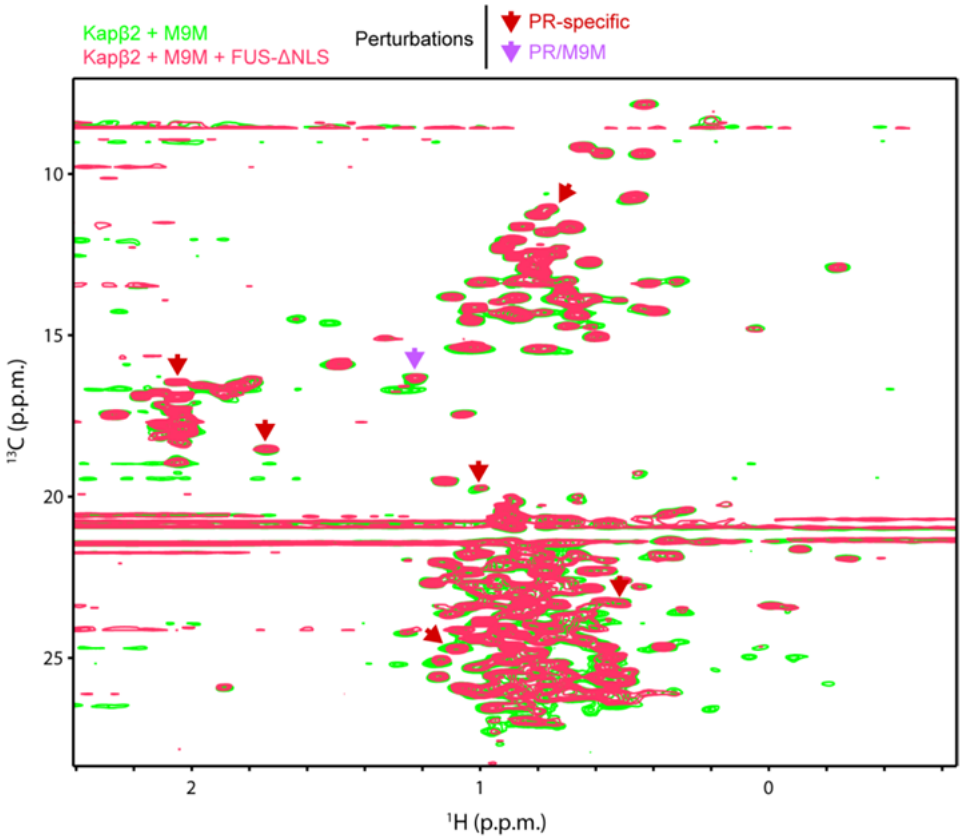


Figure 2.7 Interaction between Kapβ2 and FUS-ΔNLS investigated by solution NMR.

^{1}H-^{13}C-correlated methyl NMR spectra of [U^{2}H; Ile-1-^{13}CH3; Leu, Val-^{13}CH$_3$/^{12}CH$_3$]-labeled Kapβ2 in complex with M9M, in the absence (green) and presence (magenta) of FUS-ΔNLS. Significant representative perturbations are indicated by arrows. Perturbations only seen for PR20 are indicated by red arrows. Perturbations only seen for M9M are indicated by green arrows. Perturbations common to PR20 and M9M are indicated by the purple arrows. Perturbed resonances induced by the addition of FUS-ΔNLS were distinct from those induced by PR20.

2.4 Discussion

I showed that PRn inhibited the Kapβ2 function as a modifier of FUS phase transitions, and that PRn target the NLS-binding site of Kapβ2, which contains a negatively charged region. Recent studies on cellular systems have revealed that arginine-rich poly-dipeptides interact with NIRs resulting in NCT deficit.[11,12] I showed by a series of experiments that PRn compete with NLS for binding to Kapβ2, which may provide a molecular basis for the toxicity of positively charged arginine-rich poly-dipeptides toward NIRs.

Through our experiments, I found that PRn target the NLS-binding site of Kapβ2, which may lead to dysregulation of phase transitions of RBPs containing proline–tyrosine NLS. Besides FUS, many hnRNPs have proline–tyrosine NLS[13] and some of them, including hnRNPA1 and hnRNPA2, are relevant to ALS[14] and interact with TDP43[15] (Figure 2.8). PRn impede Kapβ2 function to modify the phase transition of hnRNPA2 containing proline–tyrosine NLS, which might correlate with pathophysiology. Moreover, it has been shown that decreased nuclear hnRNPA3, which also has proline–tyrosine NLS, leads to an accumulation of repeat RNA and poly-dipeptides in C9-ALS/FTD[16]. If dysfunction of Kapβ2 due to PRn leads to mislocalization of hnRNPA3, toxic repeat RNA and PRn, products of RAN-translation, would increase, resulting in an exacerbation of C9-ALS/FTD pathophysiology.

Given existing evidence[17,18], the binding of PRn to Kapβ2 might also affect the interaction between Kapβ2 and proteins with Arg/Gly/Gly (RGG) regions. A previous NMR study reported that RGG regions of the cold-inducible RBP (CIRBP-RGG) is recognized by Kapβ2[17]. The NMRs of Kapβ2 perturbed by CIRBP-RGG partially overlapped with those perturbed by PRn (Figure 2.9), indicating that PRn interact with the region of Kapβ2 responsible for recognizing RGG regions. The data imply that the binding of PRn to Kapβ2 may also dysregulate the localization of CIRBP and thus affect the stress response system. A recent study reported that RGG regions of FUS are capable of binding the NLS-binding site of Kapβ2[18]. The binding affinities are weaker than the interaction between PR18 and Kapβ2[8]. Thus, PRn likely inhibit the interaction between RGG regions and Kapβ2 in all respects.

Several posttranslational modifications of FUS are known to alter FUS phase transitions, including methylation, phosphorylation, and acetylation[20–22,27]. Among them, the methylation of arginine and the phosphorylation of serine are known to reduce FUS phase transitions[20–22,27]. RGG-rich regions of FUS are extensively methylated by protein arginine methyl transferases (PRMTs)[28].

Arginine methylation of FUS harboring disease-linked mutations decreases binding affinity to Kapβ2 and impairs nuclear transporting activity of Kapβ2[29]. Controversial results have been also reported; inclusions in ALS-FUS patients contain methylated FUS, but no methylated FUS is observed in inclusions of FTD-FUS patients, suggesting the possibility that there remain unknown pathways generating disease-related inclusions[21,29]. I revealed that PR and GR poly-dipeptides impede the chaperone activity of Kapβ2. Gly/Arg-rich sequence tends to be methylated by PRMTs[30]4, leading to an alternation of charge distribution. A recent study revealed that the symmetric dimethylation of GR poly-dipeptides reduces toxicity and correlates with disease duration in C9-ALS/FTD[31]. Therefore, PRn might have a stronger inhibitory effect toward Kapβ2 than GR poly-dipeptides due to the evasion of methylation, although it has yet to be shown whether PRn are methylated in C9-ALS/FTD.

The interaction sites of FUS-NLS for Kapβ2 are Arg-514, His-517, Arg-518, Arg-521, Arg-522, and Pro-525, and these amino acid residues have hydrophobic interactions with Kapβ2.PRn also has hydrophobic interactions, and because it is positively charged, we considered it to have electrostatic interactions.The FUS residue Pro-525 is important because it has numerous contacts with hydrophobic residues of Kapβ2 and also has intramolecular contacts with the adjacent FUS Tyr-526.

NLS-like peptides contain aromatic amino acids, which exhibit hydrophobic interactions with the NLS binding site.Since NMR observes the methyl group signal, the effect of the hydrophobic interaction can be observed, while the effect of the electrostatic interaction is difficult to observe.Therefore, we believe that there is a discrepancy between the NMR results and the fluorescence measurements.

In summary, I showed how PRn affect Kapβ2 function as a phase modifier using multiple biochemical and biophysical methods. In addition to the ability of PRn to stabilize self-association of LC-domains, PRn are found to bind Kapβ2 and impede their function as phase modifiers. Our current study offers additional mechanistic insights on *C9orf72*-related neurodegeneration.

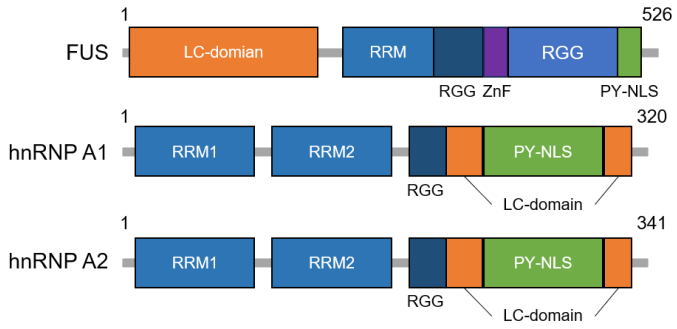


Figure 2.8 Structure of FUS, hnRNP A1 and hnRNP A2

There are LC domain and PY-NLS at RBP.

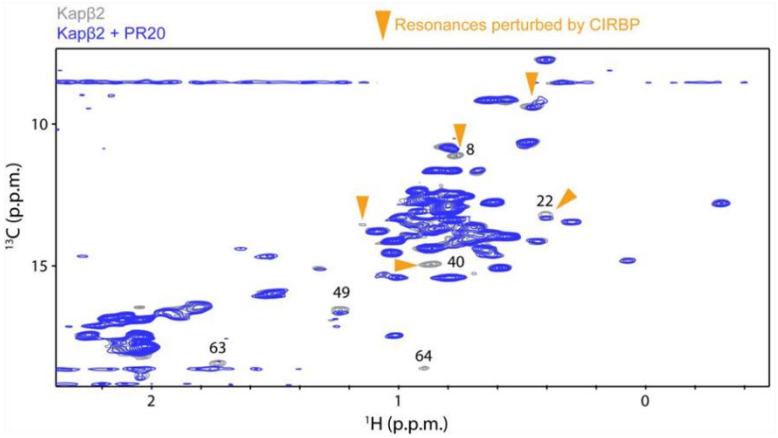


Figure 2.9 Comparison of NMR perturbations of Kapβ2 induced by PR20 and CIRBP-RGG.

^{1}H-^{13}C-correlated methyl NMR spectra of [U-^{2}H; Ile-1-$^{13}CH_3$; Leu, Val$^{13}CH_3$/C$^{2}H_3$]-labeled Kapβ2 in the absence (grey) and presence (blue) of PR20. The resonances that showed significant perturbations by interaction with CIRBP-RGG1 are indicated by arrowheads.

References

1. Mori, K. *et al.* Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. *Acta Neuropathol.* **126**, 881–893 (2013).
2. Zu, T. *et al.* RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. *Proc. Natl. Acad. Sci. U. S. A.* **110**, (2013).
3. Odeh, H. M. & Shorter, J. Arginine-rich dipeptide-repeat proteins as phase disruptors in C9-ALS/FTD. *Emerg. Top. Life Sci.* **4**, 293–305 (2020).
4. Kwon, I. *et al.* Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science (80-.).* **345**, 1139–1145 (2014).
5. Saio, T., Guan, X., Rossi, P., Economou, A. & Kalodimos, C. G. Structural Basis for Protein Antiaggregation Activity of the Trigger Factor Chaperone. *Science (80-.).* **344**, 1250494 (2014).
6. Delaglio, F. *et al.* NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
7. Rosenzweig, R., Farber, P., Velyvis, A., Rennella, E. & Latham, M. P. ClpB N-terminal domain plays a regulatory role in protein disaggregation. (2015) doi:10.1073/pnas.1512783112.
8. Nanaura, H. *et al.* C9orf72-derived arginine-rich poly-dipeptides impede phase modifiers. *Nat. Commun.* 1–12 (2021) doi:10.1038/s41467-021-25560-0.
9. Zhang, Z. C. & Chook, Y. M. Structural and energetic basis of ALS-causing mutations in the atypical proline-tyrosine nuclear localization signal of the Fused in Sarcoma protein (FUS). *Proc. Natl. Acad. Sci. U. S. A.* **109**, 12017–12021 (2012).
10. Cansizoglu, A. E., Lee, B. J., Zhang, Z. C., Fontoura, B. M. A. & Chook, Y. M. Structure-based design of a pathway-specific nuclear import inhibitor. *Nat. Struct. Mol. Biol.* **14**, 452–454 (2007).
11. Hayes, L. R., Duan, L., Bowen, K., Kalab, P. & Rothstein, J. D. C9orf72 arginine-rich dipeptide repeat proteins disrupt karyopherin-mediated nuclear import. *Elife* **9**, (2020).
12. Hutten, S. *et al.* Nuclear Import Receptors Directly Bind to Arginine-Rich Dipeptide Repeat Proteins and Suppress Their Pathological Interactions. *Cell Rep.* **33**, (2020).
13. Kimura, M. *et al.* Extensive cargo identification reveals distinct biological roles of the 12 importin pathways. *Elife* **6**, e21184 (2017).
14. Boeynaems, S., Bogaert, E., Van Damme, P. & Van Den Bosch, L. Inside out: the role of nucleocytoplasmic transport in ALS and FTLD. *Acta Neuropathol.* **132**, 159–173 (2016).
15. Ryan, V. H. *et al.* Mechanistic View of hnRNPA2 Low-Complexity Domain Structure,

Interactions, and Phase Separation Altered by Mutation and Arginine Methylation. *Mol. Cell* **69**, 465-479.e7 (2018).

16. Mori, K. *et al.* Reduced hn RNPA 3 increases C9orf72 repeat RNA levels and dipeptide-repeat protein deposition . *EMBO Rep.* **17**, 1314–1325 (2016).
17. Bourgeois, B. *et al.* Nonclassical nuclear localization signals mediate nuclear import of CIRBP. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 8503–8514 (2020).
18. Gonzalez, A. *et al.* Mechanism of karyopherin-β2 binding and nuclear import of ALS variants FUS(P525L) and FUS(R495X). *Sci. Rep.* **11**, 3754 (2021).
19. Keller, B. A. *et al.* Co-aggregation of RNA binding proteins in ALS spinal motor neurons: evidence of a common pathogenic mechanism. *Acta Neuropathol.* **124**, 733–747 (2012).
20. Murray, D. T. *et al.* Structure of FUS Protein Fibrils and Its Relevance to Self-Assembly and Phase Separation of Low-Complexity Domains. *Cell* **171**, 615-627.e16 (2017).
21. Hofweber, M. *et al.* Phase Separation of FUS Is Suppressed by Its Nuclear Import Receptor and Arginine Methylation. *Cell* **173**, 706-719.e13 (2018).
22. Monahan, Z. *et al.* Phosphorylation of the FUS low-complexity domain disrupts phase separation , aggregation , and toxicity. **36**, 2951–2967 (2017).
23. Cao, Q., Boyer, D. R., Sawaya, M. R., Ge, P. & Eisenberg, D. S. Cryo-EM structures of four polymorphic TDP-43 amyloid cores. *Nat. Struct. Mol. Biol.* **26**, 619–627 (2019).
24. Gremer, L. *et al.* Fibril structure of amyloid-β(1–42) by cryo–electron microscopy. *Science (80-.).* **358**, 116–119 (2017).
25. Tyzack, G. E. *et al.* Widespread FUS mislocalization is a molecular hallmark of amyotrophic lateral sclerosis. *Brain* **142**, 2572–2580 (2019).
26. Bolognesi, B. *et al.* The mutational landscape of a prion-like domain. *Nat. Commun.* **10**, 4162 (2019).
27. Qamar, S. *et al.* FUS Phase Separation Is Modulated by a Molecular Chaperone and Methylation of Arginine Cation-π Interactions. *Cell* **173**, 720-734.e15 (2018).
28. Ong, S.-E., Mittler, G. & Mann, M. Identifying and quantifying in vivo methylation sites by heavy methyl SILAC. *Nat. Methods* **1**, 119–126 (2004).
29. Dormann, D. *et al.* Arginine methylation next to the PY-NLS modulates Transportin binding and nuclear import of FUS. *EMBO J.* **31**, 4258–4275 (2012).
30. Najbauer, J., Johnson, B. A., Young, A. L. & Aswad, D. W. Peptides with sequences similar to

glycine, arginine-rich motifs in proteins interacting with RNA are efficiently recognized by methyltransferase(s) modifying arginine in numerous proteins. *J. Biol. Chem.* **268**, 10501–10509 (1993).

31. Gittings, L. M. *et al.* Symmetric dimethylation of poly-GR correlates with disease duration in C9orf72 FTLD and ALS and reduces poly-GR phase separation and toxicity. *Acta Neuropathol.* **139**, 407–410 (2020).

32. Baborie, A. *et al.* Accumulation of dipeptide repeat proteins predates that of TDP-43 in frontotemporal lobar degeneration associated with hexanucleotide repeat expansions in C9ORF72 gene. *Neuropathol. Appl. Neurobiol.* **41**, 601–612 (2015).

Ⅲ. Conserved loop of a phase modifier endows condensates with fluidity

Abstract

In chapter 3, the effects of excess amounts of PRn on Kapβ2 are examined based on the accumulation of PRn in cells to reproduce the toxic effects of PRn in vivo. When I titrated an excess amount of PRn to Kapβ2, Kapβ2 aggregated, suggesting that an excess amount of PRn alters the physical properties of Kapβ2. Therefore, to investigate the effect of the change in physical properties of Kapβ2 by excess amount of PRn on the FUS droplet release mechanism, I used the fluorescence recovery after optical fading (FRAP) method, which measures the fluidity of the protein. The results showed that the flowability of FUS droplets in the condition with excess amounts of PRn and Kapβ2 decreased, indicating that excess amounts of PRn inhibit the FUS droplet release canism of Kapβ2. Considering that PRn interacts with the NLS binding site, which is rich in negatively charged amino acid residues in Kapβ2 in Chapter 2, PRn may interact with the loop site, which is rich in conserved residues and negatively charged, in addition to the NLS binding site. In order to clarify whether PRn interacts with loop or not, I performed titration experiments between loop and PRn using NMR method. The results confirmed the chemical perturbation, suggesting that loop and PRn interact with each other. Next, to investigate the effect of PRn interaction with loop on the FUS droplet release mechanism of Kapβ2, I compared the flowability of FUS droplets under the condition of adding Kapβ2Δloop, which is deficient in loop relative to FUS, and under the condition of adding excess PRn and Kapβ2. The results showed that the flowability of FUS droplets decreased by the same degree, suggesting that the interaction of PRn with loop inhibits the function of loop, leading to inhibition of the FUS droplet release mechanism of Kapβ2.

3.1Introduction

In chaper2, I have focused on dysregulation of nuclear transport and liquid-liquid phase-separation as possible mechanisms of PRn toxicity, both involved with an importin Karyopherinβ2 (Kapβ2; also known as Transportin 1)[1]. In nuclear transport, Kapβ2 recognizes the proline-tyrosine nuclear localization signal (PY-NLS) of the cargo proteins and enters the nucleus[2,3], followed by binding of the low molecular weight G protein Ran (RanGTP) to release cargo in the nucleus[4,5]. To work as phase modifier, Kapβ2 selectively recognizes RNA-binding proteins (RBPs) containing PY-NLS, such as fused in sarcoma (FUS), and works to resolve phase separation[6]. In this regulation, Kapβ2 recognize multiple regions of RBPs, including PY-NLS, LC region, RGG region, RRM domain, and ZnF domain[1,6]. A recent study revealed that PRn binds to the NLS-binding site of Kapβ2 and inhibits its function as a phase modifier, providing a possible mechanism for *C9orf72*-related ALS pathogenesis[7]. However, the detailed mechanism has remained to be elucidated. Although PRn are known to accumulate in the nucleus and inhibit RNA metabolism and protein expression[8,9], which is considered to be a key factor for understanding progressive pathogenesis of *C9orf72*-related ALS pathogenesis, the mechanism of PRn accumulation in the nucleus and dose-dependent toxicity of PRn in the nucleus remains unknown.

In this chapter, I show that PRn target conserved flexible loop of Kapβ2 to impede its activity as a phase modifier. The binding sites of PRn on Kapβ2 were identified by nuclear magnetic resonance (NMR), showing that PRn binds multiple sites on Kapβ2 including the conserved acidic loop. Refractive index (RI) imaging of the RBP condensates showed PRn binding to the conserved acidic loop alters the physical properties of Kapβ2 and accordingly induces the accumulation and sequestration of Kapβ2 in the RBP condensates. The accumulation of Kapβ2 in the RBP condensates is proposed to impede the nuclear import of RBPs, resulting in the accumulation of RBPs in the cytosol, which is one of the major characteristics seen in the neurons in ALS patients[10–12].

3.2 Material and Methods

3.2.1 Refractive index measurement inside FUS droplet using the holotomography microscope

The LLPS-droplet formation was initiated by mixing 10 μM of MBP-FUS with 4% polyethylene glycol 8000 in a buffer containing 20 mM HEPES-NaOH, pH 7.4, 150 mM NaCl, 10% glycerol, and 2 mM DTT. After 20 min, 10 μM of Alexa594-Kapβ2 and/or 10 or 20 μM of PR_{20} were added to the mixture and incubated for 17 min. The droplets were observed using a holotomography microscope with laser-induced fluorescence system[13] (Tomocube, Inc.). The RI and fluorescence image show the XY cross section of the droplet center. The average RI inside the droplet and their radius were calculated by enclosing the RI image in a circle using TomoStudio version 3.2.8 software (Tomocube, Inc.) and plotted by the Igor Pro version 6.36 software (WaveMetrics). The statistical significance of differences was examined by one-way analysis of variance with Tukey honestly significant difference (HSD) post hoc testing. All statistical tests were performed using KaleidaGraph version 4.5.1 software (Synergy Software) at a significance level of $\alpha = 0.05$.

3.2.2 Fluorescence recovery after photobleaching

Droplet formation was triggered by addition of 4%(w/v) PEG8000 to FUS solutions containing 9 μM MBP-FUS, 1 μM MBP-FUS-ATTO488, 20 mM HEPES-NaOH pH 7.4 at room temperature, 150 mM NaCl, 2 mM DTT, and 10%(w/v) glycerol in the absence or presence of 20 μM PR20, and the FUS solution was well-mixed and immediately dropped onto a glass bottom 10-compartment cell culture slide (CELLview, Cat. #543079, Greiner Bio-One, Kremsmünstern, Austria). FUS droplets were observed with the 473 nm laser line of a confocal microscope (FV1200, Olympus, Tokyo, Japan) that was equipped with a UPLSAPO40X2 objective lens (NA 0.95). A specific spot in FUS droplet was bleached 100% transmission at 20th and 21st frames, and time-lapse images before and after photobleaching were collected (0.5 sec frame rate, 260 frames). Fluorescence intensity of the region of interest was then calculated using FV10-ASW software (Olympus). The images were processed using ImageJ/Fiji[14]and GIMP (GNU Image Manipulation Program) software (http://gimp.org). Fluorescence intensity before photobleaching was set to 100% and immediately after photobleaching was set to 0%, and half times of fluorescence recovery and apparent diffusion coefficients were calculated from the normalized fluorescence intensity using Microsoft Excel (Microsoft, Roselle, IL, USA), based on previous studies[15–17].

3.2.3 Expression and purification of protein sample

Kapβ2 loop (321-371) expression construct was cloned into pET21b vector (Cat. no. 69741-3CN; Novagen, Madison, Wisconsin, USA) and fused to GB1-His^6 tags. The crystal structure of Kapβ2 (PDB code 2Z5J) shows that the distance between Asp321 and Glu371 is ~12 Å. away from the other terminal of loop. The construct of GB1-loop T44C,108C were prepared. The mutant was constructed by site-directed mutagenesis using a PrimeSTAR mutagenesis basal kit (Takara Bio).
The crystal structure of Kapβ2 (PDB code 2Z5J) shows that the distance between Asp321 and Glu371 is ~12 Å. The GB1-loop distance between Thr44 and the terminal of loop was ~12 Å. I introduced an T44C mutation and 108C mutation to GB1-loop. The 15N-labelled GB1-loop protein was overexpressed in E. coli BL21(DE3) cells grown in M9 minimal medium with 15NH4Cl (CIL) as the sole nitrogen source and induced with 0.5 mm IPTG at OD600 ~ 0.6, followed by ~16 h of incubation at 25 °C. Cells were collected and re-suspended in the lysis buffer containing 50mM Tris-HCl pH 8.0, 500 mM NaCl. Cells were disrupted by a sonicator and centrifuged at 15000 rpm for 30 min. Proteins were purified by Ni-NTA affinity chromatography (QIAGEN) and gel filtration chromatography (HiLoad 26/60 Superdex 75 pg, GE Healthcare). Finally, the construct was diluted to a concentration of 10–20 μM and incubated with 1 mM 5,5′-dithiobis(2-nitrobenzoic acid) (DTNB) for 2 h at room temperature to form an intramolecular disulfide bond. After incubation, DTNB was removed by dialysis.

3.2.4 NMR spectroscopy

NMR samples are prepared in 50 mM Tris buffer (pH 8.0), 150 mM NaCl. The protein concentration was 190 mM. NMR experiments were performed on Bruker 600 MHz and 800 MHz NMR operated with Topspin software at 25 °C. The spectra were processed using the NMRPipe program, and data analysis was performed with Olivia (https://github.com/yokochi47/ Olivia).

3.2.5 Dnc-PR

Fluorescence measurements were performed using a Cytation5 Imaging Reader (BioTek Instruments, Inc.). Solutions (60 μL) containing 100 nM TPE-functionalized PEG-*b*-PLLs, proteins (0-500 nM for kapβ2; 0-2000 nM for PRn), 150 mM NaCl, and 2 mM DTT in 20 mM Tris-HCl buffer (pH = 7.5) were prepared in each well of a 384-well NBS™ black microplate (Corning Inc.) using an Andrew+ liquid handling robot (Andrew Alliance SA, Geneva, Switzerland). After incubation (35 °C,

10 min), the fluorescence spectrum ($\lambda_{ex}/\lambda_{em}$ = 330 nm/372–700 nm) or the fluorescence intensity ($\lambda_{ex}/\lambda_{em}$ = 330 nm/480 nm) were recorded at 35 °C.

3.3 Results

3.3.1 Excess DPRs dysregulates Kapβ2 to cause phase-transition of FUS droplet

The abnormal Kapβ2 accumulation into the FUS droplet in the presence of excess PR may affect the physical properties of the droplet. To investigate the internal structure of the FUS droplet in the presence of excess PR, holographic microscopy (optical diffraction tomography: ODT) measurements were performed (Figure 3.1A). ODT observation showed that the internal density of FUS, as indicated by refractive index (RI) value, increased by the addition of excess PR20 only in the presence of Kapβ2 (Figure 3.1B), indicating that the density of FUS droplet increased in the presence of Kapβ2 and excess PR. Fluorescence imaging of the FUS droplet in the presence of Alexa594-labeled Kapβ2 and PR20 showed incorporation of Kapβ2 in the FUS droplet (Figure 3.1B). Interestingly, higher contrast of the fluorescence intensity in the droplet was observed for the condition containing excess amount of PR20, showing that excess PR20 enhanced the incorporation of Kapβ2 in the FUS droplet, resulting higher density of the droplet. Note that in the presence of Kapβ2 with 1 eq PR20, the density of FUS droplet was significantly lower than that in the absence of Kapβ2 (Figure 3.1A). Fluorescence recovery after photobleaching (FRAP) measurements of FUS droplets in the presence and absence of PR and Kapβ2 showed reduced fluorescence recovery in the presence of both 2 eq PR and Kapβ2, indicating reduced fluidity of FUS droplets (Figure 3.2, 3.3). These results suggest that excess PR binds the 2nd site on Kapβ2, which accumulates PR-bound Kapβ2 into the FUS droplet and alters the physical properties of the FUS droplet.

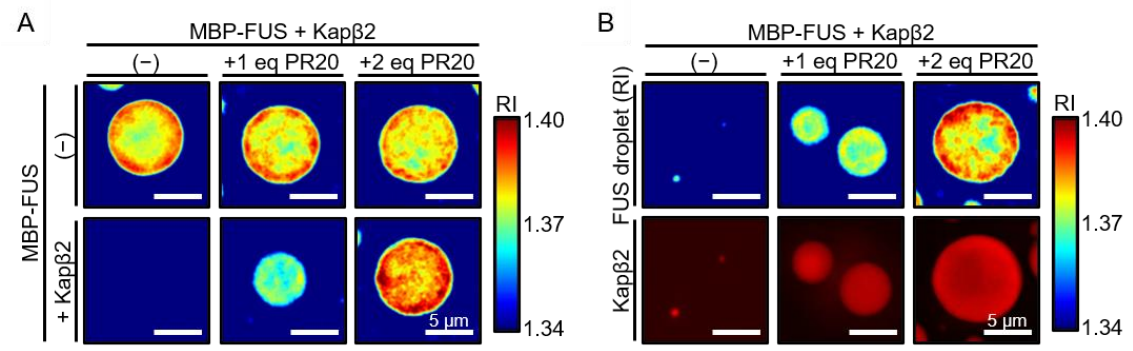


Figure 3.1 Holographic microscopic observation

(A) Refractive index analysis of FUS droplets at different concentrations of PR peptide in the absence and presence of Kapβ2 by holographic microscopy.

(B) Kapβ2 inclusion into FUS droplets with or without PR peptide.

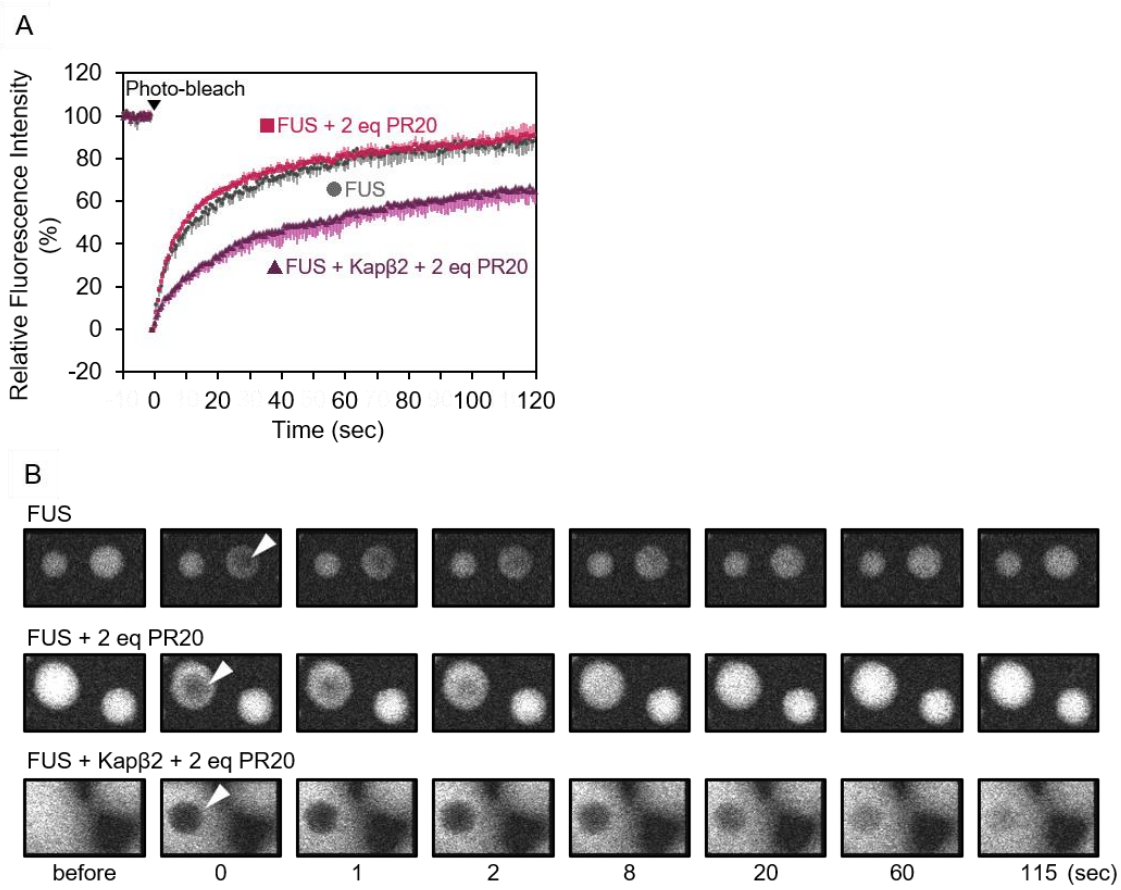


Figure 3.2 FRAP experiment of FUS droplet

FRAP experiment of FUS droplet (A) and confocal fluorescent images of FUS droplet before and after photobleaching (B). Droplet formation was triggered by addition of 4%(w/v) PEG8000 to FUS solutions containing 9 μM MBP-FUS and 1 μM MBP-FUS-ATTO488. Photobleaching was performed in the area indicated by the white arrowheads on the droplets 30 minutes after the droplet formation, and images were taken over time. FUS droplet without Kapβ2 and PR20 (upper panels), FUS droplet with 20 μM PR20 (middle panels), FUS droplet with 10 μM Kapβ2 and 20 μM PR20 (lower panels), respectively.

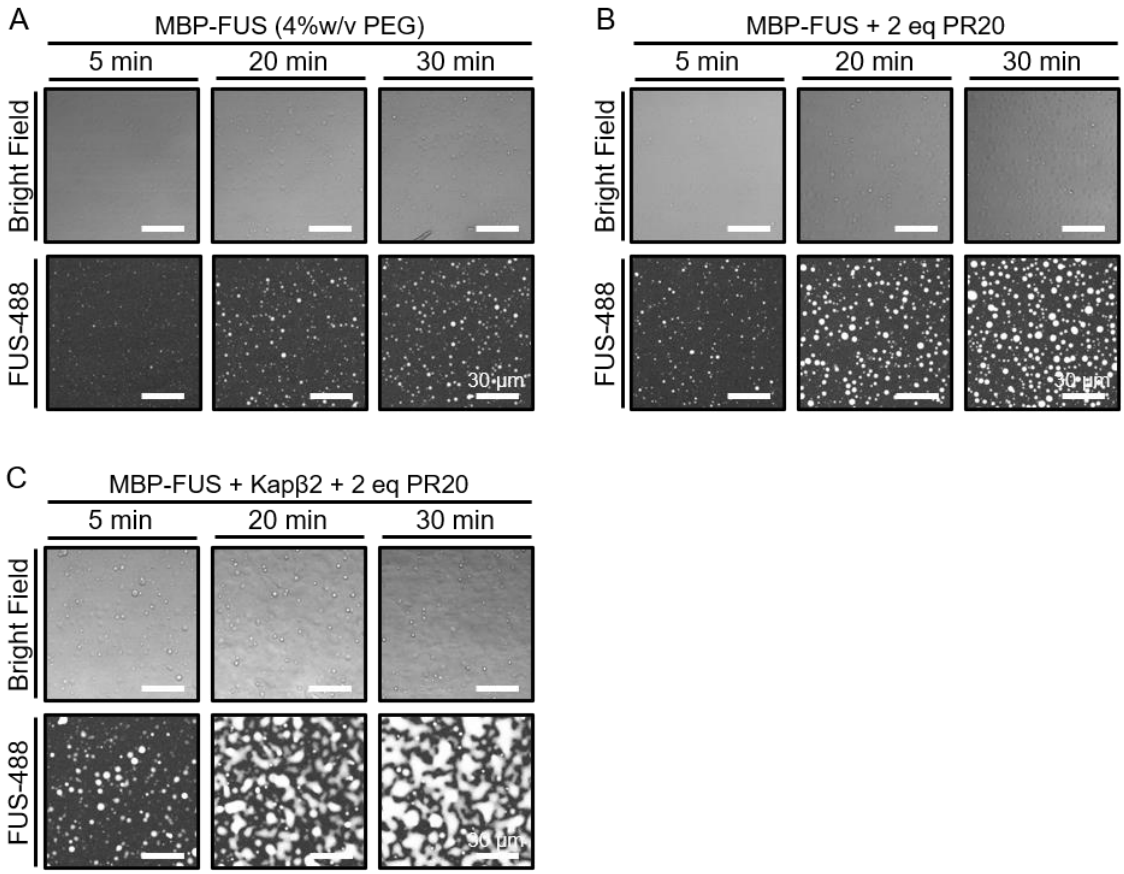


Figure 3.3 FUS droplet shape and aging

(A) Confocal fluorescent images of FUS droplet. Droplet formation was triggered by addition of 4%(w/v) PEG8000 to FUS solutions containing 9 µM MBP-FUS and 1 µM MBP-FUS-ATTO488. Bright Field, differential interference contrast image; FUS-488, fluorescent signal excited with a 473 nm laser. Scale bars: 30 µm.

(B) Confocal fluorescent images of FUS droplet with 20 µM PR20.

(C) Confocal fluorescent images of FUS droplet with 10 µM Kapβ2 and 20 µM PR20.

3.3.2 The conserved acidic loop of Kapβ2 is responsible for weak binding with DPRs

Previously I showed that PRn tightly binds the NLS binding site of Kapβ2 that consists of negatively charged cavity[7]. As for second binding site of PRn on Kapβ2, I suspected negatively-charged regions of Kapβ2, as suggested by chemical tongue analysis exploiting cation-polymers. In addition to the NLS binding site, Kapβ2 has a negatively charged 51-residue loop containing 24 acidic amino acids (Figure 3.4 A, B). The abundance of acidic residues in the loop is highly conserved from yeast to human (Figure 3.4 C). To test whether PR directly interacts with the conserved acidic loop of Kapβ2, I prepared a ^{15}N-labeled "GB1-loop" in which the Kapβ2 loop sequence was fused to the C-terminus of protein GB1 domain (GB1) T44C, 108C mutant and disulfide-bond was formed between the Cys residues at the C-terminus of the loop sequence and T44C, 108C on GB1. ^{1}H^{15}N-HSQC spectra, in the absence and presence of PR20 showed that PR20 induced significant chemical shift changes and peak broadening to the several resonances from the loop (Figure 3.5 A, B), showing the direct interaction between PR20 and the Kapβ2 loop.

The NMR spectrum of Kapβ2 Δloop showed that most of the resonances match to those from Kapβ2 WT, with few exceptions corresponding to the resonances derived from the loop (Figure 3.6), confirming that the overall structure of Kapβ2 is retained in the Δloop mutant.

In addition, the affinity of Kapβ2 Δloop to Dnc-PR was estimated as $K_D = 3.0\times10^2$ nM (Figure 3.6), suggesting that the 1st binding is preserved also in Kapβ2 Δloop. To investigate the interaction between loop and PRn, fluorescence measurements of GB1-loop and PRn were performed. The results show that the affinity is K_D = 666 nM, indicating that loop and PRn interact (Figure 3.7). The affinity of GB1-loop, which is not disulfide-bonded, was K_D = 153 nM, indicating that the root negative charge is larger in the loop structure, and that it binds more tightly to PRn (Figure 3.8).

Collectively, our data showed that PR interacts not only with NLS binding site but also with the conserved acidic loop of Kapβ2 as a 2nd binding site.

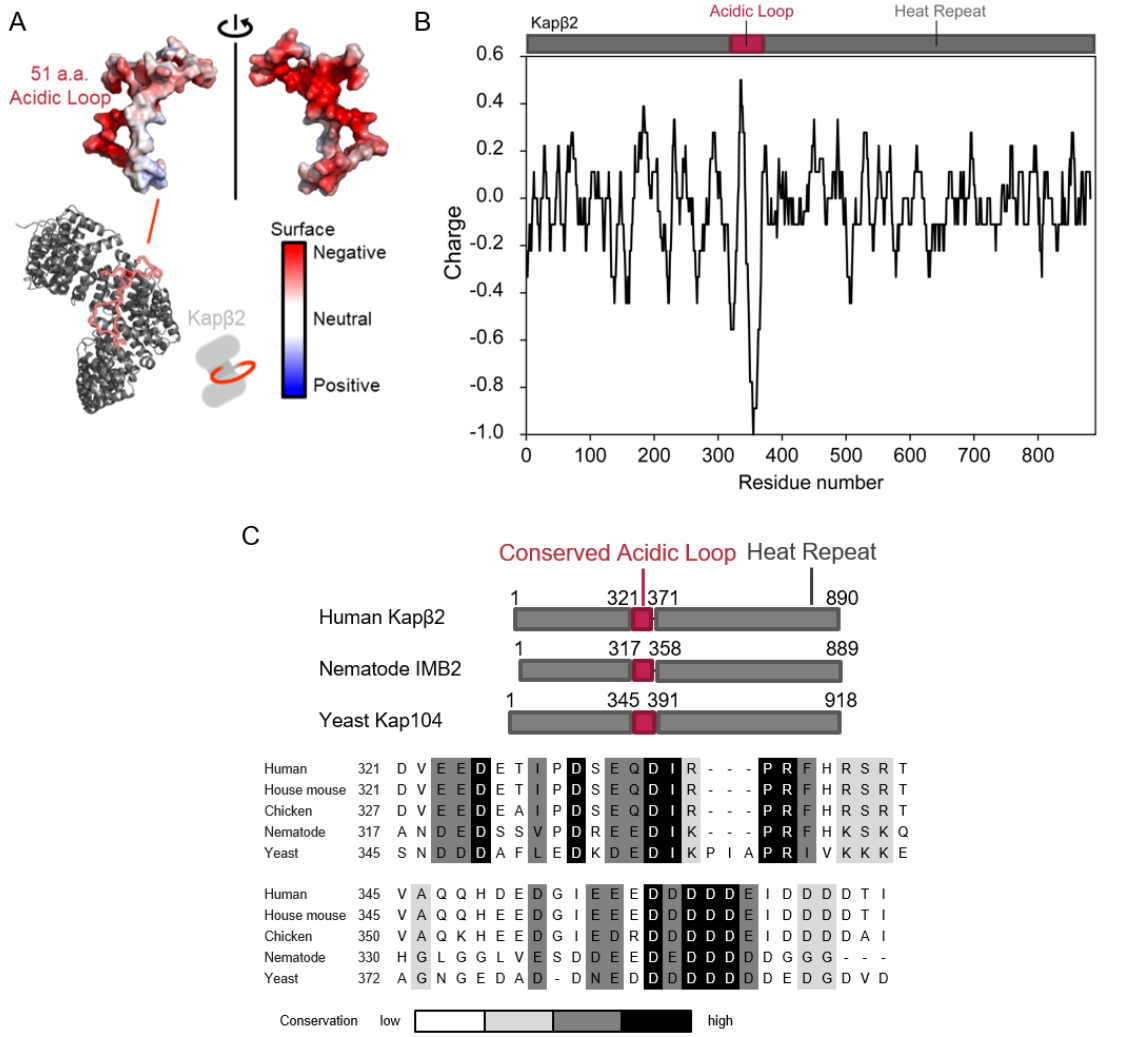


Figure 3.4 The loop region of Kapβ2 constitutes the 2nd binding site of PR peptide.

(A) Crystal structure of Kapβ2 and invisible flexible loop. The crystal structure (PDB code: 2z5j) is shown in gray cartoon. The invisible acidic loop consisting of 51 residues is shown in red dotted lines.

(B) Charge of Kapβ2

(C) The schematic representation of Kapβ2 and its homologs and multiple alignment of their conserved acidic loops. The full length of Kapβ2 homologs is indicated by a gray box and the acidic loops within it by a red box. Alignment of the amino acid sequence of conserved acidic loops was made from the genes of Kapβ2 homologs in each species (human: NP_002261.3, mouse: NP_848831.2, chicken: XP_424806.3, nematode: NP_496987.1, yeast: NP_594385.1). Highly conserved residues are highlighted in gray and black boxes.

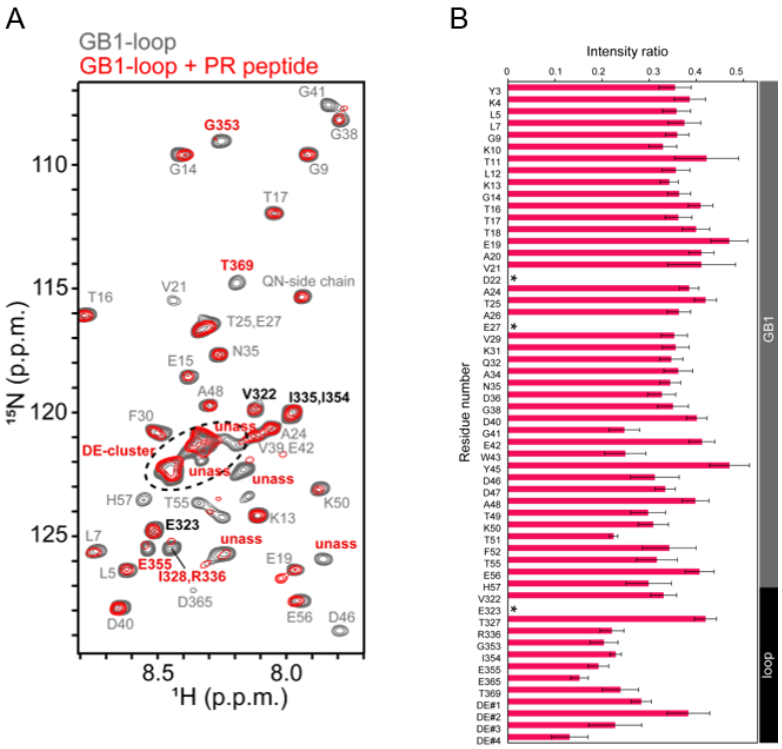


Figure 3.5 The loop region of Kapβ2 constitutes the 2nd binding site of PR peptide.

^{15}N HSQC spectra of Kapβ2 in the absence and presence of PR peptide. Interaction between GB1-loop and PR poly-dipeptide investigated by solution NMR. 1H-15N-HSQC spectra of GB1 (gray), GB1-loop in the absence (blue) and presence (red) of PR20. The resonances showing significant perturbations are indicated by arrow heads. Significant perturbations were observed for several resonances, indicating that PR poly-dipeptides bind to specific regions of loop. The "GB1-loop" consists of GB1 T19C mutant and the conserved acidic loop sequence of Kapβ2, and C-terminal of the loop sequence was linked by using disulfide bond to structurally mimic the acidic loop of Kapβ2 (Extended Data Fig. Sxx 3-1 A).

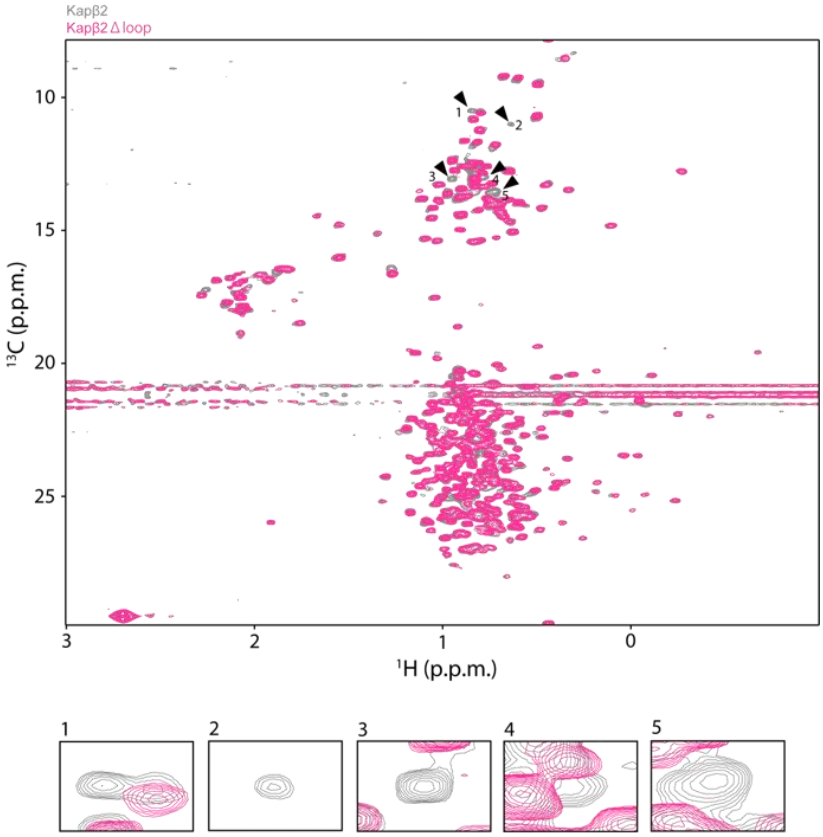


Figure 3.6 Interaction between PR20 and Kapβ2 Δloop.

Interaction between Kapβ2 Δloop and PR poly-dipeptide investigated by solution NMR. ^{1}H-^{13}C-correlated methyl NMR spectra of [U-^{2}H; Ile-δ1-$^{13}CH_3$; Leu, Val-$^{13}CH_3$/$^{12}CH_3$]-labeled Kapβ2 Δloop in the absence (gray) and presence (pink) of PR20. Significant perturbations were observed for several resonances, indicating that PR poly-dipeptides bind to specific region(s) of Kapβ2 Δloop. The perturbed resonances are indicated by numbers and arrow head.

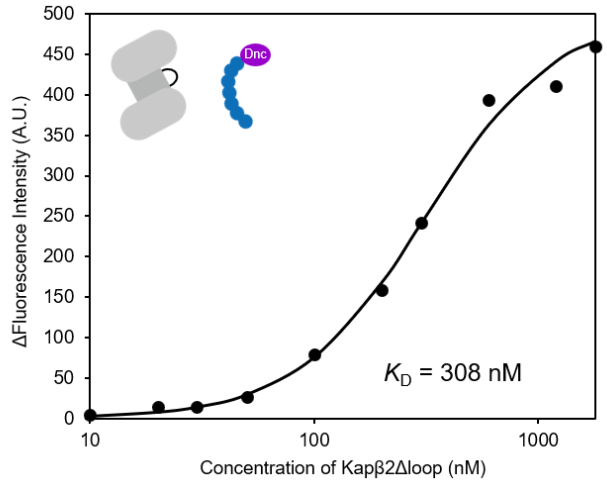


Figure 3.7 The specific binding between PR20 and Kapβ2.

The *in vitro* interaction analysis between Kapβ2 and Dnc-PR peptide. Dnc-PR20 was titrated with Kapβ2 and the change in fluorescence intensity at 550 nm was measured. The concentration with half maximum fluorescence intensity (K_D) was determined from the fitting curve and is indicated by the green triangle. Data are indicated as mean ± SD from three independent experiments.

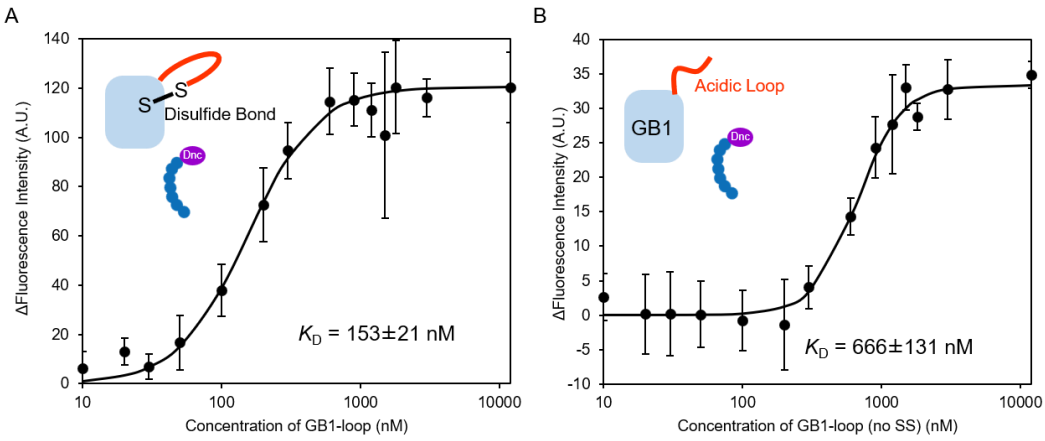


Figure 3.8 The specific binding between PR20 and GB1-loop.

The *in vitro* interaction analysis between GB1-loop and Dnc-PR peptide. Dnc-PR20 was titrated with Kapβ2 and the change in fluorescence intensity at 550 nm was measured. The concentration with half maximum fluorescence intensity (K_D) was determined from the fitting curve and is indicated by the green triangle. Data are indicated as mean ± SD from three independent experiments.

Figure 3.9 The binding between PR20 and GB1.

The *in vitro* interaction analysis between GB1 and Dnc-PR peptide. The concentration with half maximum fluorescence intensity (K_D) was determined from the fitting curve and is indicated by the green triangle. Data are indicated as mean ± SD from three independent experiments.

3.3.3 The conserved acidic loop of Kapβ2 plays essential role in phase modification

Our experiments demonstrate that Kapβ2 dysfunction and retention inside the droplet is caused by the 2nd binding of PR. This fact implies that the conserved acidic loop provides an essential function of Kapβ2 as phase modifier. To investigate the function of the loop, I again performed ODT measurements for Kapβ2 Δloop-Alexa594 (Figure 3.8 A). The results showed that Kapβ2 Δloop was incorporated into the FUS droplet even without PR (Figure 3.8 A, right panels), and the density of the droplet was increased (Figure 3.8 A, +Δloop −PR). FARP measurements showed that the addition of Kapβ2 Δloop to FUS droplet decreased the fluidity inside FUS droplets (Figure 3.8 B, C, D). These observations clearly show that the conserved acidic loop is an essential part for maintaining the phase-separating chaperone ability of Kapβ2.

The fluorescence recovery rate of FUS+Kapβ2Δloop immediately after photofading was faster than that of FUS+Kapβ2+2 eq PR20.This is thought to be a result of the loss of both the NLS binding site and loop function in FUS+Kapβ2+2 eq PR20, whereas the NLS binding site functioned in FUS+Kapβ2Δloop. The above functional analyses of Kapβ2 with excess PR suggest that the 2nd binding of PR following the 1st binding to the NLS binding site is the initiation of Kapβ2 dysfunction

and that the incorporation of the Kapβ2 and multiple-PR complex into the droplet leads to the irreversible breakdown of the Kapβ2 phase-separation regulation.

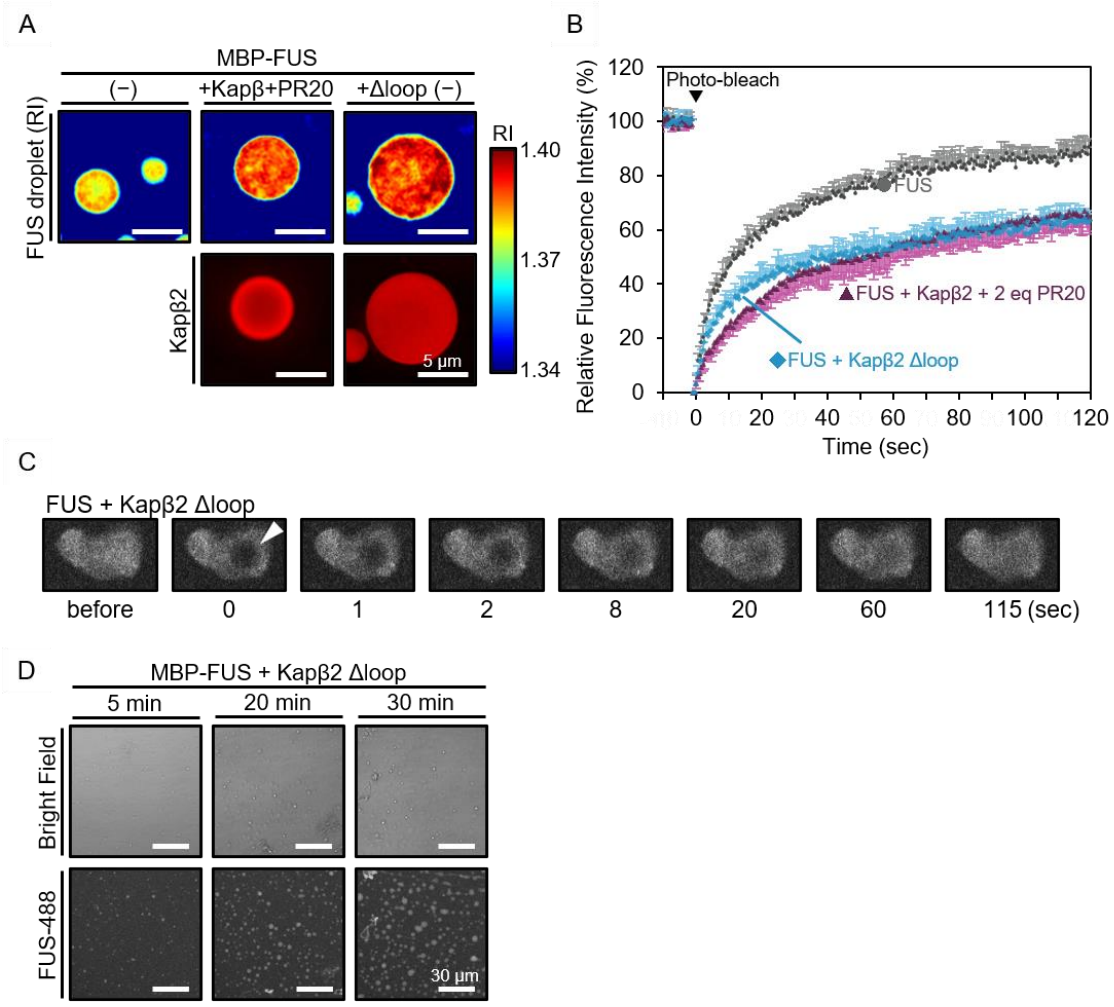


Figure 3.8 FUS droplet shape and aging

(A) Effect of Kapβ2 Δloop mutation on Kapβ2 inclusion into FUS droplet.

(B) FRAP experiment of FUS droplet in the presence of Kapβ2 Δloop. Data are indicated as mean of four−six technical replicates, ±SD. ● gray circle: FUS; ◆ blue diamond: FUS, 1 eq Kapβ2 Δloop.

(C) Confocal fluorescent images of FUS droplet with 10 μM Kapβ2 Δloop before and after photobleaching

(D) Confocal fluorescent images of FUS droplet with 10 μM Kapβ2 Δloop.

3.4 Discussion

PRn produced from *C9orf72* repeat expansion is highly neurotoxic[18], but the mechanism of toxicity was not fully understood. One important pathway of PR toxicity is nuclear transport receptors including Kapβ2. I previously showed that tight binding of PRn to the NLS-binding site of Kapβ2 inhibits interaction with RBPs and accordingly inhibits the activity as a phase modifier [7]. The previous study also suggested that PRn bind to the regions of Kapβ2 other than NLS-binding site, but the detailed interaction modes had remained to be elucidated. Given the fact that recognition of RBPs using multiple sites is important for the activity of Kapβ2 as a phase modifier[6], it has been anticipated that multiple binding of PRn to Kapβ2 is a key to elucidate the inhibitory mechanism of Kapβ2 by PRn and accordingly the toxicity of PRn. This study identified the conserved acidic loop of Kapβ2 as a second binding site for PRn (Figure 3.4, 3.5). I also showed that the excess amount of PRn binds Kapβ2 loop and accordingly sequesters Kapβ2 in the droplet (Figure 3.8), which would alter its availability in the cell. Thus, our results highlight a scheme of PRn toxicity through its interaction with Kapβ2 as follows[18]. In normal condition, Kapβ2 binds RBPs including FUS and transports RBPs into the nucleus, followed by unloading of RBPs through interaction with RanGTP. RBPs in the nucleus are often responsible for consisting of condensates including FUS for nuclear paraspeckle[19]. On the other hand, in the presence of PRn, PRn binds to the NLS-binding site of Kapβ2, and is transported into the nucleus. The binding of RanGTP to the NLS binding site dissociates PRn that accumulate in the nucleus. The accumulated PRn in the nucleus binds not only to the NLS binding site of Kapβ2 but also to the second binding site that includes the conserved acidic loop of Kapβ2 (Figure 3.4). Binding of PRn to the second site alter the physical properties of Kapβ2, resulting the incorporation of Kapβ2 in the RBP condensates (Figure 3.2) and increased density and viscosity of the condensate (Figure 3.8). Oligomerization of PR-bound Kapβ2 as shown by a previous study[18] may also affect to the physical property of Kapβ2. Thus, binding of multiple PRn to Kapβ2 suppresses function as a phase modifier and furthermore sequesters Kapβ2 in the gel-like RBP condensates, which possibly leads to insoluble aggregates. Incorporation of Kapβ2 into the RBP condensates can reduces the amount of active Kapβ2, resulting in the disruption of the nuclear transport pathway, which contributes RBP accumulation in the cytoplasm.

Our results suggest that the conserved acidic loop has the critical role in the second PRn binding to Kapβ2 (Figure 3.4, 3.5), which is consistent with the recent simulation study that showed the binding of PR to multiple regions of Kapβ2 including NLS-binding site, loops, and outer surface

areas[20]. The conserved acidic loop of Kapβ2 has also been shown to be responsible for the recognition of arginine-glycine(-glycine) (RG/RGG)–rich region of CIRBP[21]. Given that RG/RGG-rich region is seen in the Kapβ2-client proteins including FUS, competition between the client proteins and PRn for the conserved acidic loop can also explain the dysregulation of Kapβ2. Although I showed multiple binding of PRn to Kapβ2 using 18 or 20 repeats of the dipeptides, I expect that longer PRn as a single chain can simultaneously bind to multiple regions of Kapβ, including NLS-blinding site and the acidic loop. This explains the higher toxicity of longer repeat expansion at *C9orf72* [22].

I showed that PRn inhibit Kapβ2 at multiple layers, not only occupying the recognition sites for RBPs but also altering the physical properties of Kapβ2 and sequestering Kapβ2 into the RBP condensates. Previous report showing FUS mislocalization to the cytoplasm induced when the nuclear import pathway is compromised[23]. Given the fact that mislocalization of FUS is one of the common aspects of ALS, the PRn-induced Kapβ2 deficiency leading FUS mislocalization and possibly cytosolic FUS deposition should be closely related to the mechanism of ALS pathogenesis.

Droplet formation and resolution are all controlled by Kapβ2-FUS and FUS-FUS interactions. Previous studies using MD simulations have reported that in sufficiently large and dilute systems, high-density regions coalesce to form spherical droplets. On the other hand, in actual cells, FUS is localized in the nucleus, and the expression level and localization of FUS changes in response to stimuli, and changes in its local concentration are also important for control. Since it is difficult to measure the exact concentration in terms of expression, the free energy argument may not be accurate enough.

Given the mechanism of irreversible malfunction of Kapβ2 revealed in this study and the fact that aging is a major risk factor for ALS[24], reduction and shielding of PR may be a strategy to overcome the toxicity expression of PR. JigSAP I used in this study was shown to adsorb PR (Figure 3.5). JigSAP is a new bio-friendly material and has already been used in vivo as an extracellular matrix in mice[25]. In the future, such bio-friendly materials may be one of a strategy for treatment of neurodegenerative diseases.

This study has revealed the mechanism by which PR accumulates in the nucleus and exerts toxic effects. Structural and biochemical approaches have also revealed the disease relevance and functional importance of the conserved acidic loop, and further clarification of the relationship between the mechanism of PR toxicity and FUS mutations is expected to elucidate the mechanism of familial ALS onset and progression. Thus, FUS and Kapβ2 are key players in the background of

familial ALS caused by PR toxicity.

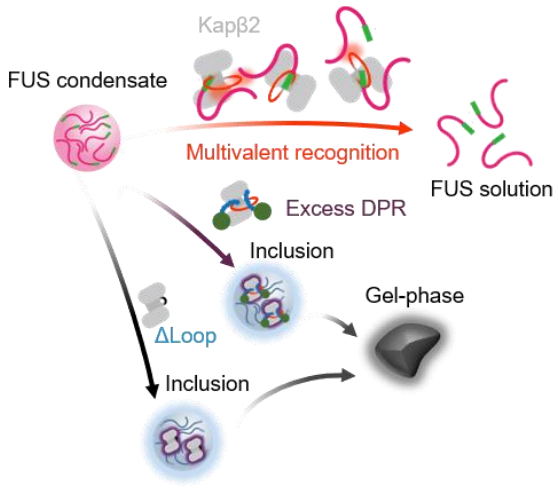


Figure 3.9 Binding of multiple PR peptides disrupts LLPS regulation by Kapβ2 in the nucleus.

Reference

1. Springhower, C. E., Rosen, M. K. & Chook, Y. M. Karyopherins and condensates. *Current Opinion in Cell Biology* vol. 64 112–123 at https://doi.org/10.1016/j.ceb.2020.04.003 (2020).
2. Lee, B. J. *et al.* Rules for Nuclear Localization Sequence Recognition by Karyopherinβ2. *Cell* 126, 543–558 (2006).
3. Zhang, Z. C. & Chook, Y. M. Structural and energetic basis of ALS-causing mutations in the atypical proline-tyrosine nuclear localization signal of the Fused in Sarcoma protein (FUS). *Proc. Natl. Acad. Sci. U. S. A.* 109, 12017–12021 (2012).
4. Weis, K. Regulating access to the genome: Nucleocytoplasmic transport throughout the cell cycle. *Cell* 112, 441–451 (2003).
5. Dirk, G. & Kutay, U. T b c n c. (1999).
6. Yoshizawa, T. *et al.* Nuclear Import Receptor Inhibits Phase Separation of FUS through Binding to Multiple Sites. *Cell* 173, 693-705.e22 (2018).
7. Nanaura, H. *et al.* C9orf72-derived arginine-rich poly-dipeptides impede phase modifiers. *Nat. Commun.* 1–12 (2021) doi:10.1038/s41467-021-25560-0.
8. Kwon, I. *et al.* Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science (80-.).* 345, 1139–1145 (2014).
9. Mizielinska, S. *et al.* C9orf 72repeat expansions causeneurodegeneration inDrosophilathrough arginine-rich proteins. 345, 1192–1195 (2014).
10. Dormann, D. *et al.* ALS-associated fused in sarcoma (FUS) mutations disrupt transportin-mediated nuclear import. *EMBO J.* 29, 2841–2857 (2010).
11. Gal, J. *et al.* Nuclear localization sequence of FUS and induction of stress granules by ALS mutants. *Neurobiol. Aging* 32, 2323.e27-2323.e40 (2011).
12. Ito, D., Seki, M., Tsunoda, Y., Uchiyama, H. & Suzuki, N. Nuclear transport impairment of amyotrophic lateral sclerosis-linked mutations in FUS/TLS. *Ann. Neurol.* 69, 152–162 (2011).
13. Baek, Y. S. & Park, Y. K. Intensity-based holographic imaging via space-domain Kramers–Kronig relations. *Nat. Photonics* 15, 354–360 (2021).
14. Schindelin, J. *et al.* Fiji: An open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682 (2012).
15. Axelrod, D., Koppel, D. E., Schlessinger, J., Elson, E. & Webb, W. W. Mobility measurement by analysis of fluorescence photobleaching recovery kinetics. *Biophys. J.* 16, 1055–1069 (1976).

16. Nehls, S. *et al.* Dynamics and retention of misfolded proteins in native ER membranes. *Nat. Cell Biol.* 2, 288–295 (2000).
17. Kemmer, G. & Keller, S. Nonlinear least-squares data fitting in Excel spreadsheets. *Nat. Protoc.* 5, 267–281 (2010).
18. Hutten, S. *et al.* Nuclear Import Receptors Directly Bind to Arginine-Rich Dipeptide Repeat Proteins and Suppress Their Pathological Interactions. *Cell Rep.* 33, (2020).
19. Ederle, H. & Dormann, D. TDP-43 and FUS en route from the nucleus to the cytoplasm. *FEBS Lett.* 591, 1489–1507 (2017).
20. Jafarinia, H., Van der Giessen, E. & Onck, P. R. Molecular basis of C9orf72 poly-PR interference with the β-karyopherin family of nuclear transport receptors. *Sci. Rep.* 12, 21324 (2022).
21. Bourgeois, B. *et al.* Nonclassical nuclear localization signals mediate nuclear import of CIRBP. *PNAS* 117, (2020).
22. DeJesus-Hernandez, M. *et al.* Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron* 72, 245–256 (2011).
23. Dormann, D. & Haass, C. TDP-43 and FUS: A nuclear affair. *Trends in Neurosciences* vol. 34 339–348 at https://doi.org/10.1016/j.tins.2011.05.002 (2011).
24. Niccoli, T., Partridge, L. & Isaacs, A. M. Ageing as a risk factor for ALS/FTD. *Hum. Mol. Genet.* 26, R105–R113 (2017).
25. Yaguchi, A. *et al.* Efficient protein incorporation and release by a jigsaw-shaped self-assembling peptide hydrogel for injured brain regeneration. *Nat. Commun.* 12, 1–12 (2021).

Ⅳ. Measurement of protein conformation distribution using Single-Pair Small Angle X-Ray Scattering (SAXS)

Abstract

The difficulty in evaluating the conformational distribution of proteins in solution often hinders mechanistic insights. One possible strategy for visualizing conformational distribution is distance distribution measurement by single-pair small-angle X-ray scattering (SAXS), in which the scattering interference from only a specific pair of atoms in the target molecule is extracted. Despite this promising concept, with few applications in synthetic small molecules and DNA, technical difficulties have prevented its application in protein conformational studies. This study used a synthetic tag to fix the lanthanide ion at desired sites on the protein and used single-pair SAXS with contrast matching to evaluate the conformational distribution of the multi-domain protein enzyme MurD. These data highlighted the broad conformational and ligand-driven distribution shifts of MurD in solution. This study proposes an important strategy in solution structural biology that targets dynamic proteins, including multi-domain and intrinsically disordered proteins.

4.1 Introduction

In Chapter 4, based on the result in Chapter 3 that PRn reduced the fluidity of FUS droplets, I focus on the internal structure of FUS droplets and clarify the structural changes from droplet formation to aggregation in order to clarify the mechanism of FUS aggregation from the perspective of FUS structure. Since FUS is a intrinsically disordered protein (IDP) that does not have a specific structure, it does not have a specific structure when it forms a droplet from a dispersed state, and is considered to have a conformational distribution with multiple structures. Although this conformational polymorphism is thought to change when FUS aggregates, no method has been established to clarify the conformational polymorphism of IDP in solution.

Conformational distribution is the key to understanding how proteins, especially IDP, function in solution. IDP often regulate their activity by changing the population of multiple conformational states with varying domain orientations, contingent on ligand binding, post-translational modifications, and environmental changes[1–4]. However, evaluating the conformational distribution of IDP in solution is not straightforward[5,6]. Although recent technical developments in cryo-EM have expanded the application of the method to the structural determination of proteins with conformational variations, its application to those with continuous variations remains challenging. Additionally, the impact of freezing proteins for detection needs to be considered[7]. Solution nuclear magnetic resonance (NMR) is useful for the structural study of proteins in solution, especially with the use of paramagnetic probes[8,9]. On the other hand, NMR often suffers from resonance averaging because of the exchange among multiple conformational states, making data interpretation difficult. Small-angle X-ray scattering (SAXS) is one of the few methods for quantitatively assessing the conformational polydispersity of macromolecules, including IDP and multi-domain proteins with flexible linkers[10]. However, the structural information obtained from typical SAXS analysis is of low resolution because each X-ray scattering between all the atom pairs in the protein sums up to make a single SAXS curve; accordingly, the information is averaged. One way to obtain the conformational distribution of a protein is to exploit atoms with higher electron densities than those consisting of proteins. Scattering interference between heavy atoms on the protein can be extracted by contrast-matching[11] or reference subtraction[12,13] to determine the distance distribution between heavy atoms. If heavy atoms are fixed at specific sites on a protein, the distance distribution of the heavy atoms reflects the conformational distribution of the protein. Previous studies using SAXS have demonstrated the measurement of the distance distribution for colloidal gold particles fixed at both ends of double-stranded DNA[12,13]. A

limitation was the size of the gold particle, 14 Å. The gold particle size is adequate for measuring the length of 10~35 bp DNA, with distance ranging from 50 to 140 Å, but is often too large to assess conformation changes in proteins. A demonstration using smaller particles was reported using two iodine atoms connected by a PEG chain, in which the distance distribution for a single pair of iodine atoms was obtained by reference subtraction[14]. However, single-pair SAXS has not yet been used in protein conformational studies. Although SAXS distance measurements have been demonstrated for a calcium-binding protein substituted with four lead ions, the distance distribution was not obtained in this study presumably because of the complexity of the analysis owing to the four metals, resulting in six possible pair distances[11]. Furthermore, its application is limited to metalloproteins; however, a new strategy applicable to non-metalloproteins is anticipated.

In this study, I evaluated the conformational distribution of a multi-domain protein by single-pair SAXS using Lu^{3+} attached to two specific sites on the protein using a tag. Lu^{3+} has the highest number of electrons among the lanthanide ions; thus, providing a larger X-ray scattering intensity compared to atoms consisting of proteins. Another advantage of lanthanide ions is that the techniques used to fix the ion to a specific site of the target protein have been extensively studied, especially in the field of biomolecular NMR[8]. Among a number of lanthanide-binding tags, I used Caged Lanthanide NMR Probe 5 (CLaNP-5)[15,16] attached to the protein through two arms to reduce the mobility of the tag with respect to the protein; accordingly, this provides more defined and reliable information about domain conformational states. X-ray scattering arising from Lu^{3+} was observed under contrast-matched conditions using 65% aqueous sucrose buffer, whose electron density matched that of the protein, resulting in the suppression of scattering from protein atoms at low angles[11,17].

To demonstrate the conformational analysis of a multi-domain protein using single-pair SAXS, I used a multi-domain protein, MurD. MurD consists of three domains and is one of the ATP-driven Mur ligases responsible for peptidoglycan biosynthesis. Conformational states and ligand-driven conformational changes in MurD have been well-characterized in previous studies using X-ray crystallography, NMR, molecular dynamics (MD) simulations, and electron spin resonance (ESR)[18–23]. This makes MurD one of the best targets to demonstrate the application of single-pair SAXS. First, I evaluated the integrity of the strategy by distance distribution measurement for the two lutetium ions (both fixed on the same domain of MurD) and confirmed that the result corresponded to the distance estimated from previous studies. Next, I conducted a conformational study of the full-length MurD with lutetium ions in domains 2 and 3. The data showed that MurD exists in multiple open

conformations. On the other hand, inhibitor binding eliminates the conformations toward the closed conformation, highlighting the utility of single-pair SAXS for protein conformational studies in solution.

4.2 Material and Methods

4.2.1 Preparation of CLaNP-5

CLaNP-5 was synthesized, purified, and chelated with Lu^{3+} as described in previous reports[9,16,24]..

4.2.2 Protein sample preparation

Escherichia coli full-length MurD (1–437) and domains 1–2 (1–302) were cloned into pGBHPS[25], expressed in *E. coli* strain BL21 (DE3), and purified as described in a previous report[9]. For SAXS measurements, MurD domains 1–2 M145C/D148C/C151A/E260C/K262C ($D12_{145\text{-}260}$) and MurD E260C/K262C/N360C/D362C ($MurD_{260–360}$) were prepared using a previously described procedure[9]. Note that $D12_{145\text{-}260}$ contains a mutation in C151A to avoid intramolecular disulfide bond formation with D148C. The CLaNP-5 tag was attached to the protein by mixing the protein and tag at a 1:2.2 ratio for 15 min on ice, followed by gel filtration. The data showed that the tags were attached to a specific position of MurD with high efficiency. Although MurD has seven cysteine residues (e.g., C20, C99, C151, C208, C227, C368, and C413), all cysteine residues have a thiol group buried in the protein and do not react with the CLaNP-5 tag[9]. To replace the protein buffer with 65% (w/v) sucrose buffer, MurD and 65% sucrose buffer were mixed in a 1:1 volume and dialyzed overnight[11]. For more accurate contrast match, 30 uL of protein solution and 1.5 mL of 65% sucrose buffer were added to a sitting drop plate (well capacity: 1.5 mL, post capacity: 40 µl, plate dimension: approx. 15.0 cm × 10.8 cm × 2.2 cm, HAMPTON RESEACH., Ltd., Osaka, Japan) and incubated overnight.

4.2.3 Flow-SAXS measurement

Contrast-matched SAXS for $D12_{145\text{-}260}$ was recorded at the beamline BL-10C of the Photon Factory (PF; Tsukuba, Japan). The X-ray wavelength was 1.1 Å and the sample-detector distance was 1.0 m. Data were collected using a PILATUS 300 K detector at twenty 30-sec exposures per flow. The data from six flows were collected and averaged. A silver behenate standard was used to locate the beam center and calibrate the scattering angle values. Data reduction was performed using the SAngler software[26]. The measurements were performed in a 1.25 mm path length flow cell with fused quartz windows (2.5 mm × 6 mm × 0.02 mm, Unisoku Co., Ltd., Osaka, Japan). Silicone rubber was used to hold the windows in place. The flow rate was set to 0.33 µl min^{-1} to reduce irradiation damage. The cell temperature was set to 20°C. Buffer scattering was recorded before and after each sample. After

the sample measurements, the cells were washed with a detergent solution and water.

As a reference, SAXS for D12 without lanthanide ion in 65% sucrose solution was recorded at the beamline BL-10C of the Photon Factory (PF; Tsukuba, Japan). The X-ray wavelength was 1.1 Å and the sample-detector distance was 1.0 m. Data were collected using a PILATUS3 2M detector at twenty 30-sec exposures per flow. The data from twelve flows were collected and averaged. A silver behenate standard was used to locate the beam center and calibrate the scattering angle values. Data reduction was performed using the SAngler software[26]. The measurements were performed in a 1.25 mm path length flow cell with fused quartz windows (2.5 mm × 6 mm × 0.02 mm, Unisoku Co., Ltd., Osaka, Japan). Silicone rubber was used to hold the windows in place. The flow rate was set to 0.33 µl min^{-1} to reduce irradiation damage. The cell temperature was set to 20°C. Buffer scattering was recorded before and after each sample. After the sample measurements, the cells were washed with a detergent solution and water.

Contrast matched SAXS for $MurD_{260–360}$ was recorded at beamline BL-6A of the Photon Factory (PF; Tsukuba, Japan). The X-ray wavelength was 1.5 Å and the sample-detector distance was 1.0 m. Data were collected using a PILATUS3 1M detector with twenty 20-sec exposures per flow. The data from five flows were collected and averaged. A silver behenate standard was used to locate the beam center and calibrate the scattering angle values. Data reduction was performed using the SAngler software[26]. The measurements were performed in a 1.25 mm path length flow cell with fused quartz windows (2.5 mm × 6 mm × 0.02 mm, Unisoku Co., Ltd., Osaka, Japan). Silicone rubber was used to hold the windows in place. The flow rate was set to 6 µl min^{-1} and measured at 0°C to reduce irradiation damage. Buffer scattering was recorded before and after each sample. After the sample measurements, the cells were washed with a detergent solution and water.

4.2.4 SEC-SAXS measurement

SEC-SAXS for MurD and D12 was recorded at the beamline BL-10C of the Photon Factory (PF; Tsukuba, Japan). The X-ray wavelength was 1.1 Å and the sample-detector distance was 1.0 m. Data were collected using a PILATUS 2M detector at twenty 30-sec exposures per sample. A silver behenate standard was used to locate the beam center and calibrate the scattering angle values. Data reduction was performed using the SAngler software[26]. The measurements were performed in a 1.25-mm path length flow cell with fused quartz windows (2.5 mm × 6 mm × 0.02 mm, Unisoku Co., Ltd., Osaka, Japan). Silicone rubber was used to hold the windows in place. The flow rate was set at 0.05

ml min^{-1}. The cell temperature was set to 20°C.

4.2.5 Distance distribution analysis

To extract distance distribution information from the SAXS data, the data from contrast-matched SAXS were fitted with *Equation 1* using a modified version of the MATLAB program used in a previous study[11]. The original program was used to study gold nanocrystal-labelled DNA fragments[12,13]. In the fitting of MurD data, a linear baseline correction was added. To consider the reliable range of the distance information, the SAXS curves were regenerated from the *P*(r) function. The *P*(r) function was assumed to be normal distribution. The peak top corresponded to mean (μ), and standard deviation (σ) was obtained from FWHM=2.35σ. The SAXS curve was estimated for varying μ, with the range up to ±5 Å to search for the fitted region.

4.2.6 SEC-MALS measurement

SEC–MALS was performed using a DAWN HELEOS8+ (Wyatt Technology Corporation, Santa Barbara, CA, USA), high-performance liquid chromatography pump LC-20AD (Shimadzu, Kyoto, Japan), refractive index detector RID-20A (Shimadzu), and UV–vis detector SPD-20A (Shimadzu), located downstream of the Shimadzu liquid chromatography system connected to a PROTEIN KW-803 gel filtration column (Cat. no. F6989103; Shodex, Tokyo, Japan). Differential RI (Shimadzu) downstream of MALS was used to determine protein concentrations. The running buffer used contained 20 mM Tris-HCl (pH 7.2) and 100 mM NaCl for D12, and 20 mM Tris-HCl (pH7.2) and 200 mM NaCl for MurD. Approximately 100 μL of the sample was injected at a flow rate of 1.0 mL min^{-1}. Data were analyzed using ASTRA version 7.0.1 (Wyatt Technology Corporation). Molar mass analysis was also performed over half the width of the UV peak top height.

4.2.7 Circular dichroism spectrometer

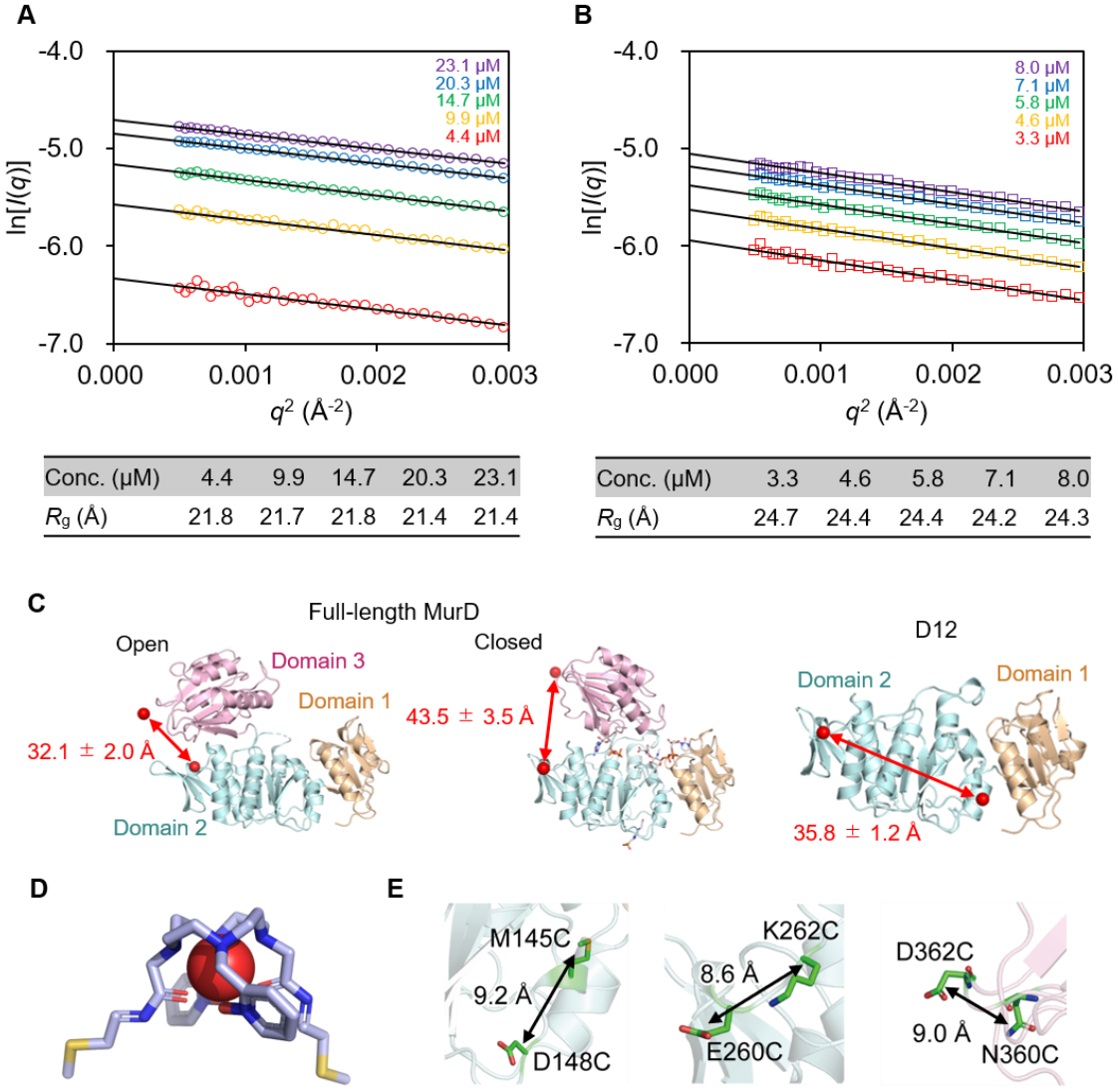
The CD spectra were recorded using a JASCO J-1500 CD spectrometer (Tokyo, Japan) with 0.1 mm path length cuvettes at 25°C in 20 mM Tris-HCl (pH 7.2), 200 mM NaCl, and 65% (w/v) sucrose buffer. Each spectrum represents an integration of four consecutive scans from 190 to 260 nm at 1.0 nm intervals, with a scan speed of 20 nm/min. The concentrations of MurD and MurD in 65% sucrose buffer were 34 μM and 27 μM, respectively.

4.3 Results

4.3.1 MurD exists in a dispersed monomeric state

In this study, full-length MurD and MurD domain 1-2 (D12) were used to demonstrate single-pair SAXS. Prior to single-pair SAXS, MurD and D12 were subjected to size exclusion chromatography (SEC)-SAXS to test the oligomeric state and dispersity in solution. SAXS measurements followed by Guinier analysis on D12 and MurD resulted in an R_g of ~21 Å and 23~24 Å, respectively (Figure 4.1A, B and 4.2). The data are consistent with the predicted R_g values of 20.8 Å for D12 and 24.0 Å for MurD as obtained from the crystal structure (PDB ID:1e0d) of the MurD monomer using HullRad[27]. The region comprising residues 1–302 was used to predict the R_g value of D12. The reasonable match between the experimental and predicted R_g values suggests that D12 and MurD exist as stable monomers under SAXS conditions. The monomeric states of D12 and MurD were also corroborated by size-exclusion chromatography with multi-angle light scattering (SEC-MALS), which showed molar mass values consistent with D12 and MurD existing as monomers (Figure 4.3). Thus, the data showed that D12 and MurD exist as monomers in solution and are suitable for single-pair SAXS.

To investigate the distance distributions of the multi-domain protein MurD, I measured the distance between two lutetium ions fixed on domains 2 and 3 using SAXS (Figure 4.1C). Following the procedures described in previous reports[9,16] pairs of amino acid residues whose C_β atoms are located in a distance of 8–10 Å were selected and mutated to cysteine residues for attachment of CLaNP-5 via disulfide bonds (Figure 4.1D and 4.1E). Two MurD variants were constructed for SAXS measurements: MurD domain 1–2 M145C/D148C/C151A/E260C/K262C ($D12_{145\text{-}260}$) and full-length MurD E260C/K262C/N360C/D362C ($MurD_{260–360}$) (Figure 4.1C-E). Because $D12_{145\text{-}260}$ has two Lu^{3+} ions in domain 2, the distance distributions for $D12_{145\text{-}260}$ are expected to reflect the local conformational variation of the tag and the Cys residues bridged to the tag. On the other hand, the distance distribution for $MurD_{260–360}$, with Lu^{3+} ions in domains 2 and 3 was expected to reflect the conformational variations of domain 3 with respect to domain 2, in addition to local conformational variations (Figure 4.1C).



Conc. (μM)	4.4	9.9	14.7	20.3	23.1
R_g (Å)	21.8	21.7	21.8	21.4	21.4

Conc. (μM)	3.3	4.6	5.8	7.1	8.0
R_g (Å)	24.7	24.4	24.4	24.2	24.3

Figure 4.1 SAXS characterization and construct designs of MurD.

(A, B) Guinier approximation, ln[$I(q)$] vs. q^2, of D12 and MurD scattering data in SEC-SAXS. The straight line gives the slope of data points from the least-squares method.

(C) Lutetium tagging of MurD. The positions of the metals in MurD domains 1–2 M145C/D148C/C151A/E260C/K262C ($D12_{145\text{-}260}$) and MurD E260C/K262C/N360C/D362C ($MurD_{260–360}$) are shown as red spheres in the crystal structures of MurD (PDB:1e0d, 3uag). Domains 1, 2, and 3 are shown in yellow, blue, and pink, respectively. The distance between the lanthanoid ions, as estimated by crystal structures and pseudocontact shift (PCS) NMR analysis[106] is indicated by a red arrow.

(D) Chemical structure of the Caged Lanthanide NMR Probe 5 (CLaNP-5) tag.

(E) Close-up views of the positions of the lanthanoid ions fixed on MurD. The residues that were mutated to cysteine for ligation with the CLaNP-5 tag are indicated by red sticks.

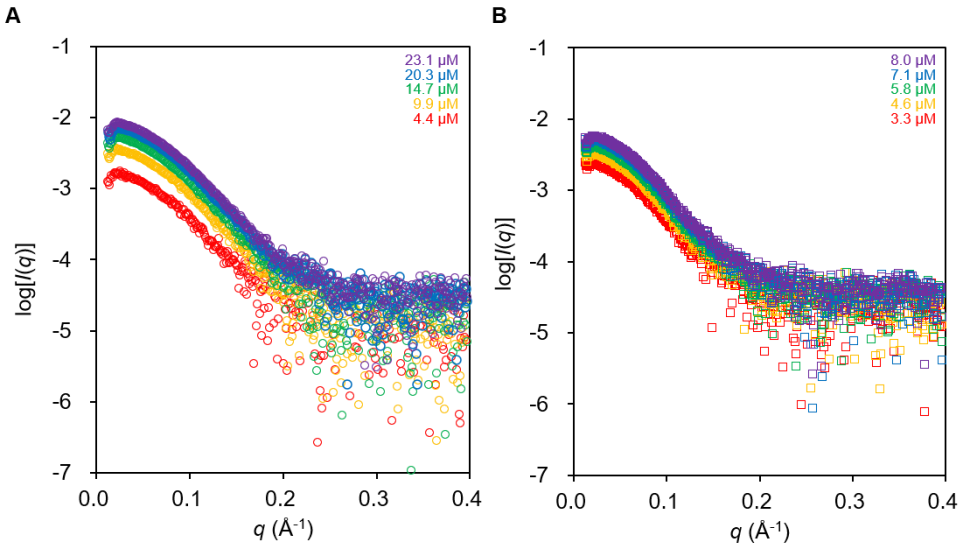


Figure 4.2 Monitoring monodisperse state at each concentration of D12 and MurD by using SEC-SAXS.

(A) SAXS curves for D12 in different concentration. The concentration of each is 4.4 μM (red), 9.9 μM (orange), 14.7 μM (green), 20.3 μM (blue) and 23.1 μM (purple).

(B) SAXS curves for MurD in different concentration. The concentration of each is 3.3 μM (red), 4.6 μM (orange), 5.8 μM (green), 7.1 μM (blue), 8.0 μM (purple)

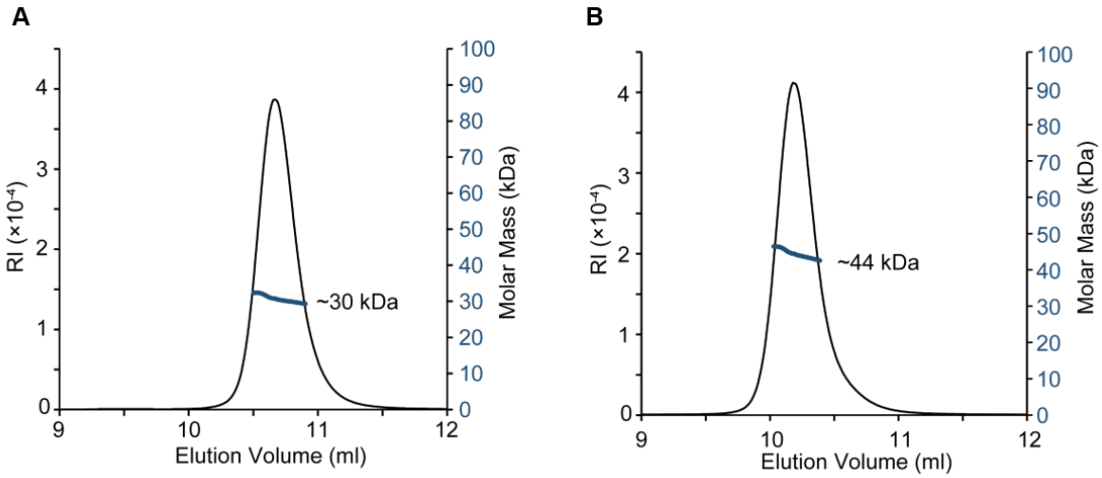


Figure 4.3 Monitoring monodisperse state of D12 and MurD by using SEC–MALS. The sample is D12 (A) and MurD (B).

4.3.2 Single pair SAXS for Lu^{3+}– Lu^{3+}distance measurement on $D12_{145-260}$

To test whether the inter-lutetium distance information could be extracted, I performed SAXS measurement on $D12_{145-260}$, in which CLaNP-5 tags containing Lu^{3+} were attached to the rigid regions of domain 2 (Figure 4.1C). To eliminate scattering from protein atoms, I next performed contrast-matched SAXS using a buffer containing 65% sucrose, which has the same electron density as the protein[11]. Prior to contrast-matched SAXS, the effect of 65% sucrose buffer on the structures of D12 and MurD was evaluated using circular dichroism (CD). Essentially identical CD spectra in the absence and presence of 65% sucrose were observed for D12 and MurD, indicating that the structures of D12 and MurD were preserved, even in 65% sucrose buffer (Figure 4.5).

Contrast-matched SAXS for the $D12_{145-260}$ data showed an oscillating curve profile characteristic of scattering interference between the two atoms (Figure 4.4A). As a reference experiment, I also obtained the SAXS data for D12 without Lu^{3+} ions (Figure 4.6). Although the data show slightly rugged profile presumably reflecting the non-uniform electron density distribution of the protein[28,29], the disturbance was much less significant compared to the oscillating curve profile for $D12_{145-260}$ attached with Lu^{3+} ions. The data of the contrast-matched SAXS for $D12_{145-260}$ attached with Lu^{3+} ions were fitted according to Equation 1:

$$I(q) = {A_e}^2 f^2 \frac{\sin(q.r)}{q.r} \quad (1)$$

where $I(q)$ is the scattering intensity [$q = 4\pi\sin(\theta)/\lambda$, θ = one-half of the scattering angle; λ = X-ray wavelength], A_e is the scattering amplitude of an atom, f is the scattering factor, and r is the atomic distance. The data were fitted using Equation 1 with the maximum entropy procedure coded in MATLAB, which was previously used for the analysis of gold nanocrystal-labeled DNA fragments[12,13]. The fitting using a distance range of 0–200 Å resulted in a major peak around 37 Å, with a couple of minor peaks above 50 Å (Figure 4.4B and 4.7A).

The Lu^{3+}– Lu^{3+} distance on $D12_{145-260}$ [as estimated from crystal structure and pseudocontact shift (PCS) analysis][9,30] was 35.8 ± 1.2 Å; this agreed with the major peak of SAXS-derived distance distribution curve. Because two Lu^{3+} ions are attached to the same domain, the width of the distribution should be mainly accounted for by the local conformational variation of the tag and the cysteine residues holding the tag. Additionally, the previous study exploiting ESR with double electron-electron resonance (DEER) measurement for $D12_{145-260}$ showed the distance distribution having a peak top at 36.4 Å and full width at half maximum (FWHM) 7.7 Å[30]; these are highly consistent with the SAXS-derived distance distribution, supporting the reliability of the distance measurement in this method.

Note that the minor peaks that appeared in the 50–150 Å range can be explained by intermolecular contributions, given that the average distance between the centers of protein at the concentration of contrast-matched SAXS measurement (700 μM) is estimated as ~130 Å. Thus, the data from contrast-matched SAXS on $D12_{145\text{-}260}$ are highly consistent with the expected values and the results from other measurement techniques; therefore, I concluded that single-pair SAXS using contrast matching successfully extracted the distance distribution of Lu^{3+}– Lu^{3+} on the protein.

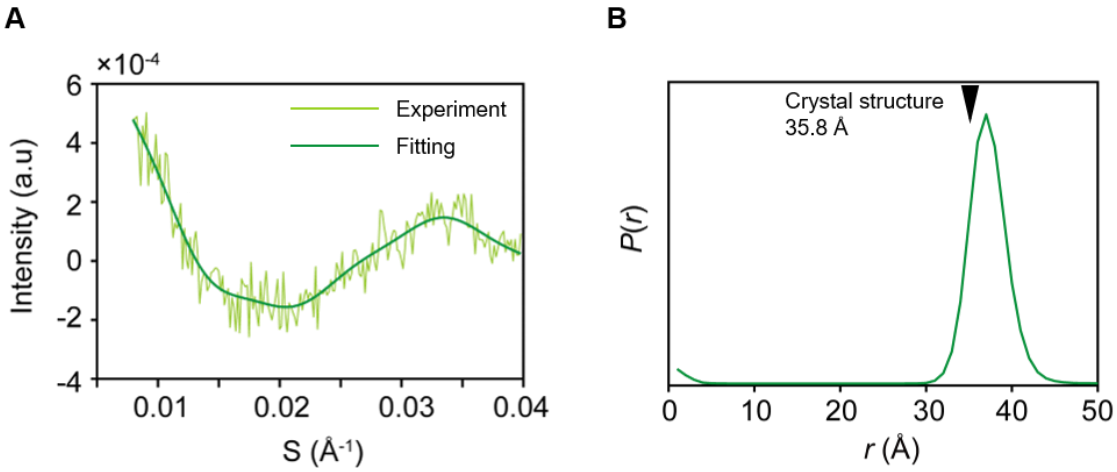


Figure 4.4 The distance between two points within $D12_{145\text{-}260}$ was measured using flow-SAXS.
(A) SAXS scattering data and fitting.
(B) Lu^{3+}– Lu^{3+} distance distribution. The arrow indicates the distance between the lutetium ions expected from the crystal structures of MurD domain1-2 (PDB: 3uag, 35.8 Å). The data were analyzed using the MATLAB program.

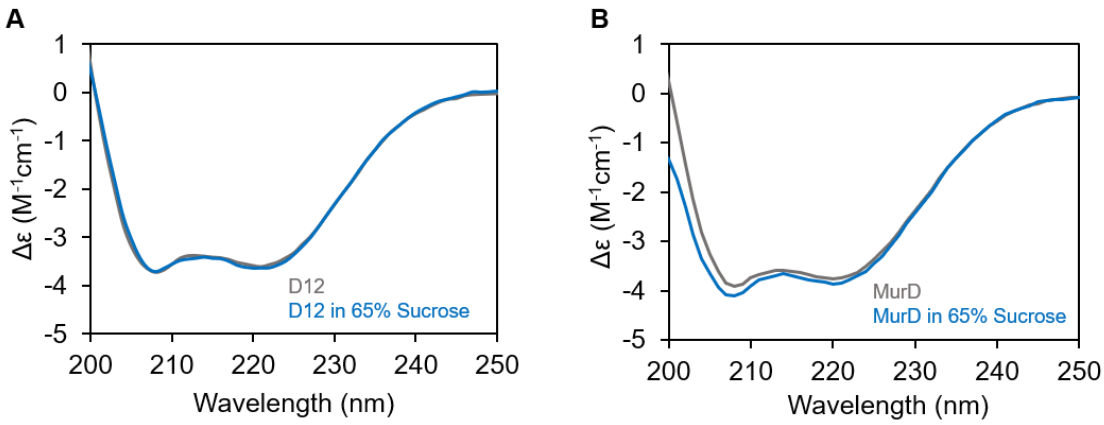


Figure 4.5 Monitoring 65% Sucrose-dependent secondary structural change of D12 (A) and MurD (B) by using Circular Dichroism spectrometer.

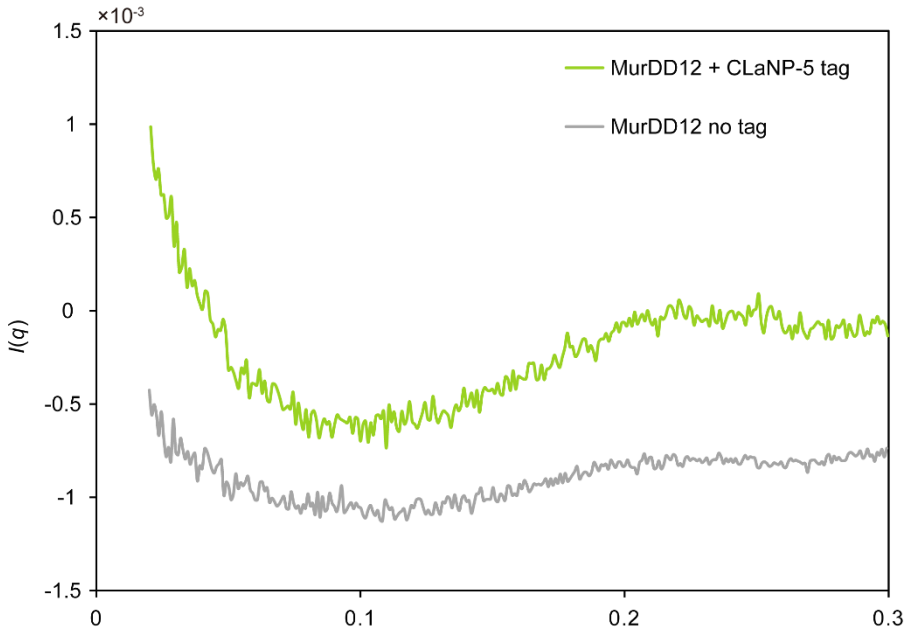


Figure 4.6 Absolute scattering data for $D12_{145\text{-}260}$ attached with the Lu^{3+} tags (light green) and MurD D12 with no tag (gray).

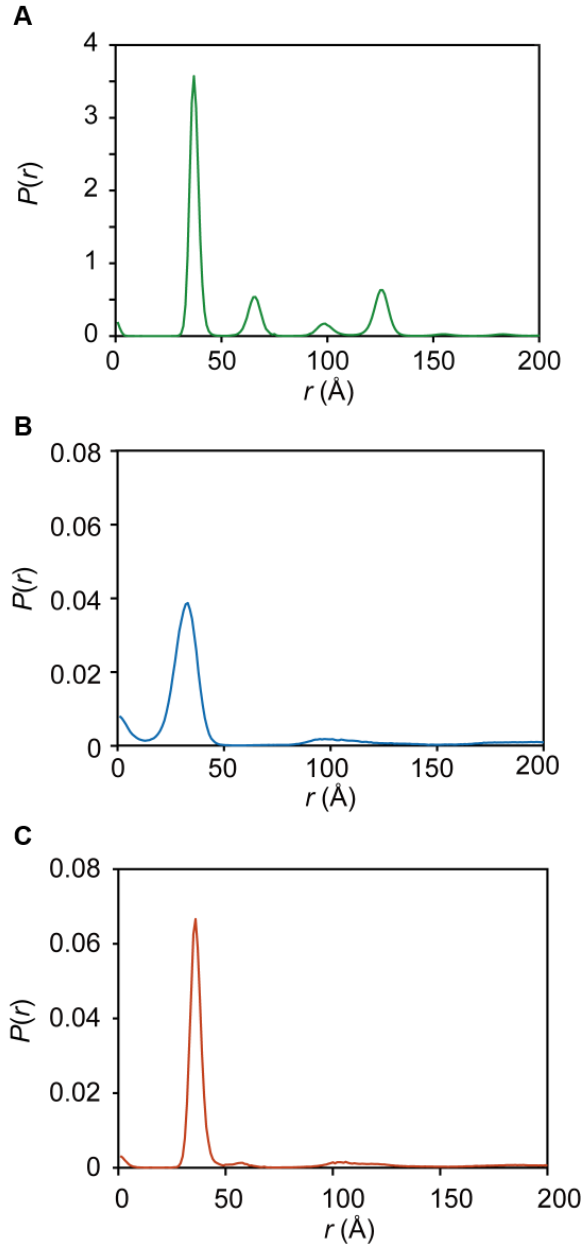


Figure 4.7 Lu^{3+}–Lu^{3+} distance distributions shown in 0–200 Å range. The data were analyzed using the MATLAB program. The sample is $D12_{145\text{-}260}$ (A), $MurD_{260\text{-}360}$ (B) and Inhibitor binding $MurD_{260\text{-}360}$ (C).

4.3.3 Conformational ensemble of MurD investigated by Lu^{3+}– Lu^{3+} single pair SAXS on $MurD_{260-360}$

The Lu^{3+}– Lu^{3+} distance for $MurD_{260–360}$ was measured to monitor the relative position of domain2 and 3 of MurD in the absence and presence of the inhibitor N-({3-[({4-[(Z)- (2,4-dioxo-1,3-thiazolidin-5-ylidene)methyl]phenyl}amino)methyl]phenyl}carbonyl)-d-glutamic acid[20]. Because the two lutetium ions were attached to domains 2 and 3, the inter-lutetium distance distribution should reflect the position of domain 3 with respect to domain 2 (Figure 4.1C). The inter-lutetium distance was expected to increase in the closed conformation (Figure 4.1C).

Contrast-matched SAXS of ligand-free $MurD_{260–360}$ resulted in an oscillating SAXS curve, reflecting Lu^{3+}– Lu^{3+} -derived scattering interference (Figure 4.8A). As is the case of $D12_{145-260}$, the data was fitted using Equation 1 with the maximum entropy procedure coded in MATLAB. The fitting resulted in a broader distance distribution peak than that obtained for $D12_{145-260}$ (Figure 4.4B and 4.8B, blue). The position of the peak top is close to the distance expected from the crystal structure of MurD in the open conformation (PDB:1e0d, 32.1 ± 2.0 Å)[9,22,30] (Figure 4.1C and 4.8B, blue), suggesting that MurD exists in a variety of conformational states, including the open conformation seen in the crystal structure as a major state and other associated states.

Contrast-matched SAXS of inhibitor-bound $MurD_{260-360}$ also exhibited an oscillating SAXS curve. Nonetheless, the profile (especially the oscillation frequency and amplitude) was different from that of ligand-free MurD (Figure 4.8A). The fitting resulted in a distance distribution with a major peak at a longer distance and a narrower width compared to ligand-free $MurD_{260-360}$ (Figure 4.8B, 4.7B, and 4.7C). The narrower width of the peak for inhibitor-bound $MurD_{260–360}$ indicated that inhibitor-bound MurD exists in more defined conformational spaces. The peak position at longer distances in the complex with the inhibitor suggests that the conformation of MurD changed from open to closed.

To verify the fitting analysis, theoretical SAXS curves were calculated from the normal distribution, as represented by Equation 2:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (2)$$

where μ is mean of expectation and equals to the peak top of the distance distribution, and σ is standard deviation with the relationship FWHM = 2.35σ. The theoretical SAXS traces corresponding to the distance distributions for $MurD_{260-360}$ in the absence and presence of the inhibitor (σ = 5.10 Å, μ=33 Å and σ = 2.55 Å, μ=36 Å) replicated the experimental data well (Figure 4.8 and 4.9). Next, the reliable range of the distance distribution, apparent distance resolution to distinguish structures, was examined

by comparing the experimental and calculated SAXS curves. The experimental SAXS data for $D12_{145-260}$ were well reproduced by the calculated curve from the normal distribution of the distance corresponding to Figure 4.4B: $\mu = 37$ Å and $\sigma = 2.12$ Å (Figure 4.10A and B). The calculated curves with varying μ showed that the those with μ in a range of 37 ± 3 Å fit within the noise level of the experimental data whereas those with $\mu = 37 \pm 4$ Å protrude from the experimental curve, suggesting that the reliable range of the peak top can be assumed as 37 ± 3 Å. The same evaluation was performed for the data of $MurD_{260-360}$ in the absence and presence of the inhibitor (Figure 4.10C-F), resulting the estimated reliable range of the peak top for $MurD_{260-360}$ in the absence and presence of the inhibitor 33 ± 2 Å and 36 ± 2 Å, respectively. With the above evaluations, I concluded that Lu^{3+}– Lu^{3+} single-pair SAXS allowed us to measure the distance distribution of MurD and inhibitor-induced changes in the distance distribution. Collectively, distance distribution measurements by single-pair SAXS revealed the conformational distribution and ligand-driven conformational changes of the multi-domain protein MurD in solution, thus highlighting the usefulness of this method for evaluating the conformational distribution of proteins in solution.

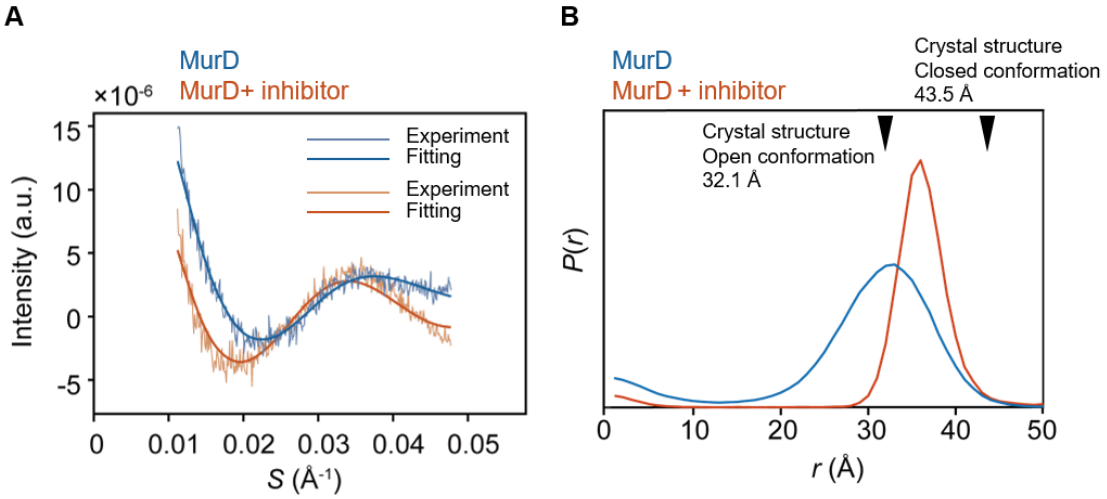


Figure 4.8 The distance between two points within $MurD_{260–360}$ was measured using flow-SAXS. (A) SAXS scattering data for $MurD_{260–360}$ (blue) and inhibitor-bound $MurD_{260–360}$ (red). (B) Lu^{3+}– Lu^{3+} distance distributions. The insets show a wider distance distribution range of ≤50 Å. The arrows indicate the distance between the lutetium ions expected from the crystal structures of MurD in open (PDB: 1e0d, 32.1 Å) and closed (PDB: 3uag, 43.5 Å) conformations. The data were analyzed using the MATLAB program.

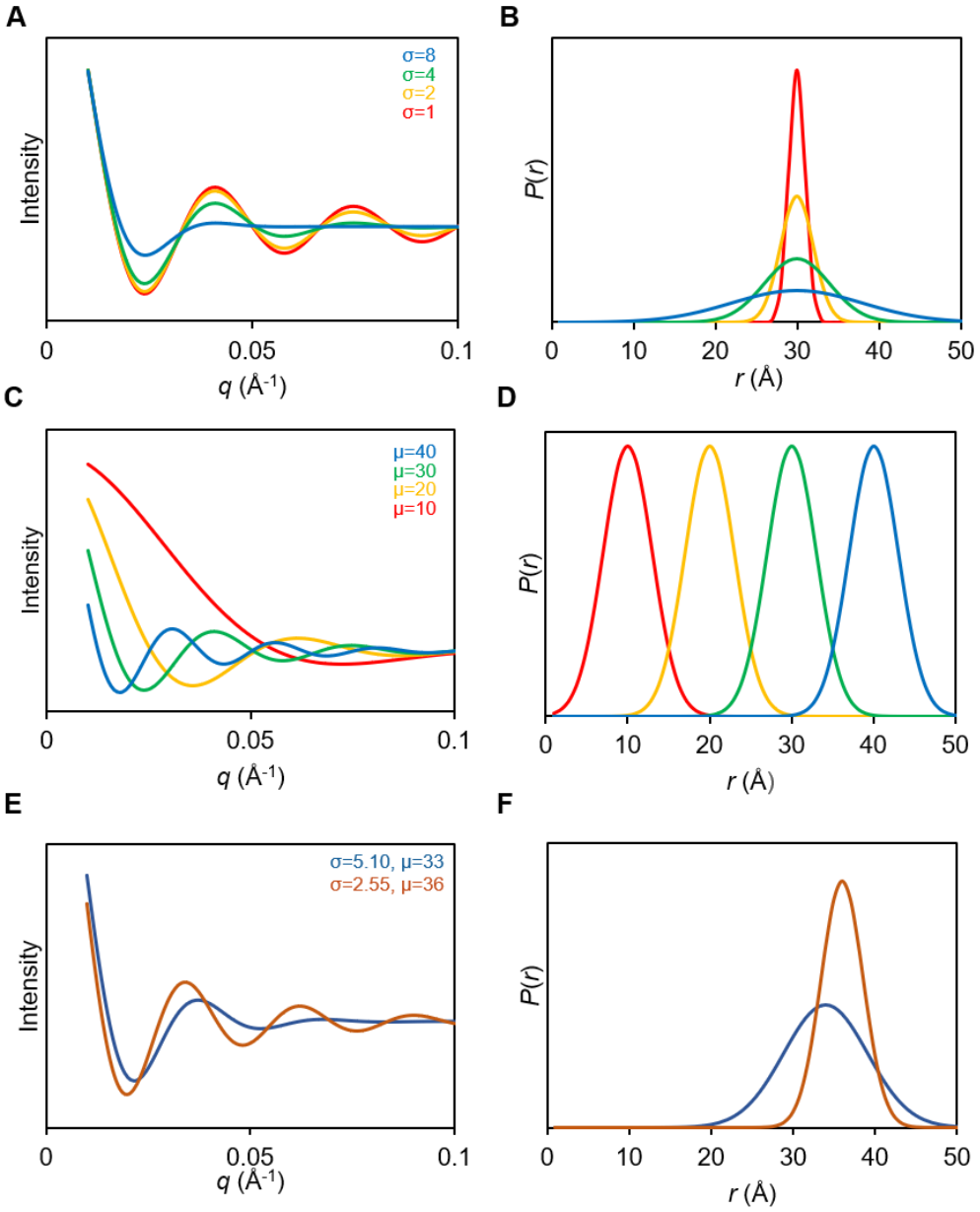


Figure 4.9 Superimposition of theoretical SAXS traces (A), (C), (E) and distance distributions (B), (D), (F).

(A, B) Variations in standard deviations of the Gaussian distribution: σ = 1 (red), 2 (orange), 4 (green), 8 (blue).

(C, D) Variations in means of the Gaussian distribution: μ = 10 (red), 20 (orange), 30 (green), 40 (blue).

(E, F) The theoretical SAXS traces corresponding to the distance distributions for $MurD_{260-360}$ in the absence (red) and presence of the inhibitor (blue) (σ = 5.10 Å, μ=33 Å and σ = 2.55 Å, μ=36 Å).

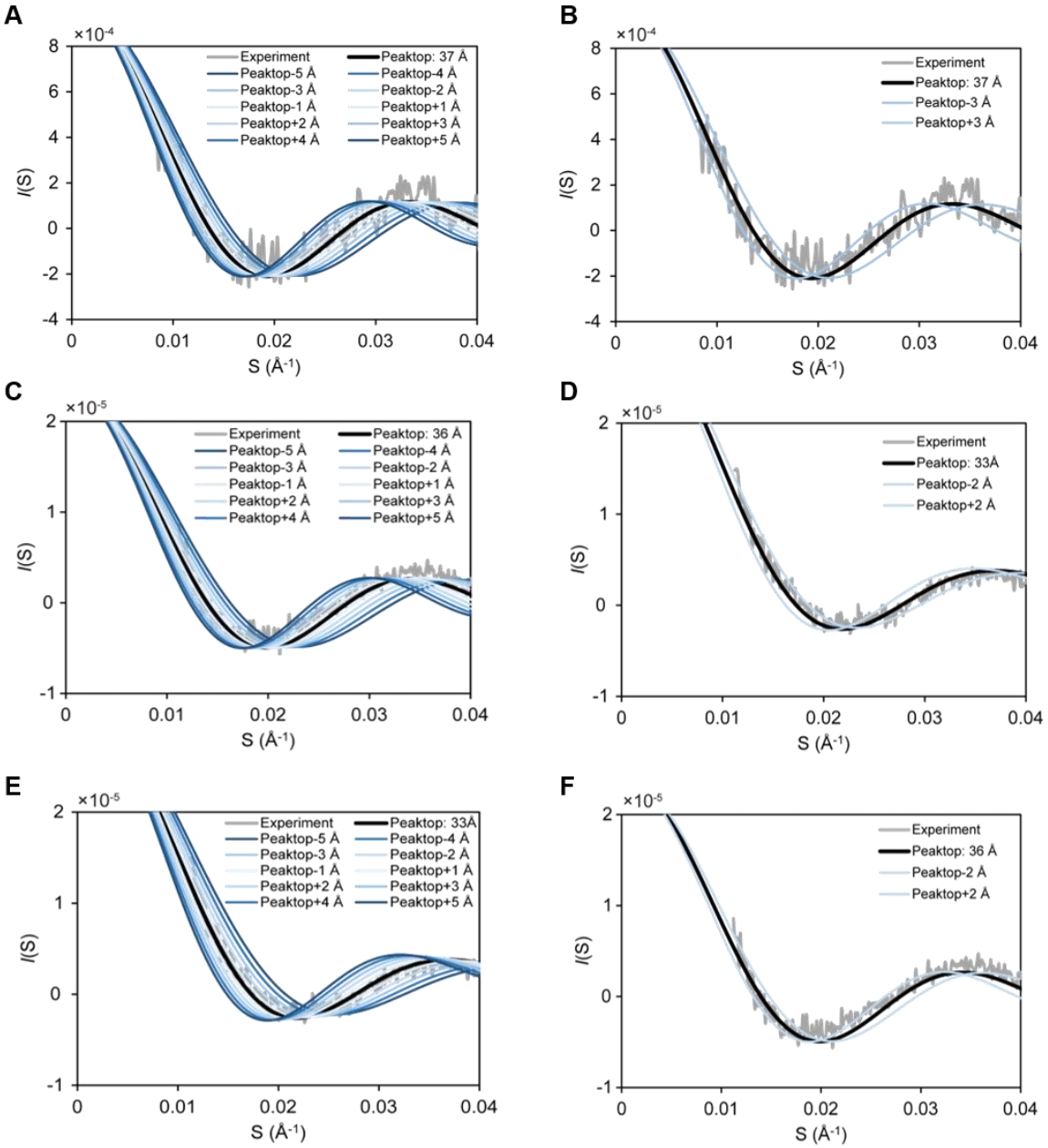

A
×10-4
Experiment
Peaktop: 37 Å
Peaktop-5 Å
Peaktop-4 Å
Peaktop-3 Å
Peaktop-2 Å
Peaktop-1 Å
Peaktop+1 Å
Peaktop+2 Å
Peaktop+3 Å
Peaktop+4 Å
Peaktop+5 Å
I(S)
S (Å-1)
B
×10-4
Experiment
Peaktop: 37 Å
Peaktop-3 Å
Peaktop+3 Å
I(S)
S (Å-1)
C
×10-5
Experiment
Peaktop: 36 Å
Peaktop-5 Å
Peaktop-4 Å
Peaktop-3 Å
Peaktop-2 Å
Peaktop-1 Å
Peaktop+1 Å
Peaktop+2 Å
Peaktop+3 Å
Peaktop+4 Å
Peaktop+5 Å
I(S)
S (Å-1)
D
×10-5
Experiment
Peaktop: 33Å
Peaktop-2 Å
Peaktop+2 Å
I(S)
S (Å-1)
E
×10-5
Experiment
Peaktop: 33Å
Peaktop-5 Å
Peaktop-4 Å
Peaktop-3 Å
Peaktop-2 Å
Peaktop-1 Å
Peaktop+1 Å
Peaktop+2 Å
Peaktop+3 Å
Peaktop+4 Å
Peaktop+5 Å
I(S)
S (Å-1)
F
×10-5
Experiment
Peaktop: 36 Å
Peaktop-2 Å
Peaktop+2 Å
I(S)
S (Å-1)

Figure 4.10 Evaluation of the reliable distance range for the distance information. The experimental data were compared with theoretical SAXS traces assuming Gaussian distribution with varying mean (μ) corresponding to the peak top of the distance distribution peak.
(A, B) The experimental SAXS data of $D12_{145\text{-}260}$ superimposed with theoretical SAXS trace calculated from Gaussian distribution with σ = 2.12 Å (FWHM = 5 Å) and μ = 37 Å that correspond to the major peak in the distance distribution (Figure 4.4B). The theoretical SAXS traces with μ shifted up to ± 5 Å (A) or μ shifted to ± 3 Å (B) are also shown. Considering the noise level, the calculated curves with μ shifted up to ± 3 Å can fit in the experimental data.
(C, D) The experimental SAXS data of apo $MurD_{260\text{-}360}$ superimposed with theoretical SAXS traces calculated from Gaussian distribution with σ = 5.10 Å (FWHM = 12 Å) and μ = 33 Å that correspond to the major peak in the distance distribution (Figure 4.8B). The theoretical SAXS traces with μ shifted up to ± 5 Å (C) or μ shifted to ± 2 Å (D) are also shown. Considering the noise level, the calculated curves with μ shifted up to ± 2 Å can fit in the experimental data.
(E, F) The experimental SAXS data of apo $MurD_{260\text{-}360}$ superimposed with theoretical SAXS traces calculated from Gaussian distribution with σ = 2.55 Å (FWHM = 6 Å) and μ = 36 Å that correspond to the major peak in the distance distribution (Figure 4.8B). The theoretical SAXS traces with μ shifted up to ± 5 Å (E) or μ shifted to ± 2 Å (F) are also shown. Considering the noise level, the calculated curves with μ shifted up to ± 2 Å can fit in the experimental data.

4.4 Discussion

Despite the importance of the conformational distribution of proteins for their function, their evaluation in solution is often not straightforward. This study demonstrates that single-pair SAXS using lanthanide ions as a scattering source can be used to evaluate the distance distribution of a multi-domain protein in solution. Scattering interference between lanthanide ions was selectively observed by suppressing those from protein atoms using contrast-matched SAXS in 65% sucrose buffer. The distance distribution obtained for $D12_{145\text{-}260}$ with the two lanthanide ions in domain 2 showed a major peak with peak top at 37 Å with estimated reliable range of ± 3 Å and was consistent with the distance expected from the crystal structure if the local conformational variation around the tag was considered, indicating the ability of this method to measure the distance between two specific points on the protein (Figure 4.4). Distance measurement for $MurD_{260\text{-}360}$ in the ligand-free state showed a broader distance distribution than that observed for $D12_{145\text{-}260}$, showing a variety of conformational states of MurD. This is consistent with the observations of previous studies using PCS-NMR, ESR, and MD simulation[9,30,31] (Figure 4.11). The position of the peak top, 33 Å with estimated reliable range of ± 2 Å, is also consistent with the crystal structure in the open conformation and previous data from ESR and MD simulation (Figure 4.1C and 4.11A). $MurD_{260\text{-}360}$ in complex with the inhibitor showed a narrower distance distribution at longer distances compared to ligand free $MurD_{260\text{-}360}$, indicating that the inhibitor induces a conformational change toward more closed and defined conformations. It should be noted that the position of the peak top for inhibitor-bound $MurD_{260–360}$ was located at 36 Å with estimated reliable range of ± 2 Å (Figure 4.10E and F) that is distinct from the crystal structure of MurD in the closed conformation[19] (Figure 4.1C) and the previous ESR distance measurement for frozen protein at 10 K (Figure 4.11B)[30]. These observations suggest the most populated state of inhibitor-bound $MurD_{260\text{-}360}$ in solution is different from that seen in the ESR measurement or in the crystal structure. Although the conformational difference of MurD between the crystal and solution has been suggested by other methods (including ESR-DEER measurement[30] and MD simulation[31]), the current data from single-pair SAXS provide the first experimental views of MurD conformation in aqueous solution. However, a current limitation is the difficulty in visualizing the actual conformational state of the inhibitor-bound $MurD_{260\text{-}360}$ in solution, mostly because of the limited structural information from distance measurement for a single pair. Further study including the multiple sets of distance measurement for lanthanide ions attached on the other positions would be anticipated.

Using a lanthanide-binding tag, this study successfully demonstrated the application of contrast-matched single-pair SAXS as a tool for visualizing the conformational distribution of a specific protein site. One of the advantages of this strategy is that the distance distribution can be measured for any desired point on the protein as long as the lanthanide can be fixed by the tag. Although I demonstrated the application of the CLaNP-5 tag, the same strategy can be exploited with other lanthanide-binding tags, including single-arm tags for easier design of the fixing points and tags with non-reducing linkages with the protein for measurement under reduced conditions[8]. These facts highlight the generality and significance of this method for protein structural studies. For application of this method to other proteins, one may concern about the impact of 65% sucrose to the structure of the protein. It should be important to evaluate the structure of the target protein in 65% sucrose by other methods including CD.

A possible future application of single-pair SAXS is in the structural study of FUS. Since FUS forms droplets starting from the LC domain, the conformational distribution of FUS inside the droplet is revealed by adding metal ions in the LC domain. By combining information on the distance distribution between two points in an aqueous solution with information obtained from NMR, spectroscopy, and MD simulations, more detailed structural information on FUS can be obtained.

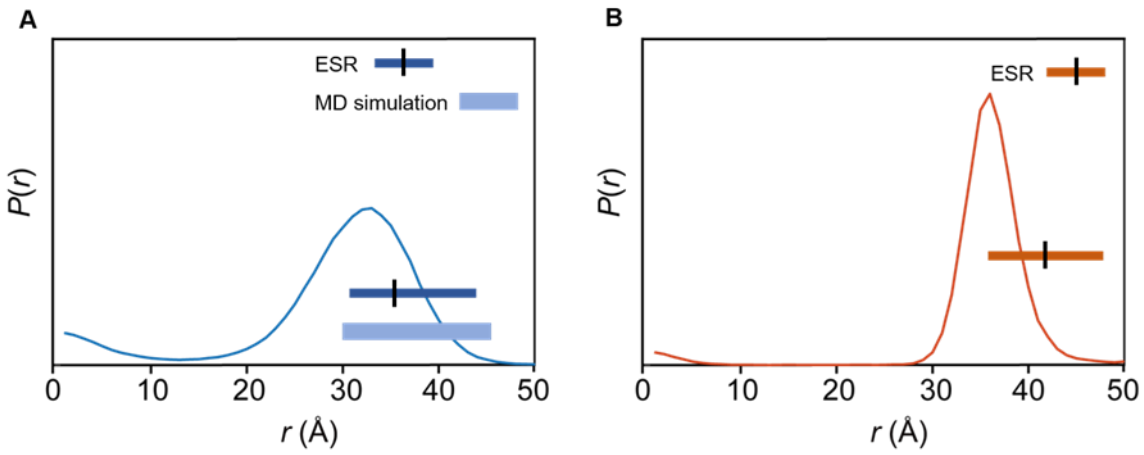


Figure 4.11 Comparison of the distance distribution of $MurD_{260-360}$ from SAXS, ESR, and MD simulation.

(A) SAXS-derived Lu^{3+}– Lu^{3+} distance distribution of apo $MurD_{260-360}$ (blue line), shown with the distance range derived from ESR distance measurement (blue bar with black line) and MD simulation (pale blue box).

(B) SAXS-derived Lu^{3+}– Lu^{3+} distance distribution of apo $MurD_{260-360}$ (dark orange line), shown with the distance range derived from ESR distance measurement (dark orange bar with black line). The black line in the bar representing ESR-derived distance range indicates the position of the peak top. The distance range from ESR distance measurement is defined by the region at half-height of the major peak. The distance information for ESR distance measurement and MD simulation on MurD is based on the previous report[106].

Reference

1. 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.* **80**, 26–63 (2014).
2. Han, J. H., Batey, S., Nickson, A. A., Teichmann, S. A. & Clarke, J. The folding and evolution of multidomain proteins. *Nat. Rev. Mol. Cell Biol.* **8**, 319–330 (2007).
3. Ma, B., Tsai, C. J., Haliloğlu, T. & Nussinov, R. Dynamic allostery: Linkers are not merely flexible. *Structure* **19**, 907–917 (2011).
4. Levitt, M. Nature of the protein universe. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 11079–11084 (2009).
5. Skora, L., Mestan, J., Fabbro, D., Jahnke, W. & Grzesiek, S. NMR reveals the allosteric opening and closing of Abelson tyrosine kinase by ATP-site and myristoyl pocket inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* **110**, (2013).
6. Kurt Wüthrich, Z. NMR with Proteins and Nucleic Acids. 11–13 (1985).
7. Egelman, E. H. The Current Revolution in Cryo-EM. *Biophys. J.* **110**, 1008–1012 (2016).
8. Saio, T. & Ishimori, K. Accelerating structural life science by paramagnetic lanthanide probe methods. *Biochim. Biophys. Acta - Gen. Subj.* **1864**, 129332 (2020).
9. Saio, T. *et al.* Ligand-driven conformational changes of MurD visualized by paramagnetic NMR. *Sci. Rep.* **5**, 1–11 (2015).
10. Kikhney, A. G. & Svergun, D. I. A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Lett.* **589**, 2570–2577 (2015).
11. Grishaev, A., Anthis, N. J. & Clore, G. M. Contrast-matched small-angle x-ray scattering from a heavy-atom-labeled protein in structure determination: Application to a lead-substituted calmodulin-peptide complex. *J. Am. Chem. Soc.* **134**, 14686–14689 (2012).
12. Mathew-Fenn, R. S., Das, R., Silverman, J. A., Walker, P. A. & Harbury, P. A. B. A molecular ruler for measuring quantitative distance distributions. *PLoS One* **3**, 1–9 (2008).
13. Mathew-Fenn, R. S., Das, R. & Harbury, P. A. B. Remeasuring the double helix. *Science (80-.).* **322**, 446–449 (2008).
14. Taguchi, Y., Saio, T. & Kohda, D. Distance Distribution between Two Iodine Atoms Derived from Small-Angle X-ray Scattering Interferometry for Analyzing a Conformational Ensemble of Heavy Atom-Labeled Small Molecules. *J. Phys. Chem. Lett.* **11**, 5451–5456 (2020).

15. Popovych, N., Tzeng, S. R., Tonelli, M., Ebright, R. H. & Kalodimos, C. G. Structural basis for cAMP-mediated allosteric control of the catabolite activator protein. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 6927–6932 (2009).

16. Keizers, P. H. J., Saragliadis, A., Hiruma, Y., Overhand, M. & Ubbink, M. Design, synthesis, and evaluation of a lanthanide chelating protein probe: CLaNP-5 yields predictable paramagnetic effects independent of environment. *J. Am. Chem. Soc.* **130**, 14802–14812 (2008).

17. Lipfert, J. & Doniach, S. Small-angle X-ray scattering from RNA, proteins, and protein complexes. *Annu. Rev. Biophys. Biomol. Struct.* **36**, 307–327 (2007).

18. Smith, C. A. Structure, Function and Dynamics in the mur Family of Bacterial Cell Wall Ligases. *J. Mol. Biol.* **362**, 640–655 (2006).

19. Bertrand, J. A. *et al.* Determination of the MurD mechanism through crystallographic analysis of enzyme complexes. *J. Mol. Biol.* **289**, 579–590 (1999).

20. Zidar, N. *et al.* Discovery of novel 5-benzylidenerhodanine and 5-benzylidenethiazolidine-2, 4-dione inhibitors of MurD ligase. *J. Med. Chem.* **53**, 6584–6594 (2010).

21. Tomašić, T. *et al.* Dual inhibitor of MurD and MurE ligases from escherichia coli and staphylococcus aureus. *ACS Med. Chem. Lett.* **3**, 626–630 (2012).

22. Bertrand, J. A. *et al.* 'Open' structures of MurD: Domain movements and structural similarities with folylpolyglutamate synthetase. *J. Mol. Biol.* **301**, 1257–1266 (2000).

23. Kotnik, M. *et al.* Structural and Functional Characterization of Enantiomeric Glutamic Acid Derivatives as Potential Transition State Analogue Inhibitors of MurD Ligase. *J. Mol. Biol.* **370**, 107–115 (2007).

24. Keizers, P. H. J., Desreux, J. F., Overhand, M. & Ubbink, M. Increased paramagnetic effect of a lanthanide protein probe by two-point attachment. *J. Am. Chem. Soc.* **129**, 9292–9293 (2007).

25. Kobashigawa, Y., Kumeta, H., Ogura, K. & Inagaki, F. Attachment of an NMR-invisible solubility enhancement tag using a sortase-mediated protein ligation method. *J. Biomol. NMR* **43**, 145–150 (2009).

26. Shimizu, N. *et al.* Software development for analysis of small-angle x-ray scattering data. *AIP Conf. Proc.* **1741**, (2016).

27. Fleming, P. J. & Fleming, K. G. HullRad: Fast Calculations of Folded and Disordered Protein and Nucleic Acid Hydrodynamic Properties. *Biophys. J.* **114**, 856–869 (2018).

28. Grant, T. D. Ab initio electron density determination directly from solution scattering data. *Nat.*

Methods **15**, 191–193 (2018).

29. Jeffries, C. M. *et al.* Preparing monodisperse macromolecular samples for successful biological small-Angle X-ray and neutron-scattering experiments. *Nat. Protoc.* **11**, 2122–2153 (2016).
30. Saio, T. *et al.* Conformational ensemble of a multidomain protein explored by Gd3+ electron paramagnetic resonance. *Biophys. J.* **120**, 2943–2951 (2021).
31. Nakagawa, H. *et al.* Conformational dynamics of a multidomain protein by neutron scattering and computational analysis. *Biophys. J.* **120**, 3341–3354 (2021).

Ⅴ. Conclusion

FUS is a intrinsically disordered protein that functions through self-associating, eliminating solvents, and forming droplets. The molecular chaperone Kapβ2 interacts with FUS and inhibits its self-assembly, thereby regulating droplet formation. Dysregulation of FUS self-assembly leads to the formation of FUS aggregates that are toxic to the cell. Thereby, the mechanism for regulation and dysregulation of FUS assembly is a key to unveil the molecular mechanism of ALS. However, the weak interaction as well as dynamic feature of FUS and Kapβ2 has prevented the understanding of the detailed mechanis.

In this thesis, the mechanism of mechanism for regulation and dysregulation of FUS assembly through Kapβ2 was investigated by exploiting solution measurement techniques including NMR. For further investigation of dynamic structure of IDPs including FUS, I have also proceeded with technical development exploiting SAXS. In Chapters II and III, I analyzed the interaction of Kapβ2 with molecules that inhibit the regulatory function of FUS droplet formation. In Chapter 4, I discussed the development and application of a new methodology to determine what structural changes are induced in the molecules inside the droplet by the FUS interaction. The results and conclusions drawn from experimental data are highlight below.

PRn interacts with the NLS binding site of Kapβ2 (Chapter II)

ALS disease produces Pro-Arg poly-dipeptide (PRn), an abnormally elongated repeat sequence in the ALS-related gene. In Chapter 2, I attempted to identify the interaction site between Kapβ2 and PRn, assuming that this PRn interacts with Kapβ2 to inhibit the droplet regulatory function of FUS. NMR measurements revealed that the addition of PRn caused a change in the chemical shift of certain NMR signals, some of which were attributed to the nuclear translocation signal (NLS) binding site in Kapβ2. It has been reported that FUS also binds to this NLS binding site in Kapβ2. Furthermore, the dissociation constants of PRn and FUS for Kapβ2 were comparable. Therefore, it was suggested that PRn competitively binds to the NLS binding site of FUS, thereby inhibiting the binding of Kapβ2 to FUS, resulting in excessive self-aggregation of FUS and inducing aggregation of FUS droplets.

PRn interacts with the loop region of Kapβ2 (Chapter III)

In Chapter III, I examined the diffusion rate of FUS in a droplet due to the interaction

between PRn and Kapβ2, since it has been reported that the diffusion rate inside a droplet decreases when the droplet aggregates. To measure the density inside the FUS droplets, I used Holographic microscopy. Compared to the addition of one equal amount of PRn and Kapβ2, the addition of an excess amount of PRn increased the FUS droplet density. Next, to quantify the density of FUS droplets, I used fluorescence recovery after optical fading (FRAP), which can track the diffusion rate of molecules in the droplets. Addition of Kapβ2 and excess PRn resulted in slower recovery of fluorescence from photo-bleach compared to FUS alone, indicating a reduction in the rate of diffusion within the droplet, and thus an enhancement of FUS self-association. On the other hand, no such decrease in diffusion rate was observed with PRn alone, indicating that PRn may promote FUS self-assembly by interacting with Kapβ2 rather than FUS. These results suggest that PRn not only inhibits FUS binding to the NLS binding site of Kapβ2, but also promotes FUS self-association by interacting with Kapβ2 at another site. As the PRn interaction site of Kapβ2 that promotes FUS self-assembly, I focused on the loop region, which, like the NLS binding site, contains many negatively charged amino acid residues.

To confirm the interaction between PRn and the loop region, I isolated only the loop region and measured its ^{1}H-^{15}N HSQC spectrum. The NMR signals of many negatively charged amino acid residues in the loop region showed chemical shift perturbations due to the addition of PRn, indicating an electrostatic interaction. In other words, PRn interacts with the loop region in addition to the NLS binding site, suggesting that PRn reduces the diffusion rate of FUS.

Development and Application of a New Methodology for Investigating Structural Changes of Molecules Inside FUS Droplets (Chapter IV)

In Chapter IV, I discuss the development and application of a new methodology to examine the conformational changes induced in the molecules inside a droplet and how they induce aggregation. I focused my attention on the fact that conformational changes of a protein can also be traced as changes in the distance between two specific amino acid residues in its sequence or their distribution. Since the distance between two points can be calculated from scattering interferences in small-angle X-ray scattering (SAXS), I attempted to fix metal ions with high electron density near two specific amino acid residues and extract only the scattering interferences originating from those metal ions from the protein-derived scattering interferences. First, to establish the method, I used the multidomain

protein MurD, which has a reported conformation and can fix electron-dense lanthanide metal ions at specific sites, and obtained the distance distribution (P(r)) function between these two points by SAXS scattering interferometry. From the obtained P(r) function, the distance between the two points was estimated, which agrees within NMR structure determination. By applying such a method, it is possible to estimate the structure and changes in its distribution from the distance between specific residues, even for intrinsically disordered proteins such as FUS, which have a variety of conformations. Furthermore, it is expected to be possible to trace what conformational changes are induced in FUS inside the droplet during aggregation.

The application of this method to FUS will be discussed. FUS forms droplets through low complexity domains. By attaching metal ions to FUS LC and performing SAXS measurements using a strategy similar to MurD, structural information can be obtained from the distance distribution of the metal. By estimating the structure inside the droplet from the intermolecular distance of FUS, I expect to be able to trace what kind of structural changes are induced in the FUS inside the droplet during aggregation.

Reference

1. Han, T. W. *et al.* Cell-free formation of RNA granules: Bound RNAs identify features and components of cellular assemblies. *Cell* **149**, 768–779 (2012).