Reconstitution of the Core Ccr4-Not Complex and Evidence for Possible Interaction between YTHDF2 and Human CNOT9

Yutong Gong

Thesis submitted to the University of Nottingham for the degree of

Master of Research

November 2024

Abstract

In eukaryotic cells, shortening and removal of the poly(A) tail of cytoplasmic mRNA (deadenylation) is a key step in mRNA degradation. Studies have shown that the Ccr4-Not complex plays a fundamental role in regulating gene expression at the posttranscriptional level. The different parts of this complex are connected by flexible linkers, which act like bridges to hold the structure together and enable it to work properly. However, it's currently unclear whether these linkers have preferred conformations or have considerable flexibility. To purify and generate different sub-complexes of Ccr4-Not for biochemical and structural investigation, and further investigate the remodeling and structural dynamics of the Ccr4-Not complex, we focused on purifying and analyzing two important subcomplexes of the core Ccr4-Not complex: the Nuclease/CNOT9 module and the Nuclease/CNOT9/NOT module. I demonstrated the practicality and efficiency of the strategy of co-expressing protein complexes and purifying proteins using Immobilised metal affinity chromatography (IMAC). Secondly, I investigated recruitment of Ccr4-Not complex by YTHDF2 RNA binding protein via CNOT9, since YTHDF2 protein plays a central role in the degradation of m^6A-modified mRNA, which is mainly achieved through interaction with the Ccr4-Not complex. In order to explore the interaction between CNOT9 and YTHDF2 domain and further figure out the mechanistic basis of mRNA degradation mediated by YTHDF2 through the Ccr4-Not complex, I demonstrated through peptide pull-down experiments that human CNOT9 can interact with YTHDF2.

Acknowledgements

I would first like to express my appreciation to my primary supervisor, Dr Sebastiaan Winkler, for giving me the opportunity and big support to complete my MRes. I would also like to thank the other supervisor Dr Lodewijk Dekker, for giving me a lot of useful suggestion and providing the experiment tools for purify my target protein. I would also like to thank Folkert Van Werven and Wei Chan, Cell Fate and Gene Regulation Laboratory, for the nice collaboration of the project "interaction between YTHDF2 and CNOT9". I would like to thank Dr Sebastiaan Winker for discussing different problems with me and guiding me to solve the problems. I would like to thank Fahad Almulhim and Ashkan Bakhtiari for assistance with purifying protein by IMAC and making my plasmids. I would also like to thank the members of GRRB Lab, Hilary, Maria, Angie, Merryn, Alfredo, Sasha, Rohum, Silvia, Sina, Grace, Chai, Esra, Olubusola, Chris, Ammar and Sabesan. They really helped me a lot and made the lab atmosphere very friendly and relaxed, which made me feel very warm when I first arrived in the UK. I would also thank Dr Hilary for the Laboratory safety training when I started my lab work. And I would like to thank our technician Maria for helping us with a lot of daily issue, especially when we moving the lab. I also would like to thank my roommates Coco, Ellen and Irene for their company and encouragement. Finally I need to thank my family, Mum, Dad and my grandparents for supporting me in pursuing my dreams and continuing my studies abroad.

Table of Contents

Abbreviations List

Ccr4	Carbon catabolite repression 4
RNA	Ribonucleic acid
mRNA	Messenger RNA
miRNA	MicroRNA
snRNA	Small nuclear RNA
BSA	Bovine serum albumin
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
E.coli	*Escherichia coli*
EEP	endonuclease-exonuclease phosphatase
eIF4	Elongation initiation factor 4
IMAC	Immobilised metal affinity chromatography
IPTG	Isopropyl -D-1-thiogalactopyranoside
LB	Lysogeny Broth
LRR	Leucin-rich repeat
MCS	Multi cloning site
NMD	Nonsense mediated decay
m^6A	N^6-methyladenosine

m^7G	7-methylguanylate
TTP	Tristetraprolin
PCR	Polymerase Chain Reaction
Poly (A)	Poly adenosine tail
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
YTH	YT521-B homology
UTR	Untranslated region
PARN	poly(A)-specific ribonuclease
PVDF	polyvinylidene fluoride

Chapter 1

Introduction

Chapter 1. Introduction

1.1 Gene regulation in eukaryotic cells

Gene regulation encompasses a complex array of processes that cells employ to precisely control the production of RNA and proteins, which are essential for cellular function. The survival and development of multicellular organisms are also intricately linked to the regulation of gene expression. The initial phase of gene regulation is transcription, a highly regulated process in which DNA is transcribed into messenger RNA (mRNA) (or pre-mRNA in eukaryotes) by RNA polymerase enzymes. In eukaryotic cells, transcription is mediated by three distinct types of RNA polymerases, designated as RNA polymerase I, II, and III (RNAPI-III) [1]. Among these, RNA polymerase II (RNAPII) is specifically responsible for synthesizing precursor messenger RNA (pre-mRNA), an immature RNA transcript that requires further processing. The pre-mRNA undergoes a series of critical post-transcriptional modifications to become a functional and mature mRNA molecule. These modifications are essential for the stability, export, and translation of mRNA, and include the addition of a 5' cap (capping), the removal of non-coding sequences through splicing, and the addition of a polyadenylated tail at the 3' end (polyadenylation). Each of these modifications plays an essential role in ensuring the accurate expression of genes, thereby contributing to the intricate regulation of gene expression necessary for cellular homeostasis and organismal development.

The initial process, known as capping, involves modification at the 5' end of the mRNA molecule. This process is initiated by the enzyme guanylyl

transferase, which catalyzes the formation of a unique 5'-to-5' triphosphate linkage by attaching a guanine residue to the 5' terminus of the nascent mRNA. Subsequently, this guanine residue undergoes methylation through the action of methyltransferase, resulting in the formation of the 5' cap. The addition of this 5' cap is essential for the stabilization of the mRNA, protecting it from exonucleolytic degradation and enhancing its translational efficiency [2].

The second essential process in mRNA post-transcriptional modification is the removal of non-coding sequences from pre-mRNA, known as splicing, which is facilitated by a ribonucleoprotein complex called the spliceosome. The splicing process is orchestrated by the spliceosome, a complex assembly composed of multiple protein subunits and small nuclear RNA (snRNA) molecules. The primary function of the spliceosome is to recognize and excise introns, the non-coding regions of the precursor mRNA, while ligating the exons, which are the coding sequences, to form a mature mRNA transcript. The precise and regulated removal of introns by the spliceosome is essential for the generation of a functional mRNA transcript that accurately reflects the coding potential of the gene [3]. Alternative splicing further adds complexity to gene expression, enabling a single gene to give rise to multiple mRNA isoforms, which in turn can produce distinct protein variants with diverse functional properties. This mechanism plays a pivotal role in cellular differentiation, development, and adaptation, and its dysregulation is often implicated in various diseases, including cancer and neurodegenerative disorders.

The third major modification in the production of mature mRNA is

polyadenylation, a process involving the addition of a poly(A) tail to the 3' end of the mRNA molecule. This poly(A) tail is an essential element in the regulation of mRNA stability, with extensive research indicating that mRNA molecules possessing longer poly(A) tails exhibit increased stability and prolonged half-lives [4]. The addition of poly(A) does not involve a template and is carried out by poly(A) polymerases. The fully processed and mature mRNA is then exported from the Nucleus to the cytoplasm, where it can either be translated into protein products or targeted for degradation. This intricate sequence of modifications ensures that only fully functional and properly regulated mRNA molecules are available for translation, thereby maintaining cellular integrity and protein homeostasis.

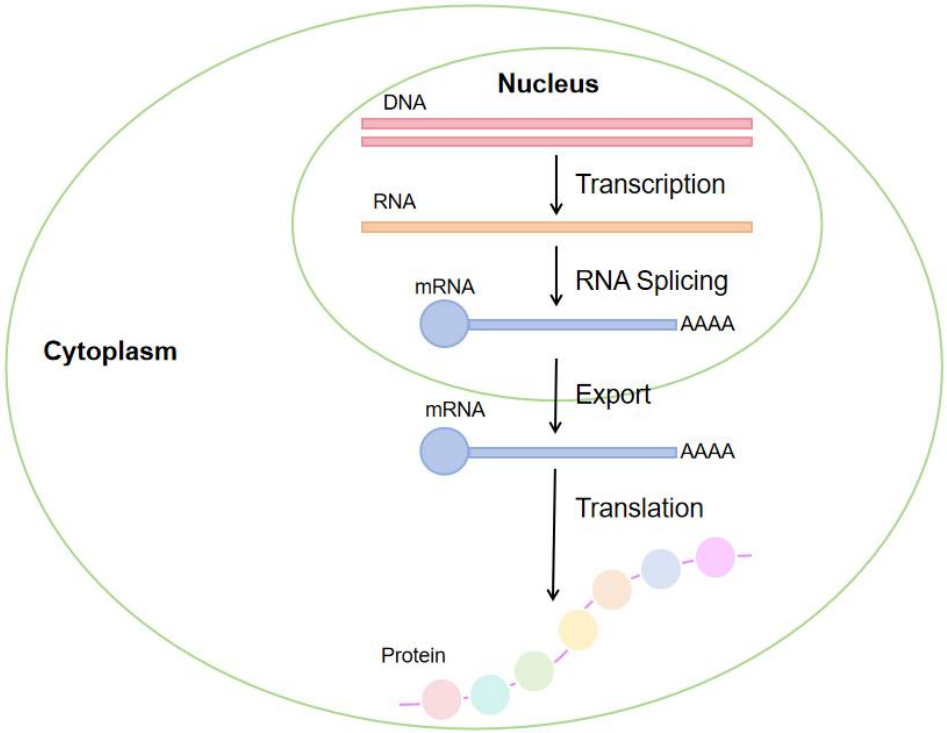


Figure1.1 Gene expression in eukaryotic cells

1.2 mRNA degradation

Messenger RNA (mRNA) is a single-stranded ribonucleic acid (RNA) molecule that carries the coding information from DNA, serving as the direct template for protein synthesis. It acts as the intermediary between the genetic instructions in DNA and the cellular machinery responsible for translating those instructions into functional proteins. It also serves as an intermediary in the central dogma of molecular biology, facilitating the transfer of genetic information from DNA to proteins by being translated by ribosomes during the process of protein synthesis. Consequently, the intracellular concentration and composition of proteins are influenced by the abundance and stability of mRNA molecules within the cell. Therefore, the regulation of mRNA decay, which encompasses the mechanisms controlling mRNA degradation and turnover, constitutes a critical processes in modulating the overall levels of gene expression. Regulating mRNA decay rates influences the duration and extent to which mRNA is available for translation, thereby directly impacting the quantity and quality of proteins synthesized within the cellular environment. Effective management of mRNA stability ensures appropriate protein expression, which is essential for maintaining cellular function and responding to various physiological demands [5, 7].

In order to respond to ever-changing environmental conditions, the regulation of gene expression enables cells to dynamically adjust protein production in response to external and internal stimuli. This adaptive capacity is particularly essential in single-celled organisms, which must

continuously modulate their gene expression profiles to thrive in diverse and often hostile environments. While in complex multicellular organisms, gene regulation is more refined due to the need for cell specialization and tissue coordination. Besides responding to external stimuli, they must maintain internal balance and regulate development, requiring more complex mechanisms like epigenetic modifications and alternative splicing. Similarly, in multicellular organisms, precise regulation of gene expression is crucial for normal development, cellular differentiation, and overall physiological health. Cells adapt to both significant environmental changes and minor fluctuations by finely tuning gene expression, which is essential for survival, growth, and homeostasis. mRNA decay often involves RNA-binding proteins or miRNA-mediated silencing complexes that recognize specific sequence elements in the 3' untranslated region of mRNA, recruiting the RNA degradation machinery to modulate mRNA stabilityas well as regulate translation and, consequently, protein production [6, 8].

Complex	Protein Names in Yeast	Human Homologs	Function
5′→3′ decay			
5′-End Processing Enzymes	Dcp1	DCP1A, DCP1B	Decaping complex
	Dcp2	DCP2	
	Xrn1(Kem1)	XRN1	
Lsm Complexes	Lsm 1-7	Lsm 1-7	Required for decapping; heptameric
Deadenylation			
Pan2-Pan3	Pan2-Pan3	Pan2-Pan3	Involved in the first wave of deadenylation, during which approximately half of the poly(A) tail is removed through a slow, distributive activity.
Ccr4-Not Complex	Ccr4	CNOT6/CNOT6L	Main deadenylase in Saccharomyces cerevisiae; inhibited by PABP
	Caf1	CNOT7, CNOT8	
	Not1	CNOT1	
	Caf40	CNOT9	
	Mot2	CNOT4	
	Not5	CNOT3	
	Not2	CNOT2	

Table1.1 Some examples of human homologs of mRNA decay (5′→3′) factors and the enzymes and their yeast orthologues

1.3 The pathway of mRNA degradation

In general, RNA molecules undergo degradation at the end of their functional lifespan. Since rRNAs are embedded in assembled ribosomes, they tend to be exceptionally stable: under normal growth conditions ribosomes are long-lived and rRNA constitutes the vast majority of total cellular RNA (on the order of ~80 %). By contrast, many other RNA species – including excised intron lariats and spacer fragments generated during pre-rRNA or pre-mRNA processing – as well as most messenger RNAs (mRNAs) are comparatively short-lived [71, 72]. In general, RNA is degraded at the end of its "useful life": rRNAs have very long lifetimes, whereas intronic and spacer RNAs are degraded almost immediately, and mRNA half-lives are tightly controlled. Accordingly, unstable or aberrant RNAs are rapidly detected and routed into decay pathways. Cells employ dedicated surveillance mechanisms to recognize defective or unneeded transcripts and ensure their prompt elimination. For example, the nonsense-mediated decay (NMD) pathway targets mRNAs harboring premature stop codons, non-stop decay (NSD) eliminates transcripts lacking an in-frame termination codon, and no-go decay (NGD) resolves messages on which ribosomes have stalled. Each of these translation-coupled quality-control pathways ultimately channels the aberrant mRNA into the general RNA decay machinery (for example the exosome and cytoplasmic exonucleases) for rapid exonucleolytic degradation. [72, 73] Intracellular RNA degradation is mediated by three principal classes of ribonucleases (RNases): endonucleases, which cleave RNA internally; 5′ exonucleases, which hydrolyze RNA from the 5′ end; and 3′ exonucleases,

which degrade RNA from the 3′ end. The majority of genomes encode numerous RNases, often with overlapping specificities, reflecting the redundancy characteristic of RNA degradation systems. This redundancy suggests that multiple RNases can target the same RNA species, enhancing the overall efficiency and robustness of the degradation pathway.

Many enzymes and cofactors involved in RNA processing and degradation exhibit multifunctionality. For instance, in Saccharomyces cerevisiae, the 5′ exonuclease Rat1 and the 3′ exonuclease within the exosome complex—a conserved multi-protein complex responsible for the 3′-to-5′ degradation of various RNA species—facilitate the breakdown of RNA transcripts produced by RNA polymerases I, II, and III. In addition to degradation, these enzymes participate in RNA processing events that generate mature RNA ends and stabilize RNA species. Similarly, in bacteria, these factors are involved in both RNA maturation and the degradation of stable RNAs and mRNAs [42]. The multifunctional roles of ribonucleases underscore the critical importance of mechanisms that specifically recognize and target aberrant RNAs and RNA-protein complexes. This specificity is frequently mediated by cofactors, many of which have been identified.

ATP-dependent RNA helicases constitute a large family of proteins integral to nearly all RNA processing and degradation pathways. Both the eukaryotic exosome complex and the bacterial degradosome exhibit 3′ exonuclease and endonuclease activities and interact with helicase family members, such as Mtr4 and Ski2 in the exosome, and RhlB in the degradosome [43]. RNA helicases can undergo significant conformational changes upon ATP binding and hydrolysis, facilitating translocation along nucleic acids, potentially

unwinding secondary structures, or displacing bound proteins and/or RNA. Alternatively, helicases may function as positional markers, remaining anchored while recruiting or signaling the degradation machinery.

All exonucleases encounter challenges when attempting to initiate degradation near stable stem-loop structures, which can impede their access to RNA ends [42]. To overcome this obstacle, certain RNA polymerases—such as the non-canonical poly(A) polymerases in the TRAMP complex—add short, unstructured poly(A) tails to the 3′ end of the RNA. These tails help unfold secondary structures and create a more accessible entry point for exonucleases. This strategy appears to be conserved across eukaryotes, suggesting an evolutionarily preserved mechanism for enhancing RNA degradation efficiency. [74] In Saccharomyces cerevisiae and Schizosaccharomyces pombe, the TRAMP polyadenylation complex plays a critical role as a cofactor of the exosome complex [44, 45]. TRAMP marks aberrant nuclear RNAs with short poly(A) tails, which in turn recruit the exosome for degradation. Homologous complexes in human cells also perform similar functions by adding poly(A) or A-rich tails to defective RNAs, promoting their elimination [46].

mRNA decay is a tightly regulated process that involves several key steps to ensure precise control of gene expression. The major stages of mRNA decay are as follows [51, 52, 53]: The first step of mRNA decay typically involves the removal of the poly(A) tail, a process known as deadenylation (Figure 1.2). This process is accomplished by a deadenylase complex, such as the Ccr4-Not complex or PARN, which shortens the poly(A) tail. Loss of this tail destabilizes the mRNA, leading to its subsequent decay. Following

deadenylation, two mechanisms can degrade mRNA: 5'→3' decay or 3'→5' decay following decapping. In the decapping pathway, the Lsm1-7 complex binds to the 3' end of the mRNA transcript and induces decapping by the DCP1-DCP2 complex, resulting in the removal of the 7-methylguanylate (m^7G) cap structure at the 5' end of the mRNA. This renders the mRNA susceptible to degradation by the 5'→3' exoribonuclease XRN1. In the 3′-5′ degradation pathway, the exosome complex (usually assisted by the Ski complex in eukaryotes or the degradosome in bacteria) degrades the mRNA from the 3′ end after deadenylation, and the remaining cap structure is hydrolyzed by the scavenger decapping enzyme DcpS.

As mentioned above, deadenylation is the first step in the decay of a large number of mRNAs, and several characterized eukaryotic deadenylases have been discovered [49, 50], including PAN2-PAN3, CCR4-NOT, and PARN (poly(A)-specific ribonuclease), each with unique properties. Notably, the structure-function analysis of the Ccr4-Not complex is still incompletely characterized. Table 1 lists some mRNA decay factors and the enzymes that perform each step in the decay.

Moreover, the Ccr4-Not complex interacts with a variety of other factors that inhibit translation or induce mRNA decapping when recruited to the translation template, thereby initiating mRNA degradation pathways. It collaborates with the decapping complex to link deadenylation with decapping, facilitating a seamless transition to mRNA degradation. Additionally, it interacts with transcription regulators and chromatin-modifying proteins, influencing gene expression beyond mRNA decay. The complex carries deadenylases that are central to the degradation

of mRNA, playing a key role in general mRNA decay following normal translation termination, nonsense-mediated decay (NMD) of aberrant transcripts, and the targeted degradation of mRNA in response to external signals that alter gene expression programs. As such, the Ccr4-Not complex is a major regulator of eukaryotic gene expression, orchestrating multiple layers of post-transcriptional and translational control [10, 11]. Its activation is further associated with a range of downstream effects, including the methylation of histones, which further links its function to the broader epigenetic regulation of gene expression.

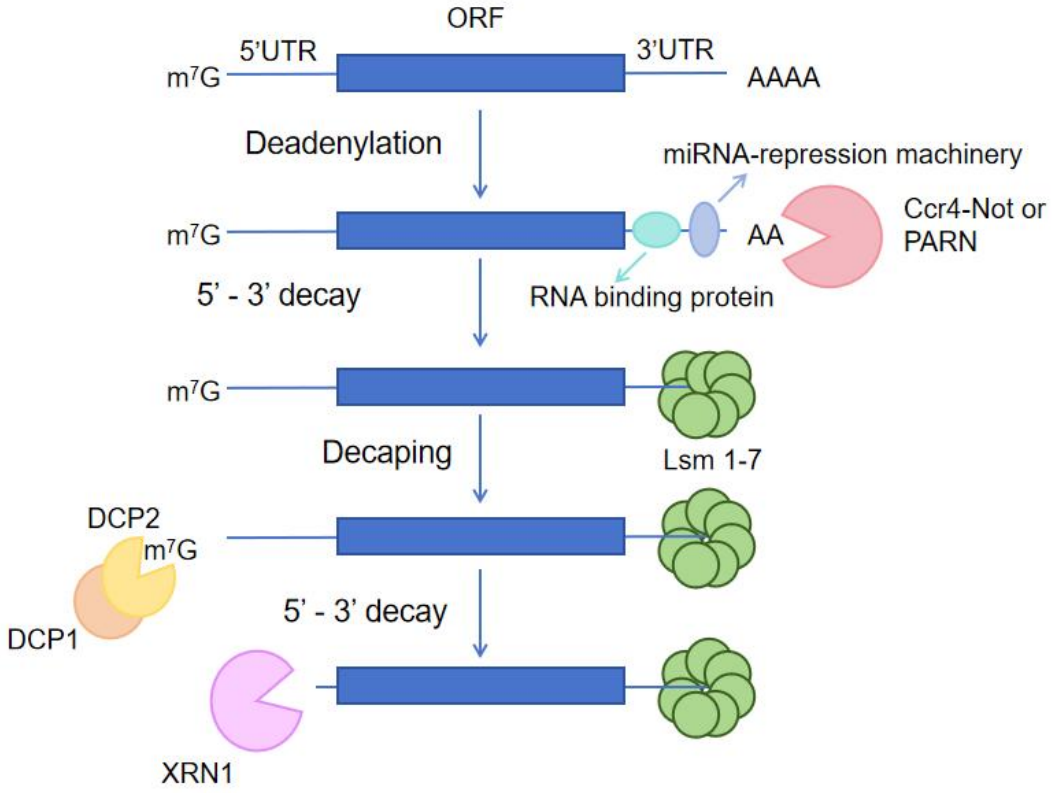


Figure1.2 Main pathway of deadenylation-dependent 5'-3' mRNA decay. Most mRNAs decay via a deadenylation-dependent pathway. This process is

accomplished by a deadenylase complex (represented in this figure by the Ccr4-Not complex or PARN), which shortens the poly(A) tail. Loss of this tail destabilizes the mRNA, leading to its subsequent decay. In the 5'-3' direction of degradation, after deadenylation, the mRNA undergoes decapping, where the 7-methylguanylate (m^7G) cap structure at the 5' end of the mRNA is removed. This step is mediated by a decapping enzyme complex, including the DCP1-DCP2 complex. Removal of the cap exposes the mRNA to degradation by exonucleases and further promotes its degradation. Once the mRNA is deadenylated and decapped, it is degraded by exonucleases.

1.4 Overview of Ccr4-Not complex

1.4.1 The Ccr4-Not Complex in Eukaryotic Gene Regulation

The Ccr4-Not complex is a large, essential, and evolutionarily conserved L-shaped multi-subunit assembly in eukaryotes, playing a critical role in various cellular processes. As a multifunctional complex, it serves as the principal regulator of deadenylation-dependent mRNA turnover, a key step in the post-transcriptional regulation of gene expression. The Ccr4-Not complex exerts its function by catalyzing the shortening of the poly(A) tail of messenger RNA (mRNA), thereby promoting deadenylation and subsequent mRNA decay. In yeast cells, in addition to its role in mRNA turnover, the complex NOT4 also exhibits ubiquitin-protein ligase activity, contributing to the regulation of protein stability and degradation in the cytoplasm. Unlike in yeast, where Not4 is a constitutive CCR4-NOT subunit, human CNOT4 (and in other metazoans) is not stably integrated into the core complex. instead it is a

loosely associated, non-constitutive factor that binds transiently via specific interfaces (for example, a conserved CNOT9/Caf40-binding motif in its C-terminal region) [75, 76].

Recent advances in research have progressively unveiled the diverse physiological roles mediated by the Ccr4-Not complex, shedding light on its involvement in a range of essential cellular functions. For example, Ccr4-Not subunits interact with RNA polymerase II and chromatin modifiers to fine-tune transcription (both initiation and elongation). In the cytoplasm the complex drives mRNA turnover: its catalytic deadenylase subunits (homologous to yeast Ccr4p and Caf1p, known as human CNOT6/CNOT6L and CNOT7/CNOT8) shorten mRNA poly(A) tails to trigger decay. Likewise, Ccr4-Not directly impacts protein ubiquitination: the Not4/CNOT4 subunit contains a RING E3 ligase domain that ubiquitylates specific protein targets, linking the complex to regulated protein degradation. Beyond these roles, Ccr4-Not influences cell-cycle control and proliferation. For instance, mammalian Ccr4-Not regulates expression of IGFBP5, an insulin-like growth factor binding protein that enforces p53-dependent growth arrest and senescence. [77, 78, 79]

Furthermore, numerous studies have identified associations between mutations in components of this micro-complex and the onset of various pathological conditions, emphasizing its clinical relevance [5, 6, 7, 9, 41]. In the nervous system, de novo variants in the scaffold subunit CNOT1 cause a neurodevelopmental syndrome characterized by developmental delay, intellectual disability, hypotonia and speech deficits. In the immune/cancer context, CNOT3 has emerged as a tumor suppressor: exome studies found

CNOT3 mutations in ~8% of adult T-cell acute lymphoblastic leukemia cases, and loss of CNOT3 promotes leukemia in model systems. Beyond cancer, CCR4-NOT abnormalities have been implicated in inflammatory and developmental disorders [79, 80, 81, 82]. Given the complex's role in cellular homeostasis and its implication in disease, understanding the structural characteristics and functional mechanisms of the Ccr4-Not complex is of paramount importance. Elucidating these aspects not only enhances our comprehension of its regulatory functions but also holds potential for the development of therapeutic strategies targeting diseases associated with its dysregulation.

In the cytoplasm, the Ccr4-Not complex plays a multifaceted role in regulating mRNA translation and maintaining protein synthesis fidelity. It associates with actively translating ribosomes (polysomes) and localizes to RNA granules, such as stress granules and processing bodies (P-bodies), where it modulates mRNA fate in response to cellular conditions. One of the key functions of the Ccr4-Not complex is its involvement in co-translational quality control [83]. The complex monitors ribosome progression during translation and responds to translational stalling events. This interaction facilitates the recruitment of the Ccr4-Not complex to the stalled ribosome, leading to the deadenylation and subsequent decay of the problematic mRNA, thereby preventing the accumulation of defective proteins [84, 85]. Additionally, the Not4 subunit, which possesses E3 ubiquitin ligase activity, ubiquitinates specific ribosomal proteins, such as eS7, upon detecting stalled translation. This ubiquitination event serves as a signal for the recruitment of downstream quality control factors that can resolve the stalled ribosome or target the aberrant nascent polypeptide for degradation [85]. Through these

mechanisms, the Ccr4-Not complex ensures the integrity of protein synthesis by coupling the detection of translational errors with appropriate mRNA and protein quality control responses. Its role in monitoring translation and enforcing quality control underscores its significance in maintaining cellular homeostasis, especially under stress conditions that can disrupt normal translational processes.

1.4.2 Structure of Ccr4-Not complex

Thanks to cryo-EM and X-ray technique, previous studies have demonstrated that the Ccr4-Not complex is structured around a central L-shaped flat protein, Not1/CNOT1, with other subunits associated with functional modules assembling onto this scaffold. In human cells, the Ccr4-Not complex contains two exonucleases, CNOT7/CNOT8 and CNOT6/CNOT6L, along with the non-catalytic proteins CNOT1, CNOT2, CNOT3, and CNOT9, all of which are conserved across eukaryotes. The Ccr4-Not complex comprises two enzymatic modules. One of the modules is a deadenylase, consisting of Ccr4/CNOT6 and Caf1/CNOT7. This big module binds to the central MIF4G domain of Not1/CNOT1 [12, 13] (Figure 1.3). The other module is a ubiquitination unit formed by the Not4/CNOT4 RING E3 ligase, which associates with the C-terminal region of Not1/CNOT1, proximal to the docking site of the Not5-Not2/CNOT3-CNOT2 heterodimer. Interestingly, the association of Not4 with the complex is stable in yeast, but CNOT4 is not a core component of mammalian Ccr4-Not [86]. Therefore, the core Ccr4-Not complex consists of a minimum of six structured proteins.

The Ccr4-Not complex comprises several well-characterized core

subunits—CNOT1, CNOT2, CNOT3, CNOT6/6L, CNOT7/8, and CNOT9—that collectively orchestrate mRNA deadenylation and decay (Figure 1.3). Beyond these essential components, the complex interacts with auxiliary proteins that, while not indispensable for its assembly, are crucial for its regulatory functions. Among these, the BTG/TOB family of proteins—including BTG1, BTG2, TOB1, and TOB2—serve as pivotal modulators of mRNA turnover and cellular proliferation. BTG/TOB proteins are characterized by a conserved N-terminal BTG domain that facilitates direct interaction with the Caf1 (CNOT7/8) catalytic subunits of the Ccr4-Not complex. This interaction enhances the deadenylase activity of Caf1, promoting the removal of poly(A) tails from target mRNAs and thereby accelerating mRNA decay. Additionally, BTG/TOB proteins engage with the cytoplasmic poly(A)-binding protein PABPC1 through specific motifs—PAM2 in TOB1/TOB2 and Box C in BTG1/BTG2—effectively bridging the Ccr4-Not complex to polyadenylated mRNAs. This tripartite interaction not only facilitates targeted mRNA deadenylation but also plays a role in regulating translation and maintaining mRNA stability [14, 19]. The BTG/TOB proteins act as antiproliferative agents. Their overexpression has been shown to inhibit cell cycle progression, an effect that is contingent upon their interaction with the Caf1 subunits. Conversely, the absence or downregulation of BTG/TOB proteins can lead to enhanced cellular proliferation, underscoring their role in cell cycle control. Through these mechanisms, BTG/TOB proteins extend the functional repertoire of the Ccr4-Not complex, linking mRNA decay processes to broader cellular outcomes such as growth regulation and response to proliferative signals [15].

Beyond these interactions, various RNA-binding proteins play significant

roles in the regulation of the CNOT complex. Proteins such as Nanos2, Pumilio, Smaug, Tristetraprolin, and those containing YTH domains act as adapter proteins, engaging with the CNOT complex [14]. These RNA-binding proteins often function through recognition of specific RNA sequences or structures, thereby guiding the Ccr4-Not complex to its target mRNAs [16,17,18]. Among these, particularly noteworthy group of RNA-binding proteins includes those containing the YTH domain. The YTH domain is known for its role in recognizing and binding N^6-methyladenosine (m^6A)-modified RNA. This modification is critical for various aspects of RNA metabolism, including splicing, translation, and degradation. Among the YTH domain-containing proteins, YTHDF2 has been extensively studied for its role in mRNA degradation. YTHDF2 selectively binds to m^6A-modified mRNAs and facilitates their decay by recruiting the CCR4–NOT deadenylase complex. Specifically, the N-terminal region of YTHDF2 interacts directly with the SH domain of the CNOT1 subunit within the CCR4–NOT complex. This interaction is crucial for the accelerated deadenylation of m^6A-containing mRNAs, leading to their subsequent degradation. The deadenylation process is primarily mediated by the CAF1 and CCR4A/B subunits of the CCR4–NOT complex [87]. The m^6A modification serves as a dynamic marker that influences mRNA stability. By recruiting YTHDF2 and the CCR4–NOT complex, m^6A marks promote the rapid turnover of specific mRNAs, thereby fine-tuning gene expression in response to various cellular stimuli. This mechanism underscores underscores the intricate network of protein-RNA interactions that fine-tune gene expression and highlights the adaptability of Ccr4-Not complex in response to different regulatory signals [25, 27].

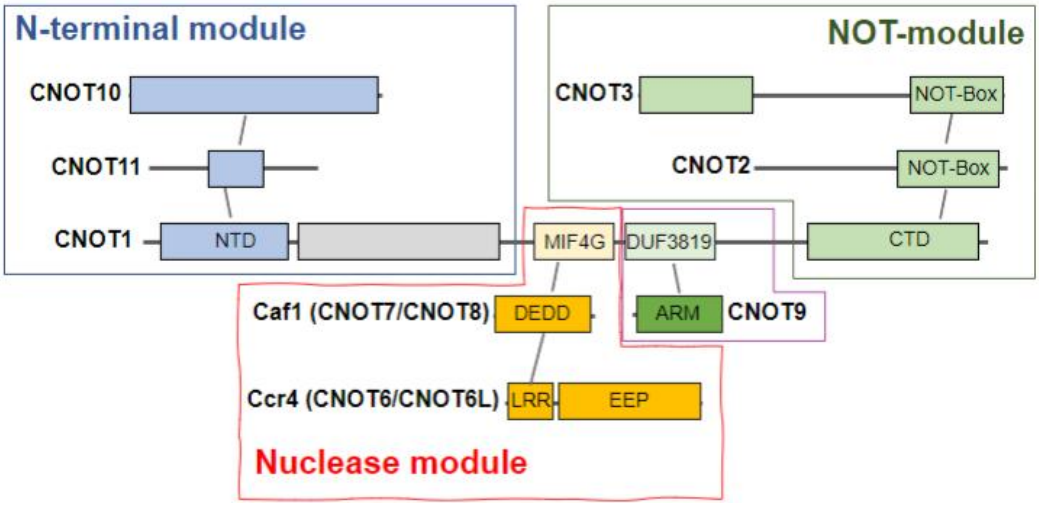


Figure 1.3 Architecture of the Core Ccr4-NOT complex

1.4.3 The Ccr4-Not complex in mRNA degradation and gene expression regulation

Based on all the information mentioned above, recent investigations have illuminated the Ccr4-Not complex's involvement in several mRNA decay pathways, including general mRNA deadenylation, nonsense-mediated mRNA decay (NMD), targeted mRNA decay, and deadenylation-independent decapping [86]. The degradation of mRNA typically initiates with the shortening of the poly(A) tail, a process mediated by the deadenylase activity of the Ccr4-Not complex. Following deadenylation, the removal of the 5' cap structure by decapping enzymes exposes the mRNA to exonucleolytic degradation. In one pathway, the exposed mRNA is degraded in the 5' to 3' direction by the exonuclease XRN1. Alternatively, in another pathway, after deadenylation, the mRNA body is degraded from the 3' to 5' direction by cytoplasmic exosomes.

Based on the insights mentioned before, further research is necessary to elucidate the precise mechanisms governing the activation and recruitment of the Ccr4-Not complex in the context of mRNA decay, as understanding these processes is essential for a comprehensive grasp of gene expression regulation in eukaryotic cells . A comprehensive understanding of the diverse roles played by the Ccr4-Not complex in mRNA decay pathways and related processes is essential for elucidating its overall function in RNA metabolism. Moreover, this understanding has broader implications for our knowledge of cellular function and the intricate regulation of gene expression, highlighting the complex's significance in maintaining cellular homeostasis and preventing the accumulation of defective mRNAs [11].

1.5 Nuclease module: CNOT1(MIF4G)-CNOT6L-CNOT7

The deadenylase module attaches to the central MIF4G region of CNOT1, which includes HEAT repeats [88]. Consequently, CNOT7/8 associates with CNot1 through pre-formed structural motifs comprised of conserved hydrophobic residues, thus granting the RNA molecule full access to the active site of CNOT7/8. The leucine-rich repeat (LRR) domain of CNOT6/6L binds to CNOT7/8, which interacts with CNOT6/6L via a surface characterized by elongated loops and α-helices [13, 88]. This binding event induces local conformational changes within CNOT7/8, altering its structure compared to its unbound state. These conformational changes are crucial for optimizing the catalytic activity of CNOT7/8 and its interaction with RNA substrates. The structural rearrangements facilitate the proper positioning of

the RNA molecule in the active site of CNOT7/8, thereby enhancing the efficiency and specificity of the deadenylation process. Additionally, these changes may impact the overall stability and functionality of the Ccr4-Not complex, influencing its role in mRNA degradation and gene expression regulation [89].

CNOT7 is a catalytic subunit of the Ccr4-Not complex and functions as a deadenylase, which means it specifically removes poly(A) tails from mRNA molecules. By shortening the poly(A) tail through deadenylation, CNOT7 initiates mRNA decay, leading to reduced mRNA stability and eventual degradation. Although CNOT7 itself is a non-specific RNA-binding protein in the sense that it can target a variety of transcripts, the specificity of its action on particular mRNA molecules is provided by sequence-specific RNA-binding proteins that interact with the Ccr4-Not complex. For instance, proteins like BTG/TOB1 bind to poly(A)-binding protein (PABPC1), which in turn recruits CNOT7 to specific mRNAs [26, 90]. This interaction acts as an adapter mechanism, ensuring that CNOT7 deadenylates the correct polyadenylated transcripts.

Additionally, CNOT1 functions as a scaffold within the Ccr4-Not complex, organizing the various components, including CNOT7, to form a functional deadenylation machinery [9]. This scaffold role of CNOT1 ensures proper assembly and stability of the complex, allowing CNOT7 and other subunits to carry out their roles in deadenylation and mRNA decay.

Extensive research has elucidated that CNOT7/8 and CNOT6/CNOT6L perform distinct yet partially overlapping functions within cellular systems. In vertebrates, the Ccr4 nuclease is represented by two paralogs, CNOT6 and

CNOT6L. Both paralogs possess the catalytic endonuclease-exonuclease phosphatase (EEP) domain as well as the leucine-rich repeat (LRR) domain, which facilitates interactions with the CNOT7/8 subunit. [26]

The CNOT7/8 subunit further engages with the BTG/TOB proteins through a dedicated binding interface distinct from those interacting with the MIF4G domain of CNOT1 or the LRR domain of CNOT6/CNOT6L. The BTG/TOB proteins, which include BTG1, BTG2, and TOB1, are known for their role in regulating cell proliferation and differentiation. They interact with the CNOT7/8 complex via the BTG domain, contributing to the regulation of mRNA stability and degradation [89].

The BTG/TOB proteins act as modulators of the Ccr4-Not complex by influencing its deadenylation activity and facilitating the recruitment of additional cofactors. The interaction between BTG/TOB proteins and CNOT7/8 thus integrates mRNA decay with broader regulatory networks governing cellular processes such as growth, stress responses, and differentiation. This interaction highlights the complex interplay between different components of the mRNA decay machinery and their role in maintaining cellular homeostasis. (Figure 1.4)

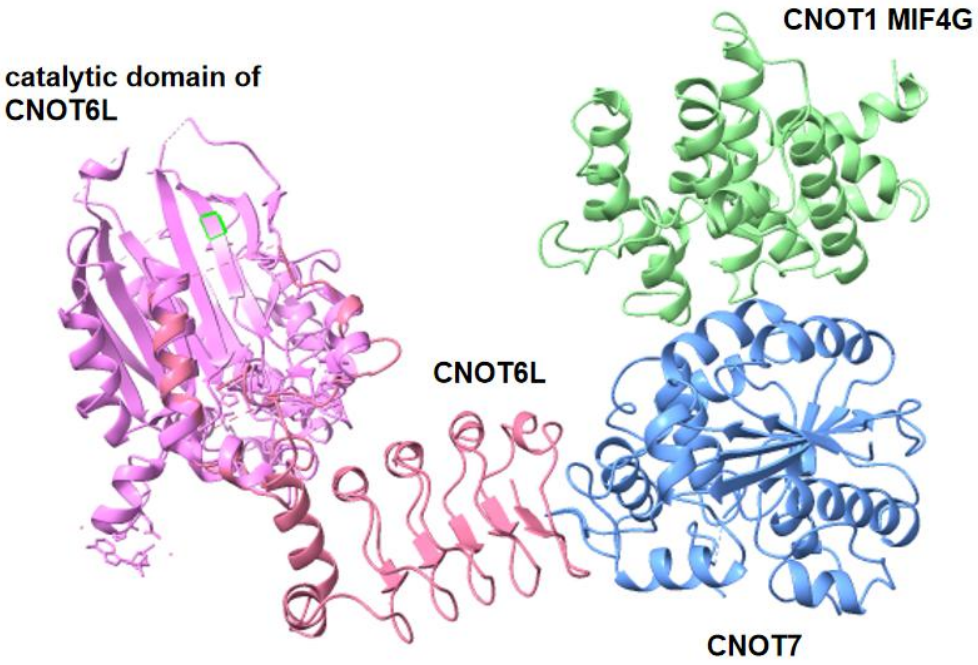


Figure 1.4 Protein structure of different part of the human core CCR4-Not complex. PDB code: 7VOI. Structure of the nuclease module: CNOT1(MIF4G)-CNOT6L-CNOT7. Structure of CNOT7 (blue) bound to CNOT1 MIF4G (green) while CNOT6L (pink) with the catalytic domain (purple) bound to CNOT7.

1.6 CNOT1(MIF4G-DUF3819)-CNOT9 module

The CNOT9 (Caf40) subunit associates with a central coiled-coil domain of CNOT1, referred to as CN9BD or DUF3819. This DUF3819 domain is positioned adjacent to the MIF4G (middle portion of eIF4G, typically composed of 80-90 amino acids) region of CNOT1, which provides a structural platform facilitating the binding of CNOT9 (RQCD1/Caf40) and

plays a crucial role in maintaining the structural integrity and proper functioning of the Ccr4-Not complex (Figure 1.5). [16, 91]. CNOT9 is capable of mediating interactions with various RNA-binding factors, such as GW182 protein (which is miRNA-associated protein), TTP protein (which is ARE-binding protein), and Roquin protein (which is stem-loop binding protein) [17, 18]. Previous studies have established that the central regions of the CNOT1 and CNOT9 subunits enhance the functionality of the Ccr4-Not complex by stimulating the enzymatic activity of its nuclease module [10]. The aforementioned domains (CNOT1 and CNOT9) constitute the central module of the reconstituted human Ccr4-Not complex. This central module comprises the central region of CNOT1, which includes the MIF4G and DUF3819 domains, the CNOT9 subunit, and a trimeric nuclease module consisting of BTG2, Caf1 (CNOT7), and Ccr4 (CNOT6).

CNOT9, an essential subunit of the Ccr4-Not complex, plays a crucial role in the recruitment of this multifunctional machinery to specific mRNA targets for degradation. It achieves this by interacting with various sequence-specific RNA-binding proteins, which guide the complex to particular transcripts. These RNA-binding proteins often contain conserved peptide motifs, such as tryptophan-rich (W) or proline-rich sequences, that are recognized by CNOT9, facilitating the association of the Ccr4-Not complex with its mRNA targets. For example, Tristetraprolin (TTP) binds to AU-rich elements in mRNAs and interacts with CNOT9 to recruit the Ccr4-Not complex, thus promoting the degradation of specific transcripts [3, 4]. This interaction highlights CNOT9's role as a mediator between the Ccr4-Not complex and RNA-binding proteins, ensuring precise mRNA targeting for decay.

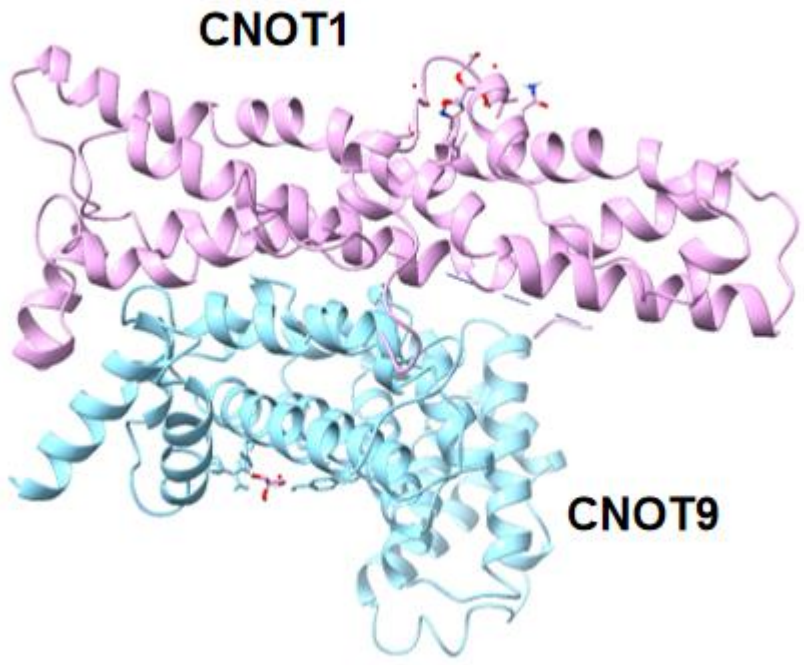


Figure 1.5 Protein structure of different part of the human core CCR4-Not complex. PDB code: 4CT6. Structure of CNOT1(MIF4G-DUF3819)-CNOT9 module. CNOT9 (light blue) bound to the CNOT1 (pink).

1.7 NOT-module: CNOT1 with the NOTboxes of CNOT3 and CNOT2

The C-terminal domain of CNOT1 acts as a rigid scaffold, comprised of two vertically aligned HEAT-like repeats. The CNOT2-C and CNOT3-C subunits

form a heterodimer via their SH3-like NOT-box domains (Figure 1.6). This heterodimer is stabilized by an interlocked arrangement of peptide regions lacking defined secondary structure and anchored to the surface of CNOT1. These assembled peptides form an extended interaction interface that is tailored to each binding surface [23]. Mutations in this specific interface and changes in the ratio of endogenous proteins can disrupt the complex assembly and alter the mRNA degradation process.

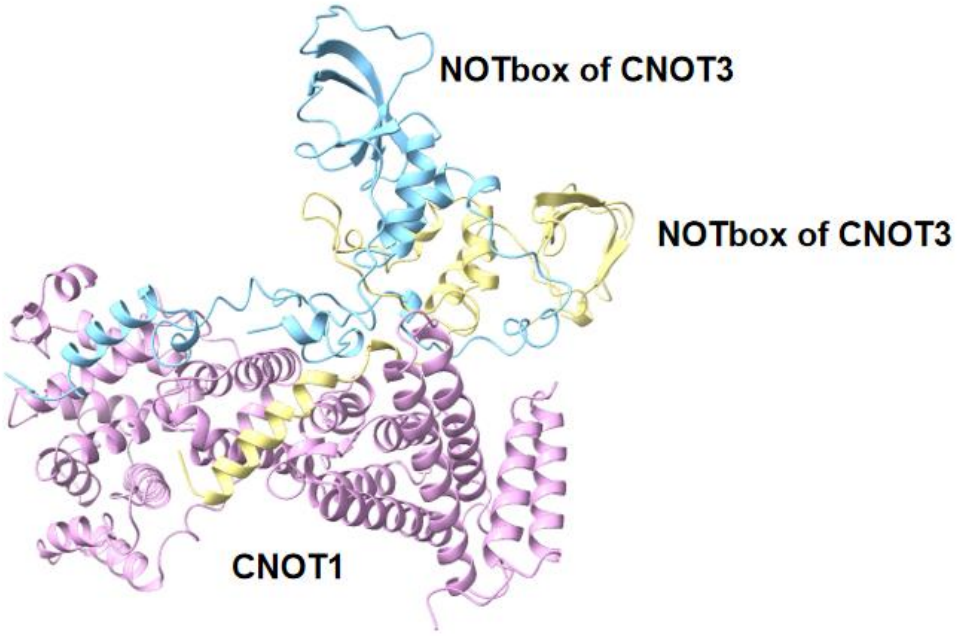


Figure 1.6 Protein structure of different part of the human core CCR4-Not complex. PDB code: 4C0D. Structure of the NOT-module. NOTbox of CNOT3 (yellow) and NOTbox of CNOT2 (blue) bound to the CNOT1 (pink).

1.8 YTH domain Proteins

1.8.1 Overview of YTH domain

As mentioned above, various RNA-binding proteins play important roles in regulating the CNOT complex. Among them, proteins containing the YTH domain act as adaptor proteins and bind to the CNOT complex. Within cellular systems, RNA molecules undergo a multitude of post-transcriptional modifications, which play an essential role in regulating their function and stability. Analogous to the well-characterized modifications of proteins and DNA, these RNA modifications represent sophisticated regulatory mechanisms that orchestrate RNA behavior and are integral to a wide array of essential biological processes. Among these modifications, N^6-methyladenosine (m^6A) is recognized as the most prevalent internal modification across various classes of eukaryotic RNA, including messenger RNA (mRNA), transfer RNA (tRNA), ribosomal RNA (rRNA), and long non-coding RNA (lncRNA). The extensive research conducted in recent years has revealed that m^6A RNA methylation is intricately involved in all stages of the RNA lifecycle, including RNA processing, nuclear export, translation regulation, and RNA degradation. These findings underscore the multifunctional nature of m^6A, which exerts both positive and negative effects on RNA metabolism, thereby influencing gene expression at multiple levels [26].

Essentially, m^6A modifications are recognized by specific proteins containing the YT521-B homology (YTH) domain. These YTH domain-containing proteins act as molecular readers that interpret m^6A marks and subsequently guide a variety of protein complexes to modulate RNA signaling pathways. These pathways include RNA metabolism, splicing, folding, and translation,

demonstrating the regulatory versatility of m^6A modifications. Moreover, the dynamic and reversible nature of m^6A modification is achieved through the interplay between methyltransferases, which add methyl groups, and demethylases, which remove them, thereby allowing for a finely tuned regulation of mRNA function. The YTH domain family proteins (Figure 1.7) are widely distributed across human tissues and are recognized as key mediators of the functional outcomes associated with m^6A-modified RNA [46]. These reader proteins include members of the YTH N^6-methyladenosine RNA-binding proteins (YTHDFs) and the YTH domain-containing proteins (YTHDCs). These proteins are instrumental in mediating a variety of critical biological processes through their specific recognition of m^6A modifications [27].

The structural basis for m^6A recognition by the YTH domain lies in its conserved aromatic cage, a specialized binding pocket that enables selective interaction with m^6A-modified RNA [28]. YTH domain proteins identify and decode the information carried by m^6A marks in a methylation-dependent manner. In humans, YTHDC1 and YTHDC2, along with YTHDF1-3, constitute the primary intracellular YTH domain proteins. YTHDC2 is primarily localized within the cytoplasm of meiotic spermatocytes, whereas YTHDF1-3 predominantly engage with m^6A-modified RNA in the cytoplasm [29, 30, 31]. Although these proteins share a common YTH domain, they also possess unique structural domains that confer distinct functional roles within the cell, highlighting the complexity and specificity of m^6A-mediated regulatory mechanisms in RNA biology.

YTHDF1 plays a critical role in promoting the translation of m^6A-modified

RNA. Research has demonstrated that the primary function of YTHDF1 is to enhance the assembly of ribosomes on m^{6}A-modified mRNA and interact with initiation factors, thereby accelerating the translation of these transcripts and regulating translation dynamics. YTHDF1 selectively recognizes and binds to m^{6}A-modified RNA at GRAC motifs (where R represents G or A) while exhibiting specificity by not interacting with other sequences near the m^{6}A sites. This selectivity highlights the precise nature of YTHDF1's binding to its target RNA. Furthermore, YTHDF1 is essential in facilitating both the initiation and elongation phases of m^{6}A RNA translation [29].

In contrast, the primary function of YTHDF2 is to regulate the stability of m^{6}A-modified RNA. YTHDF2 primarily governs the lifespan and stability of target mRNAs. The C-terminal domain of YTHDF2 is responsible for binding to m^{6}A RNA, while the N-terminal domain directs the YTHDF2-mRNA complex to RNA degradation, transitioning m^{6}A-modified RNA from a translational state to one primed for degradation. Additionally, YTHDF2 recruits the Ccr4-Not complex to destabilize and further degrade target mRNA. For instance, YTHDF2 can recognize m^{6}A-modified sites in the 3' untranslated region (UTR) of epidermal growth factor receptor (EGFR) mRNA, promoting its degradation and thereby exerting anti-tumor effects. Moreover, YTHDF2 inhibits hematopoietic stem cell proliferation by degrading mRNAs associated with stem cell self-renewal [32, 33]. It also targets transcripts with high levels of m^{6}A, such as those encoding autophagy-related proteins ATG5 and ATG7, leading to mRNA degradation and subsequent reduction in autophagy and adipogenesis [34]. These findings suggest that YTHDF2 is involved in various biological processes by

destabilizing critical gene transcripts.

YTHDF3 functions to promote both the translation and degradation of m^{6}A-modified RNA. Its role shares similarities with, yet differs from, YTHDF2. YTHDF3 enhances protein expression by recognizing m^{6}A-modified mRNA, and regulating the translation of its own mRNA [35]. Additionally, YTHDF3 interacts with ribosomes, participating in the early stages of translation [36].

YTHDC1 is primarily involved in the regulation of RNA splicing. The YTH domain within YTHDC1 specifically recognizes m^{6}A modifications, with a preference for the G(m^{6}A)C sequence. YTHDC1 can identify m^{6}A-modified chromatin-associated regulatory RNAs (CarRNAs), including promoter-associated RNAs (PaRNAs), enhancer RNAs (eRNAs), and transposable element transcripts (repeat RNAs). This recognition facilitates the decay and processing of CarRNAs, thereby modulating chromatin states and transcription [37].

YTHDC2, possessing a broad spectrum of functions, regulates the stability of m^{6}A-modified mRNA by recognizing m^{6}A marks and recruiting RNA degradation mechanisms [38]. It also bridges m^{6}A-modified mRNA and ribosomes, promoting efficient translation. Among YTH domain-containing proteins, YTHDC2 is distinct in that it includes a helicase domain, which plays an essential role in meiotic processes [39, 40].

Above all, given that research indicates YTHDF2 exerts its function by recruiting the Ccr4-Not complex [32, 33, 34], we have initiated a focused investigation into the specific molecular interactions between the YTHDF2 domain and the Ccr4-Not complex, with particular emphasis on its interaction with the CAF40 protein, a key subunit of the Ccr4-Not

mini-complex.

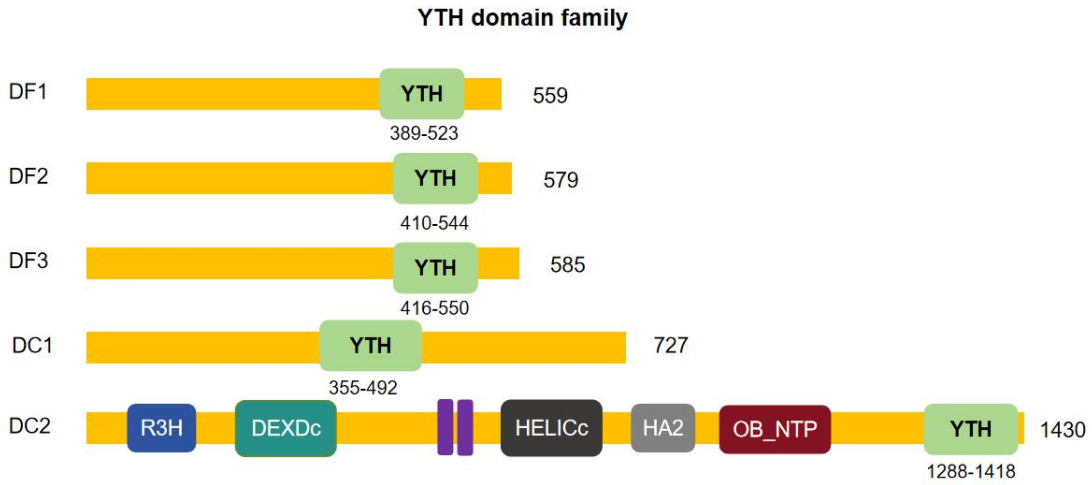


Figure 1.7 Domain structure and functions of human YTH proteins. Schematic representation of domains of human YTH proteins: DF1, DF2, DF3, DC1, and DC2.

1.8.2 Evidence for possible interaction YTHDF2 domain and CNOT9

CNOT9(Caf40) is known to interact with other proteins and peptide motifs beyond YTHDF2. For instance, CNOT9 binds to proteins containing tryptophan (W) motifs or proline-rich sequences, as seen in its interactions with proteins like TTP (Tristetraprolin) and Tob, both of which are involved in promoting mRNA degradation [3, 13]. These additional interactions suggest that CNOT9 acts as a versatile adaptor, recognizing distinct peptide motifs to mediate various regulatory roles within the Ccr4-Not complex. In the context of m^6A-modified mRNA degradation, CNOT9 has been implicated in a direct interaction with YTHDF2, an RNA-binding protein that selectively

recognizes N^6-methyladenosine (m^6A) modifications. YTHDF2 is known to bind m^6A-marked mRNAs, promoting their degradation as part of the m^6A regulatory pathway. It is hypothesized that by engaging CNOT9, YTHDF2 effectively recruits the Ccr4-Not complex to initiate deadenylation, leading to the degradation of these modified mRNAs. Investigating how YTHDF2 fits into this network of interactions could uncover new regulatory nodes within gene expression control and mRNA decay, potentially revealing mechanisms relevant to diseases such as cancer or neurodegenerative disorders. [2, 48, 92]

1.9 Aim and objectives

The structural organization of the core Ccr4-Not complex has been increasingly characterized through advanced imaging techniques and structural studies. Recent years have seen a significant expansion in the understanding of the three-dimensional structure and architectural organization of the Ccr4-Not complex [6, 7, 9], driven by an increasing number of structural studies. However, the precise interconnections and spatial orientation of these modules within the core complex remain areas of active research. One unresolved question is whether the various domains exhibit a specific, defined spacial orientation or if their assembly is more stochastic and flexible. Given the complexity and dynamic nature of the Ccr4-Not complex, understanding the extent of this flexibility is crucial. If the complex exhibits significant conformational flexibility, it is essential to determine the degree of this flexibility and identify any potential conformational constraints that might influence its functional state.

The argument for a preferred orientation stems from the necessity for precise and coordinated interactions between the complex's subunits to perform its regulatory functions effectively and the three dimensional structure of the complex indicates a defined 'L-shaped' structure. A defined three-dimensional arrangement between domains could facilitate stable and specific interactions, optimizing the complex's ability to engage with target RNAs and regulatory factors. Although the overall L-shaped form of the complex determined by previous research [7], to some extent, flexibility may also be required. Such structural constraints would likely enhance the efficiency and specificity of mRNA degradation and gene expression regulation, providing a rationale for the complex to adopt a relatively fixed spatial arrangement despite potential flexibility. Further structural studies and functional assays are needed to elucidate these aspects and confirm whether a preferred orientation exists within the L-shaped architecture of the core Ccr4-Not complex.

Generating different sub-complexes of Ccr4-Not is essential for understanding the flexibility within the Ccr4-Not complex, which is also essential for gaining insights into its functional mechanisms. This thesis aims to contribute to this growing body of knowledge by focusing specifically on the linkers and angular configurations within the multifunctional, multi-subunit Ccr4-Not complex. Special emphasis is placed on the structural and functional relationships of the nuclease module and the NOT module, key components of the complex. The first objective is to express and purify two or three modules connected by flexible linkers for the biochemical and structural characterisation of Ccr4-Not.

The second objective is to find evidence for possible interaction between

YTHDF2 and the CNOT9 subunit of the Ccr4-Not complex. This is an attractive hypothesis because it addresses a key gap in understanding the precise molecular mechanisms by which YTHDF2 recruits the Ccr4-Not complex to degrade m6A-modified RNA. Techniques such as peptides-pull down can be used. By dissecting this interaction at the molecular level, we seek to uncover the underlying mechanisms through which YTHDF2 mediates mRNA decay via the Ccr4-Not complex. Understanding these interactions is essential, as it will provide insights into the broader regulatory networks governing mRNA stability and turnover, and how alterations in these processes may contribute to disease pathogenesis.

Chapter 2

Materials and Methods

Chapter 2. Materials and Methods

2.1 Recombinant DNA techniques

2.1.1 Restriction enzyme digestion of DNA

In order to create compatible ends for cloning and to verify newly created plasmids, DNA plasmid was digested by restriction endonuclease. All restriction enzymes and buffers were supplied by NEB and followed the protocol they provided.To digest the plasmid, 1 μg of plasmid DNA was digested in 50 μl total reaction volume (Table 2.6). The reaction was carried out for 2 - 6 hours at 37°C incubator.

Volume (μL)	Reagent/Component
-	1 μg template DNA
5 (10 % of the total volume)	rCutSmart™ Buffer
2.5	Restriction enzyme Ⅰ
2.5	Restriction enzyme Ⅱ
Up to 50	Nuclease free H_2O

Table 2.6 Standard restriction enzyme digestion reaction

2.1.2 Plasmids

A full-length cDNA sequence encoding human CNOT3 and CNOT2 was generated (from SnapGene) and subsequently subcloned into the bacterial expression plasmid pCDF_Duet1 (pCDFDuet-1 Sequence and Map from SnapGene) following standard procedures. The BamHⅠ and NdeⅠ

restriction sites were used for this subcloning, with CNOT3 and CNOT2 inserted into two cloning site. Besides, A full-length cDNA encoding human CNOT1 and CNOT9 was generated and subcloned into the bacterial expression plasmid pACYCDuet1 (pACYCDuet-1 Sequence and Map from SnapGene) following standard protocols. The BamHⅠ and BglⅡ restriction sites were used for the subcloning, with the long fragment of CNOT1 (containing the C-terminal domain, CNOT1 domain and N-terminal domain) inserted into MCS-1 and CNOT9 into MCS-2. The oligonucleotides required for the site-directed mutagenesis were designed using the Agilent QuikChange Primer Design tool (Table 2.2).

The CNOT1 MIF4G and DUG3819 part is introduced by using the Inverse PCR to delete the C-Terminal domain of CNOT1 in pACYC_Duet1_CNOT1L_CNOT9 plasmid, which means the plasmid only contains N-terminal region, MIF4G and DUF3819 part of CNOT1 in MCS-1 and CNOT9 in MCS-2. And the His-Tag of the original pQE80L-CNOT7 plasmid is deleted by conducting the Inverse PCR to get the plasmid pQE80L-CNOT7-NoHis (Table 2.1).

The existing CNOT9-ARM plasmid (from SnapGene) is amplified using bacterial (Escherichia coli strain DH5α) transformation techniques to achieve high-purity, high-concentration target plasmids (Table 2.1).

Name	**Abbreviation**	**MCS[1] 1**	**MCS[1] 2**
pCDF_Duet1_CNOT3_CNOT2_only Notboxes	pCDF_Duet1	Notbox of CNOT3	Notbox of CNOT2
pACYC_Duet1_CNOT1L_CNOT9	pACYC_Duet1	CNOT1 Long	CNOT9

pACYC_Duet1_CNOT1M-CNOT9	pACYC_Duet1	CNOT1M	CNOT9
pMA-T_CNOT1	pMA-T	CNOT1	-
pQE80L-CNOT7-NoHis	pQE80L	CNOT7-NoHis	-
pQE80L-CNOT9-ARM	pQE80L	CNOT9-ARM	-

Table 2.1 List of plasmids

Name	Sequence (5'-3')
pACYCDuetUP 1	GGATCTCGACGCTCTCCCT
DuetDOWN 1	GATTATGCGGCCGTGTACAA
DuetUP2	TTGTACACGGCCGCATAATC
T7 Terminator	GCTAGTTATTGCTCAGCGG
pMA-T FW	GCCTCTTCGCTATTACGCCAG
pMA-T RV	TTAGGCACCCCAGGCTTTAC
CNOT3NOT-FW	(Phos)CTGACCAAAGAACAGCTG
MCS1-RV	(Phos)GCCCATGGTATATCTCCT
CNOT2NOT-FW	(Phos)ACCAATATTCCGCAAGGTATG
MCS2-RV	(Phos)TGCCATATGTATATCTCCTTC
CNOT1M-OEPCR-FW	ACCGAGCGTATCGTGGAGC
CNOT1M-OEPCR-RV	TGCAACATCATCTTAGGTCGCCCACGCTTGT
CNOT1C-OEPCR-FW	GGGCGACCTAAGATGATGTTGCACAGATTTATGATAAATGCATCAC
del1522	AGCGTGGGCGACCGGATCCGATGATG
del1522-antisense	CATCATCGGATCCGGTCGCCCACGCT

CNOT1L-IPCR-FW	(Phos)GATGATGTTGCACAGATTTATGATAAATGC
CNOT1L-IPCR-RV	(Phos)GGTCGCCCACGCTTGTTT
CNOT1-MC FW	CGTGGGCGACCGATGATGTTGCACAGATTTATGATAAATGC ATCAC
CNOT1-MC RV	TGCAACATCATCGGTCGCCCACGCTTGTTTC
M13-FW	GTAAAACGACGGCCAG
DUF-Stop-RV	TTTTTGTCGACTTAGAAGCCGGTCGGTTGGCTCAG
NoHis-FW	(Phos)ATGCCGGCCGCCACC
NoHis-RV	(Phos)AGTTAATTTCTCCTCTTTAATGAATTCTGTGTGAAA TTGT

Table 2.2 List of primers

2.1.3 Dpn I enzyme digestion of PCR product

Dpn I enzyme digestion is a good method to selectively remove parental, non-mutated DNA after PCR. The restriction endonuclease Dpn1 can effectively recognize and cleave adenine-methylated GmATC sequences but cannot cleave non-methylated GATC sequences. A standard reaction set up is indicated in Table 2.3.

Volume (µL)	**Reagent**
2	Dpn1 enzyme (10 U/µL)
X (10% of Y)	rCutSmart™ Buffer
20 or 30	PCR Sample
Up to Y	Nuclease free H_2O

Table 2.3 Standard reaction set up of Dpn1 enzyme digestion

2.1.4 DNA fragment ligation

T4 DNA ligase was carried out to create recombinant plasmids. Ligation of vector to insert used approximately 100 ng of digested vector DNA and a 3-fold excess insert concentration (Table 2.4). The reaction was carried out overnight at room temperature or 4℃.

Volume (µL)	Reagent/Component
-	100 ng template DNA
-	3-fold excess insert DNA
2	10× T4 DNA Ligase Buffer
1	T4 DNA Ligase enzyme
Up to 20	Nuclease free H_2O

Table 2.4 Standard ligation reaction

2.1.5 DNA transformation

10 µL of the DNA ligation product were mixed with NEB DH5α competent E. coli cells (100 µL) and subjected to heat shock at 42°C for 45 seconds, followed by immediate cooling on ice for 2 minutes. Subsequently, 1 mL of LB medium was added to the mixture, which was then incubated in a shaking incubator at 37°C for one hour. The culture was centrifuged at 4000 rpm for 1-2 minutes to pellet the cells, the supernatant was discarded, and then the pellet was resuspended in the residual medium. Finally, the resuspended cells were plated onto LB agar containing the appropriate antibiotic and incubated overnight at 37°C.

2.2 Microbiological techniques

2.2.1 Reagents, buffers, and stock solutions

Lysogeny Broth (LB) : 10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl, pH 7.2 (NaOH). LB agar included 15 g/L bacterial agar. The medium was sterilized using an autoclave and stored at room temperature.

Streptomycin 1000X stock solution : 100 mg/mL in water. Stored at -20°C.

Ampicillin 1000X stock solution : 100 mg/mL in 50% ethanol/water. Stored at -20°C.

Chloramphenicol 1000X stock solution: 34 mg/mL in 100% ethanol, stored at -20°C.

2.2.2 Escherichia coli strain

Escherichia coli strain DH5α and XL10-GOLD were used for DNA manipulations; Escherichia coli strain BL21 (DE3) was used for protein expression. The agar plates containing appropriate antibiotic can be made in advance and stored at 4°C for a few weeks but need to be preheated at 37°C for 30 minutes before the liquid bacterial culture was spread on them.

2.2.3 Transformation of competent cells

2.2.3.1 Transformation of E.coli strain DH5α

For transformation of DH5α Competent E. coli, 50 µL of competent cells were used for plasmid transformation and 100 µL of competent cells were used for ligation or mutagenesis transformation. For transforming plasmids, around 500 ng of DNA was used, and for transforming ligations or mutagenesis reactions, 10 µL of ligation reaction mix was used. Water bath was switched on and set temperature at 42°C, while heat/shaking block was turn on and set temperature at 37°C. Competent cell was thawed on ice for 5 minutes and then added with DNA or ligation product. Then the tube of cells was kept on ice for 30 minutes, followed by heat shocking at 42°C in water bath for 45 seconds and then placed on ice for 2 minutes. LB media (900 µL) was added to the competent cells before incubating for an hour at shaking incubator at 250 rpm and 37°C. For plasmid transformation, 100 µL of the cell suspension was spread onto LB agar plates (containing the appropriate antibiotic). For ligations or mutagenesis transformations, the culture was spun for 1-2 minutes at 4000 rpm until pellet form, around 800 µL of the supernatant was removed and the cells were re-suspended in the remaining 100 µL culture. Finally, the resuspended pellet was spread onto LB agar plates (containing the appropriate antibiotic). After that, the plates were incubated upside-down overnight in incubation at 37°C.

2.2.3.2 Transformation of E.coli strain XL10-GOLD

For transformation of XL10-GOLD Competent E. coli, most of the previous procedures were the same as the transformation of DH5α Competent E. coli. The only thing different was that the tube of cells was added with 5 µL or 10 µL of plasmid DNA or ligation product (around 10 ng/µL) and kept on ice for 30 minutes, followed by heat shocking at 42°C in water bath for 30 seconds

and then placed on ice for 10 minutes.

2.2.3.3 Transformation of E.coli strain BL21 (DE3)

For transformation of BL21 (DE3) Competent E. coli, 50 µL of competent cells were used for plasmid transformation. Water bath was switched on and set temperature at 42°C, while heat/shaking block was turn on and set temperature at 37°C. Competent cell was thawed on ice for 5 minutes and then added with 500ng of plasmid DNA product. Then the tube of cells was kept on ice for 30 minutes, followed by heat shocking at 42°C in water bath for 45 seconds and then placed on ice for 2 minutes.

2.3 Molecular biology

2.3.1 Reagents, buffers, and stock solutions

10× TBE : 80mM Tris base (121.1 g), 80mM boric acid (61.8 g), 1 mM EDTA (7.4 g), pH 8.0, dissolved completely in 1 L pure water, stored at room temperature.

Oligonucleotides : Primers were dissolved in TE to make100 µM solution and stored at -20°C.

2.3.2 DNA plasmid preparation

Monarch Spin Plasmid Miniprep Kits (NEB #T1110) were used for the small scale preparation of plasmid DNA. A single bacterial colony was picked and used to inoculate 10 mL LB in appropriate antibiotic, grow overnight in shaking incubator at 250 rpm, 37°C. The overnight culture was then spun in

the centrifuge for 10 minutes at 4000rpm, 4℃ until pellet form.The plasmid DNA was then extracted using Monarch Spin Plasmid Miniprep Kits. Details and procedures of manufacturer's instructions can be found in the Monarch Spin Plasmid Miniprep protocol. Plasmids were eluted in 30-50 μL of elution buffer. The concentration of plasmid DNA was measured using a Nanodrop ND-1000 spectrophotometer, then stored at -20℃.

2.3.3 Determination of DNA concentration

The concentration of plasmid DNA was determined using the NanoDrop ND-100 UV-Vis Spectrophotometer (NanoDrop Technologies). The A260/A280 value was used to estimate the purity of DNA. A ratio between 1.8 – 2.0 was acceptable for DNA sample.

2.3.4 Polymerase Chain Reaction (PCR)

All the following different PCR were performed using a Peqlab Primus 96 advanced thermal cycler and was performed using HighField® Phusion DNA polymerase. Desalted oligonucleotides used as primers were purchased from Sigma Genosys.

2.3.4.1 Inverse PCR

Inverse PCR was operated to delete part of the DNA sequence in the sample DNA plasmid. A standard reaction set up is indicated in Table 2.1. Sample were initially denatured at 95℃ for 3 minutes followed by 30 cycles of a denaturing step 95℃ for 15 seconds, annealing at T_m℃ for 30 seconds, then elongation at 72℃ for X minutes (15 - 30 seconds / kb). At the end of this

process, a final elongation step for 10 minutes at 72°C was included and then the sample can be hold at 15°C. The primers used are shown in Table 2.3. To check the PCR product was the correct size, some of the reaction was loaded onto a 1 % agarose gel.

2.3.4.2 QuikChange PCR

The QuikChange PCR is used to make point mutations, switch amino acids, and delete or insert single or multiple amino acids. Primers were designed using the Agilent web-based primer design tool: https://www.agilent.com/store/primerDesignProgram.jsp, which was optimized todesign mutagenic primer sequences specifically for the QuikChange kits. Primers were selected to be between 25 and 45 bases in length with a melting temperature (T_m) of ≥75°C. Optimal primer sets for simultaneous mutagenesis should have similar melting temperatures. A standard reaction set up is indicated in Table 2.1. Samples were initially denatured at 95°C for 3 minutes followed by 30 cycles of a denaturing step 95°C for 15 seconds, annealing at T_m°C for 30 seconds, then elongation at 72°C for X (15 - 30 seconds / kb) minutes. At the end of this process, a final elongation step for 10 minutes at 72°C was included and then the sample can be hold at 15°C. The primers used are shown in Table 2.3. To check the PCR product was as expected, some of the reaction product should be send to sequence.

2.3.4.3 Overlap PCR

Overlap PCR, also known as overlap extension PCR, involves the process of

joining two sequences, designated as A and B. The original normal primers are used to amplify the sequences containing these complementary regions. As a result, during the annealing phase, these two initially separate sequences can hybridize through their complementary regions, facilitating their connection into a single contiguous sequence.Each gene is subjected to a first round of PCR amplification using its respective primers. The resulting products are purified using column purification, yielding intermediate fragments A' and B'.

Each gene is subjected to a first round of PCR amplification using its respective primers. The resulting products are purified using column purification, yielding intermediate fragments A' and B'. The first round of PCR amplification conditions are as follows: an initial denaturation at 95°C for 30 seconds, followed by 35 cycles of denaturation at 95°C for 15 seconds, annealing at T_m°C for 30 seconds, and extension at 72°C for 30 minutes.

In the second round of PCR, intermediate fragments A' and B' are used as templates. After measuring their concentrations, equal molar amounts of A' and B' are combined, ensuring the total DNA amount does not exceed 100 ng. Primers A1 and B2 are used, while other components remain unchanged as it is shown in table 2.1. The second round of PCR amplification conditions are as follows: an initial denaturation at 95°C for 30 seconds, followed by 35 cycles of denaturation at 95°C for 15 seconds, annealing at T_m°C for 30 seconds, and extension at 72°C for X (15 - 30 seconds / kb) minutes, with a final extension step at 72°C for 10 minutes. The primers used are shown in Table 2.3.

The PCR products were subjected to 1 % agarose gel electrophoresis to verify

the correct length of the amplified gene fragments. If the fragment lengths are correct, the appropriate bands are excised from the gel and purified. The purified DNA is then transformed and sequenced for verification.

2.3.4.4 Gradient PCR

The Gradient PCR is a technology that entails configuring a thermal cycler so that instead of giving the same temperature throughout the entire thermal block of a thermal cycler, different columns or rows within the thermal block are maintained at distinct temperatures. Consequently, this allows for the simultaneous evaluation of multiple annealing temperatures (T_m) in a single run. This method eliminates the need to perform multiple, separate PCR experiments at different temperatures, each lasting approximately one hour, to determine the optimal annealing temperature. (Temperature Gradient: Set a temperature gradient 10°C above and below the Tm value to determine the optimal annealing temperature.) Once the optimal annealing temperature is identified, sample were initially denatured at 95°C for 3 minutes followed by 30 cycles of a denaturing step 95°C for 15 seconds, annealing at Tm°C for 30 seconds, then elongation at 72°C for X minutes (15 - 30 seconds / kb). At the end of this process, a final elongation step for 10 minutes at 72°C was included and then the sample can be hold at 15°C. Standard reactions were set up as detailed in Table 2.5.

Volume (µL)	Reagent

10	10 mM dNTPs
1	5× Phusion Buffer HF
2.5	1 μM forward primer
2.5	1 μM reverse primer
X	100 ng DNA template
1.5	DMSO
0.5	Phusion DNA polymerase
Up to 50	Nuclease free H_2O

Table 2.5 Standard reaction set up of PCR using the Phusion DNA polymerase

2.3.5 Purification of PCR product

Monarch Spin PCR and DNA Cleanup Kits (5 μg) (NEB #T1130) were used to prevent buffer retention and salt carry-over with optimized column design as well as rapidly and reliably purify up to 5 μg of concentrated, high-quality DNA from PCR reactions. The instructions of the experiment were provided on the protocol of Monarch Spin PCR and DNA Cleanup Kit.

2.3.6 Agarose gel electrophoresis

A 1 % agarose gel (w/v) was prepared by dissolving 0.5 g of agarose in 50 mL of 1× TBE buffer. 5 μL of ethidium bromide was added to the gel solution before it was poured into a gel tank equipped with an appropriate comb to form a solid gel. Each DNA sample was mixed with 6 fold Ficoll dye and pipetted into the different wells respectively. Samples were loaded alongside 5 μl of NEB standard 1 Kb DNA ladder and run at 90V for one hour. Bands

were then visualised using an ultraviolet transluminator (BioRad Gel Doc 2000) and BioRad Quantity One computer software.

2.3.7 Gel extraction and purification of DNA

To purify the fragment of DNA from agarose gel, the DNA bands were cut from the agarose gel using a knife. Monarch Spin DNA Gel Extraction Kits (NEB #T1120) were used to purify the DNA from the reaction mixture. The instructions of the experiment were provided on the protocol of Monarch Spin DNA Gel Extraction Kit.

2.4 Protein expression and purification

2.4.1 Reagents, buffers and stock solutions

Lysogeny Broth (LB) : 10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl, pH 7.2 (NaOH). LB agar included 15 g/L bacterial agar. The medium was sterilized using an autoclave and stored at room temperature.

AKTA™ start system and Frac30 fraction collector

Computer installed with UNICORN™ start 1.0 control software

Column : GE Healthcare HiTrap TALON crude 1 ml, from Scientific Laboratory Supplies Ltd

HiPrep™ 16/60 Sephacryl™ S-200 HR 120mL, from Scientific Laboratory Supplies Ltd

Fraction tubes: 1.5mL tubes, 15mL tubes

1mL Sample loop

IPTG (Isopropyl β-2-1-thiogalactopyranoside)

Streptomycin 1000X stock solution : 100 mg/mL in water. Stored at -20°C.

Ampicillin 1000X stock solution : 100 mg/mL in 50% ethanol/water. Stored at -20°C.

Chloramphenicol 1000X stock solution: 34 mg/mL in 100% ethanol, stored at -20°C.

NaCl (Sodium chloride), β-mercapthoethanol, Tris-HCl, Imidazole

Buffer solution:

Buffer A (Binding buffer): 20 mM Tris-HCl pH 7.9, 250 mM NaCl, 1 mM β-mercaptoethanol, 5% glycerol and 20 mM imidazole.

Buffer B (Elution buffer): 20 mM Tris-HCl pH 7.9, 250 mM NaCl, 1 mM β-mercaptoethanol, 5% glycerol and 250 mM imidazole.

Buffer C (Wash buffer): 20 mM Tris-HCl pH 7.9, 250 mM NaCl, 1 mM β-mercaptoethanol, 5% glycerol and 40 mM imidazole.

Gel filtration buffer (GF buffer): 25mM sodium phosphate pH 7.5, 150mM NaCl

2.4.2 Protein expression

The plasmid was transformed into BL21(DE3) cells. The transformed cells were spread on selection plates (containing the appropriate antibiotic) and incubated overnight at 37°C. A single colony was selected and inoculated into 1 mL of LB broth (containing the appropriate antibiotic). The culture was incubated in a shaking incubator at 37°C and 250 rpm for 2 to 6 hours. The 1 mL starter culture was then transferred into 50 mL of LB broth containing the antibiotic and incubated overnight at 37°C and 250 rpm in a shaking incubator.

1 mL of the overnight culture was used to inoculate 1 liter of liquid medium with the antibiotic. The culture was incubated at 37°C with shaking until the OD600 reached between 0.8 and 1.2. Then IPTG was added to the culture to a final concentration of either 40 μM or 400 μM (250 μL of IPTG solution), and protein expression was induced by incubating the culture overnight at 28°C and 125 rpm.

The culture was then centrifuged at 4°C and 4000 rpm for 20 to 30 minutes. The supernatant was discarded, and the cell pellet was resuspended on ice in 5 to 10 mL of Buffer A. The resuspended cells were frozen at -80°C for at least 30 minutes. Then the frozen cells were thawed in ice water and vortexed thoroughly once thawed. The cells were sonicated for a total of 30 minutes, with 30-second pulses followed by 30-second pauses. The OD600 was measured before and after sonication to ensure effective lysis. Finally, the lysate was transferred into 1.5 mL tubes and centrifuged at 4°C and 14,000 x g for 30 minutes. The supernatant was carefully transferred to a falcon tube

for further processing.

2.4.3 Immobilized metal affinity chromatography (IMAC)

The AKTA start system was set as the table 2.7. The entire system was washed with pure water for 15 minutes per channel and then washed with Buffer C for 15 minutes per channel. Then the running system was started and the purified protein sample was collected. Finally, the protein solution were measured using appropriate methods, such as spectrophotometry or BCA assay, with the expectation of obtaining an obvious peak in the data. The protein products were subjected to SDS-PAGE to verify the correct size of the amplified protein.

Phase	**Variable**	**Value**
Method Settings	Column Volume (CV)	0.962
	Column	HiTrap TALON crude, 1mL
	Flow Rate (mL/min)	1.0
Prime and Equilibration	Equilibration Volume (CV)	5.00
Sample Application	Sample Volume (mL)	12.00
Wash out unbound	Wash Column With (CV)	10.00
Elution and Fractionation	Elution Isocratic B concentration (%B)	100.0
	Elution Fixed Fractionation Volume	1.0

	(mL)	
	Elution Isocratic Volume (CV)	10.00
Prime and Equilibration	Equilibration Volumn (CV)	5.00

Table2.7 System method setting of protein purification using AKTA start

2.4.4 Gel filtration

The gel filtration system was set as the Table 2.8. Wash the whole system for 15 minutes using GF buffer (Flow rate should be less than 1mL/min). The sample loop was washed with 10 mL Buffer A with syringe and then the sample loop was pre-filled with 1mL protein sample. Then the running system was started and the purified samples were collected by 15 mL tubes. Finally, these samples were subjected to SDS-PAGE to verify the purity or concentrated using Spin column.

Phase	**Variable**	**Value**
Method Settings	Column Volume (mL)	120
	Column	HiPrep 16/60 Sephacryl S200 HR
	Pressure Limit (MPa)	0.15
	Flow Rate (mL/min)	0.5
Prime and Equilibration	Equilibration Volume (CV)	1.2
Sample Application	Apply sample using loop	-

Elution and Fractionation	Sample Volume (mL)	1.00
	Elution Isocratic B concentration (%B)	0
	Elution Fixed Fractionation Volume (mL)	4.0
	Elution Isocratic Volume (CV)	1.0

Table 2.8 System method setting of gel filtration using AKTA start

2.5 Operation of protein samples

2.5.1 Reagents, buffers and stock solutions

PD-10 desalting columns pre-packed with Sephadex G-25 (Cytiva)

Spin columns suitable for protein concentration (e.g., Amicon Ultra-0.5, Microcon, Vivaspin)

Buffer solution: Phosphate-buffered saline (PBS) or other appropriate buffer

Protein solution to be concentrated

Centrifuge tubes or collection tubes

Centrifuge with appropriate rotor for spin columns

2.5.2 PD-10 desalting assay

The top cap of the PD-10 column was removed, and the sealed end was cut to allow drainage. The column was equilibrated by adding 25 mL of PBS or an appropriate buffer, which was allowed to drain completely by gravity. Up to 2.5 mL of the protein solution was applied to the spin column, ensuring the maximum recommended volume was not exceeded. Then the spin column was capped and placed into a collection tube. After the sample has entered the packed bed, 3.5 mL of the buffer solution was added to elute the protein. The eluent was collected in a centrifuge tube or collection tube. The first 3.5 mL collected will contain the concentrated protein.

2.5.3 Concentration of protein samples

A spin column with an appropriate molecular weight cutoff (MWCO) for the protein of interest was selected and pre-washed by adding the buffer solution and centrifuging according to the manufacturer's instructions. The flow-through was discarded. The wash step was repeated if necessary to ensure the membrane was equilibrated with the buffer. The protein solution was added to the spin column, ensuring that the maximum recommended volume for the column was not exceeded. Then the spin column was capped and placed into a collection tube. The column was centrifuged at the recommended speed and duration (typically 10-15 minutes at 14,000 x g), or as specified by the manufacturer, until the desired concentration was achieved. The volume of the concentrate was monitored to avoid over-concentration or drying out the protein. After centrifugation, the spin column was carefully removed from the centrifuge and then inverted into a

fresh centrifuge tube and centrifuged again briefly to recover the concentrated protein solution from the membrane. The volume and concentration of the recovered protein solution were measured using appropriate methods, such as spectrophotometry or Bradford protein assay.

2.6 Sodium dodecyl sulphate-polyacryl-amide gel electhrophoresis (SDS-PAGE)

2.6.1 Reagents, buffers and stock solutions

Novex Tris-Glycine Mini Protein Gels, 10%, 1.0 mm, Wedge Well format, Wells: 15-well

SDS-PAGE Running buffer : 25 mM Tris, 192 mM glycine, 0.1% SDS, pH 8.3

SDS sample buffer (Laemmli buffer) : 62.5 mM Tris-HCl, 2% SDS, 10% glycerol, 5% β-mercaptoethanol, 0.01% bromophenol blue, pH 6.8

Color Prestained Protein Standard, Broad Range (10-250 kDa)

Protein samples

Staining solution: Coomassie Brilliant Blue R-250

Electrophoresis apparatus and power supply

2.6.2 Gel electrophoresis, staining and destaining

Protein samples were mixed with SDS sample buffer and heated at 95°C for 5 minutes to denature the proteins. The samples were then cooled to room temperature before loading. The polymerized gel was placed into the electrophoresis apparatus, and the wells were rinsed with running buffer. The running buffer was added to the upper and lower reservoirs of the electrophoresis apparatus and then protein samples and a protein ladder were loaded into the wells using a micropipette. The gel was run at a constant voltage (e.g., 120 V) until the dye front reached the bottom of the gel.

After electrophoresis, the gel was carefully removed from the apparatus and placed into a staining solution containing Coomassie Brilliant Blue R-250. The gel was stained for 30 ~ 60 minutes with gentle agitation. The gel was then destained using pure water until the protein bands were clearly visible against a clear background.

2.7 Western Blot

2.7.1 Reagents, buffers and stock solutions

SDS-PAGE gel

PVDF (polyvinylidene fluoride) membrane

HRP Anti-6X His tag antibody, HRP Anti-mouse conjugated antibody

1× TBST : 20mM Tris-base, 150mM NaCl, 50mM KCl, 0.2% Tween-20, adjusted pH to 7.6.

Transfer buffer : 25 mM Tris, 192 mM Glycine, 20% Methanol (v/v)

Blocking buffer : 5% non-fat milk or BSA in TBST

2.7.2 Membrane transfer and immunodetection

Reaction samples separated in SDS-PAGE gels for western blotting were transferred to from the gel to a PVDF membrane by electrophoresis in 1x transfer buffer using wet transfer apparatus. A transfer sandwich is assembled, including filter paper, the gel, the membrane, and additional filter paper. The transfer is performed using a transfer apparatus, typically at 25 V for 1-2 hours (4°C). Then the membrane is blocked to prevent non-specific binding of antibodies. Blocking is performed by incubating the membrane in blocking buffer for 1 hour at 4°C.

After that, the membrane was incubated with a primary antibody specific to the protein of interest. The antibody was diluted in blocking buffer (1:1000), and the membrane is incubated for 1-2 hours at room temperature or overnight at 4°C with gentle shaking. Then the membrane is washed 3-5 times with TBST to remove unbound primary antibody, each wash is conducted for 5-10 minutes with gentle agitation. After that, the membrane is incubated with a suitable secondary antibody (recognize the host species of the primary antibody), which diluted according to the instruction, for 1 hour with constant rocking. Then the membrane is washed 3 times with TBST for 10 minutes each. Finally, the membrane is exposed to X-ray film or imaged using a chemiluminescence detection system. The protein bands corresponding to the target proteins are visualized and analyzed. The bands

are compared to the protein ladder to determine molecular weight.

2.8 Bradford Protein Assay

2.8.1 Reagents, buffers and stock solutions

Bradford reagent (Coomassie Brilliant Blue G-250)

Protein samples to be quantified

Bovine Serum Albumin (BSA) standard for calibration

96-well microplate or cuvettes for measurement

Microplate reader or spectrophotometer capable of reading at 595 nm

Distilled water or buffer for dilutions

2.8.2 Determination of Protein concentration

First step is to prepare standard curve and protein sample. A series of BSA standard solutions were prepared by diluting a stock solution in distilled water or buffer to create a range of known protein concentrations (e.g., 0, 2, 5, 10, 20, 40 μg/mL). 10-20 μL of each BSA standard solution was added to the wells of a microplate or into cuvettes. In addition, 10-20 μL of each protein sample was added to the appropriate wells or cuvettes.

After that, 200 μL of Bradford reagent was added to each well or cuvette

containing the standard solutions and protein samples. The contents were mixed gently by pipetting up and down or by shaking the microplate gently. Then the reaction was allowed to incubate at room temperature for 5 to 10 minutes to allow color development (remember do not incubate more than 1 hour). The absorbance of each well or cuvette was measured at 595 nm using a microplate reader or spectrophotometer.

Finally, the absorbance values of the BSA standards were used to construct a standard curve, plotting absorbance versus protein concentration. The protein concentrations of the samples were determined by comparing their absorbance values to the standard curve.

2.9 Peptide pull-down assay

2.9.1 Reagents, buffers and stock solutions

Dynabeads: Common choices were Dynabeads™ Protein A or Protein G (depending on the antibody type).

YTHDF2 peptides: Target peptides to bind to the CNOT9 protein.

CNOT9 protein: Ppurified protein samples.

Binding buffer: 50 mM Tris-HCl pH7.5, 300 mM NaCl, 0.1 % NP-40, 1x Complete Protease Inhibitors.

Wash buffer: Used to remove non-specific binding proteins (e.g., usually PBS containing 0.1% Triton X-100).

2.9.2 Peptides binding assay

15 µL of Dynabeads was taken into a fresh tube (typically 15 µL, depending on the experiment scale). The Dynabeads were separated using a magnetic rack, the storage buffer was dwascarded, and they were washed 1-2 times with binding buffer by resuspending. 1.5 µg peptides or no peptide control were added to the Dynabeads. Reactions were incubated for 1 -2 hours with rotation at 4 °C in around 450 µl Binding buffer, allowing the peptides to bind. 1.5 µg purified tagged fusion protein (molarity value: around 1 ~ 3 mM) was then added to the Dynabeads bound with YTHDF2 peptides. Incubation was carried out with gentle rotation for 1-2 hours at 4 °C. After incubation, the Dynabeads were washed 3 times (10min/time) with wash buffer to remove non-specifically bound proteins. Dynabeads were then harvested using a magnet and consequently washed (3x10 minutes) in binding buffer. Samples were then subjected to SDS-PAGE and followed by Coomassie blue staining or Western Blot to check the results.

Chapter 3

Reconstitution of the CNOT7-CNOT1M-CNOT9 modules and the CNOT7-CNOT1C-CNOT9-NOT Modules

Chapter 3. Reconstitution of the CNOT7-CNOT1M-CNOT9 modules and the CNOT7-CNOT1C-CNOT9-NOT Modules

3.1 Introduction

As I mentioned in previous section, the conserved core of the Ccr4-Not complex consists of at least two modules:

CNOT1M (MIF4G-DUF3819)-CNOT9-CNOT7 module and the NOT module, which consists of the NOT-boxes of CNOT3 and CNOT2. These modules are integral to the complex's functionality, enabling interactions with various regulatory proteins and facilitating its diverse regulatory roles.

Recent advances in structural biology have provided profound insights into the organization and dynamics of these modules within the Ccr4-Not complex. The CNOT1M module, encompassing the MIF4G and DUF3819 domains of CNOT1, plays a crucial role in the structural and functional integrity of the complex. This module, along with the CNOT9 and CNOT7 module, is involved in critical interactions that underpin the complex's enzymatic and regulatory functions. The relationship between these modules is mediated by a network of flexible linkers, which are crucial for the structural organization and functional dynamics of the complex [1].

Understanding the nature of these linkers is a key area of investigation. These linkers may exhibit either defined three-dimensional arrangements or

substantial flexibility, which impacts how the Ccr4-Not complex interacts with its substrates and regulatory proteins. Flexibility within these linkers could facilitate the complex's ability to adopt multiple conformations, potentially enhancing its capacity to interact with a wide range of RNA substrates and cofactors. Conversely, a more rigid structural organization could confer specificity and efficiency in its regulatory functions, allowing for precise control over mRNA metabolism and gene expression.

This thesis aims to delve into the reconstitution and structural dynamics of the CNOT1M (MIF4G-DUF3819)-CNOT9-CNOT7 module and the NOT module (Figure 3.1A,B), with a particular emphasis on the role of linkers connecting these structural elements (linkers between CNOT1 and CNOT9, linkers between CNOT1 and CNOT7, linkers between CNOT7 and CNOT6/CNOT6L). By elucidating the spatial relationships and conformational flexibility of these linkers, we seek to advance our understanding of the Ccr4-Not complex's architecture and its implications for mRNA metabolism and gene expression regulation. Such insights are essential for unraveling the complex's functional mechanisms and exploring its potential as a therapeutic target for diseases linked to its dysregulation.

Two strategies were used to express a CNOT1M (MIF4G-DUF3819)-CNOT9-CNOT7 protein complex while only one method was used to express CNOT1L-CNOT9-Notbox of CNOT3 and CNOT2 protein complex (Figure 3.2). First, we aimed to co-express the CNOT1M-CNOT9 and un-tagged CNOT7 in E.coli (CNOT1M-CNOT9 is tagged with His-tag). Using this method, the amount of protein complex was not ideal. So second strategy involved reconstitute the complex with purified CNOT1-CNOT9 and

un-tagged CNOT7. Although it was possible to remove the majority of the contaminants by gel filtration. The strategy of expressing NOT module was co-expressing the CNOT1L-CNOT9 and NOTbox of CNOT3 and CNOT2 protein complex in same E.coli. By using immobilized metal affinity chromatography (IMAC) to purify these protein, a complex with the expected stoichiometric ratio was obtained.

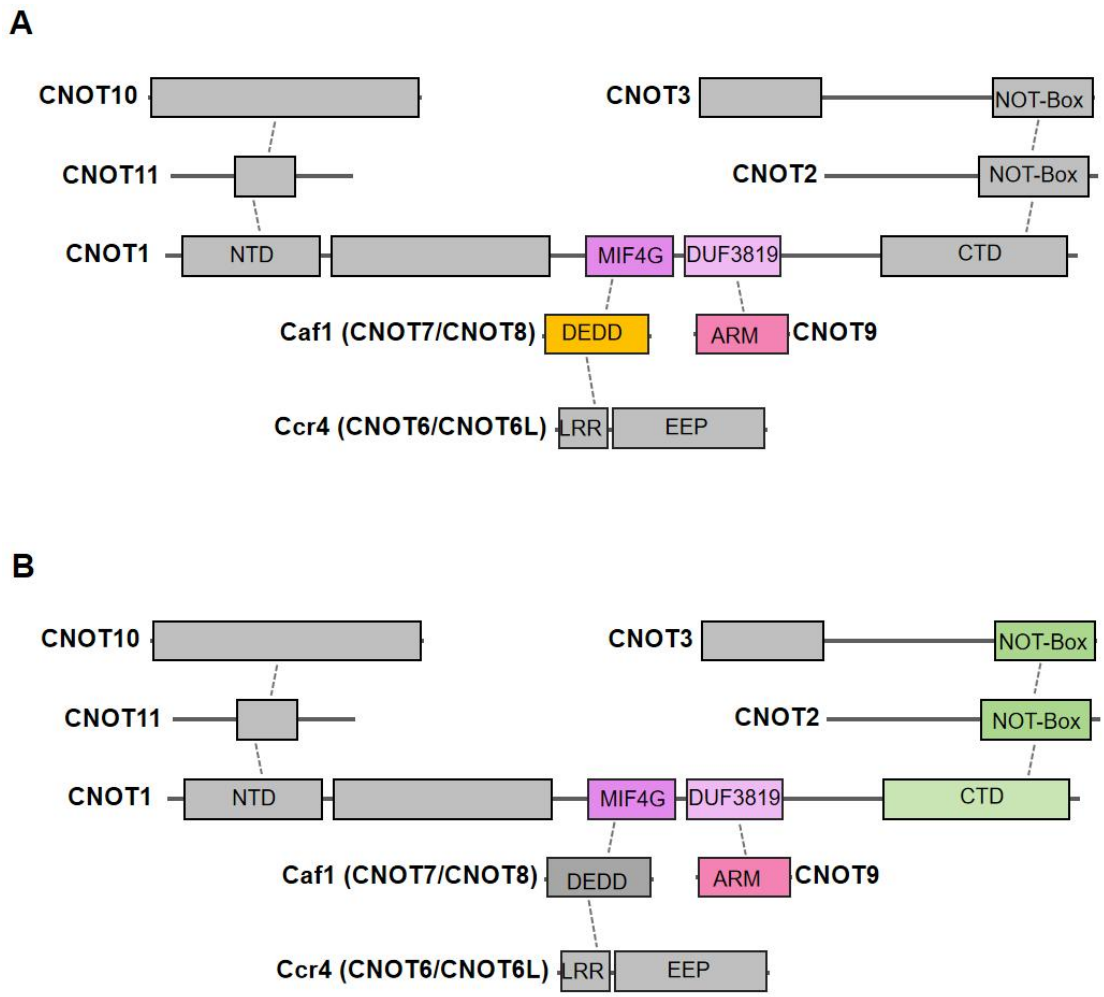


Figure 3.1 Structure of the CNOT1M (MIF4G-DUF3819)-CNOT9-CNOT7 module and the NOT module. (A) Highlighted part is the structure of the CNOT1M (MIF4G-DUF3819)-CNOT9-CNOT7 module. **(B)** Highlighted part is

the structure ofthe NOT module.

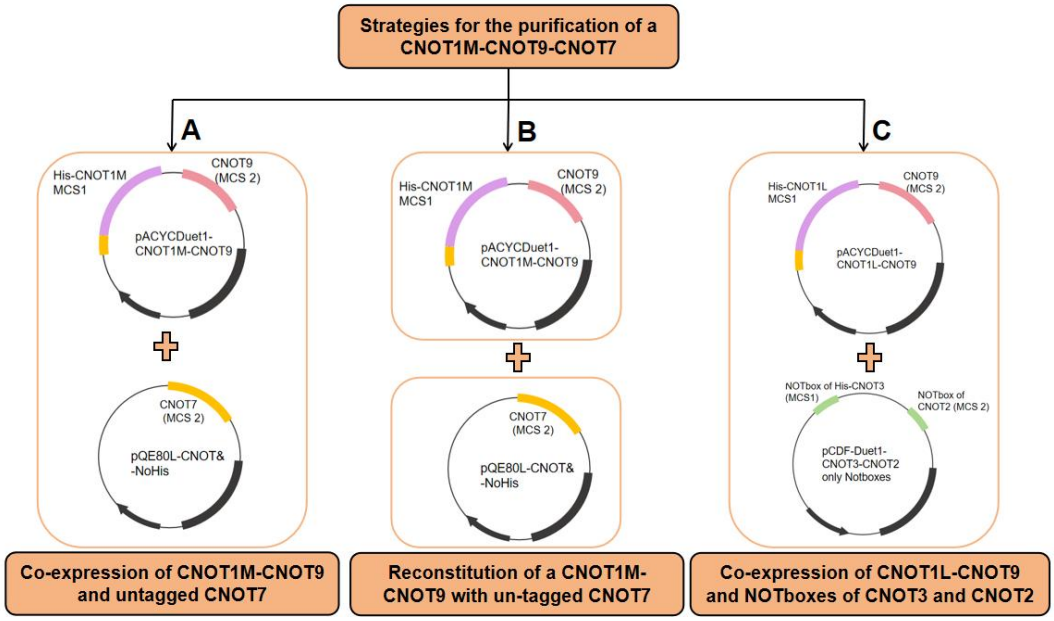


Figure 3.2 Strategies for the purification of a CNOT1M-CNOT9-CNOT7 protein complex and CNOT1L-CNOT9 and Notboxes of CNOT3 and CNOT2 protein complex. (A) Co-expression and purification of His•CNOT1M-CNOT9 and un-tagged CNOT7. **(B)** Reconstitution of a CNOT1M-CNOT9 complex with purified un-tagged CNOT7 complex. **(C)** Co-expression and purification of CNOT1L-CNOT9 and NOTboxes of CNOT3 and CNOT2 protein complex.

3.2 Co-expression and purification of His-tagged CNOT1M-CNOT9 and CNOT7 protein complex

3.2.1 Generation of plasmids

To reconstitute the His-tagged CNOT1M(MIF4G-DUF3819)-CNOT9 and

CNOT7 protein complex, we generated the following plasmids: pACYCDuet1-CNOT1M-CNOT9 (containing the His-tag on CNOT1M) and pQE80L-CNOT7-NoHis (Table 3.1).

Based on previous work in my group[1,2], our plasmid pACYCDuet1-M01 contained a stop codon between CNOT1M and CNOT1C fragment and it will interfere with protein translation process. So it was necessary to delete the stop codon. Besides, plasmid pMA-T-CNOT1 was also used to get the CNOT1-C terminal fragment through Inverse PCR using CNOT1C-OEPCR-FW and pMA-T-RV primers (Table 2.3). Initially, Overlap PCR was conducted to delete the stop codon in plasmid pACYCDuet1-M01. Primers CNOT1M-OEPCR-FW and CNOT1M-OEPCR-RV were used (Table 2.3) to get the CNOT1M fragment. However, based on the results, the fragments CNOT1M and CNOT1C cannot be aligned efficiently, which means the Overlap PCR was failed. We suspected that it was because the annealing temperature of different primers was difficult to be unified, which led to the result that two fragments cannot bind with each other. Then we attempted to using QuikChange PCR to delete the stop codon between CNOT1M and CNOT1C. Primers del1522 and del1522-antisense were used to carry out this experiment (Table 2.3). But the sequence results showed that the deletion of the stop codon was not specific. Besides, we also tried to using Inverse PCR to directly get rid of the CNOT1C fragment and the stop codon. Primers CNOT1L-IPCR-FW, CNOT1-MC RV, CNOT1-MC FW and CNOT1L-IPCR-RV were used (Table 2.3). Unfortunately, based on the results of all these experiments, stop codon still cannot be deleted efficiently. Above all, no further attempts were made to delete the stop codon directly in plasmid pACYCDuet1-M01. Instead, we ordered long CNOT1 full length cDNA as a synthetic DNA

fragment from Genescript. We got the plasmid PUC57-CNOT1L (containing the full length of CNOT1) (Figure3.4A). Next we tried to get plasmid pACYCDuet1-CNOT1M-CNOT9 using Inverse PCR and restriction digestion.

The DNA fragments which cut by Bgl II and BamH I can be ligated together because their overhangs are compatible. Both enzymes generate sticky ends with the sequence GATC, as shown in the image (Figure 3.3). When fragments cut by these enzymes are mixed, their sticky ends with complementary bases (GATC) can pair through base pairing, allowing ligation by DNA ligase. However, after ligation, the hybrid site (BamH I/Bgl II) is no longer recognized by either enzyme due to the differences in the remaining sequences outside the overhang. Based on this theory, CNOT1M fragment can be inserted into the vector with correct direction. Therefore, I successfully got the plasmid pACYCDuet1-CNOT1M-CNOT9 (with His-tag on CNOT1M), details of making the plasmid was described below.

Since the CNOT1M and CNOT9 cDNA in our group were codon-optimised for expression in E. coli (Genscript). So we need t remove this stop codon. Initially, two plasmids were generated, plasmid PUC57-CNOT1L with CNOT1 long fragment inserted into it and plasmids pACYCDuet1-M01 with CNOT1M (with stop codon) inserted into MCS 1 and CNOT9 in MCS 2. Firstly, the CNOT1M (MIF4G and DFU3819) fragment was prepared from plasmids PUC57-CNOT1L using Inverse PCR, primers M13-FW and DUF-stop-RV were used (Table 2.3). After that, the CNOT1M (MIF4G and DFU3819) cDNA was PCR amplified and then digested using restriction enzymes Bgl II (5') and Sal Ⅰ (3'), so that we got the insert fragments which was around 1600bp. Next, the restriction digested CNOT1M (MIF4G and DFU3819) cDNA (1566 bp) was

subcloned into MCS 1 of vector pACYCDuet1-CNOT9 (~ 6 kb), which was digested using the BamHⅠ and SalⅠ restriction sites to delete the CNOT1M (with stop codon)fragment. The presence of the CNOT1M (MIF4G and DFU3819) cDNA in MCS 1 of pACYCDuet1-CNOT9 was confirmed by digesting the plasmids using BamH I and SalⅠ. Digestion of pACYCDuet1-CNOT1M-CNOT9 using BamH I and SalⅠ resulted in two fragments (~ 6 kb and ~ 1600 bp). Subsequently, the CNOT1M (MIF4G and DUF3819) cDNA was subcloned into the Bgl II and SalⅠ restriction sites of MCS 1 (Figure 3.4B). Then in order to verify the correct ligation of the cDNA fragments, the plasmids were send to the Source Bioscience company to sequence. The sequence results showed the sequence as expected.

Because the plasmid pACYCDuet1-CNOT1M-CNOT9 contains the His-tag, we constructed pQE80L-CNOT7 without His-tag in order to co-express both plasmids. Original plasmids pQE80L-His•CNOT7 contains the His-tag, so we conducted Inverse PCR to deleted the His-tag. Primers NoHis-FW and NoHis-RV were used (Table 2.3). The plasmids were verified by DNA sequencing (Figure 3.4C).

Plasmid	**MCS1**	**MCS2**
PUC57-CNOT1L	CNOT1 long fragment	-
pMA-T-CNOT1	CNOT1	-
pACYCDuet1-M01	CNOT1M (with stop codon)	CNOT9
pACYCDuet1-CNOT9	-	CNOT9
pACYCDuet1-CNOT1M-CNOT9	CNOT1M (MIF4G and	CNOT9

	DUF3819)	
pQE80L-CNOT7-NoHis	CNOT7	-

Table 3.1 Plasmids

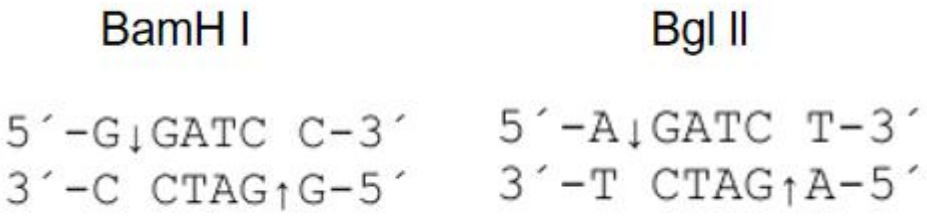
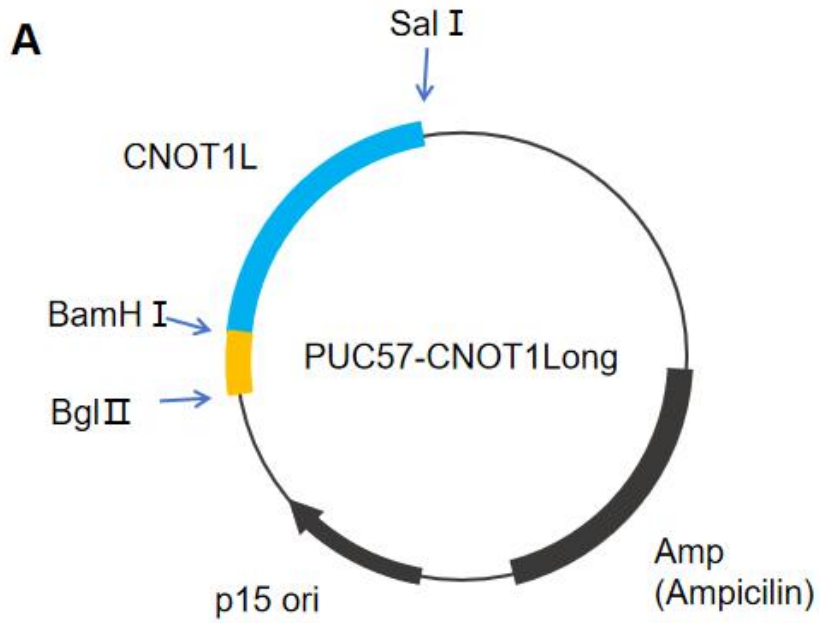
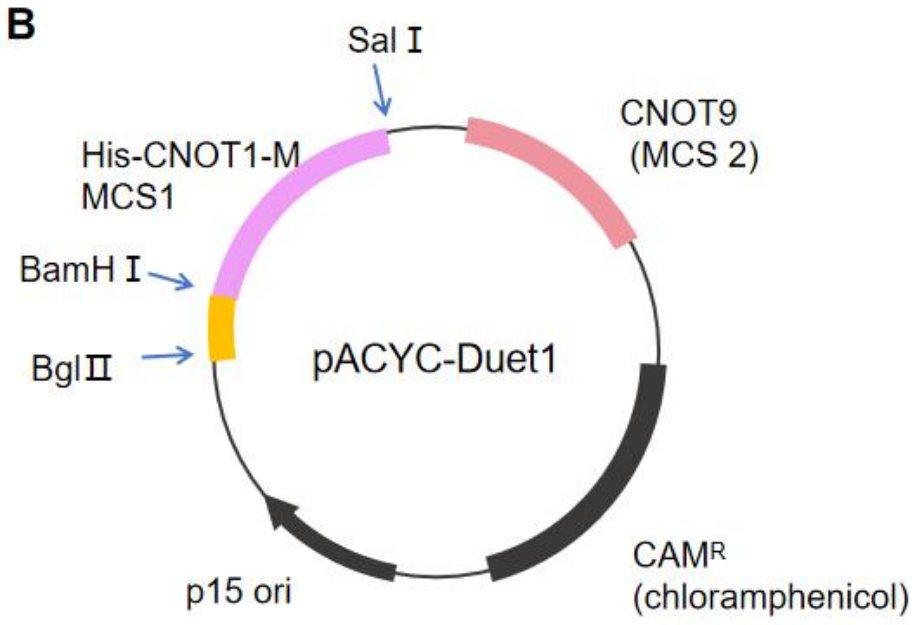


Figure 3.3 5' overhangs of DNA fragments cut by restriction enzyme Bgl II and BamH I.


A
Sal I
CNOT1L
BamH I
BglII
PUC57-CNOT1Long
Amp
(Ampicilin)
p15 ori


B
Sal I
CNOT9
(MCS 2)
His-CNOT1-M
MCS1
BamH I
BglII
pACYC-Duet1
CAMR
(chloramphenicol)
p15 ori

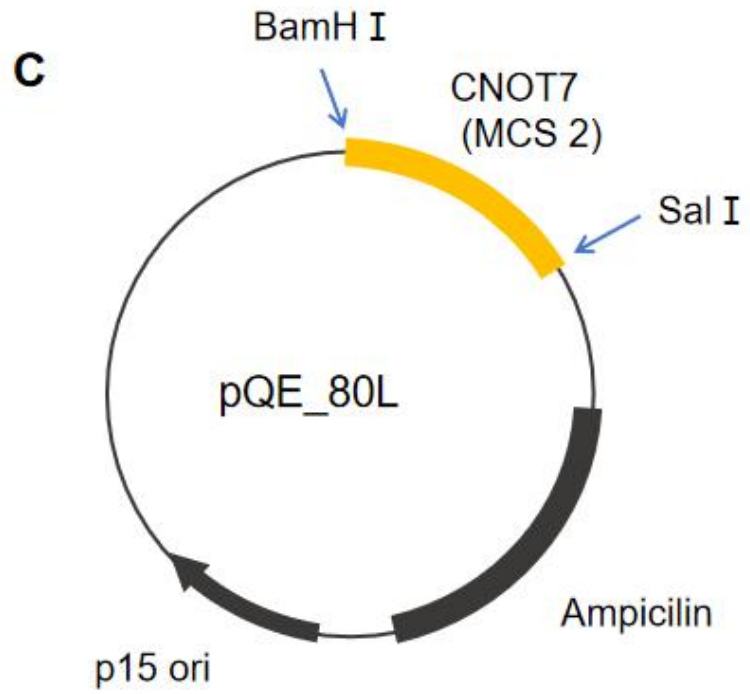


Figure 3.4 Generation of bacterial expression vectors containing the CNOT1 long fragment cDNA, bacterial expression vectors containing His-tagged CNOT1M and CNOT9 cDNAs, and bacterial expression vectors containing the CNOT7 cDNA without His-tag. The CNOT1 long fragment was ordered from Genescrip. The pACYCDuet1 and pQE80L vectors were used to generate plasmids pACYCDuet1-CNOT1M-CNOT9 and pQE80L-CNOT7-NoHis. **(A)** Generation of plasmid PUC57-CNOT1L. We ordered long CNOT1 cDNA as a synthetic DNA fragment from Genescript and then transformed to amplify the target plasmid. **(B)** Generation of plasmid pACYCDuet1-CNOT1M-CNOT9. CNOT1M (BamHI-SalI) was subcloned into multiple cloning site 1 (MCS 1) of pACYCDuet1 (BamHI-SalI). **(C)** Generation of plasmid pQE80L-CNOT7-NoHis. His-tag was deleted from the plasmid pQE80L-CNOT7 through Inverse PCR.

3.2.2 Protein expression and purification

Protein expression was initiated by transforming plasmids pACYCDuet1-CNOT1M-CNOT9 and pQE80L-CNOT7-NoHis into E. coli BL21(DE3) cells. In order to co-transform the plasmids together, for each plasmid, a single colony of E. coli BL21(DE3) was used to inoculate a starter culture for protein expression. The pre-culture was then grown in 1 L of LB containing 35 μg/ml chloramphenicol and 50 μg/ml ampicilin at 37°C until the OD(600 nm) was around 1. Then, IPTG (final concentration of 0.25 mM) was added into the culture to induce protein expression overnight at 28°C. After that, the cells and protein were collected using binding buffer containing 20 mM imidazole (Buffer A described in 2.4.1).

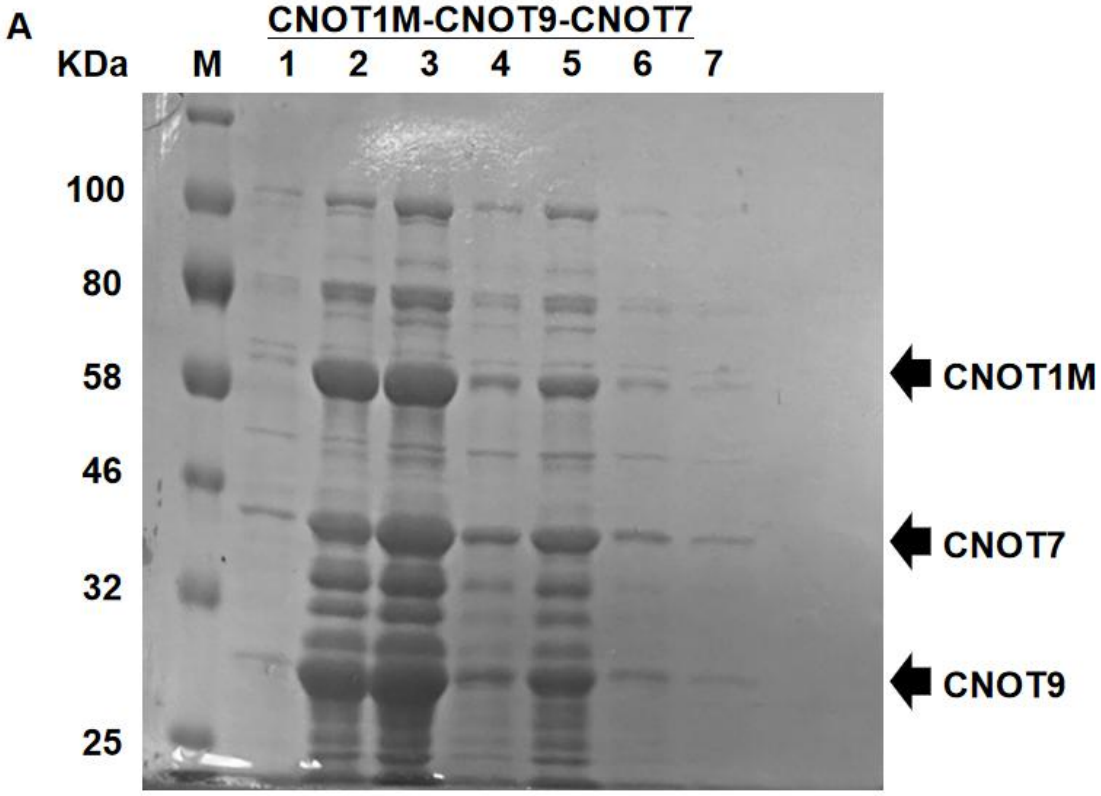
Protein purification was carried out using immobilized metal affinity chromatography (IMAC). The column was washed using 15 mL of pure water and 15 mL of wash buffer containing 40 mM imidazole (Buffer C described in 2.4.1) with a flow rate of 5 mL per minutes. Then the soluble lysate was applied to a His-Trap column with a flow rate of approximately 1 mL per minute. After that, the flow rate was stillset at 1 mL per minute and the protein was eluted with 5 ml of elution buffer containing 250 mM imidazole (Buffer B described in 2.4.1) and collected in 1 ml fractions. The purity of the protein complex was verified using SDS-PAGE and staining with Coomassie Brilliant Blue. The result of SDS-PAGE analysis demostrates that the bands of CNOT1M and CNOT9 as well as CNOT7 were showed in the gel of correct size (Figure 3.5A, B). The CNOT1M (MIF4G and DUF3819) protein was observed as a single band at ~60 kDa and the CNOT9 protein appeared as a single band

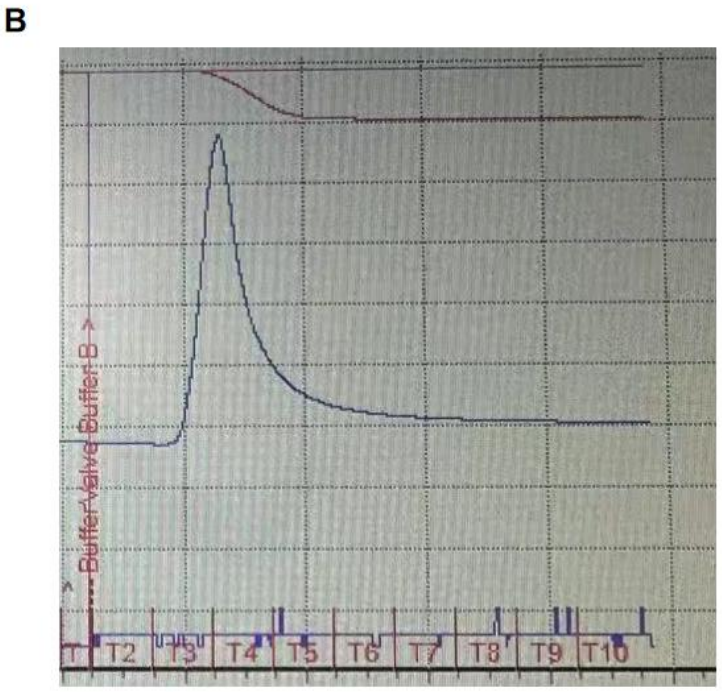
at ~28 kDa, while the CNOT7 protein was observed as a single band at ~42 kDa. However, there were still a lot of impurities present in the fractions containing the CNOT1M-CNOT9-CNOT7 protein complex (Figure 3.5A).

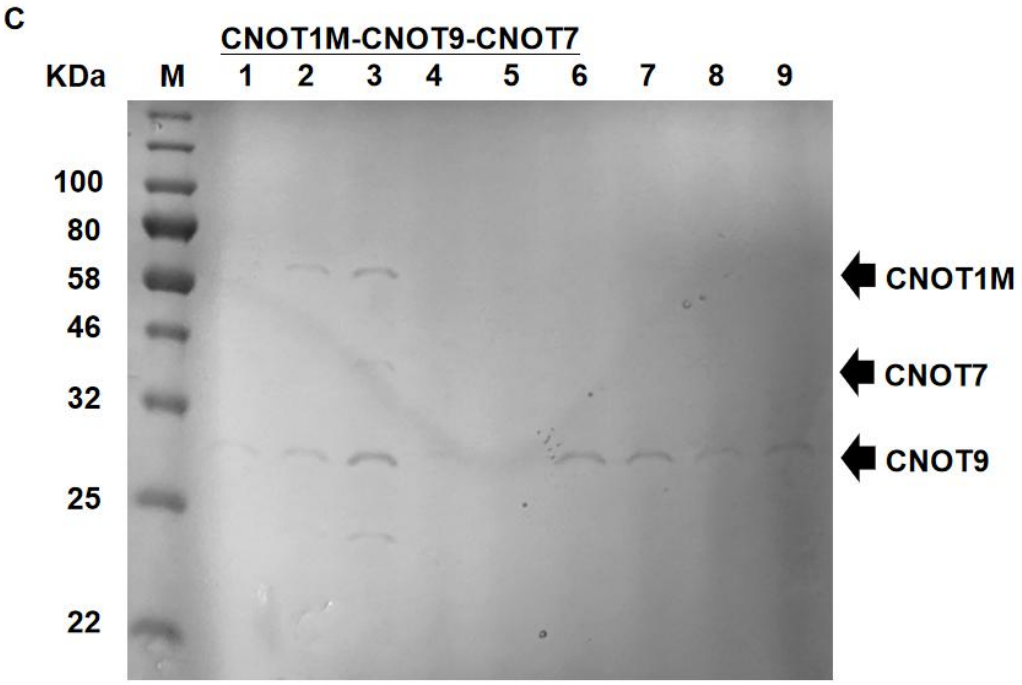
In order to improve the purity of the protein complex, we tried to optimize the conditions of the purification process. Firstly, we modified the binding buffer by increasing the volume of wash buffer and also decreasing the concentration of imidazole in wash buffer. Increasing the volume of wash buffer to 30 mL could increase the purity of the elution fractions. Besides, using the binding buffer without imidazole significantly increases the yield of protein in the elution fractions.

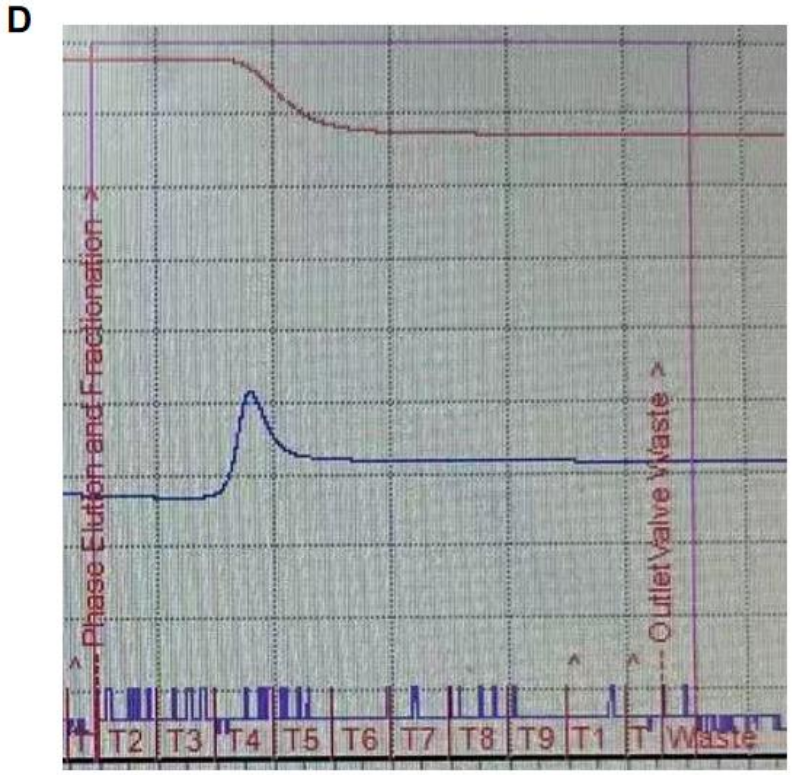
In addition, it was tried to re-purify the protein complex using gel filtration. Firstly, elution buffer containing 250 mM imidazole was removed from the protein fractions using a PD10 desalting column. Then, we carried out gel filtration to separate the protein of interest. The protein sample wasinjected into a 1 mL sample loop using a syringe at flow rate of approximately 2-3 drops per second. Then the sample was applied to a HiPrep 16/60 Sephacryl column at a flow rate of 0.5 mL/min. The column was washed out using buffer without imidazole (20 mM Tris-HCl pH 7.9, 250 mM NaCl, 1 mM β-mercaptoethanol and 5% glycerol). Then the sample was collected using 15 mL tubes (fraction size: 4 mL). SDS-PAGE analysis showed that the fractions were more pure after gel filtration, but there was still a significant level of contaminating proteins present (Figure 3.5C, D). After that, viva-spin column was used to concentrate the protein sample. However, the concentration of the proteins was still not high enough for analyzing structure. According to the Nanodrop, the concentration was around 1.01 mg/mL.

Besides, we also suspect if the deletion of the stop codon process in pACYCDuet1-M01 plasmid or the deletion of the His-tag process in pQE80L-CNOT7 plasmid would influence the expression of the protein complex and further reduce the amount of the protein. So we used plasmids pACYCDuet1-M01 pACYCDuet1-CNOT1M-CNOT9, pQE80L-CNOT7 and pQE80L-CNOT7-NoHis to carry out protein expression separately. Plasmids pACYCDuet1-M01 pACYCDuet1-CNOT1M-CNOT9, pQE80L-CNOT7 and pQE80L-CNOT7-NoHis were transformed into E. coli BL21(DE3) cells separately. For each plasmid, a single colony of E. coli BL21(DE3) was used to inoculate a starter culture for protein expression. The pre-culture was then grown in 1 L of LB containing 35 μg/ml chloramphenicol or 50 μg/ml ampicilin at 37°C until the OD(600 nm) was around 1. Then, IPTG (final concentration of 0.25 mM) was added into the culture to induce protein expression overnight at 28°C. SDS-PAGE analysis of these elution fractions indicated that different plasmids didn't show too much difference (Figure 3.5E), which meant that the deletion process didn't led to the decreasing of the protein yield.


A
CNOT1M-CNOT9-CNOT7
KDa
M
1
2
3
4
5
6
7
100
80
58
46
32
25
CNOT1M
CNOT7
CNOT9


B
Buffer Valve Buffer B
T1
T2
T3
T4
T5
T6
T7
T8
T9
T10


C
CNOT1M-CNOT9-CNOT7
KDa
M
1
2
3
4
5
6
7
8
9
100
80
58
46
32
25
22
CNOT1M
CNOT7
CNOT9


D
Phase Elution and Fractionation
Outlet Valve Waste
T2
T3
T4
T5
T6
T7
T8
T9
T1
Waste

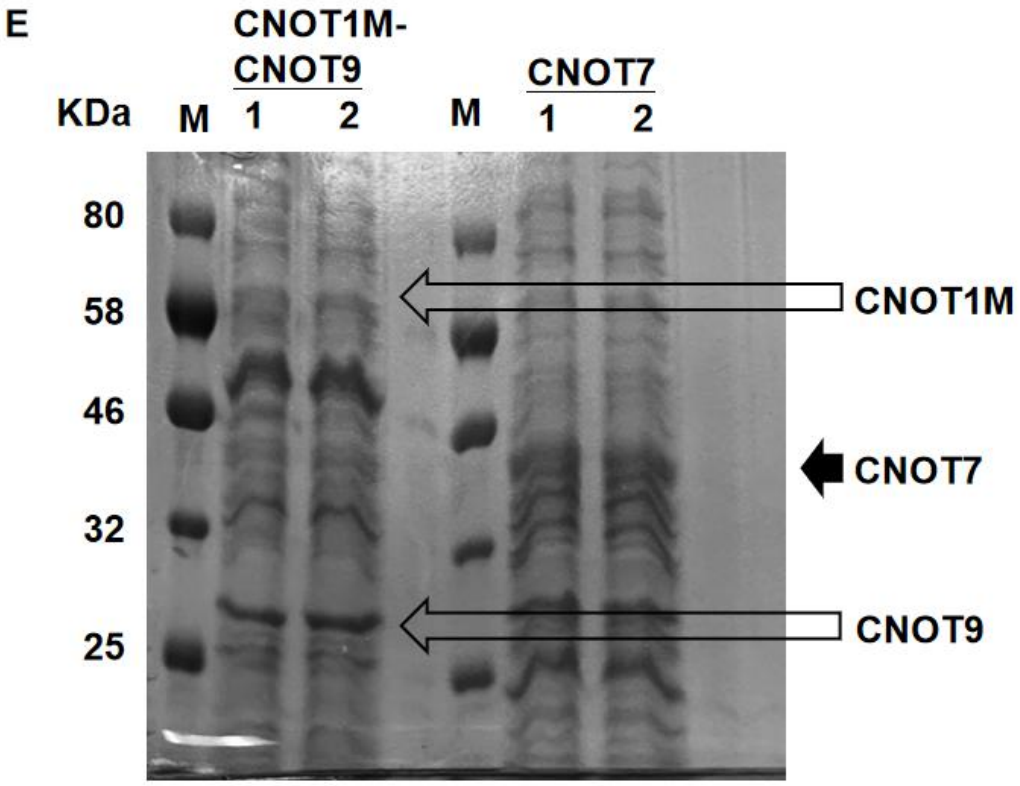


Figure 3.5 Expression and purification of His•CNOT1M-CNOT9-CNOT7 protein complexes. (A) Purification of His•CNOT1M-CNOT9-CNOT7 protein complexes by immobilized metal affinity chromatography (IMAC). Indicated are the marker (M) and protein elution fractions (1-7). **(B)** The elution profile of the elution fractions in Figure 3.5(A). And elution fractions 1-7 was T2-T8 in the elution profile. **(C)** Gel filtration of the purified His•CNOT1M-CNOT9-CNOT7 protein complexes sample. Indicated are the marker (M) and protein elution fractions (1-9). (D) The elution profile of the elution fractions in Figure 3.5(C). And elution fractions 1-9 was T2-T10 in the elution profile. **(E)** Comparison of the protein expression of plasmids pACYCDuet1-M01 pACYCDuet1-CNOT1M-CNOT9, pQE80L-CNOT7 and

pQE80L-CNOT7-NoHis. Indicated are the marker (M) and protein elution fractions (1-2).

3.3 Reconstitution of nuclease complex using purified CNOT1M-CNOT9 and CNOT7

As an alternative strategy to reconstitute the CNOT7-His-CNOT1M-CNOT9 complex, His-CNOT1M-CNOT9 and CNOT7 were expressed separately, which means the two plasmids were transformed separately. Subsequently, the soluble lysates were mixed before purification using IMAC. Expression of CNOT7 was initiated by transferring plasmid pQE80L-CNOT7-NoHis into BL21(DE3). A single colony of E. coli BL21 containing plasmid pQE80L-CNOT7-NoHis was used to inoculate a starter culture for protein expression. The pre-culture was then grown in 1 L of LB containing 100 µg/ml ampicillin at 37°C until the OD(600 nm) was between 0.8 - 1. To induce expression of the protein, IPTG (final concentration of 0.25 mM) was added to the culture for shaking overnight at 28°C. Cells were harvested by centrifugation and the cell pellet was resuspended in Buffer B (described in 2.4.1). Bacterial suspension was lysed on ice using a sonicator. Then the cell lysate containing CNOT7 was frozen and kept at -80°C until further use. Then, we tried to express His-tagged CNOT1M-CNOT9 complex using plasmid pACYCDuet1-CNOT1M-CNOT9. Plasmid pACYCDuet1-CNOT1M-CNOT9 was transformed in E. coli BL21(DE3), and a single colony was used to inoculate a starter culture. Same steps as above were carried out to purify the proteins and then we got the cell lysate of CNOT1M-CNOT9.

After thawing the cell lysate containing CNOT7 and CNOT1M-CNOT9, they were mixed completely by gently rotate. The crude lysate (total lysate) was centrifuged to remove insoluble material. The supernatant was collected using binding buffer containing 20 mM imidazole (Buffer A described in 2.4.1) as soluble lysate. Then protein purification was carried out using immobilized metal affinity chromatography (IMAC). We used the purified sample to conduct the SDS-PAGE to compare the difference between the co-expression and purification of His-tagged CNOT1M-CNOT9 and CNOT7 protein complex and the reconstitution of nuclease complex using purified CNOT1M-CNOT9 and CNOT7. Based on the result, the yield of the protein was much lower than the other strategy (Figure 3.6).

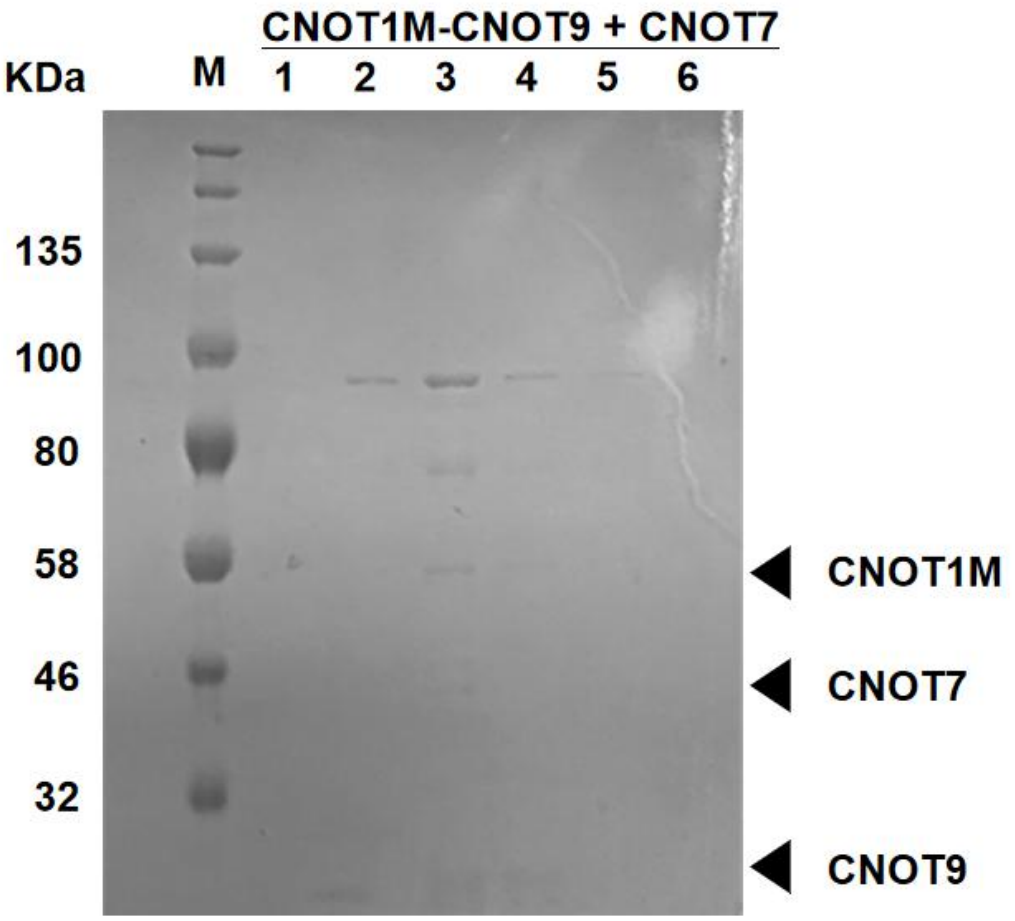


Figure 3.6 SDS-PAGE image of the reconstitution of nuclease complex by mixing purified His-tagged CNOT1M-CNOT9 and un-tagged CNOT7. We used the purified protein to conduct SDS-PAGE. The reconstitution sample of nuclease complex by mixing purified His-tagged CNOT1M-CNOT9 and CNOT7 was loaded the 10% SDS gel. Indicated are the marker (M) and protein elution fractions (1-6).

3.4 Co-expression and purification of His-tagged CNOT1L-CNOT9 and NOTboxes of CNOT3 and CNOT2 protein

complex

3.4.1 Cloning strategy

To investigate and analyse the reconstitution and structure of the His-tagged CNOT1L-CNOT9 and NOTboxes of CNOT3 and CNOT2 protein complex, we used plasmids pCDFDuet1-CNOT3-CNOT2-only NOTboxes and pACYCDuet1-CNOT1L-CNOT9 (Table 3.2).

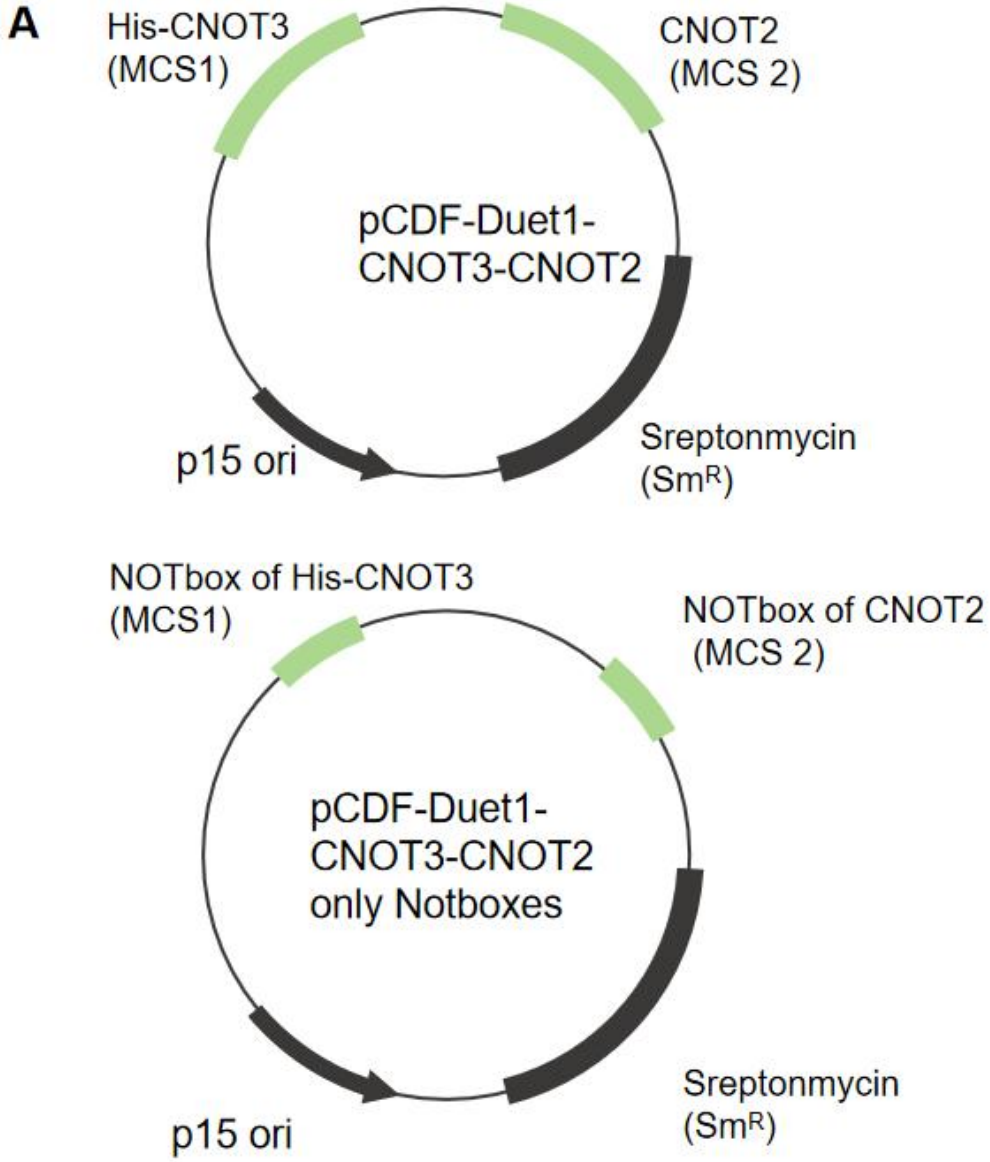
To construct plasmid pCDFDuet1-CNOT3-CNOT2-only NOTboxes, we used pCDFDuet1-CNOT3-CNOT2 as the initial plasmid (Figure 3.7 A). Construction of the plasmid was carried out by Inverse PCR. Firstly, the other part of the CNOT3 except the NOTbox was deleted by Inverse PCR using CNOT3NOT-FW and MCS1-RV primers (Table 2.3). Then, agrose gel electrophoresis was operated to confirm if the deletionof the fragment was successful. The gel image reveled that the band of the plasmids was in expected size (~6000bp). To purify the PCR product, we conducted gel extraction and PCR product cleanup. After that, we conducted Inverse PCR again using CNOT2NOT-FW and MCS2-RV primers (Table 2.3) to delete the other part of CNOT2 except the NOTbox from the ligation product. Same experiments were carried out to check the plasmids pCDFDuet1-CNOT3-CNOT2-only NOTboxes were in expected size (~4800bp). The PCR product were purified using gel extraction kit and PCR product cleanup kit. Then we ligated it and obtained the plasmid pCDFDuet1-CNOT3-CNOT2-only NOTboxes (Figure 3.7A). The plasmid was verified by DNA sequencing.

In order to get the plasmid pACYCDuet1-CNOT1L-CNOT9, plasmids

PUC57-CNOT1L and pACYCDuet1-CNOT1M-CNOT9 were used as template. The restriction enzymes BamHⅠand SalⅠ were used to digest the plasmid pACYCDuet1-CNOT1M-CNOT9 to get the vector pACYCDuet1-CNOT9 (~ 6kb). Then the plasmid PUC57-CNOT1L was digested using the restriction enzymes BglⅡ and SalⅠ or BamHⅠ and SalⅠto got CNOT1 long fragment (~ 3 kb), and then the CNOT1 long fragment was inserted into the BamHⅠand SalⅠ restriction sites of vector pACYCDuet1-CNOT9. To confirm the correct ligation of the cDNA fragment in MCS 1, plasmid pACYCDuet1-CNOT1L-CNOT9 was digested using BamHⅠ and SalⅠ enzymes. Finally, PCRs were conducted to amplify MCS 1 of plasmid pACYCDuet1-CNOT1L-CNOT9 for sequencing. To amplify MCS 1, ACYCDuetUP1 and ACYCDuetDOWN primers were used. As a result, we successfully obtained the plasmid pACYCDuet1-CNOT1L-CNOT9 (Figure 3.7B).

Plasmid	**MCS1**	**MCS2**
PUC57-CNOT1L	CNOT1 long fragment	-
pCDFDuet1-CNOT3-CNOT2	CNOT3	CNOT2
pCDFDuet1-CNOT3-CNOT2-only NOTboxes	NOT box of CNOT3	NOT box of CNOT2
pACYCDuet1-CNOT1L-CNOT9	CNOT1 long fragment	CNOT9
pACYCDuet1-CNOT9	-	CNOT9

Table 3.2 Plasmids using for colony strategy


A
His-CNOT3
(MCS1)
CNOT2
(MCS 2)
pCDF-Duet1-
CNOT3-CNOT2
p15 ori
Sreptonmycin
(SmR)
NOTbox of His-CNOT3
(MCS1)
NOTbox of CNOT2
(MCS 2)
pCDF-Duet1-
CNOT3-CNOT2
only Notboxes
p15 ori
Sreptonmycin
(SmR)

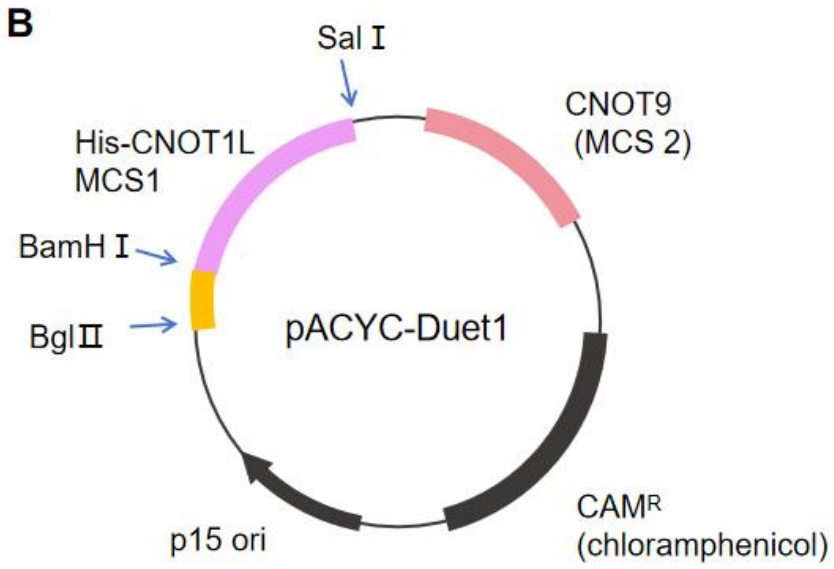


Figure 3.7 Generation of bacterial expression vectors containing the only NOTbox of CNOT3 and CNOT2. The pCDFDuet1 and pACYCDuet1 vectors were used to generate plasmids pCDFDuet1-CNOT3-CNOT2-only NOTboxes and pACYCDuet1-CNOT1L-CNOT9. **(A)** Generation of plasmid pCDFDuet1-CNOT3-CNOT2-only NOTboxes. NOTbox of CNOT3 was subcloned into multiple cloning site 1 (MCS 1) of pCDFDuet1 while NOTbox of CNOT2 was subcloned into multiple cloning site 2 (MCS 2) of pCDFDuet1. **(B)** Generation of plasmid pACYCDuet1-CNOT1L-CNOT9. CNOT1 long fragment (BamH I-Sal I) was subcloned into multiple cloning site 1 (MCS 1) of pACYCDuet1 (BamH I-Sal I).

3.4.2 Protein expression and purification

Protein expression was initiated by transforming plasmids

pCDFDuet1-CNOT3-CNOT2-only NOTboxes and pACYCDuet1-CNOT1L-CNOT9 into E. coli BL21(DE3) cells. Then protein purification was carried out using immobilized metal affinity chromatography (IMAC). The column was washed using 15 mL of pure water and 15 mL of wash buffer containing 40 mM imidazole (Buffer C described in 2.4.1) with a flow rate of 5 mL per minutes. Then the soluble lysate was applied to a His-Trap column through a tubing with a flow rate of approximately 1 mL per minute. After that, the flow rate was transfered to 1 mL per minute and the protein was eluted with 5 ml of elution buffer containing 250 mM imidazole (Buffer B described in 2.4.1) and collected in 1 ml fractions.

The purity of the protein complex was verified using SDS-PAGE and staining with Coomassie Brilliant Blue. However, the result of SDS-PAGE analysis demostrates that the bands of CNOT1L and CNOT9 as well as NOTbox of CNOT3 and CNOT2 were showed in the gel of incorrect size. The CNOT1L protein should have been observed around 65 kDa and the CNOT9 protein should have been observed around 29 kDa, while the NOTbox of CNOT3 protein should have been observed as a single band at ~30 kDa and NOTbox of CNOT2 protein at ~25 kDa. However, SDS-PAGE analysis indicated that the protein complex CNOT1L, Notbox of CNOT2, CNOT9 were in the correct size while the Notbox of CNOT3 did not appear to show the correct stoichiometry. (Figure 3.8).

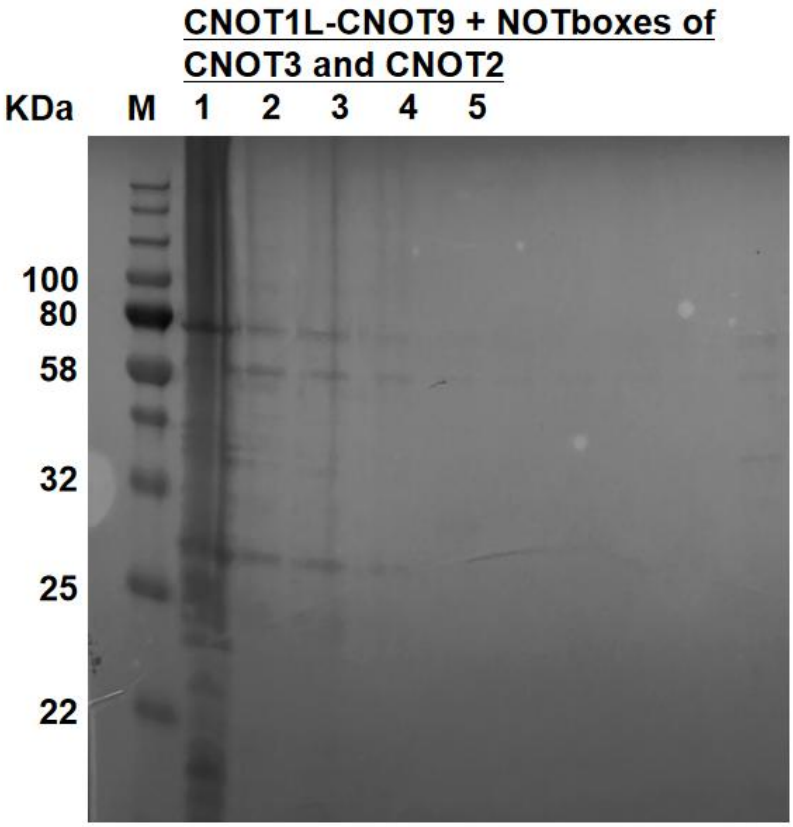


Figure 3.8 Expression and purification of His-tagged CNOT1L-CNOT9 and NOTboxes of CNOT3 and CNOT2 protein complex. Purification of His•CNOT1L-CNOT9 and NOTboxes of CNOT3 and CNOT2 protein complexes by immobilized metal affinity chromatography (IMAC). Indicated are the marker (M) and protein elution fractions (1-5).

3.5 Discussion

In this chapter, different methods are described for the purification of human CNOT1M-CNOT9-CNOT7 mini complexes. Initially, we tried to co-express His•CNOT1M-CNOT9 and un-tagged CNOT7 (Chapter 3.2). When these proteins were co-expressed in BL21 (DE3) and purified using immobilized

metal affinity chromatography (IMAC), we did not obtain samples of sufficient purity. Then we tried to use gel filtration to separate our targeted protein complex based on the size of protein. The results revealed that we successfully obtained the purified protein complex, but the concentration was low. Alternatively, we tried to reconstitute the complex by mixing purified His•CNOT1-CNOT9 and un-tagged CNOT7 (Chapter 3.3). This strategy was successful, but the yield of the target protein decreased sharply. The comparison of the yield of target protein revealed that the first method was better to use.

Moreover, we tried a different extraction method to increase the protein yield. It has been reported that Imidazole is widely used in protein purification, particularly in affinity chromatography for proteins that contain a histidine (His) tag. Different concentration of imidazole can influence the effect it works in protein purification. Low concentration of imidazole (5–20 mM) in the binding or wash buffer helps prevent non-specific proteins from binding without affecting the His-tagged proteins. While Moderate to high concentrations of imidazole (100–500 mM) are used in the elution buffer to release the His-tagged proteins from the resin. Because higher concentration of imidazole may elute target protein complex, decreasing the concentration of imidazole in wash buffer can result in increased yield and purity of target protein. However, while this method increased the yield of the protein of interest, the purified protein was still insufficiently pure. Following the co-expression and purification of full length of CNOT1L-CNOT9 and NOT module of CNOT3 and CNOT2 (Chapter 3.4), which was subsequently carried out, it was purified with binding buffer without imidazole based on the experiment mentioned above.

In summary, two different strategies were employed to co-express and purify human His•CNOT1M-CNOT9-CNOT7. This work represents our attempt to purify distinct subcomplexes of the Ccr4-Not complex, with the aim of further investigating the linker between different protein modules within this complex. The purification strategy described is adapted from a previously reported method used to purify Ccr4-Not nuclease module, containing central region of CNOT1 and CNOT9 (Lorenzo Pavanello et al., 2018). The successful isolation of the human His•CNOT1M-CNOT9-CNOT7 complexes marks a efficient step in the structural analysis of the core Ccr4-Not complex.

Chapter 4

Evidence for Possible Interaction between the YTHDF2 Domain and the Human CNOT9 Subunit

Chapter 4. Evidence for Possible Interaction between the YTHDF2 Domain and the Human CNOT9 Subunit

4.1 Introduction

YTHDF2, which primarily localizes to the cytoplasm, regulates mRNA stability by binding m^6A-modified transcripts, targeting them for degradation. The C-terminal domain of YTHDF2 binds to m^6A sites, while the N-terminal domain recruits degradation machinery, including the Ccr4-Not complex. One of the critical components of this complex is the CNOT9 subunit, also referred to as CAF40 in yeast cells, which is essential for the assembly and functionality of the Ccr4-Not complex [32, 35, 36]. The putative interaction between YTHDF2 and CNOT9 is of particular interest, as it underpins the recruitment of the Ccr4-Not complex to m^6A-modified mRNAs, thus facilitating their degradation. This interaction is crucial for regulating various biological processes, such as cell proliferation and autophagy, by controlling the stability of key mRNAs. For instance, YTHDF2 targets mRNAs with m^6A modifications in the 3' untranslated region (UTR), such as EGFR and autophagy-related transcripts, promoting their degradation and modulating processes like tumor suppression and stem cell regulation [37, 38].

Recent advancements in predictive modeling, such as AlphaFold 3, offer a rapid and accurate means of assessing potential protein-protein interactions [57, 58, 59]. Using AlphaFold Server (https://alphafoldserver.com/), we can predict the structural compatibility between peptide motifs in YTHDF2 and

the surface of CNOT9 as well as Pho92 and CAF40 (interaction in yeast cells). Preliminary AlphaFold-generated models suggest that certain YTHDF2 motifs may complement the binding surface of CNOT9, revealing potential points of interaction that align with known structural and functional domains within each protein (Figure 4.1 A, B, C, D). These motifs are hypothesized to function as docking sites, enabling YTHDF2 to physically interact with CNOT9 and thereby position target mRNAs in proximity to the Ccr4-Not complex's deadenylase machinery. This hypothesis is further supported by functional parallels in other m^6A readers and their interactions with components of the Ccr4-Not complex, which have been shown to enhance mRNA degradation processes [60]. Validating this predicted interaction experimentally will not only confirm AlphaFold's structural insights but also help elucidate how the interplay between YTHDF2 and CNOT9 contributes to the broader regulatory function of the Ccr4-Not complex in mRNA metabolism.

This project focuses on demonstrating the YTHDF2-CNOT9 interaction (Figure 4.2). We carried out peptides pull-down experiments and SDS-PAGE to figure out whether YTHDF2 will interact with CNOT9 protein. By characterizing the structural and functional dynamics of this interaction, we aim to uncover the mechanistic basis of YTHDF2-mediated mRNA degradation via the Ccr4-Not complex. Specifically, we seek to determine if YTHDF2 interacts with CNOT9 to recruit the Ccr4-Not complex, enhancing mRNA destabilization and degradation. A detailed understanding of this interaction will offer significant contributions to the field of post-transcriptional gene regulation, elucidating how m^6A-modified mRNAs are selectively targeted for degradation and how dysregulation of these processes can contribute to diseases such as cancer.

To investigate YTHDF2-CNOT9 interactions, I employed a peptide pull-down assay to investigate how CNOT9 interacts with specific domains of YTHDF2. This approach enables us to define the precise regions of YTHDF2 responsible for binding CNOT9, thereby uncovering the molecular determinants governing their interaction. In this experimental setup, synthetic peptides corresponding to distinct regions of the YTHDF2 protein, particularly those harboring known or predicted binding motifs, were used to conduct the following experiment. These peptides represent various motifs of YTHDF2, such as the m6A recognition domain and potential interaction domains identified through structural modeling or previous literature [55]. Cell lysates containing the Ccr4-Not complex, or purified CNOT9 protein, then are incubated with these immobilized peptides to facilitate binding. Following incubation, the peptide-bound complexes are washed to remove non-specific interactions, and CNOT9 binding are analyzed by Western blotting using specific antibodies against CNOT9. This allows us to identify which YTHDF2 domains are involved in the interaction and the relative strength of these interactions, as well as provide a comprehensive understanding of how CNOT9 recognizes and binds YTHDF2.

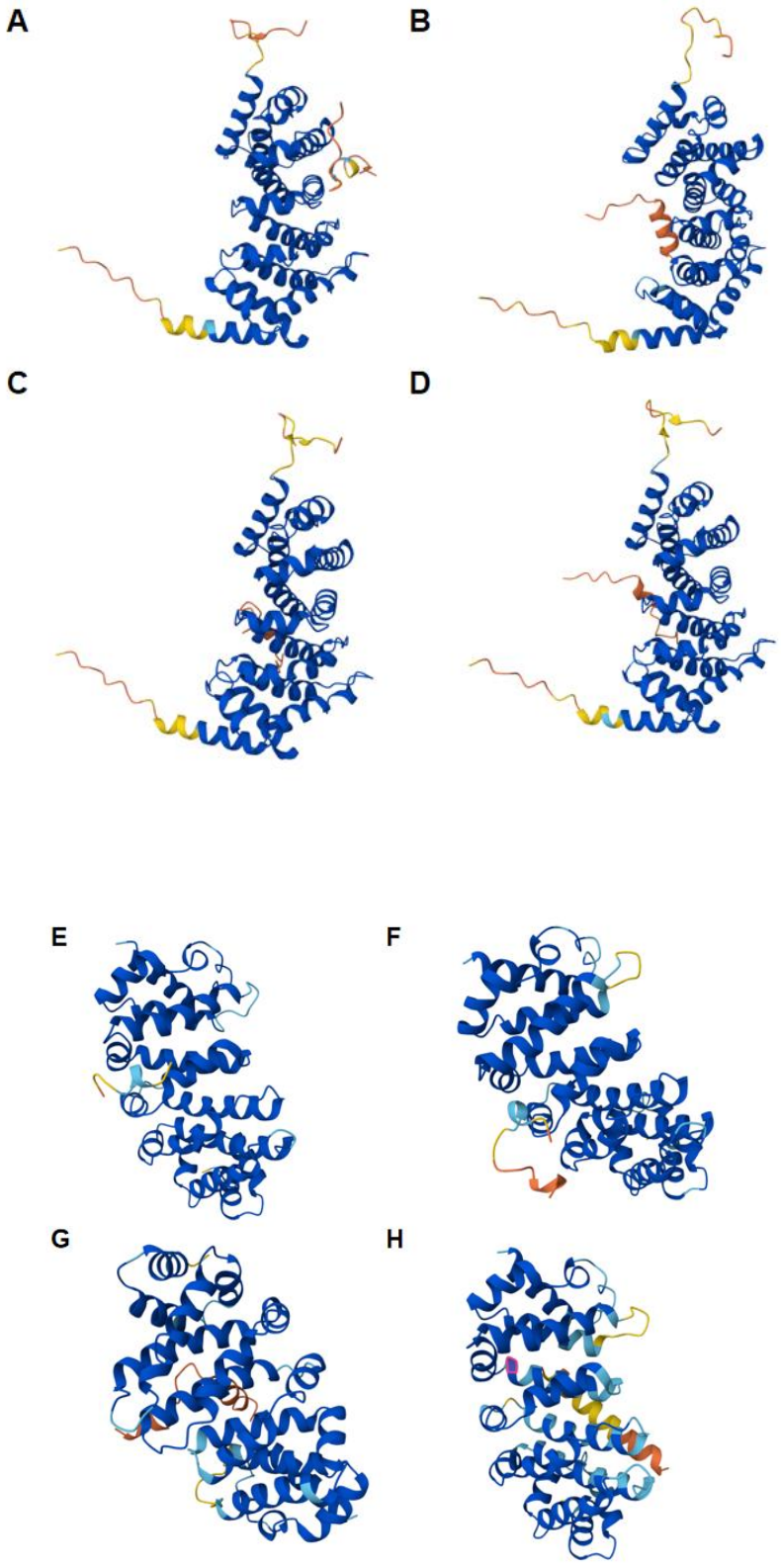

A
B
C
D
E
F
G
H

Figure 4.1 AlphaFold prediction of interaction between different motif of human YTHDF2 and CNOT9. Possibility of interaction: Very high (plDDT > 90) is deep blue, confident (90 > plDDT > 70) is light blue, low (70 > plDDT > 50) is yellow, very low (50 > plDDT) is orange. **(A)** CNOT9 with YTHDF2 65-85. **(B)** CNOT9 with YTHDF2 65-85 W75A). **(C)** CNOT9 with YTHDF2 124-144. **(D)** CNOT9 with YTHDF2 124-144 (W134A). **AlphaFold prediction of interaction between different motif of Yeast Pho92 and CAF40. (E)** CAF40 with Pho 1-10. **(F)** CAF40 with Pho 1-15. **(G)** CAF40 with Pho 1-25. **(H)** CAF40 with Pho 28-52.

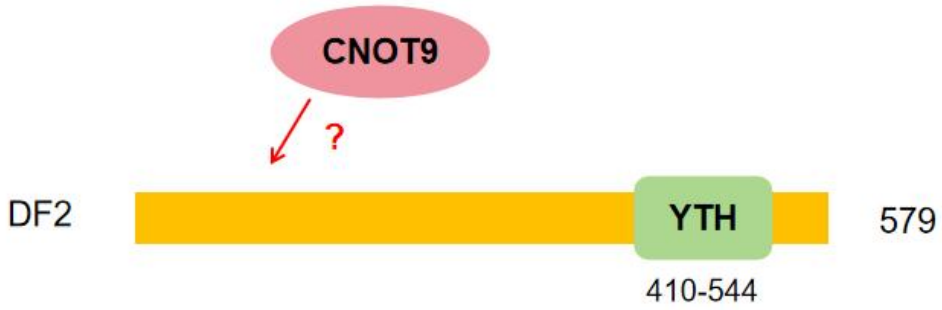


Figure 4.2 Possible interaction between different motif of human YTHDF2 and CNOT9.

4.2 Expression and purification of human CNOT9 subunit

Protein expression was initiated by transforming the plasmid pQE80LCNOT9-ARM (Figure 4.3) into E. coli BL21(DE3) cells. A single colony of E. coli BL21(DE3) was used to inoculate a starter culture for subsequent protein production. The pre-culture was expanded in 1 L of LB Broth medium containing 50 μg/mL ampicillin at 37°C until the optical density (OD) at 600

nm reached approximately 1. At this point, protein expression was induced by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.25 mM, followed by incubation at 28°C overnight. Cells were harvested, and protein was extracted using a wash buffer containing 20 mM imidazole (Buffer A described in 2.4.1). Analysis of the total lysate prior to purification revealed the successful synthesis of the target protein, as evidenced by a band corresponding to the expected molecular weight (Figure 4.4A).

For protein purification, immobilized metal affinity chromatography (IMAC) was employed. The column was washed sequentially with 25 mL of pure water and 25 mL of wash buffer containing 20 mM imidazole with a flow rate of 5 mL per minutes. The soluble lysate was then loaded onto a His-Trap column with tubing at a flow rate of approximately 1 mL/min. Subsequently, the flow rate was reduced to 1 mL/min, and the protein was eluted using an elution buffer containing 250 mM imidazole, with 1 mL fractions collected. The purity of the eluted protein was assessed by SDS-PAGE, followed by Coomassie Brilliant Blue staining. The SDS-PAGE analysis revealed a prominent band corresponding to CNOT9 (~29 kDa), confirming the presence of the target protein (Figure 4.4B). However, significant levels of co-purifying contaminants were still present in the fractions containing the CNOT9.

To enhance the purity of the protein complex, various optimizations of the purification protocol were undertaken. First, the volume of wash buffer was increased, and the imidazole concentration was adjusted. Increasing the wash buffer volume to 30 mL resulted in improved purity of the eluted fractions.

Additionally, eliminating imidazole from the wash buffer markedly improved the yield of the target protein in the elution fractions. After that, elution buffer containing 250 mM imidazole was removed from the protein fractions using a PD10 desalting column. Then the protein sample was concentrated through viva-spin column. The final concentration was measured using Bradford Protein Assay. The final concentration was 10 mg/mL.

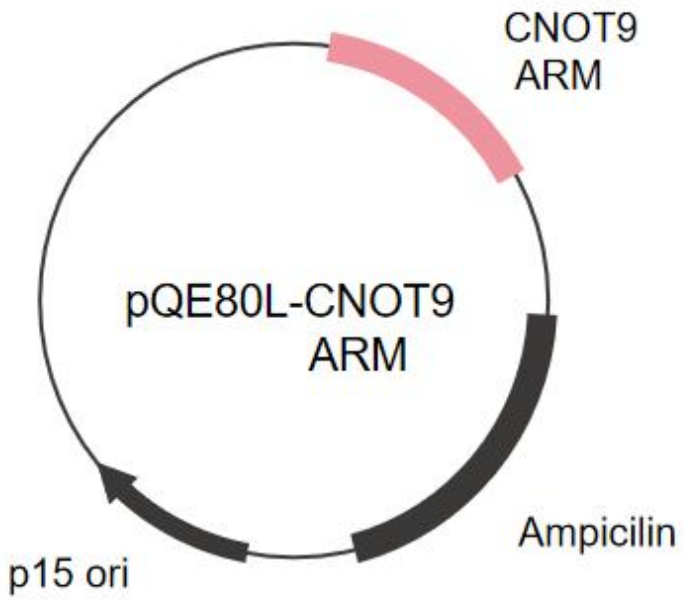
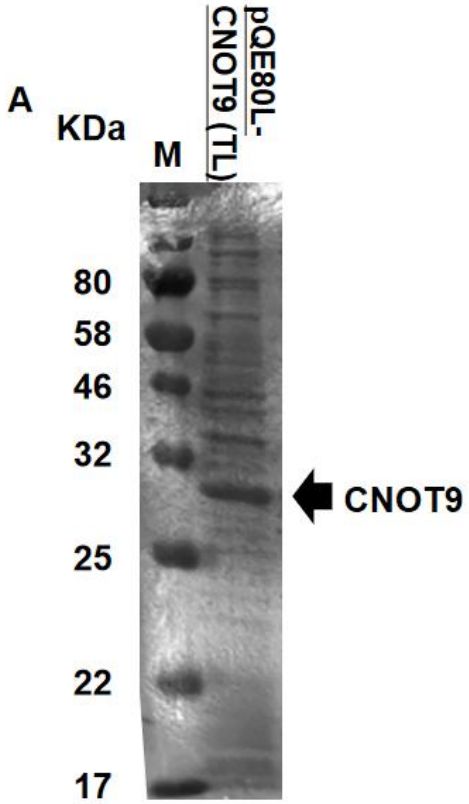


Figure 4.3 Generation of bacterial expression vectors containing the CNOT9 cDNAs. The pQE80L vector was used to generate plasmids pQE80L-CNOT9-ARM.


A
KDa
M
pQE80L-CNOT9 (TL)
80
58
46
32
25
22
17
CNOT9

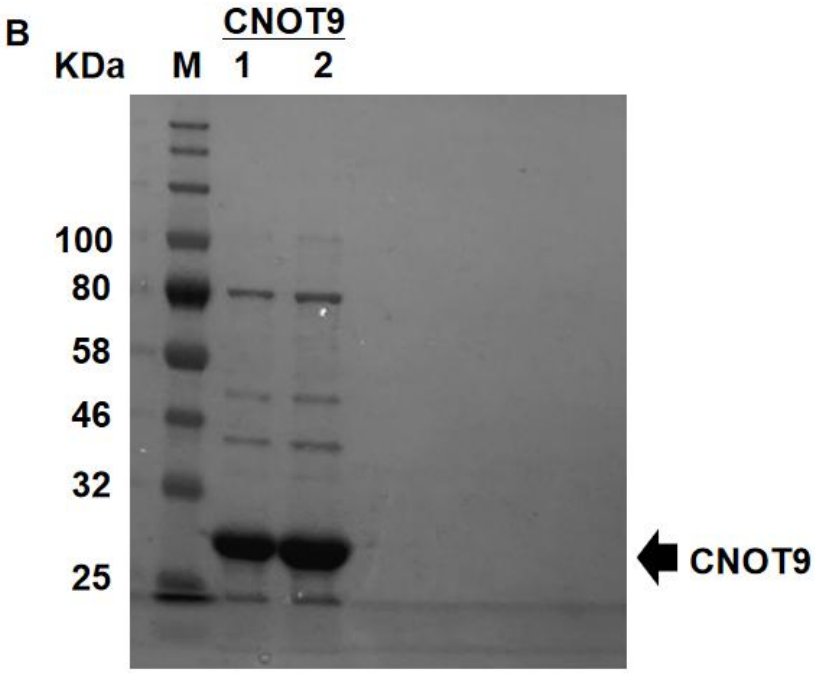

B
CNOT9
KDa
M
1
2
100
80
58
46
32
25
CNOT9

Figure 4.4 Expression and purification of CNOT9 protein complexes. (A) Lysates were obtained from 1 litre of culture of BL21(DE3) after expression was induced for 12 h with 0.25 mM IPTG. Lysates (10 µL) were analysed by 10% SDS-PAGE followed by staining with Coomassie Briliant Blue. Indicated are the marker (M), total lysate (TL). **(B)** Purification of CNOT9 protein complexes by immobilized metal affinity chromatography (IMAC). Indicated are the marker (M) and protein elution fractions, sample 1 (1µL) and sample 2 (2 µL).

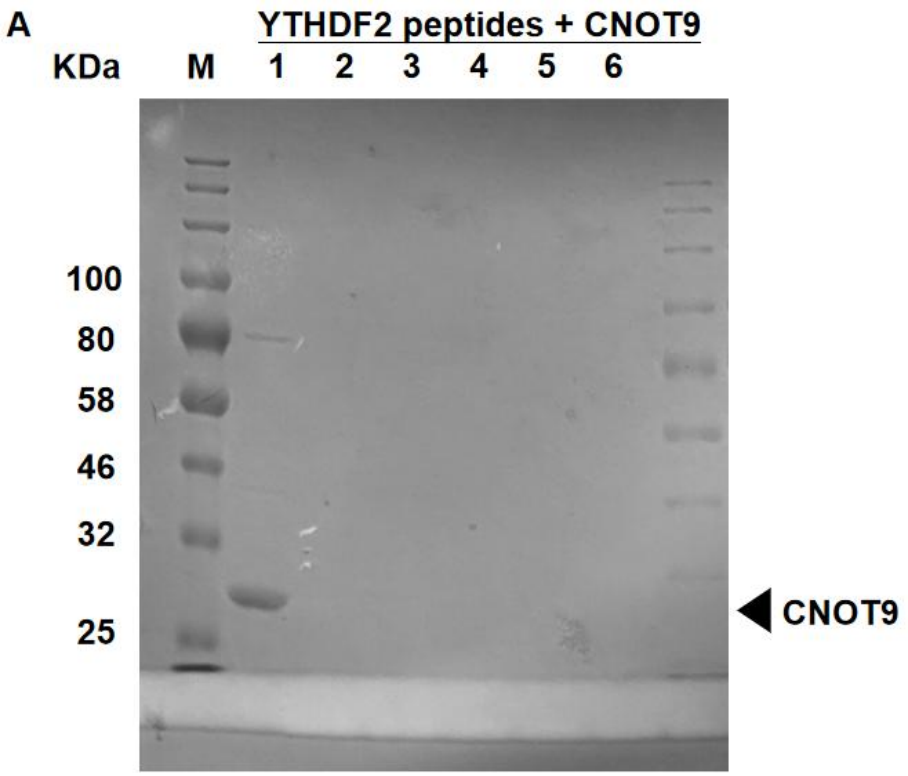
4.3 pull-down of YTHDF2 peptides with CNOT9

The YTHDF2 peptides (YTHDF2 65 - 85aa, YTHDF2 65 - 85aa (W75A), YTHDF2 124 -144aa, YTHDF2 124 -144aa (W134A)) were initially dissolved in 30 - 50 µL of 100% DMSO and then further diluted to 10% DMSO with pure water. Also, 0.2% NP-40 was applied when the peptides didn't dissolve completely. 15 µL of Dynabeads™ M-280 was taken into a fresh tube and was separated using a magnetic rack so that the storage buffer was discarded, and they were washed by resuspending with protein binding buffer (described in 2.9.1). After that, 1.5 µg YTHDF2 peptide or no peptide control were added to the Dynabeads and then they were incubated to allow binding for 1 hour with rotation at 4 °C in 450 µl binding buffer. After that, 20 µg CNOT9 protein (with His-tag) (1mg/mL) was then added to the Dynabeads bound with YTHDF2 peptide and incubated with gentle rotation for 1-2 hours at 4 °C. Finally, Dynabeads were then harvested using a magnet and consequently washed with wash buffer and binding buffer respectively (Chapter 2.9). Then the interaction system with 2 µg CNOT9 protein (10% of the entire CNOT9

protein sample, served as control) were then subjected to SDS-PAGE and followed by Coomassie blue staining or carried out Western Blot experiment to check the pull-down result. Coomassie blue stained gel was then photographed using the imager (Figure 4.5A). However, the image of the gel didn't show any bands.

In order to detect the interaction between YTHDF2 domain and CNOT9 in a more sensitive method, we therefore conducted Western Blot to observe the immunoblotting of the reaction. The reaction sample, no peptide control and 0.2 μg CNOT9 protein (1% of the entire CNOT9 proteinsample, served as control) were subjected to SDS-PAGE. After electrophoresis, we conducted membrane transfer to transfer the protein form the gel to the PVDF membrane and the transfer is performed in 1x transfer buffer at 25 V for 1 hour. Then the membrane was blocked to prevent non-specific binding of antibodies using 5% non-fat milk or BSA in TBST at 4°C for 1 hour. After that, the membrane is incubated with 6×-His Tag Monoclonal Antibody (1:1000 ditution in 5% non-fat milk/TBST) overnight at 4°C with gentle shaking. After washing the mambrane 3 times with TBST to remove unbound primary antibody, the membrane was incubated with anti-mouse HPR conjugated antibody for 1 hour with constant rocking. Then the membrane was washed 3 times with TBST for 10 minutes each. After exposing to X-ray film, the protein bands corresponding to the target proteins are visualized and analyzed (Figure 4.5B). As shown in the Figure 4.5B, band was only detected in the interaction of CNOT9 with peptide YTHDF2 124-144aa and YTHDF2 65 - 85aa (W75A), which was not observed in the interaction of CNOT9 with peptide YTHDF2 65-85aa and YTHDF2 124 - 144aa (W134A). The data revealed that YTHDF2 domain can interact with the human CNOT9 protein.

More specifically, according to the result of WesternBlot, the YTHDF2 124 - 144aa and YTHDF2 65 - 85aa (W75A) peptide in the YTHDF2 domain was able to interact with human CNOT9.



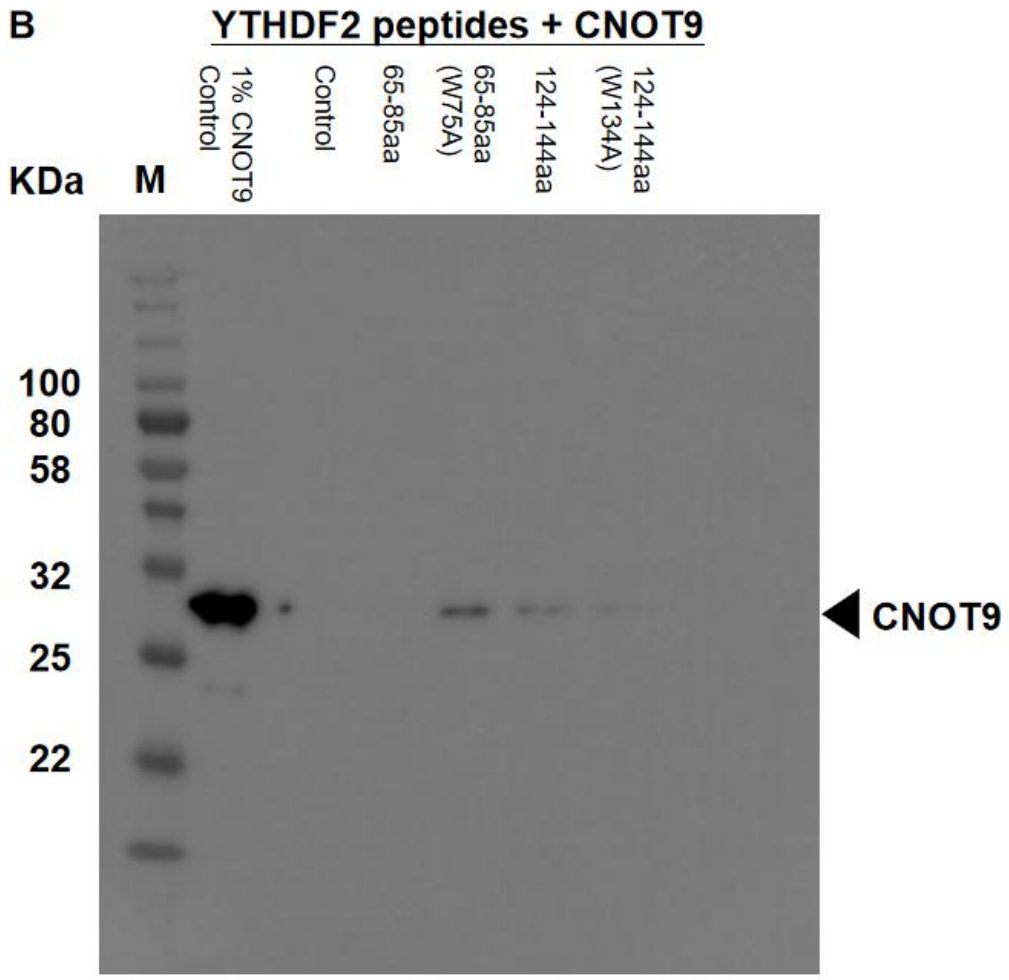


Figure 4.5 YTHDF2 peptides pull-down with human CNOT9 protein. (A) 10% of the CNOT9 protein sample was loaded in the sample 1. Interaction sample without adding peptide was loaded in the sample 2. Sample 3 ~ sample 6 was interaction between CNOT9 and different domain of YTHDF2, which was YTHDF2 65 - 85aa, YTHDF2 65 - 85aa (W75A), YTHDF2 124 - 144aa, YTHDF2 124 - 144aa (W134A) respectively. **(B)** 1% of the CNOT9 protein sample was loaded in the sample 1. Interaction sample without adding peptide was loaded in the sample 2. Sample 3 ~ sample 6 was

interaction between CNOT9 and different domain of YTHDF2, which was YTHDF2 65 - 85aa, YTHDF2 65 - 85aa (W75A), YTHDF2 124 - 144aa, YTHDF2 124 - 144aa (W134A) respectively.

4.4 Discussion

Having successfully purified human CNOT9 protein (Chapter 4.2), it was possible to conduct peptide pull-down assays to explore the interaction between human CNOT9 and YTHDF2 domain. I tried to observe the interaction through SDS-PAGE and coomassie blue staining. However, there were no bands visible in the gel. Then we turned to carried out western blot to show the immunoblotting of the reaction, which was more sensitive and precise.

In summary, peptide pull-down and western blot were employed to explore the interaction between the human CNOT9 and YTHDF2 domain. Previous researches have revealed that YTHDF2 recruits the CCR4-NOT complex through a direct interaction between the YTHDF2 N-terminal region and the SH domain of the CNOT1 subunit [55], and that this recruitment is essential for the deadenylation of m(6)A-containing RNAs by CAF1 and CCR4 [56]. To understand in more details how YTHDF2 interacts with the Ccr4-Not complex, we explored their interaction by peptide pull-down assay. We further demonstrated that human CNOT9 can interact with YTHDF2, and this interaction is achieved through the peptide YTHDF2 124 - 144aa, which suggests YTHDF2 recruits the CCR4-NOT complex through a direct interaction between its 124 - 144aa region and the CNOT9 subunit. There also showed interaction between YTHDF2 65-85 (W75A), but this is

unexpected since the interaction is supposed to be mediated by the tryptophan residue [62]. What's more, results of the AlphaFold-generated models predicted that it would be 124-144aa, rather than 65-85aa, more likely to bind the CNOT9, which means this interaction of my experiments were not so convinced. More detailed experiments should be conducted in the future. Besides, after preliminary AlphaFold-generated model prediction between yeast protein CAF40 and different motifs of Pho92 (Figure 4.1 E, F, G, H), I found the possibility of interaction between CAF40 and Pho92 was more putative (since the prediction models of interaction between CAF40 and Pho92 are in deep blue, which suggested a high possibility of interaction).

Chapter 5

Conclusion and Future Work

Chapter 5. Conclusion and Future Work

5.1 Conclusion

The CCR4–NOT complex is a pivotal multi-subunit assembly involved in mRNA deadenylation and gene expression regulation. Central to its function is the scaffold protein CNOT1, which orchestrates the assembly of various subunits, including the deadenylases CNOT6/6L and CNOT7/8, as well as the non-catalytic subunit CNOT9. Structural studies have elucidated interactions between CNOT1 and these subunits, highlighting the complex's role in post-transcriptional gene regulation. Despite these insights, the precise mechanisms governing the assembly and function of specific CCR4–NOT modules, such as CNOT1M–CNOT9–CNOT7 and CNOT1L–CNOT9–NOT, remain incompletely understood. Additionally, the interaction dynamics between CNOT9 and RNA-binding proteins like YTHDF2, which may influence mRNA decay pathways, warrant further investigation. In this study, I aimed to reconstitute and purify the CNOT1M–CNOT9–CNOT7 and CNOT1L–CNOT9–NOT modules to better understand their structure. I also sought to characterize the interaction between human CNOT9 and the YTHDF2 protein.

Specifically speaking, I explored the reconstitution and purification of the CNOT1M-CNOT9-CNOT7 and CNOT1L-CNOT9-NOT modules of the Ccr4-Not complex (Chapter 3), which play crucial roles in mRNA degradation and gene expression regulation. Our findings contribute efficiently to the structural and functional understanding of these modules, particularly in relation to the

CNOT1 scaffold, which mediates interactions between key subunits such as CNOT6/7 and CNOT4. We demonstrated two primary strategies for purifying the CNOT1M-CNOT9-CNOT7 protein complex. However, base on purification work before [21], it's not convinced that whether these complexes are right proteins. Because although all the complexes were in the correct size, the amount of CNOT1M was less than the amount CNOT9, which was completely oposite in the previous work. So futher experiments like WesternBlot need to be done to double check whether the complexes are the correct proteins. Apart from this, both co-expression and reconstitution strategies yielded purified complexes, but the co-expression method using immobilized metal affinity chromatography (IMAC) followed by gel filtration resulted in better yields and purity. The challenges I encountered, particularly in low yields and impurities, were partly mitigated by optimization of purification conditions, such as varying imidazole concentrations. As for the CNOT1L-CNOT9-NOT module, despite our efforts, achieving high purity and yield of the full-length CNOT1L-CNOT9 complex and its associated NOT modules remains a challenge. This research sets the stage for future efforts to further optimize purification strategies and explore the structural dynamics of the CNOT complex in greater detail.

The experiments also explored the interaction between human CNOT9 and YTHDF2 domain (Chapter4). The image of the immunoblotting showed that Human CNOT9 interacts with YTHDF2, and this interaction is mediated through the 124-144 amino acid region of YTHDF2. Specifically, YTHDF2 recruits the CCR4-NOT complex by directly interacting with the CNOT9 subunit via its 124-144 amino acid segmen. Although the results also indicated that there was interaction between YTHDF2 65-85aa (W75A) motif.

This is unexpected because the interaction is supposed to be mediated by the tryptophan residue [62].

In summary, these results enhance our understanding of the structural composition and interaction networks within the CCR4–NOT complex, providing a foundation for future studies aimed at elucidating the mechanisms of mRNA regulation mediated by this complex.

5.2 Future Outlook

Future research should address challenges in purifying full-length CNOT1L-CNOT9-NOT modules, with a focus on optimizing yield and purity. My work contributes to this research by advancing purification protocols and enhancing our understanding of the functional interactions within the Ccr4-Not complex. One approach for improving yield could involve exploring alternative expression systems, including eukaryotic systems that better mimic the native cellular environment of the Ccr4-Not complex, which may increase the stability and functional integrity of these proteins. In addition to purification, investigating the flexibility and functional role of the linker between CNOT1M-CNOT9-CNOT7 modules is crucial. Further studies could focus on the functional implications of linker flexibility using structural techniques to reveal high-resolution conformational changes [65]. Mutagenesis studies targeting key residues within the linker could provide crucial insights into their roles in stabilizing the structure and influencing the enzymatic activity of the Ccr4-Not complex. To complement these studies, biochemical activity assays [6] could be applied to analyze the functional

activity of the complex in greater detail. Combining these approaches would enable a more comprehensive understanding of how specific structural changes impact substrate processing and enzymatic efficiency. This integrated strategy could reveal how the interplay between structural modules and linker regions drives mRNA degradation pathways, offering valuable insights into the complex's overall functionality.

The purified protein domains will undergo structural analysis at the Diamond Light Source Research Institution (https://www.diamond.ac.uk/Home.html) using small-angle X-ray scattering (SAXS) to obtain detailed insights into the structural characteristics of the protein linker region within the Ccr4-Not complex [63]. SAXS will allow me to investigate the flexibility, shape, and dimensions of the linker in solution, complementing our understanding of how it impacts the spatial arrangement and interactions between the CNOT1L-CNOT9-NOT modules. Following SAXS, advanced structural biology techniques, including cryo-electron microscopy (cryo-EM), X-ray crystallography, will be considered to gain high-resolution structural details. Cryo-EM [64], in particular, will be instrumental in visualizing the complex in various conformational states, allowing us to observe the linker region's flexibility and its role in modulating the overall architecture of the Ccr4-Not complex. Cryo-EM's ability to capture dynamic structural states theoretically may provide invaluable insights into how these structural changes impact mRNA degradation activity and substrate binding. Additionally, cross-linking mass spectrometry (XL-MS) and Förster resonance energy transfer (FRET) will be applied to explore spatial relationships and dynamic interactions between domains [65]. XL-MS will help map interaction sites by cross-linking proximal residues, providing distance constraints that can inform models of

the complex's structure. FRET can further enhance our understanding of the linker's flexibility by measuring distance changes between fluorescently labeled domains, revealing how structural rearrangements may affect the complex's enzymatic functions. By integrating these techniques with previous structural studies, we aim to comprehensively map the spatial organization and domain-specific interactions within the Ccr4-Not complex. This integrative approach will not only clarify the functional relevance of the linker region but also advance our understanding of how the complex's structural dynamics contribute to its regulatory roles in mRNA metabolism.

The molecular mechanisms driving Ccr4-Not-mediated deadenylation remain incompletely understood. Based on the successful purification of human CNOT9 and the preliminary results of its interaction with the motifs of YTHDF2 by peptide pull-down assays and western blot analysis, several avenues for future research has emerged. First, structural studies are needed to further elucidate the molecular interface between CNOT9 and YTHDF2, with a particular focus on the identified YTHDF2 124-144aa region. Techniques such as X-ray crystallography [68, 69] could provide high-resolution information on the binding mechanism and the structural impact of this interaction on mRNA degradation [64]. In addition, mutagenesis studies targeting key residues in the 124-144aa region of YTHDF2 and CNOT9 should be performed to investigate the functional significance of this interaction and the effects of specific mutants on Ccr4-Not complex recruitment and mRNA deadenylation could be further analyzed. These studies will help to elucidate how specific domains affect the overall

function of the Ccr4-Not complex in post-transcriptional gene regulation. Moreover, relevant experiments need to be operated to find evidence for putative interaction between yeast CAF40 and Pho92 based on the prediction of AlphaFold Serve.

To further investigate the biological significance of the CNOT9-YTHDF2 interaction, additional techniques could be employed to confirm and expand upon findings that CNOT9 interacts with YTHDF2 via the 124-144 amino acid motif. Structural studies, such as cryo-electron microscopy (cryo-EM) or NMR spectroscopy, could provide detailed insights into the molecular basis of this interaction. Moreover, in vitro binding assays using purified proteins could validate the direct interaction and assess its affinity and specificity.

In vivo studies using cell lines with targeted knockouts or knock-ins of CNOT9-YTHDF2 interaction partners would also be invaluable. For example, generating cell lines with point mutations specifically disrupting the 124-144aa motif of CNOT9 could reveal its functional role. These modified cell lines could then be analyzed to observe changes in mRNA stability, degradation rates, and cellular responses, shedding light on the interaction’s physiological relevance. Reporter assays, such as tethering assays (Doidge et al., 2012, PLOS One), could further elucidate the impact of this interaction. By recruiting YTHDF2 to specific mRNAs, these assays would allow direct observation of changes in mRNA stability and degradation pathways. Combining these approaches—structural studies, in vitro assays, and advanced cell-based models—would provide a robust framework for understanding the role of the CNOT9-YTHDF2 interaction in mRNA metabolism. [66] Complementary techniques, including RNA

immunoprecipitation (RIP) and mRNA degradation assays, can help elucidate how disruption of the CNOT9-YTHDF2 interface affects the stability of m6A-modified mRNAs specifically [67]. By examining the broader CCR4-NOT complex, researchers can potentially identify interconnected pathways or regulatory networks that influence mRNA decay mechanisms. Extending these investigations to additional components of the complex may reveal how these networks coordinate to regulate gene expression at various stages of cell differentiation and stress response.

Since there are links between YTH proteins, mRNA decay and translation in the literature [62]. My findings may add a mechanism to link YTH proteins and Ccr4-Not. Beyond the molecular level, exploring the physiological relevance of CNOT9-YTHDF2 interactions in different cellular contexts, such as stress responses or pathological conditions like cancer, could offer new therapeutic insights. For instance, YTHDF2 and other YTH proteins are implicated in acute myeloid leukemia (AML) and other blood cancers, where they contribute to mRNA stability and translational control of oncogenes and tumor suppressor genes. Identifying the specific regulatory roles of YTHDF2 in these cancers could help identify context-specific vulnerabilities, offering a foundation for targeted therapies. While developing small molecules or peptides to modulate the CNOT9-YTHDF2 interaction presents a therapeutic strategy, it carries the potential risk of off-target effects and toxicity, given YTHDF2's involvement in multiple cellular pathways. Careful targeting of the interaction would be necessary to avoid undesirable interference with other YTHDF2-mediated interactions, particularly in normal hematopoiesis and other critical physiological processes.

References

[1] Cramer, P. Organization and regulation of gene transcription. Nature 573, 45–54 (2019). https://doi.org/10.1038/s41586-019-1517-4

[2] Ramanathan A, Robb GB, Chan SH. mRNA capping: biological functions and applications. Nucleic Acids Res. 2016 Sep 19;44(16):7511-26. doi: 10.1093/nar/gkw551. Epub 2016 Jun 17. PMID: 27317694; PMCID: PMC5027499.

[3] McManus CJ, Graveley BR. RNA structure and the mechanisms of alternative splicing. Curr Opin Genet Dev. 2011;21(4):373-379. doi:10.1016/j.gde.2011.04.001

[4] Passmore LA, Coller J. Roles of mRNA poly(A) tails in regulation of eukaryotic gene expression. Nat Rev Mol Cell Biol. 2022 Feb;23(2):93-106. doi: 10.1038/s41580-021-00417-y. Epub 2021 Sep 30. PMID: 34594027; PMCID: PMC7614307.

[5] Maryati M, Kaur I, Jadhav GP, Olotu-Umoren L, Oveh B, Hashmi L, Fischer PM, Winkler GS. A fluorescence-based assay suitable for quantitative analysis of deadenylase enzyme activity. Nucleic Acids Res. 2014 Mar;42(5):e30. doi: 10.1093/nar/gkt972. Epub 2013 Oct 28. PMID: 24170810; PMCID: PMC3950723.

[6] Pavanello L, Hall B, Airhihen B, Winkler GS. The central region of CNOT1 and CNOT9 stimulates deadenylation by the Ccr4-Not nuclease module. Biochem J. 2018 Nov 9;475(21):3437-3450. doi: 10.1042/BCJ20180456. PMID: 30309886.

[7] Collart MA, Panasenko OO. The Ccr4-Not Complex: Architecture and Structural Insights. Subcell Biochem. 2017;83:349-379. doi: 10.1007/978-3-319-46503-6_13. PMID: 28271483.

[8] Statello L, Guo CJ, Chen LL, Huarte M. Gene regulation by long non-coding RNAs and its biological functions. Nat Rev Mol Cell Biol. 2021 Feb;22(2):96-118. doi: 10.1038/s41580-020-00315-9. Epub 2020 Dec 22. Erratum in: Nat Rev Mol Cell Biol. 2021 Feb;22(2):159. doi:

10.1038/s41580-021-00330-4. PMID: 33353982; PMCID: PMC7754182.

[9] Faraji F, Hu Y, Yang HH, Lee MP, Winkler GS, Hafner M, Hunter KW. Post-transcriptional Control of Tumor Cell Autonomous Metastatic Potential by CCR4-NOT Deadenylase CNOT7. PLoS Genet. 2016 Jan 25;12(1):e1005820. doi: 10.1371/journal.pgen.1005820. PMID: 26807845; PMCID: PMC4726497.

[10] Chalabi Hagkarim N, Grand RJ. The Regulatory Properties of the Ccr4-Not Complex. Cells. 2020 Oct 29;9(11):2379. doi: 10.3390/cells9112379. PMID: 33138308; PMCID: PMC7692201.Inada, T.; Makino, S. Novel roles of the multi-functional CCR4-NOT complex in post-transcriptional regulation. Front. Genet. 2014, 5, 135.

[11] Petit A.-P., Wohlbold L., Bawankar P., Huntzinger E., Schmidt S., Izaurralde E., Weichenrieder O. The structural basis for the interaction between the CAF1 nuclease and the NOT1 scaffold of the human CCR4–NOT deadenylase complex. Nucleic Acids Res. 2012;40:11058–11072. doi: 10.1093/nar/gks883.

[12] Basquin J., Roudko V.V., Rode M., Basquin C., Séraphin B., Conti E. Architecture of the nuclease module of the yeast Ccr4-not complex: The Not1-Caf1-Ccr4 interaction. Mol. Cell. 2012;48:207–218. doi: 10.1016/j.molcel.2012.08.014.

[13] Shirai Y.-T., Suzuki T., Morita M., Takahashi A., Yamamoto T. Multifunctional roles of the mammalian CCR4–NOT complex in physiological phenomena. Front. Genet. 2014;5:286. doi: 10.3389/fgene.2014.00286.

[14] Winkler G.S. The mammalian anti-proliferative BTG/Tob protein family. J. Cell. Phys. 2010;222:66–72. doi: 10.1002/jcp.21919. [PubMed] [CrossRef] [Google Scholar]

[15] Mauxion F, Faux C, Séraphin B. The BTG2 protein is a general activator of mRNA deadenylation. EMBO J. 2008 Apr 9;27(7):1039-48. doi: 10.1038/emboj.2008.43. Epub 2008 Mar 13. PMID: 18337750; PMCID: PMC2323266.

[16] Yang X., Morita M., Wang H., Suzuki T., Yang W., Luo Y., Zhao C., Yu Y., Bartlam M., Yamamoto T., et al. Crystal structures of human BTG2 and mouse TIS21 involved in suppression of CAF1 deadenylase activity. Nucleic Acids Res. 2008;36:6872–6881. doi: 10.1093/nar/gkn825.

[17] Miyasaka T., Morita M., Ito K., Suzuki T., Fukuda H., Takeda S., Inoue J., Semba K., Yamamoto T. Interaction of antiproliferative protein Tob with the CCR4-NOT deadenylase complex. Cancer Sci. 2008;99:755–761. doi: 10.1111/j.1349-7006.2008.00746.x.

[18] Collart M.A. The Ccr4-Not complex is a key regulator of eukaryotic gene expression. Wiley Interdiscip. Rev. RNA. 2016;7:438–454. doi: 10.1002/wrna.1332.

[19] Sgromo, A., Raisch, T., Bawankar, P., Bhandari, D., Chen, Y., Kuzuoglu-Ozturk, D. et al (2017) A CAF40-binding motif facilitates recruitment of the CCR4-NOT complex to mRNAs targeted by Drosophila Roquin. Nat. Commun. 8, 14307

[20] Chen, Y., Boland, A., Kuzuoglu-Ozturk, D., Bawankar, P., Loh, B., Chang, C.T. et al (2014) A DDX6-CNOT1 complex and W-binding pockets in CNOT9 reveal direct links between miRNA target recognition and silencing. Mol. Cell 54, 737–750

[21] Bulbrook, D., Brazier, H., Mahajan, P., Kliszczak, M., Fedorov, O., Marchese, F.P. et al (2018) Tryptophan-mediated interactions between tristetraprolin and the CNOT9 subunit are required for CCR4-NOT deadenylase complex recruitment. J. Mol. Biol. 430, 722–736

[22] Horiuchi M, Takeuchi K, Noda N, Muroya N, Suzuki T, Nakamura T, et al. Structural basis for the antiproliferative activity of the Tob-hCaf1 complex. The Journal of biological chemistry. 2009;284(19):13244–55. Epub 2009/03/12. pmid:19276069; PubMed Central PMCID: PMC2676056.

[23] Boland A, Chen Y, Raisch T, Jonas S, Kuzuoğlu-Öztürk D, Wohlbold L, Weichenrieder O, Izaurralde E. Structure and assembly of the NOT module of the human CCR4-NOT complex. Nat Struct Mol Biol. 2013 Nov;20(11):1289-97. doi: 10.1038/nsmb.2681. Epub 2013 Oct 13. PMID:

24121232.

[24] Zhang Q, Pavanello L, Potapov A, Bartlam M, Winkler GS. Structure of the human Ccr4-Not nuclease module using X-ray crystallography and electron paramagnetic resonance spectroscopy distance measurements. Protein Sci. 2022 Mar;31(3):758-764. doi: 10.1002/pro.4262. Epub 2022 Jan 7. PMID: 34923703; PMCID: PMC8862426.

[25] Liao S, Sun H, Xu C. YTH Domain: A Family of N6-methyladenosine (m6A) Readers. Genomics Proteomics Bioinformatics. 2018;16(2):99-107. doi:10.1016/j.gpb.2018.04.002

[26] Xu Y, Zhang W, Shen F, et al. YTH Domain Proteins: A Family of m6A Readers in Cancer Progression. Front Oncol. 2021;11:629560. Published 2021 Feb 22. doi:10.3389/fonc.2021.629560

[27] Xu C, Liu K, Ahmed H, Loppnau P, Schapira M, Min J. Structural Basis for the Discriminative Recognition of N6-Methyladenosine RNA by the Human YT521-B Homology Domain Family of Proteins. J Biol Chem (2015) 290:24902–13. 10.1074/jbc.M115.680389

[28] Wang X, Lu Z, Gomez A, Hon GC, Yue Y, Han D, et al.. N6-methyladenosine-dependent regulation of messenger RNA stability. Nature (2014) 505:117–20. 10.1038/nature12730

[29] Chen M, Wei L, Law CT, Tsang FH, Shen J, Cheng CL, et al.. RNA N6-methyladenosine methyltransferase-like 3 promotes liver cancer progression through YTHDF2-dependent posttranscriptional silencing of SOCS2. Hepatology (2018) 67:2254–70. 10.1002/hep.29683

[30] Bailey AS, Batista PJ, Gold RS, Chen YG, de Rooij DG, Chang HY, et al.. The conserved RNA helicase YTHDC2 regulates the transition from proliferation to differentiation in the germline. Elife (2017) 6:e26116. 10.7554/eLife.26116

[31] Zhong L, Liao D, Zhang M, Zeng C, Li X, Zhang R, et al.. YTHDF2 suppresses cell proliferation and growth via destabilizing the EGFR mRNA in hepatocellular carcinoma. Cancer Lett (2019) 442:252–61. 10.1016/j.canlet.2018.11.006

[32] Shi Y, Fan S, Wu M, Zuo Z, Li X, Jiang L, et al.. YTHDF1 links hypoxia adaptation and non-small cell lung cancer progression. Nat Commun (2019) 10:4892. 10.1038/s41467-019-12801-6

[33] Li Z, Qian P, Shao W, Shi H, He XC, Gogol M, et al.. Suppression of m6A reader Ythdf2 promotes hematopoietic stem cell expansion. Cell Res (2018) 28:904–17. 10.1038/s41422-018-0072-0

[34] Anita R, Paramasivam A, Priyadharsini JV, Chitra S. The m6A readers YTHDF1 and YTHDF3 aberrations associated with metastasis and predict poor prognosis in breast cancer patients. Am J Cancer Res (2020) 10:2546–54.

[35] Li A, Chen YS, Ping XL, Yang X, Xiao W, Yang Y, et al.. Cytoplasmic m6A reader YTHDF3 promotes mRNA translation. Cell Res (2017) 27:444–7. 10.1038/cr.2017.10

[36] Xu C, Wang X, Liu K, Roundtree IA, Tempel W, Li Y, et al.. Structural basis for selective binding of m6A RNA by the YTHDC1 YTH domain. Nat Chem Biol (2014) 10:927–9. 10.1038/nchembio.1654

[37] Wojtas MN, Pandey RR, Mendel M, Homolka D, Sachidanandam R, Pillai RS. Regulation of m6A Transcripts by the 3'→5' RNA Helicase YTHDC2 Is Essential for a Successful Meiotic Program in the Mammalian Germline. Mol Cell (2017) 68:374–87.e12. 10.1016/j.molcel.2017.09.021

[38] Hsu PJ, Zhu Y, Ma H, Guo Y, Shi X, Liu Y, et al.. Ythdc2 is an N6-methyladenosine binding protein that regulates mammalian spermatogenesis. Cell Res (2017) 27:1115–27. 10.1038/cr.2017.99

[39] Kretschmer J, Rao H, Hackert P, Sloan KE, Höbartner C, Bohnsack MT. The m6A reader protein YTHDC2 interacts with the small ribosomal subunit and the 5'-3' exoribonuclease XRN1. RNA (2018) 24:1339–50. 10.1261/rna.064238.117

[40] Almasmoum HA, Airhihen B, Seedhouse C, Winkler GS. Frequent loss of BTG1 activity and impaired interactions with the Caf1 subunit of the Ccr4-Not deadenylase in non-Hodgkin lymphoma. Leuk Lymphoma. 2021 Feb;62(2):281-290. doi: 10.1080/10428194.2020.1827243. Epub

2020 Oct 6. PMID: 33021411.

[41] Deutscher MP. Degradation of RNA in bacteria: comparison of mRNA and stable RNA. Nucleic Acids Res. 2006 Feb 1;34(2):659-66. doi: 10.1093/nar/gkj472. PMID: 16452296; PMCID: PMC1360286.

[42] Bernstein J, Patterson DN, Wilson GM, Toth EA. Characterization of the essential activities of Saccharomyces cerevisiae Mtr4p, a 3'->5' helicase partner of the nuclear exosome. J Biol Chem. 2008 Feb 22;283(8):4930-42. doi: 10.1074/jbc.M706677200. Epub 2007 Dec 20. PMID: 18096702.

[43] Houseley J, LaCava J, Tollervey D. RNA-quality control by the exosome. Nat Rev Mol Cell Biol. 2006 Jul;7(7):529-39. doi: 10.1038/nrm1964. PMID: 16829983.

[44] Bühler M, Spies N, Bartel DP, Moazed D. TRAMP-mediated RNA surveillance prevents spurious entry of RNAs into the Schizosaccharomyces pombe siRNA pathway. Nat Struct Mol Biol. 2008 Oct;15(10):1015-23. doi: 10.1038/nsmb.1481. Epub 2008 Sep 7. PMID: 18776903; PMCID: PMC3240669.

[45] Slomovic S, Laufer D, Geiger D, Schuster G. Polyadenylation of ribosomal RNA in human cells. Nucleic Acids Res. 2006 May 31;34(10):2966-75. doi: 10.1093/nar/gkl357. PMID: 16738135; PMCID: PMC1474067.

[46] Patil DP, Pickering BF, Jaffrey SR. Reading m6A in the Transcriptome: m6A-Binding Proteins. Trends Cell Biol. 2018 Feb;28(2):113-127. doi: 10.1016/j.tcb.2017.10.001. Epub 2017 Nov 2. PMID: 29103884; PMCID: PMC5794650.

[47] Chen Y, Boland A, Kuzuoğlu-Öztürk D, et al. A DDX6-CNOT1 complex and W-binding pockets in CNOT9 reveal direct links between miRNA target recognition and silencing. Mol Cell. 2014;54(5):737-750. doi:10.1016/j.molcel.2014.03.034

[48] Sgromo A, Raisch T, Bawankar P, et al. A CAF40-binding motif facilitates recruitment of the CCR4-NOT complex to mRNAs targeted by Drosophila Roquin. Nat Commun. 2017;8:14307. Published 2017 Feb 6.

doi:10.1038/ncomms14307

[49] Garneau NL, Wilusz J, Wilusz CJ. The highways and byways of mRNA decay. Nat Rev Mol Cell Biol. 2007 Feb;8(2):113-26. doi: 10.1038/nrm2104. PMID: 17245413.

[50] Amrani, N., Sachs, M. S. & Jacobson, A. Early nonsense: mRNA decay solves a translational problem. Nature Rev. Mol. Cell Biol. 7, 415–425 (2006).

[51] Conti E, Izaurralde E. Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species. Curr Opin Cell Biol. 2005 Jun;17(3):316-25. doi: 10.1016/j.ceb.2005.04.005. PMID: 15901503.

[52] Anderson P, Kedersha N. RNA granules. J Cell Biol. 2006 Mar 13;172(6):803-8. doi: 10.1083/jcb.200512082. Epub 2006 Mar 6. PMID: 16520386; PMCID: PMC2063724.

[53] Houseley J, LaCava J, Tollervey D. RNA-quality control by the exosome. Nat Rev Mol Cell Biol. 2006 Jul;7(7):529-39. doi: 10.1038/nrm1964. PMID: 16829983.

[54] Martine A. Collart, Olesya O. Panasenko. The Ccr4–Not complex, Gene, Volume 492, Issue 1, 2012, Pages 42-53, ISSN 0378-1119, https://doi.org/10.1016/j.gene.2011.09.033.

[55] Du H, Zhao Y, He J, Zhang Y, Xi H, Liu M, Ma J, Wu L. YTHDF2 destabilizes m(6)A-containing RNA through direct recruitment of the CCR4-NOT deadenylase complex. Nat Commun. 2016 Aug 25;7:12626. doi: 10.1038/ncomms12626. PMID: 27558897; PMCID: PMC5007331.

[56] Shi R, Ying S, Li Y, Zhu L, Wang X, Jin H. Linking the YTH domain to cancer: the importance of YTH family proteins in epigenetics. Cell Death Dis. 2021 Apr 1;12(4):346. doi: 10.1038/s41419-021-03625-8. PMID: 33795663; PMCID: PMC8016981.

[57] Abramson J, Adler J, Dunger J, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature.

2024;630(8016):493-500. doi:10.1038/s41586-024-07487-w

[58] Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596(7873):583-589. doi:10.1038/s41586-021-03819-2

[59] Wee J, Wei GW. Evaluation of AlphaFold 3's Protein-Protein Complexes for Predicting Binding Free Energy Changes upon Mutation. J Chem Inf Model. 2024 Aug 26;64(16):6676-6683. doi: 10.1021/acs.jcim.4c00976. Epub 2024 Aug 8. PMID: 39116039; PMCID: PMC11351016.

[60] Zaccara S, Jaffrey SR. A Unified Model for the Function of YTHDF Proteins in Regulating m6A-Modified mRNA. Cell. 2020 Jun 25;181(7):1582-1595.e18. doi: 10.1016/j.cell.2020.05.012. Epub 2020 Jun 2. PMID: 32492408; PMCID: PMC7508256.

[61] Schäfer IB, Rode M, Bonneau F, Schüssler S, Conti E. The structure of the Pan2-Pan3 core complex reveals cross-talk between deadenylase and pseudokinase. Nat Struct Mol Biol. 2014 Jul;21(7):591-8. doi: 10.1038/nsmb.2834. Epub 2014 Jun 1. PMID: 24880344.

[62] Lobel JH, Ingolia NT. Defining the mechanisms and properties of post-transcriptional regulatory disordered regions by high-throughput functional profiling. bioRxiv [Preprint]. 2024 Feb 5:2024.02.01.578453. doi: 10.1101/2024.02.01.578453. PMID: 38370681; PMCID: PMC10871298.

[63] Kikhney AG, Svergun DI. A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. FEBS Lett. 2015 Sep 14;589(19 Pt A):2570-7. doi: 10.1016/j.febslet.2015.08.027. Epub 2015 José A. Brito, Margarida Archer,

[64] Chapter 10 - Structural biology techniques: X-ray crystallography, cryo-electron microscopy, and small-angle X-ray scattering,Editor(s): Robert R. Crichton, Ricardo O. Louro,Practical Approaches to Biological Inorganic Chemistry (Second Edition), Elsevier, 2020, Pages 375-416, ISBN 9780444642257. PMID: 26320411.

[65] O'Reilly FJ, Rappsilber J. Cross-linking mass spectrometry: methods and

applications in structural, molecular and systems biology. Nat Struct Mol Biol. 2018 Nov;25(11):1000-1008. doi: 10.1038/s41594-018-0147-0. Epub 2018 Oct 29. PMID: 30374081.

[66] Doidge R, Mittal S, Aslam A, Winkler GS. The anti-proliferative activity of BTG/TOB proteins is mediated via the Caf1a (CNOT7) and Caf1b (CNOT8) deadenylase subunits of the Ccr4-not complex. PLoS One. 2012;7(12):e51331. doi: 10.1371/journal.pone.0051331. Epub 2012 Dec 7. PMID: 23236473; PMCID: PMC3517456.

[67] Gagliardi M, Matarazzo MR. RIP: RNA Immunoprecipitation. Methods Mol Biol. 2016;1480:73-86. doi: 10.1007/978-1-4939-6380-5_7. PMID: 27659976.

[68] Papageorgiou AC, Poudel N, Mattsson J. Protein Structure Analysis and Validation with X-Ray Crystallography. Methods Mol Biol. 2021;2178:377-404. doi: 10.1007/978-1-0716-0775-6_25. PMID: 33128762.

[69] Dessau MA, Modis Y. Protein crystallization for X-ray crystallography. J Vis Exp. 2011 Jan 16;(47):2285. doi: 10.3791/2285. PMID: 21304455; PMCID: PMC3182643.

[70] Smyth MS, Martin JH. x ray crystallography. Mol Pathol. 2000 Feb;53(1):8-14. doi: 10.1136/mp.53.1.8. PMID: 10884915; PMCID: PMC1186895.

[71] Reier K, Lahtvee PJ, Liiv A, Remme J. A Conundrum of r-Protein Stability: Unbalanced Stoichiometry of r-Proteins during Stationary Phase in Escherichia coli. mBio. 2022 Oct 26;13(5):e0187322. doi: 10.1128/mbio.01873-22. Epub 2022 Aug 18. PMID: 35980033; PMCID: PMC9601097.

[72] Houseley J, Tollervey D. The many pathways of RNA degradation. Cell. 2009 Feb 20;136(4):763-76. doi: 10.1016/j.cell.2009.01.019. PMID: 19239894.

[73] Jamar NH, Kritsiligkou P, Grant CM. The non-stop decay mRNA surveillance pathway is required for oxidative stress tolerance. Nucleic

Acids Res. 2017 Jun 20;45(11):6881-6893. doi: 10.1093/nar/gkx306. PMID: 28472342; PMCID: PMC5499853.

[74] San Paolo S, Vanacova S, Schenk L, Scherrer T, Blank D, Keller W, Gerber AP. Distinct roles of non-canonical poly(A) polymerases in RNA metabolism. PLoS Genet. 2009 Jul;5(7):e1000555. doi: 10.1371/journal.pgen.1000555. Epub 2009 Jul 10. PMID: 19593367; PMCID: PMC2700272.

[75] Absmeier E, Chandrasekaran V, O'Reilly FJ, Stowell JAW, Rappsilber J, Passmore LA. Specific recognition and ubiquitination of translating ribosomes by mammalian CCR4-NOT. Nat Struct Mol Biol. 2023 Sep;30(9):1314-1322. doi: 10.1038/s41594-023-01075-8. Epub 2023 Aug 31. PMID: 37653243; PMCID: PMC7615087.

[76] Chalabi Hagkarim N, Grand RJ. The Regulatory Properties of the Ccr4–Not Complex. Cells. 2020; 9(11):2379. https://doi.org/10.3390/cells9112379

[77] Mittal S, Aslam A, Doidge R, Medica R, Winkler GS. The Ccr4a (CNOT6) and Ccr4b (CNOT6L) deadenylase subunits of the human Ccr4-Not complex contribute to the prevention of cell death and senescence. Mol Biol Cell. 2011 Mar 15;22(6):748-58. doi: 10.1091/mbc.E10-11-0898. Epub 2011 Jan 13. PMID: 21233283; PMCID: PMC3057700.

[78] Miller JE, Reese JC. Ccr4-Not complex: the control freak of eukaryotic cells. Crit Rev Biochem Mol Biol. 2012 Jul-Aug;47(4):315-33. doi: 10.3109/10409238.2012.667214. Epub 2012 Mar 15. PMID: 22416820; PMCID: PMC3376659.

[79] Vissers LELM, Kalvakuri S, de Boer E, et al. De Novo Variants in CNOT1, a Central Component of the CCR4-NOT Complex Involved in Gene Expression and RNA and Protein Stability, Cause Neurodevelopmental Delay. Am J Hum Genet. 2020;107(1):164-172. doi:10.1016/j.ajhg.2020.05.017

[80] Shirai YT, Suzuki T, Morita M, Takahashi A, Yamamoto T. Multifunctional roles of the mammalian CCR4-NOT complex in physiological phenomena. Front Genet. 2014 Aug 21;5:286. doi: 10.3389/fgene.2014.00286. PMID:

25191340; PMCID: PMC4139912.

[81] Jung JH, Lee D, Ko HM, Jang HJ. Inhibition of CNOT2 Induces Apoptosis via MID1IP1 in Colorectal Cancer Cells by Activating p53. Biomolecules. 2021;11(10):1492. Published 2021 Oct 10. doi:10.3390/biom11101492

[82] De Keersmaecker K, Atak ZK, Li N, et al. Exome sequencing identifies mutation in CNOT3 and ribosomal genes RPL5 and RPL10 in T-cell acute lymphoblastic leukemia. Nat Genet. 2013;45(2):186-190. doi:10.1038/ng.2508

[83] Matsuo Y, Inada T. Co-Translational Quality Control Induced by Translational Arrest. Biomolecules. 2023 Feb 7;13(2):317. doi: 10.3390/biom13020317. PMID: 36830686; PMCID: PMC9953336.

[84] Mammalian CCR4–NOT binds and ubiquitinates 80S ribosomes to enforce translational stalling. Nat Struct Mol Biol 30, 1254–1255 (2023). https://doi.org/10.1038/s41594-023-01079-4

[85] Buschauer R, Matsuo Y, Sugiyama T, et al. The Ccr4-Not complex monitors the translating ribosome for codon optimality. Science. 2020;368(6488):eaay6912. doi:10.1126/science.aay6912

[86] Collart MA. The Ccr4-Not complex is a key regulator of eukaryotic gene expression. Wiley Interdiscip Rev RNA. 2016;7(4):438-454. doi:10.1002/wrna.1332

[87] Du, H., Zhao, Y., He, J. et al. YTHDF2 destabilizes m6A-containing RNA through direct recruitment of the CCR4–NOT deadenylase complex. Nat Commun 7, 12626 (2016). https://doi.org/10.1038/ncomms12626

[88] Xu K, Bai Y, Zhang A, Zhang Q, Bartlam MG. Insights into the structure and architecture of the CCR4-NOT complex. Front Genet. 2014;5:137. Published 2014 May 16. doi:10.3389/fgene.2014.00137

[89] Doidge R, Mittal S, Aslam A, Winkler GS (2012) The Anti-Proliferative Activity of BTG/TOB Proteins Is Mediated via the Caf1a (CNOT7) and Caf1b (CNOT8) Deadenylase Subunits of the Ccr4-Not Complex. PLoS ONE 7(12): e51331. https://doi.org/10.1371/journal.pone.0051331

[90] Mauxion et al., 2008, "BTG/TOB factors bind to CAF1 and thereby control the deadenylation of mRNAs"

[91] Structural and Biochemical Insights to the Role of the Ccr4-not Complex and Ddx6 ATPase in Microrna Repression. Mathys, H., Basquin, J., Ozgur, S., Czarnocki-Cieciura, M., Bonneau, F., Aartse, A., Dziembowski, A., Nowotny, M., Conti, E., Filipowicz, W. (2014) Mol Cell 54: 751

[92] Xu, P., Hu, K., Zhang, P. et al. Hypoxia-mediated YTHDF2 overexpression promotes lung squamous cell carcinoma progression by activation of the mTOR/AKT axis. Cancer Cell Int 22, 13 (2022). https://doi.org/10.1186/s12935-021-02368-y