Structural Approaches to Targeting RNA Terminal Uridylytransferases



Gangshun Yi

Nuffield Department of Clinical Medicine

The Division of Structural Biology

Balliol College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Michaelmas 2022/2023

ABSTRACT

Let-7 microRNA (miRNA) plays a crucial role in the control of cell differentiation and proliferation, which also makes it a suppressor for the occurrence and development of tumors. Pre-let-7 is one of the intermediates in the maturation process of let-7 and two uridylyltransferases - ZCCHC6/TUT7, and ZCCHC11/TUT4 - play essential roles in the downstream processing of let-7 pre-miRNA. In the absence of RNA binding protein Lin28A, they perform 3' mono-uridylation, leading to pre-let-7 maturation. In the presence of Lin28A, they together form a TUT4(7)-Lin28A-pre-let-7 holo-complex and perform oligo-uridylation, which subsequently leads to exonucleolytic pre-let-7 degradation. However, with no full structural information about TUT4/7, especially the lack of a complex structure with pre-let-7, their detailed uridylation mechanism is still a mystery.

X-ray crystallography and cryo-EM are the mainstream for structure determination and have been developed tremendously in the past decades. This project combined both techniques for structural study of TUT4/7 in apo or complex forms with pre-let-7 in the presence or absence of Lin28A, in order to unravel the details of the uridylation mechanism at atomic level.

With regards to protein purification, I have established and optimized high throughput (HTP) cloning and expression methods for TUT4/7 based on the *E.coli* system. An optimized two-step purification protocol allows for the purification of both TUT4 and TUT7 in full length or truncated forms. Assays confirm both full length TUT4 and TUT7 are catalytically active. These highly purified proteins are essential milestones on the way to our goal of structural characterization of TUT4 (7) by X-ray crystallography and cryo-EM.

In the case of TUT7, I have determined the crystal structure of the TUT7 catalytic module (CM) at 2.6 Å, which can serve as a basis for fragment screening trials aiming for identifying small molecules with oncogene inhibition potential. With cryo-EM, I reconstructed four high-resolution structures. The dynamic TUT7 full length apo structure is resolved at 4 Å, showing an overall "C-shape" model with two nucleotidyltransferase domains (NTDs) connected by flexible linkers. The 3.7Å binary complex of TUT7 with pre-let-7 reveals the mono-uridylation

state, in which the pre-let-7 stem locks onto the CM, stabilizing the dynamic active-site cleft, while the pre-let-7 PRE loop region is close to the let-7 interacting module (LIM) but without a direct interaction or the interaction is transitory/dynamic. The zinc finger regions of the LIM are missing in this complex. Finally, two ternary complex structures were obtained from the TUT7/pre-let-7/Lin28A complex, determined at 3.8 Å and 3.6 Å respectively. The zinc finger in the LIM is visualized in both structures, while in conformation-I, the CM is missing, presumably due to the flexible linker regions connecting the LIM and CM, representing a pre-oligo-uridylation state. The other ternary complex in conformation-II reveals the additional densities of CM and a novel structure of zinc knuckle-1, of which the latter binds to the RNA stem, showing a second activated pre-oligo-uridylation state. The four high-resolution structures presented allow a full description of conformational changes from mono- to oligo-uridylation.

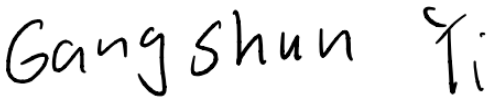
In the study of TUT4, the TUT4 CM crystalised and diffracted to 4.5 Å, indicating the possibility of obtaining well-diffracting crystals for subsequent fragment screening trials (Li, 2020); while the TUT4 LIM in complex with inhibitory nanobody S2A4 diffracted to 5 Å and is under optimization currently. Using cryo-EM, a 3.7 Å ternary complex of a truncated TUT4 (TUT4_miniF)/pre-let-7/Lin28A was determined, showing a similar assembly to one of the models of TUT7/pre-let-7/Lin28A (conformation-I). Furthermore, the cold shock domain of Lin28A was identified in another TUT4_miniF/pre-let-7/Lin28A model solved at 12 Å. The TUT4 full length protein was purified successfully recently and preliminary cryo-EM trials together with pre-let-7/Lin28A show a different shape compared to TUT4_miniF/pre-let-7/Lin28A and close to structure of the TUT7 ternary complex in conformation-II.

The structural work undertaken in this project establishes a strong basis for further understanding the uridylation mechanism at atomic level. This detailed information will provide important details for the rational design and improvement of TUT inhibitors with the potential to improve prospects for cancer patients suffering from a range of malignancies, including some of unmet need.

DECLARATION OF WORK

I hereby declare that this thesis is my own work and effort except when done in collaboration with others, and that it has not been submitted anywhere for any award. The findings from collaborative work described here are clearly stated prior to each chapter. Where other sources of information have been used, they have been acknowledged.

Signature



Gangshun Yi (printed)

ACKNOWLEDGEMENT

Time flies! It has been four years since I came to the University of Oxford and joined the group of Prof. Robert Gilbert in October 2018.

Thereby, I would like to firstly thank my parents. My biggest motivation in scientific research is the support from my parents. They give me life and also unlimited love. It is incredible that the price of growth is leaving my parents further and further from my hometown to Shanghai, China and to Oxford, the UK now. Due to the COV-19 pandemic, I missed lots of either happy or disappointed moments with them in my past three years, the last ones I would like to let slip. Hopefully, I can share more valuable moments with them in our future life.

Then, I express my great thanks to my supervisor Robert Gilbert. I still remembered the first call with him before I came to Oxford. To some extent, it was awkward due to my poor English, but I felt his patience and sincere just like the experience working with him in my past four years. Meanwhile, his enthusiasm in science and high-level guidance benefit me throughout my DPhil studies.

STRUBI is a happy and big family. I would like to thank Margaret, Weixian, Karl and Tom for their selfless lab support. I want to give thanks to Loic, Helen and Yanan for their kind support for my EM training and patient guidance on my data processing. I also appreciate Prof. Peijun and Prof. John for the ebic session support and other colleagues in STRUBI. Besides, I want to thank Mingda for his great help for both my project and life over my last three years.

Finally, I would like to thank all the ones concerning me and the ones I miss so much. It is hard to image the life will become dull without these friends I have meet over my last 28 years. I cannot thank you all enough.

Thank you all, for your support; Thank myself, for my persistence.

LIST OF ABBREVIATIONS

ADP, Adenosine diphosphate

ATP, Adenosine triphosphate

AFM, Atomic Force Microscopy

CM, Catalytic Module

dCTP, Deoxycytidine triphosphate

dGTP, Deoxyguanosine triphosphate

dUpNHpp, 2'-Deoxyuridine-5'-[(α,β)-imido]triphosphate, Sodium salt

DTT, Dithiothreitol

EM, Electron Microscopy

FSC, Fourier Shell Correlation

GFP, Green Fluorescent Protein

GraFix, Gradient Fixation

GST, Glutathione S-transferase

GSFSC, Gold Standard Fourier Shell Correlation

HTP, High Throughput

IMAC, Immobilized Metal Affinity Chromatography

LIC, Ligation Independent Cloning

LIM, Lin28 Interaction Module

kDa, Kilo Dalton

Ni-NTA, Nickel-Nitriloacetic acid

PCR, Polymerase Chain Reaction

PDB, Protein Data Bank

PEG, Polyethylene glycol

PrDOS, Protein DisOrder Prediction System

pre-let-7, precursor-let-7

SDS-PAGE, Sodium dodecyl-sulfate polyacrylamide gel electrophoresis

TECP, Tris(2-carboxyethyl) phosphine

TEM, Transmission Electron Microscopy

TEV, Tobacco Etch Virus

TUT, Terminal Uridylyltransferase

UTP, Uridine-5′-triphosphate

ZCCHC, the zinc finger CCHC-type

3' CY, 3' end Cyanine

Contents

ABSTRACT 2

DECLARATION OF WORK 4

ACKNOWLEDGEMENT 5

LIST OF ABBREVIATIONS 6

CONTENTS 8

CHAPTER 1 INTRODUCTION 11

1.1 POST-TRANSCRIPTIONAL GENE REGULATION: A BRIEF REVIEW 12

1.1.1 RNA Modification and the example of N6-methyladenosine *12*

1.1.2 Polyadenylation *15*

1.1.3 Uridylation *17*

1.2 LET-7 FAMILY: A KEY SUPPRESSOR OF ONCOGENES 20

1.2.1 Discovery of let-7 microRNAs and biogenesis *20*

1.2.2 Functions *22*

1.3. LIN28: REGULATOR OF THE SELF-RENEWAL OF STEM CELLS 27

1.3.1 Discovery and function *27*

1.3.3 Lin28A and SUMOylation *31*

1.3.4 Lin28A and cancers *32*

1.4. TUTASES : A CLASS OF NON-CANONICAL POLY(A) POLYMERASES 33

1.4.1 Discovery *33*

1.4.2 TUTases structures *34*

1.4.3 Functions *36*

1.5. X-RAY CRYSTALOGRPHY AND CRYO-EM IN STRUCTURAL BIOLOGY 42

1.5.1 X-ray crystallography *42*

1.5.2. Cryo-EM *46*

1.6. ISSUES AND AIMS 48

1.6.1 Issues *48*

1.6.2 Aims *49*

1.7. OVERVIEW OF THIS THESIS 51

CHAPTER 2 METHODS 53

2.1. HIGH THROUGHPUT (HTP) CLONING AND EXPRESSION SCREENING 54

2.1.1 Concept *54*

2.1.2 KRX stain *54*

2.1.3 Pipeline *54*

2.2 POLYMERASE CHAIN REACTION (PCR) 58

2.2.1 PCR *58*

2.2.2 Agarose gel electrophoresis *59*

2.2.3 PCR product recovery steps *59*

2.3 PLASMID EXTRACTION AND TRANSLATION 61

2.3.1 Plasmid extraction *61*

2.3.2 KRX Transfection ... *61*
2.3.3 Small scale expression screening ... *62*
2.3.4 KRX competent cells preparation ... *62*
2.4 PROTEIN PURIFICATION ... 64
2.4.1 Cell culture ... *64*
2.4.2 Ion exchange chromatography ... *64*
2.4.3 Size exclusion chromatography ... *65*
2.5. CRYSTALLOGRAPHY ... 66
2.5.1 Principles of crystallography ... *66*
2.5.2 Coarse screening ... *68*
2.5.3 Condition optimization ... *68*
2.5.4 Crystal diffraction and data collection ... *68*
2.5.5 Data processing and model building ... *69*
2.6 NEGATIVE STAINING ... 70
2.6.1 Grid preparation and data collection ... *70*
2.6.2 Data processing ... *70*
2.7. CRYO-EM ... 71
2.7.1 Sample grids preparation and data collection ... *71*
2.7.2 Data processing ... *71*
2.8. SITE-DIRECTED MUTAGENESIS ... 72
2.9. URIDYLATION ASSAY ... 73
CHAPTER 3 CHARACTERIZATION OF TUT7 ... 74
3.1. INTRODUCTION ... 75
3.2. METHODS AND RESULTS ... 77
3.2.1. Design of TUT7 in both full length and truncated forms ... *77*
3.2.2 Purification of TUT7_CMs ... *79*
3.2.3 Crystallography of TUT7_CM ... *82*
3.2.4 Purification of TUT7_LIM ... *86*
3.2.5 Purification of Lin28 ... *87*
3.2.6 Purification of TUT7 ... *91*
3.2.7 Pre-let-7 RNA preparation and assay ... *98*
3.2.8 Assembly of TUT7 with pre-let-7 in the presence or absence of Lin28A ... *100*
3.2.9 PEGylation modification of TUT7 ... *103*
3.2.10 Negative staining of TUT7 ... *104*
3.2.11 Cryo-EM imaging of TUT7 on Glacios microscope ... *105*
3.2.12 Cryo-EM imaging of TUT7 on Krios microscope ... *107*
CHAPTER 4 CHARACTERIZATION OF TUT4 ... 117
4.1. INTRODUCTION ... 118
4.2 METHODS AND RESULTS ... 120
4.2.1 Purification and crystallography of TUT4_CM ... *120*
4.2.2 Purification of TUT4 LIM ... *125*
4.2.3 Purification of S2A4 and CB11 ... *128*
4.2.4 Crystallography of TUT4LIM in complex with S2A4 ... *129*

4.2.5 Purification of TUT4_miniF construct *131*
4.2.6 Purification of TUT4_F *134*
4.2.7 Assays of TUT4 ..... *136*
4.2.8 Assembly of TUT4/pre-pre-let 7/Lin28A *137*
4.2.9 Assembly of TUT4 with Nanobodies: S2A4 ..... *140*
4.2.10 Negative staining of TUT4 ..... *141*
4.2.11 Grafix for assembling the ternary complex *142*
4.2.12 Cryo-EM structures of TUT4 from Krios microscope *144*
4.3 CONCLUSION 149

CHAPTER 5 STRUCTURAL ANALYSIS OF TUTASES 150

5.1 INTRODUCTION 151
5.2. RESULTS 153
5.2.1. CRYO-EM STRUCTURE DETERMINATION OF APO TUT7 AND TUT7/4 COMPLEXES 153
5.2.2. LOOSE ASSOCIATION BETWEEN NUCLEOTIDYLTRANSFERASE DOMAINS CHARACTERIZES THE TUT7 APO STATE 157
5.2.3. TUT7 ALONE BINDS PRE-LET-7G MIRNA VIA END-CAPTURE OF THE TARGET 3' END 158
5.2.4. LIN28A MODIFIES TARGET PRE-MIRNA RECOGNITION BY PRIOR ASSOCIATION 163
5.2.5. A MECHANISM FOR SWITCHING BETWEEN MONO- AND OLIGO-URIDYLATION IN TUT4/7 178
5.3. DISCUSSION 182

CHAPTER 6 CONCLUSION 184

6.1 HTP METHOD FOR TUTASES STUDY 185
6.4 STRUCTURAL ANALYSIS OF TUT4 187

CHAPTER 7 BIBLIOGRAPHY 190

Chapter 1 Introduction

1.1 Post-transcriptional gene regulation: A brief review

1.1.1 RNA Modification and the example of N6-methyladenosine

Ribonucleic acid (RNA) is a biological macromolecule, and serves as a bridge transferring genetic information from DNA to proteins, in a process known as the central dogma (Mansourian, 2011). Structurally, RNA consists of three components, a ribose sugar, of which the carbons are labeled 1' to 5', a phosphate group, linking 3' to the sugar, and a base (adenine (A), uracil (U), guanine (G) or cytosine (C)), which attaches to the 1' carbon of the sugar (Moore & Steitz, 2003; Nissen et al., 2000; Nombela et al., 2021; Nomura, 1987). The hydrogen bonds formed by the bases, between adenine and uracil or guanine and cytosine, can lead to formation of double stranded helical structures similar to A-form DNA, and by folding of single and multiple strands. It creates diverse structural RNAs playing critical roles in various cellular processes (Nomura, 1987).

As with biomacromolecules like DNA and proteins, RNAs can be modified by forming covalent bonds attached to functional groups (Drygin Iu et al., 1988; Drygin Iu et al., 1979). Recently emerging evidence has shown RNA modification as a critical post-transcriptional regulator of gene expression and cell fate (Barbieri & Kouzarides, 2020; Wood et al., 2021; Xue et al., 2022). This process is a highly specific and efficient way to regulate the functions of biological macromolecules. There are over 100 RNA post-transcriptional modifications on various types of RNAs (Li & Mason, 2014). The ribose sugar and nucleobases in RNAs can both be modified at different positions (Dhanoa & Housner, 2007; Faria & Ulrich, 2008). All known RNA species are modified, with ribosomal RNA (rRNA) and transfer RNA (tRNA) being most affected (Barbieri & Kouzarides, 2020). Such changes can generally be divided into terminal modifications, like 5' cap and 3' poly(A) tail, and internal modifications, such as 6-Methylguanosine (m6G), 5-methylcytosine (m5C) and pseudouridylation (Behm-Ansmant et al., 2004; Fabian & White, 2004; Jiang et al., 2007; Tomikawa, 2018; Xin et al., 2022; Zhang et al., 2022). In recent years, the finding of RNA modifications on non-coding RNAs, such as microRNAs or lncRNAs, has shown their direct functional effects on gene expression and cell fate (Norbury, 2013).

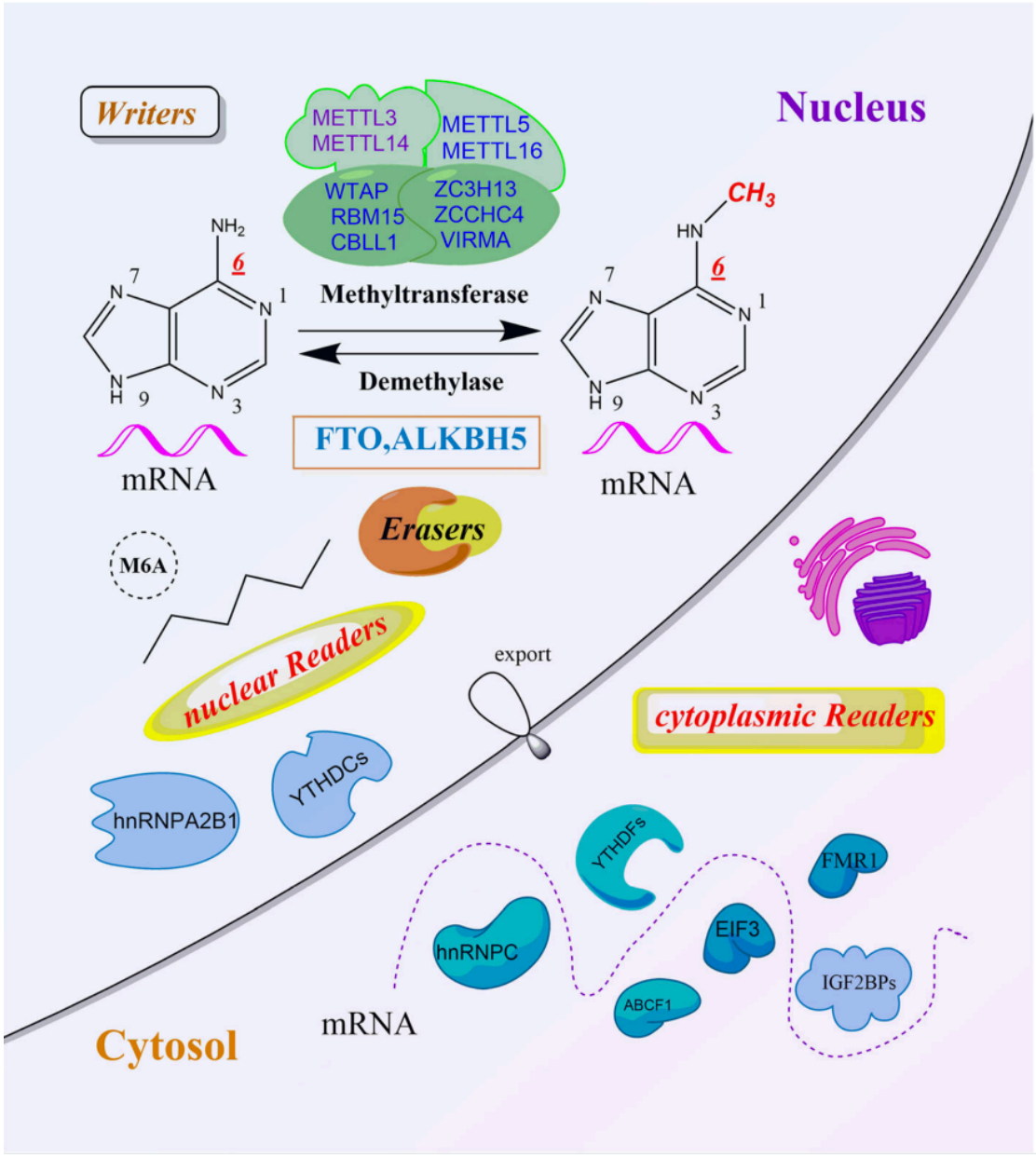


Figure 1.1 Schematic diagram of m6A RNA methylation modification process and components involved in this pathway, including writers, readers, erasers and base pairing. In writer processes, m6A METTL family members catalyze and transfer methyl groups to the 6' of adenosines. Methyladenines would be recognized by responsible readers via the YTHDF domain and YTHDF domain-containing proteins, and demethylase enzymes subsequently remove methyl at the 6' of adenosine. This figure is adapted from (Zhou et al., 2020).

The best characterized and earliest RNA modification is that occurring at position 6 of adenosine termed N6-methyladenosine (m6A) (Barbieri & Kouzarides, 2020).

N6-methyladenosine (m6A) RNA methylation is a universal modification in mRNAs, long intergenic non-coding RNA, primary-microRNA and rRNA (Alarcon et al., 2015). The whole process of m6A modification comprises writers, readers, erasers and base pairing (Luo et al., 2022). As shown in Figure 1.1, for writers, they consist of m6A METTL family members: METTL3, METTL5, METTL14, METTL16, cofactors Zinc-finger CCHC-type containing protein 13 (ZC3H13), cofactors Wilms tumor 1 associated protein (WTAP), Vir-like m6A methyltransferase associated (VIRMA; also known as KIAA1429), RNA-binding motif protein 15 (RBM15/15B) and Cbl proto-oncogene-like 1 (CBLL1; also known as HAKAI) (Zhou et al., 2020). Altogether, those enzymes catalyze and transfer methyl groups to the 6' of adenosines. The YTHDF domain and YTHDF domain-containing proteins are responsible as readers to recognize the methyl at the 6' of adenosine and subsequently demethylase enzymes: Fat mass and obesity-associated protein (FTO) and AlkB homolog 5 (ALKBH5) erase methylation from RNAs (Kim & Siddiqui, 2021). This dynamic modification on adenosines influences RNA transcription, maturation, localization and metabolism (Cai et al., 2022; Liu et al., 2020; Madugalle et al., 2020). As a vital RNA epigenetic modification, emerging evidence shows that m6A on mRNAs contributes to the self-renewal of cancer stem cells, promoting cancer proliferation and resistance to chemotherapy (chDai et al., 2018). Further studies will focus on revealing the detailed mechanism of the close link of m6A to cancers, the relationship between m6A and DNA in the regulation of gene expression, and the connections among different types of RNA modifications in the near future (Qiao et al., 2022; Zhou et al., 2020).

1.1.2 Polyadenylation

Polyadenylation is a process of adding a poly(A) tail to RNA transcripts and another example of mRNA modification (Proudfoot et al., 2002). Structurally, poly(A) tails form a single-stranded A-form-like helix, in which all adenosines are stacked on top of each other (Saenger et al., 1975; Tang et al., 2019). In the 1970s, pioneering work proved that the poly(A) tails on mRNAs play a significant role in efficient translation, with the finding that after fertilization, sea urchin eggs would have tens of cytoplasmic transcripts being polyadenylated within an extremely short time, while maternal mRNAs which have a short poly(A) chain were found to be translationally quiescent (Dworkin & Dworkin-Rastl, 1985; Slater et al., 1972). The following finding was that the polyadenylation modification on mRNAs is to control their fate and as an important way to regulate the translation processes in the cytoplasm (Richter, 1999). The Poly(A) tail modification on mRNAs also contributes to the stability of mRNAs, and acts as a master regulator of gene expression in the cytoplasm (Barnhart et al., 2013; Eads & Hand, 2003).

Polyadenylation is a complicated event, as shown in Figure 1.2. Initially, in the cell nucleus, RNA polymerase II synthesizes the transcripts of pre-mRNAs based on the DNA template. In subsequent processing, the pre-mRNAs would be modified including, 5' capping, splicing, 3' cleavage and polyadenylation (Proudfoot et al., 2002). During polyadenylation, Nuclear poly(A)-binding N protein (PABPN) controls the addition of poly(A) chain and then polyadenylated mRNAs are exported to the cytoplasm. In the cytoplasm, they initially bind to the 40S ribosome subunit, and subsequently the complete 80S ribosome will conduct translation, with the assistance of cytoplasmic poly(A)-binding protein C (PABPC) closely attached to the poly(A) tail (Mangus et al., 2003). However, removal or shortening the length of the poly(A) tail leads to PABPC mRNA degradation by exonucleases, such as CCR-NOT and PAN2-PAN3 complexes (Jonas et al., 2014; Miller & Reese, 2012; Wolf & Passmore, 2014). Thus, PABPC is considered as a key factor for mRNA stability. Emerging evidence in vivo and vitro indicates that the PABPC located at the 3' end of mRNAs prevents exonucleases accessing mRNAs, which therefore have their half-lives extended (Swartwout & Kinniburgh,

1989). Besides, polyadenylation modification is also found regulating the biogenesis of other types of RNAs, like tRNAs, rRNAs, snRNAs and snoRNAs (Reinisch & Wolin, 2007).

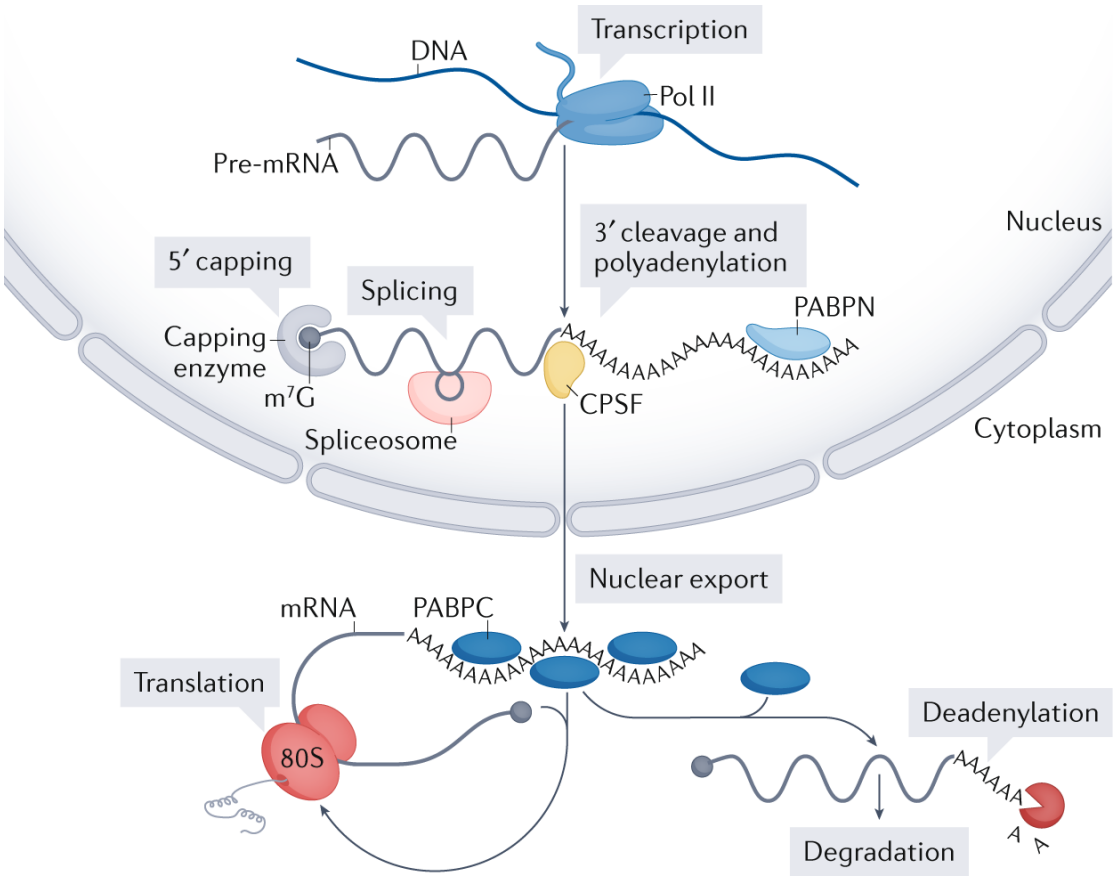


Figure 1.2 mRNA polyadenylation and mRNA poly(A) tails regulating gene expression from pre-mRNA transcription in the nucleus to degradation in the cytoplasm. In the cell nucleus, the transcripts of pre-mRNAs are synthesized by RNA polymerase II based on the DNA template, and would be modified by processes such as 5' capping, splicing, 3' cleavage and polyadenylation, in which the latter is controlled by Nuclear poly(A)-binding N protein (PABPN) for the addition of poly(A) chain and then polyadenylated mRNAs are exported to the cytoplasm. In the cytoplasm, the polyadenylated mRNAs are continually processed by 40S/80S ribosomal particles, PABPC and degraded by exonucleases, like CCR-NOT and PAN2-PAN3 complexes. This figure is taken from (Passmore & Coller, 2022).

1.1.3 Uridylation

Uridylation is similar to polyadenylation, but in a discovery about RNA modification made in rather recent years, involves addition of non-templated uridines to the 3' end of RNAs (Yu & Kim, 2020). However uridylation is different from polyadenylation, which is mainly catalyzed by canonical poly(A) polymerases (PAPs) to create a poly(A) tail, in that non-canonical polymerases uridylate RNAs by adding mono- or oligo-uridines to their 3' ends (Stewart, 2019; Weill et al., 2012). However, both poly(A) and U tails are recognized as central aspects in the regulation of gene expression and cell fate, and thus have direct influence on cancers (Posner & Laubenbacher, 2020; Yates et al., 2013). The enzymes involved are members of the terminal nucleotidyltransferases (TENTs) family of enzymes (Yu & Kim, 2020). Over the last two decades, 11 different members have been found in mammals and classified based on their substrate specificity and phylogenetic analysis (Warkocki, Liudkovska, et al., 2018; Yashiro & Tomita, 2018). As presented in Figure 1.3, structurally, all TENTs have two key types of domains, the Nucleotidyltransferase domain (NTD1) and Poly(A) polymerase-associated domain (NTD2) (Faehnle et al., 2017a). NTD2 has three conserved aspartates, which are essential for catalytic activity. They bind a magnesium ion and assist the incorporation of nucleotides at the 3' end of RNAs (Martin et al., 2000). As shown in Figure 1.4, the crystal structure of TUT7_CM in complex with RNA and UTP reveals that the two aspartates on the left as viewed (Asp 1058 and Asp 1060) bind to the Mg^{2+} ion to stabilize the β- and γ-phosphate of the uridine triphosphate (UTP) moiety, and then the third aspartate (Asp 1119) attracts the RNA 3' -hydroxyl group, allowing nucleophilic attack to the UTP α-phosphate. Subsequently, the 3' - hydroxyl on the nucleotide residue in the -1 position of the RNA substrate is in close proximity to the α-phosphate of the NTP, which nucleophilically attacks the α-phosphate, and finally the phosphodiester bond will form, releasing diphosphate (Cheng et al., 2017; Lunde et al., 2012; Munoz-Tello et al., 2012). Other domains and motifs, like KA-1—kinase associated domain (in TENT1), ZCCHC-type zinc knuckle domain, RNA recognition motifs, etc. are found in diverse TENTs, and exert various biological functions, such as substrate specificity.

Among all TENTs, Cid1 was the first identified as a non-canonical PAP in yeast and

subsequently revealed as a terminal uridylytransferase (Bai et al., 2011; Rissland et al., 2007; Rissland & Norbury, 2009). Non-canonical PAPs are conserved across different organisms, and their functions are quite diverse on different types of RNAs. A well-characterized mechanism involving 3' uridylation is the regulation of pre-let-7 miRNA (Yashiro & Tomita, 2018). A single uridine addition to group II pre-let-7 miRNAs promotes their processing by Dicer to mature let-7, while oligo-uridine addition results in their degradation by exonuclease Dis3L2 (Chang et al., 2013; Faehnle et al., 2014; Yates et al., 2013). TUT4, TUT7 uridylytransferases and Lin28A RNA binding proteins are critical enzymes and cofactors in this process.

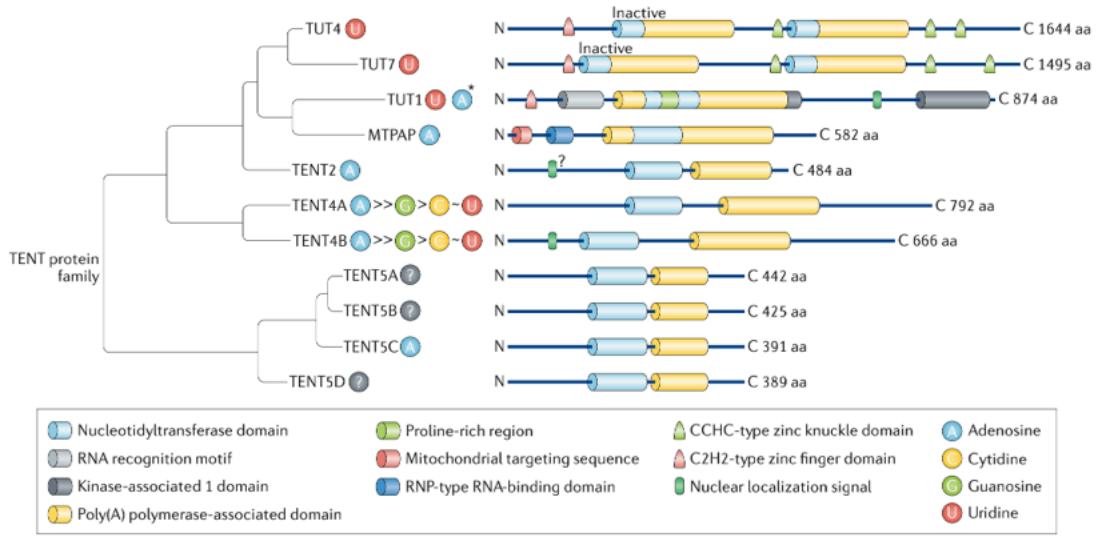


Figure 1.3 RNA terminal nucleotidetransferases in mammals. A phylogenetic tree analysis based on the specificity of their substrates and amino acid sequence alignment is presented on the left. A brief schematic diagram shows the domain architectures of each member of TENTs is shown on the right. The nucleotidytransferase domain and Poly(A) polymerase-associated domain are conserved in all TENT family members, while TUT1 is unique as a result of the additional presence of an RNA recognition motif and Kinase-associated 1 domain. This figure is adopted from (Yu & Kim, 2020).

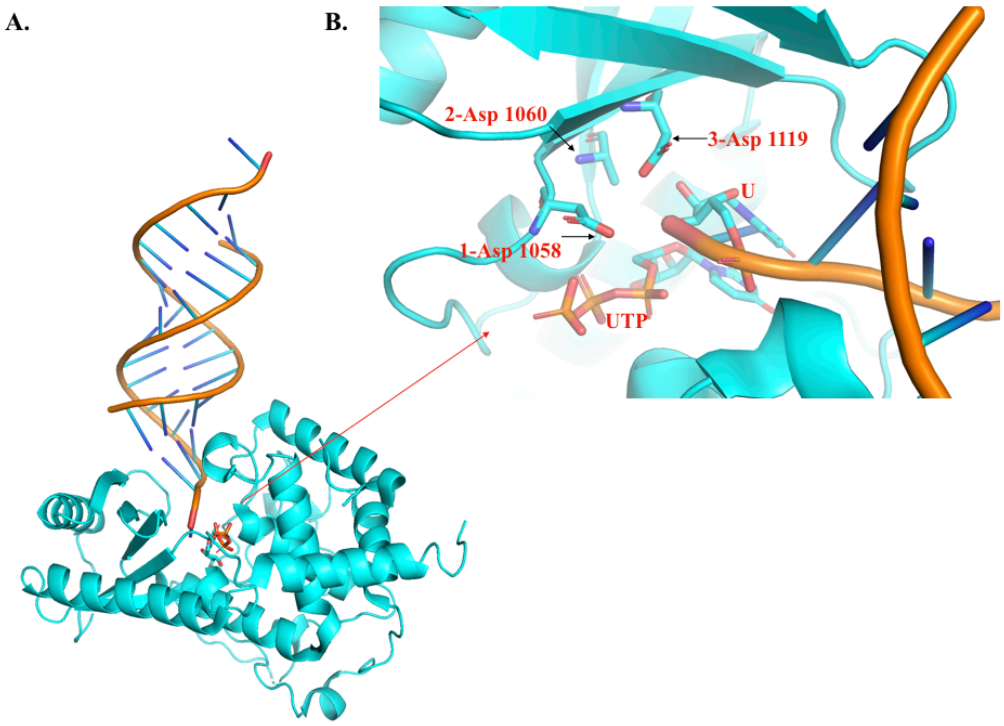


Figure 1.4 Catalytic site of the TUTases, shown for the human TUT7_CM (963-1365) and drawn from PDB:5W0O. A. Overview of crystal structure of human TUT7_CM (cyan) in complex with UTP and dsRNA (orange yellow ribbon). B. Close-up view of catalytic site of TUT7_CM, presenting three aspartates in positions 1058, 1060 and 1119 (Faehnle et al., 2017).

1.2 Let-7 family: A key suppressor of oncogenes

1.2.1 Discovery of let-7 microRNAs and biogenesis

MicroRNAs are a class of small non-coding RNAs, which have a length of 19-25 nucleotides (nts) (Hagan & Croce, 2007; Nagandla et al., 2021). Their primary function is acting as negative regulators of their targeted mRNAs (Davis-Dusenbery & Hata, 2010; Lin & Gregory, 2015). So far, there are thousands of miRNAs identified in diverse organisms regulating around 30% of total gene expression in mammals (Lewis et al., 2005). They are involved in cell differentiation, metabolism, proliferation and immune response processes, maintaining cellular homeostasis, and are often considered to link to many disease occurrences especially cancers due to their abnormal expression and dysfunction, e.g. let-7 microRNAs (miRNAs) (Balzeau et al., 2017; Bartel, 2004; Chen et al., 2019).

Let-7 miRNAs comprise a set of highly conserved members and were one of the first two miRNA types to be discovered together with Lin41 (Reinhart et al., 2000). The let-7 family was originally identified in the nematode, *C. elegans* as a regulator of developmental timing (Rougvie, 2001). Let-7 family miRNAs are evolutionarily conserved across various eukaryotic species. In *Drosophila*, there is only one let-7 microRNA, which has the same sequence to the one in *C. elegans*, but there are 12 classifications located in vertebrates, encoding 9 mature let- 7 miRNAs as shown in Figure 1.5 (Ma et al., 2021; Roush & Slack, 2008; Ruby et al., 2006). All let-7 microRNAs undergo biogenesis from primary transcripts and hairpin precursors to mature let-7 in a post-transcriptional manner (Triboulet et al., 2015). The key transcripts are precursor RNAs (pre-let 7), which are regulated by uridylytransfeases. Pre-let-7 miRNAs fall into two groups – group I with a 2-nt uridines overhang at the 3' end of pre-miRNA stem, while group II contain a 1-nt uridine overhang. Group II pre-let-7 miRNAs require the addition of a non-templated uridine to their 3' end to promote their being processed by Dicer to mature let-7 (Heo et al., 2012). There is strong evidence that this type of post-transcriptional modification has significant impact on gene expression and cell fate, either positively or negatively (Kim et al., 2015). As depicted in Figure 1.6, the synthesized RNA transcripts are initially cleaved by the Drosha/DGCR8 complex in the nucleus (Duman et al., 2022; Lin et al., 2020). The pre-let-7 hairpin product consists of 50 to 75 nucleotides (Gregory

et al., 2004) and is transported to the cytoplasm by the exportin-5/Ran-GTP complex. In the cytoplasm, group I pre-let-7 pre-miRNAs with 2-nt uridine overhang is directly recognized and specifically processed by Dicer, while the group II pre-let-7 with 1-nt overhang (90% of all pre-let-7 forms) requires the addition of uridine to the 3' end for Dicer to process efficiently (Heo et al., 2012). This uridylation reaction is mainly catalyzed by two transferases ZCCHC11 (TUT4) and ZCCHC6 (TUT7) with close mediation by RNA binding protein Lin28A. In the absence of Lin28A, mono-uridylation occurs and produces pre-miRNAs with a 2-nucleotide 3' overhang, which is ultimately recognized and processed by the Dicer/TRBP complex. This leads to the maturation of let-7 miRNA. Subsequently, the let-7 interacts with Ago2, base-pairs with its targeted mRNAs, and forms the RNA-induced silencing complex (RISC) to promote mRNA degradation and/or translational inhibition. In the presence of Lin28A, it binds to the terminal loop of pre-let-7, and then recruits the TUT4/7 to add a short polyU-tail to pre-let-7, which thus blocks the Dicer cleavage. An exonuclease Dis3l2, responsible for poly-U RNA degradation was confirmed to participate in the degradation of the uridylated pre-let-7 miRNAs, leading to a low expression level of let-7 (Chendrimada et al., 2005; Faehnle et al., 2014; Haase et al., 2005; Hutvagner et al., 2001; Yates et al., 2013).

1.2.2 Functions

1.2.2.1 Translation and regulation

Mature let-7 expression level increases following cell differentiation and proliferation. In turn, there is no detection of mature let-7 in embryonic stem cells, while pri-let-7 and pre-let-7 can be found in these cells, indicating the biogenesis of let-7 is tightly regulated by diverse factors, including post-transcriptionally (Hagan et al., 2009; Parisi et al., 2017; Rahkonen et al., 2016). Lin28 RNA binding protein was found to inhibit the generation of mature let-7 (Thomson et al., 2006). There is an intertwined relationship between them: Lin28A can bind pre-let-7 and block its biogenesis, while let-7 can silence the mRNAs of Lin28A through RISC action (Heo et al., 2008; Kim & Kim, 2012). As a result, studies began investigation of the mechanisms behind the interactions within this complex (Newman et al., 2008; Rybak et al., 2008). Several research groups have demonstrated that the high-expression of Lin28A in cancers leads to the loss of let-7 miRNAs (Albino et al., 2016; De Almeida et al., 2018).

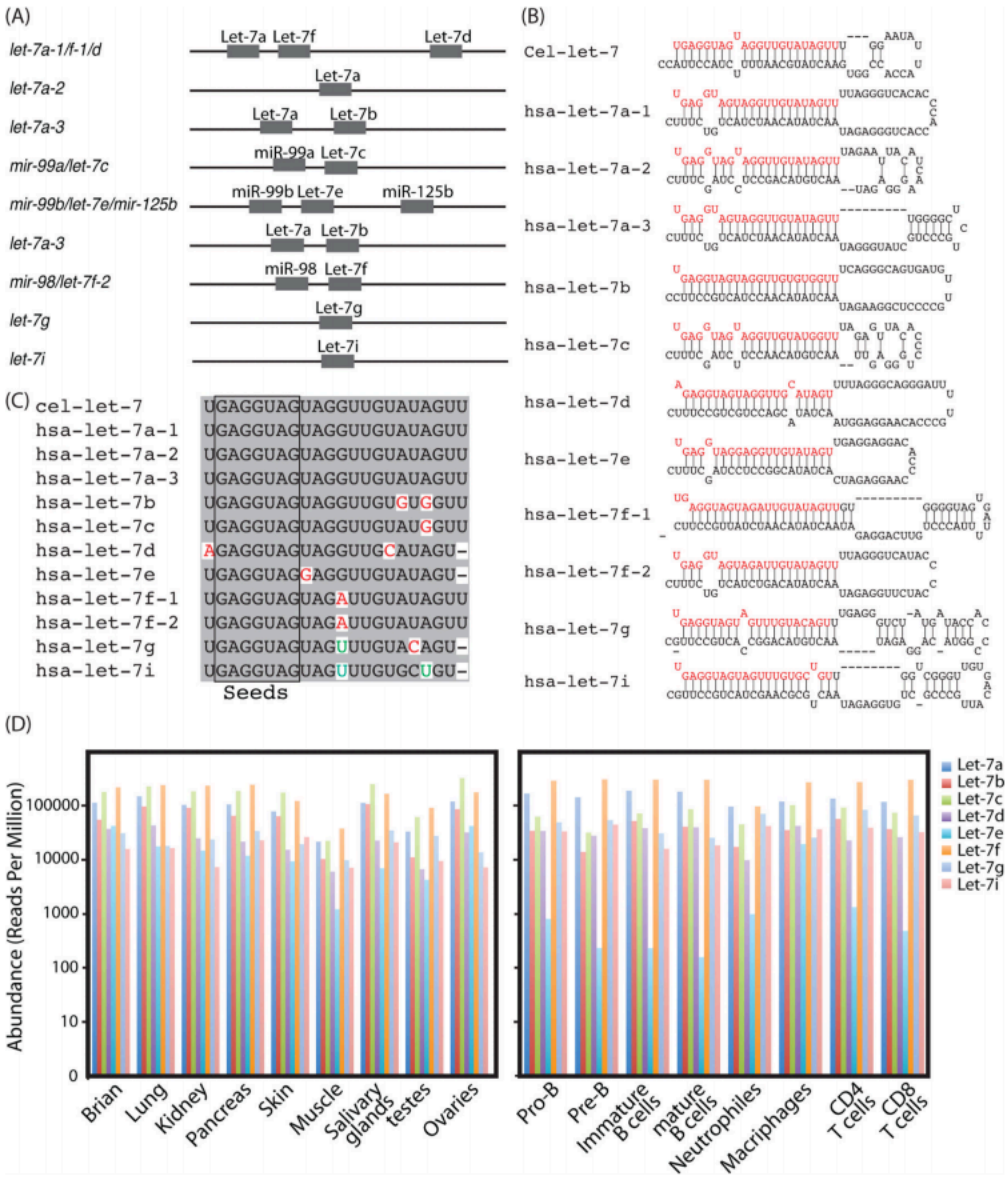


Figure 1.5 Genes of the let-7 family and their mature miRNA expression patterns in diverse tissues and cell types in mice. (A) Classification of let-7 members. (B) Secondary structure of precursor-let-7 miRNAs. (C) Sequence alignment of let-7 members. "GAGGUAG" is conserved in all let-7 miRNA family members. (D) The abundance of let-7 miRNAs expressed in diverse tissues and cell types of mice. All let-7 microRNAs have high expression level in diverse cells and organs. This figure is adapted from (Chen, 2013).

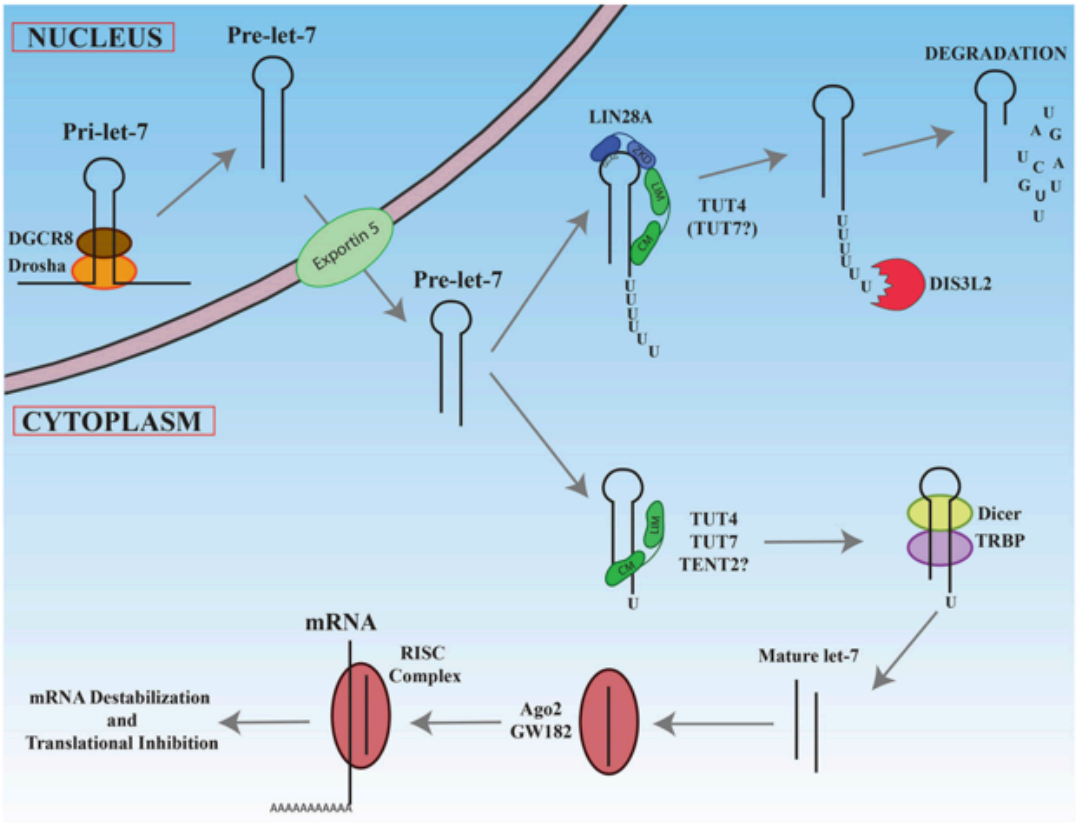


Figure 1.6 Diagram of let-7 microRNA biogenesis. RNA transcript primary-let-7 (pri-let-7) is initially processed by DGCR8/Drosha complex in the nucleus and the subsequently generated precursor-let-7 (pre-let-7) is exported into the cytoplasm. In the absence of RNA binding protein Lin28A, TUTases add a single uridine to the 3' end of pre-let-7. Mono-uridylated pre-let-7 is processed by Dicer to mature let-7, which regulates gene expression via RISC complex. In the presence of Lin28A, TUTases add oligo uridines to the 3' end of pre-let-7. This modification blocks the Dicer cleavage and oligo-uridylated pre-let-7 is recognised by the exonuclease Dis3l2, which subsequently degrades oligo-uridylated pre-let-7. This figure is taken from (Menezes et al., 2018).

1.2.2.2 Let-7 and oncogenes

Let-7 is regarded as a significant suppressor for several oncogenes, and thought to be related to many instances of cancer due to its abnormal expression and dysfunction (Yashiro & Tomita, 2018). An early study reported the reduced expression of let-7 in human lung cancers for the first time in 2004. From 143 lung cancer cases that had undergone potentially curative resection, they found the expression level of let-7 was positively correlated to the survival period of lung cancer patients. The group with low expression levels of let-7 had significantly shorter survival time after potentially curative resection than those who had a high expression level of let-7, while overexpression of let-7a or let-7f caused the inhibition of cancer cell proliferation (Takamizawa et al., 2004). The following study in 2005, showed that 60/RAS, a critical oncogene, was regulated negatively by members of the let-7 family. All human mRNAs of RAS genes have complementary sites to the let-7 sequence in their 3' UTRs, which causes the blocking of RAS gene expression after let-7 binding (Johnson et al., 2005; Steinemann et al., 2010). The importance of this inhibiting mechanism has been proved further in the RAS/MAPK pathway. RAS induces epithelial to mesenchymal transition (EMT) by downregulating E-cadherin, and subsequently let-7 can inhibit the expression of RAS genes and prevent the malignant transformation of tumor cells (Hameiri-Grossman et al., 2015; Oliveira-Mateos et al., 2019). Besides, oncogenes, like c-MYC and HMGA2, also have complementary sites with let-7 family members (Kim et al., 2009; Wu et al., 2015). Recently, emerging evidence in diverse cancers shows that the mutations of let-7 cause the overexpression of RAS, c-MYC, and HMGA2 (Ioannidis et al., 2005; Lee & Dutta, 2007; Ranjbar et al., 2018), and promote the development of tumorigenesis. Other targets such as HER2, Bcl-2 and CDC25A are also shown to be regulated by let-7 by a similar mechanism (Li et al., 2020; Singh & Saini, 2012; Wielgos et al., 2017). Thus, let-7 regulates the renewal of stem cells and the cell cycle (Fairchild et al., 2019, 2021; Rybak et al., 2008; Torres Fernandez et al., 2021). These results indicate let-7 is a critical cellular regulator of oncogenes and anti-tumor agent, which highlights mechanisms leading to its suppression as targets for better treatment of cancer patients.

miRNA	Clinical value
Let-7i	- reduced expression shortens recurrence-free survival. - reduces sensitivity to cisplatin.
	- increased expression increases sensitivity to paclitaxel.
	- reduced expression causes a worse prognosis for the patient.
Let-7a	- hyperexpression reduces overall survival, sensitivity to platinum and taxanes.
	- hyperexpression increases the sensitivity to platinum drugs. - reduces the sensitivity to taxanes.
	- reduces sensitivity to taxanes, doxorubicin, interferon α.
	- let-7a should be considered together with Lin28B and IGF-II, of which Lin28B is the main element with the worst prognostic value in hyperexpression.
Let-7d,c	- reduced expression causes a worse prognosis for the patient.
Let-7e	- reduced expression shortens recurrence-free and overall survival. - reduces sensitivity to cisplatin.
	- increases sensitivity to cisplatin.
	- increased expression reduces sensitivity to paclitaxel.
	- reduced expression causes a worse prognosis for the patient.
Let-7d	- reduced expression causes a worse prognosis for the patient.
	- proves to be an oncosuppressive miRNA.
Let-7g	- prolongs recurrence-free survival. - increases sensitivity to paclitaxel and vinblastine, does not affect sensitivity to carboplatin.
Let-7f	- reduced expression causes a worse prognosis for the patient.

Figure 1.7 Clinical roles of diverse let-7 family members. Most let-7 miRNAs have clinical significance. Let-7g, which is studied in this thesis, affects recurrence-free survival and sensitivity to paclitaxel and vinblastine. This figure is taken from (Ferneza et al., 2021).

1.3. Lin28: regulator of the self-renewal of stem cells

1.3.1 Discovery and function

Cell lineage abnormal 28 (Lin28) gene was initially found in *C. elegans* as a key regulator controlling the timing of developmental events and participating in heterochronic processes (Ambros & Horvitz, 1984). Another study proved that inactivating mutations accelerate the differentiation of subcutaneous and vulvar stem cells at the transition stage from Larval stage 2 (L2) to L3, while activating mutations promote the self-renewal of L2 larval stem cells (Moss et al., 1997)

Lin28 is highly conserved across diverse organisms. In *C. elegans,* there is only one encoded Lin28 protein, while in vertebrates, there are two paralogs, Lin28A and Lin28B (Yang & Moss, 2003), which both belong to the zinc finger CCHC-type (ZCCHC) superfamily (Wang et al., 2021). Structurally, shown in Figure 1.8, Lin28 consists of two conserved RNA-binding domains, a cold-shock domain (CSD) at the N-terminus and a tandem CCHC (Cys-Cys-His-Cys) zinc-finger (ZnF) domain at the C-terminus (Balzeau et al., 2017; Nam et al., 2011). Lin28B is longer in sequence than Lin28A, with additional putative nucleolar localization signal (NoLS) and nuclear localization signal (NLS) motifs (Balzeau et al., 2017; Chen et al., 2019; Kang et al., 2017).

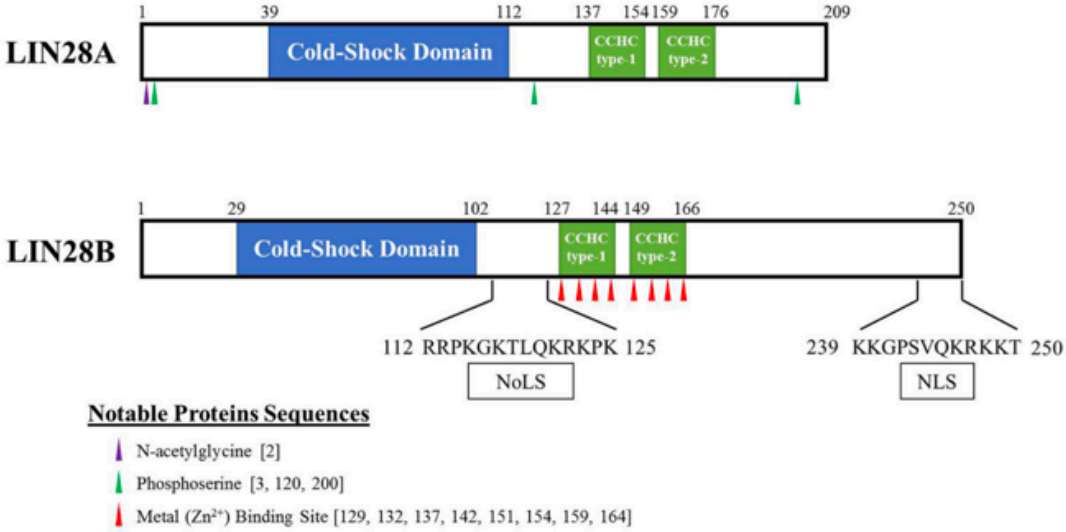


Figure 1.8 Schematic diagram of Lin28A and Lin28B proteins from human. Both proteins consist of cold shock domain (CSD) and Zinc knuckle domain, while Lin28B has additional inclusion of nucleolar localization signal (NoLS) and nuclear localization signal (NLS) motifs. This figure is adapted from (Balzeau et al., 2017).

1.3.2 Lin28A and let-7

Lin28A and let-7 were first discovered and studied together in the nematode *Caenorhabditis elegans* in the regulation of developmental timing (Ambros & Horvitz, 1984). As previously discussed, Lin28A can block the biogenesis of let-7 by recruiting uridylytansferases TUT4/TUT7 to add oligo-uridines at the 3' end of pre-let-7, which subsequently is degraded by Dis3l2 exonuclease (Chang et al., 2013). This is an indirect regulation for downstream target oncogenes. On the other hand, the CSD and Zinc Knuckles (ZK) domain of Lin28A can bind the NGAU and GGAG of precursor-let-7 (pre-let-7) respectively, aiding the recruitment of TUTases to add a U-tail. This modification prevents onwards processing by Drosha as well as Dicer, causing the low expression level of mature let-7. The Lin28/let-7 axis has been studied in diverse pathways both in vivo and vitro, Figure 1.9.

The structure of a Lin28A/pre-let-7 complex showed for the first time how the CSD and ZK domains of Lin28A interact with two distinct single-stranded (ss) regions of pre-let-7 (Loughlin et al., 2011; Nam et al., 2011). The CSD circles the loop of pre-let-7, while the ZK domain interacts with the GGAG motif at the 3' end of pre-let-7. The variable linker between the two domains of Lin28A is missing from the structure due to flexibility. The electrical potential diagram for the complex shows that positively charged surfaces on both domains are towards RNA throughout (Nam et al., 2011) (Figure 1.10). This structure itself provided useful details for potential development of therapeutics drugs aiming to affect let-7 processing.

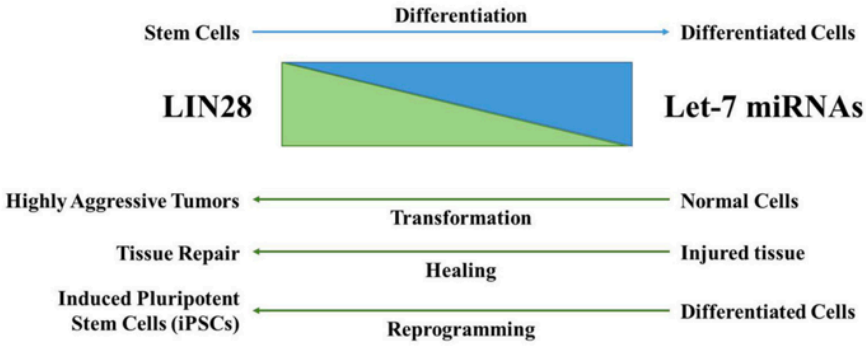


Figure 1.9 Lin28A is tightly linked with let-7 miRNAs in the regulation of biological functions. The expression level of Lin28 and let-7 miRNAs is negatively correlated. Lin28 has highest expression level in stem cells, while let-7 miRNAs express more in differentiated cells. The abnormal expression of Lin28 in differentiated cells would cause issues like transformation of normal cells to highly aggressive tumors. This figure is taken from (Balzeau et al., 2017).

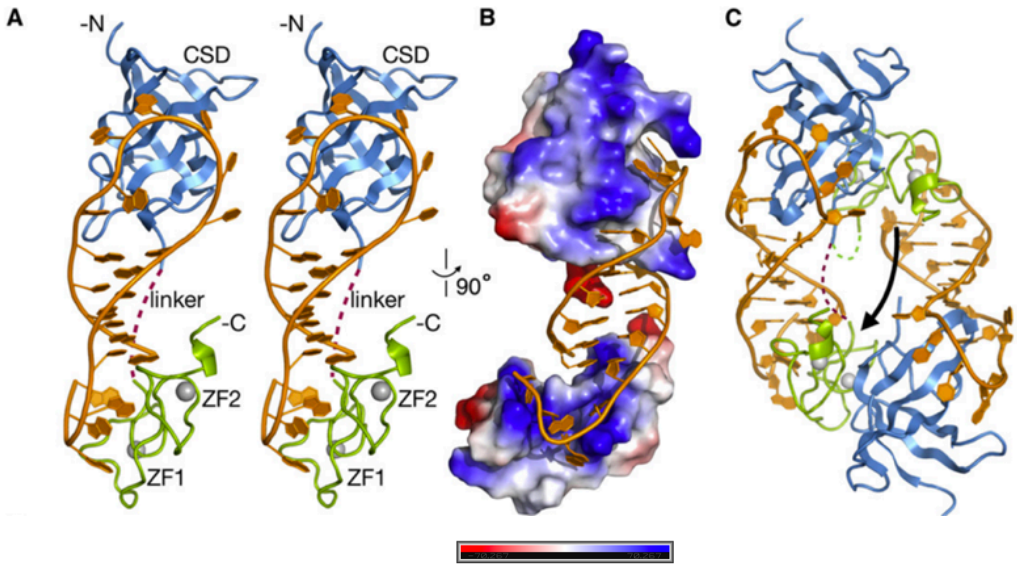


Figure 1.10 Structures of mouse Lin28A/PreE-let-7 complex. (A) The monomeric complex shown in stereo mode. A purple dotted line shows the interdomain linker regions of Lin28A. (B) Presentation of the electrostatic potential on the surface of the complex. The different color shows the electrostatic potential from negative charge (red) to positive charge (blue). (C) Dimeric crystallographic complex shows domain swapping around the PreE-let-7. The unwrapped domains are noted with a purple dotted line. This figure is taken from (Nam et al., 2011).

1.3.3 Lin28A and SUMOylation

Lin28A is subject to diverse modifications, such as methylation, phosphorylation and acetylation (Kim et al., 2014; Tsanov et al., 2017; Wang et al., 2014). These changes alter the intracellular localization, activity and stability of Lin28A (Wu et al., 2022). Recently, Jianxiu Yu' s lab proved the Lysine (K) at the position 15 of Lin28A can be SUMOylated, depicted in Figure 1.11. This modification enhances the interaction between Lin28A and pre-let-7, promotes the oligo-uridylation activity of TUT4 to pre-let-7, and thus decreases the final expression level of mature let-7. In conclusion, SUMOylation negatively regulates let-7 and may be considered as a marker for early diagnosis and prognostic evaluation of tumors (Dou et al., 2020).

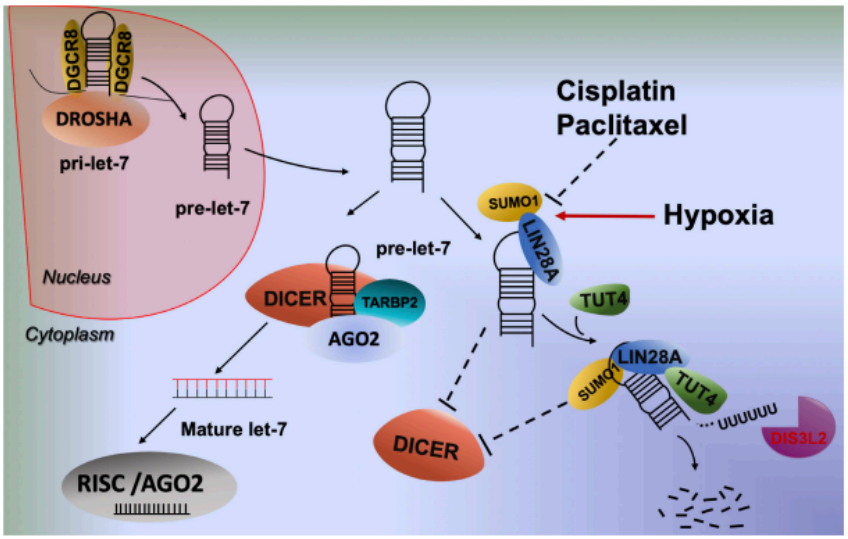


Figure 1.11 SUMOylation modification on Lin28A leads in the degradation of pre-let-7 and promotes the growth of tumor in vivo. The SUMO1 has a highly positively charged surface, increasing the binary interaction of pre-let-7 to SUMOylated Lin28A via binding to the negatively charged miRNAs. The intense interaction promotes the recruitment of TUT4 to add oligo-U tail to pre-let-7 and thus blocks the processing by Dicer to mature let-7. This uridylation process can be inhibited by cisplatin paclitaxel but promoted via hypoxia. This figure is adapted from (Dou et al., 2020).

1.3.4 Lin28A and cancers

Lin28A is a regulator of cell proliferation and differentiation via let-7 dependent and independent pathways (Balzeau et al., 2017). As shown in Figure 1.12, by blocking the regeneration of mature let-7, Lin28A upregulates the expression of RAS, c-MYC, and HMGA1/2, causing the evasion of growth suppressors, genome instability, enabling replicative immortality and finally accelerating tumorigenesis (Chen et al., 2016; Li et al., 2013; Oh et al., 2010). However, without the assistance of let-7, Lin28A still can bind the mRNAs of several oncogenes and various cell cycle factors to promote the expression of oncogenes, resulting in the development of tumors. For example, in ovarian cancer, Lin28A can bind directly to the mRNAs of bone morphogenetic proteins (BMP4) and subsequently induce the overexpression of BMP4. This process would promote cell proliferation and tumor growth (Balzeau et al., 2017). Besides, Lin28 proteins participate in cell growth, development and metabolism (Nguyen & Zhu, 2015; Shyh-Chang & Daley, 2013; Thornton & Gregory, 2012; Xu et al., 2009). In total, evidence reveals the close relationship between Lin28 and cancers and suggests a therapeutic dawn in this area for clinical research.

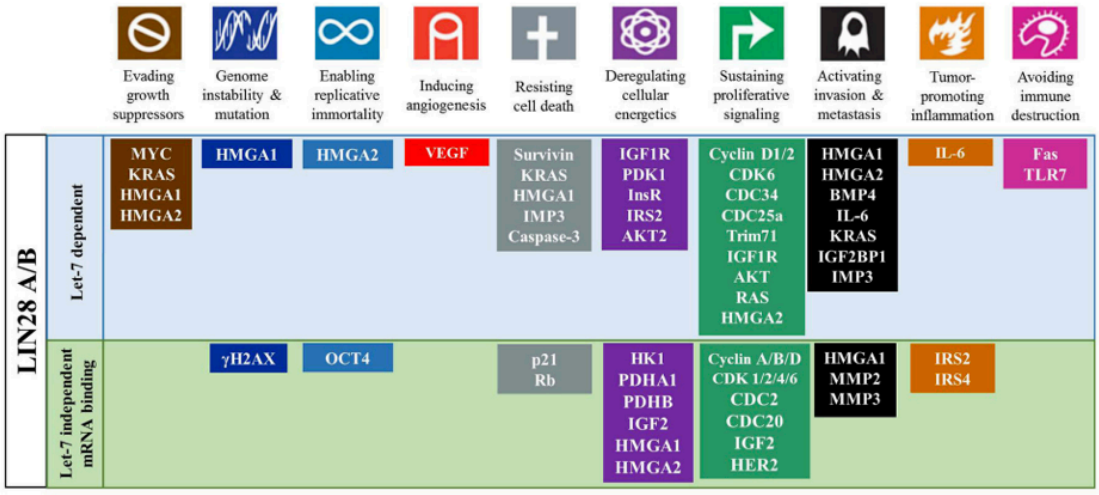


Figure1.12 Lin28 proteins have significant influence on cancer hallmarks by either let-7 dependent or let-7 independent pathways. Evading growth suppressors, inducing angiogenesis and avoiding immune destruction can be mediated only in the presence of let-7, while other pathways can be affected in the conditions with or without the participation of let-7 miRNAs. This figure is adapted from (Balzeau et al., 2017).

1.4. TUTases：a class of non-canonical poly(A) polymerases

1.4.1 Discovery

There are two homologous uridylyltransferases involved in both mono- and oligo-U addition: TUT4 and TUT7. TUT4/7 are members of the non-canonical polyA polymerase family involved in 3' uridylation of pre-let-7 miRNAs. Overall, human TUT4 and TUT7 share 48% sequence identity, and 69% sequence similarity in their catalytic module (CM) (Faehnle et al., 2017a). Unlike most uridylyltransferases, TUT4 and TUT7 are modular, multi-domain enzymes (Faehnle et al., 2017a; Lim et al., 2014; Yamashita et al., 2019). Structurally, shown in Figure 1.13, TUT4/7 consists of a classical Cys-His-His-Cys (CHHC) zinc finger (ZF), which is known to bind RNAs, DNAs, or proteins. At the N- and C-terminus of TUT4 and TUT7, there are two nucleotidyltransferase domains (NTD1 and NTD2) connected by a flexible internal region. In addition, TUT4 and TUT7 both have three Cys-Cys-His-Cys (CCHC) zinc knuckles. The C-terminal NTD2 is the catalytic module and promotes mono-uridylation, while NTD1 is responsible for the interaction with the Lin28A:pre-let-7 complex but has no transferase activity due to the loss of key aspartate residues in catalytic sites (Thornton et al., 2012).

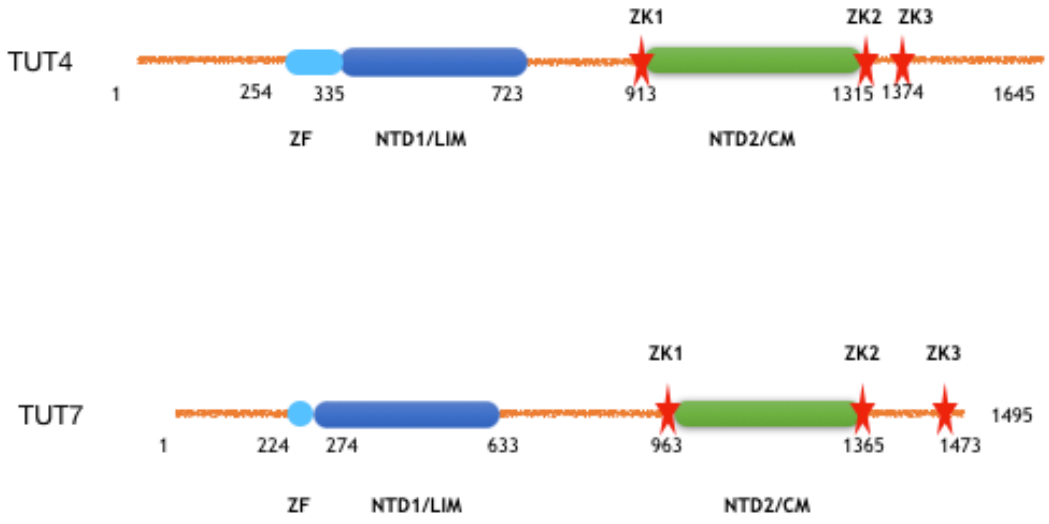


Figure 1.13 Schematic diagram of domain architectures of human TUT4 and TUT7. TUT4 and TUT7 have similar composition with two nucleotidyltransferase domains (NTD1 (blue) and NTD2 (green)), zinc finger (teal) and multiple zinc knuckle motifs (red star).

1.4.2 TUTases structures

1.4.2.1 Cid1

Caffeine-induced death-suppressor protein 1 (Cid1) is a member of the family of putative nucleotidylransferases and was initially found in a study of mitosis in fission yeast, as a regulator controlling cell cycle (Wang et al., 2000), which uridylates polyadenylated mRNAs and stimulates their decay (Rissland & Norbury, 2009). It was the first poly-U polymerase to have its structure revealed at atomic level and shows the basis of UTP selectivity and RNA binding. Structurally, a histidine residue discriminates UTP from ATP, and three patches on the enzyme surface formed by basic amino acid side chains create a critical motif, which captures the RNA (Yates et al., 2012). Later, a proposed model for the catalytic cycle was developed with the solution of higher-resolution structures in multiple states (Yates et al., 2015).

1.4.2.2 TUT7 CM

The Leemor Joshua-Tor lab determined the human TUT7 catalytic module (CM) in apo and complex forms with either dsRNA, U2, U5 or UTP (Faehnle et al., 2017). The N-terminus consists of five β-strands closely wrapped and two α-helices, while the C terminus has six α-helices and four short β-strands connected by several long loops, in which the ZK2 can be visualized with the presence of dsRNA. Between the N and C loop, a cleft is formed, which captures the UTP in the bottom of the cleft where the 3’ end of dsRNA binds to the top edge of the cleft. Structures solved with diverse substrates bound reveal that the RNA-binding pocket orients the group II pre-let-7 to favor mono-uridylation, while the ZK2 of TUT7 aids oligo-uridylation by engaging the extension of the oligo-U tail (Faehnle et al., 2017).

1.4.2.3 TUT4 LIM

The Tomita lab solved the human TUT4 Lin28A-interacting module (LIM) structure (Yamashita et al., 2019). Overall, the human TUT4_LIM (NTD1) is similar to human TUT7_CM (NTD2), consisting of very similar components including nc-palm and fingers regions. The distinguishing finding in this structure is the Zinc finger (ZF) at the N terminus, which is made of by three β-strands and five α-helices. A zinc ion binding motif consists of Cys206, Cys309,

His322 and His328 (Yamashita et al., 2019). The TUT4_LIM structure demonstrates that the ZF of TUT4_LIM is homologous to those of double-stranded RNA binding zinc-fingers; this ZF enables the ternary complex formation with pre-let-7 and Lin28A.

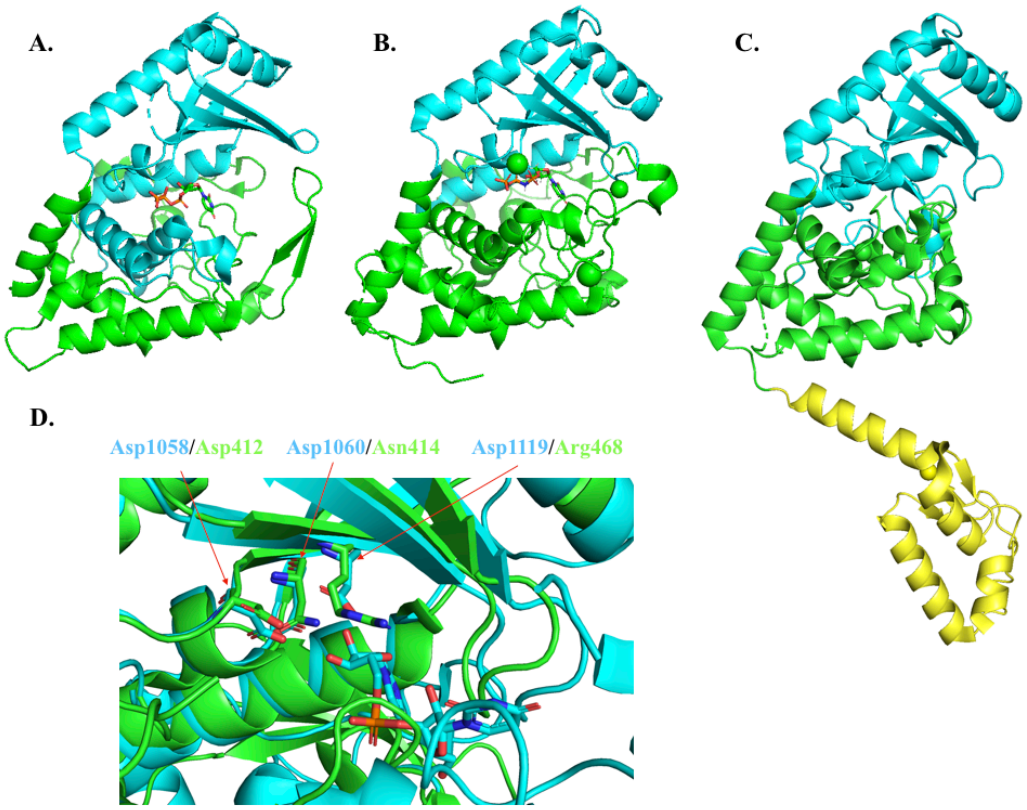


Figure 1.14 Comparison among Yeast Cid, human TUT7_CM and human TUT4_LIM A. Cid1 drawn from PDB:4e80. Blue and green color present the fingers domain and palm domain respectively. B. TUT7_CM drawn from PDB:5w0n. Blue and green color present the fingers domain and palm domain respectively. C. TUT4_LIM drawn from PDB: 6iw6. Blue, green and yellow color present the fingers domain, palm domain and zinc finger respectively. D. Super-position of the TUT4_LIM (green) and TUT7_CM (blue) with the catalytic Asp loci labelled for each. The LIM module (NTD1) loses uridylation activity due to the mutation on the second key residue from aspartate to asparagine (Asn414) compared to the Asp1060 in CM of TUT7.

1.4.3 Functions

1.4.3.1 TUTases and let-7

Let-7 expression is closely controlled by the uridylyltransferases TUT4 and TUT7 by targeting the precursor let-7 (pre-let-7) (Faehnle et al., 2017a; Thornton et al., 2012; Yu et al., 2020). As previously discussed, pre-let-7, processed by RNA II polymerase Drosha (Mondol et al., 2015), can be divided into two groups. Group I with a 2-nt overhang at the 3' end, can be processed by Dicer directly to mature let-7, while group II, accounting for 90% of transcripts, with a 1-nt overhang at the 3' end, needs to be mono-uridylated by TUT4 or TUT7, and can then be processed by Dicer. However, with the presence of Lin28A protein, TUT4/7 can add oligo-uridines to 3' of pre-let-7, which subsequently leads to exonucleolytic pre-let-7 degradation (Chang et al., 2013; Loughlin et al., 2011; Mayr et al., 2012; Norbury, 2013; Roos et al., 2015; Rybak et al., 2008; Yu & Kim, 2020).

1.4.3.2 TUTases and mRNA/histone mRNA

Besides pre-let-7, TUTases also target other types of RNAs, like mRNAs (Norbury, 2013; Yashiro & Tomita, 2018; Yu & Kim, 2020). In cells, TUT4/7 selectively recognize and uridylate mRNAs, which already have short polyA tails (less than ~25nts). Knockout experiments demonstrate that the great majority of mRNAs lose the oligo-U-tails after the deletion of TUT4/7, leading to the increasing of their half-lives in vivo (Holmquist et al., 2022; Lackey et al., 2016; Lim et al., 2014). A well-characterized system is of TUT7 regulating histone mRNAs. Histone mRNAs are not polyadenylated, ending instead with a conserved loop and their expression levels are tightly correlated with the mammalian cell cycle (Harris et al., 1991). Expression levels of histone mRNAs will surge at the beginning of S-phase while rapidly decreasing at the end of S-phase. The expression regulation of histone mRNAs mainly occurs at their 3' end by diverse mechanisms (Holmquist et al., 2022; Mossanen-Parsi et al., 2021; Potter-Birriel et al., 2021).

The whole process of histone mRNAs uridylation is shown as Figure 1.15. In the cell nucleus, the processed histone mRNAs leave a 5-nt tail in the 3' end after the stem-loop. This then recruits SLBP and 3' hExo proteins to form a ternary complex (Jaeger et al., 2006; Whitfield

et al., 2004) before the 3' hExo nuclease hydrolyzes and removes 2 nts in the 3' end (Jaeger et al., 2006). There, if 3' hExo trims more than 2 nts, TUT7 adds one uridine to the 3' end during S-phase. However, at the end of S-phase, 3' hExo trims more nucleotides and subsequently TUT7 adds a long U-tail, which marks histone mRNAs to be degraded by the exosome (Holmquist et al., 2022; Lackey et al., 2016).

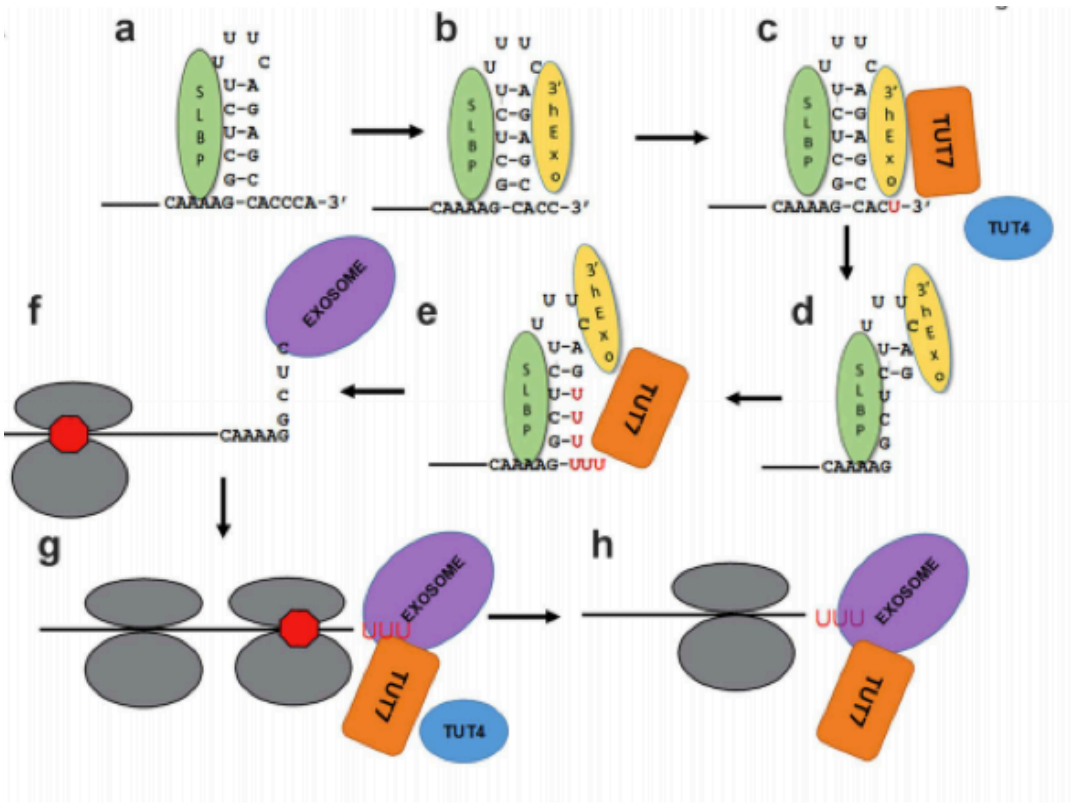


Figure 1.15 Proposed model of TUT7 and 3' hExo involved in histone mRNA degradation. Histone mRNA in the cell nucleus would be processed by SLBP/3' hExo, and then leave a 5-nt tail in its 3' end. TUT7 participates the regulation of histone mRNA accompanied by the change of the length of its 3' end tail during cell division. A long oligo-uridine tail is a sign for being degraded by the exosome. This figure is adapted from (Holmquist et al., 2022).

1.4.3.3 TUTases and Retrotransposition

Besides histone mRNAs, TUTases have also been shown to restrain retrotransposition by uridylating the LINE-1 mobile elements. Andrzej Dziembowsk' s lab proved that the 3' ends of LINE-1 mRNAs are pervasively uridylated in various human and mouse cellular models (Warkocki, Krawczyk, et al., 2018). In their proposed model, TUT7 adds a uridine tail to LINE-1 mRNAs in the cytoplasm and inhibits the initiation of reverse transcription of LINE-1 mRNAs: they are then reimported to the nucleus, whereas TUT4 uridylates LINE-1 mRNAs and destabilizes them to be degraded. All these processes are assisted by MOV10 helicase and two LINE-1 coded proteins: L1-ORF1p and L2-ORF2p, shown in Figure 1.16.

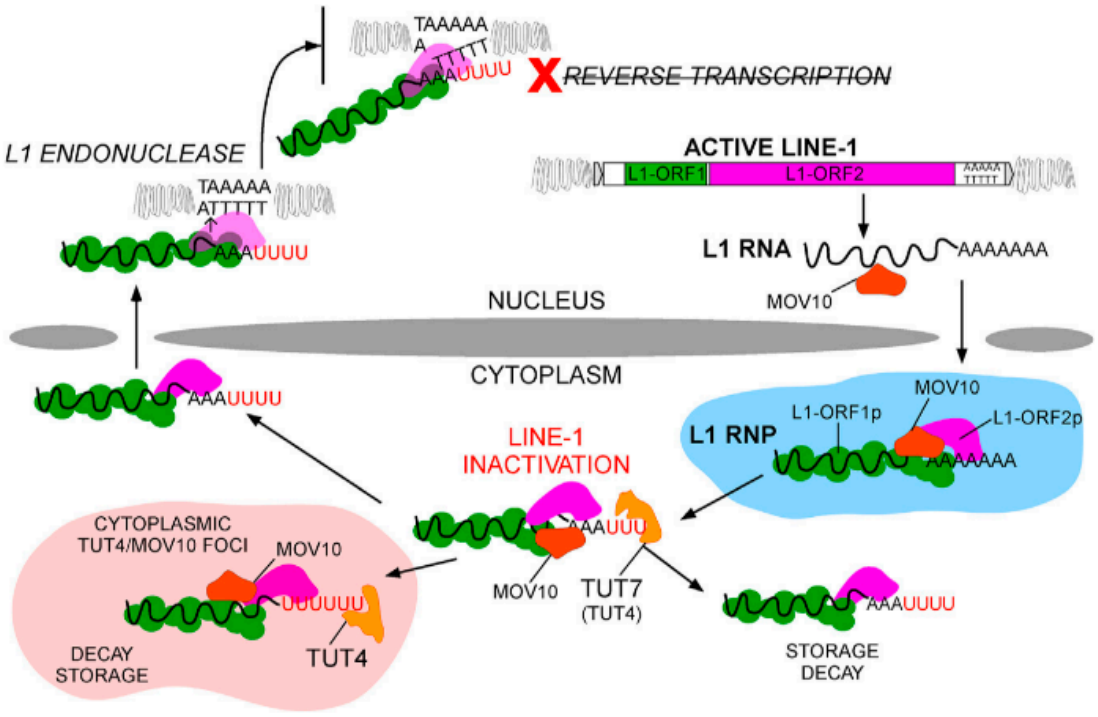


Figure 1.16 Proposed model of TUT4/7 involved in restricting L1 Retro-transposition. MOV10 assists the complex formation of LINE-1 mRNA, L1-ORF1p, L1-ORF2p and TUTases. This figure is taken from (Warkocki, Krawczyk, et al., 2018).

1.4.3.4 TUTases and Antiviral defence

TUTases participate in innate immune processes (Lin et al., 2021; Minoda et al., 2006). In *C. elegans*, the CDE-1 TUTase homologous enzyme uridylates the 3′ end of its natural pathogen, Orsay virus' (OrV) RNA genome and promotes its degradation in a manner independent of the RNAi pathway (Le Pen et al., 2018), shown in Figure 1.17. Similar trials were performed on mammalian TUT4 and TUT7. The TUT4(7) knockdown group has a higher infection level with Influenza A Virus (IAV), indicating TUT4(7) may act as an early barrier against IAV infection in mammalian cells, like CDE-1 against Orsay virus in *C. elegans* (Le Pen et al., 2018).

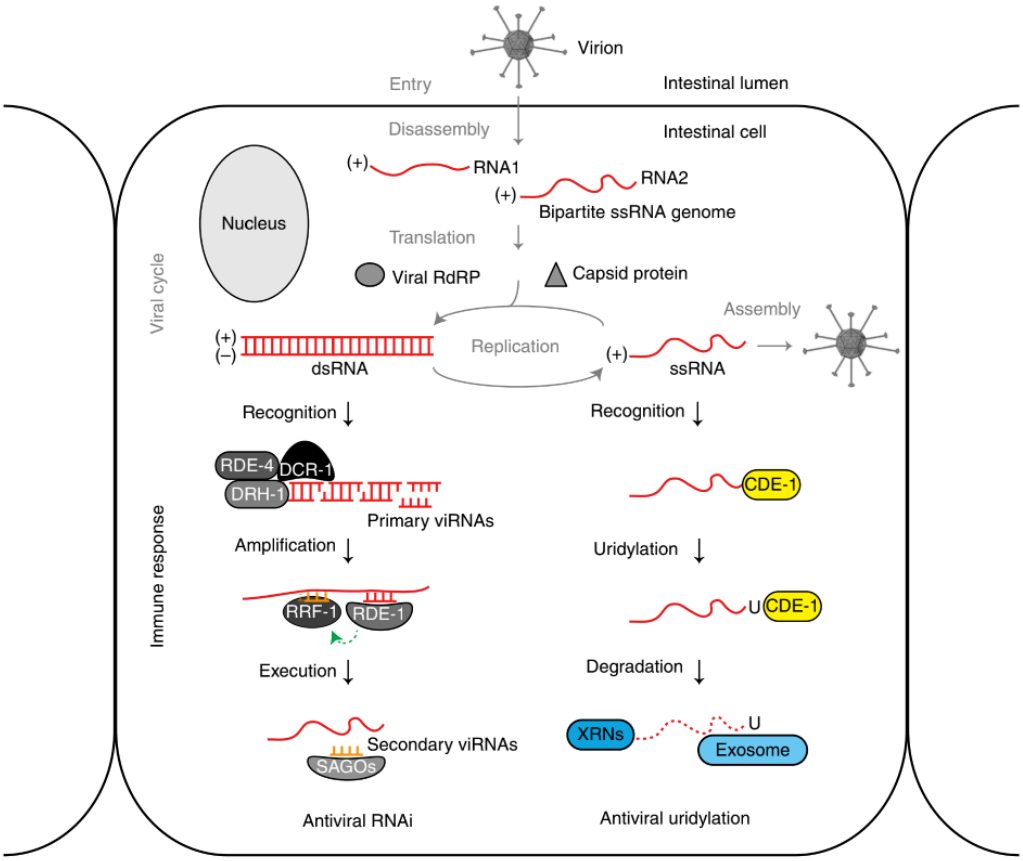


Figure 1.17 Schematic presentation shows the virus terminal uridylation is parallel to antiviral RNAi in immune-defense pathways of *C. elegans*. They both play a key role to protect against virus infection as an innate immune response. This figure is adapted from (Le Pen et al., 2018).

1.4.4.5 TUTases and the Immune system

Cytokines are a class of small proteins with molecular weight around 5-25 kDa, responsible for cell signaling regulation (Barker et al., 2018; Quagliato & Nardi, 2018). Their expression is controlled by RNA binding proteins and diverse microRNAs that target their mRNAs (Chen et al., 2011). TUT4 (ZCCHC11) was shown to be involved in cytokine expression as a regulator of the Toll-like receptor (TLR) pathway, in which TUT4 interacts with TIFA after LPS treatment and subsequently inhibits the activation of NF-kappaB (Minoda et al., 2006) shown in Figure 1.18. Besides, Joseph Mizgerd' s lab demonstrated that TUT4 directly regulates microRNA-26, which binds to the 3' UTRs of IL-6 mRNAs. The additional terminal uridines on miRNA-26 added by TUT4 facilitate silencing of the mRNAs of IL-6, reducing the expression of IL-6, and thereby TUT4 negatively regulates the expression of IL-6 (Jones et al., 2009).

1.4.4.6 TUTases and Nanobody agents

Nanobodies, also known as a single-domain antibodies, are developed from the variable domain of the heavy-chain-only antibodies found in camelids and selectively bind specific antigens (Harmsen & De Haard, 2007). Compared with antibodies and Fab fragments, nanobodies are small, highly robust, resistant to extreme environments and even fully reversible from an unfolded state caused by high temperature (Doshi et al., 2014). They have also been developed into nanobody toolkits, such as Megabody and Legobody reagents, aiding the determination of small protein structures by cryo-EM and image reconstruction (Uchanski et al., 2021; Wu & Rapoport, 2021).

Recently, the Piotr Sliz lab systematically screened and identified a series of nanobodies, which target the LIM of human TUT4 by developing a functional epitope selection approach. A functional nanobody, Nb-S2A4, tightly binds to the Lin28A/pre-let7 interaction module and inhibits both monouridylation and oligouridylation activity of TUTases to pre-let-7g by blocking the ternary interaction, shown in Figure 1.19. In vivo experiments demonstrated that cells expressing Nb-S2A4 contain higher levels of the mature let-7 microRNA species than those without Nb-S2A4 expression and thus the overexpression of Nb-S2A4 affects the cell cycle directly or indirectly (Yu et al., 2020).

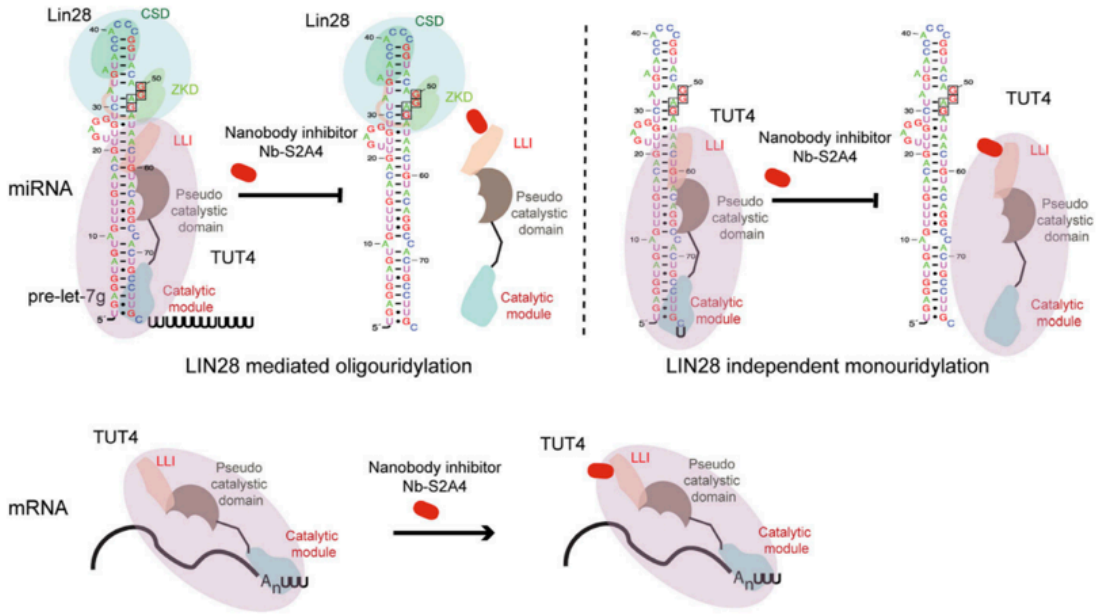


Figure 1.18 Proposed model of functional Nanobody S2A4 showing its inhibition of Lin28A mediated oligo-uridylation and Lin28A independent mono-uridylation. In Lin28 mediated oligo uridylation, S2A4 attaches to the zinc finger regions of TUT4, where a ternary complex forms with pre-let-7 and Lin28A, and thus inhibits the oligo uridylation. In Lin28 independent mono uridylation, the presence of S2A4 inhibits the binding of TUT4 to pre-let-7 and thus blocks the mono uridylation. This figure is taken from (Yu et al., 2020).

1.5. X-ray crystalogrphy and cryo-EM in structural biology

1.5.1 X-ray crystallography

1.5.1.1 History

X-ray crystallography is an experimental approach revealing atomic information about macromolecular structure (Wery & Schevitz, 1997). So far, at least 14 scientists, working in X-ray related fields, have been awarded Nobel Prizes. Roentgen was awarded the first Nobel Prize in Physics in 1901, for his breakthrough discovery of X-rays. Laue, and William and Lawrence Bragg were awarded Nobel Prizes in 1914 and 1915 respectively, honoring their contribution to the discovery of X-ray diffraction and determination of crystal structures (Shi, 2014). After that, around 10 scientists were awarded Nobel Prizes. The latest Nobel Prize related to X-ray crystallography was awarded to Kobilka in 2012 for the structure determination of GPCRs (Kobilka et al., 1987) (Figure 1.20).

1.5.1.2 Principles of X-ray diffraction

The goal of X-ray crystallography is to obtain three dimensional molecular structures from purified and crystallised samples (Smyth & Martin, 2000). In a well-diffracting crystal, the crystalline structure induces a beam of incident X-rays to diffract into multiple specific directions. The forward diffracted beams are imaged using a CCD detector (Zhao & Sakurai, 2017). Subsequently, after calculating the angles and intensities of these diffracted beams, and following Fourier transformation (a mathematical transform) (Hirschmugl & Gough, 2012; Su & Lee, 2020), a three-dimensional file of density of electrons within the crystal is produced by crystallographic software such as Phenix and CCP4 (Adams et al., 2010; Collaborative Computational Project, 1994; Hough & Wilson, 2018; Moriarty et al., 2016). The produced density map assists in locating the positions of atoms and chemical bonds using software such as Coot (Emsley & Cowtan, 2004; Goddard et al., 2005; Lam & Siu, 2017) and Phenix (Adams et al., 2002). The first determined biological macromolecule by X-ray was of myoglobin in 1957 by John Kendrew (Kendrew et al., 1958). After that, X-ray has commonly been used to solve biological components and large complexes, according to the workflow shown in Figure 1.20.

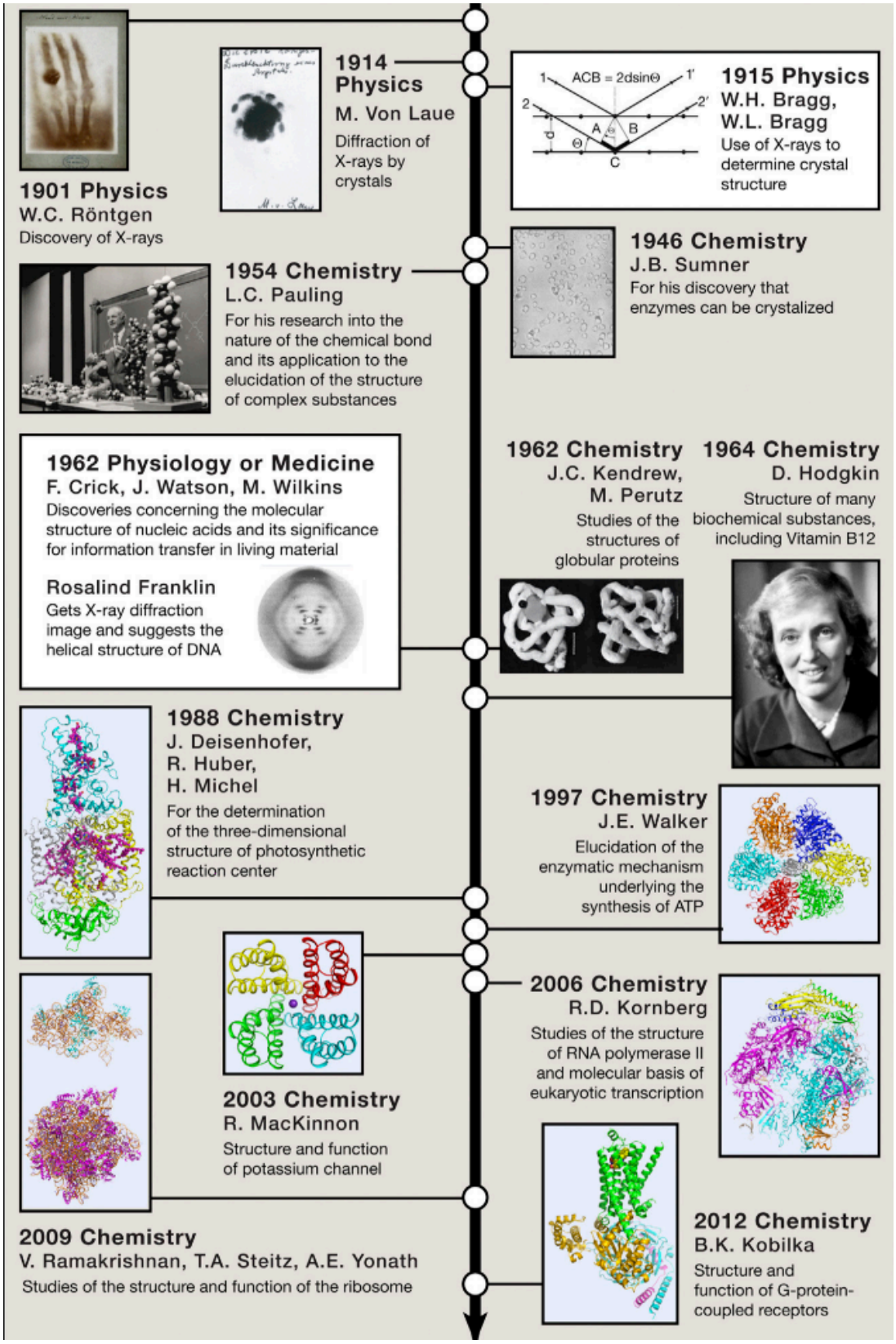


Figure 1.19 The history of X-ray crystallography marked by breakthroughs from 1914. X-ray crystallography nowadays has wide application such as small molecule-based drugs screening and identification. This figure is adapted from (Shi, 2014).

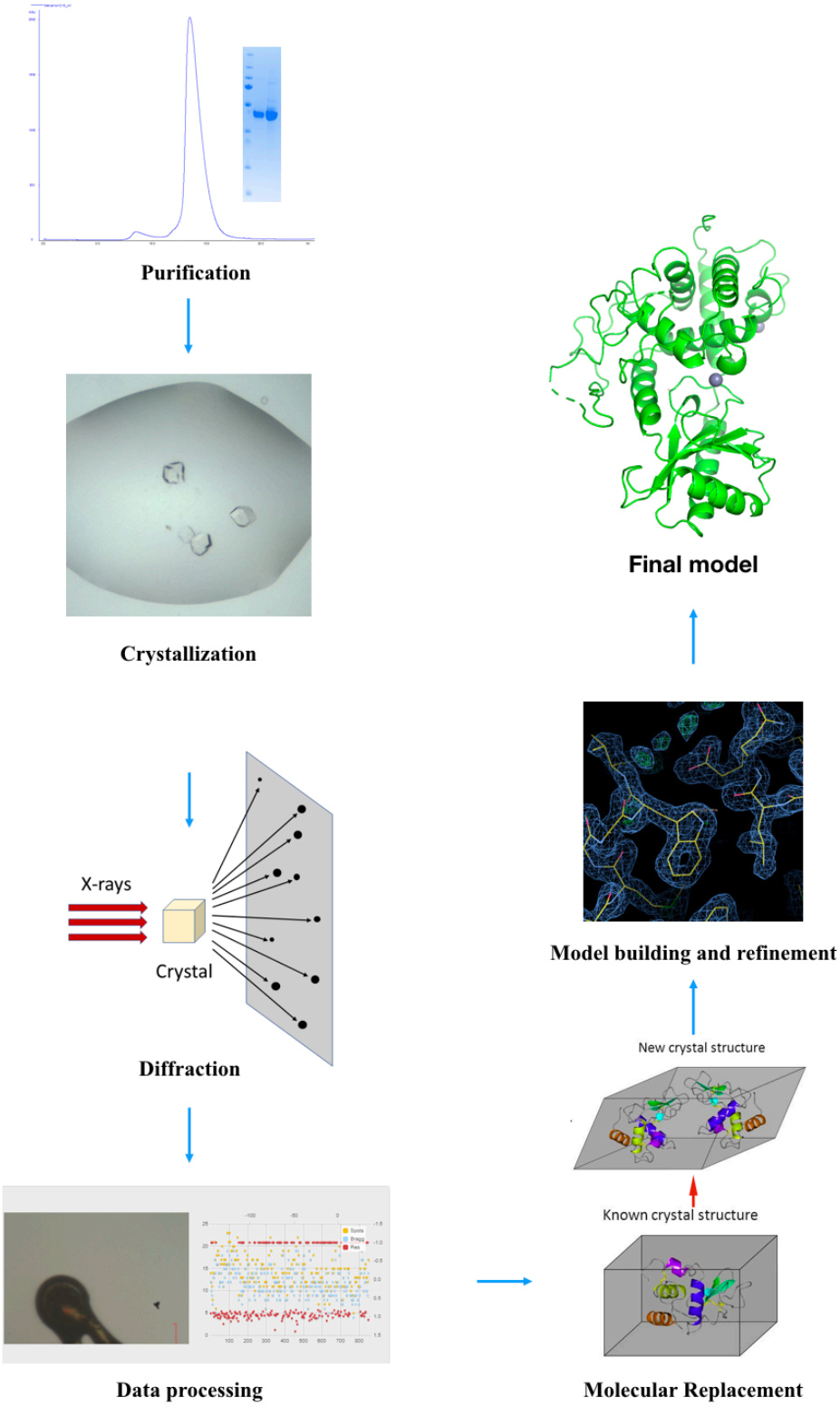


Figure 1.2: Legend on next page

Figure 1.20 The principles of X-ray crystallography and a brief flow chart showing the structure determination process for the human TUT7_CM crystal structure in this project. Initially, pure TUT7 CM with a single band as determined by staining with Coomassie blue, was obtained. Then diamond-shaped crystals were diffracted at Diamond Light Source and the data collected were automatically processed by ISPyB via combined software such as PHASER (McCoy et al., 2007), XIA2 (Gildea et al., 2022; Winter et al., 2013) and XDS (Brown et al., 2007; Kabsch, 2010). Molecular replacement was applied in this project to solve the phase issue. Finally, the TUT7 CM model was built and refined by using Coot (Emsley & Cowtan, 2004; Goddard et al., 2005; Lam & Siu, 2017) and Phenix (Adams et al., 2002).

1.5.2. Cryo-EM

1.5.2.1 History

Cryo-EM is an approach to determine macromolecular structures especially suited to larger and more complicated biological complexes (Lyumkis, 2019). It became increasingly powerful from about 1995 onwards. Then from about 2013 it became routine to determine atomic structures de novo from it. Prior to that it still played a very important role in setting crystal structures in context and understanding virus structure as well as the structures of the macromolecular machinery of cells (Frechin et al., 2022). In 2017, the Nobel Prize in chemistry was awarded to three scientists, Jacques Dubochet, Joachim Frank, and Richard Henderson, who had made significant contributions to the development of this technique. Recently, with the development of both hardware and software, like the Falcon 4 camera, Relion (Scheres, 2012) and Cryosparc software (Punjani et al., 2017; Scheres, 2012), cryo-EM has become a routine tool for structure determination of macromolecules to atomic resolution (Assaiya et al., 2021).

1.5.2.2 Principle

Single-particle analysis (SPA) is a process of imaging biological macromolecular samples by electron microscopy in a frozen-hydrated state before using the images to reconstruct 3D models (Lyumkis, 2019). A vast number of particles with different orientations are captured in the micrographs. Subsequently, the sub-regions containing particles are extracted computationally. The individual particles will be divided into different classes aiming to remove poorly imaged, damaged and spurious particles. Qualified particles are selected and used to reconstruct 3D models based on the relative angles of orientations. Depending on the resolution, the maps determined can be fitted with homologous models, predicted models, crystal and NMR structures, or used themselves for the de novo construction of a model, generating 3D structures (Wu & Lander, 2020; Xu & Dang, 2022) (Figure 1.21).

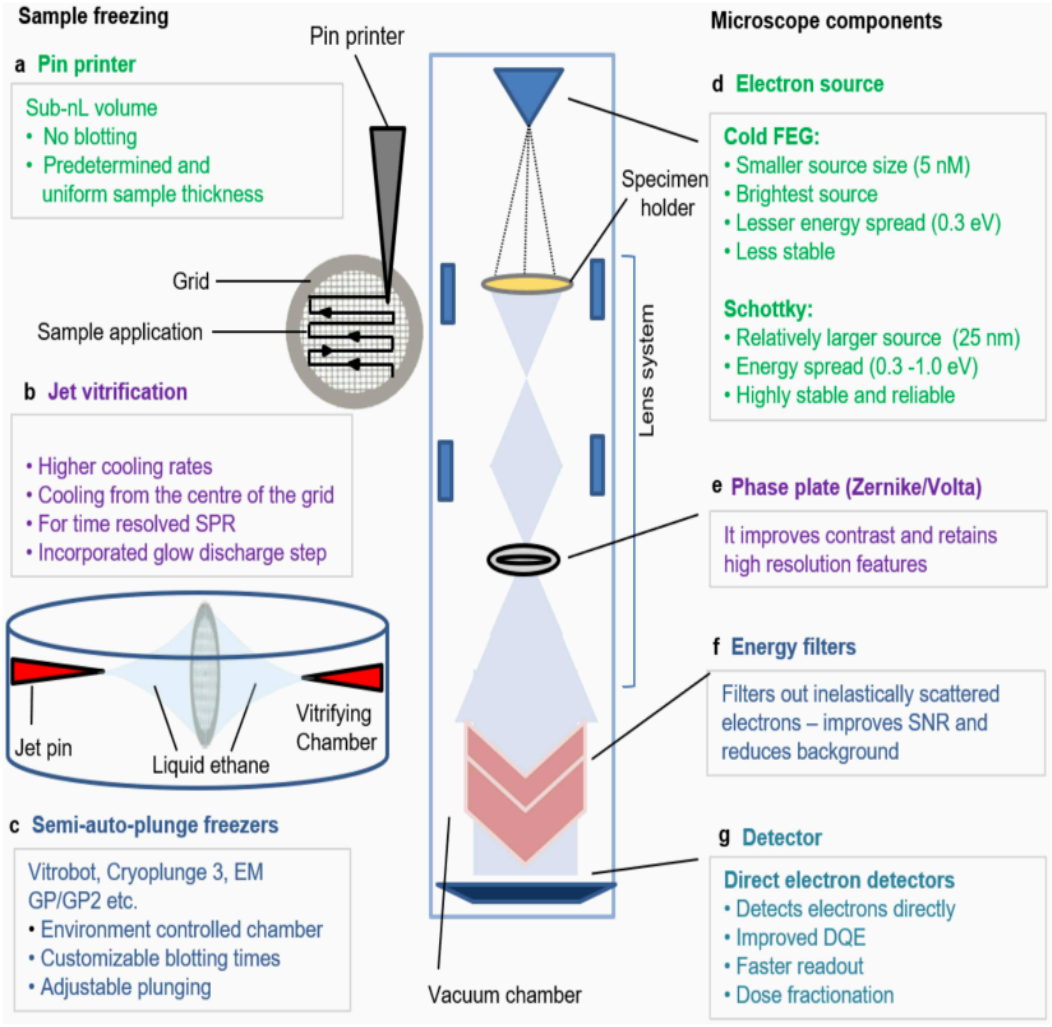


Figure 1.21 Flow chart of Cryo-EM sample preparation and the advancements in hardware and software for the cryo-EM in recent years. Besides this, the development of diverse types of grids are a key factor to solve TUT structures. In this thesis, all samples were applied onto gold grids for subsequently structural determination. This figure is taken from (Assaiya et al., 2021).

1.6. Issues and aims

1.6.1 Issues

Let-7 RNA is widely expressed in somatic cells and regulates cellular differentiation and proliferation (Yashiro & Tomita, 2018). Emerging evidence proves the importance of the TUT4(7)/Lin28/let-7 pathway in regulating the biogenesis of let-7, which finally affects the expression of diverse oncogenes. Overall, emerging evidence indicates that TUTases or Lin28 proteins may be targets for inhibition by novel drugs with clinical application.

There is a proposed model of TUT4/7 switching from mono-uridylation to oligo-uridylation states based on the solved structures of TUT4 LIM and TUT7 CM in apo and dsRNA bound forms (Faehnle et al., 2017a; Yamashita et al., 2019). The transition states modeled between mono- and oligo-uridylylation, with or without the presence of Lin28A are shown in Figure 1.23 (Yashiro & Tomita, 2018). But it is still unclear how the LIM of TUT4/7 interacts with the ZF domain of Lin28A and the GGAG motif of pre-let-7. Does this interaction change the way pre-let-7 binds the CM of TUT4/7, compared to the mono-uridylylation state? More structural work needs to be done to reveal the mysteries of the TUT4/7: pre-let-7: Lin28A pathway.

1.6.2 Aims

As detailed already, we now know significantly more than previously concerning the structure-function relationship of the catalytic regions of TUT7 and the Lin28A-interaction module. A better understanding of the mechanism of the TUTases in regulating RNA turnover would be greatly assisted by the structures of full-length enzyme, with or without RNA and Lin28A. This is a strong focus of my DPhil project. I list my strategies and aims as follows.

a) The full-length enzymes are appropriate for cryo-EM structure determination, and this is likely to prove the best route to achieving an atomic model of apo mammalian TUTases.
b) Structure determination by cryo-EM may be assisted by inclusion of pre-let-7 miRNA and Lin28A in complexes – increasing the size of the target and potentially its conformational homogeneity. More importantly, the determination of binary or ternary complex structures will provide significant insights into the mechanical basis of TUTases action on target RNAs. Our RNAs were either synthesized by collaborators in the University of Oxford or by Genscript.
c) In addition, we sought to obtain crystallised forms of sub-complexes providing atomic snapshots of interactions between TUTases, Lin28A and pre-let-7. Achieving this goal will benefit from reagents we have already generated as well as the crystallised TUT7 catalytic module.

We still know relatively little about the recognition and binding of pre-miRNA and other RNA substrates by TUT4/7, and the mechanical basis for the switch under Lin28A from mono- to oligo-uridylation. Structures of ternary complexes will reveal more information with value for both an understanding of fundamental biological mechanisms and for drug discovery programmes targeting the Lin28/Let-7 axis.

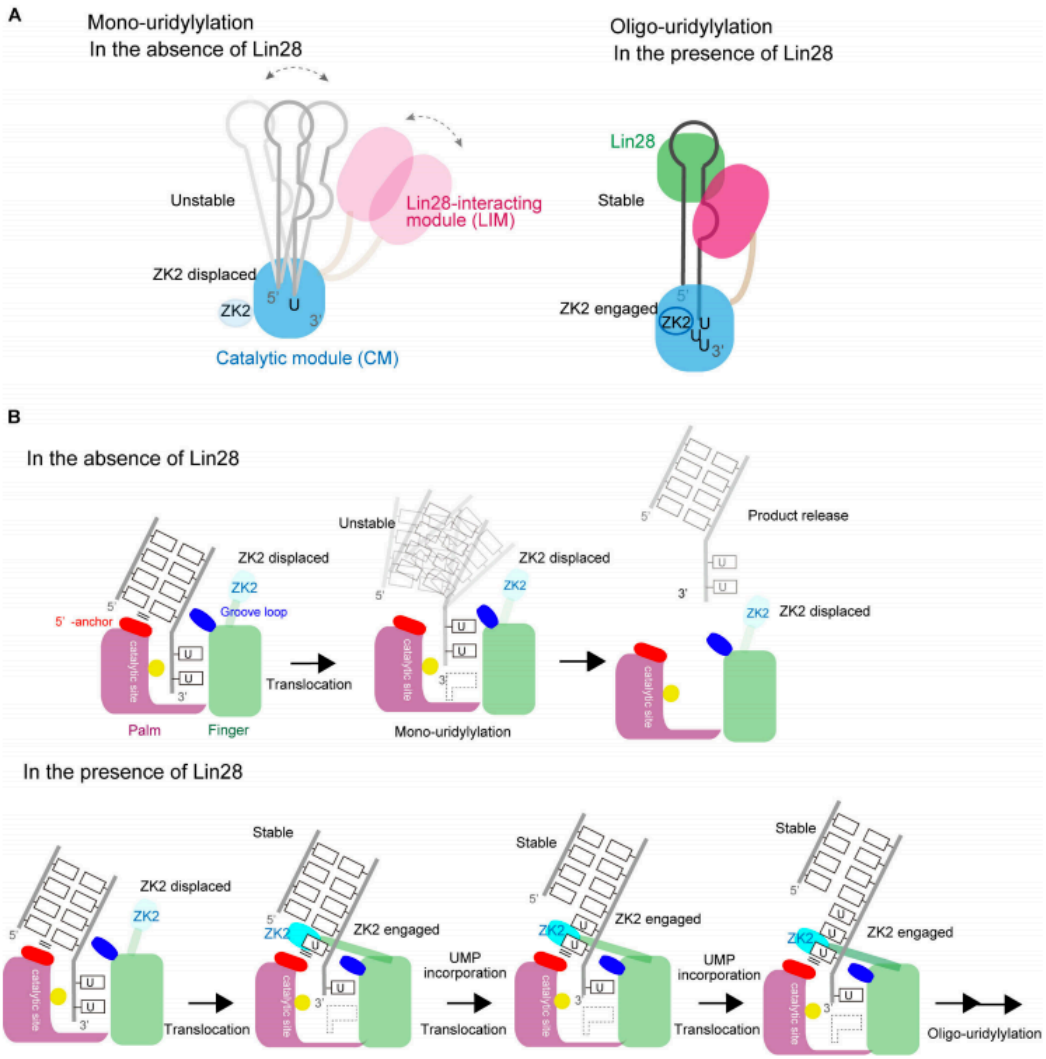


Figure 1.22 Proposed models of Lin28A mediated oligo-uridylation and Lin28A independent mono-uridylation by either TUT4 or TUT7. In mono-uridylation, TUT4(7) forms an unstable complex with pre-let7 miRNA without the participation of Lin28, while with the presence of Lin28, a stable ternary complex consisting of TUT/pre-let-7 miRNA/Lin28 benefits the addition of oligo uridines, in which the ZK2 of TUT7 is thought to aid oligo-uridylation by supporting the extension of the oligo-U tail. This figure is taken from (Yashiro & Tomita, 2018).

1.7. Overview of this thesis

This thesis provides a structural and mechanistic description of human TUT4 and TUT7 in regulating pre-let-7 miRNAs, using both X-ray crystallography and cryo-EM methods.

Charter 1, this chapter, gives a brief review of TUTases involved in the uridylation of pre-let-7, and an introduction to the techniques of X-ray crystallography and cryo-EM.

Charter 2, describes the basic techniques and experimental protocols used in this thesis. The optimized high throughout cloning and expression pipeline benefits the successful purification of a series of TUT4/7 proteins in both full length and truncated forms.

Charter 3, shows structural studies of TUT7 in different forms with or without RNA in the presence or absence of Lin28A. The TUT7_CM was solved by X-ray crystallography at 2.6 Å for subsequent fragment screening and soaking trials (Li, 2020). The high-resolution cryo-EM maps of TUT7 apo, binary complex with pre-let-7 and ternary complex with pre-let-7/Lin28A allow for description of the conformational changes involved in the reaction cycle and confirmed the ternary interactions among diverse components.

Charter 4, introduces the structural exploration of TUT4 with or without pre-let-7 and Lin28A. The high-resolution map of TUT4_miniF ternary complex with pre-let-7 and Lin28A shows a similar overall structure to the TUT7 ternary complex in one of the conformations observed for it. A low-resolution map of the TUT4_F ternary complex indicates additional features for further exploration.

Charter 5, shows a detailed structural analysis of cryo-EM structures of both TUT4 and TUT7 in various states. The ternary interface among TUT4_LIM, pre-let-7 and Lin28A is dissected, to reveal the significance of the RNA GGAG motif in the formation of ternary complex.

Charter 6, summarizes the results for the project overall, including the HTP pipeline, structure studies of TUT7(4) CM, TUT7(4) apo and pre-let-7 bound forms with or without Lin28A, and prospective future TUT7(4) studies.

In conclusion, the work undertaken in this thesis is a deep exploration into the uridylation mechanism regulated by TUT4/7, and the results reveal the detailed conformational changes at atomic level based on five high-resolution cryo-EM maps. Our research provides the structural and mechanistic framework that is recognized as wanting in studies of TENTs, which will prove invaluable in the search for an understanding of underlying mechanism informing cancer cell biology and also inhibitor drug development.

Chapter 2 Methods

2.1. High throughput (HTP) cloning and expression screening

2.1.1 Concept

Ligation Independent Cloning (LIC) is an alternative to restriction enzyme/ligase cloning techniques developed from the 1990s onwards (Aslanidis & de Jong, 1990). The key element in its application is creation of long overhangs with complementary sites between inserted genes and vectors, utilizing the 3' -5' exonuclease activity of T4 DNA polymerase. As described in Figure 2.1, the vectors are digested by BsaI enzyme and then treated by T4 DNA polymerase in buffer containing dGTP. The PCR products are also treated with T4 DNA polymerase instead with the presence of dCTP containing buffer. Digested PCR products and vectors are mixed at a ratio of 3:1 and subsequently transformed. The annealed fragments have 4 nicks, which can be repaired by *E.coli* in vivo (Gradia et al., 2017).

2.1.2 KRX stain

E.coli bacteria are commonly used to produce recombinant protein quickly in cell culture, with excellent features of well- established protocols and relatively low cost. (Zerbs et al., 2014). While the targets addressed in the work described in this thesis might have been expected to benefit from a eukaryotic expression system, the Promega Corporation proprietary strain KRX, a single step competent cell derivative of *E.coli* (Schwarzhans et al., 2017), was used here. It has mutations rendering the bacteria endA-, recA-, ompT- and ompP-, and is marketed as providing advanced features for cloning and protein expression together, the latter of which can be controlled tightly with the change of concentration of its inducer Rhamnose. These features out-compete either BL21 (DE3) or DH5a.

2.1.3 Pipeline

To obtain challenging proteins, various constructs of target genes need to be tested, and vectors containing different solubility tags are chosen. In this thesis, three vectors, pNIC-bio2 (8x His), pHGT-bio (6x His and GST) and pNIC-MBP2 (6x His and MBP) are described as in use for LIC, with the same restriction enzyme site of BsaI in common and the same complementary sequences matching to the PCR products. In this case, one pair of primers for one gene can be applied for all three vectors at the same time. The whole pipeline is described

as follows.

The primers were designed and synthesized first. PCR products were confirmed by agarose gel electrophoresis and then digested with DpnI overnight in a 37 °C oven aiming to remove template vectors. On the next day, DpnI treated PCR products were recovered, and subsequently treated by T4 DNA polymerase. Digested PCR products and vectors were mixed at a ratio of 3:1 and then transformed into KRX competent cells before 1-hour incubation at room temperature. The cells were spread onto LB-Agar plates containing 5 % sucrose and kanamycin after recovery for 5 hours at 37 °C shaker at 200 rpm. After around 18-hours' incubation at 37 °C, colonies were observed. Two colonies were picked per construct into 1.5 ml EP tubes containing 1ml LB medium with 0.2 % Rhamnose. They were placed in a 37 °C shaker at 200 rpm to incubate for 7 hours. 2 µl of each sample were diluted in 10 µl of ddH_2O for SDS-PAGE analysis to determine the levels of protein expression. Colonies with decent amounts of protein expression were selected to perform small scale expression and purification followed by large scale purification. Overall, this provides an extremely highly efficient method, enabling the testing of hundreds of alternative constructs.

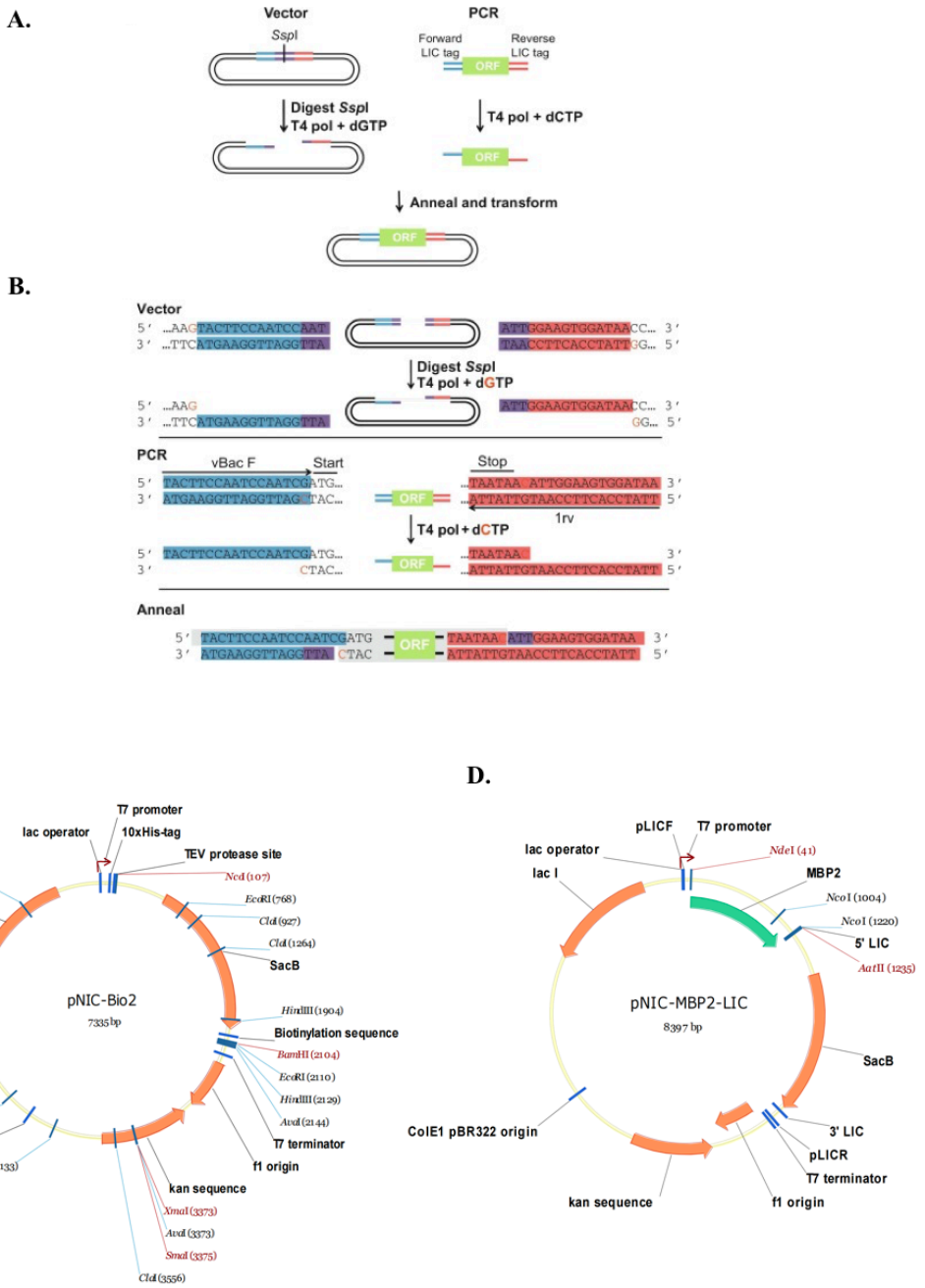
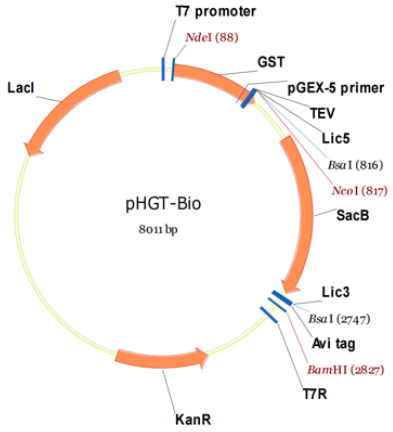


Figure continues on next page

E.



F.

Vectors	N-terminus	C-terminus
pNIC_Bio2	10xHis	Avi tag
pNIC_MBP	MBP and 6xHis	Avi tag
pHGT_Bio	6xHis and GST	Avi tag

Figure 2.1 Schematic diagram of LIC cloning A-B. Overview of LIC cloning into vectors. This figure is taken from (Gradia et al., 2017). C-E. Diagrams of pNIC-bio2, pNIC-MBP2 and pHGT-bio vectors used in this project. F. Terminal information of pNIC-bio2, pNIC-MBP2 and pHGT-bio vectors.

2.2 Polymerase chain reaction (PCR)

2.2.1 PCR

For PCR amplification, primers were designed according to target gene sequences using Snapgene software (from Dotmatics; available at snapgene.com). The PCR program setting is shown in Table 2.1

Ingredients	Volume
2x AmpClone-HiFi	25 µl
Template (plasmid 5 ng/ul)	1 µl
Primer_Forward (10 pmol)	1 µl
Primer_Reverse (10 pmol)	1 µl
ddH_2O	23 µl
Total	50 µl

95 °C	60 sec	
95 °C	10 sec	
62 °C	30 sec	Cycles of 32
72 °C	10 sec/1kb	
72 °C	5 min	
12 °C	10 min	

Table 2.1 PCR required conditions for Ligation Independence cloning (LIC) are shown as table above. These conditions are applied for all constructs preparation in this thesis.

2.2.2 Agarose gel electrophoresis

The PCR products were confirmed by agarose gel electrophoresis. To prepare the Agarose gel, 1 g of agarose powder was weighed and then put into a conical flask (250 ml) containing 100 ml 1x TBE buffer. The above flask was heated in a microwave for 2 minutes to dissolve the agarose completely. The solution cooled slowly until around 60 °C and then 10000x dye (SYBR Gold nucleic acid gel stain, Thermal Fisher) was added to 1x. The above solution was poured into the mold with comb insertion until the gel solidified completely. Then, 2 µl PCR products were taken out, mixed with 0.5 µl 5x DNA loading buffer followed by adding into agarose gel holes individually. Agarose gel electrophoresis was performed for 40 minutes at 120 V. When electrophoresis was completed, the gel was imaged, and the presence of target gene fragments was confirmed using Bio-Rad Geldoc and software.

2.2.3 PCR product recovery steps

PCR products were transferred into 1.5 ml EP tubes individually followed by adding 5x volume of CP binding buffer and mixing thoroughly. The mixture was subsequently transferred into HiBind® DNA Mini tubes respectively, which were centrifuged at 13,000 g for 1 minute. The filtrate in collection tubes were discarded, and then 700 µl DNA Washing Buffer (SPW Washing Buffer) was added into Mini tubes following by centrifugating at 13,000 g for 1 minute. The washing step was repeated twice, and the filtrate in the collection tubes was discarded. Finally, the HiBind® DNA Mini tubes were put into the empty collection tubes and centrifuged at 13,000 g for 2 minutes to remove the remaining washing buffer. After washing thoroughly, HiBind® DNA Mini tubes were put into new EP tubes, 50 µl sterilized deionized water was added into HiBind® DNA Mini tubes for incubating for 1 minute, and then the above tubes were centrifuged at 13,000 g for 1 minute. The eluted solution in the EP tubes contained PCR products fragments, which was verified by agarose gel electrophoresis and quantified by nanodrop.

If PCR products were impure with more than one band, the PCR products needed to be recovered from the agarose gel. The recovery steps were as follows. The PCR products obtained above were separated into different bands according to various molecular sizes by

agarose gel electrophoresis. The agarose gel containing the target fragment bands was cut under UV, and bands were put into 1.5 ml EP tubes and weighed. Subsequently, 1 ml of binding buffer (XP5) for 1 g agarose gel was added. Then, the above EP tubes were put into a water bath at 42 °C until the gel was completely dissolved. The gel solution was subsequently transferred into HiBind® DNA Mini tubes, which were centrifuged at 13,000 g for 1 minute. The following washing and elution steps were the same as the applied above in PCR products recovery.

2.3 Plasmid extraction and translation

2.3.1 Plasmid extraction

10 µl of glycerol stock frozen in -80 C° freezer was taken out with a pipette tip, which was placed in a 50 ml tube containing 10 ml LB medium (with corresponding antibiotics already added) followed by culturing in a 37°C shaker at 200 rpm overnight. On the next day, the above bacteria cells were centrifuged to remove the supernatant LB medium by 4,000 g for 5 minutes. The pellets were dried and then 250 µl HiBind® solution-I was added to resuspend the cell pellets followed by adding 250 µl HiBind® solution-II, which was inverted 4-5 times until the solution became clarified. Then 350 µl HiBind® solution-III was added to the above tubes and mixed by inverting them 5-10 times, and subsequently the tubes were centrifuged at 13,000 g for 10 minutes. The supernatant was taken out from the above EP tubes and transferred into HiBind® DNA Mini tubes followed by centrifuging at 13,000 g for 1 minute. The filtrate in the collection tubes was discarded. 700 µl DNA Washing Buffer (SPW Washing Buffer) was added into Mini tubes followed by centrifugation at 13,000 g for 1 minute. The washing step was repeated twice, and the filtrate in the collection tubes was discarded. Finally, the HiBind® DNA Mini tubes were put into the empty collection tubes and centrifuged at 13,000 g for 2 minutes to remove the remaining washing buffer. After washing thoroughly, HiBind® DNA Mini tubes were put into new EP tubes, 50 µl sterilized deionized water was added into HiBind® DNA Mini tubes for incubating for 1 minute, and then the above tubes were centrifuged at 13,000 g for 1 minute. The eluted solution in the EP tubes contained PCR products fragments, which was verified by agarose gel electrophoresis and quantified by nanodrop.

2.3.2 KRX Transfection

KRX competent cells frozen in a -80 °C freezer were taken out and placed on ice until they thawed completely. Then 3 µl of the prepared plasmids (the concentration of the plasmids should be higher than 5 ng/µl) or the LIC cloning reaction products were added into 100 µl of competent cells. The cells were mixed well and incubated on ice for 5 minutes. After 5 minutes of incubation on ice, the above EP tubes were placed in a 42 °C water bath and heated for 15 seconds. Subsequently, the above EP tubes were taken out and placed on ice

for 2 minutes. After the ice treatment, 1 ml of sterilized LB medium was added to the above EP tubes individually, which were subsequently placed in a 37 °C shaker to culture for about 4 hours. After 4 hours, the above EP tubes were centrifuged to remove most of the supernatant by centrifugation at 10,000 g, 25 °C for 1 minute for removing the supernatant and about 100 μl LB solution was left, which was used to gently resuspend cell pellets with a pipette tip. Finally, cells were spread on Agar-plates (with corresponding antibiotics already added), and then the plates were placed in a 37 °C oven overnight. On the next day, single colonies had grown on the plates, and could be picked for the next step.

2.3.3 Small scale expression screening

The constructed plasmids with correct sequences were transferred into RS or BL21 expression strains. On the next day, single colonies were picked and then placed in 15 ml tubes containing 4 ml of LB medium individually (with corresponding antibiotics already added). After overnight culture, 1 ml cells were taken out for seed preservation. Then, IPTG was added to the remaining 3 ml bacterial cells with a final concentration of 0.5 mM, which was cultured in a 37 °C shaker for a 3-hour induction. After induction, 1 ml of bacterial cells were centrifuged followed by removing the supernatant LB medium, and then 100 μl of protein loading buffer was added. The cell pellets were resuspended in tubes and placed in a 100°C metal bath to heat for about 10 minutes. Finally, tubes were centrifuged at 13,000 g, 25°C for 1 minute, 10 μl supernatant was taken out, and the expression was examined by SDS-PAGE.

2.3.4 KRX competent cells preparation

Reagents required: LB medium (sterilized), 100 mM $CaCl_2$ solution (sterilized and filtered), 50 % glycerol solution (sterilized). The procedures were as follows:

10 μl of glycerol stock of KRX from the -80 °C freezer was taken out using a pipette tip, which subsequently was placed in a 50 ml tube containing 20 ml LB medium (no antibiotics), followed by culturing in a 37 °C shaker at 200 rpm overnight. On the next day, 1 ml of the above bacterial cells were taken out and put it in a conical flask (250 ml) containing 100 ml fresh LB medium. The above flask was put in a 37 °C shaker at 200 rpm for culturing until the

OD value of the bacterial cells reached about 0.4. This process took around 2 hours. Then, the 100 ml cells were split into two 50 ml tubes and centrifuged at 4,000 g, 4 °C for 10 minutes. The cell pellets were dried and 50 ml of 100 mM $CaCl_2$ solution were added to the tubes. The cells were resuspended gently and placed on ice followed by a 4-hour incubation. After that, the supernatant was removed by centrifugation at 4,000 g, 4°C for 10 minutes followed by resuspending the cells with 10 ml of 100 mM $CaCl_2$ solution, and then 50 % glycerol solution was added with the final glycerol concentration of 10 % in the solution, which was placed on ice overnight. Finally, the competent cells were aliquoted into sterilized 1.5 ml EP tubes, 100 µl per tube. Then all tubes were flash-frozen with liquid nitrogen and stored in the -80 °C freezer.

2.4 Protein purification

2.4.1 Cell culture

10 µl of glycerol stock of bacteria expressing the protein of interest from the -80 °C freezer was taken out using a pipette tip, which subsequently was placed in a 50 ml tube containing 20 ml LB medium (with corresponding antibiotics already added), followed by culturing in a 37 °C shaker at 200 rpm overnight. The next day, tubes were taken out and added into fresh LB medium with corresponding antibiotics at a ratio of 1:50 into a large flask (2 L). This was placed in a 37 °C shaker at 200 rpm. When the OD_{600} value of the cells reached 0.8-1, Rhamnose was added with a final concentration of 0.1 % for a 36-hour induction at 20 °C. After that, the bacterial cells were harvested by centrifugation at 4,000 g 25 °C for 10 mins to remove the supernatant LB medium. Then the cell pellets were dried, and they can be immediately purified or frozen in the -80 °C freezer.

2.4.2 Ion exchange chromatography

To further improve the purity of proteins of interest, ion exchange chromatography is widely used to remove contaminated proteins. For the work described in this thesis a heparin column was selected due to its excellent features of separating diverse proteins with different PIs, and especially nucleic acids from target proteins. Since TUTases and Lin28A have several zinc fingers/knuckles and/or RNA binding domains, the heparin column is a good choice to remove contaminating nucleic acids and not the target proteins.

The procedures were as follows. Fi" st, ' he column was washed with de-gassed and filtered water before equilibration with buffer A: 20 mM HEPEs, pH 7.5 and 50 mM NaCl, flow rate 1 ml/min. Then, proteins were diluted with 5x sample volumes using buffer A and the diluted sample was filtered to remove precipitants before loading onto the heparin column. After loading, the column was washed with 10x column volumes, and subsequently gradient elution was performed with the setting at 1 ml/min flow rate for 25 minutes from 0 % to 100 % buffer B: 20 mM HEPEs, pH 7.5 and 1000 mM NaCl. The peaks of the UV trace at wavelength of 280 nM were collected in different fractions, which were examined by SDS-PAGE. All protein samples as determined by staining with Coomassie blue in this thesis are under reducing

conditions.

2.4.3 Size exclusion chromatography

Samples for crystallography and cryo-EM require proteins with high purity and homogeneity. Size exclusion chromatography can be utilized to separate target proteins from contaminants and aggregates based on different molecular weights. The procedures were as follows.

First, the column was washed with de-gassed and filtered water, which subsequently was equilibrated with buffer C: 20 mM HEPES, pH 7.5 and 100 mM NaCl. The speed and pressure limitation depend on different types of columns. Then the SEC program was performed with the setting at 0.4 ml/min flow rate for 60 minutes after injecting the sample, which had been already concentrated and centrifuged. The peaks of the UV trace were collected in different fractions, which were examined by SDS-PAGE.

2.5. Crystallography

2.5.1 Principles of crystallography

The growth of crystals is closely related to the solubility of proteins, and the solubility of proteins is affected by many factors including metal ions, salts, pH with diverse buffers, and so on. Most proteins are soluble under physiological conditions, but with increasing concentrations of precipitants in their buffers, proteins become less soluble, causing them to crystallise or precipitate, a phenomenon termed "salting out". However, the salting out of disordered proteins will become precipitants, and only when the conditions in the reservoirs are suitable, and the protein is sufficient ordered, can the precipitants become ordered crystals (Koopmann et al., 2012). Screening and optimizing conditions are a time-consuming and labor-intensive process since many factors affect the stability of proteins. There are several methods to obtain crystals, such as the classic batch crystallization, liquid-liquid diffusion, and gas-liquid diffusion (Ilari & Savino, 2017). In this project, the gas-liquid diffusion method is selected for the crystallization screening and optimization of the human TUT4/TUT7 catalytic module (CM). This is the most common method in crystallization experiments. The mixed droplets in wells contain different types of precipitants, in which the precipitant concentrations will gradually increase up to the concentrations in the reservoir pool.

As the concentration of precipitants continue to increase, the solubility of the proteins will continue to decrease until the protein "precipitates". If the conditions are appropriate, the ordered accumulation of proteins in the process of precipitation will become crystals, otherwise they will be disordered precipitants. Gas-liquid diffusion can be carried out in the form of hanging drops or sitting drops. The hanging drop device (Figure 2.3A) is designed to have a drop (approximately 2ul) mixed with reservoir buffer at a ratio of 1:1 placed on an inverted coverslip, suspended over the pool wells. The coverslip and the pool fluid well are sealed with Vaseline. The sitting drop crystallization device (Figure 2.3B) is designed to accommodate the mixed droplets on the shelf next to the well. Both methods require an air-tight environment so that the droplets will be equilibrated against the well reservoir over time, and not evaporate (Ilari & Savino, 2017; Van Driessche et al., 2018)

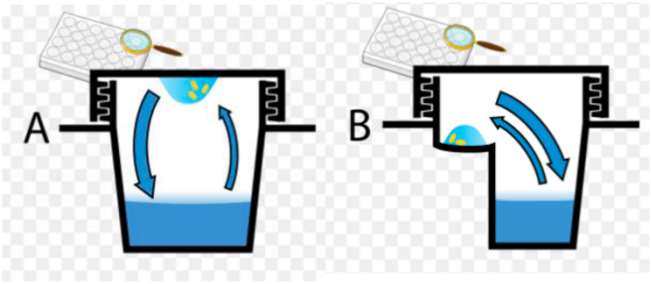


Figure 2.2 Cartoon schematic diagram of vapor diffusion. A. Hanging method. B. Sitting drop method. The latter is chosen for TUT7 CM crystallization due to its excellent features of high throughput screening and low demand of proteins. This figure is taken from (Ilari & Savino, 2017).

2.5.2 Coarse screening

The purified and frozen proteins were removed from the -80 °C freezer where they had been stored and placed on ice to thaw completely. The samples were centrifuged to remove any precipitants before crystallography trials. Different commercial coarse screen plates from Hampton research company were used. Crystallisation plates from TPP Labtech or SWISSCI companies were used in this project and crystallization set up used a Mosquito robotic liquid dispenser. Protein and reservoir buffers were mixed at a ratio of 0.1 µl: 0.1 µl, and then crystal plates were sealed with high-definition transparent tape after the program had completed prior to storage in a Rock Maker imager at 4 °C or 20 °C. Crystal plates were imaged every other day in the first week, then once a week after that.

2.5.3 Condition optimization

Diffraction of crystals was initially assessed using beamlines at the Diamond Light Source, via remote data collection protocols. High-resolution crystals were repeated, while poor-resolution crystals would be optimized on the basis of their precipitant conditions. The strategy was to slightly change the conditions around those originally identified via coarse screening. Normally, the crystalised conditions consist of diverse concentration of salt ions, a range of pH in different buffers, besides a certain concentration of precipitant such as PEG, and so on. Because the pH of the reservoir buffer and precipitants, like PEG and salts used, affect the growth of the crystal significantly, the pH gradient and precipitant concentration gradient optimization was attempted first. A 96-Well plate can be used to optimize precipitant concentration and pH range at the same time in horizontal and vertical directions. Similarly, other components in the crystal growth conditions could be optimized by this method.

2.5.4 Crystal diffraction and data collection

Before picking up crystals, materials need to be prepared, such as different sizes of loop rings, cryo-protectants, pucks and other related materials. Generally, the cryo-protectant solution was to add glycerol to the original reservoir solution at a varied concentration around 20%, but if the glycerol did not work very well, other cryo-protectant reagents could be tried, like

Ethylene Glycol, MPD, PEG400, etc. Increasing the concentration of the precipitant on the basis of the original reservoir solution was another strategy to protect the crystals. Frozen crystals were sent to and used on beamlines at the Diamond Light Source, via remote data collection protocols.

2.5.5 Data processing and model building

The data collected at Diamond Light Source were automatically processed by ISPyB via combined software such as PHASER (McCoy et al., 2007), XIA2 (Gildea et al., 2022; Winter et al., 2013) and XDS (Brown et al., 2007; Kabsch, 2010). Preliminary results were exported indicating the resolution of the crystals' diffraction. The collected two-dimensional diffraction images correspond to different crystal angles, requiring the use of mathematical Fourier transform methods to convert 2D images into a 3D electron density map for later 3D model building. A critical issue is to determine the phase of the diffracted beams giving rise to spots, and details of phasing and model building are provided in the relevant results section below, Section 3.2.3.

2.6 Negative staining

2.6.1 Grid preparation and data collection

Negative staining is an easy and quick tool to visualize proteins via electron microscopy. The proteins are stained by a heavy metal solution such as uranyl acetate. The substantial differences of densities between the stain and proteins create obvious contrast, enabling the visualization of proteins. Appropriately purified proteins were first imaged using negative-stain EM to confirm their homogeneity. 3.5 µl samples of different concentrations were loaded to glow-discharged Formvar/ carbon-coated 300 mesh copper grids (Agar Scientific). Excess sample was blotted away with filter paper after incubation for 1 min. Then grids were washed with deionized water twice followed by staining with 0.75 % uranyl acetate for 1 min. Excess stain was removed and then grids were imaged on the FEI T12 electron microscope. Images were taken at the following conditions: voltage 120kV, 63,000x magnification with a maximal electron dose of 15 electrons per $Å^2$, producing a pixel size of 3 Å/pixel using the accompanying CCD camera.

2.6.2 Data processing

Raw negative stained images were processed using both EMAN2 and Spider software. Generally, the raw images were imported into EMAN2 (Tang et al., 2007), to pick particles manually or automatically. The coordinates of all particles were extracted and exported, and subsequently were processed by Spider (Shaikh et al., 2008) to generate 2D classification images and reconstruct 3D models.

2.7. Cryo-EM

2.7.1 Sample grids preparation and data collection

The qualified samples confirmed by negative-staining EM went on to be used for cryo-EM trials. Generally, a 3.5 μl aliquot of sample was applied to commercial grids (Au C-flat, 2/1, 200 mesh) and before that the grids were glow-discharged (in a Harrick Plasma system) for 30 s at middle level plasm after 1-min evacuation. The parameters for blotting were 4 seconds with -5 force, at 4 °C and 100 % humidity using a Vitrobot (FEI). Cryo-EM movies were collected using FEI Tecnai™ Glacios transmission electron microscopes (TEMs) with an energy filter (GIF Quantum, Gatan) and a direct electron detector (K2 Summit, Gatan), and further via an FEI Titan Krios, a 300 keV ultra-high-performance, automated, three condenser transmission electron microscope with a K3 summit camera.

2.7.2 Data processing

Raw images collected from Glacios or Krios microscopes were imported into cryoSPARC (Punjani et al., 2017; Scheres, 2012). After motion correction and the finding of the contrast transfer function (CTF), qualified images were curated manually to proceed further. Small subsets of images were first chosen for manual particle picking aiming to generate templates. Subsequently, the templated picked particles were extracted to calculate 2D class averages, and particles falling into good 2D classes were chosen to reconstruct 3D models.

2.8. Site-directed mutagenesis

Figure 2.4 shows the site-directed mutagenesis strategy used in this thesis (Adachi & Fukuhara, 2012).

This method includes primers design, two rounds of PCR, PCR product purification, cloning and sequencing. There are four primers to be designed for one mutagenesis. The forward_1 refers to the starting sequence of the target gene (around 20 bp), while the reverse_1 comprises the ending sequence of the target gene (about 20 bp). The forward_2 and reverse_2 are complementary, and the sites of interest need be replaced by target sequences. Two rounds of PCRs were performed aiming to obtain final products, and the program setting for PCR is shown as Table 2.1. If the products are impure, they need to be recovered from the resolved bands out of the agarose gel electrophoresis. The subsequent procedures are similar to those given in section 2.1 after the right sequences are confirmed.

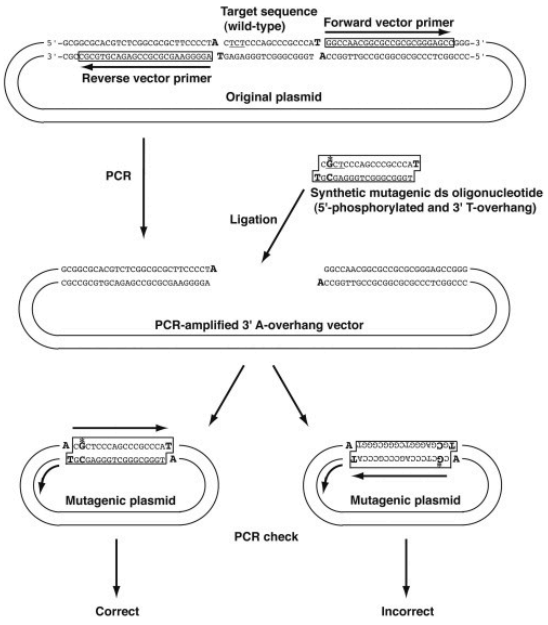


Figure 2.3 Schematic diagram of site-directed mutagenesis based on the TA strategy (The bases, T and A on PCR-amplified vector sticky ends provides direct and efficient site-directed mutagenesis). This figure is adapted from (Adachi & Fukuhara, 2012).

2.9. Uridylation assay

All ingredients listed in Table 2.3 were mixed with corresponding amounts in EP tubes TUTases were added as the last component, which initialized the start of the reaction. Then the above EP tubes were placed in a metal bath and the reaction temperature was set up at 37 °C. After incubating for 0, 5, 10, 20 minutes respectively, 10 µl of TF termination solution was added to quench the reaction. Subsequently, the temperature of the metal bath was adjusted to 95 °C, in which the nucleic acid chains in the tube were fully denatured, marking the finishing of the assays. Finally, all samples were examined by performing Urea-denaturing acrylamide gel electrophoresis and imaged by iBright 5000 (Thermo Fisher).

Ingredient	Volume
TUT4/7	1 µl (0.5 µM)
pre-let-7g-3'CY	1 µl (5 µM)
10x buffer	1 µl
ddH_2O	7 µl
Terminal buffer	10 µl

Table 2.2 Reaction system of uridylation assay. (10x buffer: 200 mM HEPES, pH 7.8, 1000 mM NaCl, 50 mM $MgCl_2$, 0.5 µM $ZnCl_2$, 50% Glycerol, 25 mM DTT, 2U RNAase inhibitor (ThermoFlsher). TF buffer: 1% SDS and 0.5 mM EDTA, pH 8.0).

Chapter 3 Characterization of TUT7

3.1. Introduction

TUT7 participates in diverse biological pathways by mono- or oligo-uridylating target RNAs including pre-let-7 miRNAs, histone mRNAs and piRNAs (Heo et al., 2009; Holmquist et al., 2022; Lim et al., 2014; Zhao et al., 2022). Previous structural studies by Leemor Joshua-Tor' s group has presented the atomic details of the TUT7 Catalytic Module (CM) in complex with various substrates including UTP, dsRNA, oligo-U2 and oligo-U5, indicating conformational changes from the transient mono-uridylation state to an oligo-uridylation state (Faehnle et al., 2017a). However, how the Lin28A interaction module (LIM) of TUT7 interacts with the zinc knuckles region of Lin28A and pre-let-7 is still unclear. Therefore, I started exploring how TUT7 catalyzes pre-let-7 with or without the presence of Lin28A by cryo-EM. Meanwhile, given the significance of TUT7 as a tumor suppressor, I used X-ray crystallography aiming to obtain well-diffracting crystals in the presence of potential inhibitors provided by a collaborator for better understanding of their inhibition mechanisms.

As a result, by a high throughput (HTP) cloning and expression screening pipeline, high quality TUT7_CM1 and _CM2 truncated proteins were obtained, and the crystal structure of TUT7_CM2 was resolved at a resolution of 2.6 Å. It was also shown that the TUT7_CM2 crystals are reproducible. These crystals have been evaluated as a potential substrate for determination of inhibitor complexes as part of a drug discovery programme.

Subsequently, I set out to explore how the full length TUT7 works on pre-let-7 in the presence or absence of Lin28A. By HTP cloning and expression screening trials and optimization of the purification protocol, the challenging full length TUT7 was successfully purified and showed its specific enzymatic activity. Negative staining results showed that it has good quality considering its homogeneity, and an obvious linker connects the LIM and CM indicating the dynamic nature of this enzyme.

I further determined the cryo-EM structures of TUT7 apo, its binary complex with pre-let7g, and ternary complex with pre-let7g and Lin28A at resolutions ranging from 3.6 Å to 4 Å. These results reveal the conformational transitions from mono- to oligo-uridylation states.

The flexible parts of TUT7 were missing in the apo structure, although the NTD1 and NTD2 of the LIM and CM respectively were resolved. In the binary complex, pre-let-7 was confirmed to bind the catalytic pocket of TUT7_CM without the participation of the LIM zinc finger. However, two distinct ternary complex structures were reconstructed. Complex 1 (conformation-I) had the CM missing (disordered), and presents a pre-oligo-uridylation let-7 state in which the pre-miRNA has been captured, while Complex 2 (conformation-II) revealed all of the integral components in a fully engaged complex, indicating an activated state but remaining pre-oligo-uridylation as a result of the presence of $CaCl_2$ and UTP α S . The zinc finger of TUT7 LIM was visualized in both structures, engaged with the pre-let-7 in the presence of two zinc knuckles of Lin28A, an interaction missing in the TUT7 binary complex structure without Lin28A.

This chapter provides the most advanced and systematic structural characterization of human TUT7 targeting pre-let-7 miRNA of which I am aware, via combined techniques of X-ray crystallography and cryo-EM, which presents a good model to explain how TUTases work on pre-let-7 miRNA.

3.2. Methods and Results

3.2.1. Design of TUT7 in both full length and truncated forms

For both X-ray crystallography and cryo-EM, proteins with high quality are necessary. TUT7 consists of multiple domains including two nucleotidyltransferase domains (inactive NTD1 and active NTD2), connected by a flexible region, three CCHC zinc knuckles motifs scattered at the C-terminus, and a C2H2 zinc finger, which is located in the LIM module (Figure 3.1A). Due to the previous failures in expressing full-length TUT7 and difficulties of obtaining decent truncated forms from either insect cells or *E. coli*, I initially attempted to express the TUT7 in HEK293 by the transient transfection-secretion method (Aricescu et al., 2006), but I failed to detect the expression of TUT7 by western blot. Aiming to obtain appropriately purified TUT7 protein, I used an HTP cloning and expression screening method based on *E.coli* strain KRX, as outlined in Chapter 2. Before that I performed a secondary structure, prediction and identified putative disordered regions of TUT7 using the online server PrDOS (Protein DisOrder prediction System) (Figure 3.1B). Based on disorder probability, I designed a large number of primers, which I used for HTP cloning and expression screening of TUT7 in diverse forms (Table 3.1)

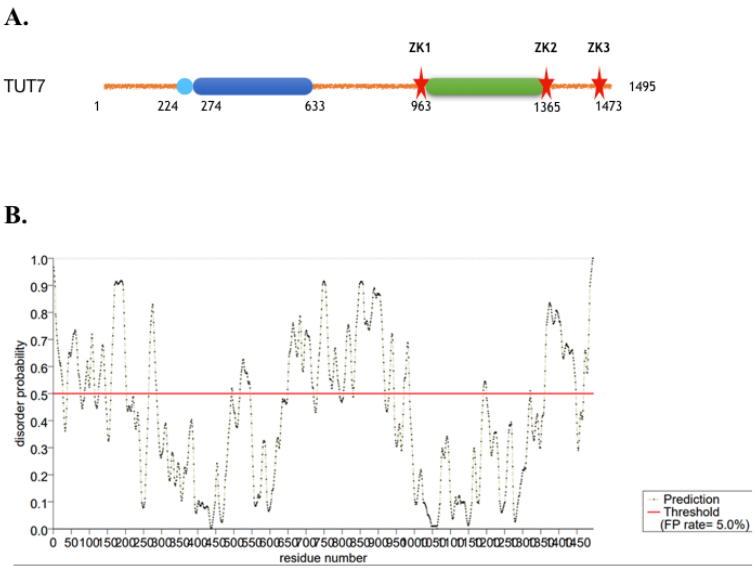


Figure 3.1 Overview of human TUT7 A. Schematic representation of domains of human TUT7, consisting of two nucleotidyltransferase domains (NTD1(blue) and NTD2(green)), zinc finger (teal) and multiple zinc knuckle motifs (red star). B. Prediction of natively disordered regions of human TUT7, analyzed by PrDOS, showing the potentially flexible regions including two terminus and internal domain (Ishida & Kinoshita, 2007).

Primers	Sequence
TUT7_1F	TACTTCCAATCCATG GGCGATACCGCCAAGC
TUT7_194F	TACTTCCAATCCATG GAGCAGGACGGCGATCTGG
TUT7_963F	TACTTCCAATCCATG GTCGTGTGTAGCCTGTGTA
TUT7_973F	TACTTCCAATCCATG CACCTGAAAAAGGACTGCC
TUT7-983F	TACTTCCAATCCATG AAGAGGATCCAGCTCGAAC
TUT7_633R	TATCCACCTTTACTGCT GTGAGGCAGGGCAAAGTAC
TUT7_643R	TATCCACCTTTACTGCT AGGCTTCAGCAGGCTGCTC
TUT7_1336R	TATCCACCTTTACTGCT GGTCAGCACGTCGGGGTCA
TUT7_1365R	TATCCACCTTTACTGCT TTTCCGCCGCATGGGGCAA
TUT7_1494R	TATCCACCTTTACTGCT GCTCTCTTGCTGTGTCCGC

Table 3.1 Primer sequences of human TUT7 used for HTP cloning.

3.2.2 Purification of TUT7_CMs

The gene of human TUT7 was synthesized to order by Thermo Fisher Scientific, and the sequences of both amino acids and the DNA are provided in the Appendix. The HTP cloning and expression screening of TUT7 CM constructs were carried out as follows.

Eight TUT7_CM truncated proteins, labelled from TUT7_CM1 to TUT7_CM8 individually, were sub-cloned into pHGT-bio vector with a TEV cleavable 6x His tag at the N-terminus (vector, a gift from SGC (now CMD), Oxford). The protein expression levels were first confirmed by SDS-PAGE analysis of the whole cell lysate after cloning and sequencing (Figure 3.2A-B). Then, they were expressed in KRX stain in TB medium containing kanamycin antibiotics at 37 °C for 8 hours followed by adding inducer Rhamnose at a final concentration of 0.1 % and culturing for another 36 hours at 22 °C. They were subsequently purified on a small scale. The Tris buffer-Ni used for Ni-NTA purification contained 20 mM Tris-base (pH 8.0), 500 mM NaCl, 5 mM $MgCl^2$, 50 µM $ZnCl^2$ and 5% Glycerol. Bacteria cells were harvested by centrifugation at 5,000g, 25 °C for 10 mins. The cell pellets were resuspended using Tris buffer-Ni at the corresponding volume (5 ml buffer for 1 g pellets), plus adding additional lysozyme, homemade benzonase and imidazole at a final concentration of 1 mg/ml, 10 µg /ml, 20 mM respectively. The lysate was sonicated in 5s-on-4s-off cycles for a total of 30 minutes and then followed by centrifugation at 18,000 g, 4 °C for 40 mins. The supernatant was poured into Ni-NTA columns (Sigma), which had been pre-equilibrated with Tris buffer-Ni containing 20 mM imidazole. Subsequently the columns were washed with 2x column volumes of Tris buffer-Ni containing 20 mM imidazole and followed by elution with 2.5 ml Tris buffer-Ni containing 300 mM imidazole, and the buffer of eluates were immediately exchanged using a PD10 column (Sigma) aiming to remove the Imidazole. Then the TEV proteases were added into protein solutions at a molar ratio of 1:20 and subsequently incubated at 4 °C overnight. The next day, the protein solutions were poured into fresh Ni-NTA columns individually which were pre-equilibrated with Tris buffer-Ni aiming to remove contaminated proteins, cleaved GST tag and TEV proteases. The flow-through fractions containing target proteins were examined by SDS-PAGE and seven TUT7_CMs were of good yield and purity (Figure 3.2C).

Then they were concentrated individually to perform size exclusion chromatography and three TUT7_CMs were sufficient with a single peak on size exclusion chromatography and a single band as determined by staining with Coomassie blue. They were candidates for the subsequent crystallography trials (Figure 3.2D).

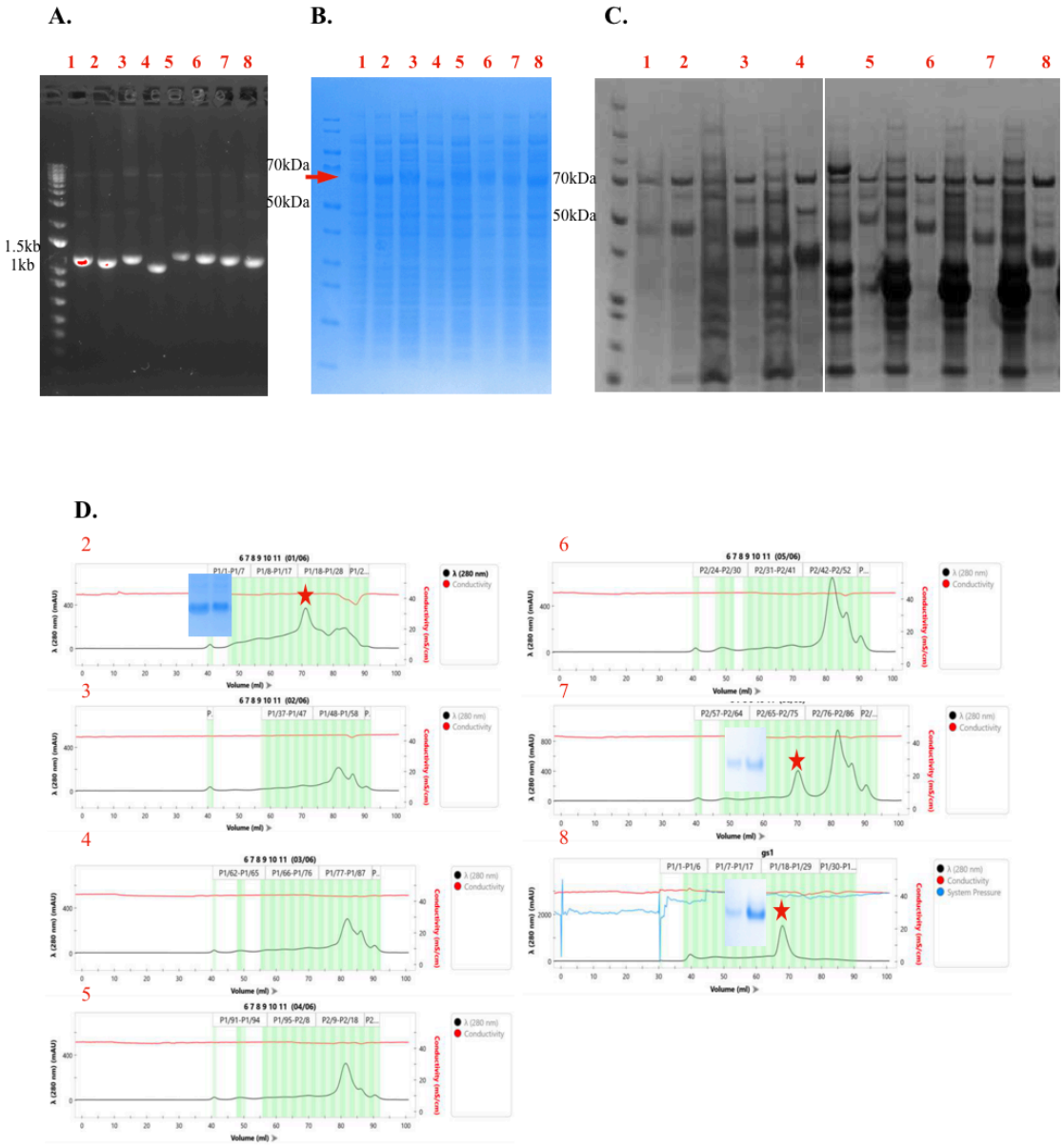


Figure 3.2 HTP cloning, expression and purification of TUT7_CMs A. Agarose gel electrophoresis analysis of PCR products of TUT7_CMs. Lanes 1-8 present TUT7_CM1(963-1336), TUT7_CM2(973-1336), TUT7_CM3(983-1336), TUT7_CM4(993-1336) TUT7_CM5(963-1365), TUT7_CM6(973-1365), TUT7_CM7(983-1365) and TUT7_CM8(1004-1365) respectively, same below. B. SDS-PAGE analysis of expression of TUT7_CMs. C. SDS-PAGE analysis of recombinant TUT7_CMs eluted from IMAC using Ni-NTA columns. D. Size exclusion chromatography profiles of TUT7_CMs.

3.2.3 Crystallography of TUT7_CM

The TUT7_CM1 was purified on a large scale (6 Liters of TB medium) using the same protocol referred to in section 3.2.2. 2 mg TUT7_CM1 from 6-liter TB medium with 95% purity were obtained via batch Ni-NTA affinity chromatography and size exclusion chromatography on Superdex 200 Increase 10/300 (GE) using HEPES buffer-GF1 containing 20 mM HEPES (pH 7.8) and 150 mM NaCl (Figure 3.3A-B). The homogeneous peak-fractions were pooled together and concentrated to 28 mg/ml for subsequent crystallization trials. Using a Mosquito nanofluidics machine and the sitting drop method, six commercial crystallization plates were selected for coarse screens. All plates were then monitored by recording images on the Rock Maker imager at 20 °C. TUT7_CM1 crystalised on the third day in various conditions. Then two conditions (condition1 with 0.1 M NaCl 0.1 M Bis-Tris pH 6.5, 1.5 M ammonium sulfate, overall pH 6.9. Condition2 containing: 0.1 M Bis-Tris pH 6, 0.4 M ammonium sulfate) were optimized via additives based on the initial reservoir buffer conditions (Figure 3.3C-D). Larger crystals appeared on the third day and were subsequently harvested at the seventh day to perform diffraction experiments at the Diamond Light Source. The best resolution of the crystals from both coarse screens and additive optimization was 4 Å processed via the multiple auto-processing pipeline ISPYB at Diamond.

To obtain high-resolution diffracting crystals, I then purified TUT7_CM2 on a large scale (6 Liters TB medium) using the same protocol as 3.2.2; TUT7_CM2 is a shortened version compared to TUT7_CM1 in its C-terminal region. 2 mg TUT7_CM2 from 6-liter TB medium with 95% purity were obtained and confirmed by size exclusion chromatography and SDS-PAGE analysis (Figure 3.4A-B). The concentration used for crystallization trials was 20 mg/ml. Six commercial crystallization condition plates from the Hampton research company were selected for coarse screens using the sitting drop method. One condition (2.8 M sodium acetate pH 7.0) with large and regular crystals was chosen to optimize its pH and salt concentrations (Figure 3.3C). Both coarse screened and optimized crystals were harvested for subsequent diffraction trials. In the optimized condition, the crystals diffracted to better than 2.6 Å consistently. The auto-processing diffraction data from this condition were exported for subsequent structure determination. The data were phased using the TUT7_CM apo structure

(PDB:5W0B) as a search model via Phenix (Adams et al., 2002). Refinement was performed using Coot (Emsley & Cowtan, 2004; Goddard et al., 2005; Lam & Siu, 2017) and Phenix (Adams et al., 2002) and the structure analyzed by Pymol (Bramucci et al., 2012). The data collection and refinement statistics for this TUT7_CM structure, refined to 2.6 Å resolution, are reported in the Appendix.

.

These crystals represent important starting points for studies of small molecule inhibition mechanisms and structure-based drug design and refinement. It is hoped that these crystals can be soaked with the inhibitors supplied by a commercial company (Figure 3.3D-E).

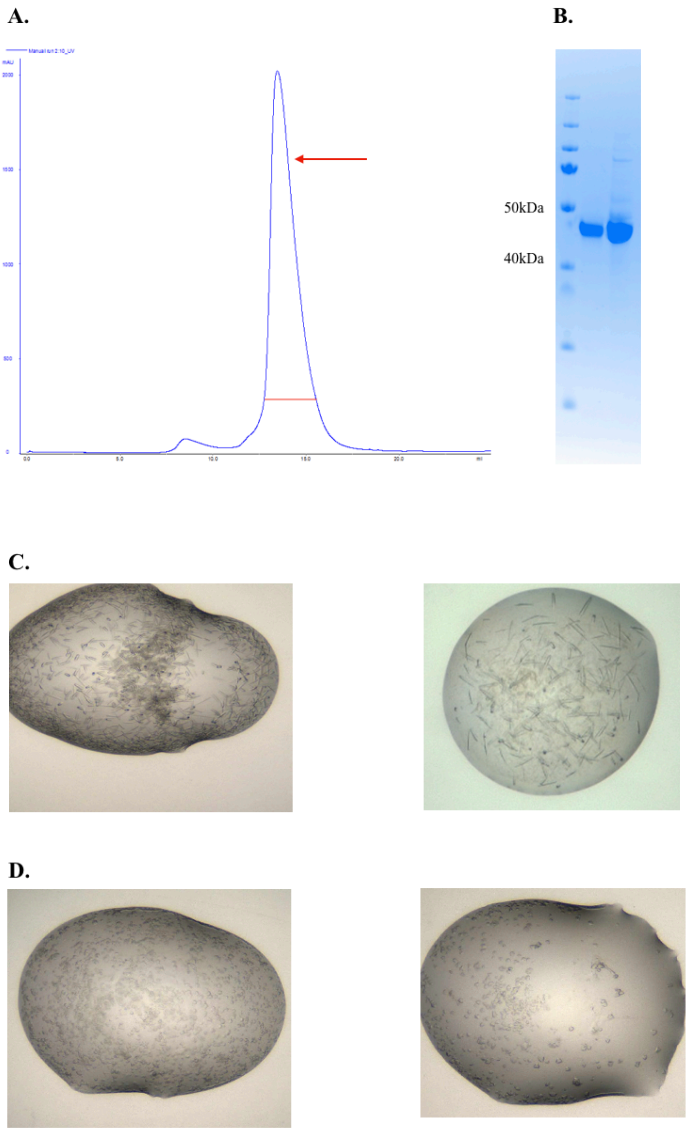


Figure 3.3 Purification and crystallization of TUT7_CM1 A-B. Size exclusion chromatography profiles and SDS-PAGE analysis of TUT7_CM1. Arrow indicates the correct homologous peak and red line cuts off the fractions used for SDS-PAGE analysis. C. TUT7_CM1 crystals of coarse screening (left) and additive optimization (right) obtaining from condition_1 (0.1 M NaCl 0.1 M Bis-Tris pH 6.5, 1.5 M ammonium sulfate, overall pH 6.9. Condition2 containing: 0.1 M Bis-Tris pH 6, 0.4 M ammonium sulfate) D. TUT7_CM1 crystals of coarse screening (left) and additive optimization (right) obtaining from condition_2 (2.8 M sodium acetate pH 7.0).

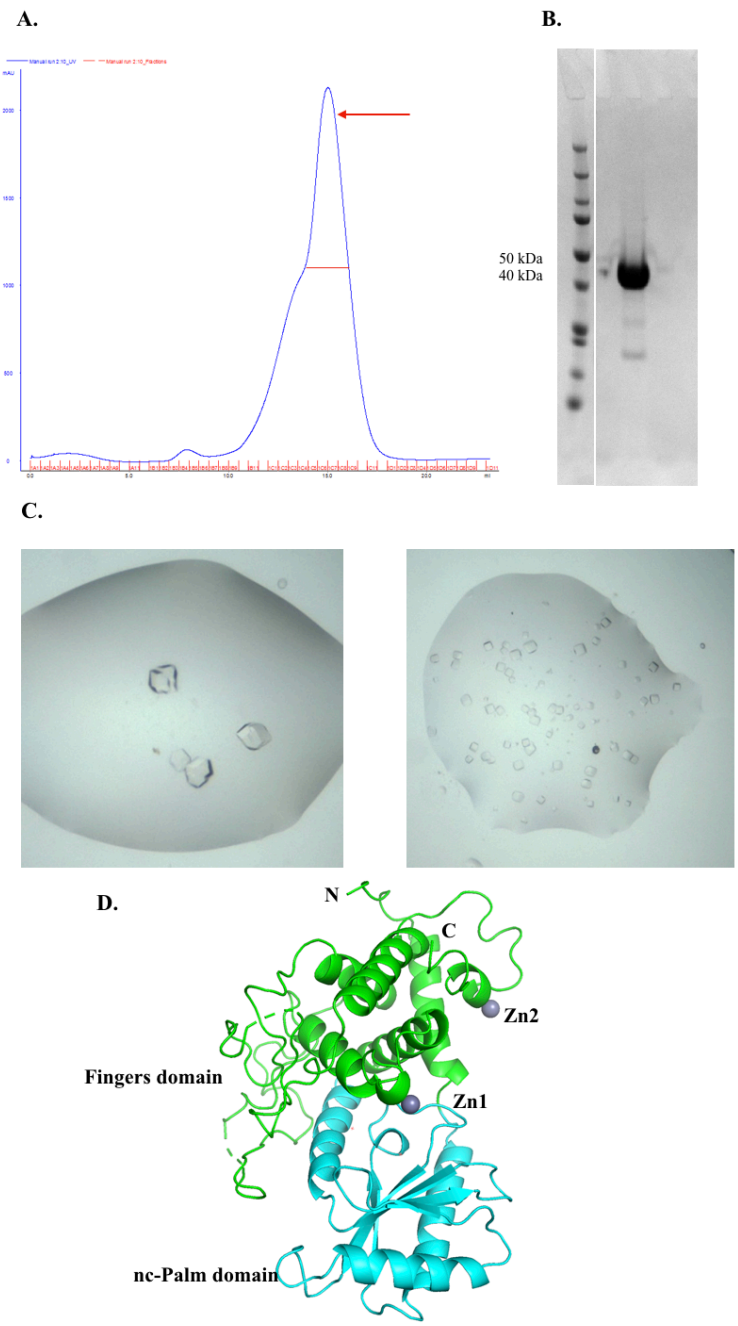


Figure 3.4 Purification and crystallization of TUT7_CM2 A-B. Size exclusion chromatography profiles and SDS-PAGE analysis of TUT7_CM2. C. Crystals of coarse screening (left) and additive optimization (right) obtaining from condition_1 D. Overall structure of TUT7_CM2. Blue and green color presents the nc-palm domain and finger domain respectively. The purple dots are zinc ions. The nc-Palm domain (cyan) consists of five β-strands closely wrapped by three α-helices, while the Fingers domain (green) has six α-helices. The two domains are connected by a long α-helix.

3.2.4 Purification of TUT7_LIM

Obtaining the LIM module of TUT7 was significant for exploring the interaction with Lin28A and pre-let-7. The group of Kozo Tomita expressed the TUT4 LIM module (residues 254-723) with a deletion of an internal putative disordered region, which crystalised and then led to a 2.4 Å crystal structure (Yamashita et al., 2019). Based on this work and amino acid alignment, two pairs of primers were design aiming to obtain TUT7_LIM with high quality.

The LIM module of TUT7 was sub-cloned into the pHGT-bio vector with a TEV cleavable 6x His GST tag at the N-terminus. The expression was successfully assessed, and subsequently small-scale purification was performed, as described in 3.2.2 (Figure 3.5A-B). 5 mg TUT7_LIM1 and 6 mg TUT7_LIM2 from 2-liter TB medium with 85% purity were obtained from IMAC using Ni-NTA columns (Figure 3.5C), which were aliquoted and flash-frozen by liquid nitrogen followed by storing in a -80 °C freezer for subsequent assay work.

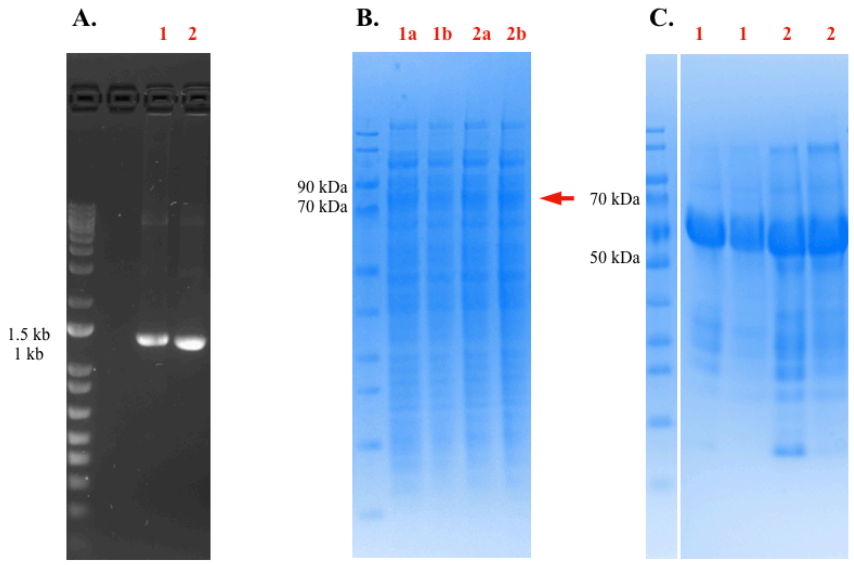


Figure 3.5 Cloning, expression and purification of TUT7_LIM A. Agarose gel electrophoresis analysis of PCR products of TUT7_LIMs. Lanes 1-2 show the TUT7_LIM1(194-643) and TUT7_LIM2 (194-633) respectively, same below. B. SDS-PAGE analysis of expression of TUT7_LIMs. A-b indicate different colonies. C. SDS-PAGE analysis of recombinant TUT7_LIMs eluted from IMAC using Ni-NTA columns.

3.2.5 Purification of Lin28

Human Lin28A_middle deletion (Lin28A_MD), Lin28A, Lin28B, Sumo_Lin28A constructs were synthesized to order by Twist Bioscience. The sequences of both amino acids and DNA are provided in the Appendix. Purification of Lin28A_MD was initially tried as follows.

The Lin28A_MD gene was amplified by PCR and sub-cloned into pHGT-bio vector with a TEV cleavable 6x His tag at the N-terminus using the LIC method. Expression could be detected by SDS-PAGE analysis and subsequently it was purified on a small scale (Figure 3.6A-B). Lin28A_MD was expressed in KRX stain *E. coli* in TB medium at 37 °C, in which Rhamnose was added at a final concentration of 0.1 % when the OD_{600} of the strain reached around 0.8-1 followed by culturing for 36 hours at 22 °C. The bacterial cells were harvested by centrifugation at 5,000 g, 4 °C for 10 minutes. Then, the cell pellets were resuspended using Tris buffer-Ni with the corresponding volume (10 ml buffer for 0.5 g pellets) plus additional lysozyme, homemade benzonase and imidazole at a final concentration of 1 mg/ml, 10 µl/ml, 20 mM respectively. The lysate was sonicated in 5s-on-4s-off cycles at 4 °C for a total of 30 minutes. After centrifugation at 18,000 g, 4 °C for 40 mins, the purification procedures following were similar to those given in section 3.2.2. Finally, the Lin28A_MD was obtained with high yield and purity after size exclusion chromatography (Figure 3.6C-E). However, the ratio of 260 nm: 280 nm was significantly high, around 2. Agar gel electrophoresis confirmed the presence of a high concentration of nucleic acid contamination (Figure 3.6F). Since both the CSD domain and ZK domain have high binding affinity to RNAs, the presence of high levels of nucleic acid contamination are not unexpected. To remove nucleic acid contamination, two additional procedures were introduced into the purification protocol of Lin28A_MD. First, PEI was added to the sonicated lysate at a final concentration of 1 mg/ml and incubated on ice for 10 mins followed by centrifugation at 20,000 g, 4 °C for 60 minutes. Another step was to perform anion exchange chromatography via a Q column (Sigma). Using the optimized protocol, most of the nucleic acid contaminations were removed, which was confirmed by Agar gel electrophoresis (Figure 3.6G-I). Although the ratio of 260 nm: 280 nm was still around 0.9 indicating the presence of a small amount of nucleic acid contaminations, the protein could be used for the following assay trials and cryo-EM experiments. Finally, 5

mg Lin28A_MD from 6-liter TB medium with 95% purity were obtained, concentrated to 2 mg/ml, aliquoted into 1.5 ml EP tubes and flash-frozen using liquid nitrogen followed by storage in a -80 °C freezer.

The recombinant Lin28A full length, Lin28B and sumo_Lin28A were subsequently constructed with a TEV cleavable 6x His tag at the N-terminus and purified by reference to the Lin28A_MD purification protocol (Figure 3.7A). The yields of Lin28A full length, Lin28B and sumo_Lin28A were 2 mg, 1 mg and 1 mg respectively with 95% purity from 2-liter TB medium using Ni-NTA affinity chromatography and ion exchange in small scale purification (Figure 3.7B-D), which were aliquoted into 1.5 ml EP tubes respectively and flash-frozen using liquid nitrogen followed by storage in a -80 °C freezer.

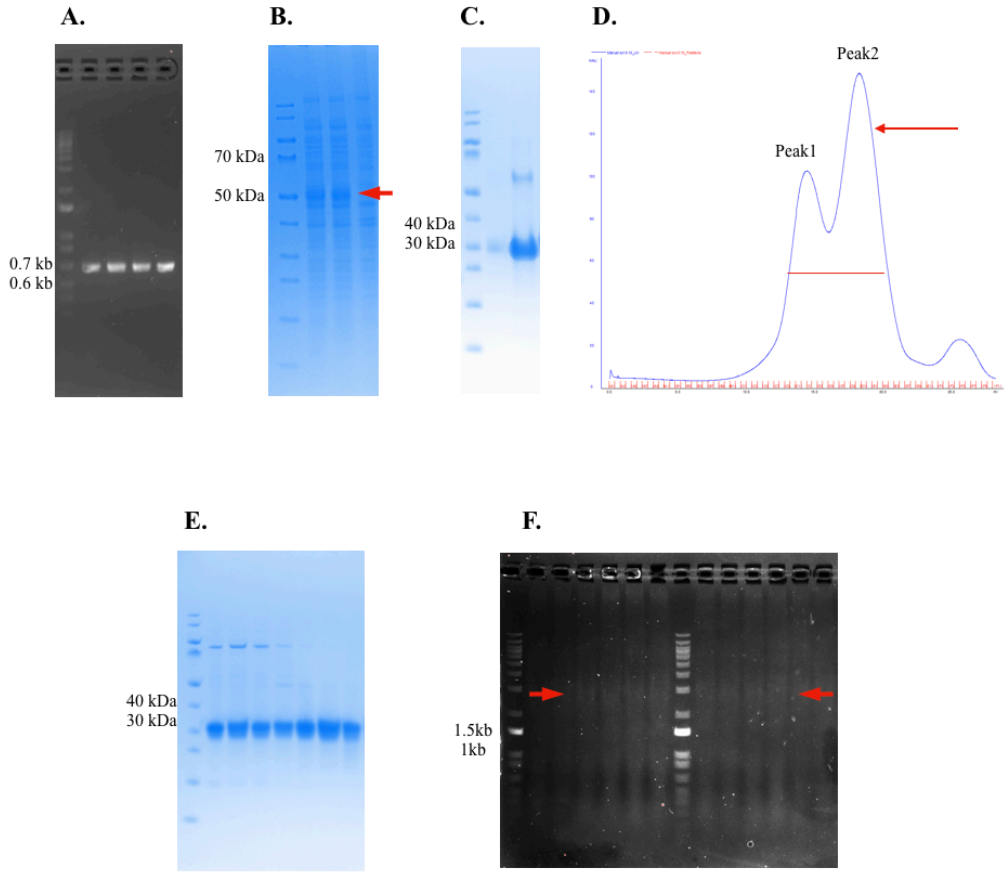


Figure continues on next page

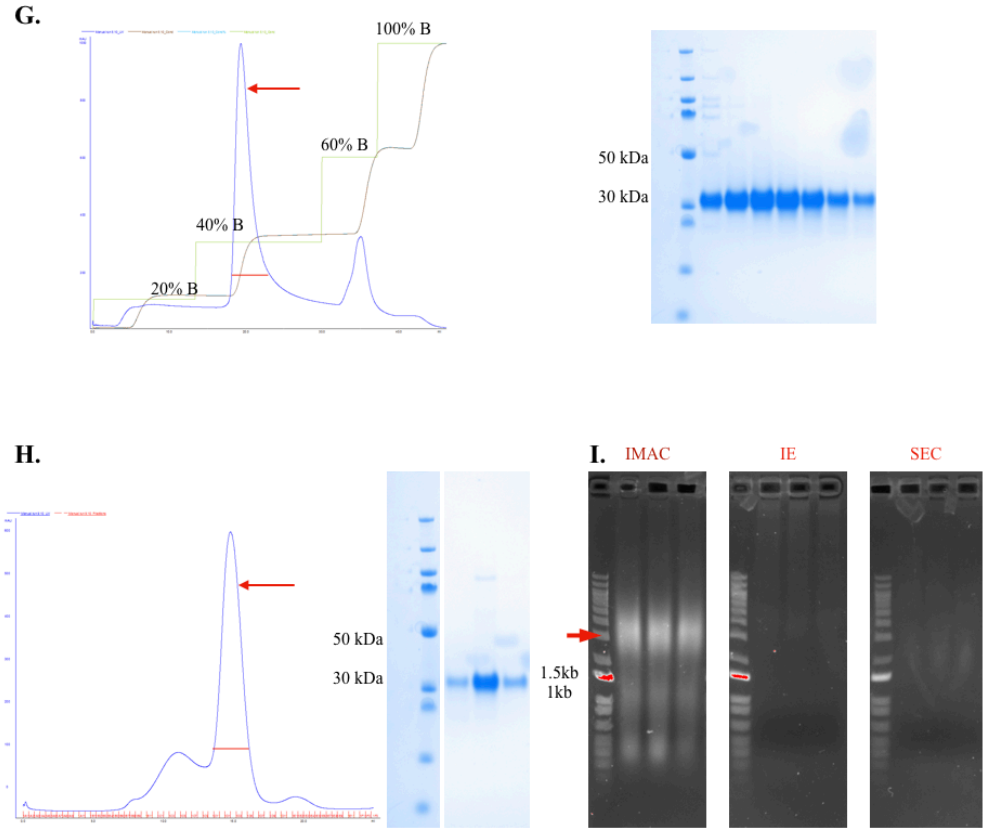


Figure 3.6 Cloning, expression and purification of Lin28_MD A. Agarose gel electrophoresis analysis of PCR products of Lin28_MD. B. SDS-PAGE analysis of expression of Lin28_MD. C. SDS-PAGE analysis of recombinant Lin28_MD eluted from IMAC using Ni-NTA column. D-E. Size exclusion chromatography profiles and SDS-PAGE analysis of Lin28_MD, Arrow indicates the correct peak and red line cuts off the fractions used for SDS-PAGE analysis, same below. The peak1 and peak2 show the monomeric and dimeric Lin28A_MD respectively. F. Agarose gel electrophoresis analysis recombinant Lin28A_MD elution from IMAC from Ni-NTA column and fractions from size exclusion chromatography. G. Ion exchange chromatography profiles and SDS-PAGE analysis of Lin28A_MD. Green curve and mud yellow curve show the concentration of buffer B and salt respectively. H. Size exclusion chromatography profiles and SDS-PAGE analysis of Lin28A_MD. I. Agarose gel electrophoresis analysis recombinant Lin28A_MD elution from IMAC using Ni-NTA column, fractions from ion exchange and fractions from size exclusion chromatography.

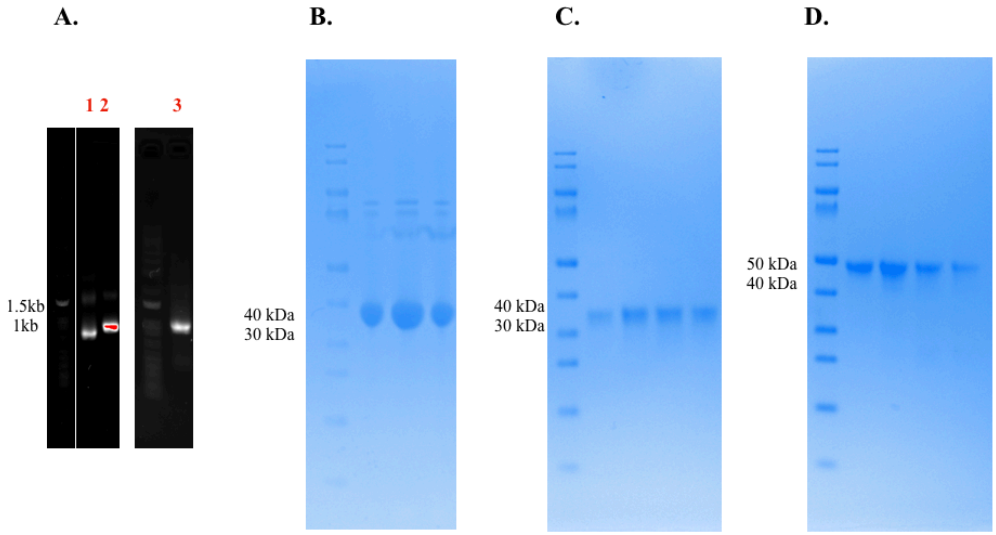


Figure 3.7 Cloning and purification of Lin28A_F, Lin28B and SUMO_Lin28A A. Agarose gel electrophoresis analysis of PCR products of Lin28A, Lin28B and sumo_Lin28A, lanes 1-3 present Lin28A (1-209), Lin28B (1-250) and Sumo_Lin28A (1-301) respectively. B-C. SDS-PAGE analysis of recombinant Lin28B, Lin28A and sumo_Lin28A from size exclusion chromatography.

Primers	Sequence
Lin28A_middleD_1F	TACTTCCAATCCATG GGCTCTGTGTCCAACC
Lin28A_middleD_209R	TATCCACCTTTACTGCT ATTCTGCGCCTCCGGCAGC
Lin28A_1F	TACTTCCAATCCATG GGCTCTGTGTCCAACC
Lin28A_209R	TATCCACCTTTACTGCT ATTCTGCGCCTCCGGCAGC
Lin28B_1F	TACTTCCAATCCATG GCAGAAGGTGGAGCTT
Lin28B_250R	TATCCACCTTTACTGCT GGTCTTCTTACGTTTCTGA
Sumo1_1F	TACTTCCAATCC ATGAGTGATCAAGAAGCTA
Sumo_Lin28A_medium_F	GGTGGTCATAGCACCGTTAAGGCGGCGGAAGAGGC
Sumo_Lin28A_medium_R	GCCTCTTCCGCCGCCTTAACGGTGCTATGACCACC

Table3.2 Primer sequences of human Lin28 used for LIC cloning.

3.2.6 Purification of TUT7

3.2.6.1 Two-step purification

TUT7 full length or large truncated forms containing the two nucleotidyltransferase domains were purified for subsequent cryo-EM trials, since both domains were essential for the interaction with pre-let-7 and Lin28A. Three constructs were made, TUT7 1-1336 (T7F_NF1), TUT7 1-1375 (T7F_NF2), TUT7 1-1495 (TUT7_F). They were all sub-cloned into pHGT-vector with a TEV cleavable 6x His tag at the N-terminus and their expression levels were successfully examined by SDS-PAGE analysis of whole cell lysate. Small scale purification used the same protocol referred to in 3.2.2. The SDS-PAGE analysis results indicated the presence of targeted proteins but with low yield and serious degradation issues. To increase the binding affinity and obtain pure protein, a single Streptavidin-tag was selected aiming to obtain pure proteins. The new Strep-tag was introduced to the C-terminus together with the design of new reverse primers. The expression level of the new batch of TUT7_NFs were examined by SDS-PAGE (Figure 3.8A-B). Meanwhile, the purification protocol was optimized into a two-step purification as described in Chapter 2 and is shown as follows.

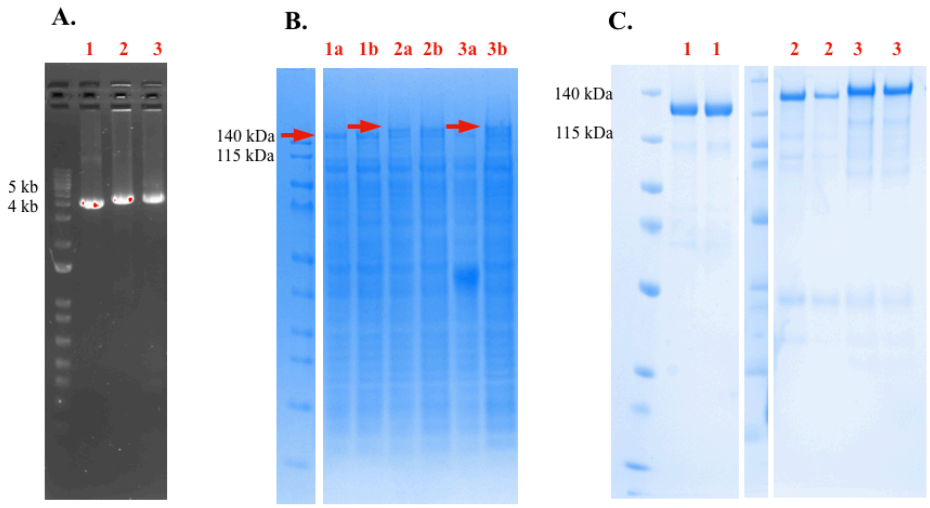


Figure 3.8 Cloning, expression and purification of TUT7 near full length (NFs) A. Agarose gel electrophoresis analysis of PCR products of TUT7_NF1, TUT7_NF2 and TUT7_F respectively in lanes 1-3, same below. B. SDS-PAGE analysis of expression of TUT7_NF1, TUT7_NF2 and TUT7_F individually. A-b indicate different colonies. C. SDS-PAGE analysis of recombinant TUT7_NF1, TUT7_NF2 and TUT7_F eluted from IMAC using Strep resins.

3.2.6.2 Purification of TUT7_MiniF

TUT7_MiniF was first expressed in KRX at 37 °C. The inducer Rhamnose was added into TB medium at a final concentration of 0.1 % when the OD of the strain reached around 0.8-1 followed by culturing for 36 hours at 20 °C. The bacterial cells were harvested by centrifugation at 5,000 g, 25 °C for 10 minutes. The pellets were resuspended using the Tris buffer-Strep (100 mM Tris, pH 7.8, 300 mM NaCl, 5 mM $MgCl_2$, 50uM $ZnCl_2$, 1 % Glycine, 1 % Sucrose, 3 % Glycerol and 0.05 % Tween-20, as optimized conditions) with the corresponding volume (10 ml buffer for 0.5g pellets), plus additional 10 µg/ml homemade benzonase and 2 mM PMSF. The lysate was sonicated followed by centrifugation at 20,000 g, 4 °C for 45 minutes. The sonicated lysate was incubated with Ni-NTA resins by rotating at 7 rpm, 4 °C for 3 hours aiming for better binding. Then the lysate was removed by centrifugation at 1,000 g, 4 °C for 2 minutes, followed by thoroughly washing the resins with Tris buffer-Strep containing 20 mM Imidazole. The fusion protein was eluted by Tris buffer-Strep plus 300 mM imidazole. Elution fractions were examined by SDS-PAGE and pooled together followed by adding TEV proteases at a molar ratio of 1: 20 and were incubated at 4 °C overnight. The next day, the digested protein was incubated with Strep resins, which were pre-equilibrated by Tris buffer-Strep. The binding procedure was done with rotating at 7 rpm, 4 °C for 5 hours, and subsequently the resins were removed by centrifugation at 1,000 g, 4°C for 2 minutes. The streptavidin resins were washed and eluted by Tris buffer-Strep or buffer containing 50 mM biotin respectively. The eluted fractions were confirmed with good purify, but the yields were too low making it hard to perform subsequent ion-exchange or size exclusion chromatography (Figure 3.8C). To obtain high yields of protein, the protocol was further optimized as follows.

The bacterial cell culture and lysate preparation procedures were the same as given above. After centrifugation, the supernatant was poured into Falcon tubes (50ml) and TEV proteases were added at a ratio of 1mg TEV proteases for 50ml supernatant followed by incubating at 4 °C overnight to cleave the 6xHis GST tag at the N-terminus. The next afternoon, streptavidin resins pre-equilibrated with Tris-buffer-Strep were added into the TEV proteases digested supernatant and binding facilitated by rotating at 7 rpm and 4 °C overnight. The following

morning, the Falcon tubes were centrifuged at 1,000 g, 4 °C for 2 mins and then the supernatant was poured out. The streptavidin resins were packaged into a column, followed by washing with 10 ml Tris buffer-Strep per 1 ml streptavidin resins. The Strep-tag fusion proteins were eluted by Tris buffer-Strep containing 50mM biotin and examined by SDS-PAGE with high yield (Figure 3.9A). The eluted fractions contained contaminated proteins and nucleic acids. Anion exchange by Heparin column was selected aiming to remove the contaminations. The eluted fractions were confirmed by SDS-PAGE, showing better purity (Figure 3.9B). Good fractions were pooled together and concentrated to perform size exclusion chromatography. The peak fractions were examined by SDS-PAGE (Figure 3.9C). The final yield of TUT7_miniF was 2 mg with 90% purity from 12-liter TB medium, which were concentrated to 2 mg/ml. The concentrated proteins were aliquoted into 1.5 ml EP tubes and flash-frozen by liquid nitrogen followed by storing in a -80 °C freezer for the subsequent assays and cryo-EM trials.

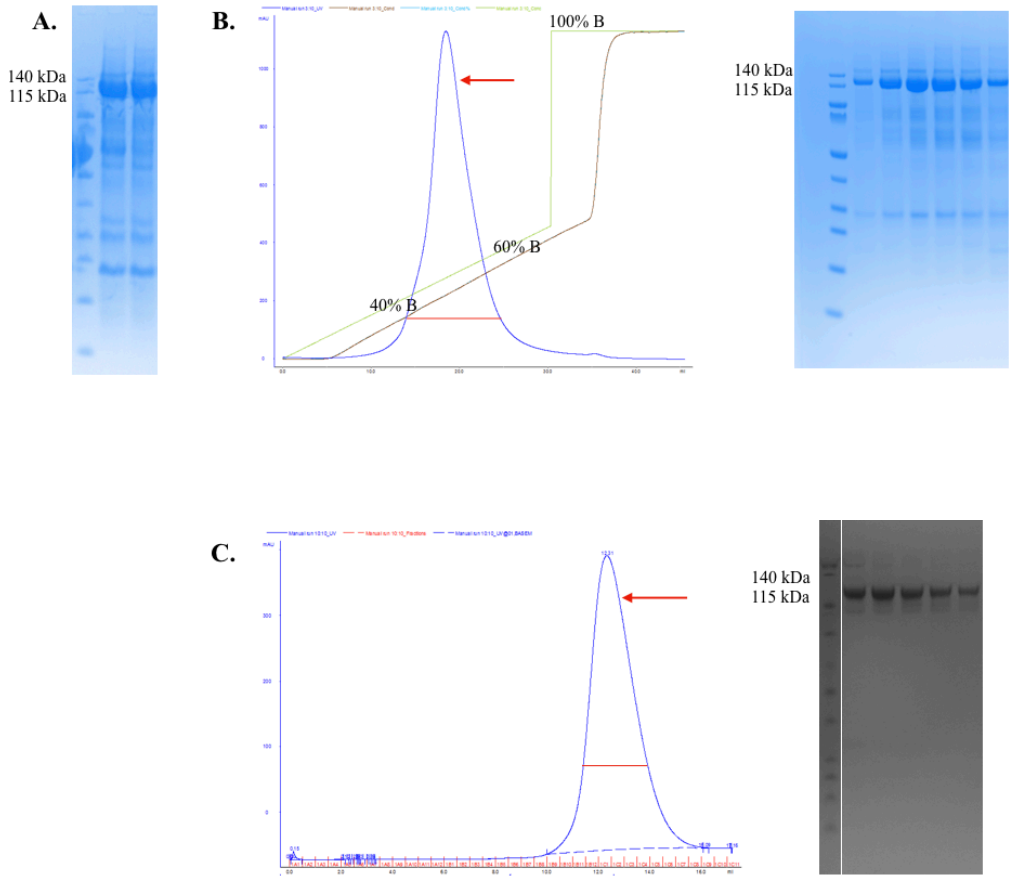


Figure 3.9 Purification of TUT7_mini full-length A. SDS-PAGE analysis of recombinant TUT7_miniF eluted from IMAC using Strep resins. B. Ion exchange chromatography profiles and SDS-PAGE analysis of TUT7_miniF. Arrow indicates the correct peak and red line cuts off the fractions used for SDS-PAGE analysis, same below. Green curve and mud yellow curve show the concentration of buffer B and salt respectively. C. Size exclusion chromatography profiles and SDS-PAGE analysis of TUT7_miniF.

3.2.6.3 Purification of TUT7_Full length

The purification protocol of TUT7_F was similar to 3.2.6.2, while differences were as follows.

Due to the low yields of TUT7_F, a large batch of bacterial cells were cultured. A twin-Strep tag was selected to replace the single Strep tag at the C-terminus to improve the binding affinity of Strep-tag fusion protein to streptavidin resins. The purification protocol was again a two-step purification. Importantly, fresh streptavidin resins from IBA commercial company were used for each batch of purification, since regenerated streptavidin resins have inefficient binding affinity confirmed according to the experience of a number of trials in the case of TUT7 full length purification. Finally, different batches of eluted fractions were pooled together for anion exchange and size exclusion chromatography. Peak fractions were examined by SDS-PAGE and pooled together followed by concentration to 1 mg/ml (Figure 3.10A-C). The final yield of TUT7_F was 1 mg with 90% purity from 12-liter TB medium. The concentrated proteins were aliquoted, flash-frozen and stored in the -80 °C freezer for subsequent assays and cryo-EM trials.

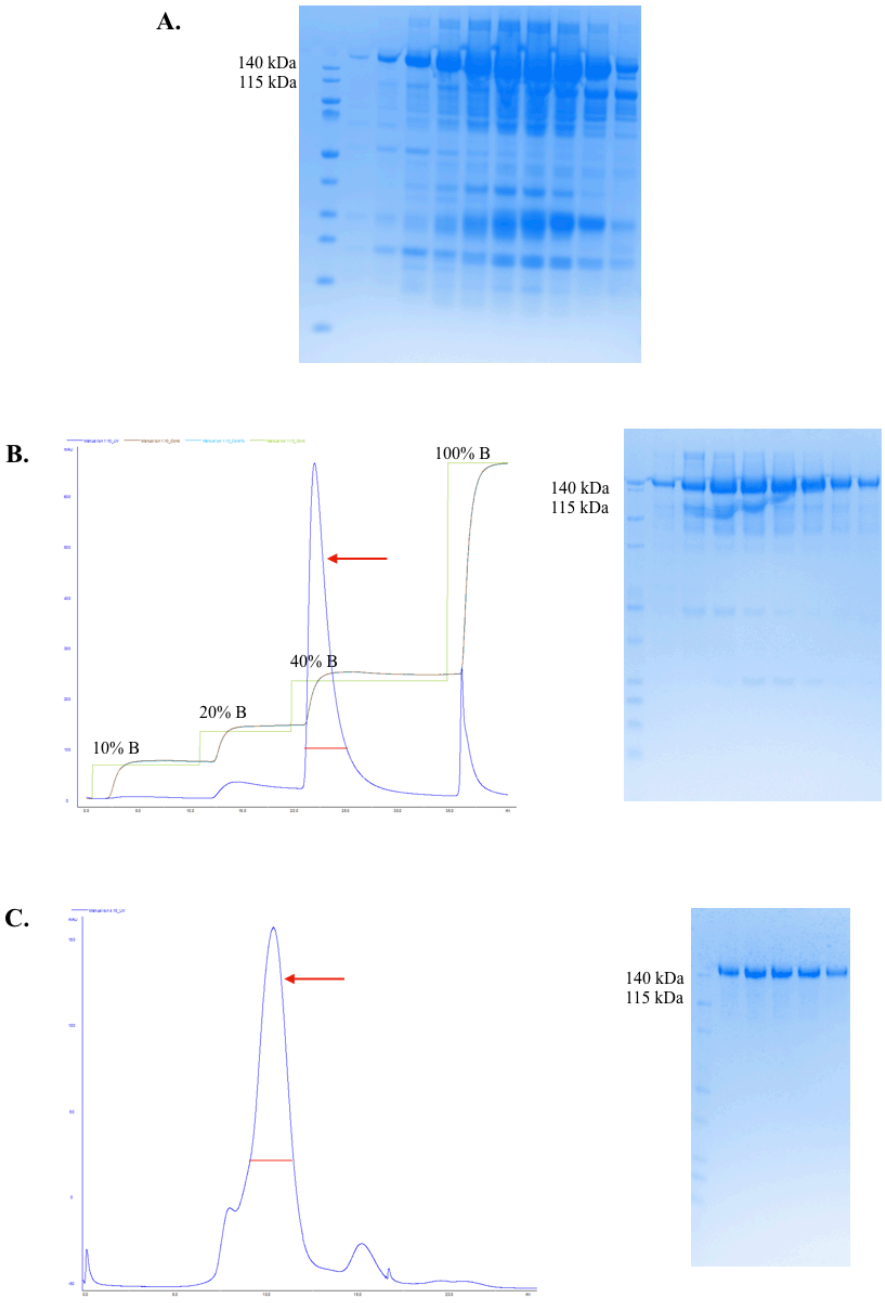


Figure 3.10 Purification of TUT7 full length A. SDS-PAGE analysis of recombinant TUT7_F eluted from IMAC using Strep resins. B. Ion exchange chromatography profiles and SDS-PAGE analysis of TUT7_F. Arrow indicates the correct peak and red line cuts off the fractions used for SDS-PAGE analysis. Green curve and mud yellow curve show the concentration of buffer B and salt respectively C. Size exclusion chromatography profiles and SDS-PAGE analysis of TUT7_F.

3.2.7 Pre-let-7 RNA preparation and assay

The pre-let-7d was supplied by collaborators Dr Afaf el-Sagheer and Professor Tom Brown from the Department of Chemistry, University of Oxford. Pre-let-7g was synthesized by the GenScript Biotech Corporated. The RNA sequences are shown in Table 3.3, and the folding of the pre-let-7 preparation is presented as follows.

The RNA powders were dissolved into RNAase free water and aliquoted into 200 µl PCR tubes (40 µl for each of them). All tubes were flash-frozen by liquid nitrogen and stored in the -80 °C freezer for future assays. RNA folding buffer consisted of 40 mM HEPES (pH7.5), 200mM NaCl, 2 mM EDTA (pH7.5). One tube containing 40 µl pre-let-7 was taken out of the -80 °C freezer and thawed at room temperature. Subsequently, tubes containing 40 µl RNA and 80 µl folding buffer respectively were put in a PCR machine with a temperature setting of 95 °C. After heating 1 minute, 40 µl heated folding buffer was added into the pre-let-7 containing tube, heated for another 2 minutes and then taken out. The tube was left at room temperature for 45 minutes and then put on ice for the following trials. The RNA was sensitive to degradation, so all tips and tubes needed to be sterilized and processed by DCEP water to remove RNA nuclease.

TUT7_F was catalytically active as confirmed by biochemical assays referred to in the protocol at 2.9 in Chapter 2. TUT7_F can only add oligo uridines to 3' of pre-let-7 in the presence of Lin28A (Figure 3.11). Catalytically dead mutant TUT7_F (TUT7_FD), purified using similar purification protocol of TUT7_F (not shown), presented neither mono-uridylation activity nor oligo-uridylation activity (Figure 3.11). The active TUT7_F was considered qualified for subsequent cryo-EM trials.

RNA	Sequence
Pre-let-7-d	AGAGGUAGUAGGUUGCAUAGUUUUAGGGCAGGGAUUU UGCCCACAAGGAGGUAACUAUACGACCUGCUGCCUUUU
pre-let-7-g	UGAGGUAGUAGUUUGUACAGUUUGAGGGUCUAUGAUA CCACCCGGUACAGGAGAUAACUGUACAGGCCACUGCC UUGC
pre-let-7-g-5'CY3	CY3-UGAGGUAGUAGUUUGUACAGUUUGAGGGUCUAUGAUA CCACCCGGUACAGGAGAUAACUGUACAGGCCACUGCC UUGC

Table 3.3 Sequences of RNAs used for assays and cryo-EM

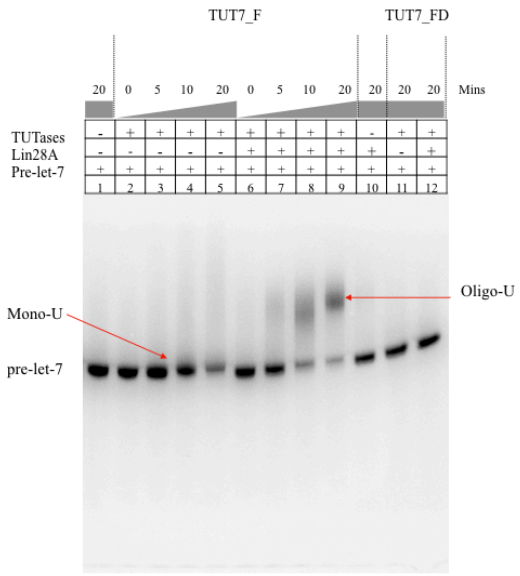


Figure 3.11 Uridylation of pre-let-7 by TUT7 in vitro. Mono-uridylation of pre-let-7 by TUT7_F in the absence of Lin28A at timepoints: 0, 5, 10, 20 mins presenting in lanes 2-5. Oligo-uridylation of pre-let-7 by TUT7_F in the presence of Lin28A at timepoints: 0, 5, 10, 20 mins presenting in lanes 6-9. Lanes 1 and 10 show negative control with or without Lin28A. Lanes 11-12 present the uridylation of pre-let-7 by TUT7_FD in the absence and presence of Lin28A. The reaction products were separated by 15% (w/v) polyacrylamide gel electrophoresis under denaturing conditions.

3.2.8 Assembly of TUT7 with pre-let-7 in the presence or absence of Lin28A

TUT7 was taken out from -80 °C freezer and melted on ice slowly. The TUT7 was mixed with the folded pre-let-7 at a molar ratio of 1:1.5 and with UTP α S with a final concentration of 100 µM plus 2mM $CaCl_2$. The mixed sample was incubated on ice overnight. The second day, the above sample was concentrated to 120 µl and injected into Superdex 200 Increase 5/150 (GE) after centrifugating at 15,000 g, 4 °C for 10 minutes. Size exclusion chromatography was performed with HEPES buffer-GF2: 20 mM HEPES (pH7.8), 100 mM NaCl, 2 mM $CaCl_2$ and 2 mM DTT and the peak fractions were examined by SDS-PAGE (Figure 3.12A-B and E).

TUT7 was incubated with pre-let-7 using the same protocol above. The next day, Lin28A was added to the TUT7/pre-let-7 mixture at a molar ratio of 1.5:1. After a 3-hour incubation on ice, the sample was concentrated to 120 µl volume and injected into Superdex 200 Increase 5/150 (GE) which was pre-equilibrated with HEPES buffer GF2: 20 mM HEPES (pH7.8), 100 mM NaCl, 2 mM $CaCl_2$ and 2 mM DTT after centrifugation at 15,000 g, 4 °C for 10 minutes. The peak fractions were examined by SDS-PAGE (Figure 3.12C and E).

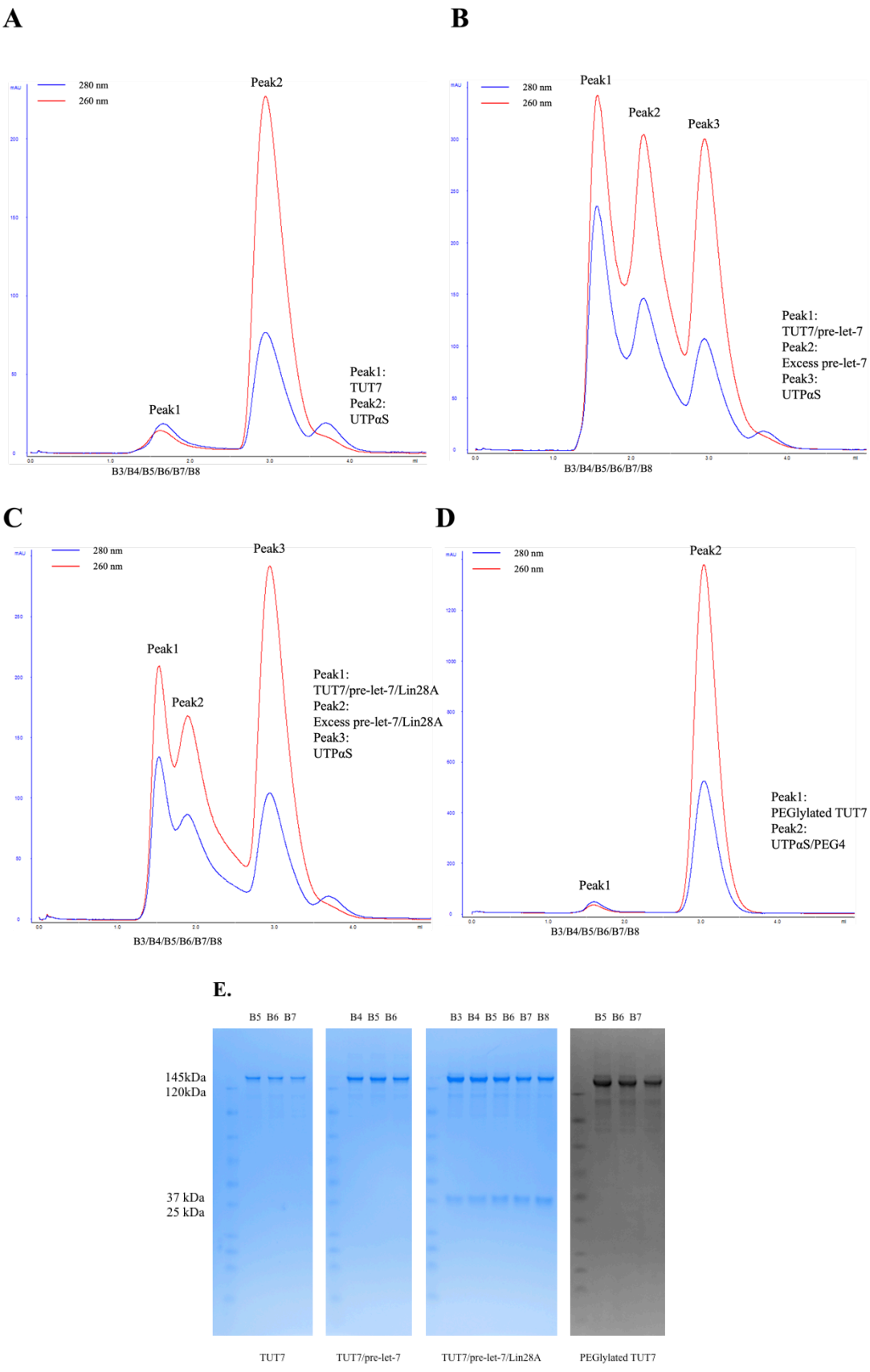


Figure 3.12: Legend on next page

Figure 3.12 Assembly of TUT7 with pre-let-7 with or without Lin28A A. Size exclusion chromatography profile of apo TUT7 in the presence of UTPαS. B. Size exclusion chromatography profile of TUT7/pre-let-7 in the presence of UTPαS. C. Size exclusion chromatography profile of TUT7/pre-let-7/Lin28A in the presence of UTPαS. D. Size exclusion chromatography profile of PEGlylated TUT7 in the presence of UTPαS. The elution volumes of peak1 for apo TUT7, TUT7/pre-let-7, TUT7/pre-let-7/Lin28A and PEGlylated TUT7 are 1.62ml, 1.54ml, 1.52ml and 1.7ml respectively. E. SDS-PAGE analysis of apo TUT7, TUT7/pre-let-7, TUT7/pre-let-7/Lin28A and PEGlylated TUT7 from size exclusion chromatography.

3.2.9 PEGylation modification of TUT7

TUT7 is not only a small protein but also consists of several putatively disordered regions. These cause alignment issues for particle images due to the low signal to noise ratio. Optimization in buffer conditions and of the protein itself were attempted. Small particles tend to stick to the air-water interface during vitrification, and in this project, a PEGylation method was selected as it prevents the particles from sticking to the air-water interface where they were prone to denaturation (Zhang et al., 2021). The PEGylation trial was as follows.

TUT7 was prepared in HEPES buffer: 20 mM HEPES (pH7.8), 100 mM NaCl, 2 mM DTT. NHS-PEG4-Azide (Thermo Fisher) was added at a final concentration of 2 mM. The reaction was incubated on ice for 2 hours, quenched by adding Tris buffer (pH 8.0) at a final concentration of 50 mM and then concentrated to 120 µl followed by size exclusion chromatography on Superdex 200 Increase 3.2/300 (GE) after centrifugation at 15,000g, 4 °C for 10 minutes. The peak fractions were examined by SDS-PAGE, and subsequent qualified fractions were selected for cryo-EM trials (Figure 3.12 D and E).

3.2.10 Negative staining of TUT7

To identify the level of TUT7 homogeneity and quality for further investigation by cryo-EM trails, I first performed negative stain EM imaging of the TUT7_miniF and TUT7_F. Particles were uniformly dispersed, and particle sizes were observed as expected (Figure 3.14A, C). Raw images were collected and around 9,000 and 15,000 particles were picked manually for TUT7_miniF and TUT7_F respectively to generate 2D class averages (Figure 3.14B, D). The features of TUT7_F suggest a ring-like conformation, indicating that the N- and C-terminal regions may be in close contact with each other. Although little information can be obtained from the negative stain EM analysis due to their low resolution, the sample quality has thereby been confirmed as good in terms of particle size and distribution.

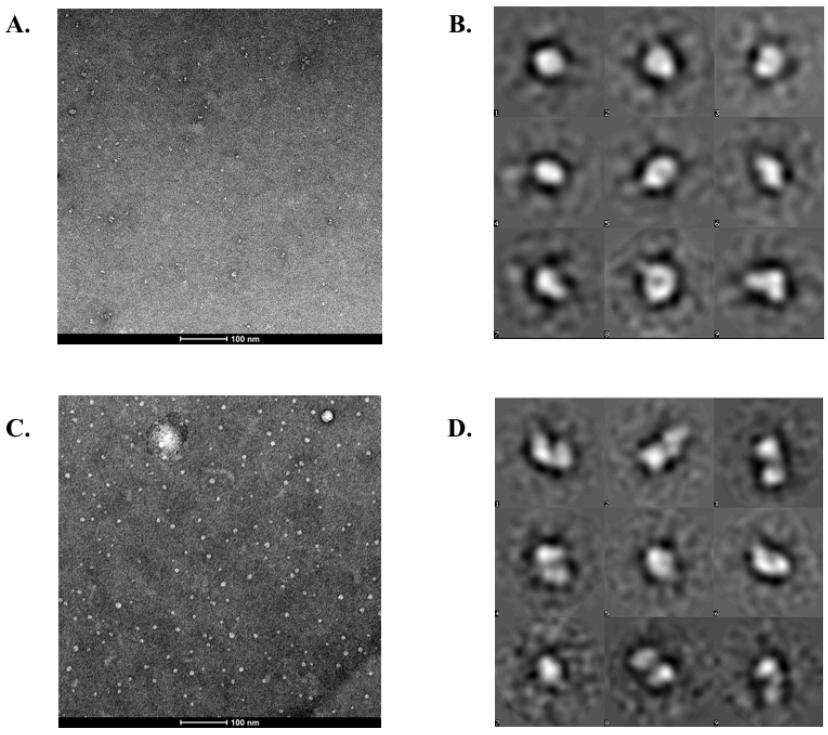


Figure 3.13 Negative staining of TUT7 A. Representative negative staining micrograph of TUT7_miniF. B. Gallery of reference-free 2D class averages of TUT7_miniF. C. Representative negative stain micrograph of TUT7_F. D. Gallery of reference-free 2D class averages of TUT7_F.

3.2.11 Cryo-EM imaging of TUT7 on Glacios microscope

Given good negative staining results, I proceeded with cryo-EM trials of TUT7. Grids with optimal ice thickness were successfully obtained (Figure 3.12A), on which the particles were sufficiently mono-dispersed for analysis. 1,310 movies were collected using the Glacios microscope at the Division of Structural Biology/Oxford Particle Imaging Centre, and around 300,000 particles were selected using manual and template-based picking and roughly 120,000 particles were selected from good classes via their 2D class averages to reconstruct a TUT7 initial 3D model (Figure 3.12B). Six different classes were then generated in subsequent hetero-refinement and one of them showed good quality with non-uniform refinement. The structure of TUT7 was obtained at a resolution of 8.6 Å. Overall, the "C-shaped" architecture of TUT7 is slightly different from the "L-shape" human dicer (Liu et al., 2018b) (Figure 3.12C-D).

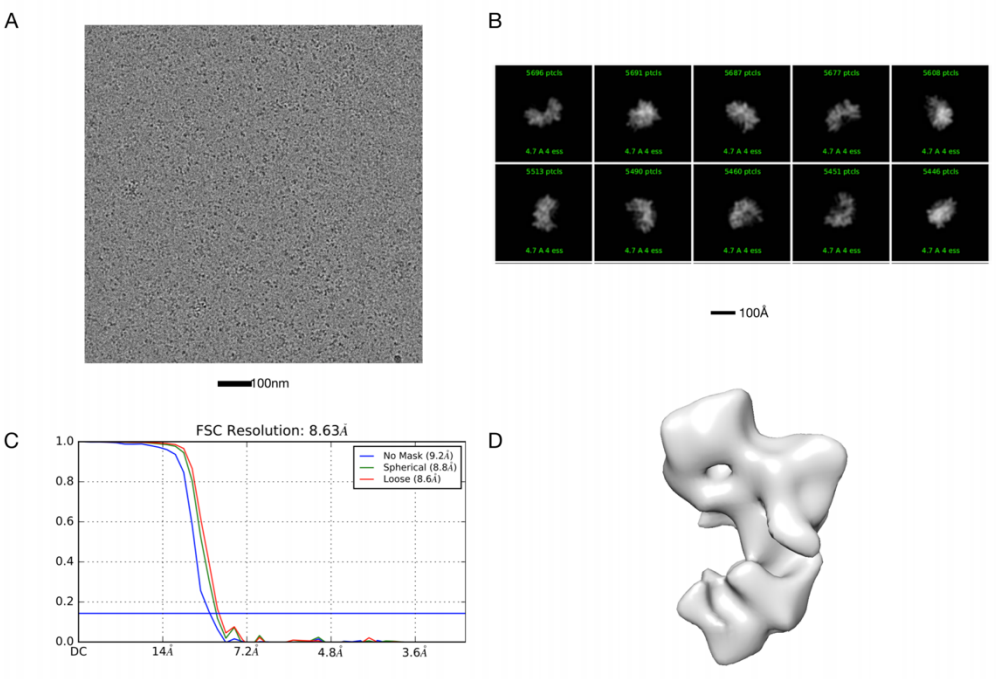


Figure 3.14 Cryo-EM analysis of TUT7 apo A. Representative negative staining micrograph of TUT7 apo. B. Gallery of reference-free 2D class averages of TUT7 apo. C. Gold standard FSC (0.413) curves for TUT7 apo reconstructions. D. Low-resolution model of TUT7 apo.

3.2.12 Cryo-EM imaging of TUT7 on Krios microscope

The high-resolution model of TUT7 apo was difficult to obtain due to difficulty in performing alignments and the damage to particles at air-water interfaces. An optimization strategy was to label the TUT7 protein with NHS-PEG4-Azide (Thermo Fisher), which improved the contrast of the particles significantly. Modified particles had better alignment than the unmodified group. Both 2D class images and 3D model thereby gained significant improvement, ultimately the TUT7 apo structure was determined at 4 Å using the dataset from the PEGylated sample Figure (3.13A-E). The obvious CM density was fitted with the solved crystal structure of TUT7_CM2. Since there was no available crystal structure of TUT7_LIM, the AlphaFold TUT7_LIM obtained from the AlphaFold Protein Structure Database was used as a starting point (Q5TAX3). Both modules, the LIM and CM, docked well into the map, except that the zinc finger in the LIM could not be fitted in due to the poor appearance of the density, indicating the flexibility of that region. Both the N-terminus (1-202 aa) and C-terminus (1365-1495) were missing, which was consistent with the disordered region prediction. The docked model was then processed by multiple rounds of real-space refinement in Phenix (Adams et al., 2002) and manual rebuilding in Coot (Emsley & Cowtan, 2004; Goddard et al., 2005; Lam & Siu, 2017)). The generated model was used for subsequent analysis. To validate the quality of models, Fourier shell correlation (FSC) between the map and the refined model was calculated (FSCsum). Then the model was compared to both first half map (FSC work) and the second half map (FSC free) aiming to detect any signs of overfitting. All figures were made using Chimera (Goddard et al., 2005), Pymol (Bramucci et al., 2012) and Coot (Emsley & Cowtan, 2004; Goddard et al., 2005; Lam & Siu, 2017).

The binary complex and ternary complex samples obtained from 3.2.8 were used for cryo-EM. Optimal grids were collected on a Krios microscope and processed by Cryosparc (Punjani et al., 2017; Scheres, 2012), outputting a 3.7 Å model for the binary complex and two ternary complex models with resolutions at 3.8 Å and 3.7 Å. For the binary complex (Figure 3.14A-E), similar to TUT7_apo, the TUT7_CM2 and Alpha-fold TUT7_LIM were applied for docking with an additional predicted pre-let-7g structure generated from ViennaRNA Web Services (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi). The docked model

was then processed by multiple rounds of real-space refinement in Phenix (Adams et al., 2002) and manual rebuilding in Coot (Emsley & Cowtan, 2004; Goddard et al., 2005; Lam & Siu, 2017). The generated model was used for subsequent analysis. With respect to the ternary complex (Figure 3.15A-H), the Lin28A zinc knuckle domain was extracted from the solved crystal structure (PDB:3TS2) and applied for docking into the ternary complex map. In addition, the predicted pre-let-7g was used and fitted into the map. The docked model showed that the stem region of the pre-let-7g, in particular, fitted fairly well to the map, but that there were significant variations from the predicted structures in other regions. In early analysis, the online normal-modes and Phenix (Adams et al., 2002) driven Namdinator server was used in map interpretation (https://namdinator.au.dk/) but as the reconstruction resolution improved a combination of manual rebuilding in Coot (Emsley & Cowtan, 2004; Goddard et al., 2005; Lam & Siu, 2017). and real-space refinement in Phenix (Adams et al., 2002) were exclusively used for subsequent analysis.

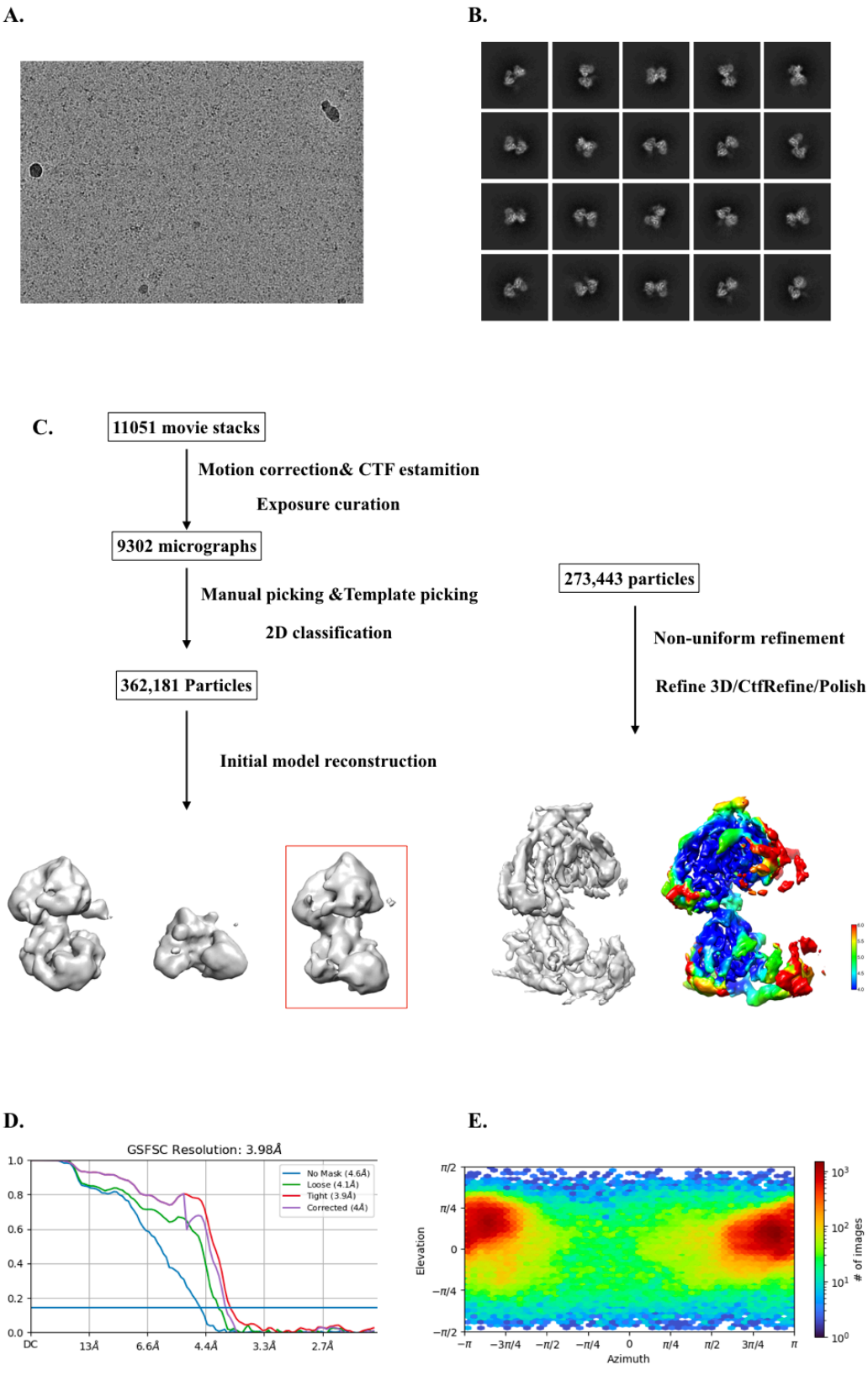


Figure 3.15: Legend on next page

Figure 3.15 Cryo-EM image processing workflow of TUT7 apo A. Representative raw micrograph of TUT7 apo. B. Gallery of reference-free 2D class averages of TUT7 apo. C. Data processing workflow of TUT7 apo. D. Gold standard FSC curves for TUT7 apo reconstructions. E. Orientation distribution heat map of TUT7 apo.

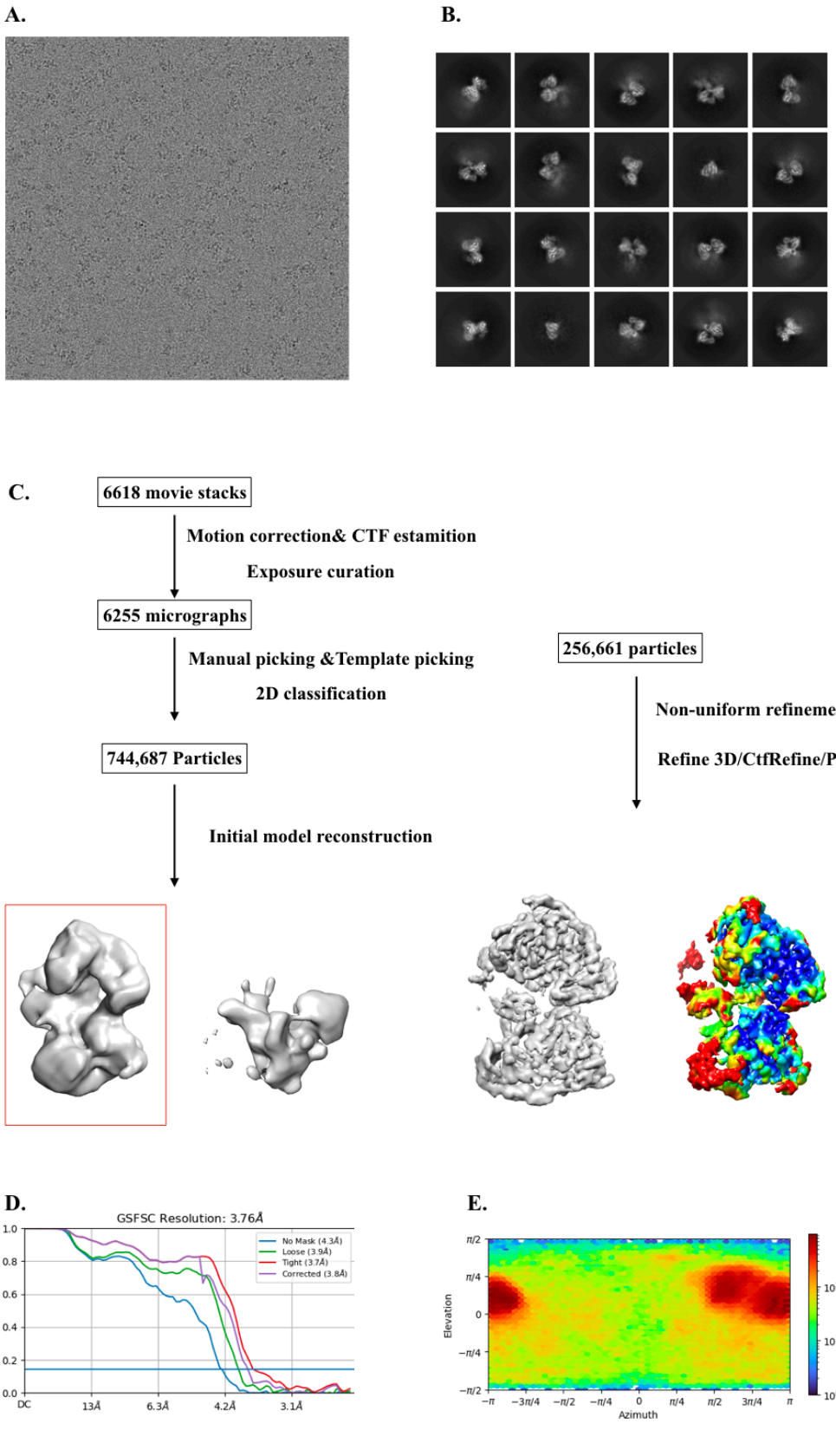
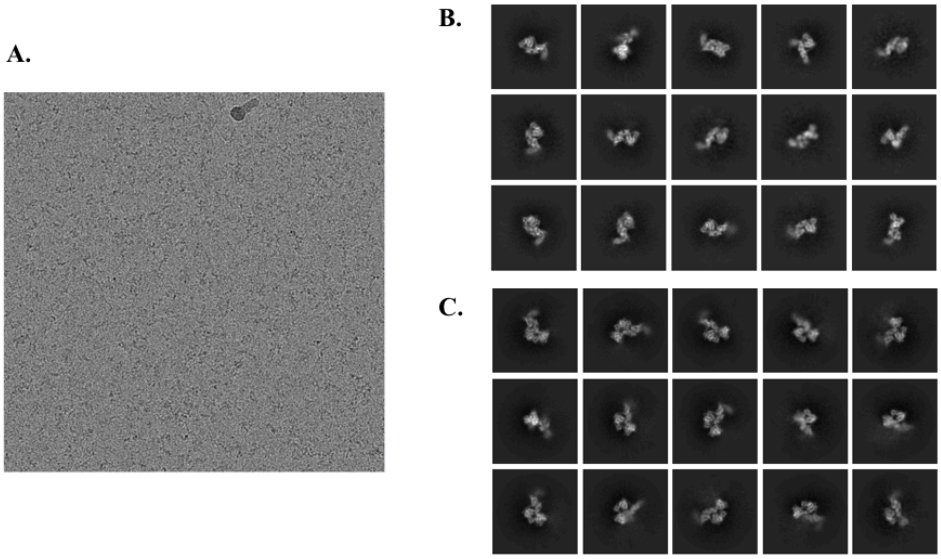


Figure 3.16: Legend on next page

Figure 3.16 Cryo-EM image processing workflow of TUT7 with pre-let-7 A. Representative raw micrograph of TUT7/pre-let-7 complex. B. Gallery of reference-free 2D class averages of TUT7/pre-let-7 complex. C. Data processing workflow of TUT7/pre-let-7. D. Gold standard FSC curves for TUT7/pre-let-7 complex reconstructions. E. Orientation distribution heat map of TUT7/pre-let-7 complex.



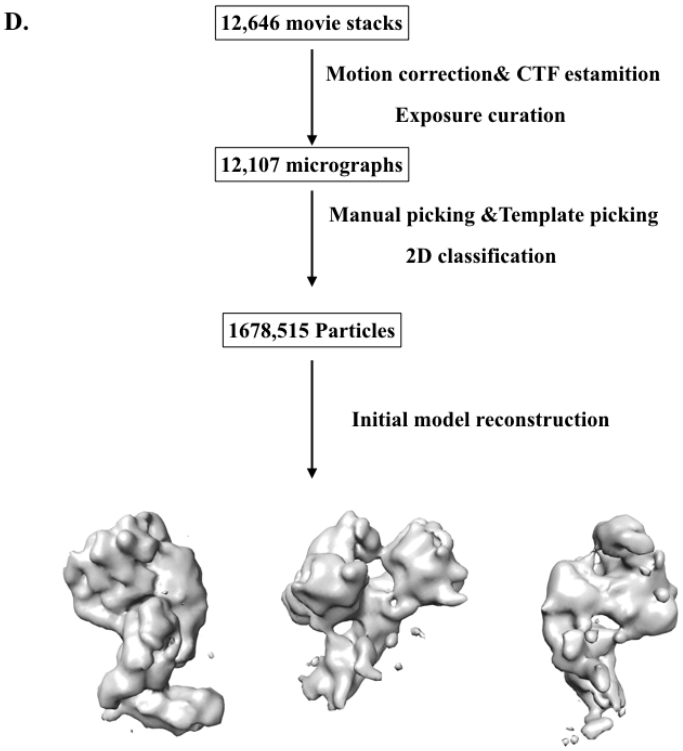


Figure continues on next page

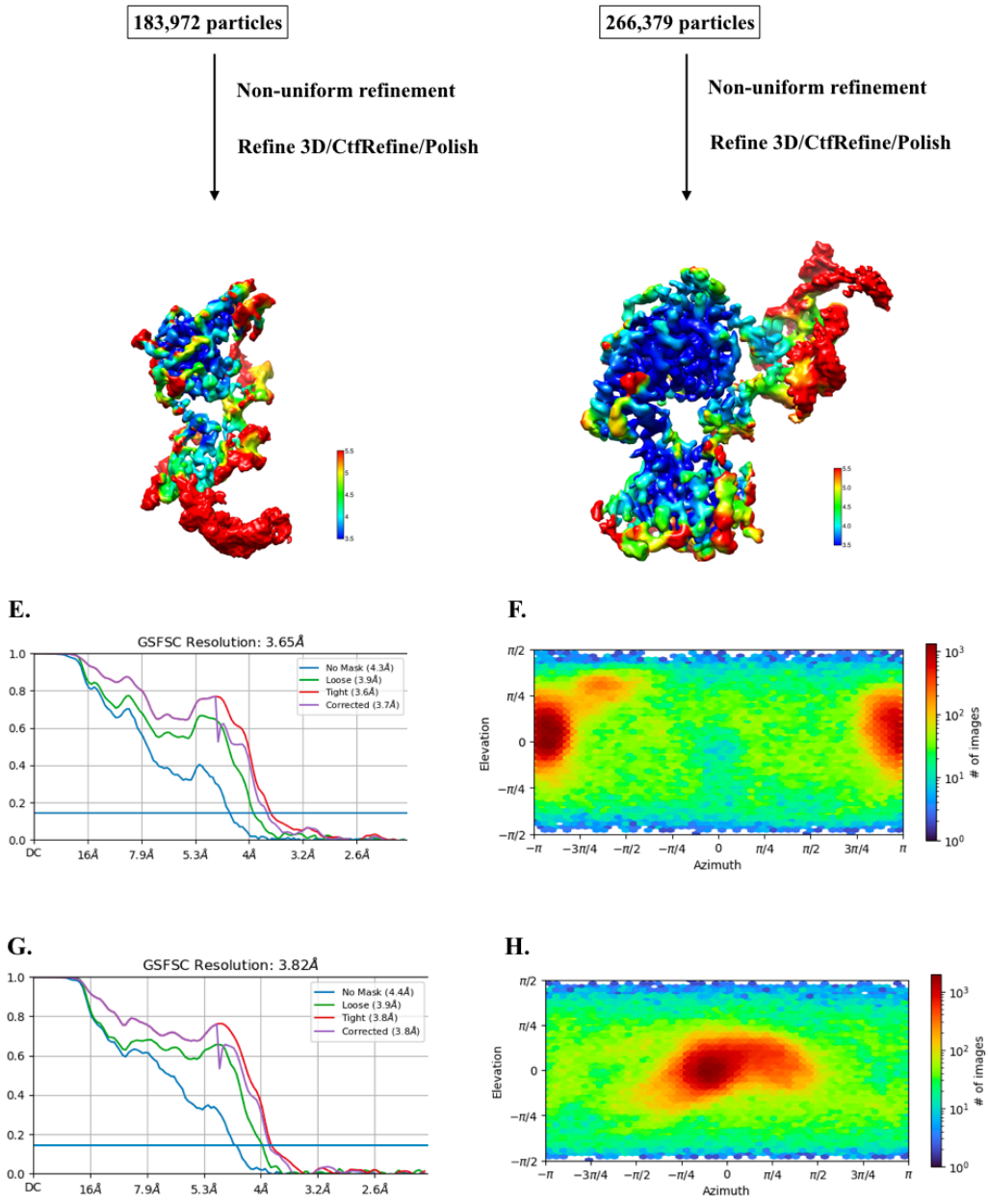


Figure 3.17 Cryo-EM image processing workflow of TUT7 with pre-let-7 and Lin28A A. Representative raw micrograph of TUT7/pre-let-7/Lin28A complex. B. Gallery of reference-free 2D class averages of TUT7/pre-let-7/Lin28A complex (related to conformation-I). C. Gallery of reference-free 2D class averages of TUT7/pre-let-7/Lin28A complex (related to conformation-II). D. Data processing workflow of TUT7/pre-let-7/Lin28. E. Gold standard FSC curves for TUT7/pre-let-7/Lin28A complex reconstructions (related to conformation-II). F. Orientational distribution heat map of TUT7/pre-let-7/Lin28A complex (related to conformation-II). G. Gold standard FSC curves for TUT7/pre-let-7/Lin28A complex

reconstructions (related to conformation-I). H. Orientation distribution heat map of TUT7/pre-let-7/Lin28A complex (related to conformation-I).

3.3 Discussion

The application of an HTP pipeline has benefited the successful purification of both full length TUT7 and diverse truncated forms. Using X-ray crystallography, the TUT7_CM was resolved at 2.6 Å, and is therefore shown as a promising candidate for fragment screening trials (Li, 2020), thus providing a structural basis for drug design and potential binding pocket prediction. Four cryo-EM structures containing TUT7 in diverse states reveal the uridylation activity switching mechanism from mono-uridylation to oligo-uridylation.

There are still unclear aspects to the regulation of oligo-uridylation by TUTases. Since the pre-let-7 we used currently belongs to group II that has a 1-nt overhang in the 3' end and the presence of 2 mM $CaCl_2$ blocks the uridylation activity, it is still unclear how TUTases work on the pre-let-7 when fully active or when oligo-uridylating to create a longer poly-U tail. Although we term conformation-II of the two ternary complexes an "engaged complex", for oligo-uridylation, it remains potentially pre-active. Uridylation activity states imaged without $CaCl_2$ and UTP α S or with pre-let-7g with different lengths of U-tails could be studied in complex with the enzyme by cryo-EM to provide further mechanistic details. We reason that TUT7 with substrates: pre-let-7+1U, pre-let-7+4U and pre-let-7+8U in the presence of Lin28A should represent an appropriate selection of the continuous state changes to simulate the whole oligo-uridylation process, showing us the different conformational changes between LIM and CM domains. The obvious relative flexibility of the LIM and CM – particularly highlighted by the capture complex (conformation-I) in which the CM is unresolved – suggests a basis for moving along an extending U-tail for a LIM-bound pre-let-7g in the process of oligo-uridylation. There are plans to study the structures of these representative complexes in the near future.

Chapter 4 Characterization of TUT4

4.1. Introduction

TUT4 is homologous to TUT7 in mammalian organisms. They both consist of multiple domains and are involved in similar cellular processes. However, they have different distributions and expression levels in diverse cells and tissues (Heo et al., 2008; Kim et al., 2015; Lim et al., 2014). Kozo Tomita' s group solved the crystal structure of the Lin28A interaction module (LIM) of human TUT4, which is similar to the human TUT7_CM structure overall, but presents a distinctive zinc finger, potentially binding RNA and Lin28A (Yamashita et al., 2019) – and, of course, is catalytically inactive. A better understanding of the mechanism of the TUTases in regulating RNA turnover would be greatly assisted by the structure of a full-length enzyme, with or without RNA and Lin28A. This chapter is a parallel study to that on TUT7 and explores the structures of TUT4 involved in uridylation regulation of pre-let-7.

I first obtained good quality TUT4_CM protein, which crystalised and diffracted to 4.5 Å. Optimization such as via varied buffer conditions and by generating truncated forms are currently being investigated. The LIM truncated protein was purified successfully. Meanwhile, the nanobodies S2A4 and CB11 were purified aiming to explore the potential interface with TUT4_LIM and unravel how S2A4 blocks pre-let-7/Lin28A access to TUT4_LIM. The two nanobodies were used in cryo-EM to stabilize the flexible TUT4, while the functional nanobody S2A4 was mixed with TUT4_LIM for crystallography. The complex crystals diffracted to 5 Å and optimization trials are being currently performed.

In cryo-EM, I found TUT4 to be more flexible than TUT7, perhaps reflecting its greater molecular weight and the associated presence of greater regions of disorder. Low-resolution models of TUT4 with or without pre-let-7 and Lin28A show similar conformational transitions from mono- to oligo-uridylation states compared to TUT7-containing cryo-EM structures (see Chapter 3).

Recently, I reconstructed a high-resolution model of TUT4_miniF with pre-let7 and Lin28A ternary complex at 3.7 Å, from which the ternary complex interface was confirmed and in which the significant "AGGAG" motif in pre-let-7 plays a crucial role in its formation. The

low-resolution model of TUT4_F ternary complex with pre-let-7 and Lin28A indicates additional densities and potentially novel conformational changes.

This chapter concludes that both TUT4_CM and TUT4_LIM/S2A4 crystals have the potential to be optimized for high resolution diffraction and become candidates for inhibitor soaking trials or providing the inhibition mechanism by solving the complex structure of TUT4_LIM/S2A4. The cryo-EM results indicate TUT4_miniF ternary complex was in the pre-oligo uridylation engaged state (equivalent to TUT7 ternary conformation-I) and there are promising signs of being about to resolve TUT4_F ternary structures in novel states, revealing a fuller story of TUT4 in uridylation regulation of pre-let-7.

4.2 Methods and Results

4.2.1 Purification and crystallography of TUT4_CM

TUT4 consists of multiple domains including two NTDs, three zinc knuckles and a zinc finger located in the TUT4_LIM. Compared to TUT7, TUT4 is larger, presenting additional challenges in protein expression and purification (Figure 4.1A-B). Table 4.1 shows all primer sequences used for HTP cloning and expression of TUT4 in diverse forms.

The gene of human TUT4 was synthesized to order by Thermo Fisher Scientific. The sequences of both amino acids and DNA are provided in the Appendix.

TUT4_CM was sub-cloned into pHGT-bio vector with a TEV cleavable 6x His GST tag at the N-terminus and purified by reference to the protocol at 3.2.2. However, the yields and purity of TUT4_CM were inadequate for subsequent crystallography trials. The optimizing step was to add a single Strep-tag onto the C-terminus of TUT4_CM, which can be included with the redesign of the reverse primer (Figure 4.2A-B). As for TUT7, a two-step purification was performed. Ultimately, high yield and pure TUT4_CM were obtained after a combination of anion exchange chromatography and size exclusion chromatography. The homogenous peak fractions were confirmed by SDS-PAGE (Figure 4.2C-E). Finally, 5 mg TUT4_CM from 10-liter TB medium were obtained with 95% purity and concentrated to 20 mg/ml for subsequent crystallography trials. Six commercial crystallisation condition plates were selected for coarse screens using a Mosquito robot set up and sitting drop plates (TTP technology). The plates were sealed and monitored in a Rock Maker imager at 20 °C.

One week after screens, small crystals appeared in conditions: 0.1 M BIS-TRIS PH5.5, 25 % w/v PEG3350, 0.2 M lithium sulfate. The initial crystals were optimized by supplementing the initial conditions with additives. Both initial and optimized crystals were harvested for diffraction trials at Diamond Light Source (Figure 4.2F-G). The best crystals diffracted to 4.5 Å. More recently, another optimization of pH and salt concentrations was attempted, and bigger crystals were obtained and are ready for diffraction testing.

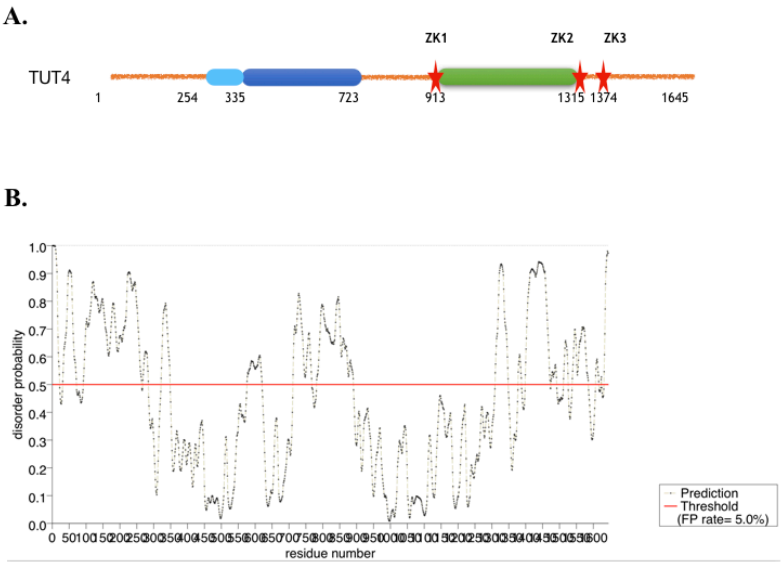


Figure 4.1 Overview of human TUT4 A. Schematic representation of domains of TUT4, consisting of two nucleotidyltransferase domains (NTD1 (blue) and NTD2 (green)), zinc finger (teal) and multiple zinc knuckle motifs (red star). B. Prediction of native disorder regions of human TUT4, showing the potentially flexible regions including two terminus and internal domain, analyzed by PrDOS (Ishida & Kinoshita, 2007).

Primers	Sequence
TUT4_1F	TACTTCCAATCCATG GAAGAGAGCAAGACCCTGAAG
TUT4_25F	TACTTCCAATCCATG GCCGTGCAAGTGATCGGCAA
TUT4_49F	TACTTCCAATCCATG AGCCCCAACAGAAACAGCAGCA
TUT4_254F	TACTTCCAATCCATG GAGATGGACTACCTGGAAAACGC
TUT4_923F	TACTTCCAATCCATG CACTCCAAGAACGACTGCCC
TUT4_943F	TACTTCCAATCCATG ACCAACAGATTCCGCGAGAT
TUT4_537R	TATCCACCTTTACTGCT CAGAGGAGGCTTCCGCTGTT
TUT4_723R	TATCCACCTTTACTGCT CTTGAAGTCCACGGTGCTCTTGT
TUT4_1284R	TATCCACCTTTACTGCT TGTCAGCACTCTGCTGTCGA
TUT4_1315R	TATCCACCTTTACTGCT CAGCAGGCTCTTTCTCTTGGG
TUT4_1400R	TATCCACCTTTACTGCT TGCACCTGATGAGGAGGCT
TUT4_1624R	TATCCACCTTTACTGCT TGTGGCGCATCTGTCTTGG
TUT4_1644R	TATCCACCTTTACTGCT CGGACACATTGCCTCTAGGAGG
S2A4_1F	TACTTCCAATCCATG catatgaaatatttactac
S2A4_112R	TATCCACCTTTACTGCT ttagctgctcacggtcacc
CB11_1F	TACTTCCAATCCATG catatgaaatatttactac
CB_112R	TATCCACCTTTACTGCT ttagctgctcacggtcacc

Table 4.1 Primer sequences of TUT4 and Nanobodies used for LIC cloning.

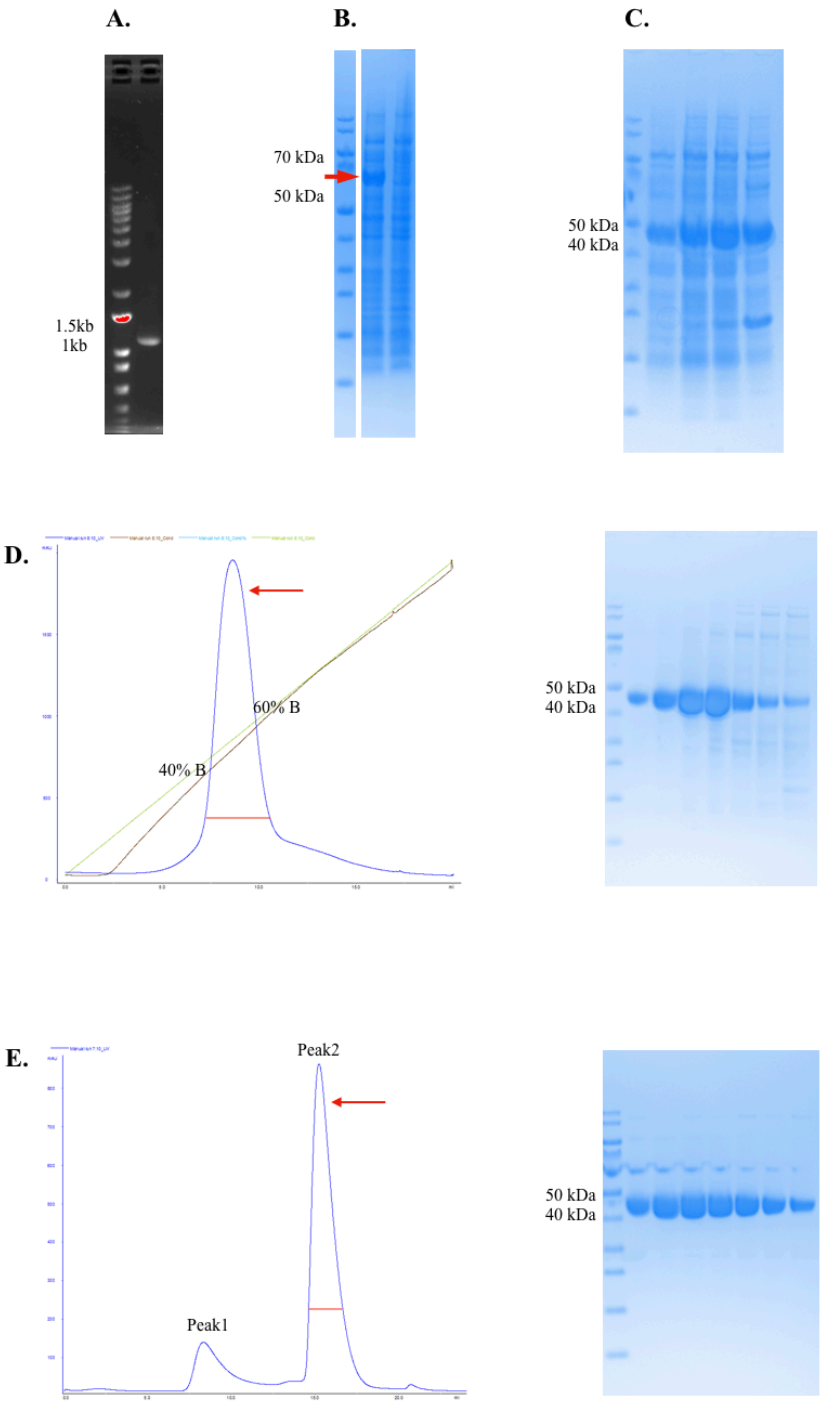


Figure continues on next page

F.

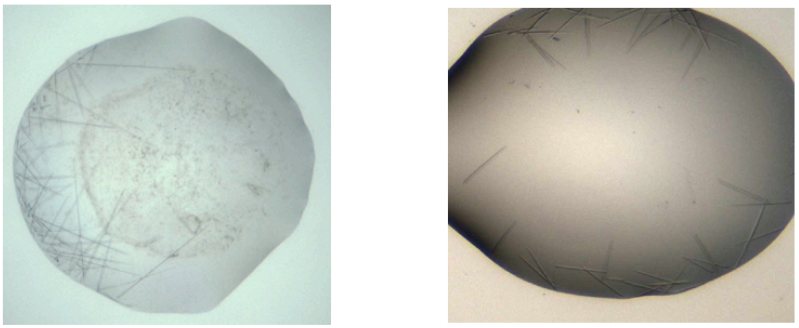


Figure 4.2 Cloning, expression, purification and crystallization of TUT4_CM A. Agarose gel electrophoresis analysis of PCR products of TUT4_CMs. B. SDS-PAGE analysis of expression of TUT4_CMs. C. SDS-PAGE analysis of recombinant TUT4_CMs eluted from IMAC using Strep resins. D. Ion exchange chromatography profiles of TUT4_CM. Arrow indicates the correct peak and the red line cuts off the fractions used for SDS-PAGE analysis, same below. Green curve and mud yellow curve show the concentration of buffer B and salt respectively. E. Size exclusion chromatography profiles and SDS-PAGE analysis of TUT4_CM1. Peak1 shows the aggregates volume and peak2 contains the soluble TUT4_CM1. F. TUT4_CM crystals of coarse screening (left) and additive optimization (right) obtained from the condition 0.1 M BIS-TRIS PH5.5, 25 % w/v PEG3350, 0.2 M lithium sulfate.

4.2.2 Purification of TUT4 LIM

The LIM of TUT4 is significant for exploring the interaction with Lin28A and pre-let-7. Kozo Tomita' s group expressed the TUT4_LIM (254-723 aa) with a deletion of an internal putative disordered region, which crystalised and then led to the determination of a 2.4 Å crystal structure (Yamashita et al., 2019). Based on this work and the secondary structure prediction, I designed a set of primers for HTP cloning and expression of human TUT4_LIM.

The LIM of TUT4 was sub-cloned into the pHGT-bio vector with a TEV cleavable 6x His GST tag at the N-terminus. Small scale purification was performed by reference to section 3.2.2 but failed to obtain pure proteins using Ni-NTA affinity chromatography and ion exchange. Then a single Strep tag was added to the C-terminus and purified using the two-step purification protocol referred to at 3.2.6.1 (Figure 4.3A-C). An issue for TUT4_LIM purification was heat shock protein contamination with a molecular weight around 70 kDa. To solve it, ATP at a final concentration of 5 mM was added to the sonicated lysate and incubated on ice for one hour followed by centrifugation at 20,000 g, 4 °C for 45 minutes. After Ni-NTA affinity chromatography, the eluted fractions were pooled together for ion exchange on an S column (Cytiva) and subsequently pure fractions were selected for size exclusion chromatography. Qualified proteins were obtained after anion exchange and size exclusion chromatography (Figure 4.3D-E). Peak fractions with better than 95% purity were pooled together and concentrated to 5 mg/ml with final yields of 5 mg from 10-liter TB medium. The concentrated proteins were aliquoted and flash-frozen with liquid nitrogen followed by storing in a -80 °C freezer for subsequent assays, crystallography and cryo-EM work.

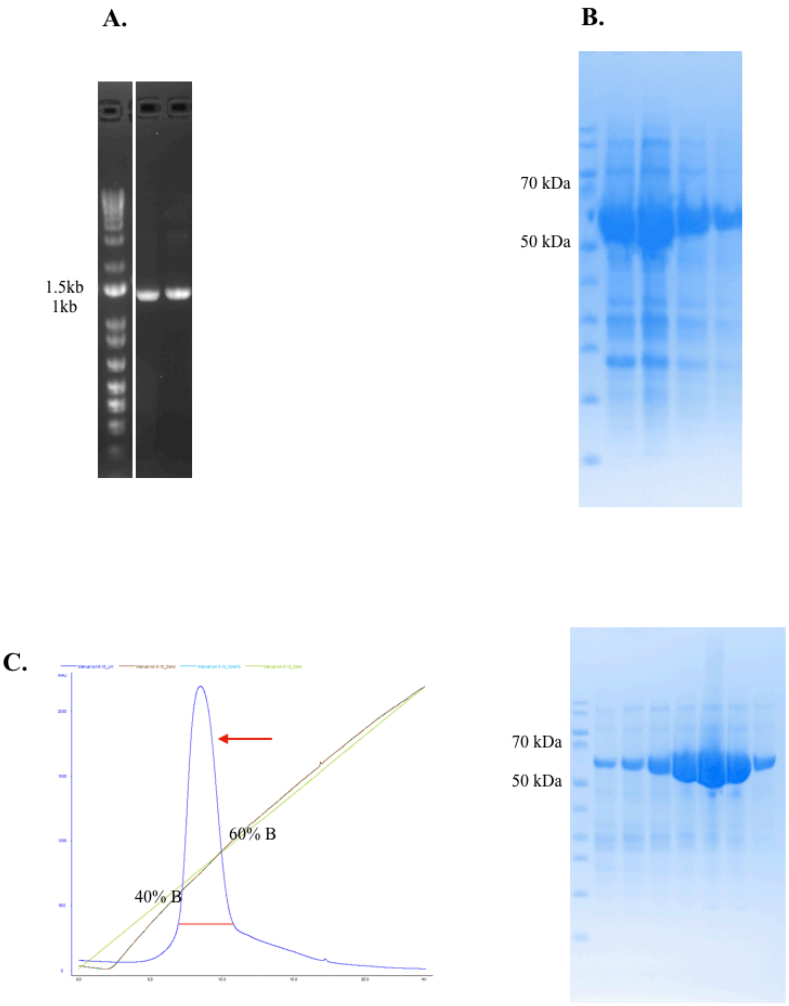


Figure 4.3: Legend on next page

Figure 4.3 Cloning and purification of TUT4_LIMs A. Agarose gel electrophoresis analysis of PCR products of TUT4_LIMs. B. SDS-PAGE analysis of recombinant TUT4_LIM1 eluted from IMAC using Ni-NTA column. C. SDS-PAGE analysis of recombinant TUT4_LIM2 eluted from IMAC using Strep resins. D. Ion exchange chromatography profiles of TUT4_LIM1, Arrow indicates the correct peak and red line cuts off the fractions used for SDS-PAGE analysis, same below. Green curve and mud yellow curve show the concentration of buffer B and salt respectively. E. Size exclusion chromatography profiles and SDS-PAGE analysis of TUT4_LIM1. Peak1 shows the aggregates volume and peak2 contains the soluble TUT4_LIM1.

4.2.3 Purification of S2A4 and CB11

The genes of the S2A4 and CB11 nanobodies were synthesized to order by Twist Bioscience. The sequences of both amino acids and DNA are provided in the Appendix. The purification of nanobodies was as follows.

The S2A4 and CB11 were sub-cloned into pINC-MBP2 vector with a TEV cleavable MBP 6xHis tag at the N-terminus (gift from the SGC (now CMD), Oxford) (Figure 4.4A-B). The purification protocol was similar to 3.2.2. Both nanobodies had high yield and purity in small scale purification. Contaminating proteins including the MBP tag and TEV proteases were removed by reverse IMAC using Ni-NTA column, followed by performing size exclusion chromatography in HEPES-buffer-GF1. The peak fractions were confirmed by SDS-PAGE (Figure 4.4C). 5 mg S2A4 with 95% purity and 3 mg CB11 with 70% purity from 3-liter TB medium were obtained and concentrated to 2 mg/ml respectively. All aliquoted tubes containing pure proteins were flash-frozen in liquid nitrogen and stored in a -80 °C freezer for the following trials.

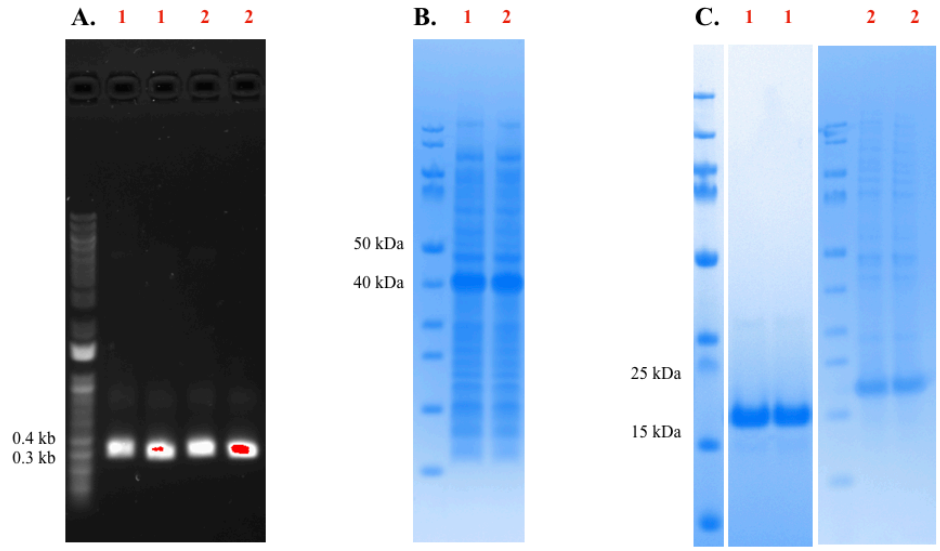


Figure 4.4 Cloning, expression and purification of nanobodies A. Agarose gel electrophoresis analysis of PCR products of S2A4 and CB11 in lanes 1-2, same below. B. SDS-PAGE analysis of expression of S2A4 and CB11. C. SDS-PAGE analysis of recombinant S2A4 and CB11 from IMAC using Ni-NTA columns.

4.2.4 Crystallography of TUT4LIM in complex with S2A4

S2A4 is a functional nanobody blocking the oligo-uridylation of TUT4 to pre-let-7 by disrupting the ternary interaction in the presence of Lin28A, which has been confirmed to bind the zinc finger of TUT4_LIM. The atomic details of a TUT4_LIM/S2A4 complex structure would provide better understanding of the inhibition mechanism of S2A4 to TUT4.

The purified TUT4_LIM and S2A4 were incubated on ice at a ratio of 1:1.5 for half an hour. The mixed sample was concentrated and centrifuged before performing size exclusion chromatography. Homogeneous peak fractions were examined by SDS-PAGE and pure fractions were selected for subsequent crystallography trials (Figure 4.5A-B). TUT4_LIM:S2A4 was concentrated to 12 mg/ml and centrifuged before performing crystallographic coarse screens. Six crystallisation condition plates from Hampton research were chosen for screening using Mosquito robotic set up and sitting drop vapor diffusion. The screen plates were stored and imaged in a Rock Maker imager at 20 °C.

Needle-shaped crystals were visualized in the second week, and optimization with additives was performed around the initial conditions but bigger crystals could not be obtained. To obtain better crystals, gradients of pH and PEG concentrations were prepared based on the initial conditions. Two plates were made and stored at 4 °C and 20 °C individually. Bigger needle-shape crystals were visualized in conditions consisting of 0.350 M Sodium di-Hydrogen Phosphate and 0.650 M di-Potassium Hydrogen Phosphate at both 4 °C and 20 °C (Figure 4.5C). Subsequently, the crystals were harvested and diffracted in the Diamond Light Source with the best resolution 5 Å. Another route to optimization would be to make shorter versions of TUT4_LIM to crystalise with S2A4 and this exploration is under further investigation.

On the other hand, TUT4_LIM in complex with CB11 was prepared using the same protocol given above for TUT4_LIM/S2A4. A 20 mg/ml sample was used for coarse screens and tiny needle-shape crystals were obtained and later harvested for diffraction at the Diamond Light Source. However, the best resolution was 8 Å for both initial and optimized crystals (not

shown). Other optimization strategies like seeding and additives were selected aiming to obtain crystals with higher quality.

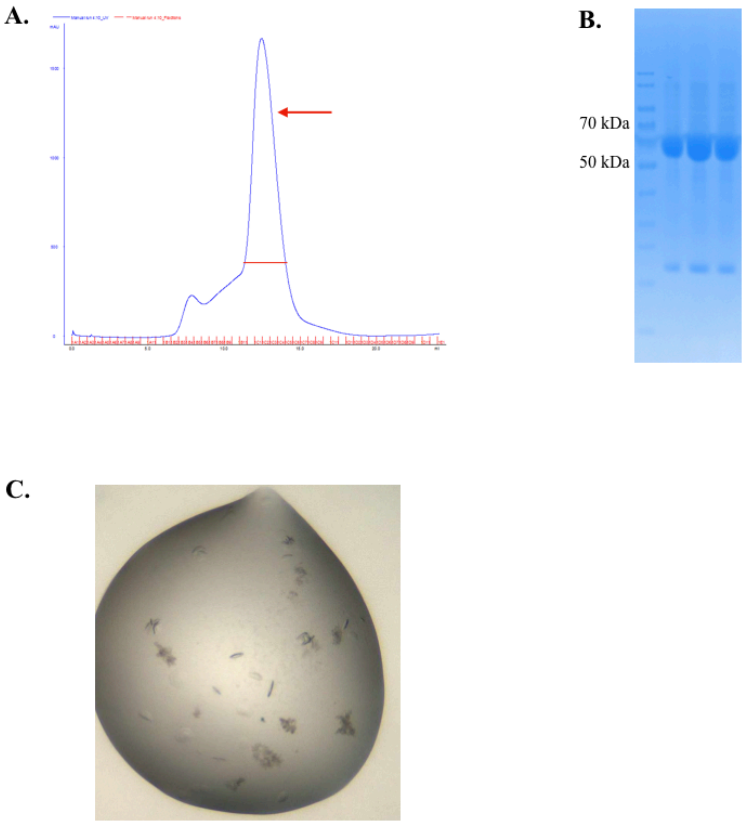


Figure 4.5 Crystallisation of TUT4_LIM in complex with Nanobody S2A4 A-B. Size exclusion chromatography profiles and SDS-PAGE analysis of TUT4_LIM/S2A4, Arrow indicates the correct peak and red line cuts off the fractions used for SDS-PAGE analysis. C. Crystals of TUT4_LIM/S2A4.

4.2.5 Purification of TUT4_miniF construct

Initially, different TUT4 near full length truncated forms were confirmed for successful expression (Figure 4.6 A-B). Then, TUT4_miniF, which consisted of the two nucleotidyltransferase domains, was amplified followed by sub-cloning into pHGT-bio vector with a TEV cleavable 6x His GST tag at the N-terminus, was chosen for large scale purification due to its good yield. The first attempt at purification was by Ni-NTA affinity chromatography using the protocol referred to at section 3.2.2. However, I failed to obtain protein of sufficient quality due to the serious degradation with multiple bands in SDS-PAGE gel as determined by Coomassie blue (not shown). Then, a Strep-tag was added to the C-terminus with design of a new reverse primer. Again, a two-step purification protocol was followed. TUT4_miniF formed oligomers (Faehnle et al., 2017a; Yamashita et al., 2019) and thus DTT was added to the TUT4_miniF containing solution at a final concentration of 3 mM during all purification steps. Pure TUT4_miniF was obtained after combined trials of anion exchange by heparin and size exclusion chromatography (Figure 4.7A-C). Peak fractions with good purity were pooled together and concentrated to 2 mg/ml. The final yield of TUT4_miniF was 3 mg with 90% purity from 12-liter TB medium after size exclusion chromatography, which subsequently were concentrated to 1.5 mg/ml. The concentrated proteins were aliquoted into EP tubes and flash-frozen by liquid nitrogen followed by storing in a -80 °C freezer for subsequent assays and cryo-EM trials.

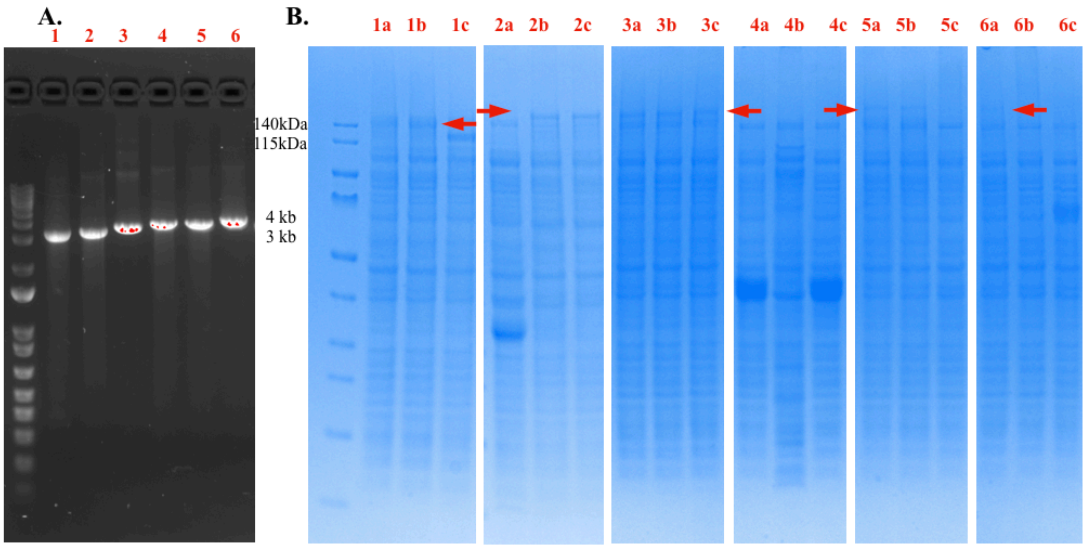


Figure 4.6 HTP cloning and expression of TUT4 near full length A. Agarose gel electrophoresis analysis of PCR products of TUT4_NFs. Lane 1-6 present TUT4_NF1 (254-1315), TUT4_NF2 (254-1400), TUT4_NF3 (254-1649), TUT4_NF4 (1-1400), TUT4_NF5 (1-1624), TUT4_NF6 (1-1649) respectively, same below. B. SDS-PAGE analysis of expression of TUT4_NFs. a-c indicates different colonies.

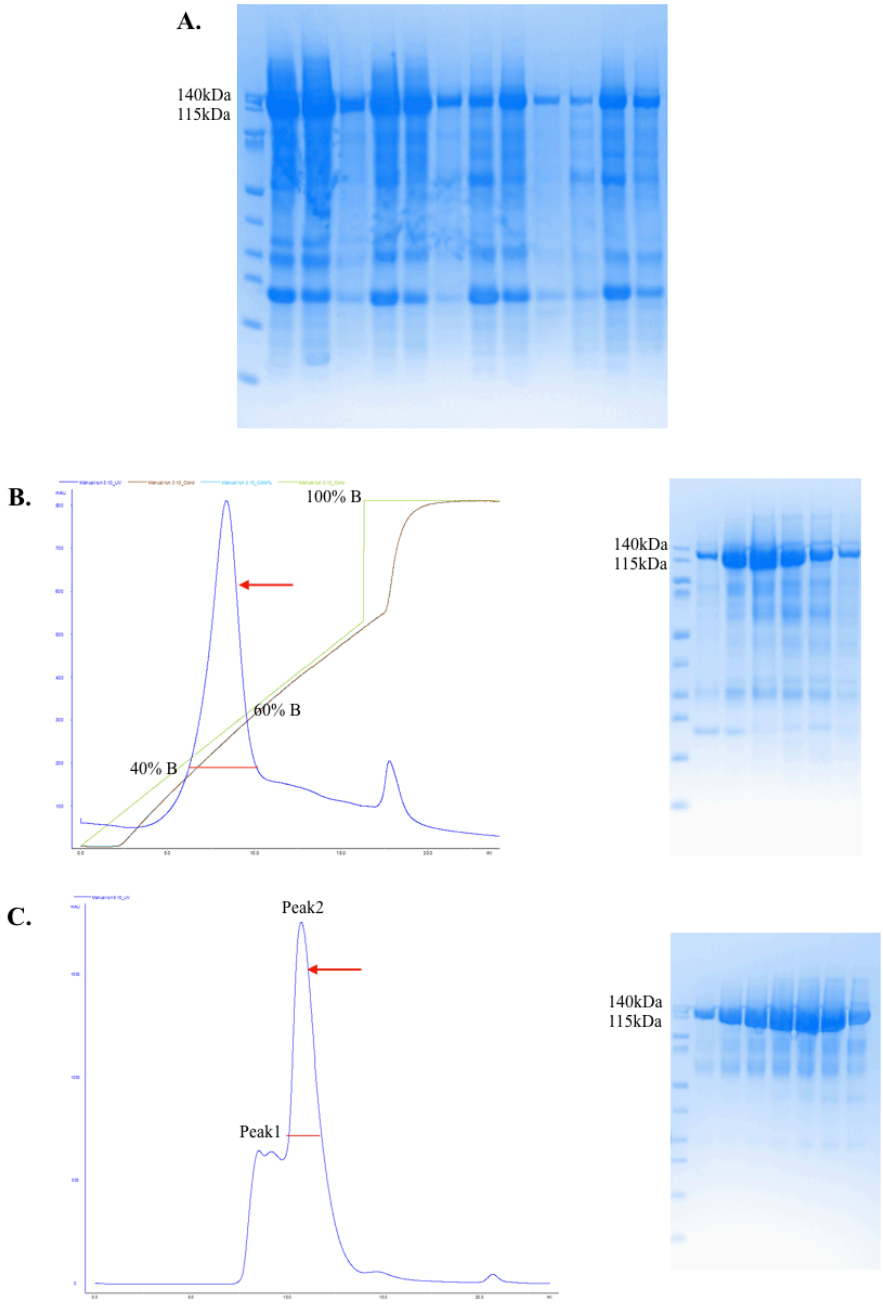


Figure 4.7 Purification of TUT4 mini full length (TUT4_miniF) A. SDS-PAGE analysis of recombinant TUT4_miniF after IMAC using Strep resins. B. Ion exchange chromatography profiles and SDS-PAGE analysis of TUT4_miniF. Arrow indicates the correct peak and red line cuts off the fractions used for SDS-PAGE analysis. Green curve and mud yellow curve show the concentration of buffer B and salt respectively. C. Size exclusion chromatography profiles and SDS-PAGE analysis of TUT4_miniF. Peak1 shows the aggregates volume and peak2 contains the soluble TUT4_miniF.

4.2.6 Purification of TUT4_F

TUT4_F gene was amplified by PCR and sub-cloned into pHGT-bio vector with a TEV cleavable 6x His GST tag at the N-terminus and a single Strep-tag at the C-terminus, which was initially purified using the two-step purification protocol detailed elsewhere. However, I failed to obtain useful amounts of purified protein due to the extremely low yields and serious degradation issues. TUT4_F has a greater molecular size and is more flexible than TUT7_F, potentially accounting for the expression issues of low yields and serious degradation with multiple bands in SDS-PAGE gel as determined by Coomassie blue (not shown). Although I elevated the purification scale to 48-liter TB medium, I still could not get enough protein. Optimization was attempted aiming to obtain a high quality TUT4_F preparation. The key modification was to add a twin-Strep-tag to the C-terminus of TUT4, which enhances the binding affinity of TUT4_F to streptavidin resins greatly, and also to decrease the concentration of NaCl to 150 mM. Considering there was a small amount of biotin in *E.coli* and its presence affected the binding process of TUT4_F to streptavidin resins, the biotin blocking solution (BioLock, IBA) was added into the lysate and incubated for one hour before centrifugation. Then after eluting from streptavidin resins, the eluted fractions were pooled together and concentrated followed by size exclusion chromatography to remove most of the aggregates (Figure 4.8A-B). The soluble peaks were examined by SDS-PAGE and pooled together to perform anion exchange chromatography via a heparin column. This step was to remove contaminated proteins and most of the nucleic acid contamination. The pure fractions confirmed by SDS-PAGE were pooled together and concentrated to perform size exclusion chromatography (Figure 4.8C). The series of chromatography steps improved the purity of TUT4_F despite decreasing the final yield obtainable. In total, several batches of purified TUT4_F were combined together for final size exclusion chromatography (Figure 4.8D). 1 mg TUT4_F was obtained with 85% purity from 48-liter TB medium after size exclusion chromatography and concentrated to 1 mg/ml. The concentrated protein was aliquoted and flash-frozen by liquid nitrogen followed by storing in a -80 °C freezer for subsequent assay and cryo-EM trials.

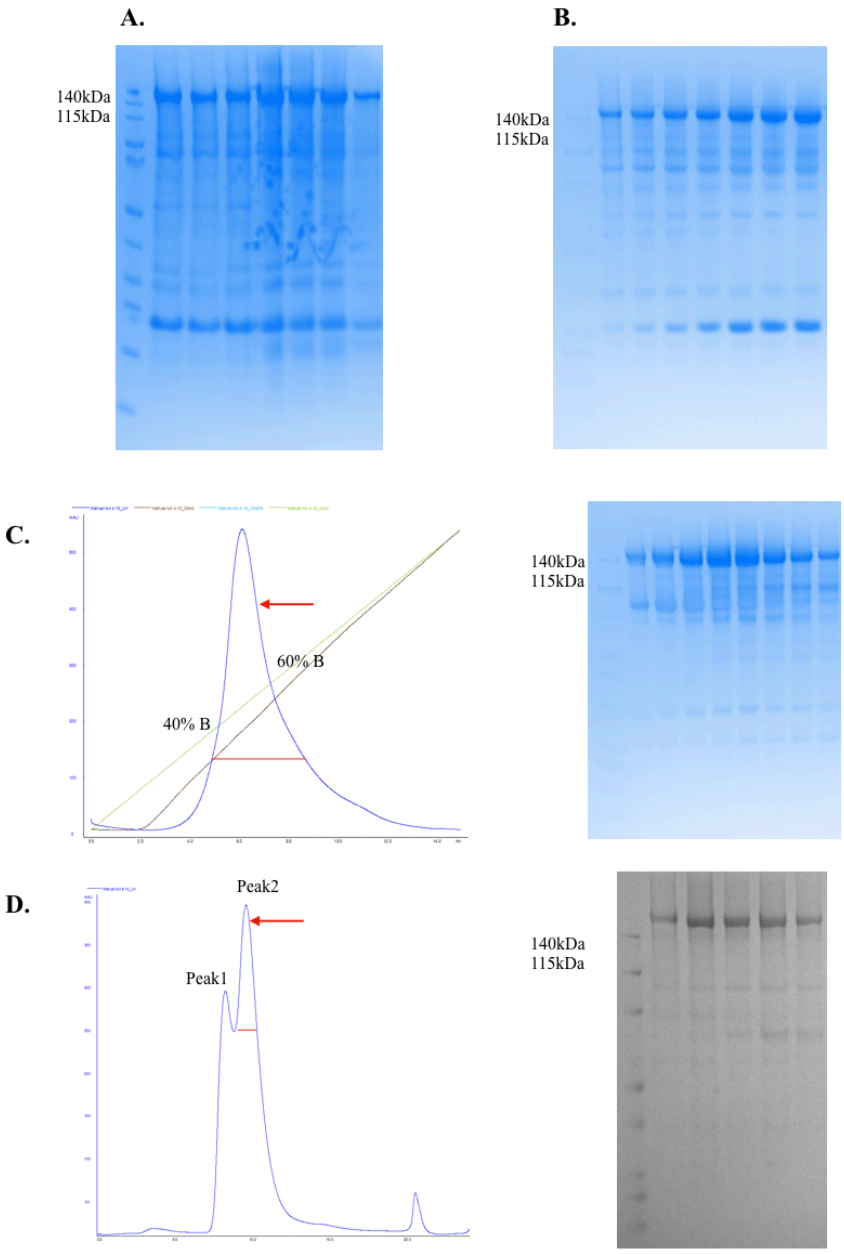


Figure 4.8 Purification of TUT4 full length A. SDS-PAGE analysis of recombinant TUT4_F eluted from IMAC using Strep resins. B. SDS-PAGE analysis of recombinant TUT4_F fractions from size exclusion chromatography. C. Ion exchange chromatography profiles and SDS-PAGE analysis of TUT4_F. Arrow indicates the correct peak and red line cuts off the fractions used for SDS-PAGE analysis, same below. Green curve and mud yellow curve show the concentration of buffer B and salt respectively. D. Size exclusion chromatography profiles and SDS-PAGE analysis of TUT4_F. Peak1 shows the aggregates volume and peak2 contains the soluble TUT4_F.

4.2.7 Assays of TUT4

TUT4_miniF was catalytically active as confirmed by biochemical assays referred to in the protocol at 2.9 in Chapter 2. TUT4_miniF can only add oligo uridines to 3' of pre-let-7 in the presence of Lin28A (Figure 3.11). Catalytic dead mutant TUT4_miniF (TUT7_FD), purified using similar purification protocol of TUT4_miniF (not shown), presented neither mono-uridylation activity nor oligo-uridylation activity (Figure 3.11). The active TUT4_miniF was thus qualified for subsequent cryo-EM trials.

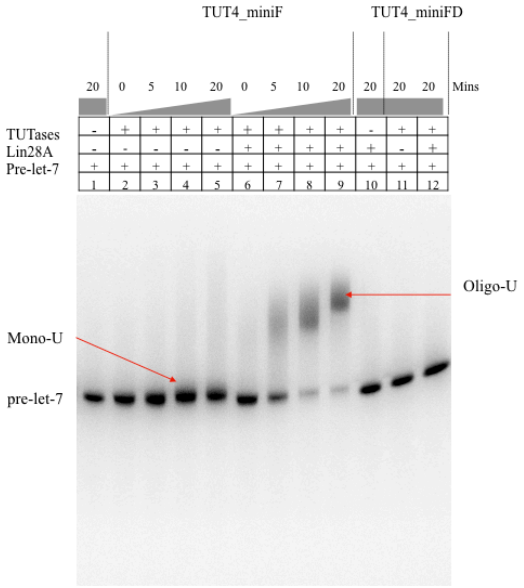


Figure 4.9 Uridylation of pre-let-7 by TUT4_minF in vitro. Mono-uridylation of pre-let-7 by TUT4_miniF in the absence of Lin28A at timepoints: 0, 5, 10, 20 mins presenting in lanes 2-5. Oligo-uridylation of pre-let-7 by TUT4_miniF in the presence of Lin28A at timepoints: 0, 5, 10, 20 mins presenting in lanes 6-9. Lane 1 and 10 show negative control with or without Lin28A. Lanes 11-12 present the uridylation of pre-let-7 by TUT4_miniFD in the absence and presence of Lin28A.The reaction products were separated by 15% (w/v) polyacrylamide gel electrophoresis under denaturing conditions.

4.2.8 Assembly of TUT4/pre-pre-let 7/Lin28A

The assembly of TUT4 with pre-let-7 with or without Lin28A uses the same protocol referred to in 3.2.8. Generally, the TUT4_miniF or TUT4_F were incubated with the folded pre-let-7 overnight plus 100 μM UTP α S and 2 mM $CaCl_2$. The second day, the mixed sample was concentrated for size exclusion chromatography or incubated with Lin28A for 3 hours followed by size exclusion chromatography. The homogeneous peak fractions of apo TUT4_miniF/ or TUT4_F, TUT4_miniF/pre-let-7 or TUT4_F/pre-let-7, and TUT4_miniF/pre-let-7/Lin28A or TUT4_F/pre-let-7/Lin28A were examined by SDS-PAGE (Figure 4.10A-C and E) (Figure 4.11A-C). Good quality fractions were selected for subsequent assays and cryo-EM trials.

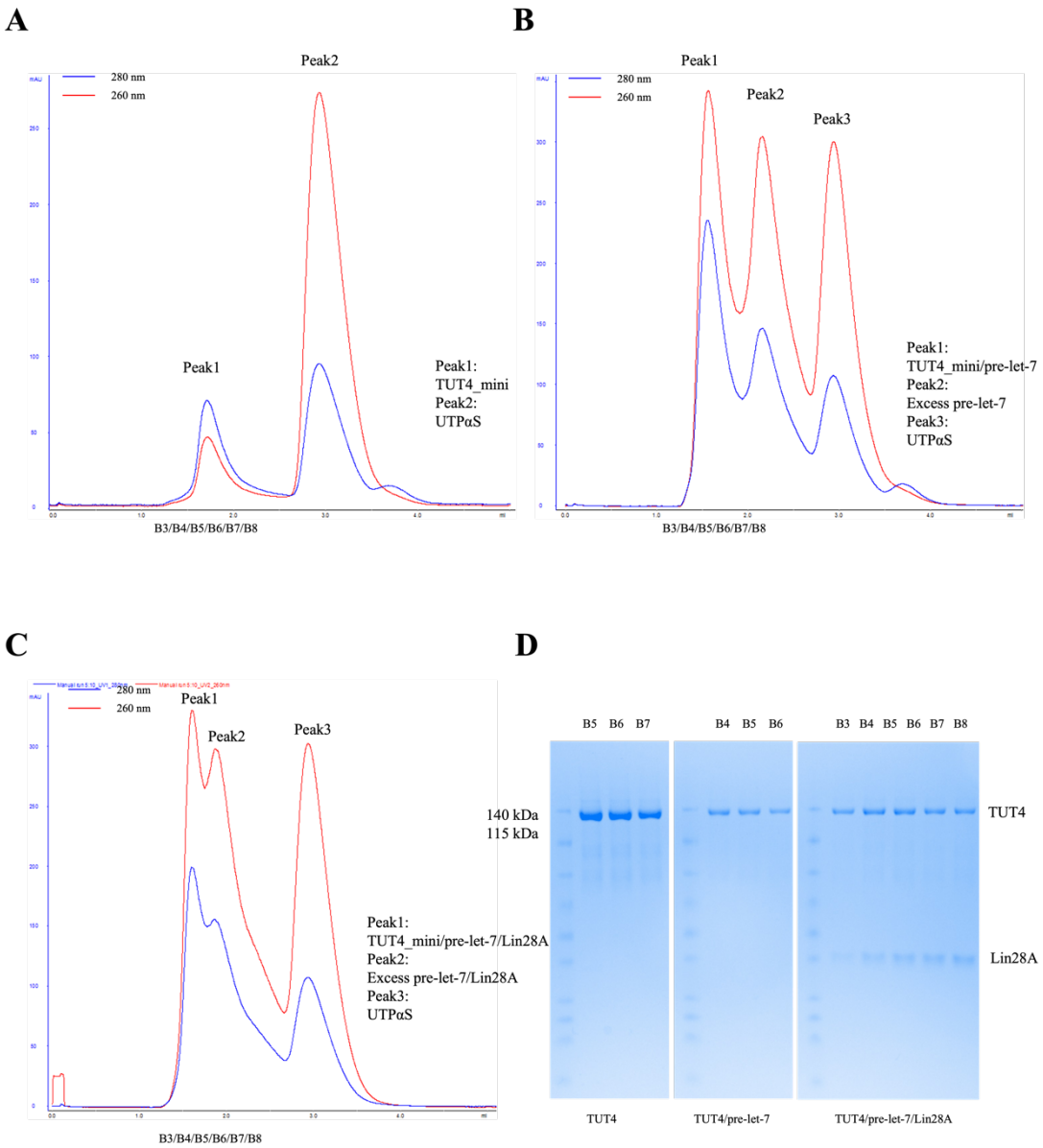


Figure 4.10 Assembly of TUT4_miniF with pre-let-7 with or without Lin28A A. Size exclusion chromatography profile of apo TUT4_miniF in the presence of UTPαS. B. Size exclusion chromatography profile of TUT4_miniF/pre-let-7 in the presence of UTPαS. C. Size exclusion chromatography profile of TUT4_miniF/pre-let-7/Lin28A in the presence of UTPαS. The elution volumes of peak1 for apo TUT4_miniF, TUT4_miniF/pre-let-7 and TUT4_miniF/pre-let-7/Lin28A are 1.71ml, 1.66ml, 1.62ml respectively. D. SDS-PAGE analysis of apo TUT4_miniF, TUT4_miniF /pre-let-7 and TUT4_miniF/pre-let-7/Lin28A from size exclusion chromatography.

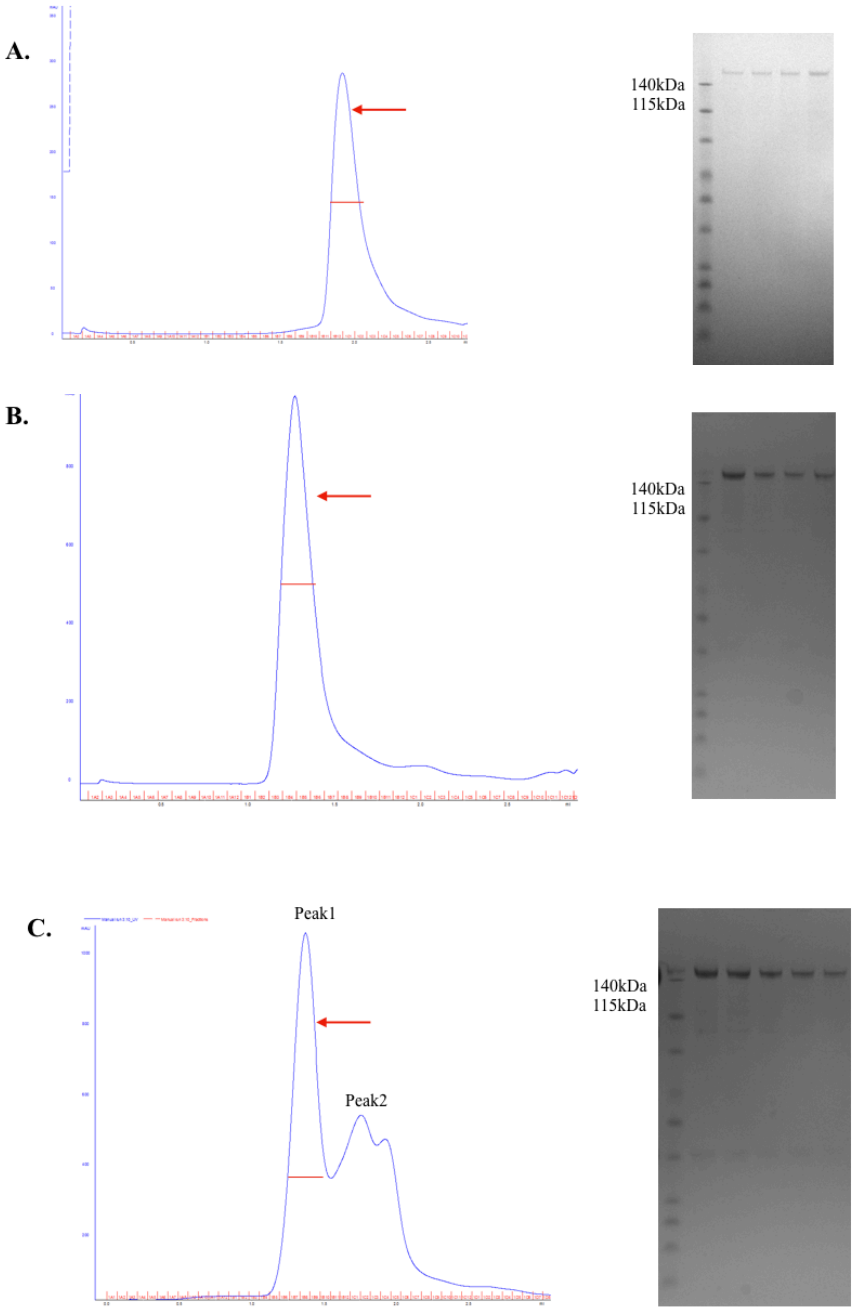


Figure 4.11 Assembly of TUT4_F with pre-let-7 with or without Lin28A A. Size exclusion chromatography profiles and SDS-PAGE analysis of apo TUT4_F. Arrow indicates the correct peak and red line cuts off the fractions used for SDS-PAGE analysis, same below. B. Size exclusion chromatography profiles and SDS-PAGE analysis of TUT4_F with pre-let-7. C. Size exclusion chromatography profiles and SDS-PAGE analysis of TUT4_F with pre-let-7 and Lin28A. The peak1 contains the TUT4_F/Lin28A/pre-let-7 ternary complex and the peak2 shows the excess Lin28A/pre-let-7.

4.2.9 Assembly of TUT4 with Nanobodies: S2A4

The TUT4 structure in complex with S2A4 should reveal its inhibition mechanism, which will be of further value in understanding TUT4/7 mechanism and druggability. To solve the complex structure, TUT4 and S2A4 complex samples was prepared for cryo-EM single particle analysis.

TUT4_miniF was incubated with S2A4 at a ratio of 1:1.5 on ice for an hour. The mixed sample was concentrated, with which size exclusion chromatography was performed after spinning down aggregates at 15,000 g, 4 °C for 10 minutes. The homogenous peak fractions were examined by SDS-PAGE (Figure 4.12 A-B) and flash-frozen with liquid nitrogen followed by storing in a -80 °C freezer for the subsequent cryo-EM trials.

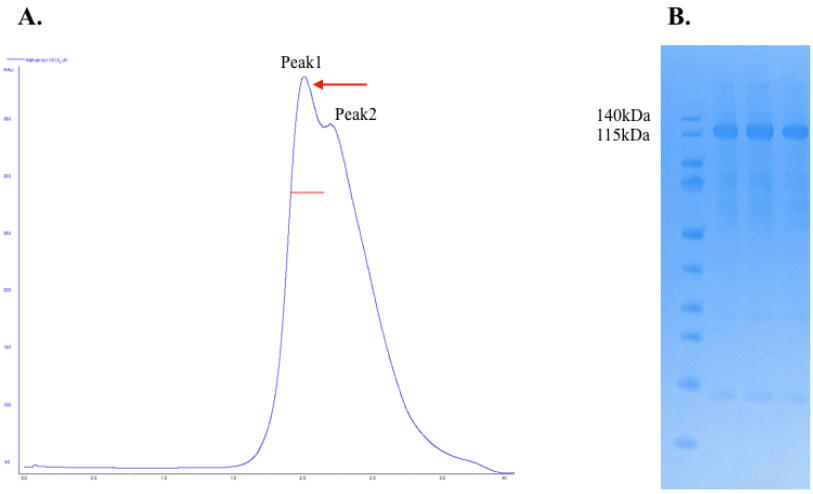


Figure 4.12 Assembly of TUT4_minniF with S2A4. Size exclusion chromatography profiles and SDS-APGE analysis of TUT4_miniF with Nanobody-S2A4. Arrow indicates the correct peak and red line cuts off the fractions used for SDS-PAGE analysis. Peak1 contains the TUT4_miniF/S2A4 and peak2 presents the excess TUT4_miniF.

4.2.10 Negative staining of TUT4

Negative stain EM trials were performed to check the quality of TUT4_miniF protein and images collected for initial data processing. Around 10,000 particles were picked manually to generate 2D class averages (Figure 4.13A-B). The gross features of TUT4_miniF appear the same as for TUT7_F. The TUT4_miniF appeared to have a high degree of homogeneity and was a good candidate for further assessment by advanced EM with Talos and Krios microscopes. I purified the TUT4_F recombinant protein more recently in a large-scale purification. Negative staining was performed to identify its quality and provide the rough features of TUT4_F. Single particles were observed and around 15,000 single particles were extracted to generate the 2D class averages (Figure 4.13C-D). The negative staining results suggested good subsequent performance under cryo-EM analysis.

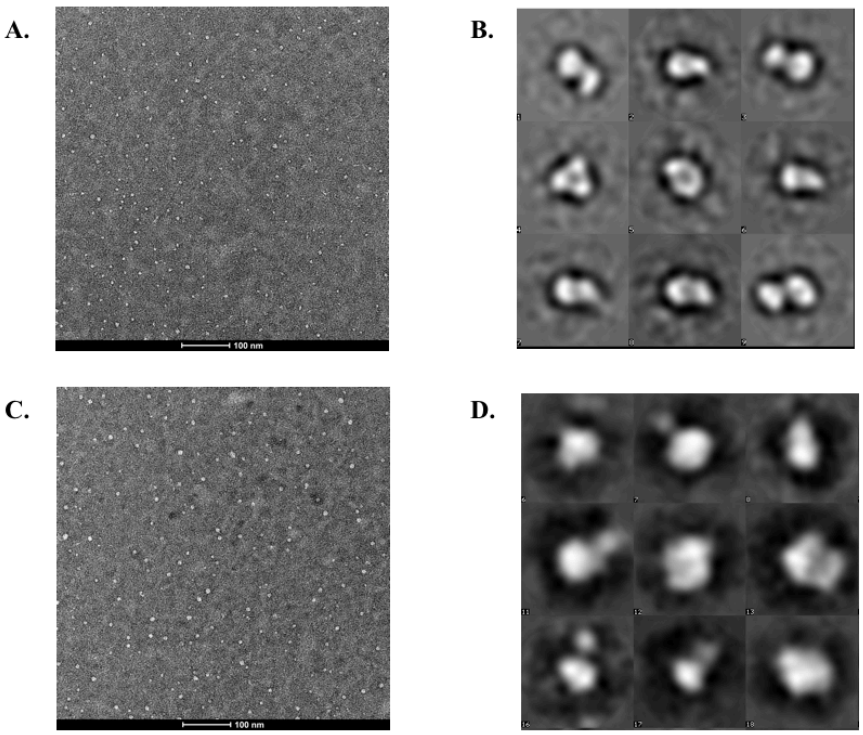


Figure 4.13 Negative staining of TUT4miniF and TUT4F A. Representative negative stain micrograph of TUT4_miniF. B. Gallery of reference-free 2D class averages of TUT4_miniF. C. Representative negative stain micrograph of TUT4_F. D. Gallery of reference-free 2D class averages of TUT4_F.

4.2.11 Grafix for assembling the ternary complex

To obtain complete ternary complex of TUT4_F with pre-let-7 and Lin28A, cross-linking with glutaraldehyde was introduced to fix the dynamic ternary complex. A gradient of glycerol and glutaraldehyde was prepared according to the molecular weight of target proteins and their complexes (Stark, 2010) (Figure 4.14A). Buffer A contained 20 mM HEPES (pH7.8), 100 mM NaCl, 2 mM $CaCl_2$, 2 mM DTT, 5% Glycerol and buffer B consists of 20 mM HEPES (pH7.8), 100 mM NaCl, 2 mM $CaCl_2$, 2 mM DTT, 30 % Glycerol plus 0.2 % glutaraldehyde. The gradient solution was made using a Gradient Master 108 with the setting of 15°, 2 mins in a total volume of 15 ml. Finally, 200 µl concentrated ternary complex, which was pre-assembled, concentrated and centrifuged to remove aggregates, was added into the tube containing the gradient solution. Ultracentrifugation was performed with the following settings: 40,000g, 4 °C for 13 hours. The next day, tubes were taken out from the centrifuge and the equilibrated solution was removed manually from the top layer to the bottom layer into 1.5 ml EP tubes, in 500 µl aliquots. Then, 50 µl buffer C (1 M Tris, pH 8.0, 100 mM NaCl, 2 mM $CaCl_2$, 2 mM DTT) was added to quench the reaction. All fractions were examined by SDS-PAGE (Figure 4.14B). The TUT4_F ternary complex samples were separated well as observed at and represented by different bands (Figure 4.14B). The fractions with proteins of expected molecular weight were selected for subsequent cryo-EM trials. Both grids made by crosslinked or non-crosslinked samples respectively were imaged on a Talos microscope. Compared to the complexes without crosslinking, those with had better contrast and distribution obvious in the raw images. A small dataset was collected, which led to good 2D classes. Further images were then collected on a Krios microscope with 3,000 movies captured, followed by processing using Cryosparc (Punjani et al., 2017; Scheres, 2012) (Figure 4.14C-D). However, poor alignment of particles was observed, causing blurring of the 2D class averages, and therefore high-resolution details could not be resolved. We hypothesized this was due to the nonspecific crosslinking for the ternary complex leading to a series of mixed states, which could not be separated well during the data processing. Other strategies like PEGylation of the ternary complex and graphene grids are currently under investigations.

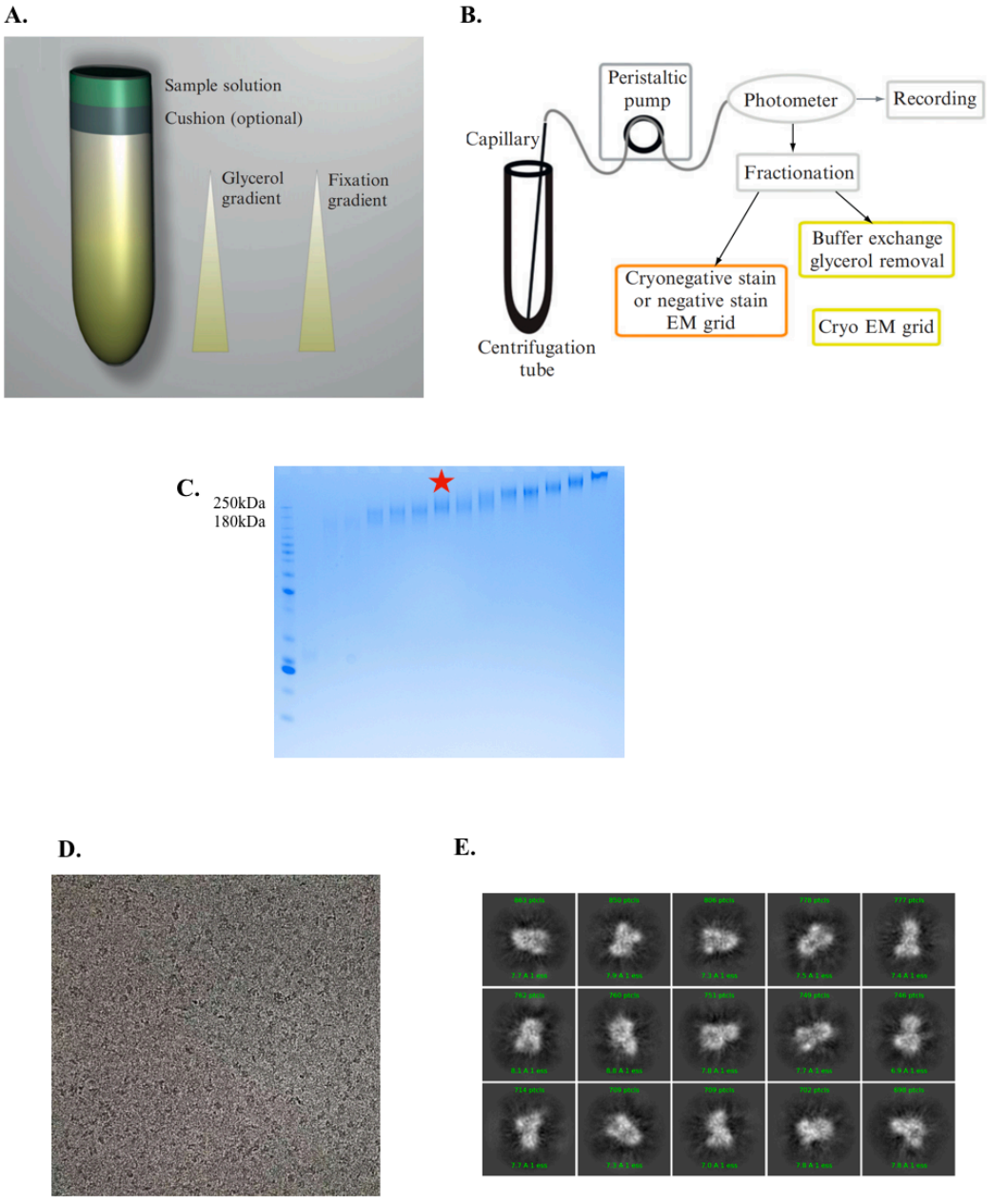


Figure 4.14 Grafix of TUT4_F with pre-let-7 and Lin28A A-B. Principle and workflow of Grafix. C. SDS-PAGE analysis of cross-linking TUT4_F/pre-let-7/Lin28A fractions from Grafix. D. Representative raw micrograph of TUT4_F ternary complex. E. Gallery of reference-free 2D class averages of TUT4_F ternary complex.

4.2.12 Cryo-EM structures of TUT4 from Krios microscope

To resolve the structure of TUT4 complexed with pre-let-7g and Lin28A, we therefore proceeded with analysis of the directly assembled complex without fixation. Good 2D class averages could be obtained from the sample. Ultimately, two structures were determined at 3.7 Å (Figure 4.15 A-E) and 12 Å (not shown), respectively. The crystal structure of TUT4_LIM, zinc domain of Lin28A and predicted pre-let-7g structure were docked into the cryo-EM maps, before multiple rounds of real-space refinement in Phenix (Adams et al., 2002) and manual rebuilding in Coot (Emsley & Cowtan, 2004; Goddard et al., 2005; Lam & Siu, 2017). The two conformations obtained show slight differences in the region of the PRE loop of pre-let-7, in which the lower resolution structure appears to show the CSD domain of Lin28A binding to the PRE loop of the pre-let7-g. The overall structure is similar to the TUT7 ternary complex (in conformation-I, capture complex) and like that structure both of the TUT4 ternary complexes lack density for the TUT4_CM.

The TUT4_miniF and TUT4_miniF/pre-let-7 were imaged on a 200 kV Talos (not shown). The processed 2D classification results of apo TUT4_minF show its similarity with TUT7 apo, in which two blobs are connected by flexible parts. The pre-let-7 stem was visualized in the binary complex of TUT4_miniF and pre-let-7 from the 2D classification results. The two samples are scheduled for data collection on the Krios microscope aiming to obtain high-resolution structures.

Finally, the serious degradation issues of the flexible TUT4 full length protein were solved. The fractions with good quality ternary complex assembly were selected for cryo-EM trials. The 2D classification results show different 2D class averages compared with both TUT4_miniF ternary complex and TUT7 ternary complex. This conclusion was confirmed by the low-resolution model, as shown in Figure 4.16A-E. The TUT4_F ternary complex is similar to the TUT7 ternary complex (in conformation-II, the engaged complex) but shows additional densities compared to it. These results indicate the urgency for obtaining a high-resolution TUT4_F/pre-let-7/Lin28A ternary complex structure, which would show us more detailed information especially for the unidentified densities.

The functional nanobody, S2A4 was selected to stabilize the flexible TUT4, for both TUT4_miniF and TUT4_F (not shown). However, no good 2D class averages were obtained from the dataset collected on the 200 kV Talos. This seems reasonable since the S2A4 was reported to bind the dynamic zinc finger of LIM so as to block the ternary interaction. I also applied another nanobody, CB11 to the EM trials which would not disrupt the interaction at low concentration. We hypothesized this nanobody may stabilize the ternary complex. However, worse 2D class averages were obtained compared to the sample without CB11 addition. More biochemistry and cryo-EM work for this complex will be explored in the near future.

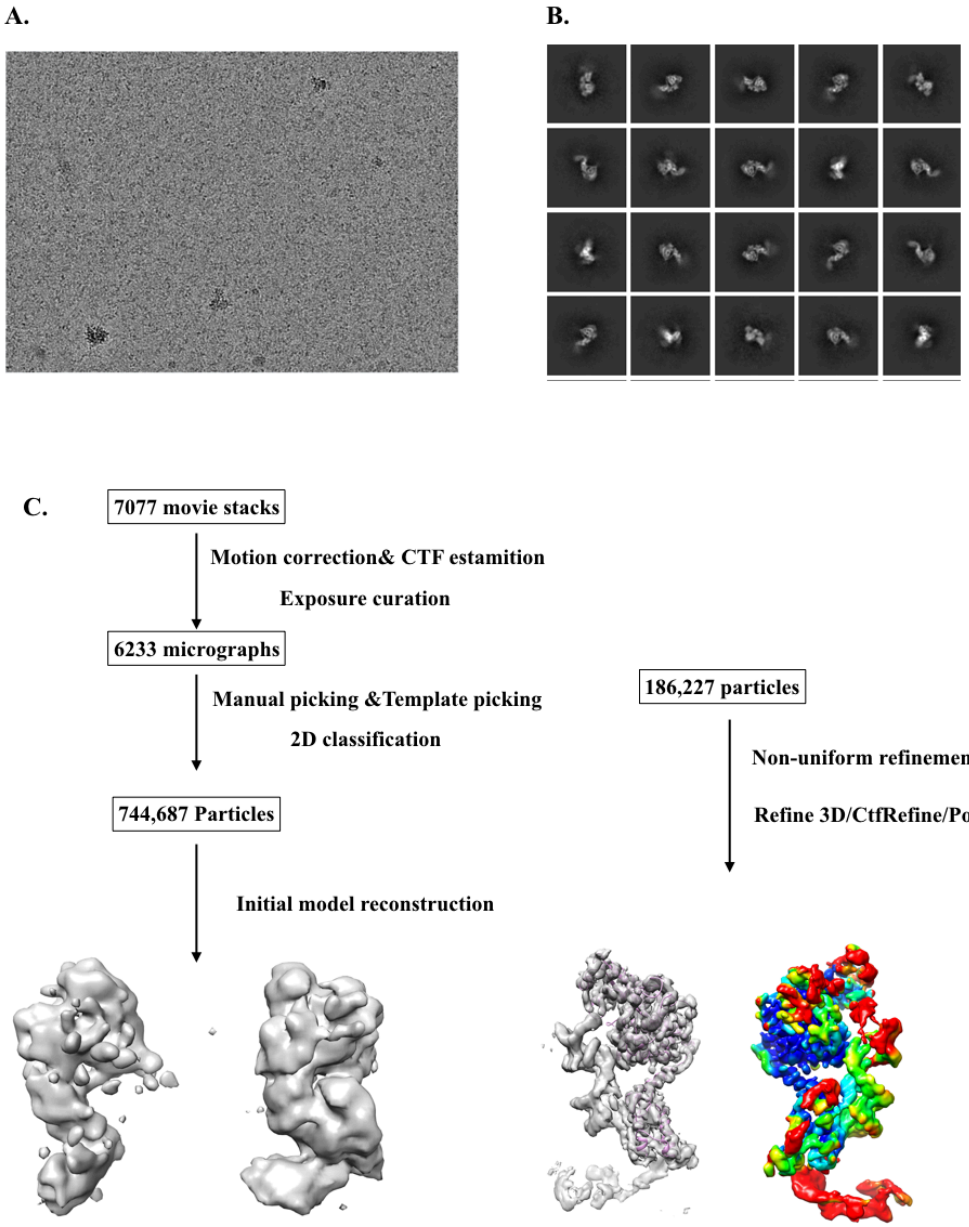


Figure 4.15: Legend on next page

Figure 4.15 Cryo-EM image processing workflow of TUT4_miniF with pre-let-7 and Lin28A A. Representative raw micrograph of TUT4_miniF/pre-let-7/Lin28A complex. B. Gallery of reference-free 2D class averages of TUT4_miniF/pre-let-7/Lin28A complex. C. Data processing workflow of T TUT4_miniF/pre-let-7/Lin28A complex. D. Gold standard FSC curves for TUT4_miniF/pre-let-7/Lin28A complex reconstructions. E. Orientation distribution heat map of TUT4_miniF/pre-let-7/Lin28A complex.

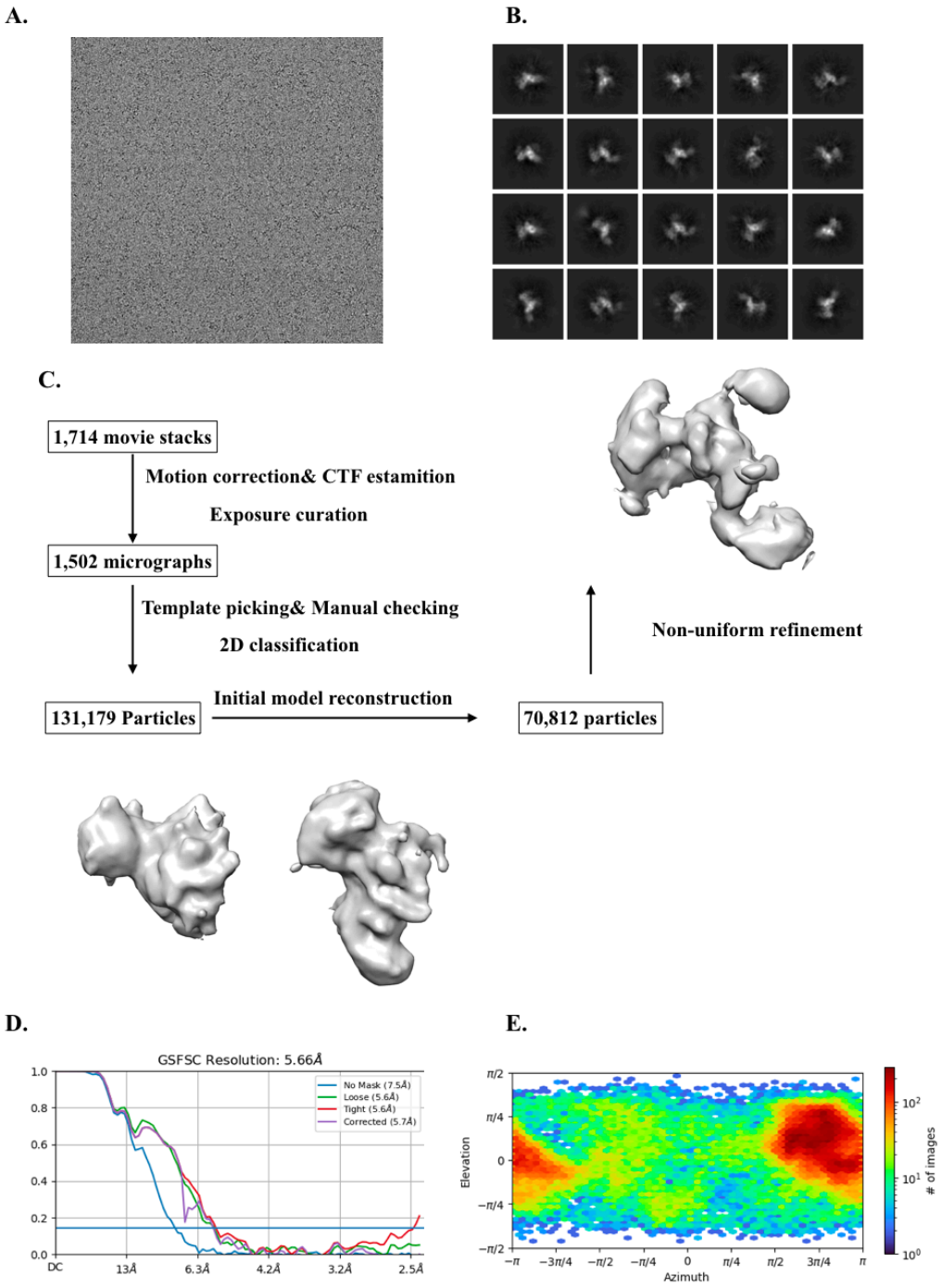


Figure 4.16 Cryo-EM image processing workflow of TUT4_F with pre-let-7 and Lin28A A. Representative negative staining micrograph of TUT4_F/pre-let-7/Lin28A complex. B. Gallery of reference-free 2D class averages of TUT4_F/pre-let-7/Lin28A complex. C. Data processing workflow of T TUT4_F/pre-let-7/Lin28A complex. D. Gold standard FSC curves for TUT4_F/pre-let-7/Lin28A complex reconstructions. E. Orientational distribution heat map of TUT4_F/pre-let-7/Lin28A complex.

4.3 Conclusion

Purified samples of the challenging TUT4 enzymes, including in both full length and truncated forms were successfully obtained via different purification methods. These expressed forms aided structural studies by X-ray crystallography and cryo-EM. Both the TUT4_CM and TUT4 LIM in complex with nanobody S2A4 diffract to 4.5 Å and 5 Å, respectively. Optimized trials are under investigation aiming for high-resolution structure determination. On the cryo-EM side, the successful determination of a truncated TUT4/pre-let-7/Lin28A ternary complex enables us to visualize the interaction interface of proteins and RNA at a residue level. Other cryo-EM trials, such as TUT4 full length with pre-let-7/Lin28A, show regions of density not observed in other ternary complexes, indicating the urgent need for the determination of this complex in more detail, and this is under investigation currently.

TUT4 and TUT7 have different distributions and expression levels in diverse cells and tissues (Heo et al., 2008; Kim et al., 2015; Lim et al., 2014). Although the structures of TUT7 in diverse states has been well studied already in work described in Chapter 3, TUT4 would be worth exploring in complex with pre-let-7g in forms containing diverse U-tails at their terminus to compare with complex structures of TUT7 for a better understanding of their uridylation mechanism. Clearly, structural methods will continue to play an important role in exploring for both TUT7 and TUT4 how the regulation of pre-let-7 maturation occurs.

Chapter 5 Structural analysis of TUTases

5.1 Introduction

Group II pre-let-7 microRNAs require the addition of a mono-uridine to their 3' end by TUTases TUT4 or TUT7, and subsequently can be efficiently processed by Dicer into mature let-7 miRNAs (Heo et al., 2012). However, in the presence of Lin28A RNA binding protein, both TUT4 and TUT7 add oligo-uridine tracts to the 3' end of a bound pre-let-7, whether of group I (already in possession of a 2-nt 3' overhang) or group II. An exonuclease Dis3l2, responsible for poly-U RNA degradation was confirmed to participate in the degradation of the oligo-uridylated pre-let-7 miRNAs, leading to a low expression level of let-7 (Chang et al., 2013; Faehnle et al., 2014; Ustianenko et al., 2013). To date, the mechanism for a switch from mono- to oligo-uridylation has remained unknown, although it has been investigated both in vivo and vitro and also been argued to involve a move from a distributive enzymatic mechanism to a processive one (Heo et al., 2012; Piskounova et al., 2011; Thornton et al., 2012; Yates et al., 2015).

To better understand the uridylation mechanism involved in the biogenesis of mature let-7, I used single-particle cryo-EM to study the structures of the TUTases TUT4 and TUT7 with or without pre-let-7 miRNA in the presence or absence of Lin28A (Chapters 3 and 4). The recombinant wild-type human TUT7 full-length and human Lin28A were purified from *E.coli*. The pre-let-7g was synthesized as the substrate for structural studies, and this pre-miRNA transcript has been broadly used in previous studies of TUT4/7 activity (Faehnle et al., 2017a; Yu et al., 2020). To reconstitute the TUT7/Lin28A/pre-let-7 holo ternary complex, the recombinant TUT7 was mixed with Lin28A and pre-let-7 and the complex then purified by gel-filtration chromatography. By a similar method, the TUT7/pre-let-7 binary complex was obtained. Using the assembled sample complexes, I embedded the TUT7 particles in diverse combinations in vitreous ice with good contrast and solved a series of cryo-EM structures of TUT7 in different functionally related states and with resolution ranging from 3.6 Å to 4 Å.

To compare with TUT7, a truncated TUT4 $^{254\text{-}1315}$, which is named TUT4_miniF in previous chapters (TUT4 presented in this chapter indicates this truncated form), was purified from *E.coli* and assembled with Lin28A and pre-let-7g before polishing by size-exclusion

chromatography. I succeeded in reconstructing two models with slight differences in the PRE-element region of the let7-g pre-miRNA. A high-resolution ternary complex was solved at 3.7 Å and was similar to the TUT7 ternary complex in a conformation (conformation-I, capture complex), in which the enzyme' s catalytic module (CM) was flexible and unresolved, unengaged with its 3' pre-miRNA target. Whereas a low-resolution module showed additional densities in the pre-element regions, to which the cold-shock domain (CSD) of Lin28A is expected to bind, and different from the binding of the Lin28A twin CCHC zinc fingers observed in the TUT7 ternary complexes in both conformation-I and conformation-II (in which the TUT7 CM is engaged and resolved at the 3' end of the pre-miRNA, an "engaged" complex).

Thus, in all, four structures determined are of TUT7; alone (apo), bound to pre-let-7g miRNA (binary complex), and in two states with both the pre-let-7g miRNA and Lin28A. The fifth structure is of human TUT4 bound with pre-let-7g and Lin28A. The cryo-EM data collection and refinement statistics are reported in the Appendix. Together, the structures I describe reveal the basis for mono- or oligo-uridylation via alternative enzymatic conformations and protein-RNA interfaces, induced either by the selective binding of a pre-let-7 miRNA, or by the initial binding of the pre-let-7 miRNA to Lin28A, followed by engagement of both the pre-miRNA and Lin28A itself directly with TUT7/4.

5.2. Results

5.2.1. Cryo-EM structure determination of apo TUT7 and TUT7/4 complexes

As detailed previously, TUT4 and TUT7 are multi-domain enzymes whose specific targeting of RNAs makes use of three zinc fingers, combined with two copies of a nucleotidyltransferase domain (NTD), one inactive in the N-terminal half of the polypeptide (pseudo-NTD) and one active in the C-terminal half (Figure 5.1A-B) (Yates et al., 2013; Yu & Kim, 2020). Structures have previously been determined for the N-terminal Lin28A-interacting module (LIM) of TUT4 including its pseudo-NTD domain (Yamashita et al., 2019), and for the C-terminal catalytic module (CM) of TUT7 (Faehnle et al., 2017b), as well as for Lin28A bound with RNA constructs representing pre-let-7d/f/g miRNAs (Nam et al., 2011; Wang et al., 2017). The human TUT4 LIM (Yamashita et al., 2019) was found to have a very similar basic architecture to that of the TUT7 CM (Faehnle et al., 2017b). The TUT4 LIM structure was however distinguished by the presence of its zinc finger domain at the N-terminus which is homologous to those of double-stranded RNA-binding zinc-fingers (Yamashita et al., 2019). The TUT7 CM structure was previously solved crystallographically in apo and complex forms with dsRNA, U2, U5 or UTP (Faehnle et al., 2017b). In these CM structures, the second TUT zinc knuckle (ZK2) could be visualized bound with both U2 and U5. The diverse substrate complexes revealed that the RNA-binding pocket orients a group II pre-let-7 to favor mono-uridylation, while the ZK2 of TUT7 is thought to aid oligo-uridylation by supporting the extension of the oligo-U tail (Faehnle et al., 2017b).

Lin28, as a key regulator controlling the timing of developmental events and participating in heterochronic processes, is highly conserved across diverse organisms (Moss & Tang, 2003). In *C. elegans,* there is only one encoded Lin28 protein, while in humans there are two paralogs, Lin28A and Lin28B (Moss & Tang, 2003; Piskounova et al., 2011). Lin28 proteins are comprised of a cold-shock domain (CSD) and a zinc knuckle domain belonging to the zinc finger CCHC-type (ZCCHC) superfamily (Figure 5.1C).

As described, to better understand the uridylation mechanism involved in the biogenesis of mature let-7, single-particle cryo-EM has been used to study the structures of TUT4 and TUT7 without and with pre-let-7g miRNA (Figure 5.1D), in the absence or presence of Lin28A.

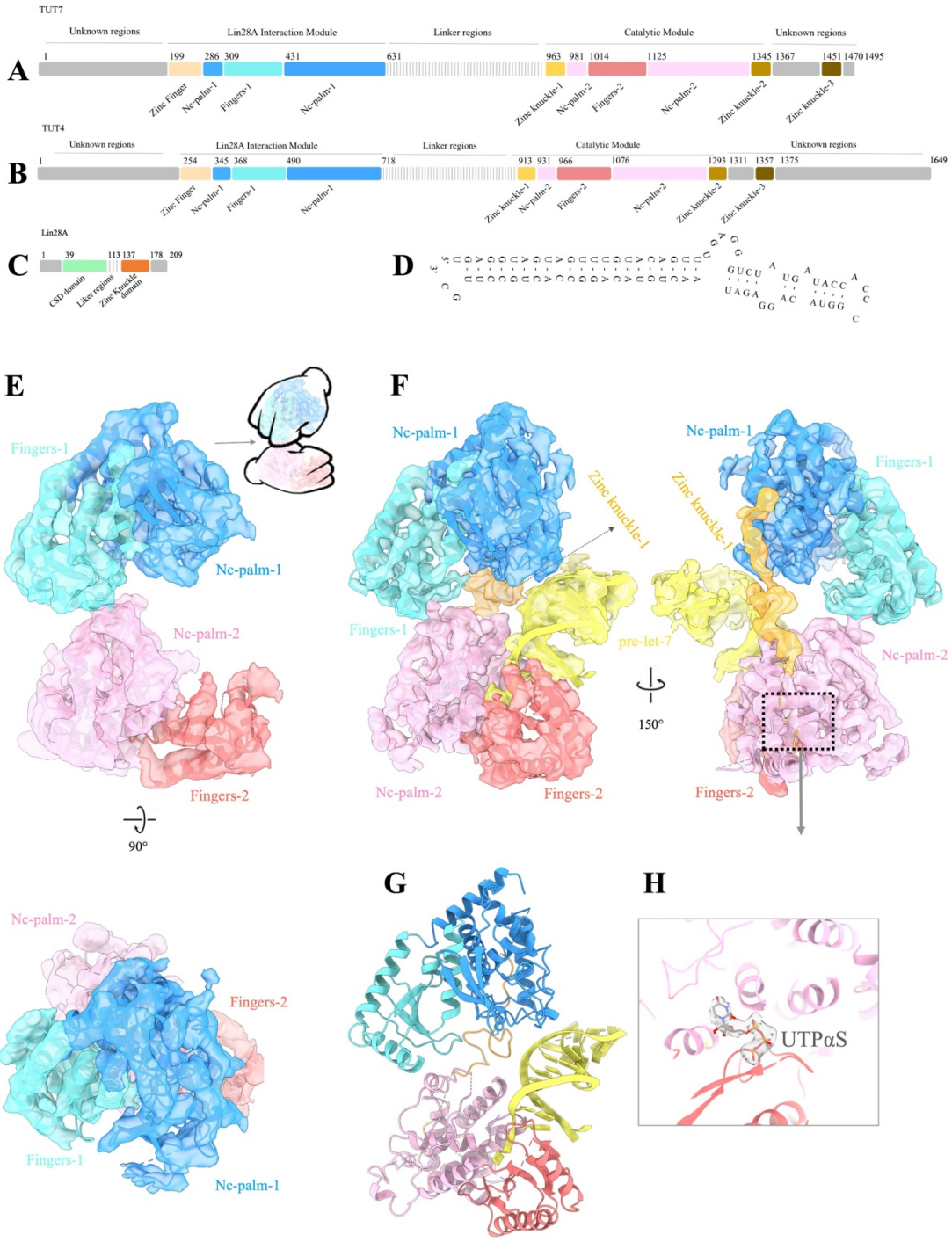


Figure 5.1: Legend on next page

Figure 5.1. Cryo-EM structures of TUT7 in apo and binary complex with pre-let-7. (A) Human TUT7 domain organization. (B) Human TUT4 domain organization. (C) Human Lin28A domain organization. (D) Nucleotide sequence of pre-let-7g, presented in a predicted secondary structure. (E) Overall structure of the TUT7 alone, shown as a 4 Å cryo-EM density map and cartoon representations in two views related by a 90° rotation, as indicated. A handed model of TUT7 apo is presented at the upper right corner.The canonical palm and fingers domains of the nucleotidyltransferase (NTD) canonical fold are colored as follows, N-terminal pseudo-NTD (from LIM) red and pink and C-terminal NTD (CM) blue and cyan. (F) Overall structure of the TUT7/pre-let-7g binary complex, shown as a 3.7 Å cryo-EM density map and cartoon representations of fitted coordinates and in two views related by a 150° rotation, as indicated. TUT7 colored as in panel E, zinc-knuckle 1 orange, pre-let-7g yellow. (G) Cartoon model of TUT7 in complex with pre-let-7g, highlighting the relative arrangements of domains in a C-shape with the pre-miRNA. (H) Zoomed-in view of UTPαS within the catalytic pocket.

5.2.2. Loose association between nucleotidyltransferase domains characterizes the TUT7 apo state

The TUT7 structure in an apo state at 4 Å revealed an overall "C-shape" conformation composed of two right-handed conformers (Figure 5.1E) distinct from the "L-shape" of human Dicer which in addition remains similar in its conformation throughout its activity cycle (Liu et al., 2018a). This difference suits Dicer's role as a "tape-measure" protein capable of recognizing the length associated with a correctly-generated pre-miRNA, hence its need for a fixed dimension, whereas the TUTs need to be able to function in two different modes (mono- and oligo-uridylation) and, further, to be able to act processively on a bound pre-miRNA substrate under the influence of Lin28A.

The two nucleotidyltransferase domains (NTDs) within TUT7 were both resolved in the apo map. Surprisingly, there was no obvious density for the backbone link connecting the LIM and CM, with only a small region of the linking density visible associated with the LIM (see later, Figure 5.3A), and several other regions were undetermined due to likely flexibility as well. Nevertheless, for maintenance of the distinctive "C-shape" structure, the two NTDs of the enzyme must interact directly, albeit relatively weakly. The interface between the LIM and CM is occupied by several non-polar and polar amino acids, indicating the likelihood of a hydrophobic lateral interaction, as observed also in Dicer through the RNase IIIa and IIIb domains (Liu et al., 2018a). This model shows us for the first time the overall "C-shape" of TUT7 (and we posit, TUT4) in a state prior to substrate binding.

5.2.3. TUT7 alone binds pre-let-7g miRNA via end-capture of the target 3' end

A 3.7 Å map of TUT7 in a binary complex with a pre-let-7g (Figure 5.1F-G) shows the RNA stem bound into the catalytic cleft of its CM, with the cytosine single nucleotide overhang at the pre-miRNA 3' end and UTPαS bound and stabilized by the critical residues; F1045, N1124, N1130, R1152, S1170, Y1171, H1286 (Figure 5.1H; Figure 5.3B) (H1286 is equivalent to H336 in the yeast homolog that is known to be critical for uracil selection (Lunde et al., 2012; Munoz-Tello et al., 2012; Yates et al., 2012)) . The let-7-g pre-element is oriented away from the CM and, lacking contact with ordered protein structure, was found to be relatively disordered, presumably due to some intrinsic flexibility. Compared to the above apo TUT7 model, a larger part of the TUT7 polypeptide region linking the LIM and CM, that includes its zinc knuckle-1 (ZK1) region, was resolved (Figure 5.3C) with a defined but dynamic interface between the LIM and CM domains (Movie1, Appendix). Interpretation of contacts made by TUT7 with the RNA stem was assisted by modeling of its ZK1, which fitted well into a clearly visualized but relatively low-resolution region of density (Figure 5.1F, Figure 5.3C). The ZK1 region had not been resolved in the previous structural study of the LIM alone (Faehnle et al., 2017b), which instead provided an atomic model for ZK2, a region of TUT7 which supports oligo-uridylation by engaging the extending U-tail (Faehnle et al., 2017b), but which is absent from our structure. The roles of all three TUT4/7 zinc knuckles were probed in a recent NMR-based study that made use of pre-let-7i as a target (Chaves-Arquero et al., 2021). While binding of TUT4 ZK2 and ZK3 to RNA was shown, it was concluded that ZK1 had lost its ability to interact with RNA due to two key mutations within its CCHC motif after the first cysteine and the single histidine, where the relevant amino acids were mutated from a variable bulky mostly hydrophobic residue to a serine (Chaves-Arquero et al., 2021). Alignment of the amino acids of the three zinc knuckles of TUT7 (Figure 5.3D) showed the residue after the first cysteine was serine, but the one after the histidine was leucine, different from the situation in TUT4 where it is another serine (Chaves-Arquero et al., 2021). Nevertheless, in the binary complex resolved, the lysine (K969) after the second TUT7 cysteine is shown to interact with the U5 and G6 of pre-let-7 using its sidechain (Figure 5.3E). This reveals that one of the functions for ZK1, at least in TUT7, is to recruit or stabilize pre-let-7 for mono-uridylation through a direct interaction with its linker loop.

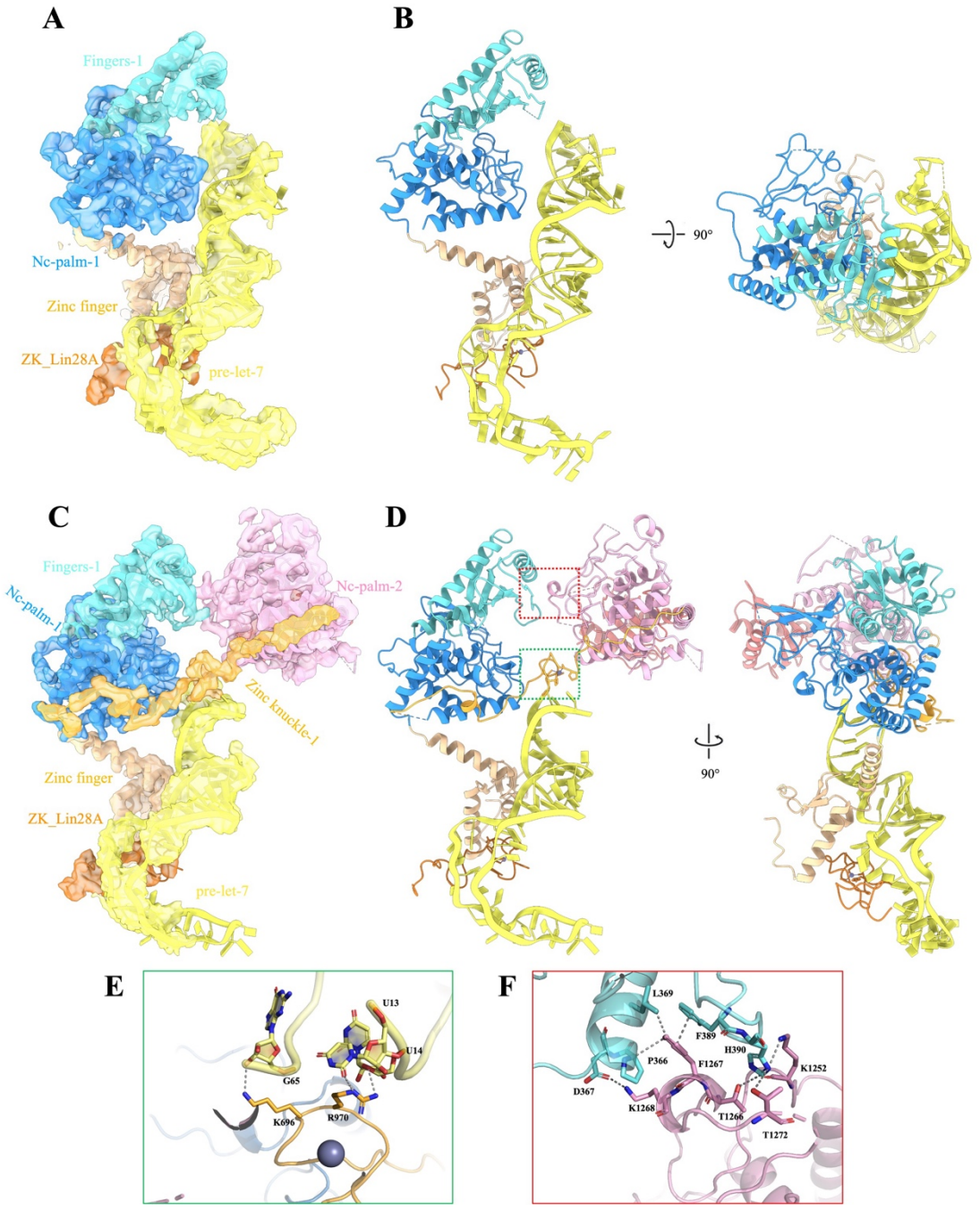


Figure 5.2: Legend on next page

Figure 5.2. Cryo-EM structures of TUT7 in complex with pre-let-7g and Lin28A. (A) Overall structure of the TUT7/pre-let-7g/Lin28A ternary complex conformation-I, shown as a 3.8 Å cryo-EM density map and cartoon representation (C-terminal NTD (CM) fingers blue and palm cyan; LIM zinc finger is colored sand, Lin28A zinc knuckle dark orange, pre-let-7g miRNA yellow. (B) Cartoon model of TUT7/pre-let-7g/Lin28A ternary complex conformation-I in two views related by a 90° rotation. Colors as in panel A. (C) Overall structure of the TUT7/pre-let-7g/Lin28A ternary complex conformation-II, shown as a 3.6 Å cryo-EM density map and cartoon representation. Colors as in panels A and B, plus TUT7 zinc knuckle 1 and linker region in bright orange, and pseudo-NTD from the LIM (fingers red and palm pink). (D) Cartoon model of TUT7/pre-let-7/Lin28A ternary complex-II in two views, related by a 90° rotation as indicated. Areas shown in expanded views below are boxed. (E) Zoomed-in view of TUT7 ZK1 bound to prelet-7. The black dotted lines link atoms within 3.8 Å of one another. (F) Zoomed-in view of domain interaction between TUT7 LIM and TUT7 CM. The black dotted lines link atoms within 3.8 Å of one another.

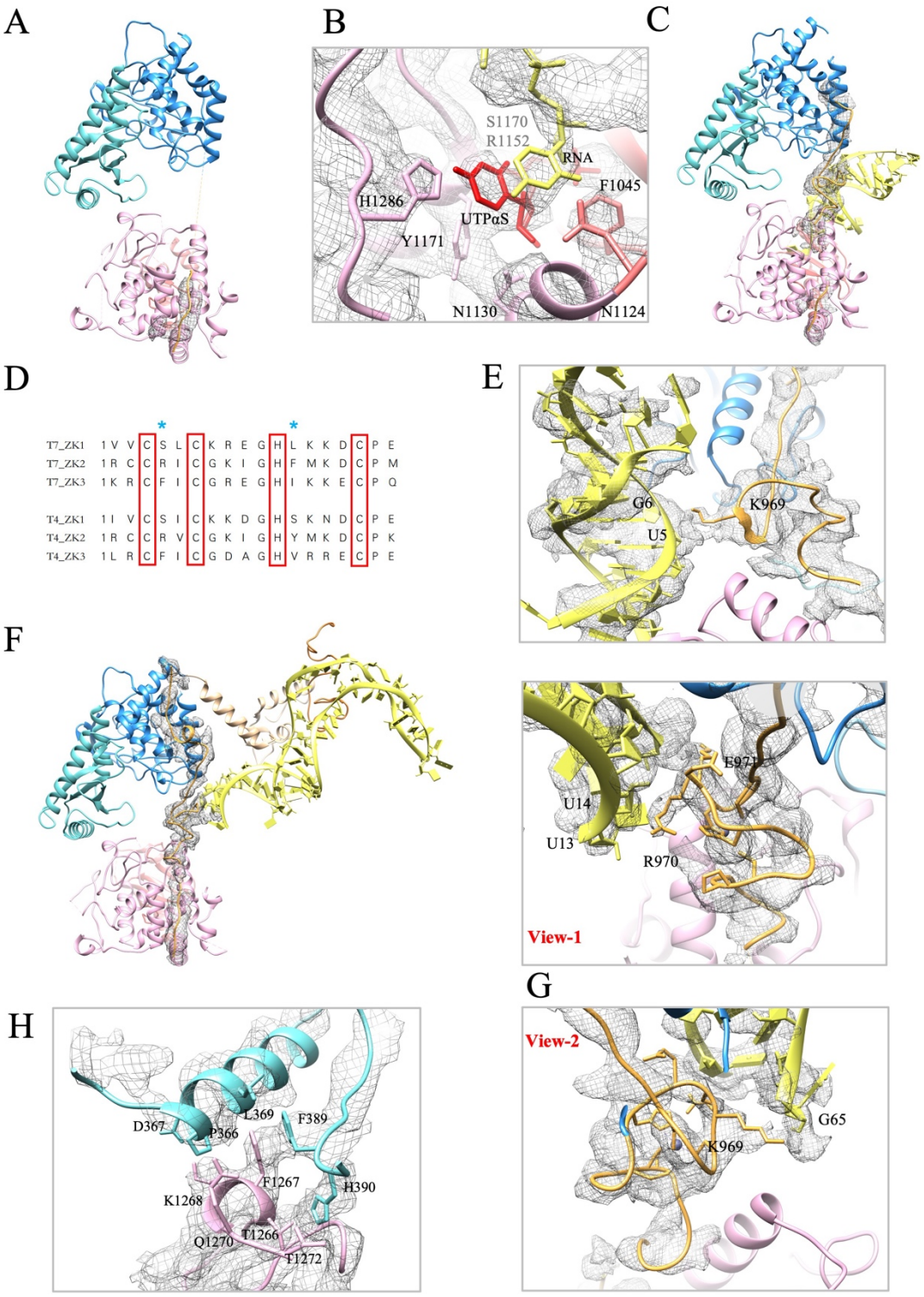


Figure 5.3: Legend on next page

Figure 5.3. Structural details of TUT7 in diverse states. (A) Cartoon models of TUT7 in apo state and bound with pre-let-7g miRNA (binary complex). Coloring as in Figures 5.1 and 5.2, with density associated with the linking polypeptide region between the LIM and CM that contains ZK1 shown in grey mesh. (B) Zoomed-in view of UTPαS within the catalytic pocket, related to Figure 5.1H. (C) Cartoon models of TUT7 bound with pre-let-7g miRNA (binary complex). (D) Sequence alignment of ZK1-ZK3 from TUT7 and TUT4; red asterisks highlight the residues after the first cysteine and first histidine in the CCHC knuckle architecture, respectively. (E) Close-up of K969 that follows the second CCHC motif cysteine interacting with pre-let-7g in the binary complex. (F) As A and B for the TUT7 ternary complex in conformation-II, again showing the linker region density in grey mesh. (G) The alternative interface found in the ternary complexes in which R970 and K969 interface with pre-let-7g miRNA. (H) Close-up of the physical interface between the LIM and CM in the ternary complex.

5.2.4. Lin28A modifies target pre-miRNA recognition by prior association

Two distinct conformations (-I and -II) of the TUT7 ternary complex were resolved at 3.8Å and 3.6Å resolutions, respectively. In conformation-I, most of the pre-let-7g could be resolved. In this conformation, the RNA stem interacts with both the TUT7 LIM and pre-element domain, while the let-7g ^{49}AGGAG^{53} region coordinates with the zinc finger of TUT7 and zinc knuckle domain of Lin28A, forming a stable ternary complex (Nam et al., 2011) (Figure 5.2A-B). In this conformation, despite a clear density for the pre-element, no density was observed for the CSD of Lin28A, presumably due to the relative flexibility between the LIM+Lin28A and pre-let-7g. Based on the position of the RNA, this complex is believed to represent the initial step in oligo-uridylation, a "capture" state in which the pre-let-7 has begun association with TUT7 but not yet completed its engagement.

In conformation-II (Figure 5.2C-D), the CM density missing in conformation-I due to intrinsic flexibility could be entirely resolved as well as the LIM-CM linker region including the ZK1 domain (Figure 5.3F). R970 forms two hydrogen bonds with U13 and U14 in the RNA stem while K969 binds the sugar backbone of G66 (Figure 5.2E, Figure 5.3F) which is different from the binary complex in which its sidechain instead interacts with U5 and G6 (Figure 5.3E). These interactions aid the translocation of pre-let-7 to close the catalytic pocket of the CM and stabilize the overall TUT7 conformation. The higher-resolution model enabled observation of intramolecular interactions at a residue level, especially for the interface between the LIM and CM. For example, a hydrophobic interaction formed by ring-stacked residues F389 and F1267, and a H390-induced polar interaction with T1266 and T1272 stabilize the interface (Figure 5.2F, Figure 5.3H), thus also confirming the role of a similar interface in the TUT7 binary complex (Figure 5.1F), and presumably the apo state also. Together, these interactions ensure maintenance of the distinctive C-shape structure of TUT4/7.

Both models of ternary complexes revealed the ZK domain of Lin28A attached to its ^{49}AGGAG^{53} recognition sequence in the pre-let-7 together with the zinc finger of the TUT7 LIM. To explore the potential conformational changes in the overall structure, a 3D

conformational variability analysis was performed. The TUT7 LIM and Lin28A were found to be stable, while the pre-element of pre-let-7g miRNA exhibited density that could be assigned to a set of flexibly-related conformations. The angle between the pre-element and RNA stem changed from around the initial (frame #1) 70° to a final (frame #20) 40° (Movie2, Appendix). The finding was repeated via the same analysis for the ternary complex in conformation-II (Movie3, Appendix). This result is consistent with a previous study, in which it was concluded that in solution, the pre-let-7 alone may adopt different conformations with various stem stabilities (Liu et al., 2018a), which is now observed in cryo-EM structures bound to TUT7. Since the CSD was thought to bind in the pre-element region this may explain why this density could not be resolved.

Altogether, comparison of the two TUT7 ternary complexes and the binary and apo structures (Fig. 5.3A-I, Figure 5.4A-H) indicates that the presence of Lin28A has changed how the pre-let-7 engages with the enzyme, directing it away from direct binding to the CM active site (as in the binary complex) and towards a two-stage process in which, first, the LIM and Lin28A capture the stem and pre-element, and then the CM engages with the 3’ end. Further details of the first of these steps – the capture complex – were provided by determination of the equivalent structure for TUT4, also with Lin28A and pre-let-7g.

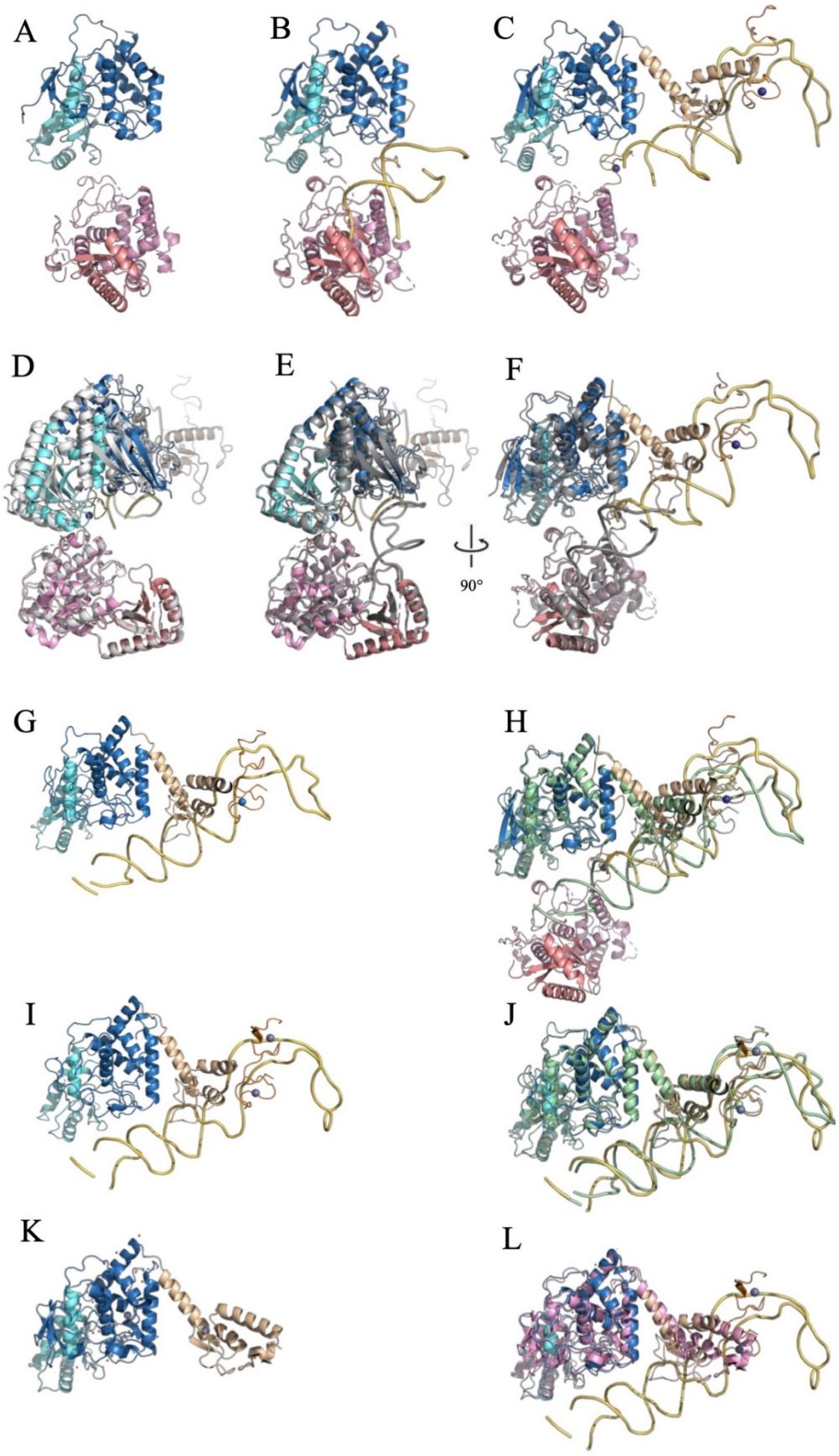


Figure 5.4: Legend on next page

Figure 5.4. Structural comparison of TUT7 and TUT4 cryo-EM and crystal structures. (A-C) Cartoon models of TUT7 in apo state, as a binary complex with pre-let-7g, and in ternary complex with pre-let-7g and Lin28A (conformation-II). Domains are colored as in the main figures and zinc ions are blue spheres. (D) Superimposition of the TUT7 apo (white) and pre-let-7-g bound cryo-EM structure from the TUT7/pre-let-7g/Lin28A ternary complex conformation-II (colored), aligning by the TUT7 CM. (E-F) Superimposition of the TUT7/pre-let-7g binary complex (deep grey) and the TUT7/pre-let-7g/Lin28A complex conformation-II (colored), in two views related by 90° as shown, and aligning by the TUT7 CM. (G) Cartoon model of TUT7/pre-let-7g/Lin28A ternary complex conformation-I. (H) Superimposition of the TUT7/pre-let-7g/Lin28A complex conformation-I (pale green) and conformation-II (colored), aligning by the TUT7. (I) Cartoon model of the TUT4/pre-let-7g/Lin28A ternary complex reported here. (J) Superimposition of the TUT7/pre-let-7g/Lin28A ternary complex conformation-I (pale green) and TUT4/pre-let-7g/Lin28A ternary complex (colored), aligning by the TUT4 LIM. (K) Crystal structure of the TUT4 LIM (PDB: 6iw6). (L) Superimposition of the crystal structure of the TUT4 LIM (pink) and TUT4/pre-let-7g/Lin28A, aligning by the TUT4.

5.2.5. TUT4 capture complex determined at higher resolution

A 3.7 Å cryo-EM structure was obtained for the TUT4/Lin28A/pre-let-7 in the oligo-uridylation capture state found also with TUT7 ternary conformation-I. This structure enabled a clearer identification of the interactions between the LIM module of TUT4, Lin28A and the pre-let-7 (Figure 5.5A). Although the CM of TUT4 is not well defined in the 3D map, density apparently derived from it can be visualized in 2D class averages of the complex, roughly oriented in the same view compared to the TUT4 ternary complex in conformation-I (Figure 5.5B), in line with this being a capture state prior to CM engagement with the open end of the pre-let-7g. The ^{49}AGGAG^{53} recognition sequence from pre-let-7g was observed trapped into a positively-charged pocket created by TUT4 and Lin28A (Figure 5.5C-D). Together with the TUT7 ternary complexes this structure therefore provides the first demonstration that TUT4 interacts directly with Lin28A, an up-to-now unresolved key question in the field. In this ternary interface, a hydrophobic patch is formed between G276, L277, I316 from TUT4 LIM (which are conserved in TUT7) and prolines 156 and 168 of Lin28A (Figure 5.5E-F, Figure 5.6). Overall, the structures of the TUT4 and TUT7 LIMs are very similar (RMSD 1.22 Å over 302 core backbone Cα atoms, 2.18 Å over all 380 pairs) (Figure 5.4I-J). Compared to the crystal structure of TUT4 LIM (Yamashita et al., 2019) (Figure 5.4K), the angle between the helices α4 in the zinc finger and α5 in the palm sub-domain is reduced by the binding of pre-let-7g (Figure 5.4L).

The pre-let-7g is observed bound to three main interfaces in the TUT4 capture complex (Figure 5.7A), which we have termed interface-1, interface-2 and interface-3 respectively (Figure 5.7B-F). In interface-1, the TUT4 LIM P667, R669, R670 and N688 interact with A64 and G65 at their deoxyribose and phosphate group (Figure 5.7C). The pre-let-7 A17 and C18 backbone regions were bound by R327 and K321 respectively in the second interface region (Figure 5.7D). Interface-3 is critical to the ternary interaction: in this region the zinc finger of the TUT4 LIM is observed to bind the ^{49}AGGAG^{53} region of pre-let-7 and the two zinc knuckles of Lin28A. The sidechains of four positively-charged residues R283, R286, H320 and K324 from the LIM participate in the interaction with G51, A52, G53 and A54 of the pre-let-7g

(Figure 5.7E-F). Mutagenesis of TUT4 (Figure 5.8A) in interfaces -2 and -3 abolished the enzyme' s capacity to oligo-uridylate pre-let-7 in the presence of Lin28A (Figure 5.8B-D), due to a much-reduced ability to form a TUT/pre-let-7g/Lin28A ternary assembly (Figure 5.8E). Although mutations in the interface-1 affected the formation of ternary complex (Figure 5.8E), they did so less than interface-2 and -3 mutations, and still permitted oligo-uridylation, albeit with a lower efficiency compared to wild-type (Figure 5.8B). However, all mutants were shown to add a single uridine to pre-let-7g in the absence of Lin28A (Figure 5.8B). It is surprising that mutations in the ZK1 region which binds to the RNA stem in the TUT7 ternary conformation-II, had no effect on oligo-uridylation activity (Figure 5.8B, E). But it is notable that the key interactions observed in TUT4 complexes involve residues conserved in TUT7, with the one notable difference that L691 in human TUT4 is P613 in TUT7, and is a polypeptide locus in interface-1 besides U16 of pre-let-7g; and otherwise only the conservative substitution of R669 in human TUT4, for K591 in TUT7 (Figure 5.6). It is notable that higher-order assemblies of TUT4/pre-let-7g/Lin28A can be observed in the shift assays testing complex assembly in the presence of mutations in the three interfaces (Figure 5.8E) and a concentration-dependent formation of these larger complexes was confirmed by titrating pre-let-7g with Lin28A and TUT4 alone and together (Figure 5.8F).

The two zinc knuckles of Lin28A interact with the ^{49}AGGAG^{53} motif of pre-let-7g, in a similar way to the previously-solved isolated complexes (Loughlin et al., 2012; Nam et al., 2011). Based on these structures, it was concluded that the double zinc knuckles motif recognizes two AG dinucleotides separated by a single nucleotide spacer (Loughlin et al., 2012; Nam et al., 2011). Perhaps due to the use of a modified RNA sequence mimicking the full pre-let-7g, the exact basis for recognition observed differed from the one identified here. In the previously solved crystal structures the Lin28A ZK1 binds with the A54 of pre-let-7g via a hydrogen bond from the Y140 sidechain, whereas we observe the tyrosine binding with G51, A52 and G53 (Figure 5.7A,E). In all, the pre-let-7 miRNA is sandwiched between Lin28A and the TUT4 LIM (Figure 5.5C-D, Figure 5.7B-F, Figure 5.9A) as it was in the TUT7/pre-let-7g/Lin28A complex (Figure 5.9B-C) and so as to engage with the pre-miRNA in a

fundamentally different way to that observed for the binary TUT7/pre-let-7g complex set for mono-uridylation (Figure 5.9D). The binding mode and overall conformation of the pre-let-7g miRNA appears to be remarkably consistent in all three Lin28A-containing complexes (Figure 5.9A-C), apparently due to the LIM/Lin28A clasp (Figure 5.7F) and other key interfaces shown by mutational data to be the specific interactions responsible for ensuring efficient and processive oligo-uridylation.

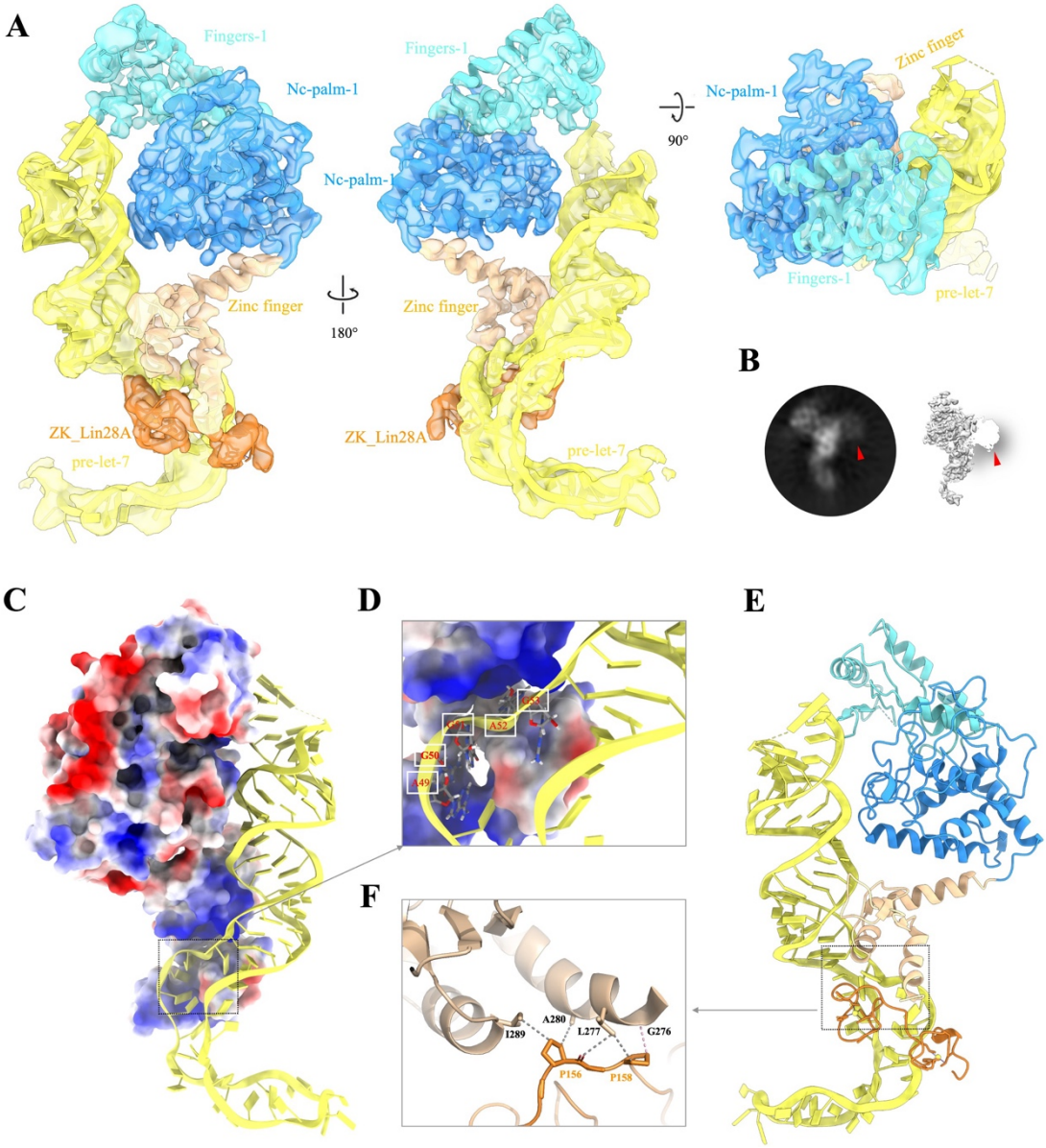


Figure 5.5: Legend on next page.

Figure 5.5. Cryo-EM structure of TUT4 in complex with pre-let-7g with Lin28A. (A) Overall structure of TUT4/pre-let-7/Lin28A, shown as a 3.7 Å cryo-EM density map and cartoon representations in four views, related by 90° and 180° rotations, as shown. Colors are as previously: pseudo-NTD (LIM) fingers and palm cyan and blue respectively, LIM zinc finger and linking density towards CM is colored sand, pre-let-7g yellow and Lin28A zinc knuckle dark orange. (B) A typical 2D class average representing the poorly-resolved density from the CM (marked by a red arrowhead). (C) Electrostatic surface potential of TUT4/pre-let-7g/Lin28A. The positively and negatively charged regions are colored from blue to red respectively. (D) Zoomed-in view of the pre-let-7g 50GGAGA54 region in a positively charged pocket created by the TUT4 zinc finger and zinc knuckle domain of Lin28A (black dashed arrows). € Cartoon model of TUT4 /prelet-7g/Lin28A ternary complex, colors as in other panels. (F) Zoomed-in view of TUT4/Lin28A interface. Black dotted lines mark distances atoms within 3.8 Å and the magenta dotted line shows the distance between two atoms of 3.8-4 Å.

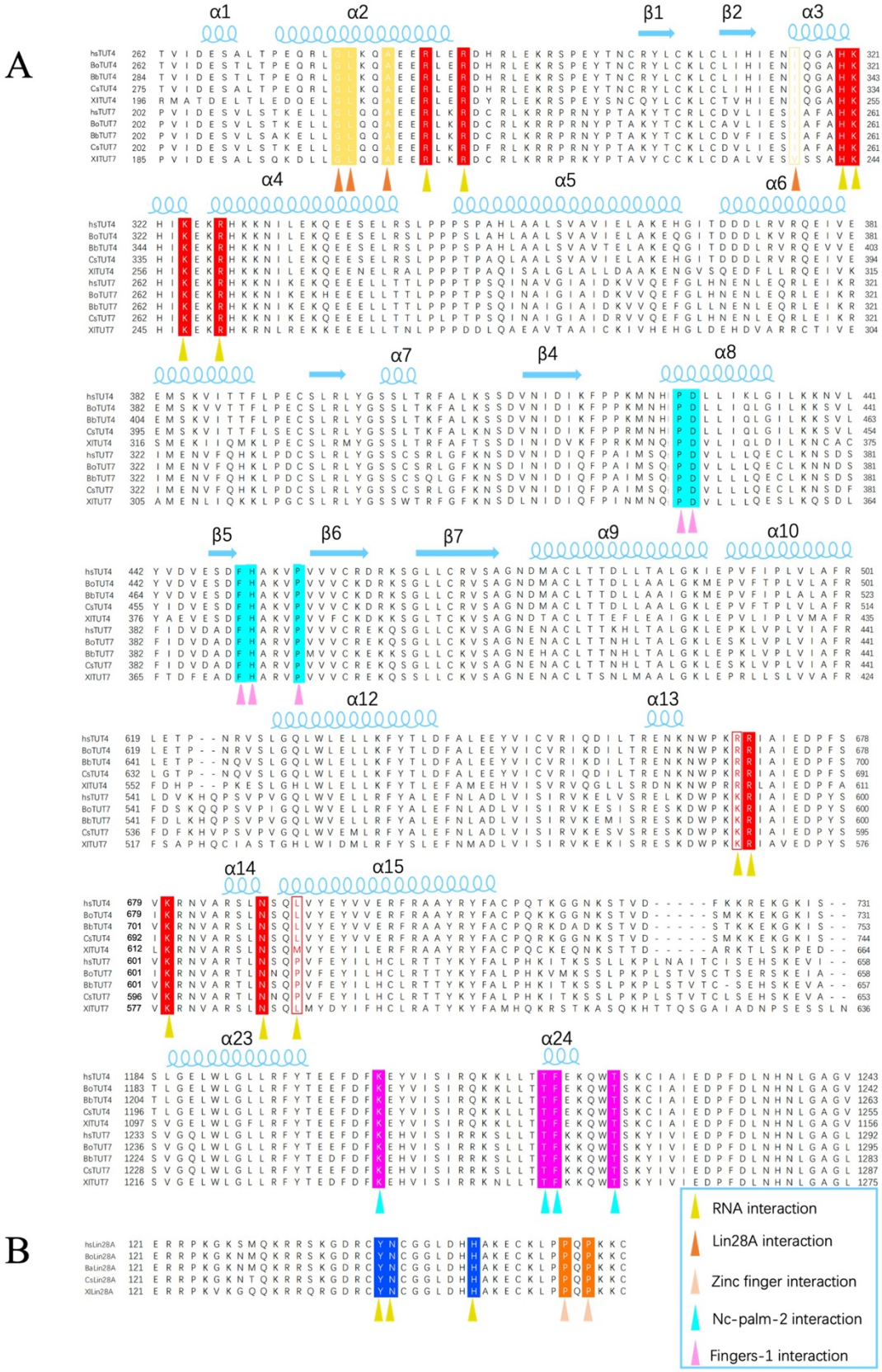


Figure 5.6: Legend on next page

Figure 5.6. Sequence alignments of TUTs and Lin28A. (A) The amino acid sequences of human TUT4 (HsTUT4) and TUT7 (HsTUT7) are aligned with those of the TUT4 or TUT7 proteins from other organisms: California sea lion, Cow, Little brown bat and Xenopus laevis. Secondary structural elements revealed by TUT7 structure determined in the present study, are depicted above the sequences, with α-helices and β-sheets shown as cartoon helices and arrows, respectively. Conserved interaction site residues are shaded while nonconserved interaction site residues are colored and unshaded. (B) The amino acid sequences of human Lin28A (HsLin28A) is aligned with those of the Lin28A proteins from other organisms: California sea lion, Cow, Little brown bat and *Xenopus laevis*.

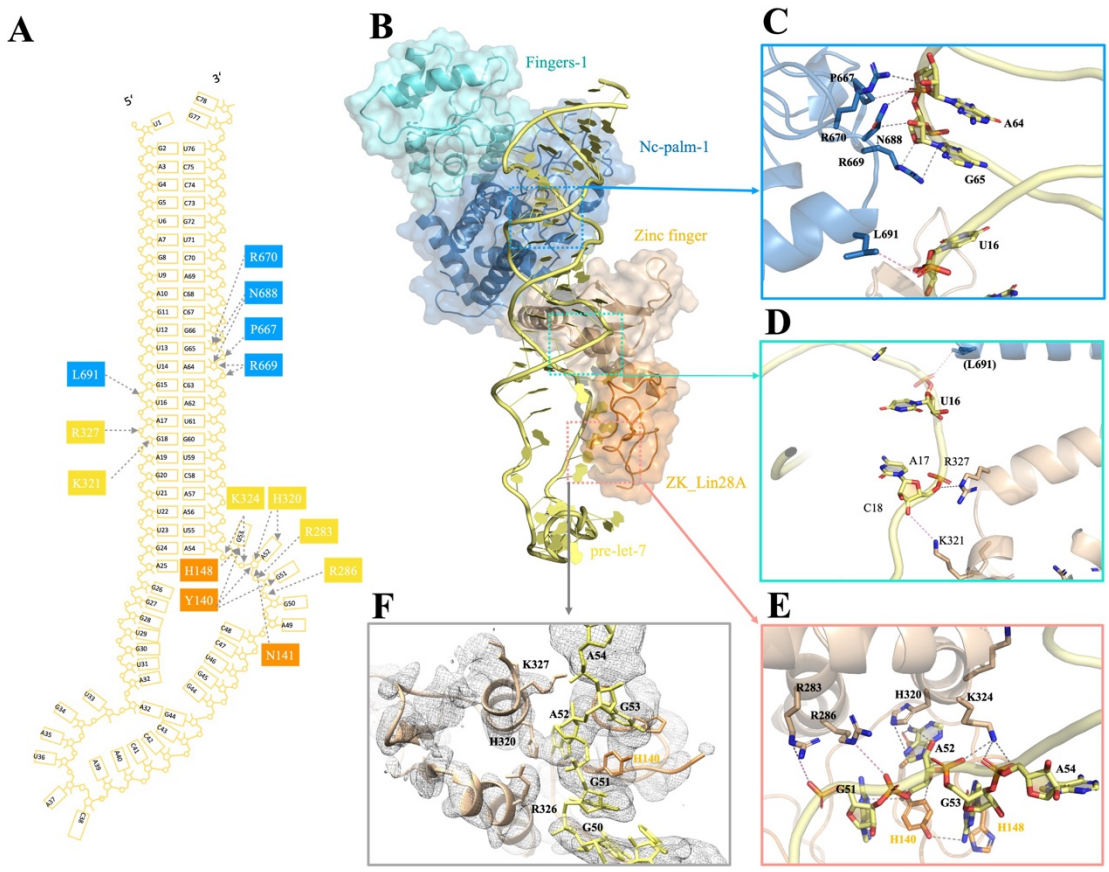


Figure 5.7. Pre-let-7 recognition by TUT4 and Lin28A. (A) A schematic of pre-let-7 recognition by TUT4 and Lin28A, based on the structure reported here (see also Figure 5.5). (B) Structure of TUT4/Lin28A bound to pre-let-7g. C-F. Zoom-in views of the different interfaces bound to pre-let-7g: interface-1 (C), interface-2 (D) and interface-3 (E, F).

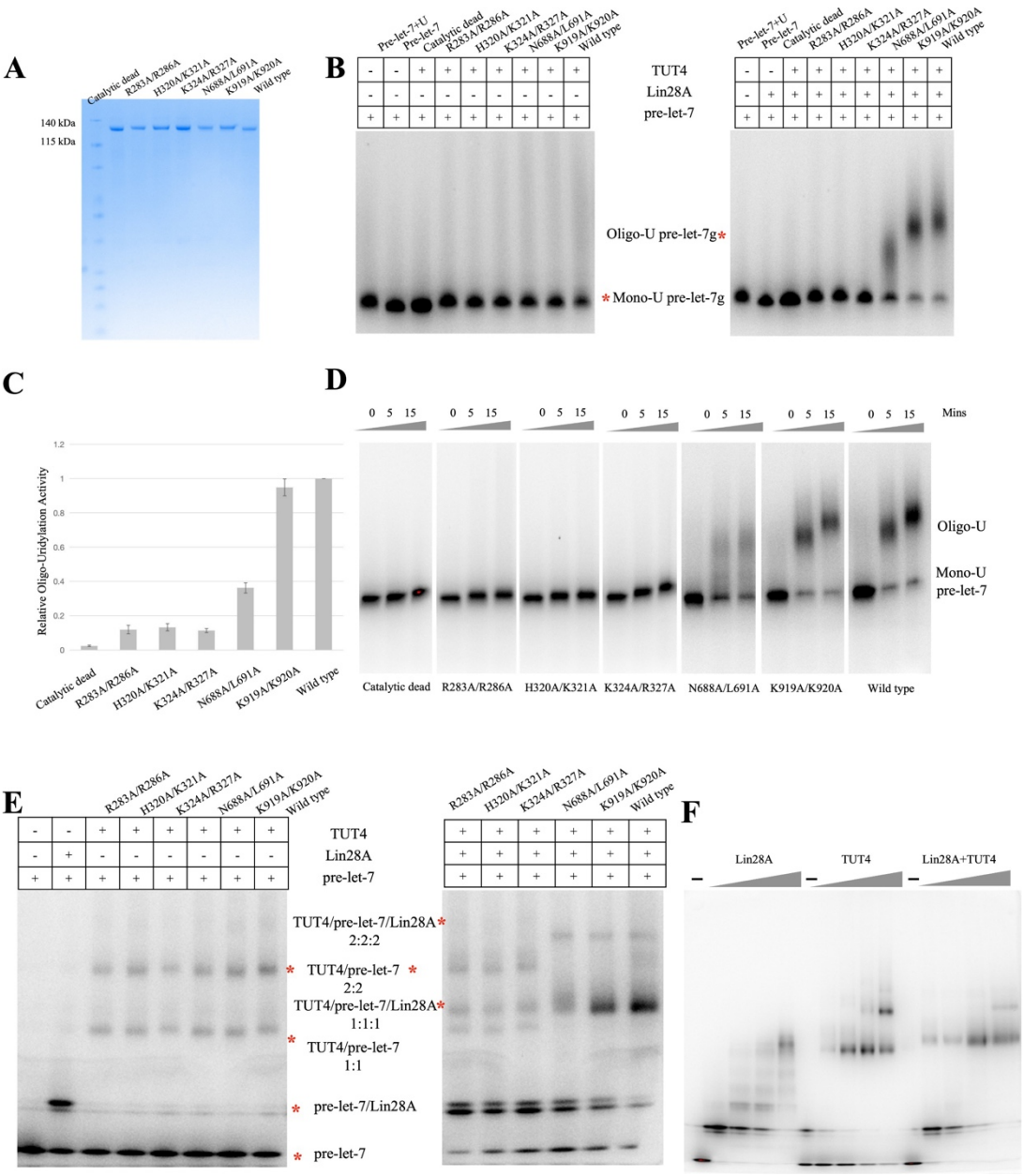


Figure 5.8: Legend on next page

Figure 5.8. Effects of mutants at the TUT4/pre-let-7g interface on uridylation activity. (A) SDS-PAGE of recombinant TUT4 wild type and its mutants. (B) In vitro uridylation of pre-let-7 by wild-type TUT4 and mutants, in the absence or presence of Lin28A, as indicated. Pre-let-7 = substrate; pre-let-7-U = synthetic mono-uridylated product as marker. (C) Relative oligo-uridylation activity of wild-type or the indicated mutant TUT4 enzymes. Error bars: standard deviations. (D) In vitro uridylation of pre-let-7g by TUT4 and its variants in the presence of Lin28A. (E) Electrophoretic mobility shift assay of pre-let-7g for wild-type TUT4 and mutants in the absence or presence of Lin28A. (F) Electrophoretic mobility shift assay of pre-let-7g with increasing concentrations of Lin28A alone, TUT4 alone, and Lin28A+TUT4 together.

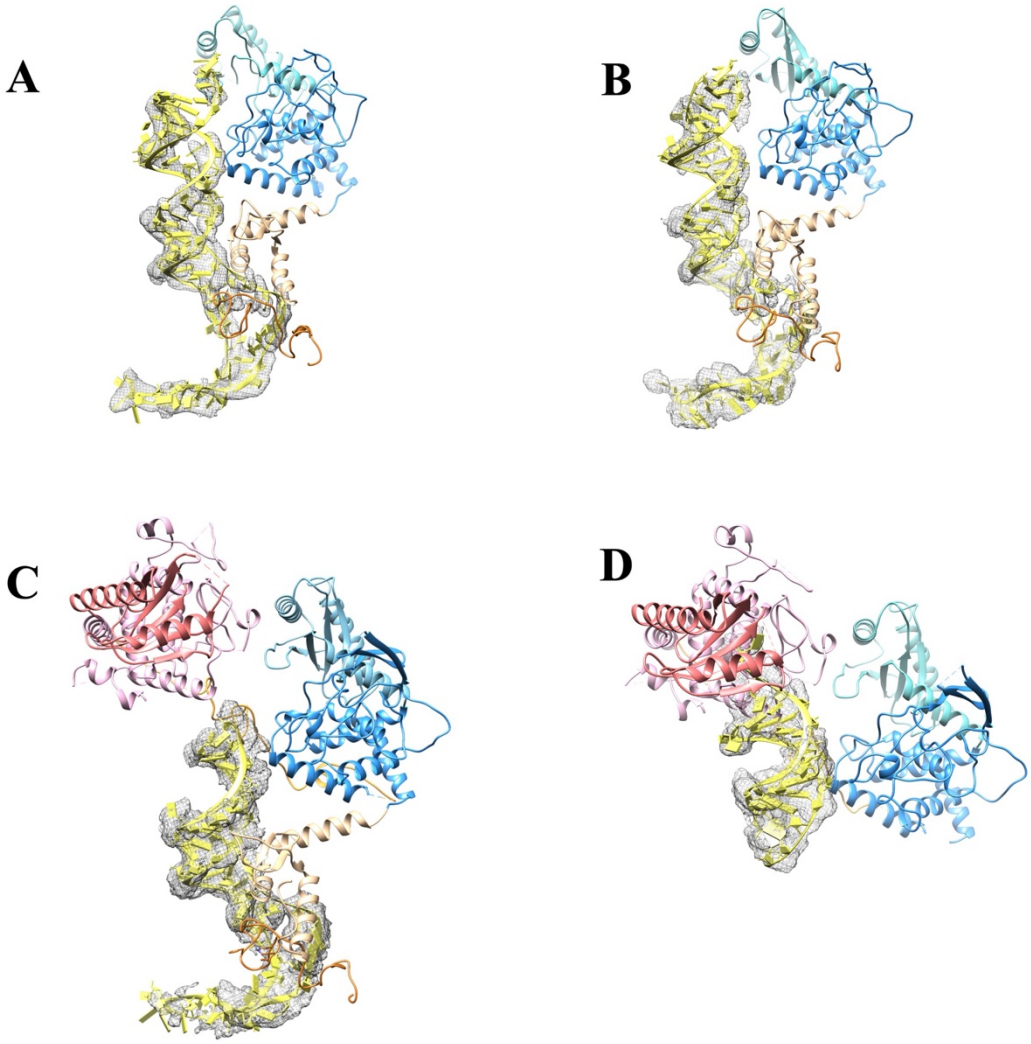


Figure 5.9. Density maps for the bound pre-let-7g in diverse states from TUT7 and TUT4 (A) Overview of the TUT4/pre-let-7g/Lin28A complex with density for the pre-miRNA alone displayed. Protein regions are colored as elsewhere. (B) TUT7/pre-let-7g/Lin28A in conformation-I. (C) TUT7/pre-let-7/Lin28A in conformation-II. (D) TUT7/pre-let-7g binary complex.

5.2.5. A mechanism for switching between mono- and oligo-uridylation in TUT4/7

Together the structures determined have shown that Lin28A induces a large conformational change, including of the CM and the Zinc finger in the LIM, on engagement with pre-let-7g miRNA and TUT4/7. After binding to pre-let-7g/Lin28A, the zinc finger in the TUT7 LIM was observed with good density, whereas it was unobserved in TUT7 apo or binary states because of mobility/disorder (Figure 5.4A-C, G). In the TUT4 ternary complex similarly the LIM was well-resolved in complex with Lin28A and pre-let-7g miRNA (Figure 5.4I). Where pre-let-7 miRNA was shown to be flexible it was previously found to form a fixed conformation within the context of a ternary Dicer complex incorporating TRBP (Liu et al., 2018a); similarly a well-ordered pre-let-7g conformation was observed here after TUT4/7 binding alongside Lin28A. Overall, the role of Lin28A appears to be to (i) directly interact with the pre-let-7 miRNA; (ii) to enable the LIM Zinc finger to also bind to the pre-let-7 RNA (since this interaction is absent in the TUT7 binary complex); (iii) to directly interact with the TUT LIM domain as well. The structures of TUT7 and TUT4 bound to pre-let-7g miRNA and to Lin28A allow us to propose a mechanistic model for mono-uridylation and for oligo-uridylation, as follows (Figure 5.10).

In the mono-uridylation process, the TUT7 alone presents a pre-state (Figure 5.1E, Figure 5.10A(i)). Next, TUT7 recruits the pre-let-7, binding its open stem (binary complex, Figure 5.1F). The catalytic pocket engages with the 3’ end of the RNA, while the ZK1 binds to the RNA stem via key residue K969 (Figure 5.3E). This stabilized state aids the addition of a single uridine, indicating the binary complex we observe as an engaged active state (Figure 5.10A(ii)). Finally, the mono-uridylated pre-let-7 becomes unstable, which subsequently leads to the release of product (Figure 5.10A(iii)). In the presence of Lin28A the order of events during pre-let-7 capture and engagement changes. The pre-let-7 firstly binds to Lin28A and then recruits the TUT enzyme via binding to the LIM as a capture complex (Figure 5.2A-B, Figure 5.5, Figure 5.7, Figure 5.9, Figure 5.10B(i-iii)), before engagement of the CM with the open stem and 3’ end (Figure 5.2C-F, Figure 5.10B(iv-v)). By altering the mode of binding, the intervention of Lin28A enables a much more extensive engagement with pre-miRNAs than in its absence. By defining different ways in which TUT4/7 interact with pre-let-7 miRNAs in the

absence and presence of Lin28A our structures provide a mechanistic basis for understanding the role played by Lin28A in driving oligo-uridylation and downstream degradation of pre-miRNAs. They also show for the first time that Lin28A forms direct contacts with TUT4/7 as well as with its target RNA.



Figure 5.10. Proposed models of mono-uridylation and oligo-uridylation. (A, i) The pre-let-7 is initially recognized by apo TUT 4 or 7 and subsequently the substrate is Mono-uridylated to add a single UMP to its 3’ end, resulting in a mature Class II pre-let-7 miRNA. (A, ii-iii) The mono-uridylated pre-let-7 is then released due to the distributive activity of TUT4/7 in the absence of Lin28A. (B, i-ii) In pre-state-0, the pre-let-7 is initially recognized by Lin28A. (B, iii) In pre-state-1, the binary complex of pre-let-7/Lin28A recruits the zinc finger of the

TUT to form a ternary complex near the critical let-7 pre-miRNA GGAG binding region. This is equivalent to TUT7 conformation-I (Figure 5.2) and to the TUT4 complex (Figures 5.5 and 5.7). (B, iv) In the active state, the 3' end of pre-let-7 is bound into the catalytic pocket of the TUT CM, stabilizing its orientation with respect to the TUT LIM, pre-let-7 miRNA and Lin28A. (B, v) The oligo-uridylated pre-let-7 is released as its attachment becomes destabilized.

5.3. Discussion

The uridylation mechanism of pre-let-7 by TUT4/7 involving the activity switch from mono-uridylation to oligo-uridylation has been proposed in previous studies (Balzeau et al., 2017; Menezes et al., 2018). The CM alone favors addition of a single uridine to the group II pre-let-7 by an unstable interaction, and this state is distributive in its activity and only forms transitory complexes without the assistance of Lin28A. Thus, in the absence of Lin28A, TUT4/7 function similarly – as previously discussed – to the homologous single-NTD yeast enzyme Cid1. Without Lin28A, the mono-uridylated pre-let-7 would release before the occurrence of oligo-uridylation, while with Lin28A oligo-uridylation occurs (Heo et al., 2012). Lin28A regulates this "activity switch", we now show, by recruiting the TUT zinc finger to the "AGGAG" motif of the pre-let-7 and forming a stable ternary complex, which supports processive oligoU addition by the CM. The ZK2 of TUTs aids this process by engaging the U tail (Faehnle et al., 2017a; Yamashita et al., 2019). Previous studies could not provide answers to questions such as, how Lin28A affects the LIM of TUTases and how the latter interacts with pre-let-7. This work provides important steps towards answering such questions.

The five solved cryo-EM structures of the TUTases in different states enabled an analysis of the uridylation mechanism and identification of the ternary interactions occurring among TUTases, pre-let-7 and Lin28A. The two TUT7 ternary complexes presented different relocating states before oligo-uridylation has occurred. However, due to the flexibility of the TUT7 enzyme, there were still around 700 aa residues which could not be reconstructed, the study of which is a very challenging mission for studies of the future. Given the ways in which engagement with pre-let-7 and with Lin28A has been shown here to affect the relative order/disorder of TUT4/7, it is quite possible that resolving the missing (to date disordered) regions of TUT4/7 will require determination of their structures with other co-factors and substrates bound, which might relate to their roles regarding other aspects of RNA silencing and mRNA turnover or to the role identified for them in suppressing viral replication, or another biological role entirely.

The findings summarized in this chapter provide a framework for understanding the effects

of TUT4/7+Lin28A on pre-miRNAs, and thus for the role played by TUT4/7 and Lin28A in a wide range of cancers (Fu et al., 2014; Hagan et al., 2009; Hallett & Hassell, 2011; Heo et al., 2009; Kim et al., 2020; Medhi et al., 2022; Piskounova et al., 2011; Thornton & Gregory, 2012; Ustianenko et al., 2018; Zhao et al., 2014), indicating the potential to tackle TUT inhibition by targeting the key residues interacting with pre-let-7 and Lin28A as a treatment regulating the expression of let-7 for cancer patients in clinical research of the near future. However, it does not show how TUT4/7 target histone mRNAs (Lackey et al., 2016), or a broader range of mRNAs (Lim et al., 2014) including that of interleukin-6 (IL6) (Jones et al., 2009). Further studies will be needed to provide an equivalent mechanistic understanding for TUT4/7 targeting these alternative substrates, and a more detailed understanding of the molecular basis for the critical and varied roles played by these enzymes in cell biology (Yu & Kim, 2020).

Chapter 6 Conclusion

6.1 HTP method for TUTases study

Due to the difficulties of expressing flexible human TENTase proteins from our past experiences of structural studies, including a range of different expression systems and cloning strategies, an efficient high throughput (HTP) technique was established and optimized for this project based on the protocol applied in the Oxford Protein Production Facility (OPPF). A series of TUT4 and TUT7 constructs in both truncated forms and full length were purified successfully with high quality and yield for subsequent functional assays and structural studies. The two-step purification protocol employed benefits target proteins with low yield and serious degradation issues. Diverse combinations of reagents were used as means for avoiding aggregate formation especially for the purification of flexible full length TUT7 and TUT4 enzymes. Encouraged by the success in obtaining TUT4/7 proteins, other challenging TENT family members like human TENT4A and TENT4B, are under exploration using this method. TENT4A and TENT4B were found to add mixed GA tails to the 3' of mRNAs of HBV together with ZCCHC14 protein and thus prevent the mRNAs from being degraded by nucleases (Lim et al., 2018). The preliminary results already show promising yield and quality in small scale purification, and initial structural results are available, while assays show that they are catalytically active (not shown).

6.2 Structural study of TUT7_CM

Leemor Joshua-Tor' s group pioneered the structural studies of TUT7_CM, and they succeeded in expressing the TUT7_CM in insect cells (sf9) for crystallographic studies (Faehnle et al., 2017a). I successfully obtained a highly expressing construct of TUT7_CM using an HTP cloning and expression method, which was purified in homogeneous form from *E.coli.* This is by far the simplest method to obtain TUT7_CM protein and lay the foundation for subsequent crystallographic studies. After optimization, repeatable crystals of TUT7_CM diffracted to 2.6 Å. Further structural work will, it is hoped, focus on revealing the inhibition mechanism of candidate inhibitors, perhaps in commercial partnership.

6.3 Structural characterization of TUT7

I successfully solved the cryo-EM structures of apo TUT7, TUT7/pre-let-7 binary and TUT7/pre-let-7/Lin28A ternary complexes using the purified TUT7 full length protein, the

construct of which was generated by an HTP cloning and expression method. The 4 Å TUT7 apo structure reveals that the "C-shape" overall architecture by the maintenance of an apparent interaction between LIM and CM. No linker is visualized, and the residues from 1-205 and 1365-1495 are undetermined structurally due to their flexible nature. The pre-let-7 RNA stabilizes the CM and reduces the dynamism of the overall structure. This benefits the alignment of particles and thus a high-resolution model of the TUT7/pre-let-7 binary complex was determined at 3.7 Å. In this complex structure, the RNA stem locks at the catalytic pocket and pre-element region of the RNA was not resolved. The zinc finger in LIM was not fully resolvable and could not be modelled but a small proportion of the 2D class averages obtained for the sample indicates the likelihood of interaction between the zinc finger of LIM and pre-let-7 or this binding is a transient state. Unexpectedly, I obtained two distinctive classes of 3D maps for the ternary complex of TUT7/pre-let-7/Lin28A. The first structure (conformation-I) has the density for the CM missing but the zinc finger and NTD1 of TUT7_LIM can be visualized, which is bound to Lin28A and pre-let-7. These positively-charged residues (R223, R226, H260, K264, E265, R267, N610) are confirmed to interact with RNA in diverse regions**.** It was hypothesized that this model presents a pre-oligo-uridylation "capture" state. In the second structure, the pre-let-7 stem is bound to the CM. Although the densities of CM are poor, the zinc finger of LIM is clearly bound to the RNA stem, while the 3' end of RNA is very close to the cleft of the CM, indicating the formation of a ternary complex representing an active "engaged" state. These RNA interacted residues confirmed in the ternary complex above (conformation-I) are visualized in this conformation-II complex. Besides, the K969 and R970 in the zinc knuckle-1 show clear interaction with the RNA stem. In this high-resolution model, a hydrophobic interaction formed by residues F389 and F1267, and a H390-induced polar interaction with T1226 and T1272 was validated and obviously stabilizing the interface. All in all, the four cryo-EM structures of TUT7 in different forms represent the conformational transitions from the apo to mono- and oligo-uridylation states. The interfaces among the TUT7 zinc finger, pre-let-7 and Lin28A are identified for the first time. The detailed atomic information provides us with a uridylation mechanism of pre-let-7 regulated by TUT7, modulated by the presence or absence of Lin28A.

The future exploration for TUT7 will focus on how TUT7 works on other RNAs in diverse pathways. For example, TUT7 in particular uridylates histone mRNAs and regulates their biogenesis together with other components such as 3' hExo (Holmquist et al., 2022). Questions like 'How TUT7 uridylates and adds oligo-uridines to hmRNAs without the assistance of Lin28A' and 'How TUT7 coordinates with 3' hExo and SLMP when processing the hmRNAs' will be interesting to answer. Other questions such as 'how TUT7 participates in the modification of piRNA localized in the nucleus' (Holmquist et al., 2022) also need answering.

6.4 Structural analysis of TUT4

On the crystallography side, both TUT4_CM and TUT4_LIM in complex with S2A4 can crystalise and diffract to 4.5 Å and 5Å respectively, indicating high-resolution structures of both could be obtained, with crystal improvement, to help understand the uridylation mechanism at atomic level and provide detailed information to reveal how the functional nanobody S2A4 interacts with LIM and then blocks the ternary interaction with pre-let-7 and Lin28A.

Furthermore, the challenging TUT4 full length protein was successfully obtained for structural studies. Although cryo-EM structures of TUT7 in diverse forms provide us with detailed information on how TUT7 works on pre-let-7, TUT4 is worth exploring at an atomic level since it has a different distribution in cells and distinct substrate specificity compared to TUT7. The cryo-EM work initially focused on the TUT4_miniF/pre-let-7/Lin28A ternary complex, for which I reconstructed two distinct models at 3.7 Å and 12 Å respectively. One of them is similar to the TUT7 ternary complex in conformation-I, in which the CM is missing, and the zinc finger interacts with pre-let-7/Lin28A. These positively-charged residues confirmed in the TUT7 ternary complex, are conserved and the same in TUT4 and interact with pre-let-7. Three hydrophobic residues, G276, L277, I316 in TUT4_LIM, together with two Prolines of Lin28A form a hydrophobic lateral contact, indicating the direct interaction between TUT4 and Lin28A. Interestingly, the second structure of the TUT4 ternary complex has additional densities on the loop region of pre-let-7, and I hypothesize it is likely to be the CSD domain of Lin28A, which was fit well with the crystal structure of Lin28A's CSD domain. Other forms

of TUT4_miniF such as apo TUT4_miniF and TUT4_miniF/pre-let-7 binary complex were imaged on the 200 kV Talos microscope, showing decent 2D class averages. They were similar to apo TUT7 and the TUT7/pre-let-7 binary complex. Krios data collection is scheduled soon aiming for high-resolution structure determination. Besides, the purified TUT4_F was assembled with pre-let-7 and Lin28A and also imaged on the 200 kV Talos, revealing a different shape with additional densities compared to the TUT4_miniF/pre-let-7/Lin28A. The missing CM part or additional unidentified parts in the N-terminus may account for the extra densities; otherwise, the TUT4 may take different conformations. Nevertheless, it is worth exploring the full-length ternary complex and comparing it with the truncated form and TUT7 ternary complex. The apo TUT4_miniF and TUT4_F both show high levels of flexibility, such that high-resolution structures would be hard to obtain. I expressed and purified two nanobodies: S2A4 and CB11, which could be used for stabilizing the TUT4 in trials of the future.

Future studies for TUT4 will focus on obtaining high-resolution ternary complex structures of TUT4_F/pre-let-7/Lin28A, especially in different functional states based on newly designed pre-let-7 substrates with different lengths of uridines at the 3’ terminus, aiming to capture different stages in oligo-uridylation. A key question to answer is, when TUT4 adds oligo-U to 3’, does uridylated pre-let-7 still bind to the CM, and how conformational changes happen between the LIM and CM during the extension of RNA. The new models should provide us with a better understanding of the uridylation mechanisms regulated by TUTases. Besides, TUT4 participates in the immune pathways such as those triggered by Toll-like receptors. It would be interesting to know if there is a direct interaction between TUT4 and TIFA after LPS simulation (Minoda et al., 2006). My focus will be on solving the complex structure by cryo-EM and understanding their physical interactions. TUT4 is also reported to work with E3 ligases such as TRIM 71 and TRIM25, which bind to TUT4 and the pre-let-7 stem respectively (Choudhury et al., 2014; Torres Fernandez et al., 2021). How do these E3 ligases affect TUT4 working on pre-let-7 and, how does TUT4 uridylate the piRNA without the assistance of Lin28A? What are the differences between the mechanisms by which TUT4 works on double stranded pre-let-7 and single-stranded piRNA? (Holmquist et al., 2022) The opportunity to

answer a series of interesting questions in the near future has been firmly established.

Chapter 7 Bibliography

Adachi, Y., & Fukuhara, C. (2012, Dec 1). TA strategy for rapid and efficient site-directed mutagenesis. *Anal Biochem, 431*(1), 66-68. https://doi.org/10.1016/j.ab.2012.08.030

Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., & Zwart, P. H. (2010, Feb). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr, 66*(Pt 2), 213-221. https://doi.org/10.1107/S0907444909052925

Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., & Terwilliger, T. C. (2002, Nov). PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr, 58*(Pt 11), 1948-1954. https://doi.org/10.1107/s0907444902016657

Alarcon, C. R., Lee, H., Goodarzi, H., Halberg, N., & Tavazoie, S. F. (2015, Mar 26). N6-methyladenosine marks primary microRNAs for processing. *Nature, 519*(7544), 482-485. https://doi.org/10.1038/nature14281

Albino, D., Civenni, G., Dallavalle, C., Roos, M., Jahns, H., Curti, L., Rossi, S., Pinton, S., D'Ambrosio, G., Sessa, F., Hall, J., Catapano, C. V., & Carbone, G. M. (2016, Jun 15). Activation of the Lin28/let-7 Axis by Loss of ESE3/EHF Promotes a Tumorigenic and Stem-like Phenotype in Prostate Cancer. *Cancer Res, 76*(12), 3629-3643. https://doi.org/10.1158/0008-5472.CAN-15-2665

Ambros, V., & Horvitz, H. R. (1984, Oct 26). Heterochronic mutants of the nematode Caenorhabditis elegans. *Science, 226*(4673), 409-416. https://doi.org/10.1126/science.6494891

Aricescu, A. R., Lu, W., & Jones, E. Y. (2006, Oct). A time- and cost-efficient system for high-level protein production in mammalian cells. *Acta Crystallogr D Biol Crystallogr, 62*(Pt 10), 1243-1250. https://doi.org/10.1107/S0907444906029799

Aslanidis, C., & de Jong, P. J. (1990, Oct 25). Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Res, 18*(20), 6069-6074. https://doi.org/10.1093/nar/18.20.6069

Assaiya, A., Burada, A. P., Dhingra, S., & Kumar, J. (2021, May 14). An overview of the recent advances in cryo-electron microscopy for life sciences. *Emerg Top Life Sci, 5*(1), 151-168. https://doi.org/10.1042/ETLS20200295

Bai, Y., Srivastava, S. K., Chang, J. H., Manley, J. L., & Tong, L. (2011, Feb 4). Structural basis for dimerization and activity of human PAPD1, a noncanonical poly(A) polymerase. *Mol Cell, 41*(3), 311-320. https://doi.org/10.1016/j.molcel.2011.01.013

Balzeau, J., Menezes, M. R., Cao, S., & Hagan, J. P. (2017). The LIN28/let-7 Pathway in Cancer. *Front Genet, 8*, 31. https://doi.org/10.3389/fgene.2017.00031

Barbieri, I., & Kouzarides, T. (2020, Jun). Role of RNA modifications in cancer. *Nat Rev Cancer, 20*(6), 303-322. https://doi.org/10.1038/s41568-020-0253-2

Barker, C. A., Kim, S. K., Budhu, S., Matsoukas, K., Daniyan, A. F., & D'Angelo, S. P. (2018, Jan 3). Cytokine release syndrome after radiation therapy: case report and review of the literature. *J Immunother Cancer, 6*(1), 1. https://doi.org/10.1186/s40425-017-0311-9

Barnhart, M. D., Moon, S. L., Emch, A. W., Wilusz, C. J., & Wilusz, J. (2013, Nov 27). Changes in cellular mRNA stability, splicing, and polyadenylation through HuR protein sequestration by a cytoplasmic RNA virus. *Cell Rep, 5*(4), 909-917. https://doi.org/10.1016/j.celrep.2013.10.012

Bartel, D. P. (2004, Jan 23). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell, 116*(2), 281-297. https://doi.org/10.1016/s0092-8674(04)00045-5

Behm-Ansmant, I., Grosjean, H., Massenet, S., Motorin, Y., & Branlant, C. (2004, Dec 17). Pseudouridylation at position 32 of mitochondrial and cytoplasmic tRNAs requires two distinct enzymes in Saccharomyces cerevisiae. *J Biol Chem, 279*(51), 52998-53006. https://doi.org/10.1074/jbc.M409581200

Bramucci, E., Paiardini, A., Bossa, F., & Pascarella, S. (2012, Mar 28). PyMod: sequence similarity searches, multiple sequence-structure alignments, and homology modeling within PyMOL. *BMC Bioinformatics, 13 Suppl 4*(Suppl 4), S2. https://doi.org/10.1186/1471-2105-13-S4-S2

Brown, D. J., Orelien, J., Gordon, J. D., Chu, A. C., Chu, M. D., Nakamura, M., Handa, H., Kayama, F., Denison, M. S., & Clark, G. C. (2007, Jun 15). Mathematical model developed for environmental samples: prediction of GC/MS dioxin TEQ from XDS-CALUX bioassay data. *Environ Sci Technol, 41*(12), 4354-4360. https://doi.org/10.1021/es062602+

Cai, T., Atteh, L. L., Zhang, X., Huang, C., Bai, M., Ma, H., Zhang, C., Fu, W., Gao, L., Lin, Y., & Meng, W. (2022). The N6-Methyladenosine Modification and Its Role in mRNA Metabolism and Gastrointestinal Tract Disease. *Front Surg, 9*, 819335. https://doi.org/10.3389/fsurg.2022.819335

Chang, H. M., Triboulet, R., Thornton, J. E., & Gregory, R. I. (2013, May 9). A role for the Perlman syndrome exonuclease Dis3l2 in the Lin28-let-7 pathway. *Nature, 497*(7448), 244-248. https://doi.org/10.1038/nature12119

Chaves-Arquero, B., Collins, K. M., Christodoulou, E., Nicastro, G., Martin, S. R., & Ramos, A. (2021, Nov 12). The distinct RNA-interaction modes of a small ZnF domain underlay TUT4(7)

diverse action in miRNA regulation. *RNA Biol, 18*(sup2), 770-781. https://doi.org/10.1080/15476286.2021.1991169

chDai, D., Wang, H., Zhu, L., Jin, H., & Wang, X. (2018, Jan 26). N6-methyladenosine links RNA metabolism to cancer progression. *Cell Death Dis, 9*(2), 124. https://doi.org/10.1038/s41419-017-0129-x

Chen, C. Z. (2013, Sep). An unsolved mystery: the target-recognizing RNA species of microRNA genes. *Biochimie, 95*(9), 1663-1676. https://doi.org/10.1016/j.biochi.2013.05.002

Chen, L., Sun, Y., Wang, J., Jiang, H., & Muntean, A. G. (2016, May 3). Differential regulation of the c-Myc/Lin28 axis discriminates subclasses of rearranged MLL leukemia. *Oncotarget, 7*(18), 25208-25223. https://doi.org/10.18632/oncotarget.8199

Chen, T., Li, Z., Tu, J., Zhu, W., Ge, J., Zheng, X., Yang, L., Pan, X., Yan, H., & Zhu, J. (2011, Feb 18). MicroRNA-29a regulates pro-inflammatory cytokine secretion and scavenger receptor expression by targeting LPL in oxLDL-stimulated dendritic cells. *FEBS Lett, 585*(4), 657-663. https://doi.org/10.1016/j.febslet.2011.01.027

Chen, Y., Xie, C., Zheng, X., Nie, X., Wang, Z., Liu, H., & Zhao, Y. (2019, Mar). LIN28/let-7/PD-L1 Pathway as a Target for Cancer Immunotherapy. *Cancer Immunol Res, 7*(3), 487-497. https://doi.org/10.1158/2326-6066.CIR-18-0331

Chendrimada, T. P., Gregory, R. I., Kumaraswamy, E., Norman, J., Cooch, N., Nishikura, K., & Shiekhattar, R. (2005, Aug 4). TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature, 436*(7051), 740-744. https://doi.org/10.1038/nature03868

Cheng, K. J., Demir, O., & Amaro, R. E. (2017, Jun 20). A Comparative Study of the Structural Dynamics of Four Terminal Uridylyl Transferases. *Genes (Basel), 8*(6). https://doi.org/10.3390/genes8060166

Choudhury, N. R., Nowak, J. S., Zuo, J., Rappsilber, J., Spoel, S. H., & Michlewski, G. (2014, Nov 20). Trim25 Is an RNA-Specific Activator of Lin28a/TuT4-Mediated Uridylation. *Cell Rep, 9*(4), 1265-1272. https://doi.org/10.1016/j.celrep.2014.10.017

Collaborative Computational Project, N. (1994, Sep 1). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr, 50*(Pt 5), 760-763. https://doi.org/10.1107/S0907444994003112

Davis-Dusenbery, B. N., & Hata, A. (2010, Nov). MicroRNA in Cancer: The Involvement of Aberrant MicroRNA Biogenesis Regulatory Pathways. *Genes Cancer, 1*(11), 1100-1114. https://doi.org/10.1177/1947601910396213

De Almeida, C., Scheer, H., Zuber, H., & Gagliardi, D. (2018, Jan). RNA uridylation: a key posttranscriptional modification shaping the coding and noncoding transcriptome. *Wiley Interdiscip Rev RNA, 9*(1). https://doi.org/10.1002/wrna.1440

Dhanoa, T. S., & Housner, J. A. (2007, Jul). Ribose: more than a simple sugar? *Curr Sports Med Rep, 6*(4), 254-257. https://doi.org/10.1007/s11932-007-0041-8

Doshi, R., Chen, B. R., Vibat, C. R., Huang, N., Lee, C. W., & Chang, G. (2014, Oct 24). In vitro nanobody discovery for integral membrane protein targets. *Sci Rep, 4*, 6760. https://doi.org/10.1038/srep06760

Dou, J., Zhang, H., Chen, R., Shu, Z., Yuan, H., Zhao, X., Wang, Y., Huang, J., Zhou, A., & Yu, J. (2020, Sep). SUMOylation modulates the LIN28A-let-7 signaling pathway in response to cellular stresses in cancer cells. *Mol Oncol, 14*(9), 2288-2312. https://doi.org/10.1002/1878-0261.12694

Drygin Iu, F., Shabanov, A. A., & Bogdanov, A. A. (1988, Aug). [An enzyme hydrolyzing the covalent bond between RNA and VPg of picornaviruses. Substrates and methods of analysis of activity]. *Biokhimiia, 53*(8), 1371-1379. https://www.ncbi.nlm.nih.gov/pubmed/2847820 (Ferment, gidroliziruiushchii kovalentnuiu sviaz' mezhdu RNK i VPg pikornavirusov. Substraty i metody analiza aktivnosti.)

Drygin Iu, F., Vartapetian, A. B., & Chumakov, K. M. (1979, Jul-Aug). [Covalent bond between RNA and protein in encephalomyocarditis virus]. *Mol Biol (Mosk), 13*(4), 777-787. https://www.ncbi.nlm.nih.gov/pubmed/224308 (Kovalentnoe soedinenie RNK i belka v viruse zntsefalomnokardita.)

Duman, G., Sariakcali, B., Ersan, S., & Bakir, S. (2022). The impact of Dicer, Drosha, and Exportin-5 levels in polycystic ovary syndrome (PCOS) diagnosis and phenotyping. *Endokrynol Pol, 73*(1), 8-15. https://doi.org/10.5603/EP.a2021.0096

Dworkin, M. B., & Dworkin-Rastl, E. (1985, Dec). Changes in RNA titers and polyadenylation during oogenesis and oocyte maturation in Xenopus laevis. *Dev Biol, 112*(2), 451-457. https://doi.org/10.1016/0012-1606(85)90417-8

Eads, B. D., & Hand, S. C. (2003, Oct). Mitochondrial mRNA stability and polyadenylation during anoxia-induced quiescence in the brine shrimp Artemia franciscana. *J Exp Biol, 206*(Pt 20), 3681-3692. https://doi.org/10.1242/jeb.00595

Emsley, P., & Cowtan, K. (2004, Dec). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr, 60*(Pt 12 Pt 1), 2126-2132. https://doi.org/10.1107/S0907444904019158

Fabian, M. R., & White, K. A. (2004, Jul 9). 5'-3' RNA-RNA interaction facilitates cap- and poly(A)

tail-independent translation of tomato bushy stunt virus mrna: a potential common mechanism for tombusviridae. *J Biol Chem, 279*(28), 28862-28872. https://doi.org/10.1074/jbc.M401272200

Faehnle, C. R., Walleshauser, J., & Joshua-Tor, L. (2014, Oct 9). Mechanism of Dis3l2 substrate recognition in the Lin28-let-7 pathway. *Nature, 514*(7521), 252-256. https://doi.org/10.1038/nature13553

Faehnle, C. R., Walleshauser, J., & Joshua-Tor, L. (2017a, Aug). Multi-domain utilization by TUT4 and TUT7 in control of let-7 biogenesis. *Nat Struct Mol Biol, 24*(8), 658-665. https://doi.org/10.1038/nsmb.3428

Faehnle, C. R., Walleshauser, J., & Joshua-Tor, L. (2017b, Jul 03). Multi-domain utilization by TUT4 and TUT7 in control of let-7 biogenesis. *Nat Struct Mol Biol*. https://doi.org/10.1038/nsmb.3428

Fairchild, C. L. A., Cheema, S. K., Wong, J., Hino, K., Simo, S., & La Torre, A. (2019, Oct 25). Let-7 regulates cell cycle dynamics in the developing cerebral cortex and retina. *Sci Rep, 9*(1), 15336. https://doi.org/10.1038/s41598-019-51703-x

Fairchild, C. L. A., Cheema, S. K., Wong, J., Hino, K., Simo, S., & La Torre, A. (2021, Jan 28). Author Correction: Let-7 regulates cell cycle dynamics in the developing cerebral cortex and retina. *Sci Rep, 11*(1), 2872. https://doi.org/10.1038/s41598-021-82224-1

Faria, M., & Ulrich, H. (2008, Apr). Sugar boost: when ribose modifications improve oligonucleotide performance. *Curr Opin Mol Ther, 10*(2), 168-175. https://www.ncbi.nlm.nih.gov/pubmed/18386229

Ferneza, S., Fetsych, M., Shuliak, R., Makukh, H., Volodko, N., Yarema, R., & Fetsych, T. (2021). Clinical significance of microRNA-200 and let-7 families expression assessment in patients with ovarian cancer. *Ecancermedicalscience, 15*, 1249. https://doi.org/10.3332/ecancer.2021.1249

Frechin, L., Holvec, S., von Loeffelholz, O., Hazemann, I., & Klaholz, B. P. (2022, Oct 12). High-resolution cryo-EM performance comparison of two latest-generation cryo electron microscopes on the human ribosome. *J Struct Biol, 215*(1), 107905. https://doi.org/10.1016/j.jsb.2022.107905

Fu, X., Meng, Z., Liang, W., Tian, Y., Wang, X., Han, W., Lou, G., Wang, X., Lou, F., Yen, Y., Yu, H., Jove, R., & Huang, W. (2014, Aug 21). miR-26a enhances miRNA biogenesis by targeting Lin28B and Zcchc11 to suppress tumor growth and metastasis. *Oncogene, 33*(34), 4296-4306. https://doi.org/10.1038/onc.2013.385

Gildea, R. J., Beilsten-Edmands, J., Axford, D., Horrell, S., Aller, P., Sandy, J., Sanchez-Weatherby, J.,

Owen, C. D., Lukacik, P., Strain-Damerell, C., Owen, R. L., Walsh, M. A., & Winter, G. (2022, Jun 1). xia2.multiplex: a multi-crystal data-analysis pipeline. *Acta Crystallogr D Struct Biol, 78*(Pt 6), 752-769. https://doi.org/10.1107/S2059798322004399

Goddard, T. D., Huang, C. C., & Ferrin, T. E. (2005, Mar). Software extensions to UCSF chimera for interactive visualization of large molecular assemblies. *Structure, 13*(3), 473-482. https://doi.org/10.1016/j.str.2005.01.006

Gradia, S. D., Ishida, J. P., Tsai, M. S., Jeans, C., Tainer, J. A., & Fuss, J. O. (2017). MacroBac: New Technologies for Robust and Efficient Large-Scale Production of Recombinant Multiprotein Complexes. *Methods Enzymol, 592*, 1-26. https://doi.org/10.1016/bs.mie.2017.03.008

Gregory, R. I., Yan, K. P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., & Shiekhattar, R. (2004, Nov 11). The Microprocessor complex mediates the genesis of microRNAs. *Nature, 432*(7014), 235-240. https://doi.org/10.1038/nature03120

Haase, A. D., Jaskiewicz, L., Zhang, H., Laine, S., Sack, R., Gatignol, A., & Filipowicz, W. (2005, Oct). TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing. *EMBO Rep, 6*(10), 961-967. https://doi.org/10.1038/sj.embor.7400509

Hagan, J. P., & Croce, C. M. (2007). MicroRNAs in carcinogenesis. *Cytogenet Genome Res, 118*(2-4), 252-259. https://doi.org/10.1159/000108308

Hagan, J. P., Piskounova, E., & Gregory, R. I. (2009, Oct). Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. *Nat Struct Mol Biol, 16*(10), 1021-1025. https://doi.org/10.1038/nsmb.1676

Hallett, R. M., & Hassell, J. A. (2011, Mar 31). E2F1 and KIAA0191 expression predicts breast cancer patient survival. *BMC Res Notes, 4*, 95. https://doi.org/10.1186/1756-0500-4-95

Hameiri-Grossman, M., Porat-Klein, A., Yaniv, I., Ash, S., Cohen, I. J., Kodman, Y., Haklai, R., Elad-Sfadia, G., Kloog, Y., Chepurko, E., Feinmesser, M., Issakov, J., Sher, O., Luria, D., Kollender, Y., Weizman, A., & Avigad, S. (2015, Oct 20). The association between let-7, RAS and HIF-1alpha in Ewing Sarcoma tumor growth. *Oncotarget, 6*(32), 33834-33848. https://doi.org/10.18632/oncotarget.5616

Harmsen, M. M., & De Haard, H. J. (2007, Nov). Properties, production, and applications of camelid single-domain antibody fragments. *Appl Microbiol Biotechnol, 77*(1), 13-22. https://doi.org/10.1007/s00253-007-1142-2

Harris, M. E., Bohni, R., Schneiderman, M. H., Ramamurthy, L., Schumperli, D., & Marzluff, W. F. (1991, May). Regulation of histone mRNA in the unperturbed cell cycle: evidence

suggesting control at two posttranscriptional steps. *Mol Cell Biol, 11*(5), 2416-2424. https://doi.org/10.1128/mcb.11.5.2416-2424.1991

Heo, I., Ha, M., Lim, J., Yoon, M. J., Park, J. E., Kwon, S. C., Chang, H., & Kim, V. N. (2012, Oct 26). Mono-uridylation of pre-microRNA as a key step in the biogenesis of group II let-7 microRNAs. *Cell, 151*(3), 521-532. https://doi.org/10.1016/j.cell.2012.09.022

Heo, I., Joo, C., Cho, J., Ha, M., Han, J., & Kim, V. N. (2008, Oct 24). Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA. *Mol Cell, 32*(2), 276-284. https://doi.org/10.1016/j.molcel.2008.09.014

Heo, I., Joo, C., Kim, Y. K., Ha, M., Yoon, M. J., Cho, J., Yeom, K. H., Han, J., & Kim, V. N. (2009, Aug 21). TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. *Cell, 138*(4), 696-708. https://doi.org/10.1016/j.cell.2009.08.002

Hirschmugl, C. J., & Gough, K. M. (2012, May). Fourier transform infrared spectrochemical imaging: review of design and applications with a focal plane array and multiple beam synchrotron radiation source. *Appl Spectrosc, 66*(5), 475-491. https://doi.org/10.1366/12-06629

Holmquist, C. E., He, W., Meganck, R. M. C., & Marzluff, W. F. (2022, Aug 30). Knock-outs of TUT7 and 3'hExo show that they cooperate in histone mRNA maintenance and degradation. *RNA*. https://doi.org/10.1261/rna.079233.122

Hough, M. A., & Wilson, K. S. (2018, Feb 1). From crystal to structure with CCP4. *Acta Crystallogr D Struct Biol, 74*(Pt 2), 67. https://doi.org/10.1107/S2059798317017557

Hutvagner, G., McLachlan, J., Pasquinelli, A. E., Balint, E., Tuschl, T., & Zamore, P. D. (2001, Aug 3). A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science, 293*(5531), 834-838. https://doi.org/10.1126/science.1062961

Ilari, A., & Savino, C. (2017). A Practical Approach to Protein Crystallography. *Methods Mol Biol, 1525*, 47-78. https://doi.org/10.1007/978-1-4939-6622-6_3

Ioannidis, P., Mahaira, L. G., Perez, S. A., Gritzapis, A. D., Sotiropoulou, P. A., Kavalakis, G. J., Antsaklis, A. I., Baxevanis, C. N., & Papamichail, M. (2005, May 20). CRD-BP/IMP1 expression characterizes cord blood CD34+ stem cells and affects c-myc and IGF-II expression in MCF-7 cancer cells. *J Biol Chem, 280*(20), 20086-20093. https://doi.org/10.1074/jbc.M410036200

Ishida, T., & Kinoshita, K. (2007, Jul). PrDOS: prediction of disordered protein regions from amino acid sequence. *Nucleic Acids Res, 35*(Web Server issue), W460-464. https://doi.org/10.1093/nar/gkm363

Jaeger, S., Martin, F., Rudinger-Thirion, J., Giege, R., & Eriani, G. (2006). Binding of human SLBP on the 3'-UTR of histone precursor H4-12 mRNA induces structural rearrangements that enable U7 snRNA anchoring. *Nucleic Acids Res, 34*(17), 4987-4995. https://doi.org/10.1093/nar/gkl666

Jiang, H., Coleman, J., Miskimins, R., Srinivasan, R., & Miskimins, W. K. (2007). Cap-independent translation through the p27 5'-UTR. *Nucleic Acids Res, 35*(14), 4767-4778. https://doi.org/10.1093/nar/gkm512

Johnson, S. M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., Labourier, E., Reinert, K. L., Brown, D., & Slack, F. J. (2005, Mar 11). RAS is regulated by the let-7 microRNA family. *Cell, 120*(5), 635-647. https://doi.org/10.1016/j.cell.2005.01.014

Jonas, S., Christie, M., Peter, D., Bhandari, D., Loh, B., Huntzinger, E., Weichenrieder, O., & Izaurralde, E. (2014, Jul). An asymmetric PAN3 dimer recruits a single PAN2 exonuclease to mediate mRNA deadenylation and decay. *Nat Struct Mol Biol, 21*(7), 599-608. https://doi.org/10.1038/nsmb.2837

Jones, M. R., Quinton, L. J., Blahna, M. T., Neilson, J. R., Fu, S., Ivanov, A. R., Wolf, D. A., & Mizgerd, J. P. (2009, Sep). Zcchc11-dependent uridylation of microRNA directs cytokine expression. *Nat Cell Biol, 11*(9), 1157-1163. https://doi.org/10.1038/ncb1931

Kabsch, W. (2010, Feb). Xds. *Acta Crystallogr D Biol Crystallogr, 66*(Pt 2), 125-132. https://doi.org/10.1107/S0907444909047337

Kang, M., Lee, K. H., Lee, H. S., Jeong, C. W., Ku, J. H., Kim, H. H., & Kwak, C. (2017). Concurrent treatment with simvastatin and NF-kappaB inhibitor in human castration-resistant prostate cancer cells exerts synergistic anti-cancer effects via control of the NF-kappaB/LIN28/let-7 miRNA signaling pathway. *PLoS One, 12*(9), e0184644. https://doi.org/10.1371/journal.pone.0184644

Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish, R. G., Wyckoff, H., & Phillips, D. C. (1958, Mar 8). A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature, 181*(4610), 662-666. https://doi.org/10.1038/181662a0

Kim, B., Ha, M., Loeff, L., Chang, H., Simanshu, D. K., Li, S., Fareh, M., Patel, D. J., Joo, C., & Kim, V. N. (2015, Jul 2). TUT7 controls the fate of precursor microRNAs by using three different uridylation mechanisms. *EMBO J, 34*(13), 1801-1815. https://doi.org/10.15252/embj.201590931

Kim, G. W., & Siddiqui, A. (2021, Mar). The role of N6-methyladenosine modification in the life cycle and disease pathogenesis of hepatitis B and C viruses. *Exp Mol Med, 53*(3), 339-345. https://doi.org/10.1038/s12276-021-00581-3

Kim, H., Kim, J., Yu, S., Lee, Y. Y., Park, J., Choi, R. J., Yoon, S. J., Kang, S. G., & Kim, V. N. (2020, Jun 18). A Mechanism for microRNA Arm Switching Regulated by Uridylation. *Mol Cell, 78*(6), 1224-1236 e1225. https://doi.org/10.1016/j.molcel.2020.04.030

Kim, H. H., Kuwano, Y., Srikantan, S., Lee, E. K., Martindale, J. L., & Gorospe, M. (2009, Aug 1). HuR recruits let-7/RISC to repress c-Myc expression. *Genes Dev, 23*(15), 1743-1748. https://doi.org/10.1101/gad.1812509

Kim, S. K., Lee, H., Han, K., Kim, S. C., Choi, Y., Park, S. W., Bak, G., Lee, Y., Choi, J. K., Kim, T. K., Han, Y. M., & Lee, D. (2014, Dec 4). SET7/9 methylation of the pluripotency factor LIN28A is a nucleolar localization mechanism that blocks let-7 biogenesis in human ESCs. *Cell Stem Cell, 15*(6), 735-749. https://doi.org/10.1016/j.stem.2014.10.016

Kim, Y., & Kim, V. N. (2012, May 25). MicroRNA factory: RISC assembly from precursor microRNAs. *Mol Cell, 46*(4), 384-386. https://doi.org/10.1016/j.molcel.2012.05.012

Kobilka, B. K., Matsui, H., Kobilka, T. S., Yang-Feng, T. L., Francke, U., Caron, M. G., Lefkowitz, R. J., & Regan, J. W. (1987, Oct 30). Cloning, sequencing, and expression of the gene coding for the human platelet alpha 2-adrenergic receptor. *Science, 238*(4827), 650-656. https://doi.org/10.1126/science.2823383

Koopmann, R., Cupelli, K., Redecke, L., Nass, K., Deponte, D. P., White, T. A., Stellato, F., Rehders, D., Liang, M., Andreasson, J., Aquila, A., Bajt, S., Barthelmess, M., Barty, A., Bogan, M. J., Bostedt, C., Boutet, S., Bozek, J. D., Caleman, C., Coppola, N., Davidsson, J., Doak, R. B., Ekeberg, T., Epp, S. W., Erk, B., Fleckenstein, H., Foucar, L., Graafsma, H., Gumprecht, L., Hajdu, J., Hampton, C. Y., Hartmann, A., Hartmann, R., Hauser, G., Hirsemann, H., Holl, P., Hunter, M. S., Kassemeyer, S., Kirian, R. A., Lomb, L., Maia, F. R., Kimmel, N., Martin, A. V., Messerschmidt, M., Reich, C., Rolles, D., Rudek, B., Rudenko, A., Schlichting, I., Schulz, J., Seibert, M. M., Shoeman, R. L., Sierra, R. G., Soltau, H., Stern, S., Struder, L., Timneanu, N., Ullrich, J., Wang, X., Weidenspointner, G., Weierstall, U., Williams, G. J., Wunderer, C. B., Fromme, P., Spence, J. C., Stehle, T., Chapman, H. N., Betzel, C., & Duszenko, M. (2012, Jan 29). In vivo protein crystallization opens new routes in structural biology. *Nat Methods, 9*(3), 259-262. https://doi.org/10.1038/nmeth.1859

Lackey, P. E., Welch, J. D., & Marzluff, W. F. (2016, Nov). TUT7 catalyzes the uridylation of the 3' end for rapid degradation of histone mRNA. *RNA, 22*(11), 1673-1688. https://doi.org/10.1261/rna.058107.116

Lam, W. W., & Siu, S. W. (2017, Jan 2). PyMOL mControl: Manipulating molecular visualization with mobile devices. *Biochem Mol Biol Educ, 45*(1), 76-83. https://doi.org/10.1002/bmb.20987

Le Pen, J., Jiang, H., Di Domenico, T., Kneuss, E., Kosalka, J., Leung, C., Morgan, M., Much, C., Rudolph, K. L. M., Enright, A. J., O'Carroll, D., Wang, D., & Miska, E. A. (2018, Sep). Terminal uridylyltransferases target RNA viruses as part of the innate immune system. *Nat Struct*

Mol Biol, 25(9), 778-786. https://doi.org/10.1038/s41594-018-0106-9

Lee, Y. S., & Dutta, A. (2007, May 1). The tumor suppressor microRNA let-7 represses the HMGA2 oncogene. *Genes Dev, 21*(9), 1025-1030. https://doi.org/10.1101/gad.1540407

Lewis, B. P., Burge, C. B., & Bartel, D. P. (2005, Jan 14). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell, 120*(1), 15-20. https://doi.org/10.1016/j.cell.2004.12.035

Li, D., Liu, Z., & Ning, G. (2020, Nov 30). [Expression of CDC25A in non-small cell lung cancer and its relationship with let-7 gene]. *Nan Fang Yi Ke Da Xue Xue Bao, 40*(11), 1622-1627. https://doi.org/10.12122/j.issn.1673-4254.2020.11.13

Li, Q. (2020). Application of Fragment-Based Drug Discovery to Versatile Targets. *Front Mol Biosci, 7*, 180. https://doi.org/10.3389/fmolb.2020.00180

Li, S., & Mason, C. E. (2014). The pivotal regulatory landscape of RNA modifications. *Annu Rev Genomics Hum Genet, 15*, 127-150. https://doi.org/10.1146/annurev-genom-090413-025405

Li, Y., Liu, H., Lai, C., Du, X., Su, Z., & Gao, S. (2013, Nov). The Lin28/let-7a/c-Myc pathway plays a role in non-muscle invasive bladder cancer. *Cell Tissue Res, 354*(2), 533-541. https://doi.org/10.1007/s00441-013-1715-6

Lim, J., Ha, M., Chang, H., Kwon, S. C., Simanshu, D. K., Patel, D. J., & Kim, V. N. (2014, Dec 4). Uridylation by TUT4 and TUT7 marks mRNA for degradation. *Cell, 159*(6), 1365-1376. https://doi.org/10.1016/j.cell.2014.10.055

Lim, J., Kim, D., Lee, Y. S., Ha, M., Lee, M., Yeo, J., Chang, H., Song, J., Ahn, K., & Kim, V. N. (2018, Aug 17). Mixed tailing by TENT4A and TENT4B shields mRNA from rapid deadenylation. *Science, 361*(6403), 701-704. https://doi.org/10.1126/science.aam5794

Lin, C. C., Shen, Y. R., Chang, C. C., Guo, X. Y., Young, Y. Y., Lai, T. Y., Yu, I. S., Lee, C. Y., Chuang, T. H., Tsai, H. Y., & Hsu, L. C. (2021, Jun 29). Terminal uridyltransferase 7 regulates TLR4-triggered inflammation by controlling Regnase-1 mRNA uridylation and degradation. *Nat Commun, 12*(1), 3878. https://doi.org/10.1038/s41467-021-24177-7

Lin, D., Fu, Z., Yang, G., Gao, D., Wang, T., Liu, Z., Li, G., & Wang, Y. (2020, Oct 15). Exportin-5 SUMOylation promotes hepatocellular carcinoma progression. *Exp Cell Res, 395*(2), 112219. https://doi.org/10.1016/j.yexcr.2020.112219

Lin, S., & Gregory, R. I. (2015, Jun). MicroRNA biogenesis pathways in cancer. *Nat Rev Cancer, 15*(6), 321-333. https://doi.org/10.1038/nrc3932

Liu, J., Dou, X., Chen, C., Chen, C., Liu, C., Xu, M. M., Zhao, S., Shen, B., Gao, Y., Han, D., & He, C. (2020, Jan 31). N (6)-methyladenosine of chromosome-associated regulatory RNA regulates chromatin state and transcription. *Science, 367*(6477), 580-586. https://doi.org/10.1126/science.aay6018

Liu, Z., Wang, J., Cheng, H., Ke, X., Sun, L., Zhang, Q. C., & Wang, H. W. (2018a, May 17). Cryo-EM Structure of Human Dicer and Its Complexes with a Pre-miRNA Substrate. *Cell, 173*(5), 1191-1203 e1112. https://doi.org/10.1016/j.cell.2018.03.080

Liu, Z., Wang, J., Cheng, H., Ke, X., Sun, L., Zhang, Q. C., & Wang, H. W. (2018b, May 31). Cryo-EM Structure of Human Dicer and Its Complexes with a Pre-miRNA Substrate. *Cell, 173*(6), 1549-1550. https://doi.org/10.1016/j.cell.2018.05.031

Loughlin, F. E., Gebert, L. F., Towbin, H., Brunschweiger, A., Hall, J., & Allain, F. H. (2011, Dec 11). Structural basis of pre-let-7 miRNA recognition by the zinc knuckles of pluripotency factor Lin28. *Nat Struct Mol Biol, 19*(1), 84-89. https://doi.org/10.1038/nsmb.2202

Loughlin, F. E., Gebert, L. F., Towbin, H., Brunschweiger, A., Hall, J., & Allain, F. H. (2012, Jan). Structural basis of pre-let-7 miRNA recognition by the zinc knuckles of pluripotency factor Lin28 [Research Support, Non-U.S. Gov't]. *Nat Struct Mol Biol, 19*(1), 84-89. https://doi.org/10.1038/nsmb.2202

Lunde, B. M., Magler, I., & Meinhart, A. (2012, Oct). Crystal structures of the Cid1 poly (U) polymerase reveal the mechanism for UTP selectivity. *Nucleic Acids Res, 40*(19), 9815-9824. https://doi.org/10.1093/nar/gks740

Luo, L., Zhen, Y., Peng, D., Wei, C., Zhang, X., Liu, X., Han, L., & Zhang, Z. (2022, Jul 18). The role of N6-methyladenosine-modified non-coding RNAs in the pathological process of human cancer. *Cell Death Discov, 8*(1), 325. https://doi.org/10.1038/s41420-022-01113-2

Lyumkis, D. (2019, Mar 29). Challenges and opportunities in cryo-EM single-particle analysis. *J Biol Chem, 294*(13), 5181-5197. https://doi.org/10.1074/jbc.REV118.005602

Ma, Y., Shen, N., Wicha, M. S., & Luo, M. (2021, Sep 14). The Roles of the Let-7 Family of MicroRNAs in the Regulation of Cancer Stemness. *Cells, 10*(9). https://doi.org/10.3390/cells10092415

Madugalle, S. U., Meyer, K., Wang, D. O., & Bredy, T. W. (2020, Dec). RNA N(6)-Methyladenosine and the Regulation of RNA Localization and Function in the Brain. *Trends Neurosci, 43*(12), 1011-1023. https://doi.org/10.1016/j.tins.2020.09.005

Mangus, D. A., Evans, M. C., & Jacobson, A. (2003). Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. *Genome Biol, 4*(7), 223. https://doi.org/10.1186/gb-2003-4-7-223

Mansourian, A. R. (2011, Feb 1). Central dogma in thyroid dysfunction: a review on structure modification of TSHR as a cornerstone for thyroid abnormalities. *Pak J Biol Sci, 14*(3), 170-181. https://doi.org/10.3923/pjbs.2011.170.181

Martin, G., Keller, W., & Doublie, S. (2000, Aug 15). Crystal structure of mammalian poly(A) polymerase in complex with an analog of ATP. *EMBO J, 19*(16), 4193-4203. https://doi.org/10.1093/emboj/19.16.4193

Mayr, F., Schutz, A., Doge, N., & Heinemann, U. (2012, Aug). The Lin28 cold-shock domain remodels pre-let-7 microRNA. *Nucleic Acids Res, 40*(15), 7492-7506. https://doi.org/10.1093/nar/gks355

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., & Read, R. J. (2007, Aug 1). Phaser crystallographic software. *J Appl Crystallogr, 40*(Pt 4), 658-674. https://doi.org/10.1107/S0021889807021206

Medhi, R., Price, J., Furlan, G., Gorges, B., Sapetschnig, A., & Miska, E. A. (2022, Mar). RNA uridyl transferases TUT4/7 differentially regulate miRNA variants depending on the cancer cell type. *RNA, 28*(3), 353-370. https://doi.org/10.1261/rna.078976.121

Menezes, M. R., Balzeau, J., & Hagan, J. P. (2018). 3' RNA Uridylation in Epitranscriptomics, Gene Regulation, and Disease. *Front Mol Biosci, 5*, 61. https://doi.org/10.3389/fmolb.2018.00061

Miller, J. E., & Reese, J. C. (2012, Jul-Aug). Ccr4-Not complex: the control freak of eukaryotic cells. *Crit Rev Biochem Mol Biol, 47*(4), 315-333. https://doi.org/10.3109/10409238.2012.667214

Minoda, Y., Saeki, K., Aki, D., Takaki, H., Sanada, T., Koga, K., Kobayashi, T., Takaesu, G., & Yoshimura, A. (2006, Jun 9). A novel Zinc finger protein, ZCCHC11, interacts with TIFA and modulates TLR signaling. *Biochem Biophys Res Commun, 344*(3), 1023-1030. https://doi.org/10.1016/j.bbrc.2006.04.006

Mondol, V., Ahn, B. C., & Pasquinelli, A. E. (2015, Aug). Splicing remodels the let-7 primary microRNA to facilitate Drosha processing in Caenorhabditis elegans. *RNA, 21*(8), 1396-1403. https://doi.org/10.1261/rna.052118.115

Moore, P. B., & Steitz, T. A. (2003). The structural basis of large ribosomal subunit function. *Annu Rev Biochem, 72*, 813-850. https://doi.org/10.1146/annurev.biochem.72.110601.135450

Moriarty, N. W., Tronrud, D. E., Adams, P. D., & Karplus, P. A. (2016, Jan). A new default restraint library for the protein backbone in Phenix: a conformation-dependent geometry goes mainstream. *Acta Crystallogr D Struct Biol, 72*(Pt 1), 176-179. https://doi.org/10.1107/S2059798315022408

Moss, E. G., Lee, R. C., & Ambros, V. (1997, Mar 7). The cold shock domain protein LIN-28 controls developmental timing in C. elegans and is regulated by the lin-4 RNA. *Cell, 88*(5), 637-646. https://doi.org/10.1016/s0092-8674(00)81906-6

Moss, E. G., & Tang, L. (2003, Jun 15). Conservation of the heterochronic regulator Lin-28, its developmental expression and microRNA complementary sites. *Dev Biol, 258*(2), 432-442. https://doi.org/10.1016/s0012-1606(03)00126-x

Mossanen-Parsi, A., Parisi, D., Browne-Marke, N., Bharudin, I., Connell, S. R., Mayans, O., Fucini, P., Morozov, I. Y., & Caddick, M. X. (2021, Feb). Histone mRNA is subject to 3' uridylation and re-adenylation in Aspergillus nidulans. *Mol Microbiol, 115*(2), 238-254. https://doi.org/10.1111/mmi.14613

Munoz-Tello, P., Gabus, C., & Thore, S. (2012, Jun 6). Functional implications from the Cid1 poly(U) polymerase crystal structure. *Structure, 20*(6), 977-986. https://doi.org/10.1016/j.str.2012.04.006

Nagandla, K., Lin, K. H., Chitra, E., & Jamli, M. (2021, Sep). Role of microRNAs as biomarkers of cervical carcinogenesis: a systematic review. *Obstet Gynecol Sci, 64*(5), 419-436. https://doi.org/10.5468/ogs.21123

Nam, Y., Chen, C., Gregory, R. I., Chou, J. J., & Sliz, P. (2011, Nov 23). Molecular basis for interaction of let-7 microRNAs with Lin28 [Research Support, N.I.H., Extramural
Research Support, Non-U.S. Gov't]. *Cell, 147*(5), 1080-1091. https://doi.org/10.1016/j.cell.2011.10.020

Newman, M. A., Thomson, J. M., & Hammond, S. M. (2008, Aug). Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. *RNA, 14*(8), 1539-1549. https://doi.org/10.1261/rna.1155108

Nguyen, L. H., & Zhu, H. (2015, Jan). Lin28 and let-7 in cell metabolism and cancer. *Transl Pediatr, 4*(1), 4-11. https://doi.org/10.3978/j.issn.2224-4336.2015.01.05

Nissen, P., Hansen, J., Ban, N., Moore, P. B., & Steitz, T. A. (2000, Aug 11). The structural basis of ribosome activity in peptide bond synthesis. *Science, 289*(5481), 920-930. https://doi.org/10.1126/science.289.5481.920

Nombela, P., Miguel-Lopez, B., & Blanco, S. (2021, Jan 18). The role of m(6)A, m(5)C and Psi RNA modifications in cancer: Novel therapeutic opportunities. *Mol Cancer, 20*(1), 18. https://doi.org/10.1186/s12943-020-01263-w

Nomura, M. (1987). The role of RNA and protein in ribosome function: a review of early reconstitution studies and prospects for future studies. *Cold Spring Harb Symp Quant Biol,*

52, 653-663. https://doi.org/10.1101/sqb.1987.052.01.075

Norbury, C. J. (2013, Oct). Cytoplasmic RNA: a case of the tail wagging the dog. *Nat Rev Mol Cell Biol, 14*(10), 643-653. https://doi.org/10.1038/nrm3645

Oh, J. S., Kim, J. J., Byun, J. Y., & Kim, I. A. (2010, Jan 1). Lin28-let7 modulates radiosensitivity of human cancer cells with activation of K-Ras. *Int J Radiat Oncol Biol Phys, 76*(1), 5-8. https://doi.org/10.1016/j.ijrobp.2009.08.028

Oliveira-Mateos, C., Sanchez-Castillo, A., Soler, M., Obiols-Guardia, A., Pineyro, D., Boque-Sastre, R., Calleja-Cervantes, M. E., Castro de Moura, M., Martinez-Cardus, A., Rubio, T., Pelletier, J., Martinez-Iniesta, M., Herrero-Martin, D., Tirado, O. M., Gentilella, A., Villanueva, A., Esteller, M., Farre, L., & Guil, S. (2019, Sep 4). The transcribed pseudogene RPSAP52 enhances the oncofetal HMGA2-IGF2BP2-RAS axis through LIN28B-dependent and independent let-7 inhibition. *Nat Commun, 10*(1), 3979. https://doi.org/10.1038/s41467-019-11910-6

Parisi, S., Passaro, F., Russo, L., Musto, A., Navarra, A., Romano, S., Petrosino, G., & Russo, T. (2017, Mar). Lin28 is induced in primed embryonic stem cells and regulates let-7-independent events. *FASEB J, 31*(3), 1046-1058. https://doi.org/10.1096/fj.201600848R

Piskounova, E., Polytarchou, C., Thornton, J. E., LaPierre, R. J., Pothoulakis, C., Hagan, J. P., Iliopoulos, D., & Gregory, R. I. (2011, Nov 23). Lin28A and Lin28B inhibit let-7 microRNA biogenesis by distinct mechanisms. *Cell, 147*(5), 1066-1079. https://doi.org/10.1016/j.cell.2011.10.039

Posner, R., & Laubenbacher, R. (2020, Feb 7). The contribution of microRNA-mediated regulation to short- and long-term gene expression predictability. *J Theor Biol, 486*, 110055. https://doi.org/10.1016/j.jtbi.2019.110055

Potter-Birriel, J. M., Gonsalvez, G. B., & Marzluff, W. F. (2021, Feb 11). A region of SLBP outside the mRNA-processing domain is essential for deposition of histone mRNA into the Drosophila egg. *J Cell Sci, 134*(3). https://doi.org/10.1242/jcs.251728

Proudfoot, N. J., Furger, A., & Dye, M. J. (2002, Feb 22). Integrating mRNA processing with transcription. *Cell, 108*(4), 501-512. https://doi.org/10.1016/s0092-8674(02)00617-7

Punjani, A., Rubinstein, J. L., Fleet, D. J., & Brubaker, M. A. (2017, Mar). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods, 14*(3), 290-296. https://doi.org/10.1038/nmeth.4169

Qiao, H., Liu, L., Chen, J., Shang, B., & Wang, L. (2022, Sep 29). The functions of N6-methyladenosine (m6A) RNA modifications in colorectal cancer. *Med Oncol, 39*(12), 235. https://doi.org/10.1007/s12032-022-01827-4

Quagliato, L. A., & Nardi, A. E. (2018, Mar 1). Cytokine alterations in panic disorder: A systematic review. *J Affect Disord, 228*, 91-96. https://doi.org/10.1016/j.jad.2017.11.094

Rahkonen, N., Stubb, A., Malonzo, M., Edelman, S., Emani, M. R., Narva, E., Lahdesmaki, H., Ruohola-Baker, H., Lahesmaa, R., & Lund, R. (2016, Nov). Mature Let-7 miRNAs fine tune expression of LIN28B in pluripotent human embryonic stem cells. *Stem Cell Res, 17*(3), 498-503. https://doi.org/10.1016/j.scr.2016.09.025

Ranjbar, R., Hesari, A., Ghasemi, F., & Sahebkar, A. (2018, Sep). Expression of microRNAs and IRAK1 pathway genes are altered in gastric cancer patients with Helicobacter pylori infection. *J Cell Biochem, 119*(9), 7570-7576. https://doi.org/10.1002/jcb.27067

Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., Horvitz, H. R., & Ruvkun, G. (2000, Feb 24). The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature, 403*(6772), 901-906. https://doi.org/10.1038/35002607

Reinisch, K. M., & Wolin, S. L. (2007, Apr). Emerging themes in non-coding RNA quality control. *Curr Opin Struct Biol, 17*(2), 209-214. https://doi.org/10.1016/j.sbi.2007.03.012

Richter, J. D. (1999, Jun). Cytoplasmic polyadenylation in development and beyond. *Microbiol Mol Biol Rev, 63*(2), 446-456. https://doi.org/10.1128/MMBR.63.2.446-456.1999

Rissland, O. S., Mikulasova, A., & Norbury, C. J. (2007, May). Efficient RNA polyuridylation by noncanonical poly(A) polymerases. *Mol Cell Biol, 27*(10), 3612-3624. https://doi.org/10.1128/MCB.02209-06

Rissland, O. S., & Norbury, C. J. (2009, Jun). Decapping is preceded by 3' uridylation in a novel pathway of bulk mRNA turnover. *Nat Struct Mol Biol, 16*(6), 616-623. https://doi.org/10.1038/nsmb.1601

Roos, M., Rebhan, M. A., Lucic, M., Pavlicek, D., Pradere, U., Towbin, H., Civenni, G., Catapano, C. V., & Hall, J. (2015, Jan). Short loop-targeting oligoribonucleotides antagonize Lin28 and enable pre-let-7 processing and suppression of cell growth in let-7-deficient cancer cells. *Nucleic Acids Res, 43*(2), e9. https://doi.org/10.1093/nar/gku1090

Rougvie, A. E. (2001, Sep). Control of developmental timing in animals. *Nat Rev Genet, 2*(9), 690-701. https://doi.org/10.1038/35088566

Roush, S., & Slack, F. J. (2008, Oct). The let-7 family of microRNAs. *Trends Cell Biol, 18*(10), 505-516. https://doi.org/10.1016/j.tcb.2008.07.007

Ruby, J. G., Jan, C., Player, C., Axtell, M. J., Lee, W., Nusbaum, C., Ge, H., & Bartel, D. P. (2006, Dec

15). Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in C. elegans. *Cell, 127*(6), 1193-1207. https://doi.org/10.1016/j.cell.2006.10.040

Rybak, A., Fuchs, H., Smirnova, L., Brandt, C., Pohl, E. E., Nitsch, R., & Wulczyn, F. G. (2008, Aug). A feedback loop comprising lin-28 and let-7 controls pre-let-7 maturation during neural stem-cell commitment. *Nat Cell Biol, 10*(8), 987-993. https://doi.org/10.1038/ncb1759

Saenger, W., Riecke, J., & Suck, D. (1975, Apr 25). A structural model for the polyadenylic acid single helix. *J Mol Biol, 93*(4), 529-534. https://doi.org/10.1016/0022-2836(75)90244-2

Scheres, S. H. (2012, Dec). RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol, 180*(3), 519-530. https://doi.org/10.1016/j.jsb.2012.09.006

Schwarzhans, J. P., Wibberg, D., Winkler, A., Kalinowski, J., & Friehs, K. (2017, Sep 28). Complete Draft Genome Sequence of Escherichia coli KRX, a Strain for Efficient Cloning and High-Yield Expression of Proteins under Control of the T7 RNA Polymerase. *Genome Announc, 5*(39). https://doi.org/10.1128/genomeA.00933-17

Shaikh, T. R., Gao, H., Baxter, W. T., Asturias, F. J., Boisset, N., Leith, A., & Frank, J. (2008). SPIDER image processing for single-particle reconstruction of biological macromolecules from electron micrographs. *Nat Protoc, 3*(12), 1941-1974. https://doi.org/10.1038/nprot.2008.156

Shi, Y. (2014, Nov 20). A glimpse of structural biology through X-ray crystallography. *Cell, 159*(5), 995-1014. https://doi.org/10.1016/j.cell.2014.10.051

Shyh-Chang, N., & Daley, G. Q. (2013, Apr 4). Lin28: primal regulator of growth and metabolism in stem cells. *Cell Stem Cell, 12*(4), 395-406. https://doi.org/10.1016/j.stem.2013.03.005

Singh, R., & Saini, N. (2012, Mar 15). Downregulation of BCL2 by miRNAs augments drug-induced apoptosis--a combined computational and experimental approach. *J Cell Sci, 125*(Pt 6), 1568-1578. https://doi.org/10.1242/jcs.095976

Slater, D. W., Slater, I., & Gillespie, D. (1972, Dec 8). Post-fertilization synthesis of polyadenylic acid in sea urchin embryos. *Nature, 240*(5380), 333-337. https://doi.org/10.1038/240333a0

Smyth, M. S., & Martin, J. H. (2000, Feb). x ray crystallography. *Mol Pathol, 53*(1), 8-14. https://doi.org/10.1136/mp.53.1.8

Stark, H. (2010). GraFix: stabilization of fragile macromolecular complexes for single particle cryo-EM. *Methods Enzymol, 481*, 109-126. https://doi.org/10.1016/S0076-6879(10)81005-5

Steinemann, D., Tauscher, M., Praulich, I., Niemeyer, C. M., Flotho, C., & Schlegelberger, B. (2010,

Sep). Mutations in the let-7 binding site - a mechanism of RAS activation in juvenile myelomonocytic leukemia? *Haematologica, 95*(9), 1616. https://doi.org/10.3324/haematol.2010.024984

Stewart, M. (2019, Mar 1). Polyadenylation and nuclear export of mRNAs. *J Biol Chem, 294*(9), 2977-2987. https://doi.org/10.1074/jbc.REV118.005594

Su, K. Y., & Lee, W. L. (2020, Jan 1). Fourier Transform Infrared Spectroscopy as a Cancer Screening and Diagnostic Tool: A Review and Prospects. *Cancers (Basel), 12*(1). https://doi.org/10.3390/cancers12010115

Swartwout, S. G., & Kinniburgh, A. J. (1989, Jan). c-myc RNA degradation in growing and differentiating cells: possible alternate pathways. *Mol Cell Biol, 9*(1), 288-295. https://doi.org/10.1128/mcb.9.1.288-295.1989

Takamizawa, J., Konishi, H., Yanagisawa, K., Tomida, S., Osada, H., Endoh, H., Harano, T., Yatabe, Y., Nagino, M., Nimura, Y., Mitsudomi, T., & Takahashi, T. (2004, Jun 1). Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. *Cancer Res, 64*(11), 3753-3756. https://doi.org/10.1158/0008-5472.CAN-04-0637

Tang, G., Peng, L., Baldwin, P. R., Mann, D. S., Jiang, W., Rees, I., & Ludtke, S. J. (2007, Jan). EMAN2: an extensible image processing suite for electron microscopy. *J Struct Biol, 157*(1), 38-46. https://doi.org/10.1016/j.jsb.2006.05.009

Tang, T. T. L., Stowell, J. A. W., Hill, C. H., & Passmore, L. A. (2019, Oct). Author Correction: The intrinsic structure of poly(A) RNA determines the specificity of Pan2 and Caf1 deadenylases. *Nat Struct Mol Biol, 26*(10), 988. https://doi.org/10.1038/s41594-019-0295-x

Thomson, J. M., Newman, M., Parker, J. S., Morin-Kensicki, E. M., Wright, T., & Hammond, S. M. (2006, Aug 15). Extensive post-transcriptional regulation of microRNAs and its implications for cancer. *Genes Dev, 20*(16), 2202-2207. https://doi.org/10.1101/gad.1444406

Thornton, J. E., Chang, H. M., Piskounova, E., & Gregory, R. I. (2012, Oct). Lin28-mediated control of let-7 microRNA expression by alternative TUTases Zcchc11 (TUT4) and Zcchc6 (TUT7). *RNA, 18*(10), 1875-1885. https://doi.org/10.1261/rna.034538.112

Thornton, J. E., & Gregory, R. I. (2012, Sep). How does Lin28 let-7 control development and disease? [Research Support, N.I.H., Extramural
Research Support, Non-U.S. Gov't
Research Support, U.S. Gov't, Non-P.H.S.
Review]. *Trends Cell Biol, 22*(9), 474-482. https://doi.org/10.1016/j.tcb.2012.06.001

Tomikawa, C. (2018, Dec 17). 7-Methylguanosine Modifications in Transfer RNA (tRNA). *Int J Mol Sci, 19*(12). https://doi.org/10.3390/ijms19124080

Torres Fernandez, L. A., Mitschka, S., Ulas, T., Weise, S., Dahm, K., Becker, M., Handler, K., Beyer, M., Windhausen, J., Schultze, J. L., & Kolanus, W. (2021, May 11). The stem cell-specific protein TRIM71 inhibits maturation and activity of the pro-differentiation miRNA let-7 via two independent molecular mechanisms. *RNA*. https://doi.org/10.1261/rna.078696.121

Triboulet, R., Pirouz, M., & Gregory, R. I. (2015, Oct 13). A Single Let-7 MicroRNA Bypasses LIN28-Mediated Repression. *Cell Rep, 13*(2), 260-266. https://doi.org/10.1016/j.celrep.2015.08.086

Tsanov, K. M., Pearson, D. S., Wu, Z., Han, A., Triboulet, R., Seligson, M. T., Powers, J. T., Osborne, J. K., Kane, S., Gygi, S. P., Gregory, R. I., & Daley, G. Q. (2017, Jan). LIN28 phosphorylation by MAPK/ERK couples signalling to the post-transcriptional control of pluripotency. *Nat Cell Biol, 19*(1), 60-67. https://doi.org/10.1038/ncb3453

Uchanski, T., Masiulis, S., Fischer, B., Kalichuk, V., Lopez-Sanchez, U., Zarkadas, E., Weckener, M., Sente, A., Ward, P., Wohlkonig, A., Zogg, T., Remaut, H., Naismith, J. H., Nury, H., Vranken, W., Aricescu, A. R., Pardon, E., & Steyaert, J. (2021, Jan). Megabodies expand the nanobody toolkit for protein structure determination by single-particle cryo-EM. *Nat Methods, 18*(1), 60-68. https://doi.org/10.1038/s41592-020-01001-6

Ustianenko, D., Chiu, H. S., Treiber, T., Weyn-Vanhentenryck, S. M., Treiber, N., Meister, G., Sumazin, P., & Zhang, C. (2018, Jul 19). LIN28 Selectively Modulates a Subclass of Let-7 MicroRNAs. *Mol Cell, 71*(2), 271-283 e275. https://doi.org/10.1016/j.molcel.2018.06.029

Ustianenko, D., Hrossova, D., Potesil, D., Chalupnikova, K., Hrazdilova, K., Pachernik, J., Cetkovska, K., Uldrijan, S., Zdrahal, Z., & Vanacova, S. (2013, Dec). Mammalian DIS3L2 exoribonuclease targets the uridylated precursors of let-7 miRNAs. *RNA, 19*(12), 1632-1638. https://doi.org/10.1261/rna.040055.113

Van Driessche, A. E. S., Van Gerven, N., Bomans, P. H. H., Joosten, R. R. M., Friedrich, H., Gil-Carton, D., Sommerdijk, N., & Sleutel, M. (2018, Apr 4). Molecular nucleation mechanisms and control strategies for crystal polymorph selection. *Nature, 556*(7699), 89-94. https://doi.org/10.1038/nature25971

Wang, L., Nam, Y., Lee, A. K., Yu, C., Roth, K., Chen, C., Ransey, E. M., & Sliz, P. (2017, Mar 14). LIN28 Zinc Knuckle Domain Is Required and Sufficient to Induce let-7 Oligouridylation. *Cell Rep, 18*(11), 2664-2675. https://doi.org/10.1016/j.celrep.2017.02.044

Wang, L. X., Wang, J., Qu, T. T., Zhang, Y., & Shen, Y. F. (2014, Jun). Reversible acetylation of Lin28 mediated by PCAF and SIRT1. *Biochim Biophys Acta, 1843*(6), 1188-1195.

https://doi.org/10.1016/j.bbamcr.2014.03.001

Wang, S. W., Toda, T., MacCallum, R., Harris, A. L., & Norbury, C. (2000, May). Cid1, a fission yeast protein required for S-M checkpoint control when DNA polymerase delta or epsilon is inactivated. *Mol Cell Biol, 20*(9), 3234-3244. https://doi.org/10.1128/MCB.20.9.3234-3244.2000

Wang, Y., Yu, Y., Pang, Y., Yu, H., Zhang, W., Zhao, X., & Yu, J. (2021, Dec). The distinct roles of zinc finger CCHC-type (ZCCHC) superfamily proteins in the regulation of RNA metabolism. *RNA Biol, 18*(12), 2107-2126. https://doi.org/10.1080/15476286.2021.1909320

Warkocki, Z., Krawczyk, P. S., Adamska, D., Bijata, K., Garcia-Perez, J. L., & Dziembowski, A. (2018, Sep 6). Uridylation by TUT4/7 Restricts Retrotransposition of Human LINE-1s. *Cell, 174*(6), 1537-1548 e1529. https://doi.org/10.1016/j.cell.2018.07.022

Warkocki, Z., Liudkovska, V., Gewartowska, O., Mroczek, S., & Dziembowski, A. (2018, Nov 5). Terminal nucleotidyl transferases (TENTs) in mammalian RNA metabolism. *Philos Trans R Soc Lond B Biol Sci, 373*(1762). https://doi.org/10.1098/rstb.2018.0162

Weill, L., Belloc, E., Bava, F. A., & Mendez, R. (2012, Jun 5). Translational control by changes in poly(A) tail length: recycling mRNAs. *Nat Struct Mol Biol, 19*(6), 577-585. https://doi.org/10.1038/nsmb.2311

Wery, J. P., & Schevitz, R. W. (1997, Oct). New trends in macromolecular X-ray crystallography. *Curr Opin Chem Biol, 1*(3), 365-369. https://doi.org/10.1016/s1367-5931(97)80075-0

Whitfield, M. L., Kaygun, H., Erkmann, J. A., Townley-Tilson, W. H., Dominski, Z., & Marzluff, W. F. (2004). SLBP is associated with histone mRNA on polyribosomes as a component of the histone mRNP. *Nucleic Acids Res, 32*(16), 4833-4842. https://doi.org/10.1093/nar/gkh798

Wielgos, M. E., Rajbhandari, R., Cooper, T. S., Wei, S., Nozell, S., & Yang, E. S. (2017, Mar). Let-7 Status Is Crucial for PARP1 Expression in HER2-Overexpressing Breast Tumors. *Mol Cancer Res, 15*(3), 340-347. https://doi.org/10.1158/1541-7786.MCR-16-0287-T

Winter, G., Lobley, C. M., & Prince, S. M. (2013, Jul). Decision making in xia2. *Acta Crystallogr D Biol Crystallogr, 69*(Pt 7), 1260-1273. https://doi.org/10.1107/S0907444913015308

Wolf, J., & Passmore, L. A. (2014, Feb). mRNA deadenylation by Pan2-Pan3. *Biochem Soc Trans, 42*(1), 184-187. https://doi.org/10.1042/BST20130211

Wood, S., Willbanks, A., & Cheng, J. X. (2021). The Role of RNA Modifications and RNA-modifying Proteins in Cancer Therapy and Drug Resistance. *Curr Cancer Drug Targets, 21*(4), 326-352. https://doi.org/10.2174/1568009621666210127092828

Wu, A., Wu, K., Li, J., Mo, Y., Lin, Y., Wang, Y., Shen, X., Li, S., Li, L., & Yang, Z. (2015, Mar 31). Let-7a inhibits migration, invasion and epithelial-mesenchymal transition by targeting HMGA2 in nasopharyngeal carcinoma. *J Transl Med, 13*, 105. https://doi.org/10.1186/s12967-015-0462-8

Wu, K., Ahmad, T., & Eri, R. (2022, Mar 27). LIN28A: A multifunctional versatile molecule with future therapeutic potential. *World J Biol Chem, 13*(2), 35-46. https://doi.org/10.4331/wjbc.v13.i2.35

Wu, M., & Lander, G. C. (2020, Oct 6). Present and Emerging Methodologies in Cryo-EM Single-Particle Analysis. *Biophys J, 119*(7), 1281-1289. https://doi.org/10.1016/j.bpj.2020.08.027

Wu, X., & Rapoport, T. A. (2021, Oct 12). Cryo-EM structure determination of small proteins by nanobody-binding scaffolds (Legobodies). *Proc Natl Acad Sci U S A, 118*(41). https://doi.org/10.1073/pnas.2115001118

Xin, S., Deng, Y., Mao, J., Wang, T., Liu, J., Wang, S., Song, X., Song, W., & Liu, X. (2022). Characterization of 7-Methylguanosine Identified Biochemical Recurrence and Tumor Immune Microenvironment in Prostate Cancer. *Front Oncol, 12*, 900203. https://doi.org/10.3389/fonc.2022.900203

Xu, B., Zhang, K., & Huang, Y. (2009, Mar). Lin28 modulates cell growth and associates with a subset of cell cycle regulator mRNAs in mouse embryonic stem cells. *RNA, 15*(3), 357-361. https://doi.org/10.1261/rna.1368009

Xu, Y., & Dang, S. (2022). Recent Technical Advances in Sample Preparation for Single-Particle Cryo-EM. *Front Mol Biosci, 9*, 892459. https://doi.org/10.3389/fmolb.2022.892459

Xue, C., Chu, Q., Zheng, Q., Jiang, S., Bao, Z., Su, Y., Lu, J., & Li, L. (2022, Apr 28). Role of main RNA modifications in cancer: N(6)-methyladenosine, 5-methylcytosine, and pseudouridine. *Signal Transduct Target Ther, 7*(1), 142. https://doi.org/10.1038/s41392-022-01003-0

Yamashita, S., Nagaike, T., & Tomita, K. (2019, Apr 29). Crystal structure of the Lin28-interacting module of human terminal uridylyltransferase that regulates let-7 expression. *Nat Commun, 10*(1), 1960. https://doi.org/10.1038/s41467-019-09966-5

Yang, D. H., & Moss, E. G. (2003, Dec). Temporally regulated expression of Lin-28 in diverse tissues of the developing mouse. *Gene Expr Patterns, 3*(6), 719-726. https://doi.org/10.1016/s1567-133x(03)00140-6

Yashiro, Y., & Tomita, K. (2018). Function and Regulation of Human Terminal Uridylyltransferases. *Front Genet, 9*, 538. https://doi.org/10.3389/fgene.2018.00538

Yates, L. A., Durrant, B. P., Fleurdepine, S., Harlos, K., Norbury, C. J., & Gilbert, R. J. (2015, Mar 11).

Structural plasticity of Cid1 provides a basis for its distributive RNA terminal uridylyl transferase activity. *Nucleic Acids Res, 43*(5), 2968-2979. https://doi.org/10.1093/nar/gkv122

Yates, L. A., Fleurdepine, S., Rissland, O. S., De Colibus, L., Harlos, K., Norbury, C. J., & Gilbert, R. J. C. (2012, Aug). Structural basis for the activity of a cytoplasmic RNA terminal uridylyl transferase. *Nat Struct Mol Biol, 19*(8), 782-787. https://doi.org/10.1038/nsmb.2329

Yates, L. A., Norbury, C. J., & Gilbert, R. J. (2013, Apr 25). The long and short of microRNA. *Cell, 153*(3), 516-519. https://doi.org/10.1016/j.cell.2013.04.003

Yu, C., Wang, L., Rowe, R. G., Han, A., Ji, W., McMahon, C., Baier, A. S., Huang, Y. C., Marion, W., Pearson, D. S., Kruse, A. C., Daley, G. Q., Wu, H., & Sliz, P. (2020, Mar 3). A nanobody targeting the LIN28:let-7 interaction fragment of TUT4 blocks uridylation of let-7. *Proc Natl Acad Sci U S A, 117*(9), 4653-4663. https://doi.org/10.1073/pnas.1919409117

Yu, S., & Kim, V. N. (2020, Sep). A tale of non-canonical tails: gene regulation by post-transcriptional RNA tailing. *Nat Rev Mol Cell Biol, 21*(9), 542-556. https://doi.org/10.1038/s41580-020-0246-8

Zerbs, S., Giuliani, S., & Collart, F. (2014). Small-scale expression of proteins in E. coli. *Methods Enzymol, 536*, 117-131. https://doi.org/10.1016/B978-0-12-420070-8.00011-8

Zhang, Y., Zhang, L. S., Dai, Q., Chen, P., Lu, M., Kairis, E. L., Murugaiah, V., Xu, J., Shukla, R. K., Liang, X., Zou, Z., Cormet-Boyaka, E., Qiu, J., Peeples, M. E., Sharma, A., He, C., & Li, J. (2022, Oct 18). 5-methylcytosine (m(5)C) RNA modification controls the innate immune response to virus infection by regulating type I interferons. *Proc Natl Acad Sci U S A, 119*(42), e2123338119. https://doi.org/10.1073/pnas.2123338119

Zhang, Z., Shigematsu, H., Shimizu, T., & Ohto, U. (2021, Oct 7). Improving particle quality in cryo-EM analysis using a PEGylation method. *Structure, 29*(10), 1192-1199 e1194. https://doi.org/10.1016/j.str.2021.05.004

Zhao, B., Han, H., Chen, J., Zhang, Z., Li, S., Fang, F., Zheng, Q., Ma, Y., Zhang, J., Wu, N., & Yang, Y. (2014, Jan 1). MicroRNA let-7c inhibits migration and invasion of human non-small cell lung cancer by targeting ITGB3 and MAP4K3 [Research Support, Non-U.S. Gov't]. *Cancer Lett, 342*(1), 43-51. https://doi.org/10.1016/j.canlet.2013.08.030

Zhao, M. Z., Lin, D. H., Zuo, H., Wei, H., Wang, X., Gou, L. T., & Liu, M. F. (2022, Nov). piRNA 3' uridylation facilitates the assembly of MIWI/piRNA complex for efficient target regulation in mouse male germ cells. *Cell Res, 32*(11), 1030-1033. https://doi.org/10.1038/s41422-022-00659-1

Zhao, W., & Sakurai, K. (2017, Jun). CCD camera as feasible large-area-size x-ray detector for x-

ray fluorescence spectroscopy and imaging. *Rev Sci Instrum, 88*(6), 063703. https://doi.org/10.1063/1.4985149

Zhou, Z., Lv, J., Yu, H., Han, J., Yang, X., Feng, D., Wu, Q., Yuan, B., Lu, Q., & Yang, H. (2020, Jun 8). Mechanism of RNA modification N6-methyladenosine in human cancer. *Mol Cancer, 19*(1), 104. https://doi.org/10.1186/s12943-020-01216-3