

Electronic Thesis and Dissertation Repository

8-9-2018 10:00 AM

The Functional Characterization of the N-terminal Domains of TUT4

Lauren E. Seidl
The University of Western Ontario

Supervisor
Heinemann, Ilka U.
The University of Western Ontario

Graduate Program in Biochemistry

A thesis submitted in partial fulfillment of the requirements for the degree in Master of Science

© Lauren E. Seidl 2018

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>

Recommended Citation

Seidl, Lauren E., "The Functional Characterization of the N-terminal Domains of TUT4" (2018). *Electronic Thesis and Dissertation Repository*. 5592.

<https://ir.lib.uwo.ca/etd/5592>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.

Abstract

Cells have the ability to adapt in response to environmental stressors by regulating RNA stability. Terminal uridylyltransferases (TUTases) have emerged as essential enzymes in post-transcriptional regulation. TUTases catalyze the untemplated addition of uridine residues to the RNA 3'-end, which generally leads to RNA degradation. Human TUTase 4 (TUT4) regulates mRNA and miRNA stability by initiating the decay of RNA through the addition of a poly(U) tail. TUT4 encodes two catalytic regions. Previously, the C-terminal catalytic motif was thought to execute uridylation activity, while the N-terminal motif was thought to be catalytically inactive. I here demonstrate that while less active than its C-terminal counterpart, the N-terminal motif is indeed capable of post-transcriptional RNA editing and displays RNA substrate specificity. I further identified one of the three catalytic aspartates required for uridylation activity. This reveals a previously unknown catalytic function of the N-terminal catalytic domains with implications for its biological function.

Keywords: TUT4, Uridylyltransferase, Nucleotidyltransferase, Non-canonical poly(A) polymerase, TUT4 N-terminal catalytic region, uridylation, post-transcriptional modification, RNA regulation, RNA decay, miRNA

Co-Authorship Statement

Chapter 1

Sections and figures from this chapter are from a published review article. The published review paper was entitled “*Tipping the balance of RNA stability by 3’ editing of the transcriptome*” in which I share co-authorship with Christina Z. Chung, Mitchell R. Mann, and Ilka U. Heinemann. I contributed equally with my co-authors to this review article, published in *Biochimica et Biophysica Acta (BBA)- General Subjects* in 2017.

Chung, C. Z., Seidl, L. E., Mann, M. R. & Heinemann, I. U. Tipping the balance of RNA stability by 3’ editing of the transcriptome. *Biochim. Biophys. Acta BBA - Gen. Subj.* **1861**, 2971-2979 (2017).

Dedication

For my parents.

Thank you for always believing in me and pushing me to achieve my full potential and for answering all of my phone calls home. I love you both.

For Brandon.

Thank you for all the support during the last 2 years and all the driving you did so we could see each other every weekend. I love you.

Acknowledgements

I would first and foremost like to thank my supervisor, Ilka Heinemann, for taking a chance on me and giving me the opportunity to complete my Master's degree in her lab. Her guidance and constant support throughout my Master's degree has helped me develop into the scientist I am today. She helped immensely with planning my experiments, editing anything I had to write, and always keeping my eye on the prize even when it wasn't always clear. I have gained so much knowledge from Ilka that I will carry on with me into my future endeavors.

Thank you to my advisory committee members, Dr. Chris Brandl and Dr. Murray Junop for all of their insight and guidance. Their advice helped shape my thesis into what it is.

Thank you to the Heinemann lab members, both past and present, especially Christina Chung. She was my mentor in the lab from the very beginning. Thank you for sharing all of your knowledge and expertise with me over the past 2 years.

To Yumin Bi, thank you for all your help during my Master's thesis. She was not only a lab tech, but a miracle worker. I don't know how I would have made it through my thesis without her expertise.

Thank you to Dr. Patrick O'Donoghue and the O'Donoghue lab members. Their insight and suggestions during lab meetings helped me greatly.

Lastly, thank you to all of my family and friends for your constant encouragement and support throughout my Master's degree. It is greatly appreciated.

Table of Contents

Abstract.....	ii
Co-Authorship Statement.....	iii
Dedication	iv
Acknowledgements	v
Table of Contents.....	vi
List of Tables.....	x
List of Figures.....	xi
List of Abbreviations.....	xiii
Chapter 1: Introduction.....	1
1.1 The flow of information through the central dogma.....	1
1.2 Poly(A) polymerases are essential enzymes for RNA regulation.....	4
1.2.1 Molecular basis of nucleotide addition	5
1.2.2 The discovery of RNA uridylation	8
1.2.3 Adenylation vs uridylation in RNA stability	8
1.2.4 PAPs and TUTases belong to the same phylogenetic family.....	12
1.3 Terminal uridylyltransferases	14
1.3.1 Terminal uridylyltransferases TUT4 and TUT7	17
1.4 Terminal uridylyltransferase 4.....	21
1.4.1 Terminal uridylyltransferase 4 is involved in mRNA degradation	23
1.4.2 Terminal uridylyltransferase 4 is a regulator of group II miRNA biogenesis.....	26
1.5 TUT4 is a possible therapeutic target in cancer.....	29
1.6 TUT4 N-terminal domain activity plays a role in the cell cycle	30

1.7	Rationale.....	31
1.8	Hypothesis and Aims	34
Chapter 2: Methods.....		35
2.1	Cloning	35
2.2	<i>Escherichia coli</i> competent cells and transformation	37
2.2.1	Chemically competent cells.....	37
2.2.2	Heat-shock transformation	37
2.3	Site-directed mutagenesis	38
2.4	Purification of TUT4 proteins	40
2.5	SDS polyacrylamide gel electrophoresis	42
2.6	Bradford assay	43
2.7	Western blot	43
2.8	TUT4 enzyme activity assay.....	45
2.8.1	End-point activity assay	45
2.8.2	TUT4 inhibitors.....	45
2.8.3	PUP time course activity assay.....	46
2.8.4	TUT4-N time course activity assay.....	46
Chapter 3: Results		48
3.1	TUT4-N was cloned into an <i>E. coli</i> expression vector	48
3.1.1	The N-terminal domain lacks an identifiable catalytic aspartate triad..	48
3.2	Purification of TUT4-N	51
3.2.1	TUT4-N was partially purified.....	51
3.2.2	Expression of TUT4-N was confirmed by Western blot and mass spectrometry	53
3.2.3	Mock purification	55

3.3 The TUT4 N-terminal domain is catalytically active	57
3.3.1 The Cid1 poly(U) polymerase is catalytically active with several RNA substrates	57
3.3.2 TUT4-N is capable of uridylation activity with select substrates.....	59
3.3.3 DMSO stabilizes TUT4-N and increases product formation.....	61
3.3.4 Mock protein activity.....	63
3.4 Mutation of proposed N-terminal active site residues and protein purification	65
3.4.1 Mutation of possible catalytically relevant active site residues determined by sequence alignment	65
3.4.2 Purification of mutants with the similar purity as WT	67
3.5 TUT4 mutants are catalytically active	70
3.6 Mutating N-terminal active site residues has no significant effect on the activity of TUT4-N.....	74
3.6.1 Cid1 adds nucleotides in a time dependent manner	74
3.6.2 Uridylation activity of TUT4-N and mutants is time dependent and some mutants display altered activity.	76
3.6.3 Alanine mutation of proposed catalytic residues has little effect on catalytic activity	79
3.7 TUT7 also displays N-terminal uridylation activity	82
Chapter 4: Discussion	88
4.1 The TUT4 N-terminal domain is catalytically active	88
4.1.1 <i>E. coli</i> cell extracts do not exhibit uridylyltransferase activity.	93
4.2 TUT7 N-terminus also shows activity similar to TUT4	94
4.3 The N-terminal active site of TUT4 likely exhibits a different active site compared to the C-terminal domains.....	97
4.4 TUT4 N-terminal activity shows RNA specificity	101

4.5 The purpose of two encoded catalytic regions.....	102
4.6 Conclusion and Future Directions.....	104
References.....	107
Curriculum Vitae.....	118

List of Tables

Table 1. Over the past two decades, the different terminal nucleotidyltransferase (TNTase) homologs have been associated with adenylation or uridylation activity on various RNA substrates.

Table 2. Primer sequences for site-directed mutagenesis.

Table 3. Initial velocities and specific activities of TUT4-N and mutants.

Table 4. Initial velocities of TUT7 N- and C-terminal catalytic domains.

Table 5. Specific activities of TUT7-N and TUT7-C proteins.

List of Figures

Figure 1. The central dogma.

Figure 2. Catalytic aspartate triad allows for the transfer of nucleotides onto the RNA 3'-end.

Figure 3. Overview of mRNA modifications by adenylation and uridylation in humans.

Figure 4. Phylogenetic tree of TNTases.

Figure 5. Non-canonical poly(A) polymerases contain conserved catalytic domains across species.

Figure 6. Structure of TUT7 as determined by Faehnle et al., 2017.

Figure 7. TUT4 multi-domain organization.

Figure 8. TUT4 polyuridylates mRNA to initiate decay pathway.

Figure 9. TUT4 involvement in miRNA maturation and degradation pathway.

Figure 10. The existence of 24 different TUT4 splicing variants.

Figure 11. TUT4 N-terminal domain layout and sequence alignment.

Figure 12. Two-step purification process of TUT4-N.

Figure 13. Verification of purified TUT4-N.

Figure 14. Mock purification of empty pGEX-6p2-His vector as a negative control.

Figure 15. Cid1 is catalytically active with a variety of substrates.

Figure 16. TUT4-N exhibits uridylation activity.

Figure 17. DMSO stabilizes TUT4-N and increases its enzymatic activity.

Figure 18. Confirmation of TUT4-N uridylation activity using Mock protein.

Figure 19. Sequence alignment of TUT4 N- and C-termini to identify catalytic residues of the N-terminal Ntr domain.

Figure 20. Purification of TUT4-N mutants.

Figure 21. Verification of TUT4-N mutants by Western blot.

Figure 22. TUT4-N mutants display RNA-specific uridylation activity.

Figure 23. Quantification of uridylated product with Cid1.

Figure 24. Product formation of uridylated 15A RNA by TUT4-N proteins.

Figure 25. Visual representation of difference of initial velocities of TUT4-N and mutants.

Figure 26. SDS-PAGE of concentrated TUT7-N and TUT7-C proteins.

Figure 27. Time course activity of TUT7-N and TUT7-C.

Figure 28. Product formation of uridylated 15A RNA by TUT7.

Figure 29. Specific activities of TUT4-N, TUT7-N, and TUT7-C.

Figure 30. Specific activities of TUT4-N and mutants.

List of Abbreviations

AGO: Argonaute
Ala (A): alanine
AMP: ampicillin
Asn (N): asparagine
Asp (D): aspartate or aspartic acid
Arg (R): arginine
ATP: adenosine triphosphate
BL21 C+: BL21 codon plus
cDNA: complementary DNA
Cid1: Caffeine induced death protein 1
CM: chloramphenicol
CPSF: cleavage and polyadenylation specificity factor
dsRNA: double stranded RNA
DNA: deoxyribonucleic acid
DTT: Dithiothreitol
eIF: eukaryotic initiation factor
ECL: enhanced chemiluminescence
E. coli: *Escherichia coli*
EDTA: ethylenediaminetetraacetic acid
eRF3: eukaryotic release factor 3
FPLC: fast protein liquid chromatography
Gld2: Germline development 2
gRNA: guide RNA
GST: glutathione S-transferase
GV: germline vesicle
HCl: hydrochloric acid
His (H): histidine
HRP: horseradish peroxidase

H. sapiens: *Homo sapiens*
Ile (I): Isoleucine
IPTG: isopropyl- β -D-1-thiogalactopyranoside
LB: Luria-Bertani
L. tarentolae: *Leishmania tarentolae*
Leu (L): Leucine
Lsm1-7: like-Sm proteins
Met (M): methionine
min.: minute
miRNA: microRNA
Mg²⁺: magnesium
MgCl₂: magnesium chloride
mRNA: messenger RNA
ncPAP: non-canonical poly(A) polymerase
NaCl: sodium chloride
Ni-NTA: nickel-nitrilotriacetic acid
nt: nucleotide
nts: nucleotides
NTP: nucleoside triphosphate
Ntr: nucleotidyltransferase
OD600: optical density at 600nm
PABP: poly(A)-binding protein
PAP: poly(A) polymerase
PAPD1: PAP-associated domain-containing protein 1
PAS: polyadenylation signal
PBS: phosphate buffer saline
PCR: polymerase chain reaction
PneumoG: pneumoviridae
Pol: RNA polymerase
Pol β : DNA polymerase β -like family
Pre-miRNA: precursor microRNA

Pri-miRNA: primary microRNA
PUP: poly(U) polymerase
PVDF: polyvinylidene difluoride
RBD: RNA binding domain
RISC: RNA-induced silencing complex
RNA: ribonucleic acid
SDS: sodium dodecyl sulfate
SDS-PAGE: sodium dodecyl sulfate polyacrylamide gel electrophoresis
sec.: second
siRNA: silencing RNA
snRNA: small nuclear RNA
S. pombe: *Schizosaccharomyces pombe*
Thg1: tRNA^{His} Guanylyltransferase
TNTase: terminal nucleotidyltransferase
Tris: tris(hydroxymethyl)aminomethane
TUTase: terminal uridylyltransferase
TUT1: terminal uridylyltransferase 1
TUT4: terminal uridylyltransferase 4
TUT7: terminal uridylyltransferase 7
UTP: uridine triphosphate
UTR: untranslated region
WT: wild-type
Zcchc11: zinc finger CCHC-type containing 11
ZnF: zinc finger

Chapter 1: Introduction

1.1 The flow of information through the central dogma

In 1958, Francis Crick explained a two-step process of the movement of genetic information from deoxyribonucleic acids (DNA) to essential protein products, which is termed the central dogma of molecular biology (Figure 1)¹. DNA, a double-stranded helix, stores our genetic information in specific nucleotide sequences, dictating which proteins will ultimately be expressed. The first step in the central dogma is the transcription of DNA. Ribonucleic acid (RNA) polymerase (Pol), an essential protein in this process, begins transcription by binding to the DNA sequence. Pol II, specifically, is responsible for replication of the genes involved in protein expression, forming messenger RNA (mRNA)². The binding of the Pol II complex results in the separation of the two nucleotide strands, which is termed the 'transcription bubble'². Nucleoside triphosphates (NTP) are required for this process, as complementary nucleotides are added to the DNA template to form the RNA chain². Transcription terminates when the mRNA sequence is complete.

During the second step of the central dogma, the newly synthesized mRNA is exported out of the nucleus and is used as a template for the expression of proteins in the process of translation. In eukaryotes, the 80S ribosome, a ribosome capable of elongation, binds to the mRNA 5'-untranslated region (UTR) to initiate translation^{3, 4, 5}. The 80S ribosome complex is formed in two steps. The methionyl transfer RNA (Met-tRNA) is bound to the 40S subunit of

the ribosome in a complex with eukaryotic initiation factors (eIF) before being recruited to the start codon of the mRNA^{3, 4, 5, 6, 7}. The 60S ribosomal subunit then joins the mRNA-bound complex forming the final 80S initiation ribosome^{3, 4, 5}. The ribosome moves down the mRNA in the 5'-3' direction adding the correct amino acid to the codon sequence.

Although the originally proposed central dogma seemed like a straight forward process, over the years many different enzymes have been discovered that are essential in regulating this process. Disruption at any level can result in error and mutations during either transcription or translation. As a means of survival, the cell has the ability to adapt to environmental stressors⁸. At the RNA post-transcriptional level, many different mechanisms of regulation are involved in making sure that RNA is tightly controlled. The addition of untemplated nucleotides to the 3'-end of RNA is an important post-transcriptional modification directly controlling RNA stability and therefore the amount of protein expressed⁸. The addition of nucleotides is catalyzed by enzymes termed nucleotidyltransferases. The most well-studied nucleotidyltransferases are poly(A) polymerases (PAPs), which utilize adenosine triphosphate (ATP) to modify RNA.

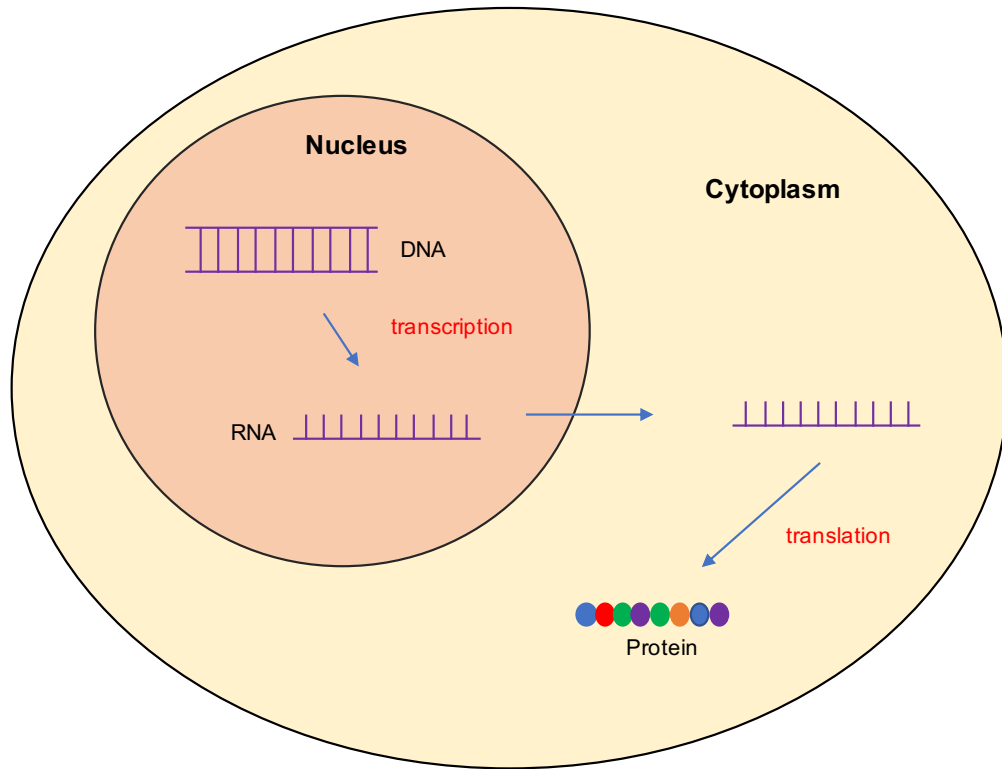


Figure 1. The central dogma. The flow of information from DNA to protein as proposed by Francis Crick in 1958. Double-stranded DNA undergoes transcription, the copying of genetic data, into single-stranded mRNA. mRNA is exported out of the nucleus before translation begins. This process translates the mRNA genetic information into protein products.

1.2 Poly(A) polymerases are essential enzymes for RNA regulation

Since the discovery of poly(A) tails on mRNAs in 1970^{9, 10}, PAPs have been well studied for their roles in the maturation process and stability of mRNA^{11, 12, 13}. PAPs, a member of the DNA polymerase β -like (pol β) family, modify mRNA by the untemplated addition of adenosine residues to the 3'-end^{12, 13}. In eukaryotes, canonical PAPs begin the poly(A) addition initially in the nucleus, following pre-mRNA cleavage, as part of the mRNA maturation process, ceasing transcription and stimulating mRNA export into the cytoplasm^{8, 13, 14, 15, 16}. The length of the poly(A) tail of mRNAs is an essential regulatory mode, where varying lengths of the poly(A) tail determines life span and stability of the mRNA¹⁷. Long poly(A) tails are important in mRNA stabilization and increase the mRNA life span¹⁷. Polyadenylation of nuclear mRNA occurs when the cleavage and polyadenylation specificity factor (CPSF) is recruited to a conserved AAUAAA polyadenylation signal (PAS) on the mRNA^{17, 18, 19, 20, 21}; This in turn allows the nuclear poly(A) binding protein (PABPN1) to bind to the mRNA, which promotes the addition of a long 3'-poly(A) tail by PAP^{21, 22}. In mammalian cells, for example, the stabilizing poly(A) tail length is approximately 250 nucleotides (nts) in length, while in yeast, the poly(A) tail can be up to 80 nts in length^{11, 17, 21, 22}. Once the 3'-end tail is shortened to less than 25 nts, due to each subsequent RNA translation event, the proteins involved in mRNA stabilization and mRNA translation can no longer bind^{17, 22, 23, 24}. The shortening of the poly(A) tail results in mRNA degradation.

Non-canonical PAPs (ncPAPs), which function in both the nucleus and cytoplasm, have been more recently identified^{25, 26, 27}. In the cytoplasm, polyadenylation elongates the pre-existing mRNA poly(A) tail, allowing the PABP to bind, forming a closed-loop and starting the process of translation^{16, 28}. Cytoplasmic mRNA poly(A) tail length is regulated carefully through both the deadenylation and polyadenylation processes to either shut down or increase translation, respectively¹⁶.

1.2.1 Molecular basis of nucleotide addition

On a molecular basis, both canonical and non-canonical PAPs catalyze nucleotide addition via a conserved carboxylate triad, composed of three aspartate residues²⁶. In canonical PAPs, this triad is located in the N-terminal catalytic domain¹³. In ncPAPs, the catalytic triad is located within a nucleotidyltransferase (Ntr) domain^{8, 29}. This aspartate triad functions to coordinate two essential metal ions in the active site¹³. The first two aspartates coordinate a single magnesium (Mg^{2+}) ion, required for stability of the NTP through interactions with the β - and γ -phosphates of the triphosphate moiety^{30, 31, 32}. The phosphates of the incoming nucleotide are also stabilized by interactions with surrounding water molecules^{30, 32}. The third aspartate coordinates the second Mg^{2+} ion for the transfer of the nucleotide onto the RNA substrate by weakening the attraction of the RNA 3'-hydroxyl to the hydrogen³¹. 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, allowing for nucleophilic attack of the α -phosphate¹², and the subsequent formation of the phosphodiester bond,

releasing diphosphate. Figure 2 shows the addition of uridine triphosphate (UTP) onto an RNA 3'-end.

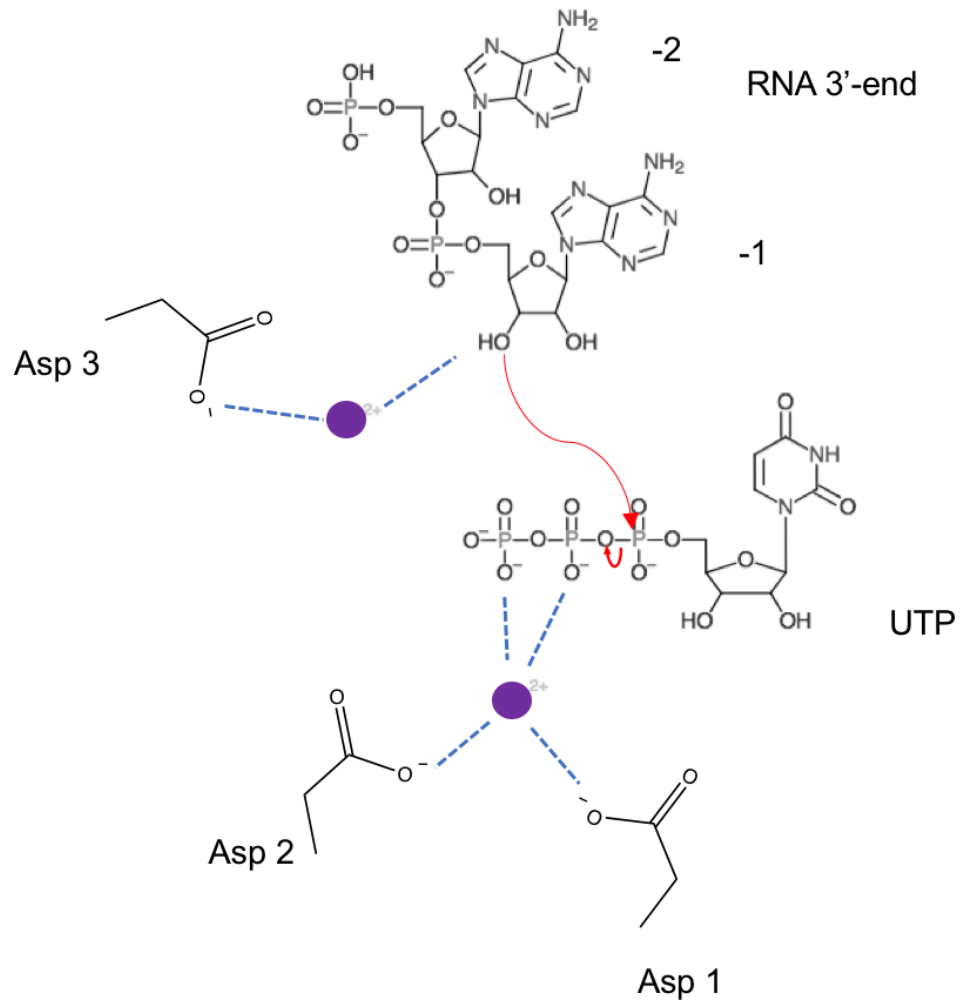


Figure 2. Catalytic aspartate triad allows for the transfer of nucleotides onto the RNA 3'-end. Aspartate 1 and 2 (Asp 1 and Asp 2) coordinate a Mg^{2+} ion (purple) to stabilize the β - and γ -phosphate of the uridine triphosphate (UTP) moiety. Aspartate 3 (Asp 3) lowers the RNA 3'-hydroxyl group affinity for hydrogen, allowing for nucleophilic attack on the UTP α -phosphate. Water molecule interactions not shown.

1.2.2 The discovery of RNA uridylation

The addition of U-residues on RNA has been known since the 1950s, but with little implication to how important this RNA modification is³³. It wasn't until shortly after poly(A) tails were discovered^{9, 10} that interest in poly(U) tails increased, and the existence of poly(U) tails on both nuclear and cytoplasmic RNA were detected^{8, 34}. The addition of uridine residues to RNA was found to be important in mitochondrial RNA editing of kinetoplasts³⁵, guide RNA (gRNA) processing³⁶, and the generation of mRNA. It wasn't until the last 15 years that uridylation on mRNA was discovered as a post-cleavage modification during mRNA maturation³⁷.

In recent years, ncPAPs which utilize nucleotides other than ATP to modify RNA were identified. These ncPAPs, termed terminal uridylyltransferases (TUTases) or poly(U) polymerases (PUPs), add UTP rather than ATP to the RNA 3'-end as a means of regulating RNA stability. Many terminal nucleotidyltransferase (TNTase) homologs have been biochemically characterized and are categorized into PAPs or TUTases (Table 1), yet the nucleotide preference of some homologs remain uncharacterized⁸.

1.2.3 Adenylation vs uridylation in RNA stability

Adenylation has long been known as a mechanism to stabilize mRNA. Monoadenylation by the ncPAP germline-development 2 (Gld2) has been recently found to stabilize microRNA (miRNA)³⁸. In contrast, recent studies found that polyuridylation marks RNAs for degradation. RNA degradation following polyuridylation has been found in mRNAs, histone mRNAs, and precursor

miRNAs (pre-miRNAs)^{27, 39, 40, 41}. Interestingly, monouridylation of pre-miRNAs is important in miRNA biogenesis⁴². The modification of a poly(A) or poly(U) tail dictates the fate of the RNA (Figure 3). UTP specificity of TUTases is due to the presence of a specific histidine residue, important for nucleotide recognition. In 2012, Bradley Lunde and colleagues discovered that histidine residue 336 (H336) in the yeast ncPAP caffeine induced death protein 1 (Cid1) allowed for preferential selectivity of UTP over ATP, which is conserved only in members of the Cid1 family that have confirmed activity with UTP³⁰. When H336 was mutated to an asparagine residue, Cid1 preference for ATP increased, with a slight decrease in UTP activity, confirming the importance of the histidine residue in UTP selectivity³⁰. More recently, Christina Z. Chung in our lab has shown that in the ncPAP Gld2, which functions with ATP, nucleotide specificity can be changed with the insertion of a histidine residue, in a position relative to H336 in Cid1⁴³. This histidine insertion in the Gld2 sequence displayed activity with UTP over ATP, with a significant decrease in ATP activity⁴³. Sequence alignment of several nucleotidyltransferases showed that this histidine residue is conserved in uridylyltransferases, while adenylyltransferases have a different amino acid in the homologous position or lack this amino acid altogether^{30, 43}.

Table 1. Over the past two decades, the different terminal nucleotidyltransferase (TNTase) homologs have been associated with adenylation or uridylation activity on various RNA substrates. Adapted from Chung, Seidl et al. 2017⁸.

	Adenylation	Uridylation
<i>H. sapiens</i> Tut1	mRNA <ul style="list-style-type: none"> • NQ01, HO1^{44, 45, 46, 47, 48, 49} 	miRNA <ul style="list-style-type: none"> • Multiple⁵⁰ • miR-24, miR-29a⁵¹ U6 snRNA ^{52, 53, 54}
<i>H. sapiens</i> Tut2/ Papd4/Gld2	mRNA <ul style="list-style-type: none"> • Multiple^{55, 56, 57, 58} • p53⁵⁹ miRNA <ul style="list-style-type: none"> • miR-122^{38, 60, 61} • miR-134³⁸ • let-7a^{38, 62} • pre-let-7a⁶² 	miRNA <ul style="list-style-type: none"> • pre-let-7a⁴³ • miR-122³⁸ • Group II pre-miRNA^{42, 63} • pre-miRNA with 5'-overhang⁶⁴
<i>H. sapiens</i> Tut4/ Zcchc11		mRNA ⁶⁵ miRNA ^{66, 67, 68} <ul style="list-style-type: none"> • multiple⁶⁸ • Group II pre-miRNA^{42, 64} • pre-miRNA with 5'-overhang⁶⁴ histone mRNA ⁶⁹
<i>H. sapiens</i> Tut7/Zcchc6		mRNA ⁶⁵ miRNA <ul style="list-style-type: none"> • multiple⁶⁸ • Group II pre-miRNA⁴² • pre-miRNA with 5'-overhang⁶⁴ histone mRNA ⁷⁰
<i>S. pombe</i> Cid1	multiple mRNA ²⁵	multiple mRNA ^{16, 39}

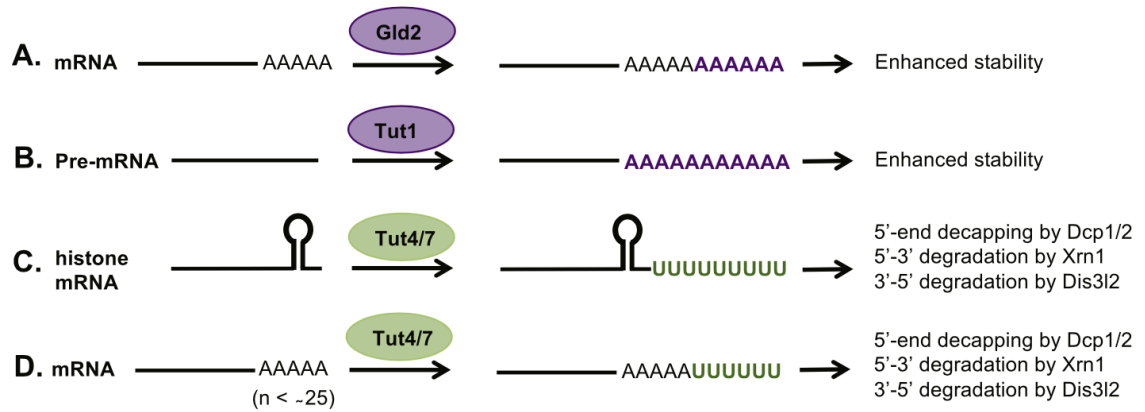


Figure 3. Overview of mRNA modifications by adenylation and uridylation in humans. Polyadenylation of **A)** mRNAs by Gld2 and **B)** pre-mRNA by TUT1 extends their half-life. **C)** TUT4 or TUT7 uridylation of histone mRNAs leads to 5'-3' and 3'-5' degradation. **D)** TUT4 or TUT7 uridylation of mRNAs with short poly(A) tails leads to 5'-3' and 3'-5' degradation. Adapted from Chung, Seidl et al. 2017⁸.

1.2.4 PAPs and TUTases belong to the same phylogenetic family

As outlined above, uridylation and adenylation have opposing effect on RNA stability. Surprisingly, a phylogenetic analysis of over 300 members of the pol β superfamily showed that the enzymes are closely related, and uridylyltransferase activity arose from adenylyltransferases multiple times during evolution (Figure 4). Both TUTases and PAPs catalyze nucleotide addition using the same two metal-ion coordinating mechanism. The two catalytically similar reactions of adenylation and uridylation have opposing biological effects on RNA stability: adenylation stabilizes mature miRNA or mRNA while uridylation destabilizes or silences. We found it striking that TUTase activity appears to emerge from a PAP ancestor at least 4 times independently in the evolution of this TNTase family (Figure 4). We showed that a single amino acid change is sufficient to convert an ATP adding enzyme into a UTP adding enzyme⁴³. The crystal structure of U-specific Cid1 revealed the molecular basis of nucleotide discrimination: a single inserted histidine residue in TUTases sterically hinders ATP from entering the active site³⁰. It is possible that these enzymes diverged from each other due to the need to regulate the same RNA substrates in an opposing manner for different biological effects. We speculate that the gene duplication events may have been related to the cells' need to regulate the same RNA substrates in a diametrically opposing manner. This enabled evolution of an "off" switch, where uridylation signals decay, from an "on" switch, where adenylation increases RNA stability (Figure 4).

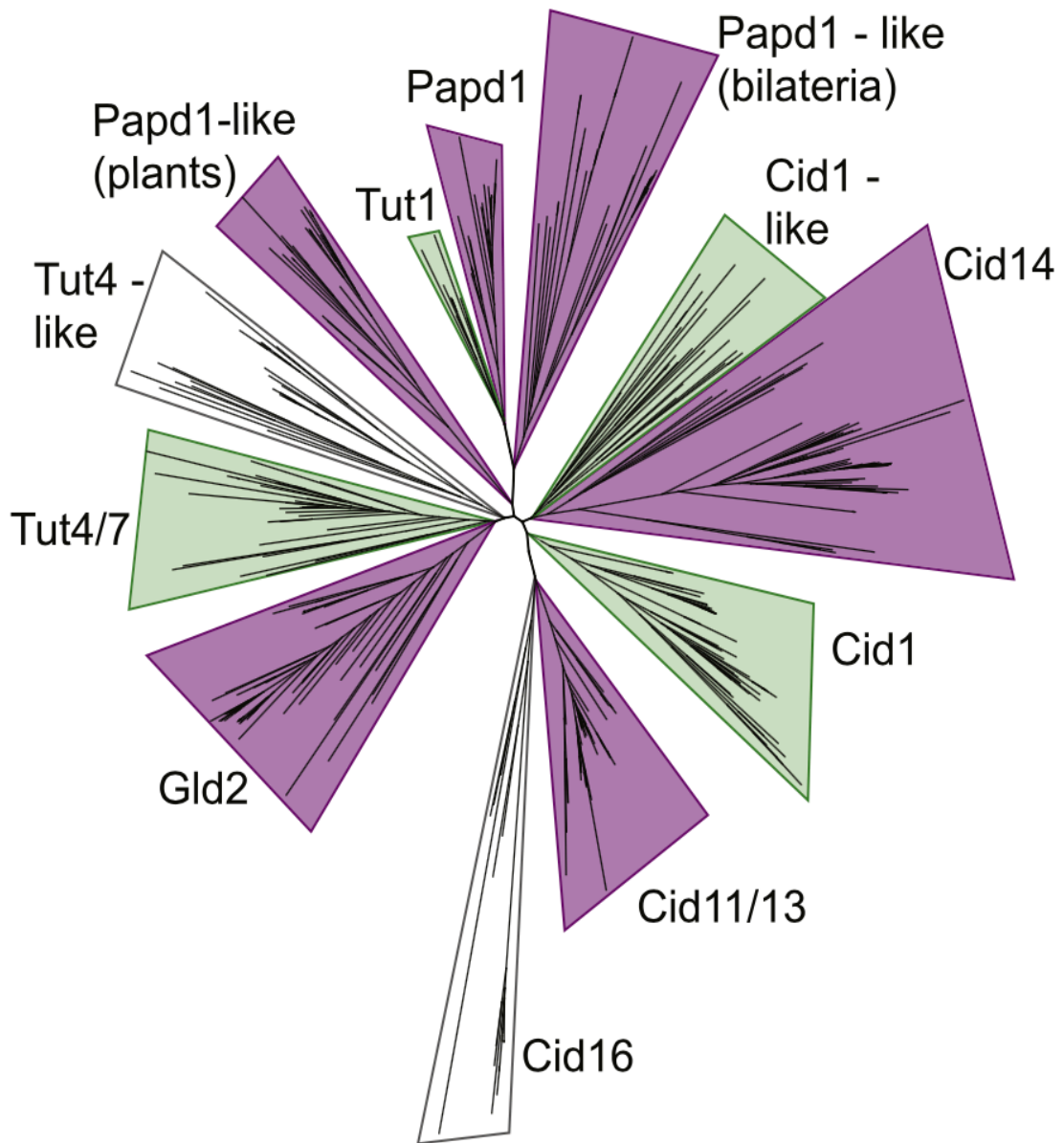


Figure 4. Phylogenetic tree of TNTases. This figure was adapted from Chung, Seidl et al. 2017⁸. TUTase activity emerged from PAP activity in at least three separate incidences. TUTases are marked in green, and PAP homologs in purple.

1.3 Terminal uridylyltransferases

Since their discovery, TUTases have gained interest in the scientific world due to their ability to add uridine residues to the RNA 3'-end. Originally discovered as a PAP in *Schizosaccharomyces pombe* (*S. pombe*), Cid1 was later identified through biochemical characterization as the first TUTase, with the ability to utilize UTP to modify RNA in yeast¹⁶. This novel discovery triggered an investigation of human homologs, which were previously thought to function as adenylyltransferases *in vivo*. Indeed, several human homologs are now known to add uridine instead of adenine residues to modify different types of RNA, including miRNA, mRNA, and small nuclear RNA (snRNA)^{24, 52, 66, 71}.

While all TUTases modify their own specific substrates, the members of the TUTase family have structural similarities. Cid1 consists of the catalytic core of a nucleotidyltransferase (Ntr) domain and poly(A) polymerase (PAP) domain, much like the human adenylyltransferase counterpart Gld2^{16, 43}. This catalytic core is also found in terminal uridylyltransferase 4 (TUT4) and terminal uridylyltransferase 7 (TUT7), which are part of the pol β superfamily (Figure 5). While Cid1, TUT4, and TUT7 have been shown to act as TUTases *in vivo*^{16, 39, 66, 67, 68, 69, 72, 73, 74}, Gld2 is a bona fide PAP^{38, 43, 59, 61, 75, 76} (Figure 5).

The ncPAP Gld2 is capable of adenylation of mRNA and miRNA, as well as uridylation of miRNA^{38, 43, 57}. In general, mRNA and miRNA adenylation by Gld2 is important for stabilization³⁸, while monouridylation of pre-miRNA by Gld2 has been implicated as part of the miRNA maturation process⁴². The yeast homolog Cid1, shortly after its discovery as a TUTase in *S. pombe*, was found to

be crucial in the degradation of mRNA through polyuridylation^{39, 77}. The poly(U) tail addition catalyzed by Cid1 is important in a 5'-3' degradation pathway in yeast³⁹. As the first discovered TUTase, this pathway in yeast was believed to be conserved in other eukaryotes^{39, 77}.

TUT1, a nuclear TUTase, is important in the stabilization of the U6 snRNA⁷⁸. 3'-end polyuridylation of U6 snRNA plays a role in the 3'-end maturation (preventing degradation), acts as a binding site for important protein complexes^{79, 80}, and allows for the U6 snRNA to be degraded after its activity in pre-mRNA splicing⁷¹.

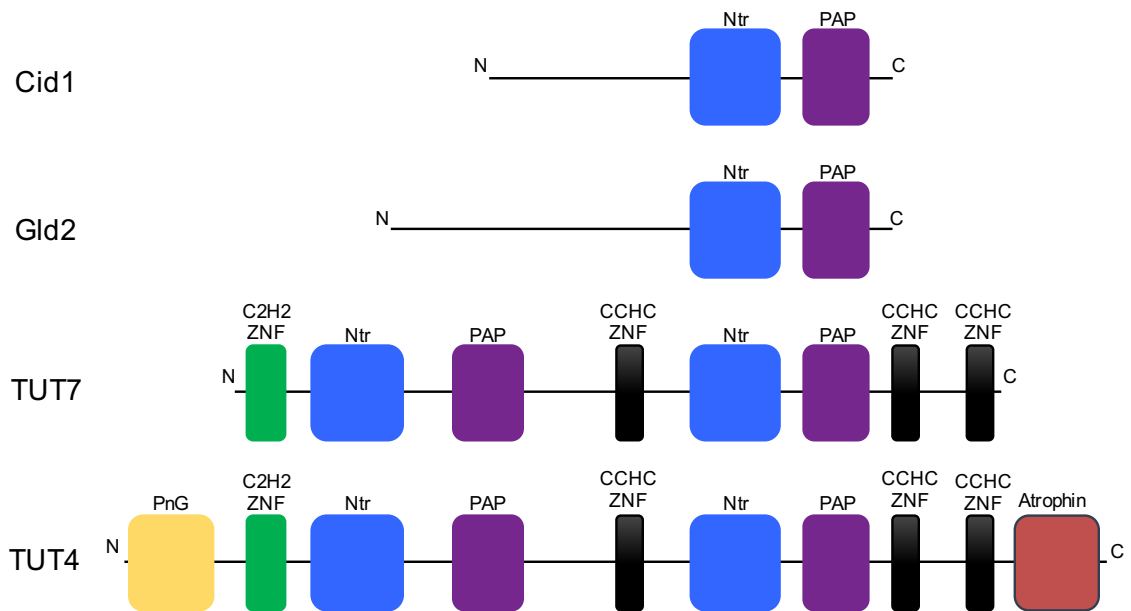


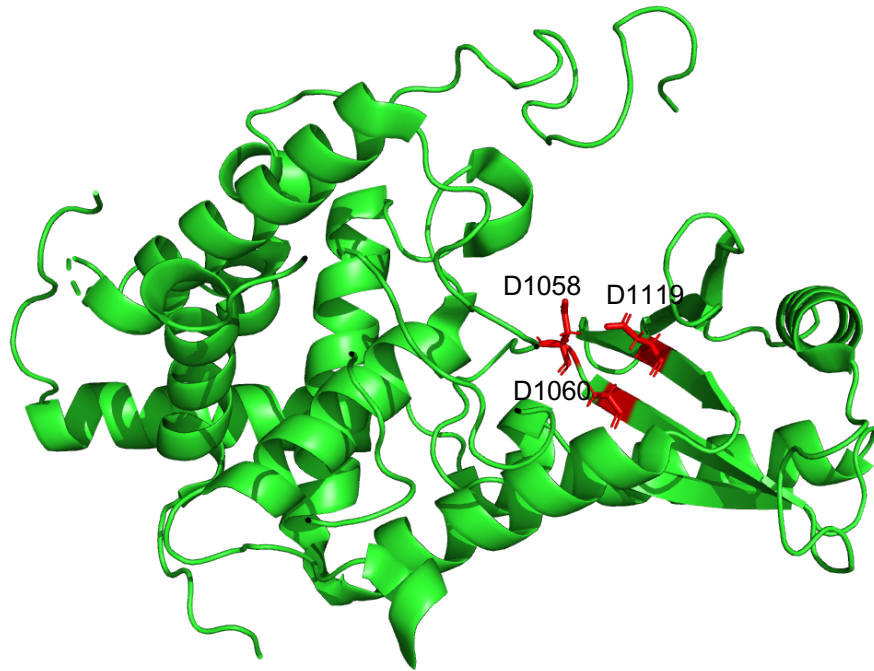
Figure 5. Non-canonical poly(A) polymerases contain conserved catalytic domains across species. Poly(A) polymerase Gld2 (*H. sapiens*) and terminal uridylyltransferases Cid1 (*S. pombe*), TUT7, and TUT4 (*H. sapiens*) all encode the catalytic nucleotidyltransferase (Ntr) and poly(A) polymerase (PAP) domains, conserved across all members of the DNA polymerase β -like family. The Ntr domain is required for nucleotide transfer while the PAP domain is important for nucleotide specificity. Adapted from Chung, Seidl et al. 2017⁸.

1.3.1 Terminal uridylyltransferases TUT4 and TUT7

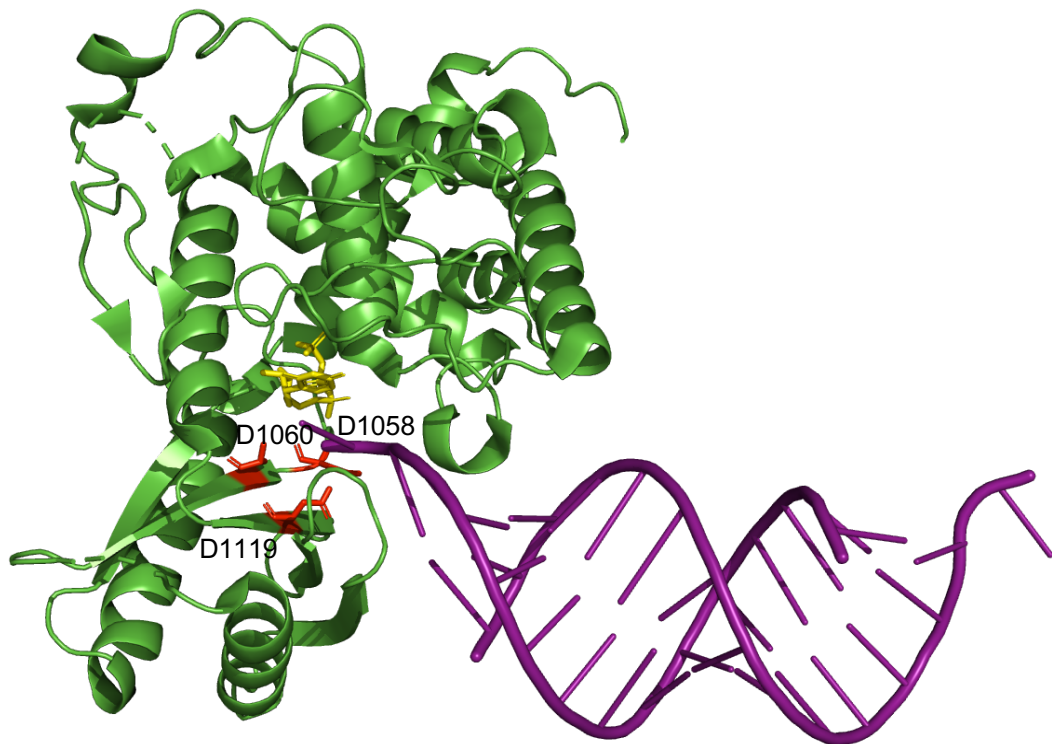
TUT4 and TUT7, two human homologs, are quite functionally and structurally redundant. Functional studies have shown that both TUT4 and TUT7 are involved in the uridylation of mRNA and miRNA for degradation, as well as for miRNA biogenesis. Both TUTases are capable of polyuridyating the 3'-end of mRNA to signal for degradation^{41, 65, 70}. Polyuridylation of mRNA has recently been shown to be crucial in embryonic development by controlling the degradation of RNA transcripts and therefore controlling the turnover of the maternal transcriptome²⁴. During group II miRNA biogenesis, both enzymes monouridylate the pre-miRNA 3'-end allowing for further maturation^{32, 42, 68}. Alternately, in a Lin28A-dependent pathway, TUT4 and TUT7 recognize Lin28A-bound pre-miRNA and polyuridylate the RNA 3'-end, leading to degradation by the exonuclease Dis3L2^{40, 68, 72}. Structurally, TUT4 and TUT7 have a similar domain layout, sharing 40% amino acid identity. Recently, Faehnle et al. have crystallized the C-terminus of TUT7 (Figure 6A), which they termed the catalytic module³². TUT7 catalytic module 2, consisting of amino acid residues 983 to 1365 with a D1060A mutation in the active site and lacking the first CCHC zinc finger, was crystallized with double stranded RNA (dsRNA) with a 1-nt 3'-end overhang, mimicking the TUT7 substrate pre-let-7 (Figure 6B)³². This crystal structure captured TUT7 in its monouridylation activity state, important in miRNA biogenesis. To observe TUT7 as it would be during poly(U) addition, TUT7 catalytic module 1, which includes the first CCHC zinc finger, was co-crystallized with non-hydrolyzable UTP (UMPNPP) and an RNA substrate containing a 2-nt poly(U) 3'-overhang (Figure 6C)³².

From the crystal structures, Faehnle et al. observed that the C-terminal catalytic region, containing the Ntr and PAP domains, formed two lobes, an N-lobe and C-lobe³². When in an unbound state, the N-lobe is in an open conformation, but when a mono(U) or poly(U) RNA substrate is bound, the N-lobe slightly shifts into a closed conformation, and due to high similarity and redundancy, Faehnle et al. conclude that TUT4 undergoes the same conformational changes³².

A)



B)



C)

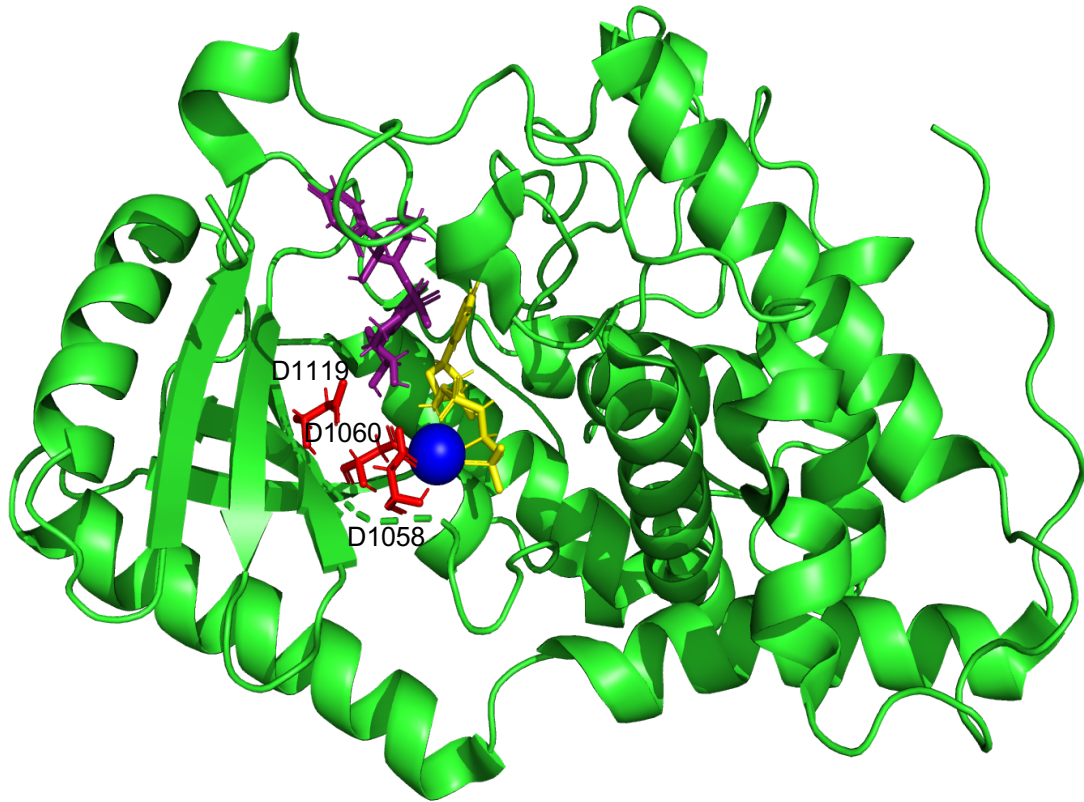


Figure 6. Structure of TUT7 as determined by Faehnle et al., 2017³². **A)** Structure of TUT7 catalytic module, amino acids 963 to 1365. Active site aspartate residues shown in red. PDB 5W0B. **B)** TUT7 catalytic module 2 (amino acids 983 to 1365) with D1060A mutation co-crystallized with dsRNA (purple) to depict structure during monouridylation activity. The active site aspartate residues are shown in red, with UTP position in yellow. PDB 5W0O. **C)** TUT7 catalytic module with D1060A mutation co-crystallized with poly(U) RNA mimic (purple) and non-hydrolyzable UMPNPP (yellow). Active site residues (red) D1058 and A1060 are shown coordinating a single magnesium ion (blue). PDB 5W0N.

1.4 Terminal uridylyltransferase 4

TUT4, also known as zinc finger CCHC-type containing 11 (Zcchc11), was first described as a TUTase in 2009⁶⁷, and is the largest human TUTase containing several important domains (Figure 7). In the C-terminus, three CCHC zinc fingers (ZnF) are thought to promote RNA binding and flank the active site^{73, 81}. The active site is composed of two conserved domains, the Ntr domain, which houses the catalytic aspartate triad, and PAP domain, which plays a role in nucleotide specificity^{73, 81}. Upstream in the N-terminus, a second catalytic region is encoded but is believed to be inactive as it lacks an identifiable aspartate triad required for activity. Also in the N-terminus is the C2H2 ZnF, known to interact with the RNA binding protein Lin28A^{73, 81}. In 2012, Thornton and colleagues showed that both the deletion of the C2H2 ZnF and a cysteine to alanine mutation in the C2H2 ZnF led to basal TUT4 activity, with no increase in uridylation activity in the presence of Lin28A⁷³. Encoded on the outer ends of the protein are pneumoviridae (PneumoG) and Atrophin-like domains on the N- and C-terminus, respectively, both with no known function⁷³. Since their discovery, a number of RNA substrates have been suggested, with miRNA and mRNA being the most common.

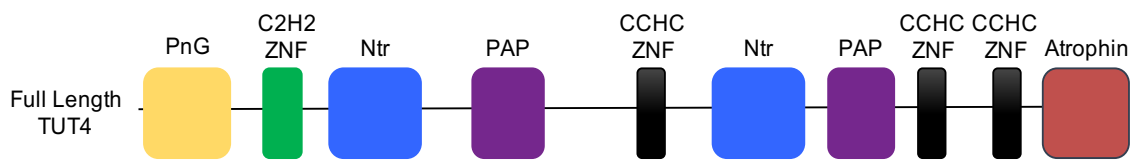


Figure 7. TUT4 multi-domain organization. TUT4, 1644 amino acids in length, contains 10 distinct domains. Three CCHC zinc fingers (ZnF) (black) in the C-terminus have been suggested to be involved in RNA binding. A C2H2 ZnF (green) in the N-terminus is involved in Lin28 binding. In the C-terminus, the nucleotidyltransferase (Ntr) (blue) and poly(A) polymerase (PAP) (purple) domains form the catalytic region. In the N-terminus, a second catalytic region containing the Ntr (blue) and PAP (purple) is encoded but is lacking an identifiable aspartate triad. On the extreme N- and C-terminus are the PneumoG (PnG) (yellow) and Atrophin-like (Atrophin) domains (red), both with no known role.

1.4.1 Terminal uridylyltransferase 4 is involved in mRNA degradation

Replication-dependent histone mRNA in eukaryotes terminates in a 3' stem loop, and is the only mRNA in eukaryotes that is known to not have a poly(A) tail^{69, 70}. Due to the lack of a poly(A) tail on histone mRNAs, the first step of mRNA degradation, deadenylation, cannot be initiated. Instead, degradation is initiated by the 3'-end addition of uridine residues to degrade the mRNA in both a 5'-3' and 3'-5' direction (Figure 8A)^{41, 69, 70}. Although other TUTases have been proposed to participate in this process, TUT4 has been shown to be involved in histone mRNA degradation⁶⁹. Schmidt et al. have shown *in vivo* association between TUT4 and histone H3 mRNA using immunoprecipitation studies and a reduction in uridylation of histone mRNA when TUT4 is knocked down⁶⁹. Polyuridylation, and as a result degradation, of histone mRNA can occur at both the end of S phase of the cell cycle, first discovered by Mullen and Marzluff, and as consequence to DNA replication inhibition^{41, 65}.

Uridylation by TUT4 also occurs on mRNAs with a poly(A) tail on the 3'-end (Figure 8B). Lim et al. discovered that uridylation occurs more frequently on poly(A) tail mRNAs shorter than 25 nts in length⁶⁵. In *in vivo* siRNA knockdown experiments of TUT4, polyuridylation was decreased by 3.7-fold, especially on mRNA with poly(A) tails consisting of 5-25 adenine residues⁶⁵. Polyuridylation of short poly(A) tails occurs on mRNA to accelerate the decay process^{24, 65, 82}. PABP regulates mRNA stability, including degradation, by interacting with eRF3, a termination factor²². Every translation event of an mRNA results in the shortening of the 3'-poly(A) tail, leading to degradation²². When the poly(A) tail becomes too short (less than 25 nts), PABP can no longer stably bind to the

poly(A) tail²⁴. This leads to the recruitment of TUT4 and polyuridylation of the mRNA 3'-end²⁴. Following polyuridylation, the Lsm1-7 (like-Sm proteins) complex, which has an affinity to bind to uridyl residues, will bind to the mRNA 3'-end^{24, 65} (Figure 8B). Binding of this protein complex to the poly(U) tail is the beginning step in the 5'-3' decay pathway^{24, 65}. Lsm1-7 binding is followed by decapping of the 5'-end, allowing Xrn1, a 5'-exonuclease, to digest the mRNA^{8, 24, 65} (Figure 8B).

Recently, Morgan et al. discovered the importance of mRNA uridylation in oocyte formation and maturation²⁴. Interestingly, TAIL-seq experiments of germline vesicle (GV) oocytes and somatic mouse tissues revealed that GV oocytes have the highest polyuridylation to monouridylation ratio²⁴. Studying the role of TUT4 (and TUT7) in oocyte maturation, Morgan and colleagues mutated mouse alleles, creating several different allele combinations, including a TUT4/7 knockout (*TUT4/7^{ckO}*) allele²⁴. They observed a lack of early embryonic development in the *TUT4/7^{ckO}* oocytes²⁴. Investigating into the role of TUT4 further, expression of catalytically inactive TUT4 protein displayed the same phenotype as the knockout mice²⁴.

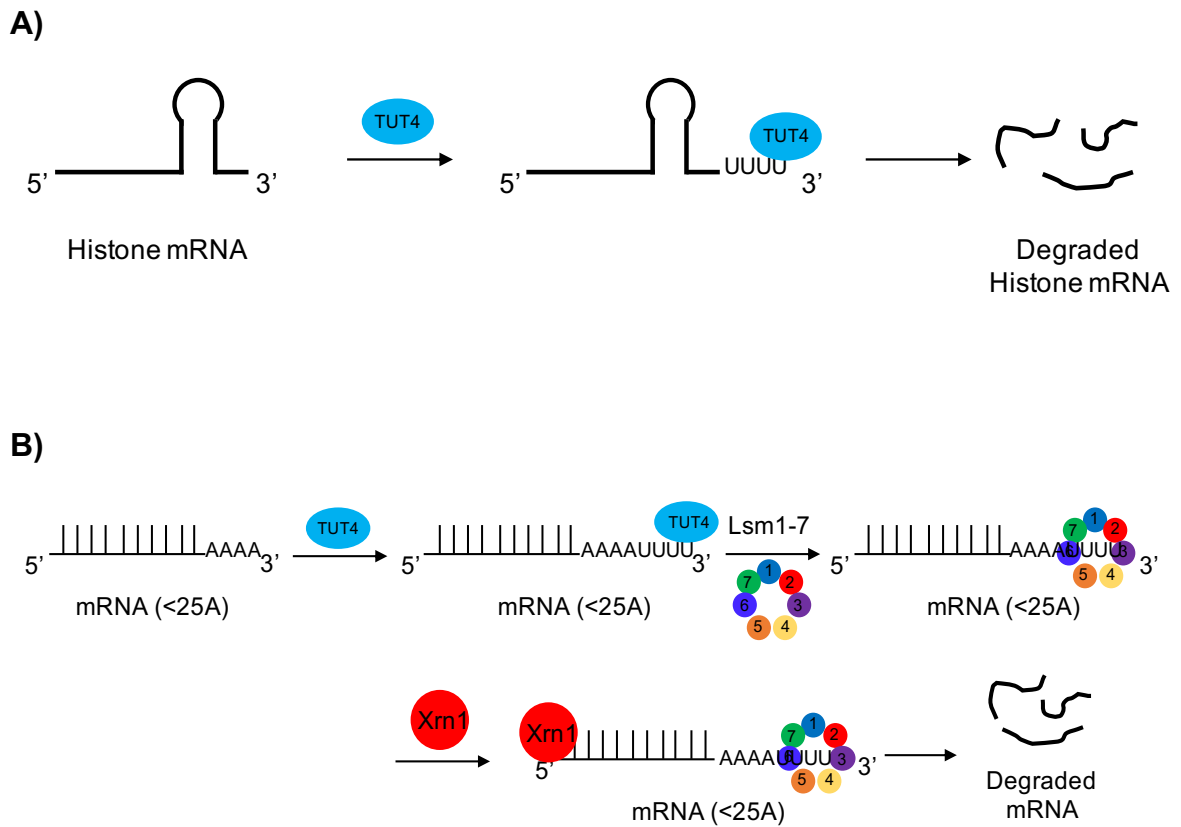


Figure 8. TUT4 polyuridylates mRNA to initiate decay pathway. A)

Polyuridylation of histone mRNA initiates a decay pathway. TUT4 uridyates the 3'-end, after the stem loop. Degradation can occur in either the 5'-3' or 3'-5' direction. **B)** Polyuridylation of mRNA with short poly(A) tails leads to degradation. TUT4 polyuridylates the 3'-end of mRNA, adding to the poly(A) tail.

Lsm1-7 complex preferentially binds to the poly(U) tail, recruiting the exonuclease Xrn1, initiating the 5'-3' decay pathway.

1.4.2 Terminal uridylyltransferase 4 is a regulator of group II miRNA biogenesis

A well-studied function of TUT4 is its involvement in the biogenesis of group II miRNA (Figure 9). Primary miRNA (pri-miRNA) is processed in the nucleus into pre-miRNA by RNase III Drosha and co-factor Pasha (DGCR8), a holoenzyme, resulting in a hairpin double-stranded RNA structure, before being exported into the cytoplasm by exportin 5^{27, 42, 64, 83, 84, 85}. Two different groups of pre-miRNAs can be exported out of the nucleus. In group I pre-miRNA, the pre-miRNA exported has a 2-nt 3'-overhang, recognized by Dicer for further processing^{86, 87}. Dicer binds the 5'- and 3'-ends of pre-miRNAs in the 5'- and 3'-pockets of Dicer's PAZ domain, cleaving 22 nts from the pre-miRNA 5'-end^{27, 42, 83}. In group II pre-miRNA, the pre-miRNA only has a 1-nt 3'-overhang. For the maturation pathway of group II pre-miRNA, TUT4 is recruited and monouridylates the 3'-end of the pre-miRNA to create the required 2-nt 3'-overhang^{42, 64} (Figure 9). Dicer is able to recognize the pre-miRNA, cleaving it into mature miRNA, allowing for the incorporation of one of the miRNA strands into an RNA-induced silencing complex (RISC) by Argonaute proteins^{83, 88}. Incorporation into a RISC complex leads to the suppression of the target gene expression⁸³.

In an alternate pathway, the RNA binding protein Lin28A binds to the group II pre-miRNA and recruits TUT4, which will polyuridylate the pre-miRNA⁸⁹. Binding between Lin28A and the pre-miRNA terminal loop occurs through a GGAG motif, conserved in pre-miRNA^{27, 42}. This interaction blocks Dicer from processing the pre-miRNA into mature miRNA^{40, 42, 89}. Recruitment of TUT4 results in the interaction between the C2H2 ZnF of TUT4 and Lin28A, resulting in polyuridylation of the pre-miRNA 3'-end⁹⁰. Polyuridylation of pre-miRNA triggers a

decay mechanism. The poly(U) tail is recognized by Dis3L2, a 3'-5' exonuclease, which will rapidly degrade the pre-miRNA^{89, 90, 91} (Figure 9). The function of TUT4 in the pre-miRNA biogenesis pathway has been well-studied in reference to the group II miRNA let-7^{27, 32, 40, 72}.

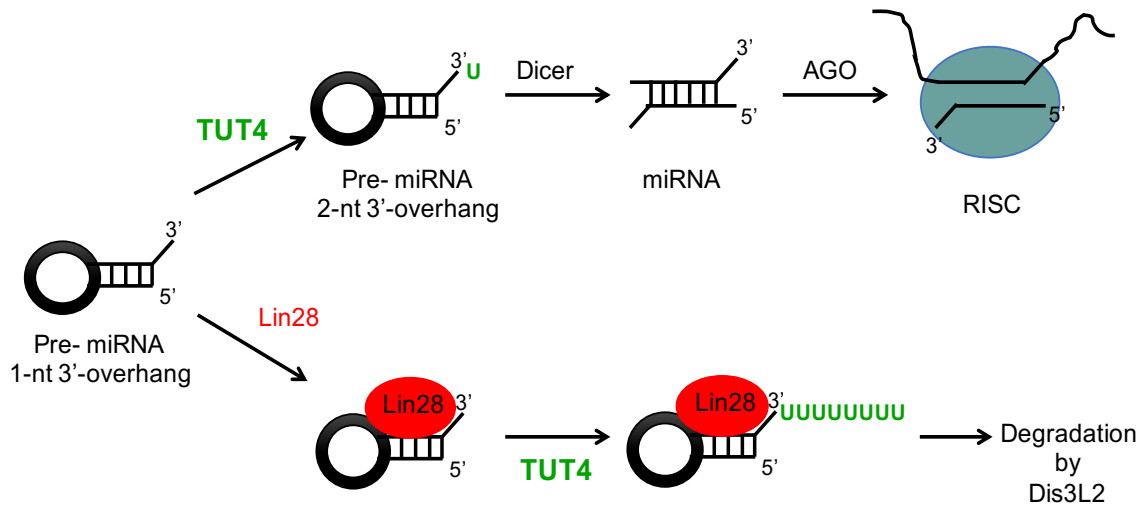


Figure 9. TUT4 involvement in miRNA maturation and degradation pathway.

Group II pre-miRNA exported out of the nucleus contains a 1-nt 3'-overhang, requiring further processing. In a monouridylation pathway, TUT4 adds a single uridine residue to the 3'-end of pre-miRNA, allowing for recognition by Dicer and further processing by Argonaute (AGO) proteins and incorporation into RNA-induced silencing complex (RISC). In an alternate, Lin28A-dependent pathway, Lin28A RNA binding protein binds to the pre-miRNA, recruiting TUT4 to polyuridylate the 3'-end of the pre-miRNA. The poly(U) tail is recognized by Dis3L2, a 3'-5' exonuclease, and degraded.

1.5 TUT4 is a possible therapeutic target in cancer

MiRNA uridylation, especially polyuridylation, must be tightly regulated to prevent oncogenesis. The uridylation pathway of the group II miRNA let-7 has been well studied due to its function as an important tumour suppressor in several oncogenic pathways including HMG2A, MYC, and RAS^{92, 93}. Lin28A, an important RNA binding protein in the miRNA degradation pathway, is a known suppressor of let-7 miRNA maturation^{92, 94}. Reduced let-7 levels, caused by an overexpression of Lin28, have previously shown to be related to several different types of cancer^{92, 94, 95, 96}. Interestingly, pri-miRNA let-7 has been shown to still be produced, while it is only the mature miRNA levels that are affected⁹⁴. When this pathway is regulated, Lin28-dependent let-7 degradation is required for normal development⁹². In cancers where Lin28A is overexpressed, Lin28A outcompetes Dicer for pre-let-7 binding, recruiting TUT4 into the let-7 degradation pathway which results in metastatic tumour formation⁹⁷. Piskounova and colleagues have previously shown that TUT4 is a potential therapeutic target in cancer cells where Lin28A is expressed⁹². When TUT4 is depleted, the invasiveness of T47D breast cancer cells with overexpressed Lin28A was reduced, due to an increase in mature let-7 miRNA levels⁹². Thus, TUT4 is a potential target for small molecule inhibitors as future chemotherapeutics.

1.6 TUT4 N-terminal domain activity plays a role in the cell cycle

Most studies on TUT4 indicate that the catalytic function of TUT4 is dependent on its C-terminal catalytic region (Ntr and PAP domains), and that the N-terminal Ntr and PAP domains are catalytically inactive. In 2011, Blahna et al. discovered a biological function of the N-terminal region of TUT4. Previously, Blahna et al. showed that TUT4 uridylated miRNA-26a (miR-26a), an miRNA which has cell cycle protein targets⁶⁷. This polyuridylation activity leads to the degradation of miR-26a, controlling the expression of the cytokine IL-6⁶⁷. This TUT4 activity intrigued Blahna et al., leading to an investigation of the potential role of TUT4 in cell proliferation. *In vivo* experiments, studying the role of TUT4 in H1299 lung epithelial cells, confirmed the importance of TUT4 in cell proliferation⁷⁴. A knockdown of TUT4 expression using siRNA showed a significant decrease in cell proliferation⁷⁴. Flow cytometry experiments, analyzing the stages of the cell cycle, showed an increase of the TUT4 knockdown cells accumulating in the G₁ phase⁷⁴. Overexpression of TUT4 had an opposite effect on H1299 cells *in vivo*, with an increase of progression into the S phase of the cell cycle⁷⁴. Immunoblots indicated TUT4 expression has a direct effect on cyclin A, a protein involved in cell cycle progression⁷⁴. Surprisingly, a catalytically dead mutant of TUT4 did not affect the cell cycle progression of H1299 cells, as well as no difference in cyclin A expression⁷⁴. H1299 cells transfected with truncations of TUT4 either missing the N- or C-terminal regions, termed Zcchc11-ΔN and Zcchc11-ΔC respectively, showed that overexpression of the TUT4 N-terminal domains had an effect on the cell cycle, similar to that of the full-length protein⁷⁴.

Interestingly, their research led to the discovery of N-terminal region activity and its effect on the cell cycle⁷⁴. Blahna and colleagues concluded that the N-terminal domains of TUT4 had a biological activity in the regulation of cell cycle development that was independent of TUT4's uridylation activity⁷⁴.

1.7 Rationale

The reason behind the existence of two encoded Ntr-PAP catalytic regions in both the N- and C-termini of TUT4 is unknown yet intriguing. In 2004, complementary DNA (cDNA) experiments performed by Toshio Ota and colleagues revealed the existence of a TUT4 isoform, termed TUT4-isoform 2⁹⁸. The discovery of TUT4-isoform 2 was part of a large-scale study, the “full-length long Japan” (FLJ), which created a collection of characterized cDNA from approximately 21,000 clones⁹⁸. This isoform, which is 719 amino acids in length, is identical to the full-length TUT4, or TUT4-isoform 1, in the first 684 amino acids. TUT4-isoform 2 differs, however, in the last 35 amino acids, from amino acid 685 to 719. In 2009, follow-up experiments, focused on the alternative splicing of cDNA, mapped out the isoform to two separate locus⁹⁹. The first 684 amino acids were located on the minus strand of chromosome 1, while the last 35 amino acids were located on the plus strand of chromosome 16⁹⁹. While intriguing, it is unclear whether this isoform exists *in vivo*.

While TUT4-isoform 2 may or may not be a naturally existing isoform, many splicing variants do occur, some only containing the N-terminal domains of TUT4. Experimental data from the NCBI Aceview database has shown the existence of 24 spliced, differentially spliced, or partial mRNA variants of TUT4,

with 14 of them resulting in the ability to encode proteins at a high level¹⁰⁰ (Figure 10). These 14 mRNA variants produce 15 different isoforms, some of them only containing the N-terminal domains¹⁰⁰. This raises questions as to what the function of the N-terminal catalytic region is since it is lacking an identifiable aspartate triad.

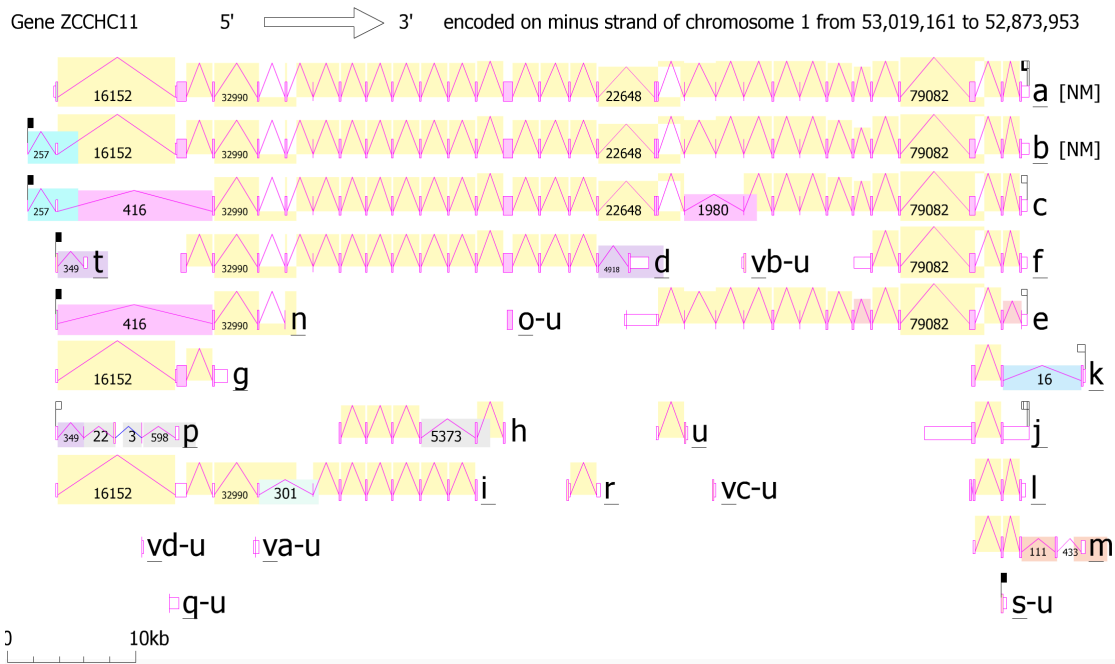


Figure 10. The existence of 24 different TUT4 splicing variants. NCBI Aceview database of naturally occurring TUT4 isoforms. Gene expression of isoforms is from 15 different primate species and 16 different tissues. TUT4 exists as 25 different isoforms, including 18 alternatively spliced and 7 unspliced variants¹⁰⁰.

1.8 Hypothesis and Aims

The importance of the activity of the N-terminal region in cell proliferation⁷⁴ and the existence of splicing variants encoding only the N-terminal domains raises the question of the function of the TUT4 N-terminal catalytic region. The N-terminal Ntr domain lacks an identifiable aspartate triad, and its catalytic function remains to be studied. I hypothesize that the N-terminal region is catalytically active and capable of uridylation activity. I will test my hypothesis with the following three aims:

Aim 1: Characterization of TUT4 N-terminal domain activity- I will clone and purify a truncated TUT4 variant encoding only the N-terminal domains. Following TUT4 N-terminal protein purification, I will use an activity assay utilizing radioactive [α -³²P]-UTP and various RNA substrates to elucidate whether the N-terminal catalytic region has the capability to uridylate RNA.

Aim 2: Mutational analysis and kinetic characterization- Once the activity of the N-terminal region is determined, I will mutate the proposed N-terminal active site residues accordingly to either diminish or introduce uridylation activity. If the TUT4 N-terminal wild-type (WT) protein or mutants display uridylation activity, a radioactive time course assay will be performed to determine the initial velocity of the active protein.

Aim 3: Characterization and kinetic analysis of TUT7 N- and C-terminal domains- Following TUT4 characterization, I will investigate the activity of the TUT7 N- and C-terminal catalytic domains, which are believed to be redundant to TUT4.

Chapter 2: Methods

2.1 Cloning

To express the N-terminal domains of TUT4 (TUT4-N), the segment encoding amino acids 1 to 678 of TUT4 was cloned into an expression vector for protein production in *Escherichia coli* (*E. coli*). The TUT4-N construct was cloned using the plasmid HsCD00347737 from the Harvard PlasmID Database as a template for gene amplification. Primers (TUT4-N FWD: 5'-ATTAGGATCCATGGAAGAGTCTAAAACC-3'; TUT4-N REV: 5'-TTTAGCGGCCGCCTATGAAAATGGATCTTCAATGG-3') were designed to amplify the gene segment consisting of nucleotides 1 to 2034 (amino acids 1 to 678) and add *Bam*HI and *Not*I restriction sites to the 5'- and 3'-ends of the PCR product, respectively. The gene segment was created using Polymerase Chain Reaction (PCR) with Phusion polymerase (New England Biolabs #M0530L) according to manufacturer's directions. The thermocycler protocol began by initially denaturing the DNA at 98°C for 30 seconds (sec.), followed by 30 cycles of [98°C for 10 sec., 55°C for 30 sec., and 72°C for 30 sec/kb] to denature, anneal and extend the DNA segment. A final extension at 72°C for 10 minutes (min.) ensures full-length amplification of the gene segment. Successful amplification was verified by gel electrophoresis on a 1% agarose gel with ethidium bromide at 140W in 1x TAE buffer (10x (pH 8.0): 40 mM Tris-acetate, 1 mM ethylenediaminetetraacetic acid (EDTA)). The PCR products were purified using the GeneJET PCR Purification Kit (Thermo Scientific K0702). For protein

expression, the PCR product was cloned into the expression vector pGEX-6p2. The pGEX-6p2 vector encodes an N-terminal Glutathione S-transferase (GST) tag. The TUT4-N PCR product and pGEX-6p2 vector were digested with restriction enzymes *Bam*HI (New England Biolabs #R3136S) on the 5'-end and *Not*I (New England Biolabs #R3189L) on the 3'-end in Cutsmart buffer (New England Biolabs # B7204S). Following digestion, the DNA was diluted with 6x-DNA loading dye and analyzed by 1% agarose gel electrophoresis. The correct DNA band was cut out of the gel and gel extracted using the GeneJET Gel Extraction Kit (Thermo Scientific K0692). The DNA concentration for both the PCR segment and the pGEX-6p2 vector were determined using the Nanodrop 2000c Spectrophotometer (Thermo Scientific). TUT4-N and pGEX-6p2 vector were ligated in a 3:1 ratio using T4 DNA ligase (New England Biolabs #M0202L) at 37°C for 1 hour, followed by heat-shock transformation into TOP10 competent cells. The cells were spread on Luria-Bertani (LB) agar plates containing 100 µg/mL ampicillin (AMP) and incubated overnight at 37°C. After transformation, up to 15 colonies are picked and resuspended in 50 mL of double distilled water (ddH₂O). PCR of each colony was performed using Taq DNA polymerase (GeneDireX MB101A-0500) according to manufacturer's directions and the 5'-pGex and 3'-pGex sequencing primers (5'-pGEX primer: 5'-GGGCTGGCAAGCCACGTTTGGTG-3'; 3'-pGEX primer: 5'-CCGGGAGCTGCATGTGTCAGAGG-3'). PCR products were analyzed by 1% agarose gel electrophoresis for confirmation of successful ligation. If ligation of the gene segment into the pGEX-6p2 vector was successful, a band at the

correct DNA size will be visualized on the agarose gel. Empty pGEX-6p2 vector DNA is run with the colony PCR DNA as a negative control. Colonies containing the PCR product were grown overnight in 5 mL of LB with 100 µg/mL AMP and isolated by miniprep using the GeneJET Plasmid Miniprep Kit (Thermo scientific K0503) before prepared DNA samples were sent to Genewiz for sequencing.

2.2 *Escherichia coli* competent cells and transformation

2.2.1 Chemically competent cells

TOP10 and BL21 codon plus (C+) *E. coli* cells were made competent by the RbCl₂ method. *E. coli* cells were grown aerobically in 250 mL of sterilized LB medium. BL21 C+ *E. coli* cells were grown with chloramphenicol (CM). When the culture reached an optical density at 600nm (OD₆₀₀) of 0.6, cells were harvested by centrifugation at 6,000 x g for 10 min. at 4°C. The cell sediment was resuspended in 100 mL of TFB1 pH 5.8 (30 mM Potassium acetate, 10 mM CaCl₂, 50 mM MnCl₂, 100 mM RbCl₂, 15% glycerol) and incubated on ice for 5 min. Following incubation, the cells were centrifuged at 3,000 x g for 5 min. at 4°C, and the cell sediment was resuspended in 2 volumes of TFB2 pH 6.5 (10 mM MOPS, 75 mM CaCl₂, 10 mM RbCl₂, 15% glycerol) and incubated on ice for 30 min. The cell suspension was divided into 40 µL aliquots in pre-chilled 1.5 mL Eppendorf tubes. The cell aliquots were stored at -80°C.

2.2.2 Heat-shock transformation

1 µL of DNA was incubated with 40 µL of chemically-competent *E. coli* cells and incubated on ice for 20 min. Following incubation on ice, the cells were

subjected to heat-shock at 42°C for 45 sec. and placed back on ice for 2 min. 1 mL of LB was added to the cells and incubated at 37°C for 1 hour at 300 rpm (Thermomix compact machine). After incubation, 100 µL of cells were spread onto an LB agar plate with the required antibiotics. The remaining suspended cells were centrifuged at 16,000 x g for 2 min. The supernatant was removed, and the cell sediment was resuspended in 100 µL of fresh LB and spread onto an LB agar plate with the required antibiotics. The plates were incubated overnight at 37°C.

2.3 Site-directed mutagenesis

Mutagenesis of D412A, N414A, D416A, and R468A, and a C-terminal 6-histidine (His₆) tag insertion were performed using site-directed mutagenesis PCR following the method outlined by Edelheit et al. 2009¹⁰¹. Forward and reverse primers were designed with complementary regions 15-20 nts upstream and downstream of the mutation site. The mutation is introduced into the sequence by PCR using primers with a mismatched codon sequence at the site of interest. This mismatch will change the coding sequence to a new amino acid after translation. For the His₆ insertion, instead of a sequence mutation, a 6-histidine sequence was inserted into the plasmid sequence at the site of interest. The primer sequences are listed in Table 2, as well as the codon change. PCR was performed with two separate single-primer reactions for the forward and reverse primers using Phusion polymerase as follows: 94°C for 2 min., [94°C for 40 sec., 55°C for 40 sec., 72°C for 30 sec./kb of plasmid] x30 cycles, and 37°C

infinite. Single-primer reactions prevent the unintended annealing of the primers and formation of “primer dimers”. Forward and reverse primer reactions were combined and incubated at the following temperatures: 95°C for 5 min., 90°C for 1 min., 80°C for 1 min., 70°C for 30 sec., 60°C for 30 sec., 50°C for 30 sec., 40°C for 30 sec., 37°C for infinity. This allows for annealing of the parental DNA strands and the newly synthesized DNA strands. Final products were *DpnI* (New England Biolabs #MR0176S) digested for 2 hours at 37°C. *DpnI* digests methylated parent DNA, ensuring only the newly synthesized PCR products remain. This was followed by transformation using the heat-shock method in TOP10 competent cells. Transformed cells were spread on LB plates with 100 µg/mL AMP and incubated overnight at 37°C. 5-10 colonies were picked and grown in 5 mL LB overnight cultures with 100 µg/mL AMP and purified by plasmid miniprep using the GeneJET Plasmid Miniprep Kit (Thermo Scientific, K0503) according to manufacturer’s instructions and sent for sequencing (Genewiz).

Table 2. Primer sequences for site-directed mutagenesis

Mutant	Primer	Sequence	Codon Change
D412A	Forward	5'-GCTCTGAAAAGTAGT GC AGTTAATATAGATATA-3'	GAT to GCA
	Reverse	5'-TATATCTATATTA ACTGC ACTACTTTTCAGAGC-3'	
N414A	Forward	5'-CTGAAAAGTAGTGATGTT GCA ATAGATATAAAA TTTCC-3'	AAT to GCA
	Reverse	5'-GGAAATTTTATATCTATT GCA ACATCACTACT TTTCAG-3'	
D416A	Forward	5'-GCTCTGAAAAGTAGTGATGTTAATATAG CA ATA AAATTCCTCCCAAGATGAATCATCC-3'	GAT to GCA
	Reverse	5'-GGATGATTCATCTTGGGAGGAAATTTTATT GCT ATATTAACATCACTACTTTTCAGAGC-3'	
R468A	Forward	5'-GTTGTGGTGTGCAGAGATCGAAAAAGTGTTT ACTTTGT GC AGTGAGTGCAGGAAACGATATGGCATGT CTCACTACTGATTTACTTACTGCC-3'	AGA to GCA
	Reverse	5'-GGGCAGTAAGTAAATCAGTAGTGAGACATGCC ATATCGTTTCCTGCACTCACT GC ACAAAGTAAACCACT TTTTCGATCTCT GCACACCACAAC-3'	
His ₆	Forward	5'-GAAGATCCATTTT CATCATCACCATCACC ACT AGGCGG CCGCATCGTG-3'	His ₆ addition
	Reverse	5'-CACGATGCGGCCGCCT AGGGTGATGGTGAT GATGT GAAAATGGATCTTC-3'	

2.4 Purification of TUT4 proteins

To purify the N-terminal domains of TUT4, TUT4-N in pGEX-6p2-His was transformed into BL21 C+ competent *E. coli* cells via heat-shock and grown on LB plates containing 100 µg/mL AMP and 34 µg/mL CM. A colony was selected and grown in 50 mL LB culture containing 100 µg/mL AMP and 34 µg/mL CM overnight. 10 mL of the overnight culture was transferred to a 1 L flask containing

LB supplemented with 100 $\mu\text{g}/\text{mL}$ AMP and 34 $\mu\text{g}/\text{mL}$ CM. Five 1 L cultures were grown, using 10 mL of cells/L from the 50 mL pre-culture, to an OD600 of 0.6, and protein expression was induced with 250 μM of isopropyl- β -D-1-thiogalactopyranoside (IPTG). The cells were collected by centrifugation in 1-L bottles at 9,000 x g for 10 minutes. The cell sediment was transferred to 50 mL falcon tubes and centrifuged at 3,000 x g for 15 minutes. The cells were stored at -80°C .

The cells were resuspended in 10 mL of lysis buffer (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 0.1 mM EDTA, 1 mM dithiothreitol (DTT), 0.1% Triton X-100, 5 mM MgCl_2 , 10% glycerol) per litre of cells and lysed via a French Pressure Cell. For French Press cell lysis, resuspended cells are applied to the French pressure cell and placed under constant pressure of 1000 psi, forcing cell suspension through a small flow outlet. Soluble protein was separated from cell debris using ultracentrifugation (Beckman Coulter Ultracentrifuge) at 150,000 x g for 1 hour and the supernatant containing the protein was collected. Proteins were first purified on the AKTA FPLC (GE) using a pre-packed GSTrap FF 5 mL glutathione column (GE 17-5131). Non-specific proteins were washed off the column with Buffer A (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 0.1 mM EDTA, 1 mM DTT, 0.1% Triton X-100, 5 mM MgCl_2 , 10% glycerol), and TUT4-N protein was eluted in a gradient elution from 0-100% Buffer B (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 5 mM MgCl_2 , 10% glycerol, 10 mM glutathione) over 45 mL. Elutions were collected in 2 mL fractions. Fractions were analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) for the presence

of TUT4.

Elution fractions containing the TUT4-N proteins were further purified using the C-terminal His₆ tag on a nickel-nitrilotriacetic acid (Ni-NTA) affinity column. The protein was loaded onto HisPur™ Ni-NTA resin (Thermo Scientific 88222) and washed with Buffer C (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 5 mM MgCl₂, 50 mM imidazole, 10% glycerol). Protein was eluted in 1 mL fractions with Buffer D (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 5 mM MgCl₂, 500 mM imidazole, 10% glycerol). Fractions were analyzed by SDS-PAGE for the presence of TUT4.

Elution fractions were concentrated using a Vivaspin 6 ultrafiltration concentrator with a 10,000 MWCO (Sartorius VS0601) at 4,000 x g and dialyzed using Biotech RC Tubing dialysis membrane (Spectra/Por® Biotech Dialysis Membrane 133342) overnight into Buffer E (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 5 mM MgCl₂, 10% glycerol). The protein was further concentrated using the Amicon® Ultra-0.5mL Centrifugal Filters Ultracel® 3K concentrator (Millipore Sigma UFC500324). All TUT4-N mutants were purified following the same procedure. Samples from each step of the purification of the TUT4-N proteins were analyzed by SDS-PAGE as described below.

2.5 SDS polyacrylamide gel electrophoresis

Protein samples were diluted in 3x sodium dodecyl sulfate (SDS) loading dye and boiled for 5 min. at 95°C. Samples were loaded on an 8% separating SDS-PAGE gel with a 5% stacking gel and run at 100V in SDS running buffer

until the loading dye front runs off the gel. The gel was stained in Coomassie Blue for an hour, followed by an incubation in destain for an hour. The SDS-PAGE gel was imaged on the Chemidoc MP imaging system (Bio-Rad).

2.6 Bradford assay

BSA protein standard (0 mg, 2.5 mg, 5 mg, 10 mg) was diluted in ddH₂O to a final volume of 100 μ L. 5 μ L of purified protein was diluted to a final volume of 100 μ L. 5x Bradford reagent (Bio-Rad, #5000006) was diluted to 1x and 1 mL of reagent was added to each Eppendorf tube and mixed well. The Eppendorf tubes were incubated for 15 min. at room temperature. Absorbance at 595 nm was measured for each standard and the purified protein. The absorbance of the standards was plotted against protein concentration to determine the standard curve. Using the linear equation of the standard curve, the concentration of the purified protein was calculated.

2.7 Western blot

Purified proteins (TUT4-N, D412A, N414A, and D416A) were run on an 8% SDS-PAGE gel, as described above. After SDS-PAGE, the gel and Whatman paper was pre-incubated in transfer buffer for 15 min. PVDF membrane (Immobilon[®] -FL Transfer Membrane IPFL00010) was activated by incubation in 100% methanol for 10 min. before subsequent incubation in transfer buffer. This pre-incubation was followed by transferring of the proteins to the PVDF membrane using the Trans-Blot[®] Turbo[™] transfer system (Bio-Rad). The gel was placed on top of the PVDF membrane between three pieces of Whatman paper

on each side. The layered membrane and gel are placed in the cassette of the Trans-Blot[®] Turbo[™] transfer system and the protein was transferred at 1.3A and 25V for 15 min. The membrane was subsequently incubated in blocking solution (5% skim milk in 1x PBS-tween (10% 10x-PBS [1.37 M NaCl, 27 mM KCl, 100 mM Na₂HPO₄, 20 mM KH₂PO₄], 0.1% Tween20, ddH₂O) for 1 hour to prevent non-specific binding, followed by incubation with the primary anti-TUT4 antibody (Proteintech 18980-1-AP) (1:200 in blocking solution) overnight at 4°C with agitation at 40 rpm on a Standard Analog Shaker (VWR). The membrane was washed 3 times in wash solution (1% skim milk in 1x PBS-tween) for 5 min. This was followed by incubation with the secondary antibody anti-rabbit IgG HRP-linked F(ab')₂ fragment (GE life sciences NA9340V) (1:500 in wash solution) for 2 hours at room temperature with agitation. The blot was washed with 1x PBS-tween 3 times for 5 min. before imaging. Antibody binding was visualized, taking advantage of the reaction catalyzed by horseradish peroxidase (HRP), which is linked to the secondary antibody. In this reaction, HRP catalyzes the oxidation of luminol in the presence of hydrogen peroxide, an oxidizing agent, which in turn emits light^{102, 103}. This light emission can be imaged and is correlated to the location of the protein immobilized on the membrane. The blots were imaged using Clarity Max[™] ECL substrate (Bio-Rad 1705062S) on the Chemidoc MP imaging system.

2.8 TUT4 enzyme activity assay

2.8.1 End-point activity assay

To test the enzymatic activity of TUT4 protein variants, a radioactive activity assay was performed, monitoring the addition of a radiolabelled nucleotide to a substrate RNA. The purified proteins (TUT4-N, D412A, N414A, D416A, and R468A) were individually incubated with 3.2 mM MgCl₂, 1 mM DTT, 1 μM RNA, and 0.5 μL of 3.3 μM [α -³²P]-UTP (Perkin Elmer EasyTides® Uridine 5'-triphosphate [α -³²P] BLU507H25OUC) at 37°C for 30 min. RNAs miR-122 (Sigma), 15A RNA (Sigma), let-7a (Sigma), and pre-let-7a (purified in lab) were used. The reactions were stopped using 2x RNA loading dye (95% v/v formamide, 0.1% w/v xylene cyanol, 0.1% w/w bromophenol blue, 10 mM EDTA). Uridylated RNA products were separated on a 12% polyacrylamide gel at 30W for 2 hours and exposed overnight on a phosphor-imaging screen. Radioactivity was imaged using the Storm 820 imager (Molecular Dynamics). For nucleotide size reference, radiolabelled RNA ladder was used (Invitrogen Decade™ Markers System 7778a).

2.8.2 TUT4 inhibitors

Known TUT4 inhibitors were used in the radioactive activity assay reactions to determine the effect on TUT4-N. These inhibitors were validated through experimentation by Lin and Gregory (2015)¹⁰⁴. The two most successful selective TUT4 inhibitors are aurothioglucose (auro) and IPA-3. In the end-point activity assay, auro (Sigma A0606-5MG) and IPA-3 (Sigma I2285-5MG) were

added to a final concentration of 0.1 μM .

2.8.3 PUP time course activity assay

Commercial PUP (New England Biolabs #M0337S) was incubated with 3.2 mM MgCl_2 , 1 mM DTT, 1 μM 15A RNA, 0.5 μL of 1.65 μM [α - ^{32}P]-UTP, and 0.95 μM cold UTP. The reaction was stopped using 2x RNA loading dye at 1, 2, 5, 10, and 20-min. time points. The reaction products were separated on a 12% polyacrylamide RNA gel and exposed overnight at -80°C on a phosphor-imaging screen. Standards of defined amounts of radiolabelled nucleotide (3.3, 33, 330, 3300 & 33000 pM) were spotted on Whatman paper and exposed with the gel. The radioactivity was imaged using the Storm scanner and the product concentration of radiolabelled RNA was quantified and plotted on a scatter plot. A dilution factor of 24 was taken into consideration for quantification purposes.

2.8.4 TUT4-N time course activity assay

12 μg (0.6 mg/mL) of TUT4-N and mutant protein was incubated in a 20 μL reaction with 3.2 mM MgCl_2 , 1 mM DTT, 1 μM 15A RNA, and 1 μL [α - ^{32}P]-UTP. The reaction was incubated for 30 sec. at 37°C before addition of TUT4 enzyme. The reaction was stopped with 2x RNA loading dye at 2, 5, 10, 20, 30, and 60-min. time points. The reaction products were separated on a 12% polyacrylamide gel at 30W, followed by exposure on a phosphor-imaging screen overnight at -80°C . Standards of defined amounts of radiolabelled nucleotide (3.3, 33, 330, 3300 & 33000 pM) were spotted on Whatman paper and exposed with the gel. The radioactivity was imaged using the Storm scanner and the

product concentration of uridylated RNA was quantified using the standards and plotted on a scatter plot. Using the slope from the graph, the initial velocities and specific activities of the TUT4-N and mutants were calculated.

Chapter 3: Results

3.1 TUT4-N was cloned into an *E. coli* expression vector

To study the N-terminal region of TUT4, a TUT4 construct containing only the N-terminal domains was cloned using the plasmid HsCD00347737 as described in Chapter 2 (Figure 11). The construct, which will be referred to as TUT4-N, encodes amino acids 1 to 678 (Figure 11A) and was cloned into a pGEX-6p2 vector with an N-terminal GST tag. For more efficient purification, a C-terminal His₆ tag was incorporated using site-directed mutagenesis, as described in Chapter 2. The TUT4-N plasmid was transformed into BL21 C+ *E. coli* cells for expression.

3.1.1 *The N-terminal domain lacks an identifiable catalytic aspartate triad*

An amino acid sequence alignment of the N- and C-terminal domains of TUT4 was performed to determine which amino acids in the N-terminal Ntr domain correspond to the catalytic aspartate triad in the C-terminal Ntr domain. The catalytic triad, common to all nucleotidyltransferases and described in detail in Faehnle et al. 2017³², of the C-terminal Ntr domain consist of aspartates D1009, D1011, and D1070. The alignment of other terminal uridylyltransferases with the N- and C-terminals of TUT4 show that this catalytic triad is conserved across species (Figure 11B). Our sequence alignment further shows that the corresponding amino acids in the N-terminal Ntr domain are D412, N414, and R468 (Figure 11B). Two of the three aspartates that are part of the catalytic triad

in the C-terminal Ntr domain are not encoded in the corresponding positions in the N-terminal Ntr domain.

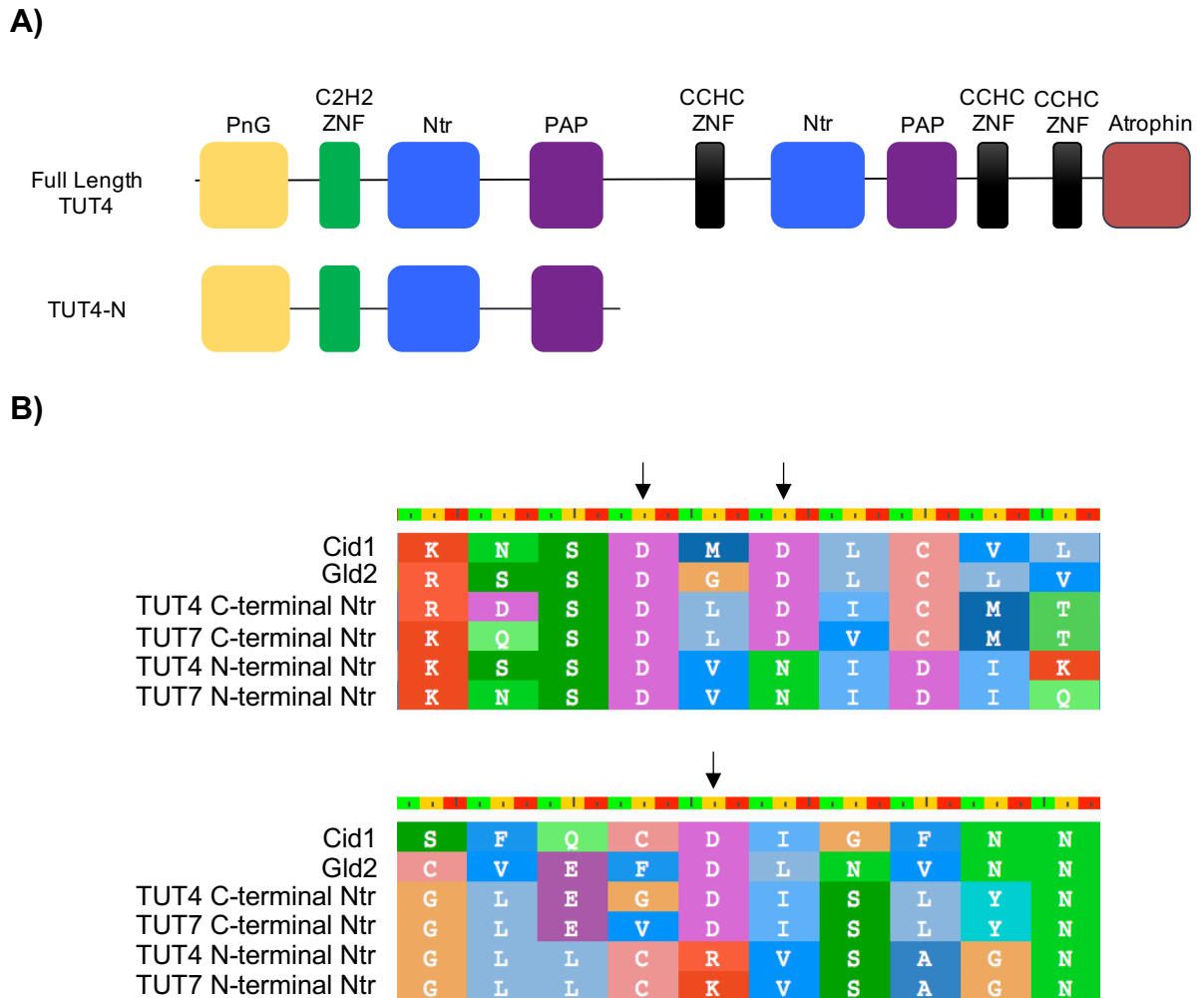


Figure 11. TUT4 N-terminal domain layout and sequence alignment. A) Full-length TUT4 encodes several zinc fingers (ZnF), and two copies of the catalytic Ntr and PAP domains. The expression construct for the N-terminal region (TUT4-N) encodes the TUT4 N-terminal catalytic region with the Ntr domain (blue) and the PAP domain (purple). The N-terminal C2H2 ZnF (green) was shown to interact with Lin28. The PneumoG (PnG) domain (yellow) has no known activity. **B)** Sequence alignment of conserved aspartate residues in TUTases. The catalytic triad is conserved in TUTases across species (Cid1 in yeast, Gld2 PAP in humans). The TUT4 (and TUT7) N-terminal Ntr domain lacks two of the aspartate residues, with an asparagine (Asn(N)) and arginine (Arg(R)) (or lysine (Lys(K)) for TUT7) in their place.

3.2 Purification of TUT4-N

3.2.1 *TUT4-N was partially purified*

TUT4-N was purified using both glutathione affinity and Ni-NTA affinity chromatography, as described in Chapter 2, taking advantage of both the N-terminal GST tag and C-terminal His₆ tag of the recombinant fusion protein. This two-step process was performed to remove as many contaminants as possible. BL21 C+ *E. coli* cells expressing the TUT4-N protein were lysed with a French Pressure Cell and proteins were separated from cell debris by ultracentrifugation. The TUT4-N protein was initially purified by glutathione affinity purification using a pre-packed GST column (GSTrap FF 5 mL) on the AKTA FPLC purification system. The protein was eluted with glutathione, and fractions containing protein of the expected relative molecular mass were identified by SDS gel electrophoresis and Western blotting with a TUT4-specific antibody. The fractions containing eluted TUT4-N protein were then further purified by Ni-NTA affinity chromatography, again followed by SDS gel electrophoresis to identify elution fractions containing recombinant TUT4-N. After Ni-NTA purification, the fractions containing TUT4-N protein were dialyzed overnight to be used for downstream applications. Samples from different steps of the purification process were run on an SDS-PAGE gel to demonstrate the elimination of non-specific, contaminating proteins (Figure 12). Ni-NTA elution fractions containing the TUT4-N protein were concentrated (Figure 13A) and a Bradford assay was performed to determine TUT4-N protein concentration. The protein yield was generally between 0.1 and 0.2 mg/ L *E. coli* culture. The protein was concentrated to 0.95 mg/mL.

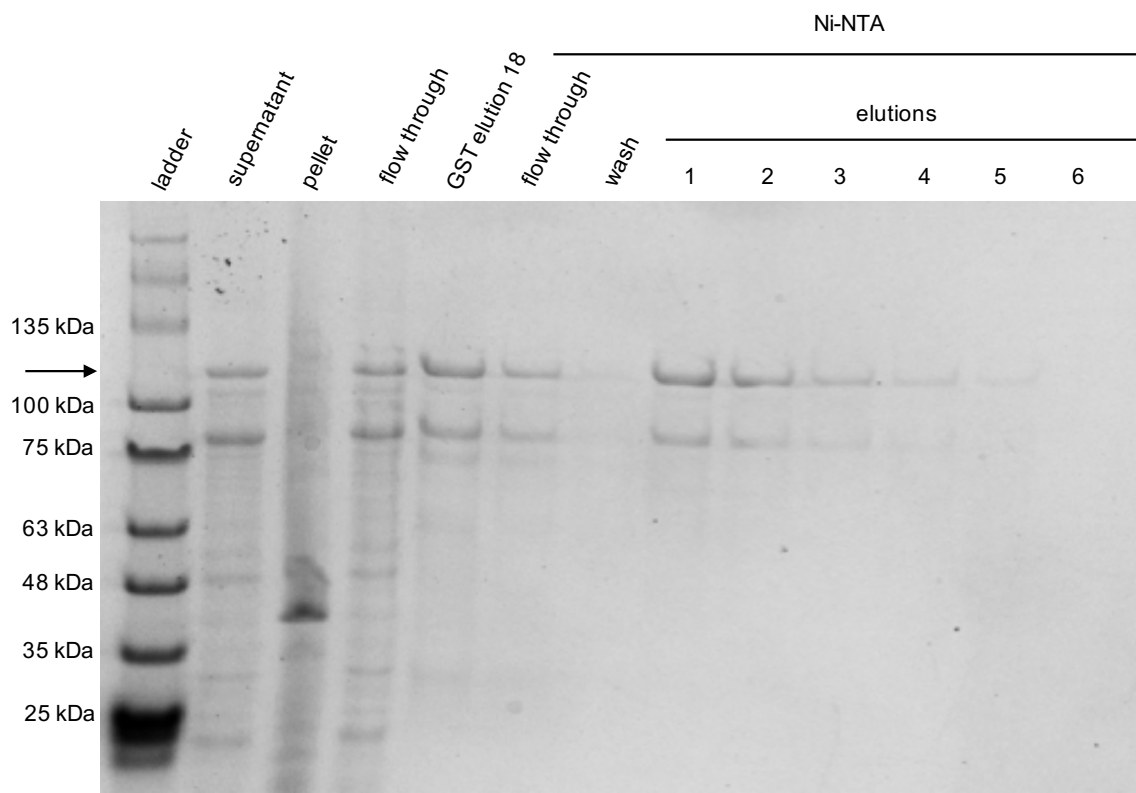


Figure 12. Two-step purification process of TUT4-N. Image of an SDS gel of the TUT4-N purification process. After an initial GST purification, the protein was further purified by His-tag affinity purification. The relative molecular mass of TUT4-N is calculated as 105 kDa, and a band is visible at that mass, indicated by an arrow.

3.2.2 Expression of TUT4-N was confirmed by Western blot and mass spectrometry

To verify the identity of purified proteins, a Western blot with an antibody specific to TUT4 was performed. The TUT4-specific antibody was used as the primary antibody. The TUT4 antibody is specific for a peptide located in the N-terminal region of TUT4, specifically amino acids 265 to 282. A band was detected at approximately 105 kDa, which corresponds to the expected relative molecular mass of GST and His-tagged TUT4-N, confirming the successful purification of TUT4-N (Figure 13B). The 105 kDa protein was excised from the gel and analyzed by mass spectrometry (London Regional Proteomics Centre) (Figure 13C). Mass spectrometry and Western blot confirmed that TUT4-N protein was in fact purified and was present at 105 kDa. A second, prominent band at a relative molecular mass of 80 kDa, seen on the SDS-PAGE and Western blot, could be a result of protein degradation.

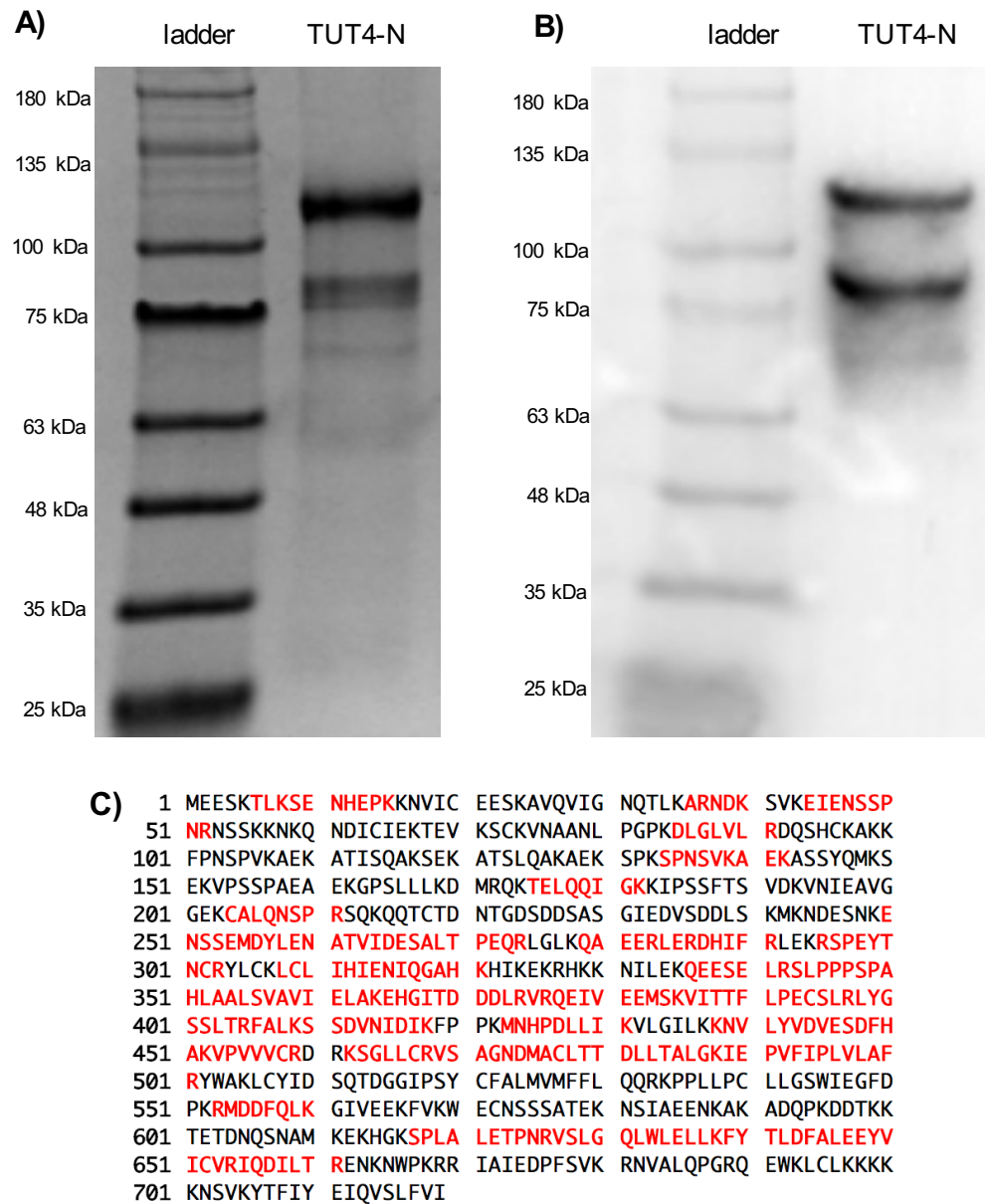


Figure 13. Verification of purified TUT4-N. **A)** SDS-PAGE and **B)** Western blot of concentrated TUT4-N protein. Western blot was performed using a TUT4-specific primary antibody. Distinct bands can be seen at 105 kDa and 80 kDa. **C)** Mass spectrometry of TUT4-N purification. The 105 kDa protein from the TUT4-N purification was analyzed by mass spectrometry. The peptides identified by mass spectrometry are highlighted in red.

3.2.3 Mock purification

Our two-step purification (GST and His₆ tag) did not yield entirely pure protein, and further purification steps like gel filtration or anion exchange chromatography led to significant degradation of TUT4-N (data not shown). Since TUT4-N expresses at very low levels (0.1-0.2 mg/L *E. coli* culture), and extended purification protocols lead to degradation of the protein, further purification of TUT4-N was not successful. To confirm that enzymatic activity is in fact due to the recombinantly purified protein and not residual contaminants with *E. coli* proteins, a mock protein purification was performed. The empty vector (pGex-6p2-His) was transformed into BL21 C+ *E. coli* cells and the protein production and purification was performed under identical conditions as described for the TUT4-N purification. The cells were lysed, and the protein was purified using the same two-step process as the TUT4-N protein (Figure 14). This mock purification was used as a negative control in the subsequent radioactive activity assay. From here on, this protein will be referred to as 'Mock'.

3.3 The TUT4 N-terminal domain is catalytically active

3.3.1 The *Cid1* poly(U) polymerase is catalytically active with several RNA substrates

To establish an activity assay using radioactive nucleotides as substrate, commercially available poly(U) polymerase (PUP) *Cid1* (NEB M0337S) was used as a positive control. *Cid1* is isolated from an *E. coli* strain carrying the PUP gene from *S. pombe*. *Cid1* is capable of utilizing UTP for RNA 3'-end labelling as described previously^{16, 57}. A radioactive activity assay was performed, and *Cid1* was incubated with several different RNA substrates and radioactive [α -³²P]-UTP for 30 min. in an end-point reaction. The reaction products were separated by gel electrophoresis on a 12% polyacrylamide gel, and the radioactive signal of the uridylated RNA was visualized using phosphor-imaging (Figure 15). 'No enzyme' and 'no RNA' control reactions were used as negative controls. Polyuridylation activity can be seen as a distinct ladder of bands for RNA substrates 15A RNA (15 nts), miR-122 (22 nts), and let-7a (22 nts). *Cid1* catalyzed the addition of UTP to a variety of RNA substrates, including pre-let-7a (70 nts), and total yeast tRNA (72-95 nts). Very little activity was observed on total RNA extracts from human cells (HEK293T and MDA-MB-231 breast cancer cells). This demonstrates that the radioactive activity assay works well in determining enzyme activity, and therefore will be used to determine TUT4-N activity.

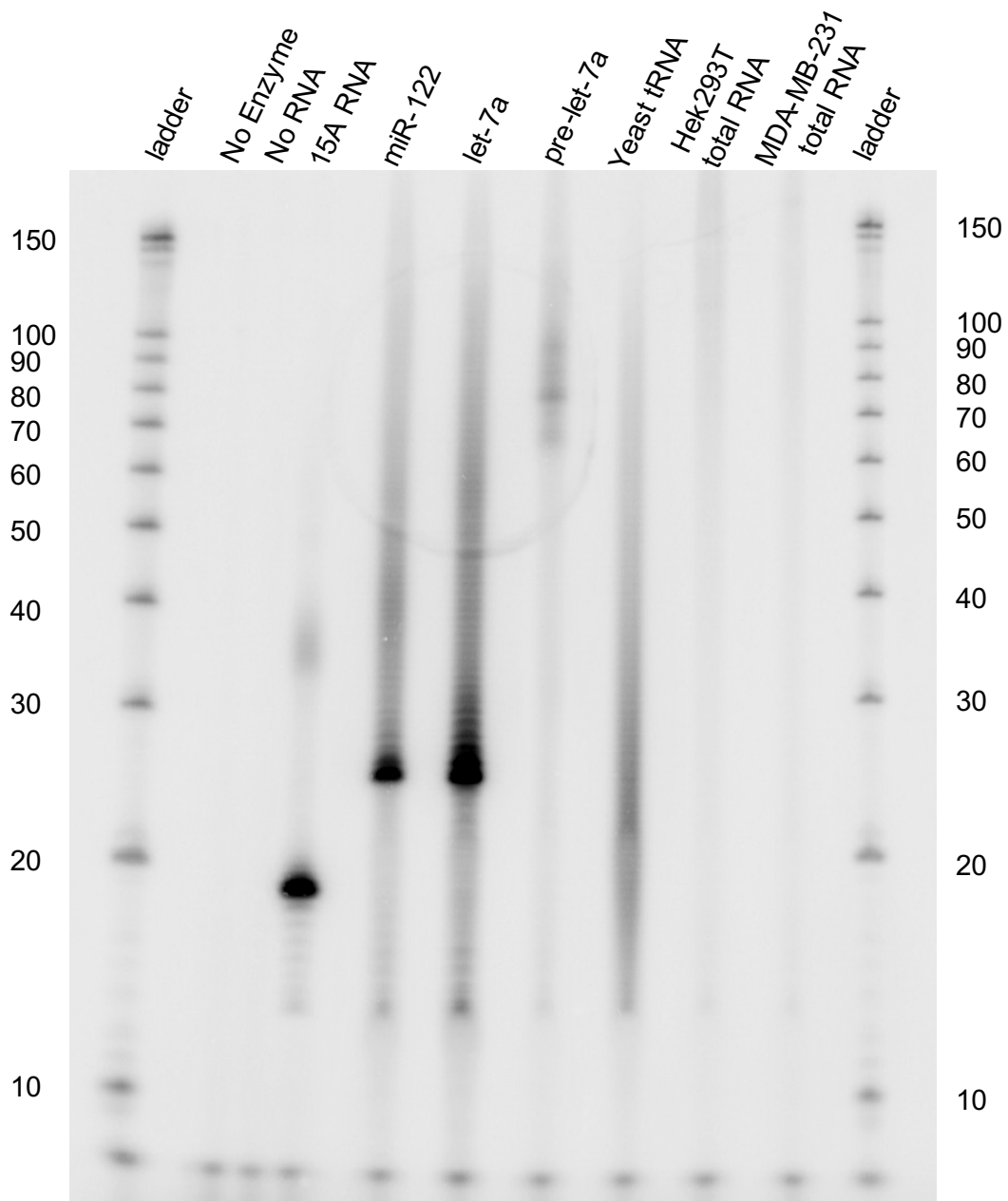


Figure 15. Cid1 is catalytically active with a variety of substrates. Cid1 is active with 15A RNA (15 nts), miR-122 (22 nts), let-7a (22 nts), pre-let-7a (70 nts), and total yeast tRNA. No enzyme control contains 1 μ M RNA substrate. Distinct polyuridylation activity can be seen as a ladder of bands for 15A RNA, miR-122, and let-7a. Radioactive bands seen in all lanes including the negative controls are common to the assay and stem from the radiolabelled nucleotide.

3.3.2 *TUT4-N is capable of uridylation activity with select substrates*

The N-terminal catalytic region of TUT4 is believed to be inactive due to the lack of an identifiable aspartate triad in the N-terminal Ntr domain. To test the activity of TUT4-N, I performed a radioactive activity assay, as described in Chapter 2. Radiolabelled UTP ($[\alpha\text{-}^{32}\text{P}]\text{-UTP}$) was incubated with TUT4-N and RNA substrate for 30 min. in an end-point reaction. Reactions without enzyme (No enzyme) or RNA (No RNA) were performed as negative controls. Four RNA substrates were used, including the miRNAs let-7a and miR-122, pre-miRNA pre-let-7a, and an mRNA poly(A) tail mimic, 15A RNA. Reactions including Lin28A, an RNA binding protein, or the previously described TUT4 inhibitors aurothioglucose (auro) or IPA-3¹⁰⁴, were performed with miR-122 to determine their effect on the N-terminal catalytic region of TUT4. TUT4-N displayed uridylation activity with two of the RNA substrates (Figure 16). Overall, the protein displayed very little catalytic activity, and high concentrations of proteins were required (0.6 mg/mL). MiR-122 and 15A RNA were uridylated by TUT4-N, as visible by the radioactive signal at the correct nucleotide lengths (~23 nts and ~16 nts, respectively). Interestingly, TUT4-N was not active with let-7a and pre-let-7a, both previously described substrates of the full-length TUT4. Lin28A and aurothioglucose displayed no effect on the activity of TUT4-N, while the addition of IPA-3 in the reaction resulted in an increased product formation.

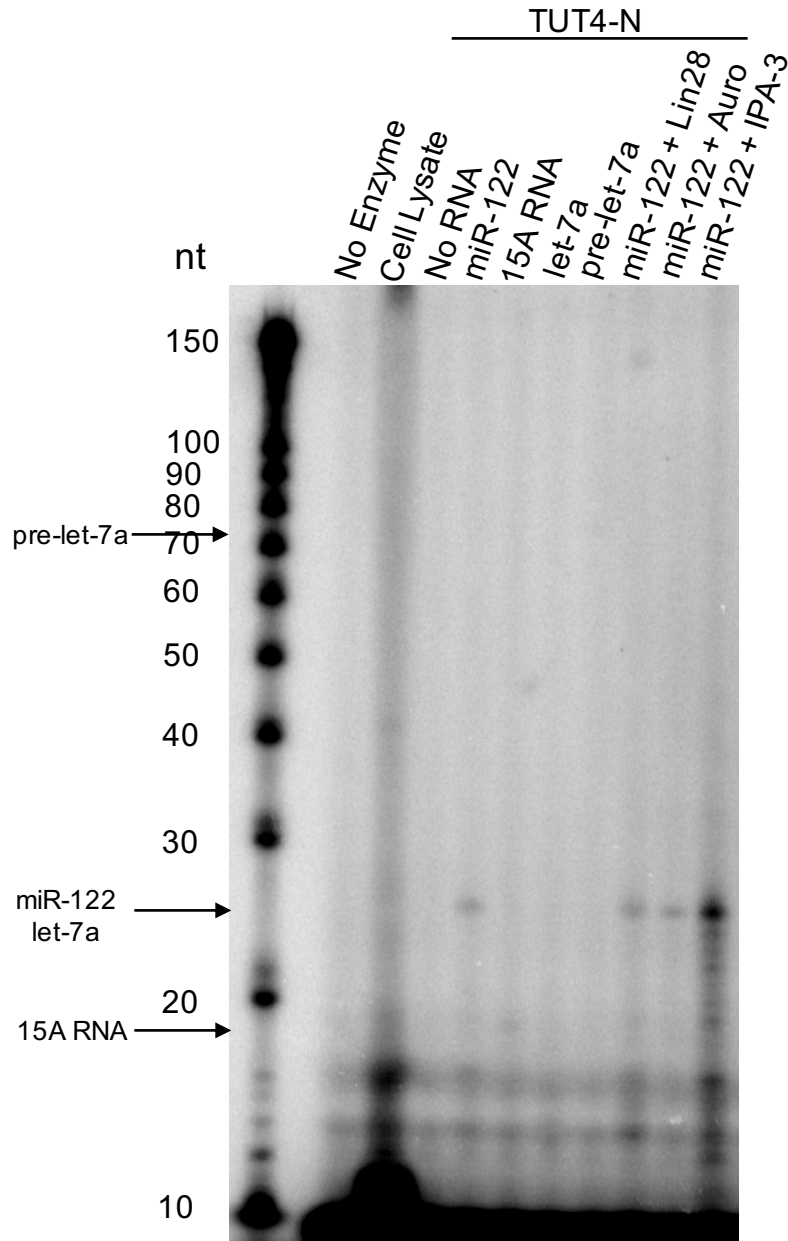


Figure 16. TUT4-N exhibits uridylation activity. TUT4-N has *in vitro* uridylation activity with miR-122 (22 nts) and 15A RNA (15 nts). A single band can be seen for both reactions at the correct nucleotide length. No activity is seen with let-7a (22 nts) or pre-let-7a (70 nts). Lin28 has no effect on TUT4-N activity with miR-122. No enzyme, no RNA, and cell lysate reactions were used as negative controls. Radioactive bands seen in all lanes including the negative controls are common to the assay and stem from the radiolabelled nucleotide.

3.3.3 DMSO stabilizes TUT4-N and increases product formation

As shown above, the addition of IPA-3 to the enzymatic reaction resulted in an increase of enzymatic activity. To test whether the increased product formation observed upon addition of IPA-3 to the reaction was due to IPA-3 or the solvent DMSO, I carried out a series of activity assays with increasing reactions of DMSO (5 to 30%). Interestingly, DMSO addition led to an increase in enzyme activity (Figure 17). As it is unlikely that DMSO participates in the catalytic reaction, DMSO likely stabilized TUT4-N and allowed for increased turnover rates.

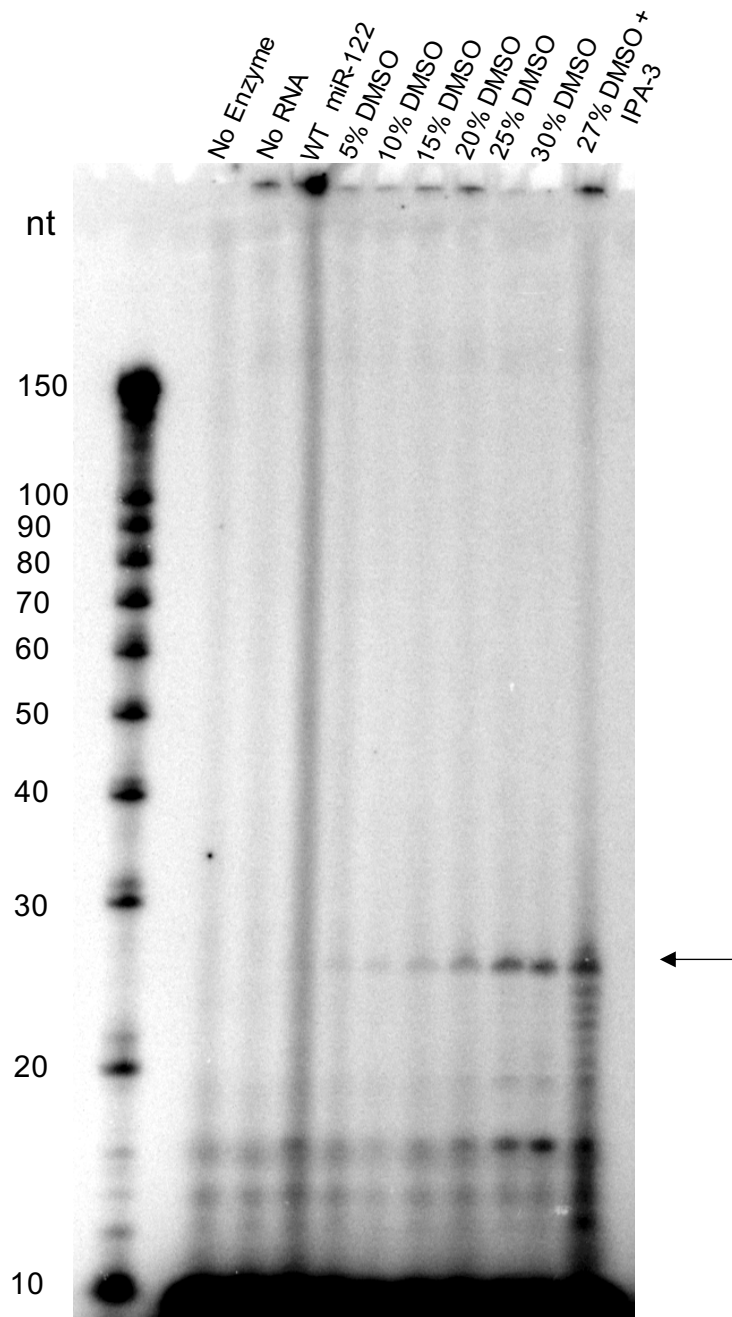


Figure 17. DMSO stabilizes TUT4-N and increases its enzymatic activity.

Uridylation activity of TUT4-N with miR-122 (22 nts) increases with increasing concentrations of DMSO (5 to 30%). The reaction product (uridylated miR-122) is indicated by an arrow. Radioactive bands seen in all lanes including the negative controls are common to the assay and stem from the radiolabelled nucleotide.

3.3.4 Mock protein activity

While TUT4-N was significantly purified in our two-step purification process, we could not exclude that the observed uridylation activity was due to potential *E. coli* protein contaminants. To determine that the activity is in fact from the TUT4-N protein, reactions were performed using Mock protein, obtained from a mock protein purification described above. Reactions containing 12 μg (1.2 mg/mL) of Mock protein, double the amount of total protein in the TUT4-N reactions, and 50 μg (5 mg/mL) of Mock protein, were performed with both miR-122 and 15A RNA. 'No enzyme' and 'no RNA' reactions were used as negative controls, while TUT4-N with 15A RNA was used as a positive control. As expected, only the TUT4-N positive control showed uridylation activity, at approximately 16 nts (Figure 18). This demonstrates that the uridylation activity is from the presence of TUT4-N and not a contaminating bacterial protein.

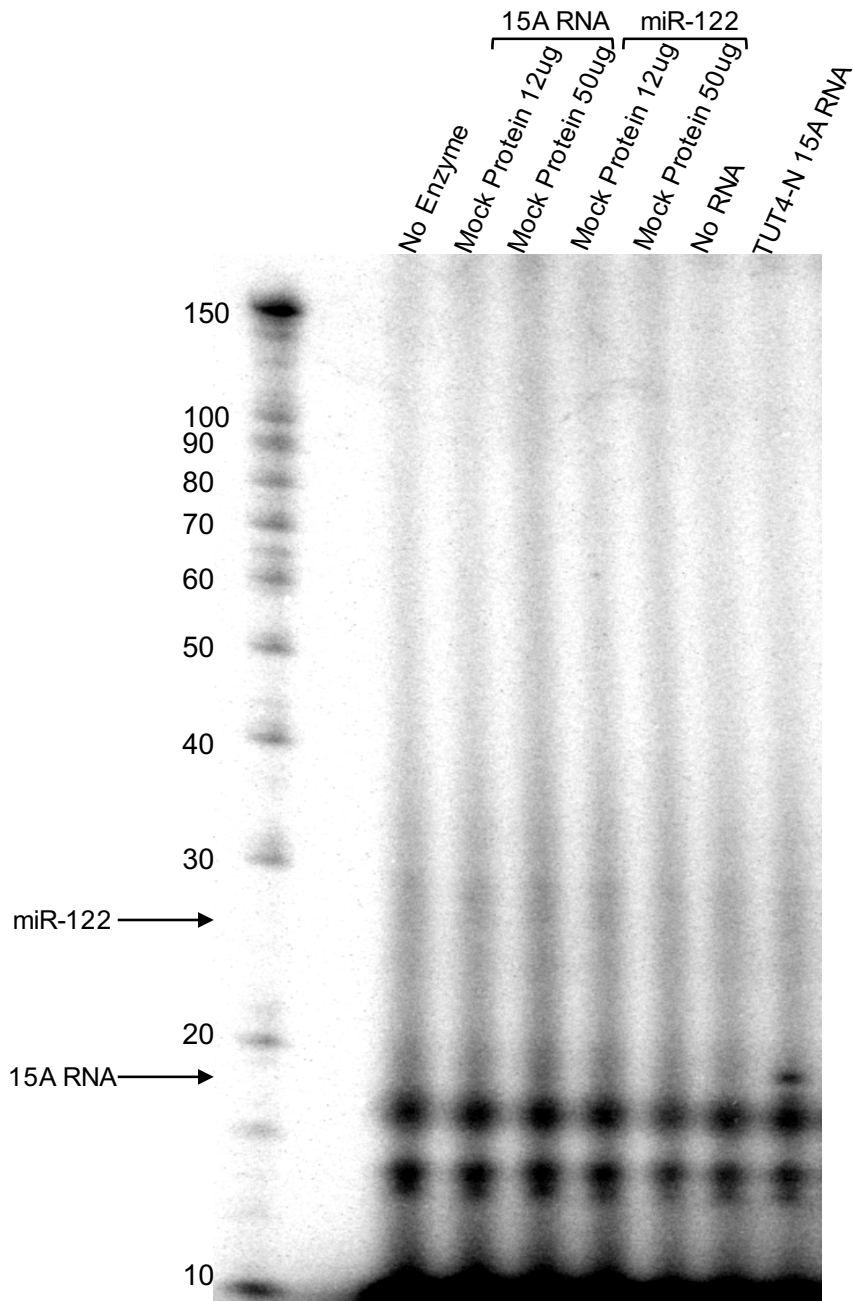


Figure 18. Confirmation of TUT4-N uridylation activity using Mock protein. No activity is seen with the Mock protein with 15A RNA (15 nts) or miR-122 (22 nts). TUT4-N activity can be seen as a distinct band in a positive control reaction with 15A RNA. ‘No enzyme’ and ‘no RNA’ reactions were used as negative controls. Representative image of experiments performed in triplicate. Radioactive bands seen in all lanes including the negative controls are common to the assay and stem from the radiolabelled nucleotide.

3.4 Mutation of proposed N-terminal active site residues and protein purification

3.4.1 Mutation of possible catalytically relevant active site residues determined by sequence alignment

Since an unexpected uridylation activity of the N-terminal domains of TUT4 was discovered, a sequence alignment was performed to determine which residues align with the C-terminal catalytic aspartate residues. Within the N-terminal Ntr domain, aspartate 412 (D412), asparagine 414 (N414), and arginine 468 (R468) are proposed to align with the catalytic triad (Figure 19). Interestingly, just downstream of N414 is an aspartate residue (D416), which we believe could potentially be involved in the uridylation activity.

To determine if these residues are in fact contributing to TUT4-N uridylation activity, site-directed mutagenesis was performed, as described in Chapter 2, to mutate these residues to alanine. In the event that D412, N414, D416, or R468 are important for activity, alanine mutations should significantly decrease or eliminate the catalytic activity. These mutants will be referred to as D412A, N414A, D416A, and R468A from here on.

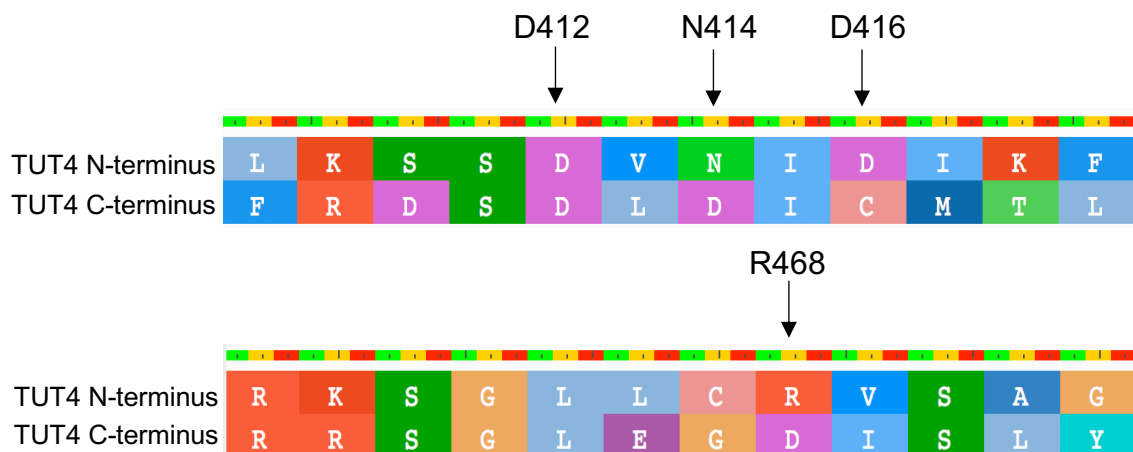


Figure 19. Sequence alignment of TUT4 N- and C-termini to identify catalytic residues of the N-terminal Ntr domain. The N-terminal catalytic region was aligned with the C-terminal catalytic region of TUT4 to determine which residues are in homologous position to the active site catalytic triad. N-terminal D412, N414, and R468 align with the C-terminal aspartates. N-terminal D416 is in close proximity to these residues and is of interest.

3.4.2 Purification of mutants with the similar purity as WT

Once mutation of the N-terminal active site residues to alanine was confirmed by sequencing, all 4 mutants were transformed into BL21 C+ *E. coli* cells. Cells were grown to an OD600 of 0.6, followed by overnight induction with 250 μ M IPTG, the same as the TUT4-N protein, as described in Chapter 2. Cells were lysed using French Press and soluble proteins were isolated and applied to the GST affinity column. Elutions containing the mutant protein were then applied to the Ni-NTA affinity column, following the same purification process as the TUT4-N protein. Purified proteins were concentrated before dialyzing overnight to remove imidazole. Samples from each step of the purification process were visualized by SDS-PAGE to ensure purification similar to the TUT4-N protein (Figure 20A). After dialysis, the mutant proteins were further concentrated, and the proteins were separated via SDS-PAGE, along with TUT4-N. All proteins were purified to the same homogeneity (Figure 20B & C) and concentrated. Final protein concentrations were determined by Bradford assay, with concentrations of 1.2 mg/mL (D412A), 0.77 mg/mL (N414A), 0.95 mg/mL (D416A), and 0.7 mg/mL (R468A). It is to be noted that while the yields for D412A, N414A, and D416A were similar to WT (0.2 mg/L *E. coli* culture), R468A showed a significantly lower expression rate at 0.03 mg/L *E. coli* culture. Western blot analysis of D412A, N414A, and D416A was performed to verify the presence of TUT4 protein (Figure 21). A TUT4-specific antibody verified the purification of the TUT4 mutants, as evident by a band at a relative molecular mass of 105 kDa.

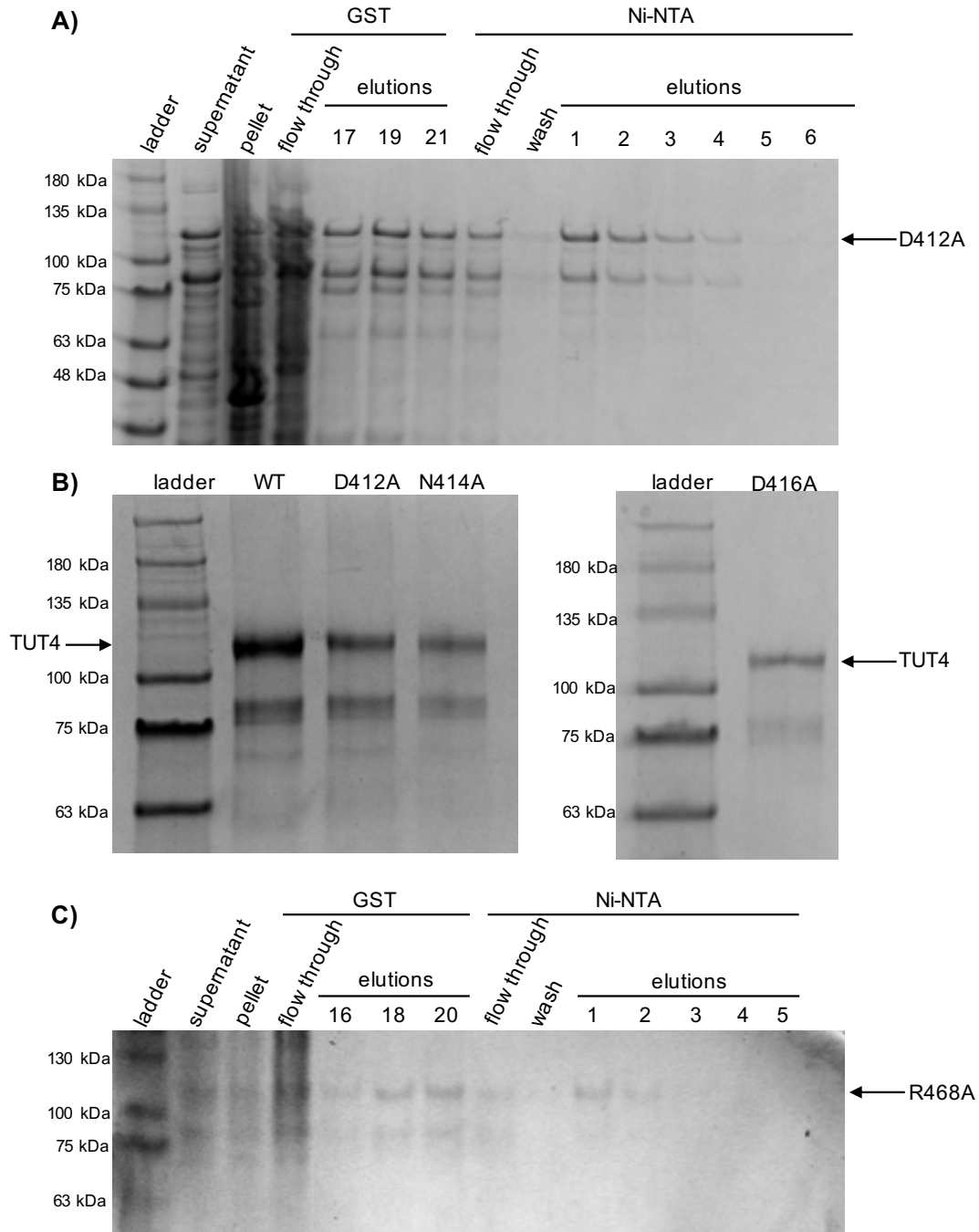


Figure 20. Purification of TUT4-N mutants. A) SDS-PAGE of the purification of mutant protein D412A. **B)** SDS-PAGE of concentrated TUT4 mutant proteins compared to purified concentrated TUT4-N WT protein. TUT4-N WT and mutants have a relative molecular mass of 105 kDa. **C)** SDS-PAGE of mutant protein R468A. R468A had significantly lower expression of 0.03 mg/L compared to WT TUT4-N.

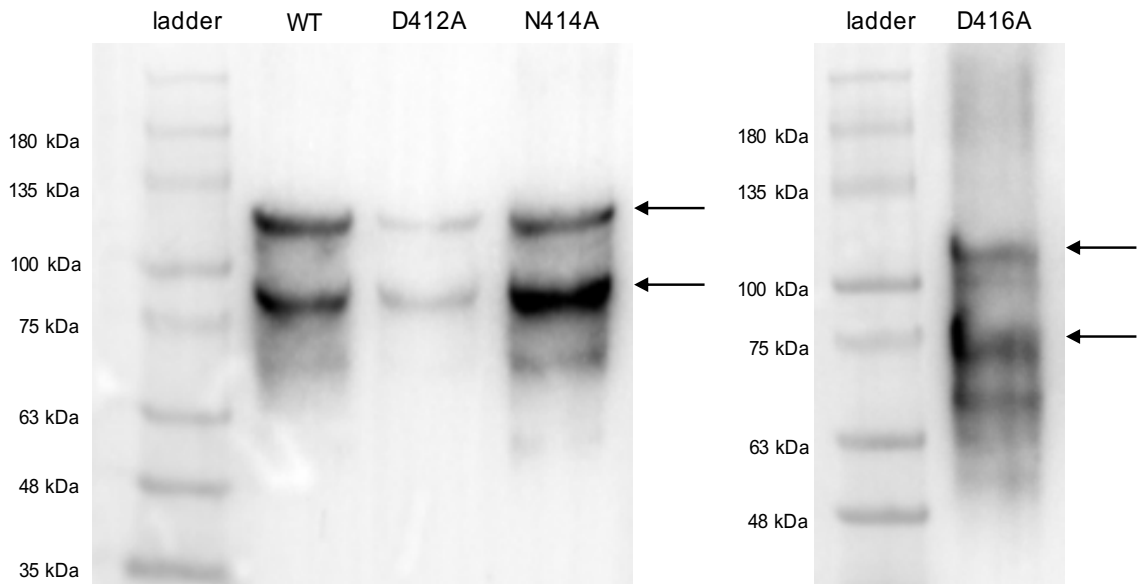


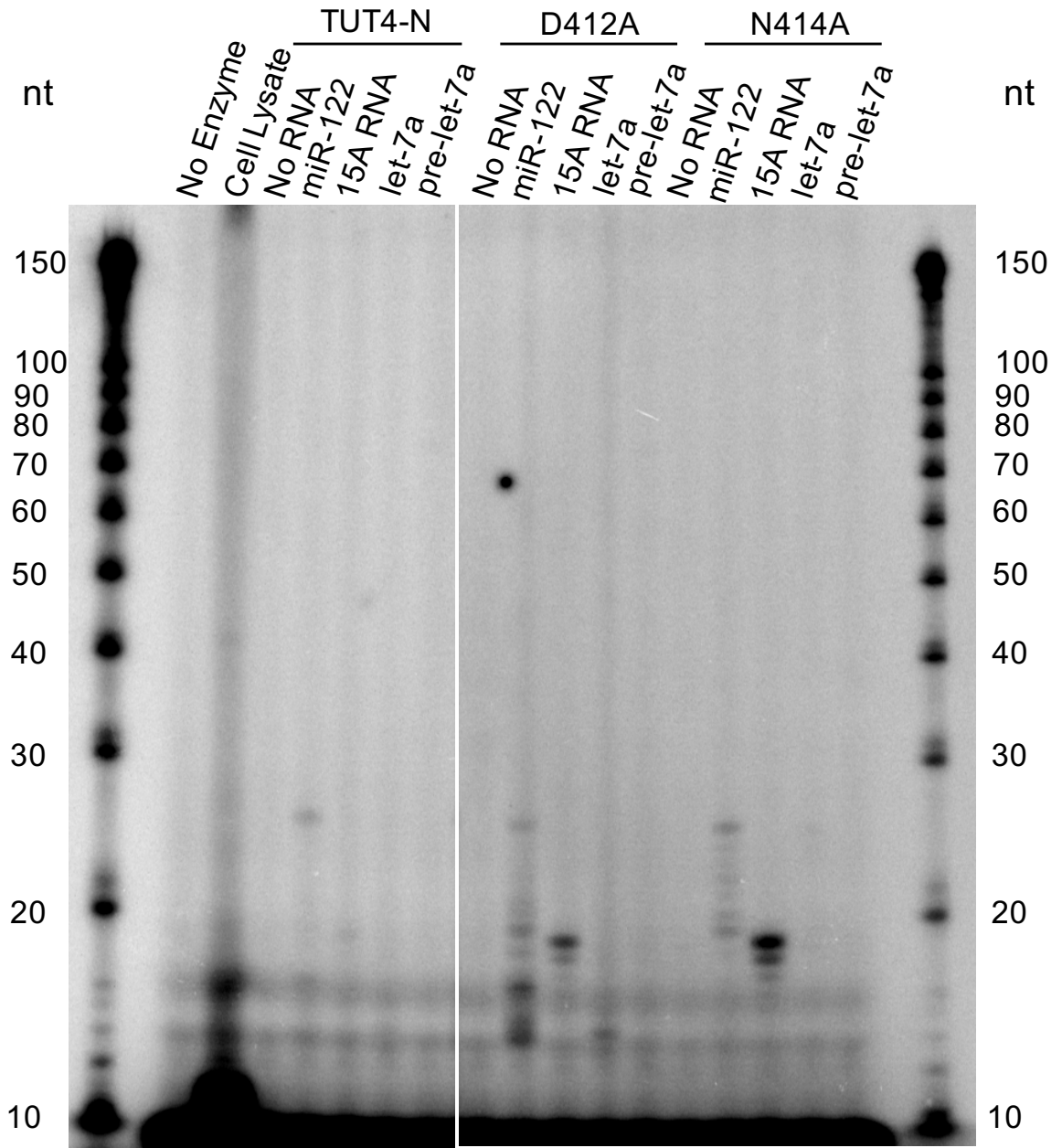
Figure 21. Verification of TUT4-N mutants by Western blot. Western blot analysis using a TUT4-specific antibody to verify the presence of TUT4-N and mutants. A band at a relative molecular mass of 105 kDa is visible for all proteins (indicated by upper arrows). A second distinct band is seen at 80 kDa (indicated by lower arrows).

3.5 TUT4 mutants are catalytically active

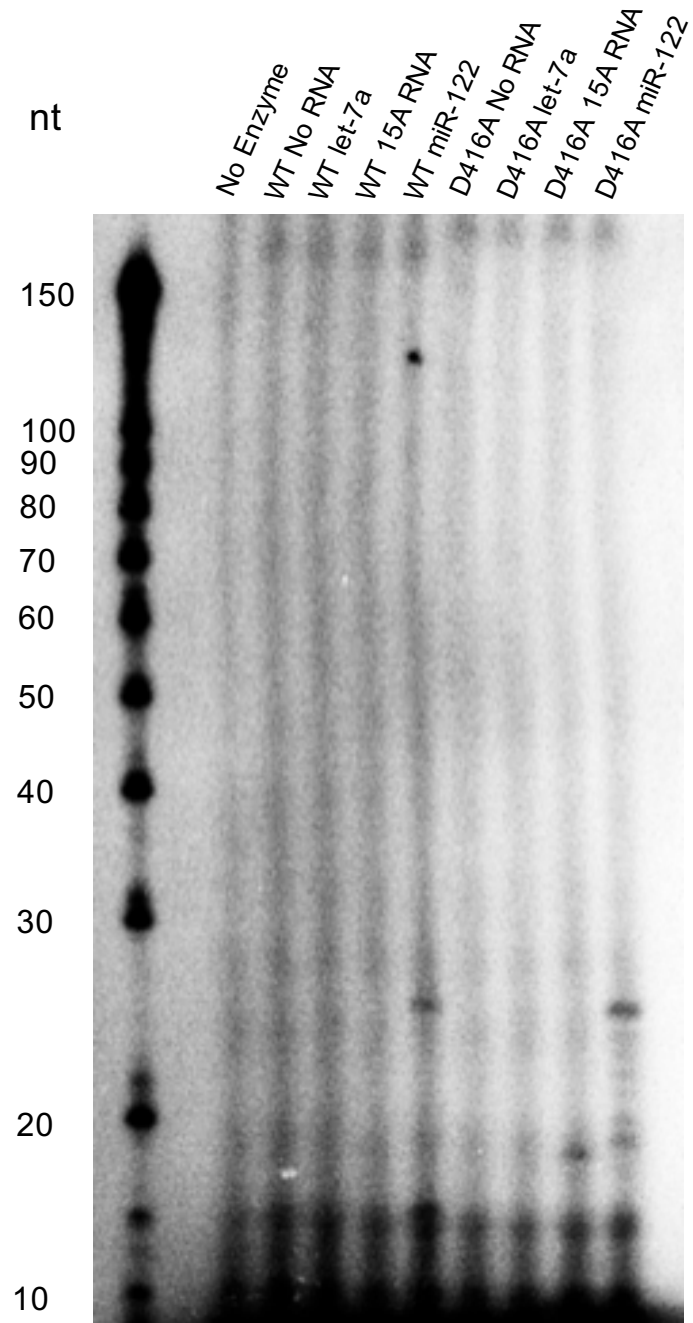
After purification of the TUT4 mutants (D412A, N414A, D416A, and R468A), radioactive activity assays were performed to determine which residues were critical for the uridylation activity. Proteins were incubated with [α - 32 P]-UTP and different RNA substrates in a 30-min. end-point reaction, as described in Chapter 2. The reaction products were separated on a 12% polyacrylamide gel and the radioactive signal was imaged on the Storm scanner. 'No enzyme' and 'no RNA' reactions were used as negative controls, while reactions with TUT4-N were used as positive controls.

Surprisingly, all four mutants displayed uridylation activity similar to the TUT4-N protein. Interestingly, the mutant proteins, as well as the WT protein, were not active with all RNA substrates but displayed RNA specificity for miR-122 and 15A RNA (Figure 22). Uridylation activity is distinct for D412A and N414A with 15A RNA and miR-122, but no activity is seen with let-7a or pre-let-7a (Figure 22A). Since no activity is seen with WT or the D412A and N414A mutants for let-7a and pre-let-7a, reactions with these RNAs were excluded for D416A and R468A mutant proteins (Figure 22B & C). Interestingly, it seems that the D412A and N414A mutants show an increase in miR-122 uridylation of degraded miR-122 substrate, as seen by a ladder of bands under the nucleotide length of the full-length miRNA (22 nts).

A)



B)



C)

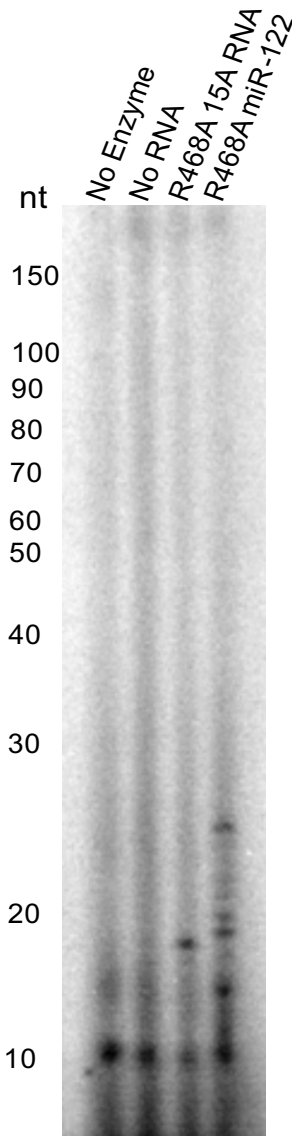


Figure 22. TUT4-N mutants display RNA-specific uridylation activity. A)

Uridylation activity of D412A and N414A. Distinct bands of uridylation activity can be seen with 15A RNA (15 nts) and miR-122 (22 nts). **B)** Uridylation activity of D416A with 15A RNA and miR-122. **C)** Uridylation activity of R468A with 15A RNA and miR-122. TUT4-N was used as a positive control, while 'no enzyme' and 'no RNA' were used as negative controls. Radioactive bands seen in all lanes including the negative controls are common to the assay and stem from the radiolabelled nucleotide.

3.6 Mutating N-terminal active site residues has no significant effect on the activity of TUT4-N

3.6.1 Cid1 adds nucleotides in a time dependent manner

To ensure that the time course activity assay, as described in Chapter 2, is an accurate means to assess nucleotide addition, Cid1 was used as a positive control. Since Cid1 is more active than the TUT4-N variants, [α - 32 P]-UTP was diluted by a factor of 24 with unlabelled UTP to a final concentration of 1 μ M total UTP. The radioactive assay was performed using 15A RNA, with the reaction being stopped at time points of 1, 2, 5, 10, and 20 min. The reactions were separated on a 12% polyacrylamide gel, and the intensities of the radioactive signals were quantified. Serial dilutions of radiolabelled UTP were spotted on Whatman paper and exposed with the gel to create a standard curve of the intensities for quantification of the uridylated RNA product. Once plotted on a graph (Figure 23), it is evident that Cid1 activity plateaus after 10 minutes at an average product formation concentration of 52 ± 10 nM. The time course assay was repeated in triplicate and standard error was calculated. Initial velocity was calculated to determine the activity of Cid1, with an average of 0.31 ± 0.1 nM/sec.

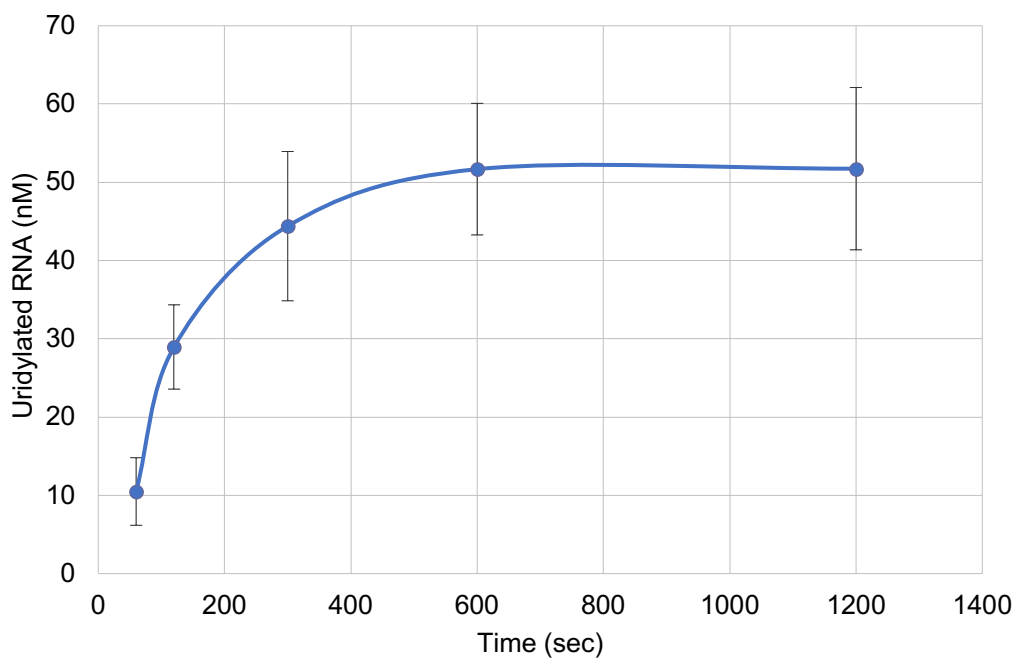


Figure 23. Quantification of uridylated product with Cid1. Commercial PUP Cid1 was used as a positive control for the quantitative radioactive time course assay. Cid1 is most active within the first 10 min., with its activity plateauing at approximately 52 ± 10 nM of uridylated product formation. The error represents one standard error.

3.6.2 Uridylation activity of TUT4-N and mutants is time dependent and some mutants display altered activity.

To determine if there is a significant change in the activity of the TUT4-N mutants compared to WT protein, a radioactive time course assay was performed with 15A RNA, as described in Chapter 2. This assay measures the product formation of uridylated RNA at different time points to quantify the activity of the TUT4-N enzymes. Since the activity of TUT4-N and mutants is much lower than Cid1, the [α - 32 P]-UTP was not diluted or supplemented with unlabelled UTP. Instead, the radioactive end-point activity assay reactions were scaled up 2-fold to ensure that the reactions could be stopped at several time points for the most accurate quantification. 12 μ g of protein was added to the reaction at a final concentration of 0.6 mg/mL for TUT4-N and mutants. The reactions were stopped at time points of 2, 5, 10, 20, 30, and 60 min. For D412A, the 2-min. time point was omitted due to a limited amount of protein. The time course assay was repeated in a technical triplicate for all TUT4-N proteins and the standard error was calculated. The product formation was plotted against time to show the linear increase of the product formation for the TUT4-N and mutant proteins (Figure 24).

The TUT4-N and mutant proteins are catalytically inefficient, with approximately 200,000-fold lower enzymatic activity compared to Cid1. Product formation continued to be linear at 60 min. The TUT4-N protein had a quantified average final product concentration of 6 ± 0.5 pM. The D412A and N414A mutants had higher average final product formation of 11 ± 0.1 pM and 19 ± 5

pM, respectively. On the other hand, the D416A mutant displayed reduced activity, with a final product formation of 1.5 ± 0.2 pM. Although the activity was not completely abolished, the product formation of D416A decreased by approximately 5-fold compared to TUT4-N.

TUT4-N variants uridylylate 15A RNA

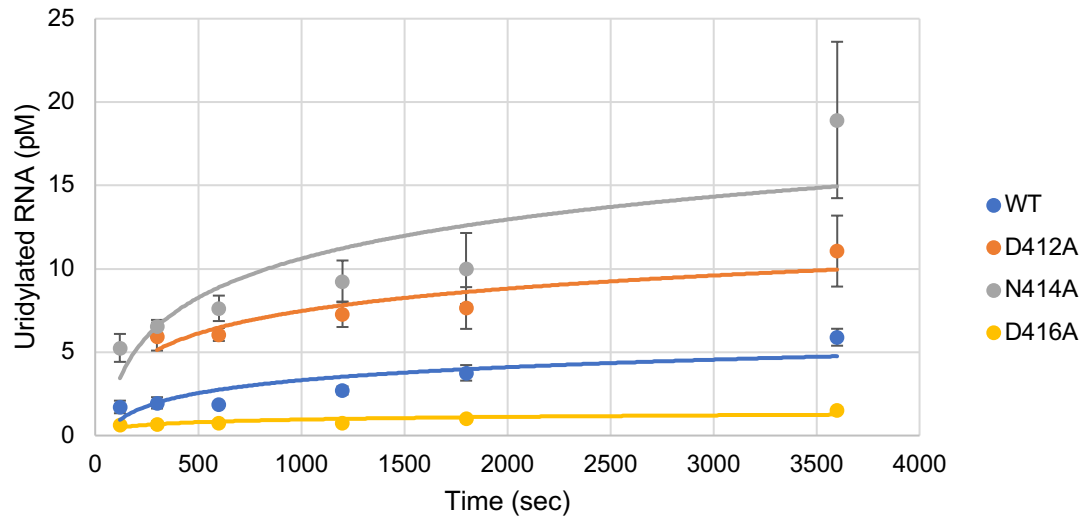


Figure 24. Product formation of uridylylated 15A RNA by TUT4-N proteins.

The product formation of uridylylated 15A RNA was quantified at different time points. The product concentration was plotted over time to compare the increase in uridylation activity of the TUT4 mutants with TUT4-N. D412A and N414A have approximately 2-fold and 3-fold increased product formation, respectively, while D416A has an approximately 5-fold decrease in product formation compared to TUT4-N. The error represents one standard error.

3.6.3 Alanine mutation of proposed catalytic residues has little effect on catalytic activity

Since there was a change in the product formation of uridylated 15A RNA when comparing TUT4-N to the mutants, the initial velocities of the reactions were calculated to determine if there is a significant difference in the activity of the TUT4-N variants. The initial velocities were determined by calculating the slope of the linear trendline for each protein (Table 3). Although there is an increase in the initial velocities of D412A and N414A (0.0016 ± 0.00012 pM/sec and 0.0037 ± 0.0013 pM/sec, respectively) compared to TUT4-N (0.0012 ± 0.00013 pM/sec), there is no significant difference between these mutants and the WT within error (p-value >0.05). D416A, with an initial velocity of 0.0002 ± 0.00003 pM/sec, is significantly lower than TUT4-N, with a 5-fold decrease in enzyme activity (p-value=0.009). The difference in initial velocities was plotted on a bar graph as a visual representation (Figure 25).

Using the initial velocities and concentration of protein, the specific activities of TUT4-N and the mutants at $0.165 \mu\text{M}$ UTP was determined. TUT4-N, with a specific activity of 0.10 ± 0.01 pM/sec/mg, is not significantly different in activity from the D412A (0.13 ± 0.005 pM/sec/mg) or N414A (0.31 ± 0.12 pM/sec/mg) mutants (p-value >0.05). The D416A mutant, with a specific activity of 0.02 ± 0.003 pM/sec/mg, is significantly less active than the TUT4-N protein (p-value= 0.009).

Table 3. Initial velocities and specific activities of TUT4-N and mutants.

Mutants D412A and N414A do not significantly differ in initial velocity (V_0) or specific activity compared to TUT4-N WT protein (p-value >0.05). D416A is 5-fold less active than TUT4-N (p-value=0.009).

TUT4-N Construct	V_0 (pM/sec)	Specific Activity (pM/sec/mg)
WT	0.0012 ± 0.00013	0.10 ± 0.01
D412A	0.0016 ± 0.00012	0.13 ± 0.005
N414A	0.0037 ± 0.0013	0.31 ± 0.12
D416A	0.0002 ± 0.00003	0.02 ± 0.003

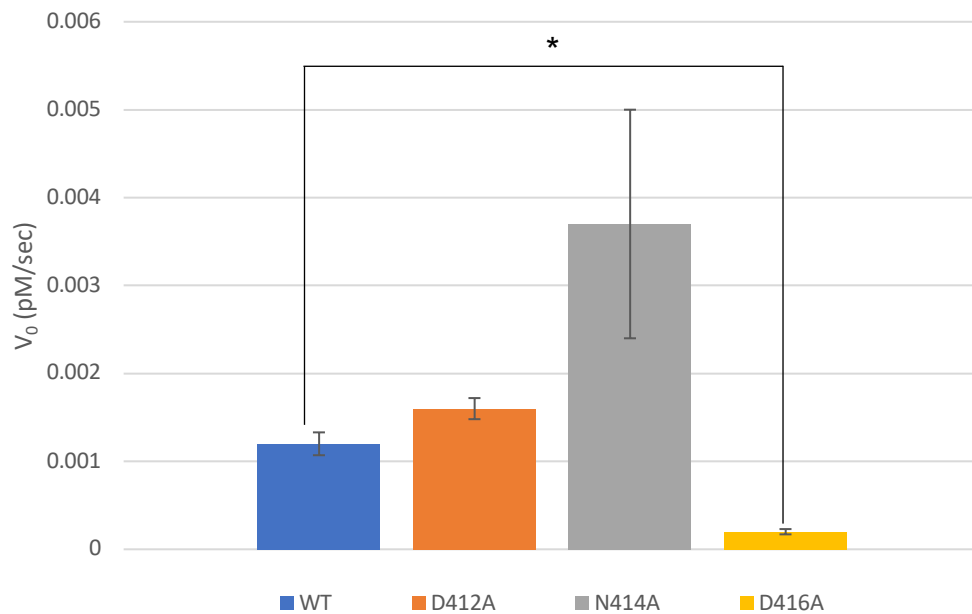


Figure 25. Visual representation of difference of initial velocities of TUT4-N and mutants. D412A and N414A show increased initial velocities compared to TUT4-N. Within error, there is no significant increase for either mutant (p-value >0.05). D416A displayed a significant 5-fold decrease in initial velocity compare to TUT4-N (p-value= 0.009). Initial velocities were determined for each trial and standard error was calculated. Error bars represent one standard error of experiments performed in triplicate. Significance is indicated by asterisk (*).

3.7 TUT7 also displays N-terminal uridylation activity

TUT7, a human homolog believed to be functionally redundant to TUT4⁸,^{65, 68, 69, 70}, has a very similar domain layout and shares 40% identity with TUT4. Similar to TUT4, TUT7 encodes two catalytic regions, one in the C-terminus believed to be responsible for its catalytic activity, and a second one in the N-terminus, believed to be catalytically inactive. Low conservation between the N- and C- terminus suggest that the duplication of the domains occurred early during evolution. The N-terminal domains may be less conserved because its catalytic function is less relevant for its biological function. The N-terminal Ntr domain of TUT7 is also lacking an identifiable catalytic aspartate triad (Figure 11). Like the C-terminal Ntr domains of TUT4 and TUT7, the N-terminal domains also share identity with one another. Because of its similarities to TUT4, we also determined if the N-terminal region of TUT7 was catalytically active. The N-terminus of TUT7 (amino acids 308 to 476) from human TUT7 (TUT7-N) and the C-terminal catalytic region (amino acids 1,017 to 1,282) of TUT7 from mouse TUT7 (TUT7-C) was cloned and purified by Sarah Lee (4th year Honours student, Heinemann lab). The human TUT7 plasmid HsCD00082112 from the Harvard PlasmID Database is lacking the C-terminal Ntr domain. Thus, mouse TUT7 (Addgene #60044) was used to clone the TUT7-C construct. Considering the impurity of the TUT7-N protein, I quantified the percentage of TUT7-N in the total protein using Image Lab. TUT7-N constitutes approximately 10% of the total protein concentration.

Using the purified TUT7 proteins (Figure 26), I performed the radioactive time course assay experiments on both the N-terminal and C-terminal TUT7 proteins with 15A RNA (Figure 27). For TUT7-N, 12 μg of protein was added to the reaction at a final concentration of 0.6 mg/mL, the same as in the TUT4-N activity assays. For TUT7-C, 1 μg of total protein was added at a final concentration of 0.05 mg/mL. When imaging the time course assay, Whatman paper with serial dilutions of radiolabelled UTP were spotted to allow for the quantification of product formation. Technical triplicates were performed for both proteins and standard error was calculated for each time point. The product formation was calculated at each time point and plotted over time to visualize the linear increase of product formation over 60 min. (Figure 28). TUT7-N protein had a final product formation of 13 ± 0.7 pM after 60 min. (Figure 28A), while for TUT7-C, 12-fold less enzyme produced a final product formation of 15 ± 3 pM after 60 min. (Figure 28B). The C-terminal catalytic region of TUT7 is significantly more active than the N-terminal catalytic region of TUT7.

Initial velocities of TUT7-N and TUT7-C were determined (Table 4). TUT7-N, at a concentration 12-fold greater than the TUT7-C protein, had a slightly lower initial velocity (0.0017 ± 0.0002 pM/sec), within error, compared to the TUT7-C protein (0.0028 ± 0.0009 pM/sec). When comparing their specific activity however, TUT7-C has a significantly higher activity (2.8 ± 0.9 pM/sec/mg) than the TUT7-N protein (0.14 ± 0.01 pM/sec/mg) (Table 5). The TUT7-C protein is 20-fold more active than the TUT7-N protein. Due to the impurity of TUT7-N, the activity of the N-terminal domains of TUT7 is likely greater than what was calculated.

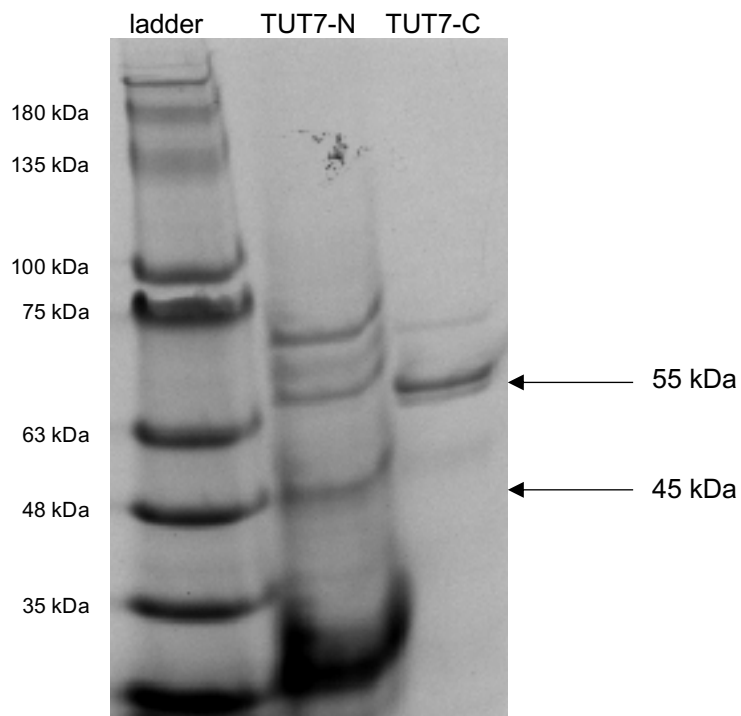


Figure 26. SDS-PAGE of concentrated TUT7-N and TUT7-C proteins. TUT7-N and TUT7-C (purified by Sarah Lee) were concentrated after purification. The relative molecular mass of TUT7-N is calculated as 45 kDa, and a band is visible at that mass, indicated by an arrow. The relative molecular mass of TUT7-C is calculated as 55 kDa, indicated by an arrow.

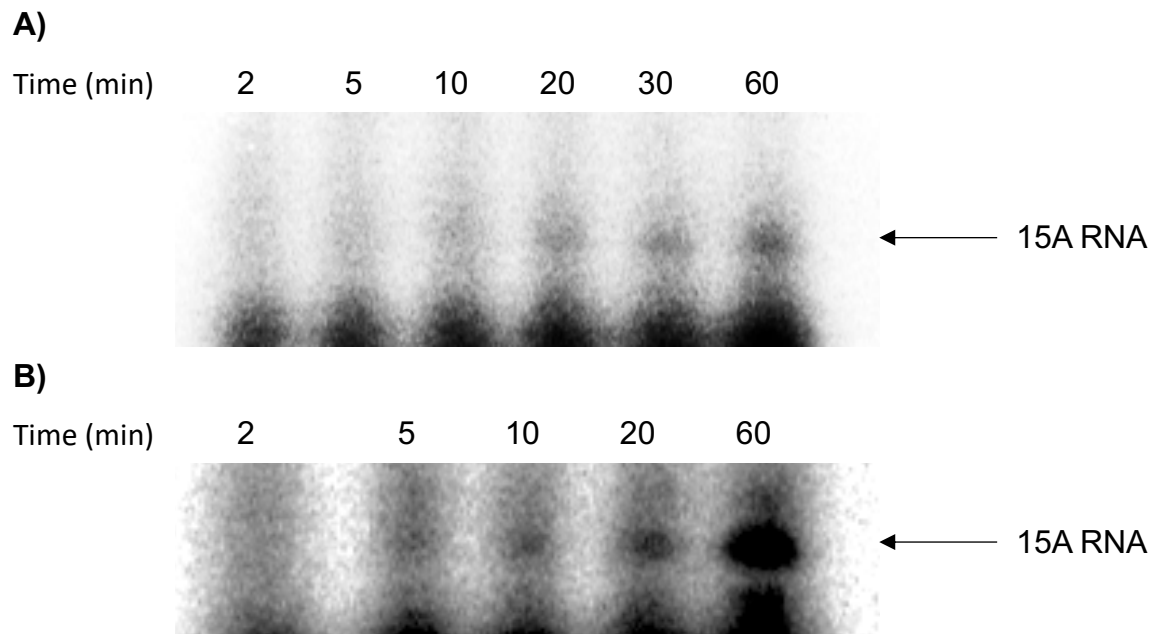
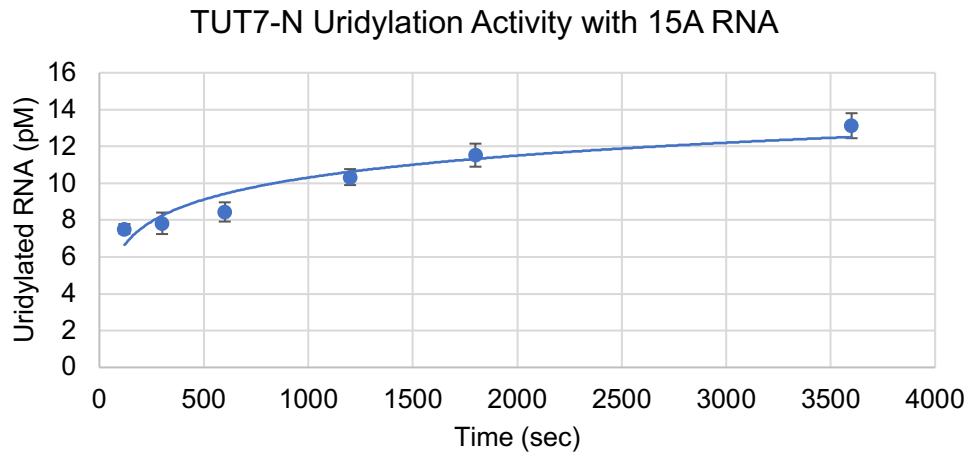


Figure 27. Time course activity of TUT7-N and TUT7-C. A) TUT7-N activity with 15A RNA over 60 min. **B)** TUT7-C activity with 15A RNA over 60 min. 15A RNA (15 nts) is indicated by an arrow.

A)



B)

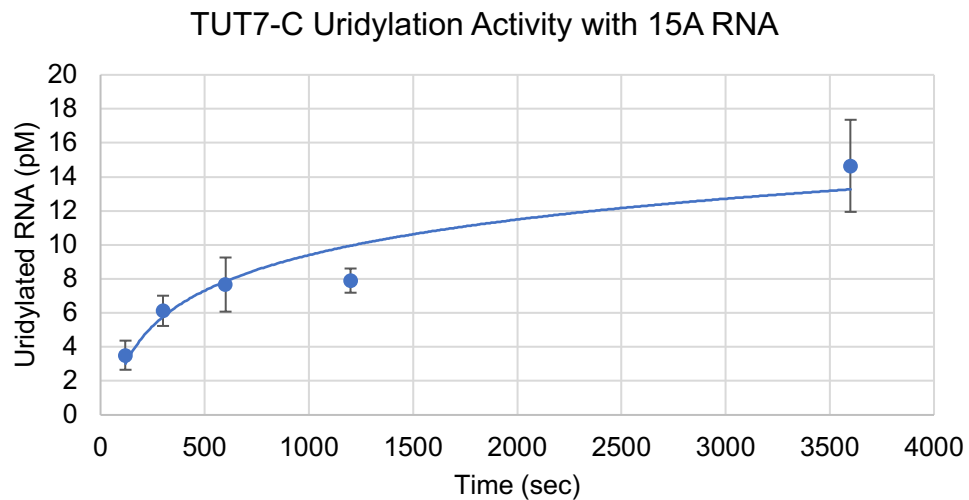


Figure 28. Product formation of uridylated 15A RNA by TUT7. The product formation of uridylated 15A RNA was quantified at different time points for **A)** TUT7-N and **B)** TUT7-C. The product concentration (y-axis) was plotted over time (x-axis) to visualize the increase in product formation over 60 min. TUT7-N (12 μg) had a final product concentration of 13 pM while TUT7-C (1 μg) had a final product concentration of 15 pM with 12-fold less enzyme. Standard error was calculated for each time point. Error bars represent one standard error.

Table 4. Initial velocities of TUT7 N- and C-terminal catalytic domains.

TUT7-N with 12-fold more enzyme (12 μg), has an initial velocity (V_0) similar to TUT7-C (1 μg), within error.

Construct	V_0 (pM/sec)
TUT7-N (12 μg)	0.0017 ± 0.0002
TUT7-C (1 μg)	0.0028 ± 0.0009

Table 5. Specific activities of TUT7-N and TUT7-C proteins. TUT7-N is significantly less efficient than TUT7-C (p-value= 0.04).

Construct	Specific Activity (pM/sec/mg)
TUT7-N	0.14 ± 0.01
TUT7-C	2.8 ± 0.9

Chapter 4: Discussion

4.1 The TUT4 N-terminal domain is catalytically active

The aim of this study was the characterization of a potential active site in the N-terminal domain of the human uridylyltransferase TUT4. For characterization of the N-terminal domains of TUT4 (TUT4-N), I used cloning techniques, described in Chapter 2, to clone the DNA sequence encoding amino acids 1 to 678 of TUT4 into a pGEX-6p2 vector with an N-terminal GST tag. Once the ligation of *Tut4-N* into the pGEX-6p2 vector was successful, a C-terminal His₆ tag was incorporated into the vector between the coding sequence and the stop codon by site-directed mutagenesis. This construct, termed TUT4-N, was transformed into BL21 C+ *E. coli* cells for bacterial expression. TUT4-N protein expression produced a purified protein yield of 0.19 mg/L; Bacterial expression less than 1 mg/L is considered low¹⁰⁵. This low expression of TUT4-N could be a result of the mRNA secondary structure. mRNA secondary structure varies sequence to sequence, which in turn affects the stability of the mRNA^{106, 107}. mRNA stability influences translation, including the rate, as well as the expression level of the protein product¹⁰⁷. Another possible effect on TUT4 protein expression is codon usage^{106, 108}. BL21 C+ *E. coli* cells contain a plasmid encoding 3 tRNAs (R, I, L) that are rare in bacterial cells, to increase protein expression. The encoded tRNAs that are required for human protein expression may not be produced at a high enough efficiency in the bacterial cell if the TUT4-N protein has a high human codon bias. Faehnle et al. produced recombinant

mouse TUT4 in insect cells, with yields of 15 mg/mL of concentrated protein³². Insect cells and expression using the baculovirus expression system, which was used by Faehnle et al., is capable of producing high amounts of recombinant proteins with the potential to include protein modifications due to its ability to efficiently process eukaryotic proteins¹⁰⁹. Other members of the pol β family have been purified in high yields for crystallization. Cid1 purification for crystallization experiments yielded 5-10 mg/mL of concentrated protein using BL21 DE3 or BL21 Star *E. coli* bacterial expression^{29, 30}; Gld2 was purified to a protein concentration of 8-15 mg/mL for crystallization experiments¹¹⁰. Using different bacterial cells more suitable for the expression of TUT4-N (e.g. BL21 Star which increases mRNA stability for increased expression), strains optimized to encode tRNAs for codons more common in eukaryotic mRNAs, or a eukaryotic expression system to allow for required eukaryotic translation modifications (e.g. human or insect cells) could increase TUT4-N protein yield.

TUT4-N was purified using affinity chromatography. After both GST and Ni-NTA affinity chromatography, the protein was purified to >85%. Two distinct bands were seen, both containing TUT4-N, verified by western blot. The 105 kDa protein band corresponds to the full-length TUT4-N protein, while the lower 80 kDa protein band is likely a degradation product due to protein instability. Cleavage within the protein sequence could be occurring in the linker regions between the domains. Other members of the pol β family of enzymes have also shown instability when recombinantly purified, and full-length TUT4 is known to be difficult to purify^{32, 111}. Purification of other TUTases has resulted in the presence of other bands after affinity purification. *Leishmania tarentolae* (*L.*

tarentolae) TUTase purification resulted in a single major band, as well as the presence of 2 minor bands after several affinity, size-exchange, and ion exchange columns¹¹¹. Further purification of TUT4-N using ion exchange or size-exchange chromatography would have resulted in a loss of protein, with the possibility of further degradation.

During the purification process, it was noticed that TUT4-N protein was being lost during both GST and Ni-NTA affinity chromatography in the flow through when the sample was being applied to the column. We speculate that this is occurring due to misfolding of the protein and aggregation. Aggregation of misfolded protein could be blocking the GST and His₆ tags, preventing binding to the glutathione and Ni-NTA beads during purification. This could also be contributing to the low levels of protein being purified. Protein misfolding and aggregation could be a result of TUT4-N being a partial protein and low stability of the protein. Also, TUT4 is a protein that has a relatively high cysteine content. Partially or fully unfolded proteins could have unpaired cysteines, which is promoting aggregation.

Using the purified TUT4-N protein, radioactive activity assays were used to elucidate the uridylation activity of the TUT4 N-terminal domains. The purified enzyme was incubated with RNA substrate and [α -³²P]-UTP in an end-point activity reaction. The N-terminal domains of TUT4 displayed uridylation activity *in vitro*. As evident in Figure 16, TUT4-N had uridylation activity when incubated with 15A RNA (15 nts), an mRNA poly(A) tail mimic, and miR-122 (22 nts), an miRNA. Although active, 6 μ g of TUT4-N protein (0.6 mg/mL) was needed in the

reaction to see activity after a 30-minute incubation.

To further analyze the activity of TUT4-N, a radioactive time course assay was performed. TUT4-N had a quantified average final product concentration of 6 ± 0.5 pM of uridylated 15A RNA after 60 minutes. The initial velocity of TUT4-N, determined by the slope of the linear increase of product formation, was calculated with an activity of 0.0012 ± 0.00013 pM/sec.

Compared to previously published kinetic data of Gld2⁴³ and Cid1³⁰, it is evident that TUT4-N is less active than the other Ntr domains containing the aspartate triad. Gld2, a poly(A) polymerase, had a V_{max} of $(2.93 \pm 0.16) \times 10^{-6}$ $\mu\text{M}/\text{sec}$ with ATP and $(24.8 \pm 1.17) \times 10^{-6}$ $\mu\text{M}/\text{sec}$ with UTP⁴³. When comparing these values to the activity of TUT4-N with UTP, the Gld2 enzyme is approximately 20,000-fold more active than the N-terminal catalytic domains of TUT4. Although the activity data presented here do not represent a full kinetic analysis, we can assume that TUT4-N is significantly less active than other TUTases. Cid1³⁰, the yeast homolog, has a catalytic efficiency of $6.6 \times 10^2 \text{ M}^{-1} \text{ sec}^{-1}$, while *Trypanosoma brucei* TUT4¹¹² (*TbTUT4*) has a catalytic efficiency of $1.5 \times 10^6 \text{ M}^{-1} \text{ min}^{-1}$. While the k_{cat}/K_M ratios for Cid1 and *TbTUT4* cannot be directly compared to our TUT4-N specific activity¹¹³, they give us an indication that TUT4-N is approximately 10^8 -fold lower in activity.

The activity of TUT4-N may be significantly lower than other homologs due to it being a partial protein. All of the domains may be required for the N-terminal catalytic domains to achieve full potential activity. Also, the C-terminal catalytic domains, when part of the full-length protein, may be the main active region, with the N-terminal catalytic domains providing a potential secondary binding and

modification site for additional substrates. Thus, investigation of the activity of the N-terminal catalytic region in the context of the full-length protein, e.g. by mutating the aspartate triad in the C-terminal active site, may result in a more stable, and potentially more active N-terminal catalytic region. Another potential reason for the low activity of TUT4-N could be the presence of the 80 kDa degradation product. This protein is a possible degradation product of TUT4-N, determined by Western blot. This smaller protein product may partially inhibit the full-length TUT4-N protein's ability to catalyze the uridylation of RNA substrate.

From the activity assays, conclusions of processivity cannot be made. Processive nucleotide addition activity in gel-based assays manifests in a ladder-like product formation. This ladder of nucleotide addition is not observed for TUT4-N with both 15A RNA or miR-122. Only a single band is visible for both RNA substrates. From the single band for both miR-122 (approximately 23 nts) and 15A RNA (approximately 16 nts) we can speculate that TUT4-N monouridylates its RNA substrates or is only capable of adding a specific, short poly(U) tail consisting of 2 or 3 uridine residues. The exact number of nucleotides added will have to be determined by methods that include sequencing, such as circularized rapid amplification of cDNA ends (cRACE). Full-length TUT4, and therefore the TUT4 C-terminal catalytic domains, has been shown to polyuridylate mRNA and pre-miRNA both *in vitro* and *in vivo*, leading to degradation^{27, 65, 70}. *In vitro*, TUT4 has been shown to have an average poly(U) tail length of 6 uridine residues after 10 minutes on RNA with a poly(A) tail shorter than 20 adenine residues⁶⁵. My data shows that TUT4-N is less processive, with only 1 or 2 uridine residues being added to 15A RNA, which mimics an mRNA

poly(A) tail. Monouridylation of mature miRNAs by TUT4 has been shown to silence miRNA activity⁶⁷. In the case that a single uridine is being added to miR-122 by TUT4-N, TUT4-N could be important in the silencing of miR-122 and mRNA stability. Without a quantitative answer on poly(U) tail length, TUT4-N function cannot be concluded. Further studies have to be performed to determine the number of uridines being added to both miR-122 and 15A RNA.

4.1.1 E. coli cell extracts do not exhibit uridylyltransferase activity.

Due to the presence of contaminating bacterial protein in my TUT4-N purification, a mock purification was performed using the empty pGEX-6p2 vector with an inserted C-terminal His₆ tag to exclude the possibility that observed activity was from contaminating *E. coli* proteins.

GST has a high expression compared to TUT4-N. The Mock protein yielded purified protein product of a concentration of 7.32 mg/mL. Western blot analysis using the TUT4-specific antibody verified that there was no TUT4 present in the Mock protein sample.

E. coli itself does not encode any known TUTases, and uridylyltransferase activity of RNA post-transcriptionally appears limited to eukaryotes^{8, 25, 41, 57}. Regardless of this, uridylyltransferase activity in eukaryotes was only discovered a few years ago^{25, 41, 67, 114}, and it cannot be excluded that the annotated adenylyltransferase PAP1 in *E. coli* may exhibit residual uridylyltransferase activity¹¹⁵. Other adenylyltransferases have previously shown nucleotide promiscuity^{16, 25, 43, 116, 117}. Thus, a mock purification with an empty vector control was performed identical to the purification of the TUT4 proteins. Figure 18 shows

that even with high amounts of Mock protein (up to 50 μg), no activity was observed with 15A RNA or miR-122, confirming that the N-terminal catalytic region is in fact capable of uridylation activity, and no contaminating bacterial protein is causing the uridylation activity observed with 15A RNA or miR-122.

Overall, as expected, no uridylyltransferase activity was observed from *E. coli* cell extracts after following the mock purification protocol. This data confirms that the observed TUTase activity stems from the recombinant purified TUT4-N protein.

4.2 TUT7 N-terminus also shows activity similar to TUT4

Similar to TUT4, the human homolog TUT7 also encodes two potential active sites^{32, 73} (Figure 5). Sarah Lee, a 4th year Honours thesis student, purified the N-terminal catalytic domains of TUT7 (TUT7-N), as well as a construct containing only the TUT7 C-terminal catalytic domains (TUT7-C). During her time in the lab, Sarah determined that the TUT7-N protein was also active and capable of uridylation activity.

Using the TUT7 purified proteins, I performed a time course assay for both the TUT7-N and TUT7-C proteins to quantify their activity with 15A RNA. The initial velocities of the N- and C-terminal TUT7 proteins were determined. TUT7-N protein had an initial velocity of 0.0017 ± 0.0002 pM/sec, while the TUT7-C protein, with 12-fold less total protein, had an initial velocity of 0.0028 ± 0.0009 pM/sec (Table 4). To compare TUT7-N and TUT7-C, the specific activities of the enzymes were determined (Table 5). TUT7-N, with a specific activity of $0.14 \pm$

0.01 pM/sec/mg, was 20-fold less efficient than TUT7-C, with a specific activity of 2.8 ± 0.9 pM/sec/mg (Figure 29). Through this quantification, it is evident that the C-terminal catalytic region of TUT7 is significantly more active and efficient at catalyzing the uridylation of RNA than the N-terminal catalytic region. Again, it has to be considered that the purified proteins are not full-length proteins, and that the active sites may exhibit more activity in the context of full-length proteins, which includes more RNA binding domains, and may be more stable or catalytically favourable.

The TUT7-N protein had a similar activity to the TUT4-N protein. When comparing the initial velocities, the activity of TUT4-N, with an initial velocity of 0.0012 ± 0.00013 pM/sec, was not significantly different than the activity of TUT7-N, with an initial velocity of 0.0017 ± 0.0002 pM/sec. Comparing the specific activities of these two enzymes, TUT4-N (0.10 ± 0.01 pM/sec/mg) and TUT7-N (0.14 ± 0.01 pM/sec/mg) were similar in uridylation efficiency with 15A RNA substrate (Figure 29). TUT7-N has greater impurities after purification compared to TUT4-N, with TUT7-N only constituting approximately 10% of the total protein. Because TUT7-N is under-represented, and the purities of TUT7-N and TUT4-N differ, TUT7-N could be more active than TUT4-N.

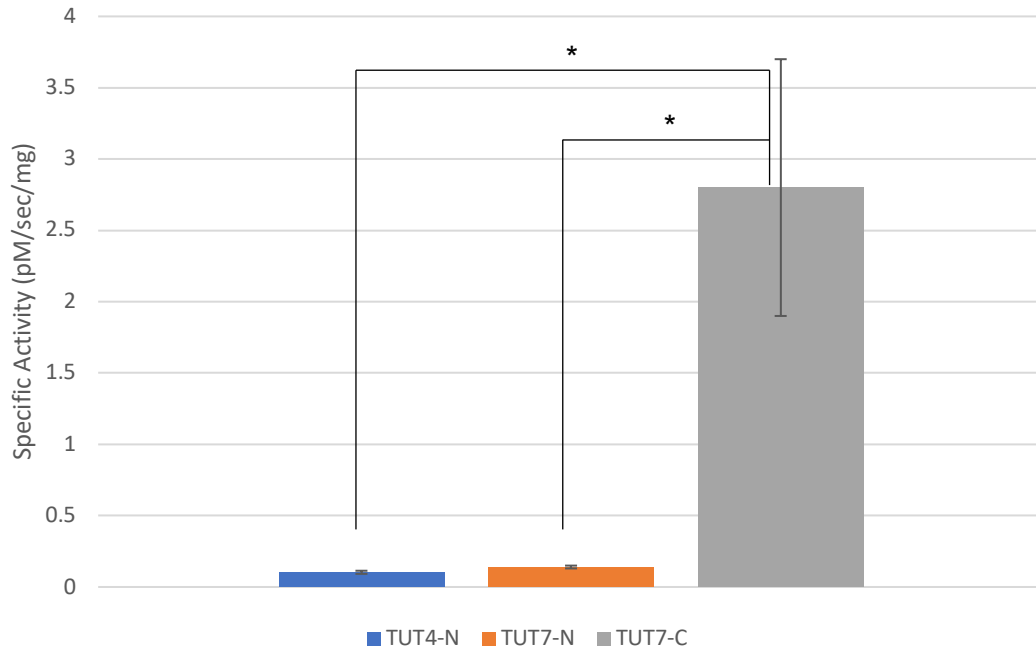


Figure 29. Specific activities of TUT4-N, TUT7-N, and TUT7-C. TUT4-N (0.10 ± 0.01 pM/sec/mg) and TUT7-N (0.14 ± 0.01 pM/sec/mg) have no significant difference in specific activity. TUT7-C (2.8 ± 0.9 pM/sec/mg) is significantly more efficient than TUT4-N and TUT7-N (p -value < 0.05), denoted by a single asterisk (*). Error bars represent one standard error.

Due to similarities in activity, the N-terminus of TUT4 and TUT7 may be similar in their roles. Both TUT4-N and TUT7-N seem to be less active than their respective C-terminal catalytic regions. This activity could change with substrate specificity. While TUT7-C is significantly more active than TUT7-N with the mRNA poly(A) tail mimic 15A RNA, as determined by the specific activity, this may be substrate specific. The activity of the N-terminal and C-terminal proteins with other RNAs, such as miR-122, will need to be determined to see if substrate specificity has an effect on activity. Similar to TUT4-N, TUT7-N and TUT7-C were less processive than full-length protein, indicating possible monouridylation or low

processivity polyuridylation (2-3 nts) (Figure 27). This polyuridylation activity differs from what has been seen for full-length TUT7, which has shown uridylation addition of greater than 3 nts⁶⁵.

Compared to other TUTases, all three proteins (TUT4-N, TUT7-N, and TUT7-C) are less active. Kinetic experiments with TUT1 and U6 snRNA determined a K_M of 55 nM of enzyme to reach half of its maximum velocity, with more than 1 pmol of uridylated RNA in 8 minutes⁷¹. The experiments shown here required a minimum of 12 μ g of TUT4-N and TUT7-N enzyme per time course reaction (0.6 mg/mL) or 1 μ g of TUT7-C per time course reaction (0.05 mg/mL) to produce pM amounts of uridylated RNA products in 60 minutes. TUT4-N, TUT7-N, and TUT7-C are all recombinant partial proteins and may require other domains from the full-length protein to achieve maximum activity. TUT1 lacking RNA binding domains has a significant decrease in uridylation activity with less than 2 fmol of uridylated RNA when 10 nM of enzyme is added⁷¹. TUT7-C, which includes only the C-terminal Ntr and PAP domains, may require the presence of RNA binding ZnFs for optimal activity, hence why the C-terminal region is less active than what is expected.

4.3 The N-terminal active site of TUT4 likely exhibits a different active site compared to the C-terminal domains

The classic active site architecture of nucleotidyltransferases includes three aspartate residues and two Mg^{2+} ions^{13, 31}. To identify residues involved in nucleotidyltransferase activity and Mg^{2+} coordination in the N-terminal TUT4 domains, I aligned the N- and C-terminal catalytic regions. In the C-terminal Ntr

domain, residues D1009, D1011, and D1070 form the catalytic triad (Figure 19). From this alignment, we were unable to identify the aspartates directly aligning to the catalytic triad in the N-terminal Ntr domain. Four amino acid residues (D412, N414, D416, and R468) were identified that may be involved in catalysis.

The four proposed residues were mutated to alanine and purified following the same purification process as the TUT4-N protein. Interestingly, one mutant, R468A expressed at much lower efficiency compared to other mutants. The protein yield for this mutant was significantly lower than the other mutants, with only 0.03 mg/L, compared to a yield of 0.19 mg/L of TUT4-N protein, or 0.12 mg/L, 0.10 mg/L, and 0.19 mg/L yield for D412A, N414A, and D416A, respectively. A single mutation can have many different effects on a protein, including folding¹¹⁸, expression, and activity¹¹⁹. In the case of R468A, it is clear that the single mutation is affecting expression. The arginine residue, protonated at physiological pH, may be involved in an important interaction with regards to proper folding. We can assume that the yield of the R468A mutant is low due to an issue with protein folding, which may lead to degradation. Therefore, due to low expression and no decrease in end-point activity, we excluded the protein from the time course assay.

Next, the enzymatic activity of these mutants was tested. Surprisingly, all four mutants were active with 15A RNA and miR-122, similar to TUT4-N (Figure 22). To further study the effect of these mutations on enzymatic activity, specific activities were determined for mutants D412A, N414A, and D416A. Compared to the TUT4-N protein, with an average product formation of 6 ± 0.5 pM of uridylated 15A RNA, the D412A and N414A mutants were more active, with 11 ± 0.1 pM

and 19 ± 5 pM of uridylated 15A RNA produced, respectively. Interestingly, the TUT4 D416A mutant produced an average of 1.5 ± 0.2 pM of uridylated 15A RNA after 60 minutes, which is 5-fold less uridylated RNA than the TUT4-N protein. All reactions were performed in triplicate and the standard error was calculated for each TUT4 protein at each time point.

Initial velocity and specific activity for each TUT4-N mutant enzyme was determined to compare the mutant activity to the TUT4-N activity. TUT4-N had an initial velocity of 0.0012 ± 0.00013 pM/sec. When comparing the mutants, D412A and N414A displayed initial velocities of 0.0016 ± 0.00012 pM/sec and 0.0037 ± 0.0013 pM/sec, respectively, while D416A had a 5-fold decrease of activity, with an initial velocity of 0.0002 ± 0.00003 pM/sec. Within error, the activity of D412A and N414A were not significantly different than the WT (p-value >0.05) while the D416A mutant was significantly less active than the TUT4-N protein (p-value=0.009) (Table 3, Figure 25).

The specific activities of D412A (0.13 ± 0.005 pM/sec/mg) and N414A (0.31 ± 0.12 pM/sec/mg) were not significantly different (p-value >0.05) in comparison to the TUT4-N specific activity (0.10 ± 0.01 pM/sec/mg). D416A, with a specific activity of 0.02 ± 0.003 pM/sec/mg, was significantly less efficient (p-value= 0.009) than TUT4-N by 5-fold (Table 3, Figure 30).

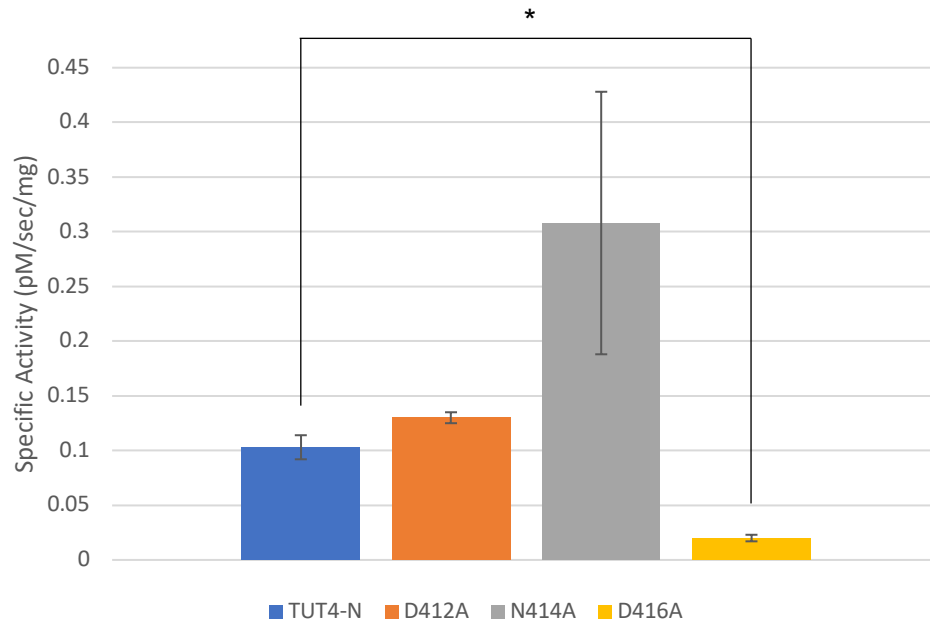


Figure 30. Specific activities of TUT4-N and mutants. D412A (0.13 ± 0.005 pM/sec/mg) and N414A (0.31 ± 0.12 pM/sec/mg) mutants did not significantly differ in activity from TUT4-N, with a specific activity of 0.10 ± 0.01 pM/sec/mg (p-value >0.05). D416A (0.02 ± 0.003 pM/sec/mg) was 5-fold less efficient than TUT4-N (p-value=0.009). The significant difference of D416A and TUT4-N is denoted by a single asterisk (*). Error bars represent one standard error.

With exception of D416A, the N-terminal residues that aligned to the C-terminal catalytic aspartates did not have an effect on the protein activity when mutated to alanine residues. Consequently, the TUT4 N-terminal region must take on a different conformational fold compared to the C-terminal domain. Other aspartate residues, or even glutamates, may be positioned in the active site, allowing for the catalysis of the nucleotide addition. There may even be only two residues involved in the N-terminal activity, hence why this catalytic region is less active than the C-terminal catalytic region. Mutation experiments performed by Hyde et al. determined that this is possible in the catalytic addition of guanine

residues to tRNA, catalyzed by tRNA^{His} guanylyltransferase (Thg1)¹²⁰. In general, the first two aspartate residues are required for positioning the metal ions required for catalysis in the G-specific nucleotidyltransferase Thg1¹²⁰. Mutation of the third carboxylate, which is a glutamate in Thg1, only resulted in a slight decrease in the polymerase activity of human Thg1 compared to the detrimental decrease in activity when the metal coordinating aspartates were mutated to alanine residues¹²⁰. Thg1 can still function, although to a lesser extent, when the glutamic acid in the active site is mutated¹²⁰. Thus, it is possible that two carboxylate residues may be sufficient for catalysis. Further mutational analysis or crystallization of the N-terminal active site will aid in the identification of the active site residues in the N-terminal catalytic domain of TUT4.

4.4 TUT4 N-terminal activity shows RNA specificity

As demonstrated in this thesis, the N-terminal domain of TUT4 is catalytically active. Surprisingly, I found substrate specificity for TUT4-N that has not been previously described for the full-length protein. Full-length TUT4 is catalytically active with mRNA⁶⁵, histone mRNA⁶⁹, and miRNA^{42, 64, 66, 67, 68} (Table 1). Interestingly, for the N-terminal domains, no activity is observed with let-7a or pre-let-7a, known substrates of the full-length TUT4 protein^{32, 42, 68}. This demonstrates that the N-terminal activity is substrate-specific and may differ from the substrate specificity of the full-length TUT4 enzyme. For full-length TUT4, it is believed that the C-terminal CCHC ZnFs are implicated in RNA binding and possibly specificity⁷³. These ZnFs may be important for RNA binding in relation to the C-terminal catalytic activity. *In vitro*, however, the C2H2 ZnF in the N-terminal

region of TUT4 may also play a role in RNA binding or specificity. Even though the analyzed TUT4 N-terminal region lacks the C-terminal ZnFs, RNA specificity and preference for miR-122, but not let-7a, was shown. In other ncPAPs, such as Cid1 and Gld2, RNA binding is not driven by the presence of ZnFs; Cid1 and Gld2 only contain two ordered domains, the Ntr and PAP domains^{25, 43}. The crystal structure of Cid1 revealed a positively charged groove on the protein surface that is thought to facilitate RNA binding³⁰. It is unknown how RNA specificity is determined in Cid1 and Gld2^{8, 43, 57, 121}, as neither protein displays a substrate specificity *in vitro*^{25, 38, 43, 121}. For the N-terminal domain, substrate specificity must be determined by binding to the active site, yet the exact mechanism remains to be elucidated.

I have shown in my data that the N-terminal catalytic region was not active with pre-let-7a and let-7a. The N-terminal catalytic region may thus play a specific role in e.g. mRNAs or specific miRNA metabolism and may not function in one of the most prevalent functions of full-length TUT4, the biogenesis of let-7a.

4.5 The purpose of two encoded catalytic regions

The activity of the full-length protein may require the presence, and therefore catalytic activity, of both the N- and C-terminal catalytic regions. In 2012, Thornton and colleagues tested the catalytic activity of varying TUT4 domains⁷³. Their experiments led to the conclusion that the N-terminal Ntr domain of TUT4 is inactive⁷³. I here show that the N-terminal domain of TUT4 displays catalytic activity, even though it is low compared to the C-terminal domain. Interestingly, in the same paper, protein variants lacking any of the N-

terminal domains, or even the second CCHC ZnF in the C-terminus also displayed no uridylation activity⁷³. Thornton and colleagues concluded that both N- and C-terminal catalytic regions were required for catalytic activity⁷³. We conducted our experiments under different conditions, including more domains in our TUT4-N protein. The inclusion of the entire TUT4 N-terminal region, and not just the N-terminal Ntr and PAP domains, could be required for uridylation activity *in vitro*. The TUT4 (and TUT7) N- and C-terminal regions, although showing activity *in vitro* when separated, may be required together to display the full extent of their activity.

While the biological function of the dual domain architecture of TUT4 and TUT7 is still unclear, other examples of dual domains have been described in nature. PAPD1, a mitochondrial adenylyltransferase, forms a dimer necessary for activity¹¹⁷. Mutation experiments determined that a stable monomer did not have PAP activity, therefore dimerization was required¹¹⁷. Thg1 is another example of a nucleotidyltransferase requiring two copies of the catalytic region for activity¹²⁰. Thg1 proteins usually exist naturally as a “dimer of dimers” and contain two copies of the catalytic domains^{120, 122, 123}. Disruption of the interaction between dimers resulted in a decrease in Thg1 activity^{120, 124}. In some plants, these “dimer of dimers” have fused, and the protein is now encoded as a tandem of two Thg1 repeats^{125, 126}. TUT4, being a large human protein, may require the two encoded catalytic regions to fulfill all of its functions in the cell.

Another example are some kinases that contain a catalytically active pseudokinase domain^{127, 128}. Recently, several pseudokinase domains which were believed to be inactive, have been shown to indeed exhibit phosphorylation

activity. In 2011, Daniela Ungureanu and colleagues discovered autophosphorylation activity in the pseudokinase domain JH2 of human JAK2, a non-receptor tyrosine kinase¹²⁷. In the case of HER3, the pseudokinase has low activity, but dimerizes with and activates the tyrosine kinases HER and HER2^{128, 129, 130}. The relevance of pseudokinase domain existence is being discovered, uncovering many biological roles.

Similar to PAPD1 and Thg1, which form dimers, or kinases with two active domains, such as JAK2 and HER3, TUT4 may require two functional catalytic regions for uridylation activity. The N-terminal catalytic domains of TUT4 may also enhance the function of the C-terminal catalytic domains, common in the dimerized enzymes discussed here.

4.6 Conclusion and Future Directions

Reiterating my hypothesis, I hypothesized that the N-terminal catalytic region of TUT4 was active and capable of uridylation activity. I here demonstrated that the TUT4 N-terminal catalytic domains are indeed catalytically active. When investigating the N-terminal domains of TUT7, uridylation activity similar to TUT4 is demonstrated.

Not only was TUT4-N catalytically active, the N-terminal domains display RNA substrate specificity, preferentially uridylating miR-122 and 15A RNA. Interestingly, no activity was observed with let-7a or pre-let-7a, which are known substrates of the full-length TUT4 protein. This implies that the TUT4 N-terminal domains may play a role in pathways other than the well-studied let-7a biogenesis pathway.

Mutation of the putative active site residues to alanine residues demonstrated no significant change in the activity of the TUT4 N-terminal domains, with the exception of D416A, which led to a significant decrease in activity. Thus, residue D416 likely plays a role in catalysis. These data indicate that the TUT4 N-terminal region may fold differently than the C-terminus. Further experiments, such as mutational analysis or a crystal structure, will elucidate which residues are involved in catalyzing the addition of uridines onto the RNA substrate.

Although TUT4-N is capable of uridylating specific RNA substrates, it would be interesting to investigate the nucleotide preference of TUT4-N. In the C-terminus, a histidine residue in the PAP domain preferentially allows UTP into the active site, by recognizing the O4 carbonyl of the uridine base as well as sterically hindering ATP from entering the active site³⁰. This amino acid is conserved in all true TUTases yet seems to be lacking in the N-terminal PAP domain of TUT4. Without the histidine residue, the TUT4 N-terminal catalytic domains, while active with UTP, may give preference to different nucleotide addition *in vivo*, including ATP. Nucleotide competition experiments and a detailed kinetic characterization will determine the NTP preference of the TUT4 N-terminus. This will elucidate other possible functions of the N-terminus of TUT4, especially for possible splicing variants which lack C-terminal domains.

Overall, the experimental results presented in this work are just the beginning of the characterization of the N-terminal domains of TUT4. The next steps to elucidate the biological functions of this activity will expand our knowledge on TUTases as a family of enzymes. The substrate specificity, and

therefore the pathways that TUT4 regulates, could be much more expansive than we originally presumed. Exploring the function of the TUT4 N-terminus could give valuable insight into other important functions of TUT4 that have yet to be elucidated.

References

1. Crick, F. On Protein Synthesis. *Symp Soc Exp Biol.* **12**, 139–163 (1958).
2. Sainsbury, S., Bernecky, C. & Cramer, P. Structural basis of transcription initiation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.* **16**, 129–143 (2015).
3. Kasinath, B. S., Mariappan, M. M., Sataranatarajan, K., Lee, M. J. & Feliers, D. mRNA Translation: Unexplored Territory in Renal Science. *J. Am. Soc. Nephrol.* **17**, 3281–3292 (2006).
4. Jackson, R. J., Hellen, C. U. T. & Pestova, T. V. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat. Rev. Mol. Cell Biol.* **11**, 113–127 (2010).
5. Aitken, C. E. & Lorsch, J. R. A mechanistic overview of translation initiation in eukaryotes. *Nat. Struct. Mol. Biol.* **19**, 568–576 (2012).
6. Zlotorynski, E. Profiling ribosome dynamics: Translation. *Nat. Rev. Mol. Cell Biol.* **17**, 535–535 (2016).
7. Schuller, A. P. & Green, R. Roadblocks and resolutions in eukaryotic translation. *Nat. Rev. Mol. Cell Biol.* (2018). doi:10.1038/s41580-018-0011-4
8. Chung, C. Z., Seidl, L. E., Mann, M. R. & Heinemann, I. U. Tipping the balance of RNA stability by 3' editing of the transcriptome. *Biochim. Biophys. Acta BBA - Gen. Subj.* **1861**, 2971-2979 (2017).
9. Kates, J. Transcription of the Vaccinia Virus Genome and the Occurrence of Polyriboadenylic acid sequences in messenger RNA. *Cold Spring Harbor Symp. Quant. Biol.* **38**, 743–752 (1970).
10. Lim, L. & Canellakis, E. S. Adenine-rich Polymer associated with Rabbit Reticulocyte Messenger RNA. *Nature* **227**, 710–712 (1970).
11. Sachs, A. The role of poly(A) in the translation and stability of mRNA. *Curr. Opin. Cell Biol.* **2**, 1092–1098 (1990).
12. Balbo, P. B. & Bohm, A. Mechanism of Poly(A) Polymerase: Structure of the Enzyme-MgATP-RNA Ternary Complex and Kinetic Analysis. *Structure* **15**, 1117–1131 (2007).

13. Chang, J. H. & Tong, L. Mitochondrial poly(A) polymerase and polyadenylation. *Biochim. Biophys. Acta BBA - Gene Regul. Mech.* **1819**, 992–997 (2012).
14. Huang, S., Deerinck, T. J., Ellisman, M. H. & Spector, D. L. In vivo analysis of the stability and transport of nuclear poly(A)+ RNA. *J. Cell Biol.* **126**, 877–899 (1994).
15. Zhao, J., Hyman, L. & Moore, C. Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol. Mol. Biol. Rev. MMBR* **63**, 405–445 (1999).
16. Rissland, O. S., Mikulasova, A. & Norbury, C. J. Efficient RNA Polyuridylation by Noncanonical Poly(A) Polymerases. *Mol. Cell. Biol.* **27**, 3612–3624 (2007).
17. Eckmann, C. R., Rammelt, C. & Wahle, E. Control of poly(A) tail length: Control of poly(A) tail length. *Wiley Interdiscip. Rev. RNA* **2**, 348–361 (2011).
18. Tian, B. & Graber, J. H. Signals for pre-mRNA cleavage and polyadenylation: Polyadenylation signals. *Wiley Interdiscip. Rev. RNA* **3**, 385–396 (2012).
19. Chan, S. L. *et al.* CPSF30 and Wdr33 directly bind to AAUAAA in mammalian mRNA 3' processing. *Genes Dev.* **28**, 2370–2380 (2014).
20. Schönemann, L. *et al.* Reconstitution of CPSF active in polyadenylation: recognition of the polyadenylation signal by WDR33. *Genes Dev.* **28**, 2381–2393 (2014).
21. Bresson, S. M., Hunter, O. V., Hunter, A. C. & Conrad, N. K. Canonical Poly(A) Polymerase Activity Promotes the Decay of a Wide Variety of Mammalian Nuclear RNAs. *PLOS Genet.* **11**, e1005610 (2015).
22. Eliseeva, I. A., Lyabin, D. N. & Ovchinnikov, L. P. Poly(A)-binding proteins: Structure, domain organization, and activity regulation. *Biochem. Mosc.* **78**, 1377–1391 (2013).
23. Baer, B. W. & Kornberg, R. D. The protein responsible for the repeating structure of cytoplasmic poly(A)-ribonucleoprotein. *J. Cell Biol.* **96**, 717–721 (1983).
24. Morgan, M. *et al.* mRNA 3' uridylation and poly(A) tail length sculpt the mammalian maternal transcriptome. *Nature* **548**, 347–351 (2017).

25. Stevenson, A. L. & Norbury, C. J. The Cid1 family of non-canonical poly(A) polymerases. *Yeast* **23**, 991–1000 (2006).
26. Schmidt, M. J. & Norbury, C. J. Polyadenylation and beyond: emerging roles for noncanonical poly(A) polymerases: Emerging roles for noncanonical poly(A) polymerases. *Wiley Interdiscip. Rev. RNA* **1**, 142–151 (2010).
27. Heo, I. *et al.* TUT4 in Concert with Lin28 Suppresses MicroRNA Biogenesis through Pre-MicroRNA Uridylation. *Cell* **138**, 696–708 (2009).
28. Villalba, A., Coll, O. & Gebauer, F. Cytoplasmic polyadenylation and translational control. *Curr. Opin. Genet. Dev.* **21**, 452–457 (2011).
29. Munoz-Tello, P., Gabus, C. & Thore, S. Functional Implications from the Cid1 Poly(U) Polymerase Crystal Structure. *Structure* **20**, 977–986 (2012).
30. Lunde, B. M., Magler, I. & Meinhart, A. Crystal structures of the Cid1 poly (U) polymerase reveal the mechanism for UTP selectivity. *Nucleic Acids Res.* **40**, 9815–9824 (2012).
31. Cheng, K., Demir, Ö. & Amaro, R. A Comparative Study of the Structural Dynamics of Four Terminal Uridyl Transferases. *Genes* **8**, 166 (2017).
32. Faehle, C. R., Walleshauser, J. & Joshua-Tor, L. Multi-domain utilization by TUT4 and TUT7 in control of let-7 biogenesis. *Nat. Struct. Mol. Biol.* **8**, 658-665 (2017).
33. Canellakis, E. S. Incorporation of radioactive uridine-5'-monophosphate into ribonucleic acid by soluble mammalian enzymes. *Biochim. Biophys. Acta* **23**, 217–218 (1957).
34. Korwek, E. L., Nakazato, N., Edmonds, M. & Venkatesan, S. Poly(uridylic acid) sequences in messenger ribonucleic acid of HeLa cells. *Biochemistry (Mosc.)* **15**, 4643–4649 (1976).
35. Seiwert, S. D., Heidmann, S. & Stuart, K. Direct Visualization of Uridylate Deletion In Vitro Suggests a Mechanism for Kinetoplastid RNA Editing. *Cell* **84**, 831–841 (1996).
36. Blum, B., Bakalara, N. & Simpson, L. A model for RNA editing in kinetoplastid mitochondria: 'guide' RNA molecules transcribed from maxicircle DNA provide the edited information. *Cell* **60**, 189–198 (1990).

37. Shen, B. & Goodman, H. M. Uridine addition after microRNA-directed cleavage. *Science* **306**, 997 (2004).
38. D'Ambrogio, A., Gu, W., Udagawa, T., Mello, C. C. & Richter, J. D. Specific miRNA Stabilization by Gld2-Catalyzed Monoadenylation. *Cell Rep.* **2**, 1537–1545 (2012).
39. Rissland, O. S. & Norbury, C. J. Decapping is preceded by 3' uridylation in a novel pathway of bulk mRNA turnover. *Nat. Struct. Mol. Biol.* **16**, 616–623 (2009).
40. Heo, I. *et al.* Lin28 Mediates the Terminal Uridylation of let-7 Precursor MicroRNA. *Mol. Cell* **32**, 276–284 (2008).
41. Mullen, T. E. & Marzluff, W. F. Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5' to 3' and 3' to 5'. *Genes Dev.* **22**, 50–65 (2008).
42. Heo, I. *et al.* Mono-Uridylation of Pre-MicroRNA as a Key Step in the Biogenesis of Group II let-7 MicroRNAs. *Cell* **151**, 521–532 (2012).
43. Chung, C. Z., Jo, D. H. S. & Heinemann, I. U. Nucleotide specificity of the human terminal nucleotidyltransferase Gld2 (TUT2). *RNA* **22**, 1239–1249 (2016).
44. Li, W., Laishram, R. S. & Anderson, R. A. The novel poly(A) polymerase Star-PAP is a signal-regulated switch at the 3'-end of mRNAs. *Adv. Biol. Regul.* **53**, 64–76 (2013).
45. Laishram, R. S. & Anderson, R. A. The poly A polymerase Star-PAP controls 3'-end cleavage by promoting CPSF interaction and specificity toward the pre-mRNA. *EMBO J.* **29**, 4132–4145 (2010).
46. Gonzales, M. L., Mellman, D. L. & Anderson, R. A. CKI α Is Associated with and Phosphorylates Star-PAP and Is Also Required for Expression of Select Star-PAP Target Messenger RNAs. *J. Biol. Chem.* **283**, 12665–12673 (2008).
47. Laishram, R. S., Barlow, C. A. & Anderson, R. A. CKI isoforms α and ϵ regulate Star-PAP target messages by controlling Star-PAP poly(A) polymerase activity and phosphoinositide stimulation. *Nucleic Acids Res.* **39**, 7961–7973 (2011).

48. Li, W. *et al.* Star-PAP Control of BIK Expression and Apoptosis Is Regulated by Nuclear PIPK1 α and PKC δ Signaling. *Mol. Cell* **45**, 25–37 (2012).
49. Mohan, N., Sudheesh, A. P., Francis, N., Anderson, R. & Laishram, R. S. Phosphorylation regulates the Star-PAP-PIPK1 α interaction and directs specificity toward mRNA targets. *Nucleic Acids Res.* **43**, 7005–7020 (2015).
50. Knouf, E. C., Wyman, S. K. & Tewari, M. The Human TUT1 Nucleotidyl Transferase as a Global Regulator of microRNA Abundance. *PLoS ONE* **8**, e69630 (2013).
51. Zhu, D., Lou, Y., He, Z. & Ji, M. Nucleotidyl transferase TUT1 inhibits lipogenesis in osteosarcoma cells through regulation of microRNA-24 and microRNA-29a. *Tumor Biol.* **35**, 11829–11835 (2014).
52. Trippe, R. A highly specific terminal uridylyl transferase modifies the 3'-end of U6 small nuclear RNA. *Nucleic Acids Res.* **26**, 3119–3126 (1998).
53. Trippe, R., Richly, H. & Benecke, B. J. Biochemical characterization of a U6 small nuclear RNA-specific terminal uridylyltransferase: Characterization of U6 terminal uridylyltransferase. *Eur. J. Biochem.* **270**, 971–980 (2003).
54. Trippe, R. Identification, cloning, and functional analysis of the human U6 snRNA-specific terminal uridylyl transferase. *RNA* **12**, 1494–1504 (2006).
55. Barnard, D. C., Ryan, K., Manley, J. L. & Richter, J. D. Symplekin and xGLD-2 Are Required for CPEB-Mediated Cytoplasmic Polyadenylation. *Cell* **119**, 641–651 (2004).
56. Kim, J. H. & Richter, J. D. Opposing Polymerase-Deadenylase Activities Regulate Cytoplasmic Polyadenylation. *Mol. Cell* **24**, 173–183 (2006).
57. Kwak, J. E. & Wickens, M. A family of poly(U) polymerases. *RNA* **13**, 860–867 (2007).
58. Rouhana, L. Vertebrate GLD2 poly(A) polymerases in the germline and the brain. *RNA* **11**, 1117–1130 (2005).
59. Glahder, J. A. & Norrild, B. Involvement of hGLD-2 in cytoplasmic polyadenylation of human p53 mRNA. *APMIS* **119**, 769–775 (2011).

60. Burns, D. M., D'Ambrogio, A., Nottrott, S. & Richter, J. D. CPEB and two poly(A) polymerases control miR-122 stability and p53 mRNA translation. *Nature* **473**, 105–108 (2011).
61. Katoh, T. *et al.* Selective stabilization of mammalian microRNAs by 3' adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2. *Genes Dev.* **23**, 433–438 (2009).
62. Lan, H., Chung, A. & Yu, X. MicroRNA and nephropathy: emerging concepts. *Int. J. Nephrol. Renov. Dis.* **6**, 169-179 (2013).
63. Kim, S. J. *et al.* MicroRNA let-7a suppresses breast cancer cell migration and invasion through downregulation of C-C chemokine receptor type 7. *Breast Cancer Res.* **14**, R14 (2012).
64. Kim, B. *et al.* TUT7 controls the fate of precursor microRNAs by using three different uridylation mechanisms. *EMBO J.* **34**, 1801–1815 (2015).
65. Lim, J. *et al.* Uridylation by TUT4 and TUT7 Marks mRNA for Degradation. *Cell* **159**, 1365–1376 (2014).
66. Jones, M. R. *et al.* Zcchc11 Uridylates Mature miRNAs to Enhance Neonatal IGF-1 Expression, Growth, and Survival. *PLoS Genet.* **8**, e1003105 (2012).
67. Jones, M. R. *et al.* Zcchc11-dependent uridylation of microRNA directs cytokine expression. *Nat. Cell Biol.* **11**, 1157–1163 (2009).
68. Thornton, J. E. *et al.* Selective microRNA uridylation by Zcchc6 (TUT7) and Zcchc11 (TUT4). *Nucleic Acids Res.* **42**, 11777–11791 (2014).
69. Schmidt, M. J., West, S. & Norbury, C. J. The human cytoplasmic RNA terminal U-transferase ZCCHC11 targets histone mRNAs for degradation. *RNA* **17**, 39–44 (2011).
70. Lackey, P. E., Welch, J. D. & Marzluff, W. F. TUT7 catalyzes the uridylation of the 3' end for rapid degradation of histone mRNA. *RNA* **22**, 1673–1688 (2016).
71. Yamashita, S., Takagi, Y., Nagaïke, T. & Tomita, K. Crystal structures of U6 snRNA-specific terminal uridylyltransferase. *Nat. Commun.* **8**, 15788 (2017).
72. Hagan, J. P., Piskounova, E. & Gregory, R. I. Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. *Nat. Struct. Mol. Biol.* **16**, 1021–1025 (2009).

73. Thornton, J. E., Chang, H. M., Piskounova, E. & Gregory, R. I. Lin28-mediated control of let-7 microRNA expression by alternative TUTases Zcchc11 (TUT4) and Zcchc6 (TUT7). *RNA* **18**, 1875–1885 (2012).
74. Blahna, M. T., Jones, M. R., Quinton, L. J., Matsuura, K. Y. & Mizgerd, J. P. Terminal Uridyltransferase Enzyme Zcchc11 Promotes Cell Proliferation Independent of Its Uridyltransferase Activity. *J. Biol. Chem.* **286**, 42381–42389 (2011).
75. Kwak, J. E. *et al.* GLD2 poly(A) polymerase is required for long-term memory. *Proc. Natl. Acad. Sci.* **105**, 14644–14649 (2008).
76. Kwak, J. E., Wang, L., Ballantyne, S., Kimble, J. & Wickens, M. Mammalian GLD-2 homologs are poly(A) polymerases. *Proc. Natl. Acad. Sci.* **101**, 4407–4412 (2004).
77. Rissland, O. S. & Norbury, C. J. The Cid1 poly(U) polymerase. *Biochim. Biophys. Acta BBA - Gene Regul. Mech.* **1779**, 286–294 (2008).
78. Perumal, K. & Reddy, R. The 3' end formation in small RNAs. *Gene Expr.* **10**, 59–78 (2002).
79. Achsel, T. *et al.* A doughnut-shaped heteromer of human Sm-like proteins binds to the 3'-end of U6 snRNA, thereby facilitating U4/U6 duplex formation in vitro. *EMBO J.* **18**, 5789–5802 (1999).
80. Vidal, V. P., Verdone, L., Mayes, A. E. & Beggs, J. D. Characterization of U6 snRNA-protein interactions. *RNA N. Y. N* **5**, 1470–1481 (1999).
81. De Almeida, C., Scheer, H., Zuber, H. & Gagliardi, D. RNA uridylation: a key posttranscriptional modification shaping the coding and noncoding transcriptome: RNA uridylation. *Wiley Interdiscip. Rev. RNA* **9**, e1440 (2018).
82. Chang, H., Lim, J., Ha, M. & Kim, V. N. TAIL-seq: Genome-wide Determination of Poly(A) Tail Length and 3' End Modifications. *Mol. Cell* **53**, 1044–1052 (2014).
83. Feng, Y., Zhang, X., Graves, P. & Zeng, Y. A comprehensive analysis of precursor microRNA cleavage by human Dicer. *RNA* **18**, 2083–2092 (2012).
84. Lee, Y. *et al.* The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**, 415–419 (2003).

85. Yi, R. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev.* **17**, 3011–3016 (2003).
86. Zhang, H. Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. *EMBO J.* **21**, 5875–5885 (2002).
87. Zhang, H., Kolb, F. A., Jaskiewicz, L., Westhof, E. & Filipowicz, W. Single Processing Center Models for Human Dicer and Bacterial RNase III. *Cell* **118**, 57–68 (2004).
88. Ha, M. & Kim, V. N. Regulation of microRNA biogenesis. *Nat. Rev. Mol. Cell Biol.* **15**, 509–524 (2014).
89. Suzuki, H. I., Katsura, A. & Miyazono, K. A role of uridylation pathway for blockade of let-7 microRNA biogenesis by Lin28B. *Cancer Sci.* **106**, 1174–1181 (2015).
90. Faehnle, C. R., Walleshauser, J. & Joshua-Tor, L. Mechanism of Dis3L2 substrate recognition in the Lin28–let-7 pathway. *Nature* **514**, 252–256 (2014).
91. Ustianenko, D. *et al.* Mammalian DIS3L2 exoribonuclease targets the uridylated precursors of let-7 miRNAs. *RNA* **19**, 1632–1638 (2013).
92. Piskounova, E. *et al.* Lin28A and Lin28B Inhibit let-7 MicroRNA Biogenesis by Distinct Mechanisms. *Cell* **147**, 1066–1079 (2011).
93. Büssing, I., Slack, F. J. & Großhans, H. let-7 microRNAs in development, stem cells and cancer. *Trends Mol. Med.* **14**, 400–409 (2008).
94. Balzeau, J., Menezes, M. R., Cao, S. & Hagan, J. P. The LIN28/let-7 Pathway in Cancer. *Front. Genet.* **8**, (2017).
95. Zhou, J., Ng, S. B. & Chng, W. J. LIN28/LIN28B: An emerging oncogenic driver in cancer stem cells. *Int. J. Biochem. Cell Biol.* **45**, 973–978 (2013).
96. Viswanathan, S. R. *et al.* Lin28 promotes transformation and is associated with advanced human malignancies. *Nat. Genet.* **41**, 843–848 (2009).
97. Romero-Cordoba, S. L., Salido-Guadarrama, I., Rodriguez-Dorantes, M. & Hidalgo-Miranda, A. miRNA biogenesis: Biological impact in the development of cancer. *Cancer Biol. Ther.* **15**, 1444–1455 (2014).
98. Ota, T. *et al.* Complete sequencing and characterization of 21,243 full-length human cDNAs. *Nat. Genet.* **36**, 40–45 (2004).

99. Wakamatsu, A. *et al.* Identification and Functional Analyses of 11 769 Full-length Human cDNAs Focused on Alternative Splicing. *DNA Res.* **16**, 371–383 (2009).
100. Thierry-Mieg, D. & Thierry-Mieg, J. Aceview: a comprehensive cDNA-supported gene and transcripts annotation. *Genome Biol.* **7(Suppl 1)**, (2006).
101. Edelheit, O., Hanukoglu, A. & Hanukoglu, I. Simple and efficient site-directed mutagenesis using two single-primer reactions in parallel to generate mutants for protein structure-function studies. *BMC Biotechnol.* **9**, 61 (2009).
102. Harwood, A. J. & Durrant, I. *Protocols for Gene Analysis.* **31**, 147-161 (Humana Press, 1994).
103. Kapeluich, Y. L., Rubtsova MYu, & Egorov, A. M. Enhanced chemiluminescence reaction applied to the study of horseradish peroxidase stability in the course of p-iodophenol oxidation. *J. Biolumin. Chemilumin.* **12**, 299–308 (1997).
104. Lin, S. & Gregory, R. I. Identification of small molecule inhibitors of Zcchc11 TUTase activity. *RNA Biol.* **12**, 792–800 (2015).
105. Rosano, G. L. & Ceccarelli, E. A. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front. Microbiol.* **5**, (2014).
106. Khow, O. & Suntrarachun, S. Strategies for production of active eukaryotic proteins in bacterial expression system. *Asian Pac. J. Trop. Biomed.* **2**, 159–162 (2012).
107. Faure, G., Ogurtsov, A. Y., Shabalina, S. A. & Koonin, E. V. Role of mRNA structure in the control of protein folding. *Nucleic Acids Res.* **44**, 10898–10911 (2016).
108. Kurland, C. G. Codon bias and gene expression. *FEBS Lett.* **285**, 165–169 (1991).
109. Jarvis, D. L. Chapter 14 Baculovirus–Insect Cell Expression Systems. in *Methods in Enzymology* **463**, 191–222 (Elsevier, 2009).
110. Nakel, K., Bonneau, F., Eckmann, C. R. & Conti, E. Structural basis for the activation of the *C. elegans* noncanonical cytoplasmic poly(A)-polymerase GLD-2 by GLD-3. *Proc. Natl. Acad. Sci.* **112**, 8614–8619 (2015).

111. Aphasizhev, R. *et al.* Trypanosome Mitochondrial 3' Terminal Uridylyl Transferase (TUTase). *Cell* **108**, 637–648 (2002).
112. Stagno, J., Aphasizheva, I., Rosengarth, A., Luecke, H. & Aphasizhev, R. UTP-bound and Apo Structures of a Minimal RNA Uridylyltransferase. *J. Mol. Biol.* **366**, 882–899 (2007).
113. Carrillo, N., Ceccarelli, E. & Roveri, O. Usefulness of kinetic enzyme parameters in biotechnological practice. *Biotechnol. Genet. Eng. Rev.* **27**, 367–382 (2010).
114. Aphasizheva, I., Aphasizhev, R. & Simpson, L. RNA-editing Terminal Uridylyl Transferase 1: Identification of functional domains by mutational analysis. *J. Biol. Chem.* **279**, 24123–24130 (2004).
115. Raynal, L. C., Krisch, H. M. & Carpousis, A. J. Bacterial poly(A) polymerase: an enzyme that modulates RNA stability. *Biochimie* **78**, 390–398 (1996).
116. Laishram, R. S. Poly(A) polymerase (PAP) diversity in gene expression - Star-PAP vs canonical PAP. *FEBS Lett.* **588**, 2185–2197 (2014).
117. Bai, Y., Srivastava, S. K., Chang, J. H., Manley, J. L. & Tong, L. Structural Basis for Dimerization and Activity of Human PAPD1, a Noncanonical Poly(A) Polymerase. *Mol. Cell* **41**, 311–320 (2011).
118. Alfalah, M., Keiser, M., Leeb, T., Zimmer, K. & Naim, H. Y. Compound Heterozygous Mutations Affect Protein Folding and Function in Patients With Congenital Sucrase-Isomaltase Deficiency. *Gastroenterology* **136**, 883–892 (2009).
119. Reva, B., Antipin, Y. & Sander, C. Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic Acids Res.* **39**, e118 (2011).
120. Hyde, S. J. *et al.* tRNA^{His} guanylyltransferase (THG1), a unique 3'-5' nucleotidyl transferase, shares unexpected structural homology with canonical 5'-3' DNA polymerases. *Proc. Natl. Acad. Sci.* **107**, 20305–20310 (2010).
121. Read, R. L., Martinho, R. G., Wang, S. W., Carr, A. M. & Norbury, C. J. Cytoplasmic poly(A) polymerases mediate cellular responses to S phase arrest. *Proc. Natl. Acad. Sci.* **99**, 12079–12084 (2002).

122. Nakamura, A. *et al.* Structural basis of reverse nucleotide polymerization. *Proc. Natl. Acad. Sci.* **110**, 20970–20975 (2013).
123. Desai, R. *et al.* Minimal requirements for reverse polymerization and tRNA repair by tRNA^{His} guanylyltransferase. *RNA Biol.* 1–9 (2017).
doi:10.1080/15476286.2017.1372076
124. Jackman, J. E. & Phizicky, E. M. Identification of Critical Residues for G₋₁ Addition and Substrate Recognition by tRNA^{His} Guanylyltransferase †. *Biochemistry (Mosc.)* **47**, 4817–4825 (2008).
125. Heinemann, I. U., Randau, L., Tomko, R. J. & Söll, D. 3'-5' tRNA^{His} guanylyltransferase in bacteria. *FEBS Lett.* **584**, 3567–3572 (2010).
126. Heinemann, I. U., Nakamura, A., O'Donoghue, P., Eiler, D. & Soll, D. tRNA^{His}-guanylyltransferase establishes tRNA^{His} identity. *Nucleic Acids Res.* **40**, 333–344 (2012).
127. Ungureanu, D. *et al.* The pseudokinase domain of JAK2 is a dual-specificity protein kinase that negatively regulates cytokine signaling. *Nat. Struct. Mol. Biol.* **18**, 971–976 (2011).
128. Byrne, D. P., Foulkes, D. M. & Evers, P. A. Pseudokinases: update on their functions and evaluation as new drug targets. *Future Med. Chem.* **9**, 245–265 (2017).
129. Shi, F., Telesco, S. E., Liu, Y., Radhakrishnan, R. & Lemmon, M. A. ErbB3/HER3 intracellular domain is competent to bind ATP and catalyze autophosphorylation. *Proc. Natl. Acad. Sci.* **107**, 7692–7697 (2010).
130. Yarden, Y. & Sliwkowski, M. X. Untangling the ErbB signalling network. *Nat. Rev. Mol. Cell Biol.* **2**, 127-137 (2001).

Curriculum Vitae

Lauren E. Seidl

Post-Secondary Education

- 2016-2018 **Master of Science Candidate, Biochemistry**
The University of Western Ontario, London, Ontario, Canada
Title: The Functional Characterization of the N-terminal Domains of TUT4
- 2012-2016 **Bachelor of Science, Honours Biochemistry with Thesis**
University of Windsor, Windsor, Ontario, Canada
Thesis: Determining the Interaction between Polo-like Kinase 4 and Nucleophosmin

Scholarships and Awards

- 2016-2018 **Western Graduate Research Scholarship**
- Awarded to eligible full-time graduate students
- 2016 **Distinction**
- Awarded for achieving above an 80% average overall in Bachelor's Degree
- 2016 **Dean's Honour Roll**
- Awarded for achieving above an 80% average in the final year of Bachelor's degree

- 2012 **University of Windsor Entrance Scholarship**
- Awarded for achieving above an 85% average overall in secondary school
- 2012 **Ciociaro Club of Windsor Bursary**
- Awarded as a member of the Ciociaro Club of Windsor for academic excellence in secondary school

Research Experience

- 2016-2018 **Research-Based Master's Thesis**, Department of Biochemistry
The University of Western Ontario, London, Ontario, Canada
Supervisor: Dr. Ilka Heinemann
- 2016 **Research Assistant**, Department of Biology
University of Windsor, Windsor, Ontario, Canada
Supervisor: Dr. John Hudson
- 2015-2016 **Undergraduate Thesis Project**, Department of Biology
University of Windsor, Windsor, Ontario, Canada
Supervisor: Dr. John Hudson

Posters and Presentations

- 2018 **Oral Presentation**, Third Annual Toronto RNA Enthusiasts' Day
(TREN D) 2018, Toronto, Ontario, Canada (July 31, 2018)
- Title: The Functional Characterization of TUT4 N-terminal Domains

- Authors: Lauren E. Seidl, Matthew A. Turk, Ilka U. Heinemann
- 2018 **Oral Presentation**, Graduate Student Seminars, Department of Biochemistry (June 13, 2018)
- Title: The Functional Characterization of TUT4 N-terminal Domains
 - Authors: Lauren E. Seidl
- 2017 **Poster Presentation**, RiboClub 2017 in Association with the Swiss National Center of Competence in RNA and Disease, Sherbrooke, Quebec, Canada (September 24-28, 2018)
- Title: TUT4 and its Isoforms: Potential Targets for Let-7 Downregulated Cancers
 - Authors: Lauren E. Seidl, Matthew A. Turk, and Ilka U. Heinemann
- 2017 **Poster Presentation**, Second Annual Toronto RNA Enthusiasts' Day (TRENd) 2017, Toronto, Ontario, Canada
- Title: Differential Expression of Two RNA Editing TUT Isoforms May Play a Role in Breast Cancer Development
 - Authors: Lauren E. Seidl, Matthew A. Turk, and Ilka U. Heinemann
- 2017 **Poster Presentation**, Harold B. Stewart Memorial Lecture and Research Showcase, London, Ontario, Canada (January 19, 2018)
- Title: Characterizing Human Terminal Uridyltransferase 4

- Authors: Lauren E. Seidl, Yumin Bi, Christina C. Chung, and Ilka U. Heinemann

Publications

Seidl, L. E., Lee, K., Chung, C. Z., Bi, Y., Heinemann, I. U. The N-terminal domain of the human uridylyltransferase TUT4 displays distinct catalytic activity. Manuscript in preparation. (2018)

Jaramillo, J. E., Ellis, M., Chung, C. Z., **Seidl, L. E.**, Jo, D. H. S., Mann, M. M., Turk, M. A., Bi, Y., Haniford, D., Heinemann, I. U. RNA Surveillance by Uridylation dependent RNA decay in *Schizosaccharomyces pombe*. *Nucleic Acid Research*. (2018) Under Revision. NAR-00513-A-2018.

Chung, C. Z., **Seidl, L. E.**, Mann, M. R. & Heinemann, I. U. Tipping the balance of RNA stability by 3' editing of the transcriptome. *Biochim. Biophys. Acta BBA - Gen. Subj.* **1861**, 2971-2979 (2017).