

7-11-2019 10:30 AM

## Regulation of RNA stability by terminal nucleotidyltransferases

Christina Z. Chung, *The University of Western Ontario*

Supervisor: Heinemann, Ilka U., *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree  
in Biochemistry

© Christina Z. Chung 2019

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

 Part of the [Biochemistry Commons](#)

---

### Recommended Citation

Chung, Christina Z., "Regulation of RNA stability by terminal nucleotidyltransferases" (2019). *Electronic Thesis and Dissertation Repository*. 6255.  
<https://ir.lib.uwo.ca/etd/6255>

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](mailto:wlsadmin@uwo.ca).

## Abstract

The dysregulation of RNAs has global effects on all cellular pathways. The regulation of RNA metabolism is thus tightly controlled. Terminal RNA nucleotidyltransferases (TENTs) regulate RNA stability and activity through the addition of non-templated nucleotides to the 3'-end. TENT-catalyzed adenylation and uridylation have opposing effects; adenylation stabilizes while uridylation silences or degrades RNA. All TENT homologs were initially characterized as adenylyltransferases; the identification of caffeine-induced death suppressor protein 1 (Cid1) in *Schizosaccharomyces pombe* as an uridylyltransferase led to the reclassification of many TENTs as uridylyltransferases. Cid1 uridylates mRNAs that are subsequently degraded by the exonuclease Dis-like 3'-5' exonuclease 2 (Dis3L2), while the human homolog germline-development 2 (Gld2) has been associated with adenylation of mRNAs and miRNAs and uridylation of Group II pre-miRNAs. Mechanisms regulating these enzymes and the extent of TENT activity on cellular RNA homeostasis remain largely unknown. In this thesis, the regulation of human Gld2 and the role of the yeast Cid1/Dis3L2-mediated RNA decay pathway were investigated. An enzyme kinetic study revealed that Gld2 is a true adenylyltransferase with only weak activity for UTP. A detailed phylogenetic analysis revealed that uridylyltransferases arose multiple times during evolution through a single histidine insertion in the active site of adenylyltransferases. Insertion of the critical histidine into Gld2 changed its nucleotide preference from ATP to UTP. Next, the regulation of Gld2 through site-specific phosphorylation in the predicted disordered N-terminal domain was investigated using phosphomimetic substitutions at specific serine (S) residues. Two sites (S62, S110) increased Gld2 activity while one site (S116) drastically reduced 3'-adenylation activity. Mass spectrometry and *in vitro* activity assays identified protein kinases A (PKA) and B (Akt1) as kinases that specifically phosphorylate Gld2 at S116 to obliterate nucleotide addition activity similarly to the S116E phosphomimetic mutant. Finally, RNA deep sequencing of *cid1* and *dis3L2* *S. pombe* deletion strains revealed that the role of Cid1 is redundant in uridylation-dependent mRNA decay while Dis3L2 is the bottleneck to RNA decay. Deletion of either gene increases the accumulation of misfolded proteins but only the *dis3L2* deletion up-regulates stress response proteins.

Overall, this thesis demonstrates how terminal nucleotidyltransferases regulate RNA stability.

## Keywords

RNA editing; miRNA; miRNA maturation; nucleotidyltransferase; terminal uridylyltransferase; adenylation; uridylation; phosphorylation; enzyme kinetics; post-translational modification; RNA degradation; mixed RNA tails; unfolded protein response

## Summary for Lay Audience

Ribonucleic acids (RNAs) play important roles in protein production and regulating cellular processes such as cell proliferation. Dysregulation of RNA expression, maturation, and/or degradation is associated with multiple human diseases such as cancer and cardiovascular disease. RNAs can be regulated through the addition of adenine or uridine nucleotides to its 3'-end. The proteins that perform these additions are known as terminal RNA nucleotidyltransferases (TENTs). All TENTs were initially thought to add adenine residues (adenylyltransferases), but more extensive studies revealed that some TENTs preferred to add uridine residues (uridylyltransferases). The addition of adenine is associated with stability while uridine addition is associated with silencing/degradation. Thus, the simple addition of different nucleotides can change the fate of an RNA molecule. The yeast TENT uridylylates RNAs which are recognized and degraded by the exonuclease Dis-like 3'-5' exonuclease 2 (Dis3L2). On the other hand, its human counterpart, germline-development 2 (Gld2), has been associated with RNA adenylation and uridylation. Mechanisms regulating these proteins and the extent of TENT activity on cellular RNA homeostasis remain largely unknown. In this thesis, the regulation of human Gld2 and the role of the yeast Cid1/Dis3L2-mediated RNA decay pathway were investigated. First, Gld2 was shown to be a true adenylyltransferase. The simple insertion or deletion of the amino acid histidine in the active site was shown to change the nucleotide preference of TENTs. Secondly, Gld2 was shown to be regulated through phosphorylation of specific serine residues (S). Two sites (S62, S110) increased Gld2 activity while one site (S116) drastically reduced activity. Two cancer-related kinases, protein kinases A (PKA) and B (Akt1), were identified to phosphorylate Gld2 at S116 to obliterate nucleotide addition activity. This discovery provided the first link between cancer-related kinases and RNA regulation. Finally, deletion of either the Cid1 or Dis3L2 genes in yeast revealed that Cid1 is redundant in uridylation-dependent mRNA decay while Dis3L2 is the bottleneck to RNA decay. Deletion of the Dis3L2 gene elicited a larger change in the RNA population. Overall, this thesis demonstrates how terminal nucleotidyltransferases regulate RNA stability.

# Co-Authorship Statement

## Chapter 1

Sections of the text and some figures were adapted from the published review titled “*Tipping the balance of RNA stability by 3' editing of the transcriptome.*” The review is co-authored by Christina Z. Chung, Lauren E. Seidl, and Mitchell R. Mann. All authors contributed to creating the figures and writing the manuscript.

Chung CZ\*, Seidl LE\*, Mann MR\*, & Heinemann IU. 2017. Tipping the balance of RNA stability by 3' editing of the transcriptome. *Biochimica et Biophysica Acta*. **1861**: 2971-2979 (\*These authors contributed equally).

## Chapter 2

Text and figures are from the published paper titled “*Nucleotide specificity of the human terminal nucleotidyltransferase Gld2 (TUT2).*” David H.S. Jo constructed the wildtype Gld2 expression plasmid and contributed to the phylogenetic analysis. Ilka U. Heinemann performed the phylogenetic analysis and sequence alignment. All other experiments were carried out by Christina Z. Chung. Both I.U. Heinemann and C.Z. Chung contributed to the writing of the manuscript.

Chung CZ, Jo DHS, & Heinemann IU. 2016. Nucleotide specificity of the human terminal nucleotidyltransferase Gld2 (TUT2). *RNA N. Y. N.* **22(8)**: 1239-1249.

## Chapter 3

Text and figures are from the published paper titled “*Gld2 activity is regulated by phosphorylation in the N-terminal domain.*” Patrick O'Donoghue performed the multiple sequence alignment and Xuguang Liu carried out the mass spectrometry analysis. Nileeka Balasuriya cloned, expressed, and purified all the Akt1 variants and performed the dot plot kinase activity assays. All other experiments were carried out by Christina Z. Chung. C.Z. Chung, N. Balasuriya, X. Liu, P. O'Donoghue, and I.U. Heinemann contributed to the writing of the manuscript.

Chung CZ, Balasuriya N, Manni E, Liu X, Li SSC, O'Donoghue P, & Heinemann IU. 2019. Gld2 activity is regulated by phosphorylation in the N-terminal domain. *RNA Biol.* doi:10.1080/15476286.2019.1608754.

## **Chapter 4**

Text and figures are from the published paper titled “*RNA surveillance by uridylation-dependent RNA decay in Schizosaccharomyces pombe.*” David H.S. Jo cloned the Cid1 expression plasmid and performed the cRACE experiment. D.H.S. Jo, Julia E. Jaramillo, Lauren E. Seidl, Matthew A. Turk, and Christina Z. Chung purified Cid1. The Cid1 activity assay was performed by C.Z. Chung and L.E. Seidl and the yeast spotting assays were performed by Daniel Y.N. Bour. The Northern blots were performed by J.E. Jaramillo and Yumin Bi and the RT-qPCR by C.Z. Chung and J.E. Jaramillo. C.Z. Chung performed the sedimentation assay and western blot and Ilka U. Heinemann carried out the RNA sequencing and data analysis. Michael J. Ellis contributed to the analysis of the sequencing data. All authors contributed to the writing of the manuscript.

Chung CZ\*, Jaramillo JE\*, Ellis MJ, Bour DYN, Seidl LE, Jo DHS, Turk MA, Mann MR, Bi Y, Haniford DB, Duennwald ML, & Heinemann IU. 2018. RNA surveillance by uridylation-dependent RNA decay in *Schizosaccharomyces pombe*. *Nucleic Acids Res.* **47(6)**: 3045-3057 (\*These authors contributed equally).

## Dedication

*For my loving and supportive parents, Betsy and Lawrence Chung.*

## Acknowledgements

I would like to thank my supervisor Dr. Ilka U. Heinemann for her support and guidance. Her enthusiasm and optimism have continuously pushed me forward and I am grateful for all the opportunities she has given me.

Many thanks to my advisory committee, Dr. Murray Junop, Dr. Stanley Dunn, and Dr. David Haniford, for their guidance and suggestions.

All members of the Heinemann lab, past and present, have my thanks for making my time there enjoyable and all the amazing discussions. Special thanks to Dr. Patrick O'Donoghue and all the past and present members of the O'Donoghue lab for all the support over the years.

Finally, many people in the Biochemistry department have helped me from my time there as an undergraduate student. I would like to extend my thanks to all of you.



# Table of Contents

Abstract .....	ii
Summary for Lay Audience .....	iv
Co-Authorship Statement.....	v
Dedication .....	vii
Acknowledgements .....	viii
Table of Contents .....	ix
List of Tables .....	xv
List of Figures .....	xvi
List of Appendices .....	xviii
List of Abbreviations .....	xx
Chapter 1 .....	1
1 Introduction .....	1
1.1 Central Dogma of Molecular Biology .....	1
1.2 Regulation of mRNA .....	3
1.2.1 mRNA transcription in eukaryotes .....	3
1.2.2 Role of the poly(A) tail .....	5
1.2.3 5'-3' and 3'-5' degradation of mRNAs .....	6
1.3 miRNAs: Functions in Gene Expression .....	10
1.3.1 Role of miRNAs in regulating gene expression .....	10
1.3.2 miRNA maturation pathway .....	11
1.3.3 Regulation of miRNA through untemplated nucleotide addition .....	12
1.4 Terminal RNA Nucleotidyltransferases .....	15
1.4.1 Adenylation vs. uridylation .....	16
1.4.2 Human uridylyltransferases .....	18

1.4.3	Gld2, the minimal human nucleotidyltransferase .....	20
1.4.4	The founder uridylyltransferase Cid1 .....	21
1.5	Scope of Thesis .....	22
1.6	References .....	24
Chapter 2	.....	34
2	Nucleotide specificity of the human terminal nucleotidyltransferase Gld2 (TUT2) ...	34
2.1	Introduction .....	34
2.2	Materials and Methods .....	37
2.2.1	Gld2 cloning and site-directed mutagenesis .....	37
2.2.2	Gld2 expression and purification .....	38
2.2.3	Size exclusion chromatography .....	38
2.2.4	RNA substrate preparation .....	39
2.2.5	Determination of enzymatic activity and substrate range .....	40
2.2.6	Determination of enzyme kinetics .....	40
2.2.7	Phylogenetic analysis .....	41
2.3	Results .....	41
2.3.1	Gld2 displays RNA substrate promiscuity .....	41
2.3.2	Nucleotide preference of recombinant Gld2 .....	43
2.3.3	Gld2 is an adenylyltransferase .....	44
2.3.4	Nucleotidyltransferase specificity arose multiple times during evolution .....	46
2.3.5	Insertion of a histidine residue confers UTP specificity .....	48
2.4	Discussion .....	51
2.4.1	Gld2 is an adenylyltransferase .....	51
2.4.2	Gld2 monoadenylates small RNA substrates .....	53

2.4.3	Convergent evolution of Tutase activity by histidine insertion in the PAP active site .....	54
2.5	Acknowledgements .....	56
2.6	References .....	56
Chapter 3	.....	60
3	Gld2 activity is regulated by phosphorylation in the N-terminal domain.....	60
3.1	Introduction.....	60
3.2	Materials and Methods.....	63
3.2.1	Multiple sequence alignment .....	63
3.2.2	Plasmids .....	64
3.2.3	Gld2 protein production and purification .....	64
3.2.4	Western blotting.....	64
3.2.5	Nucleotide addition assay .....	65
3.2.6	Fluorescence anisotropy.....	65
3.2.7	Identification of potential kinases .....	66
3.2.8	Dot plot kinase activity assays .....	66
3.2.9	Kinase activity assays using SDS gels.....	66
3.2.10	Isolation of phosphorylated Gld2 for downstream assays .....	67
3.2.11	Phosphorylation of Gld2 using HEK 293 cell extract .....	67
3.2.12	Mass spectrometry .....	68
3.3	Results.....	68
3.3.1	Phosphorylation of the Gld2 N-terminal domain by HEK 293 cells .....	69
3.3.2	Gld2 N-terminal domain phosphomimetic variants regulate catalytic activity.....	71
3.3.3	Gld2 phosphomimetic substitutions impact RNA substrate affinity .....	73
3.3.4	PKA and Akt1 site-specifically phosphorylate Gld2 at S116.....	75

3.3.5	Phosphorylation of Gld2 at S116 abolishes nucleotide addition activity .	78
3.4	Discussion .....	79
3.4.1	Gld2 activity is regulated by phosphorylation .....	79
3.4.2	Oncogenic protein kinases signal to miRNA regulation.....	82
3.4.3	Relevance of Akt1-dependent regulation of miRNAs to disease .....	83
3.4.4	Conclusion .....	84
3.5	Acknowledgements.....	85
3.6	References.....	85
Chapter 4	.....	90
4	RNA surveillance by uridylation-dependent RNA decay in <i>Schizosaccharomyces pombe</i> .....	90
4.1	Introduction.....	90
4.2	Materials and Methods.....	93
4.2.1	Yeast strains and growth conditions .....	93
4.2.2	Spotting assays .....	94
4.2.3	Cid1 cloning, purification, and activity assays .....	94
4.2.4	Circular rapid amplification of cDNA ends (cRACE) and Northern blotting .....	95
4.2.5	RT-qPCR.....	96
4.2.6	Yeast sedimentation assay and Western blot .....	96
4.2.7	RNA sequencing and data analysis .....	97
4.3	Results.....	97
4.3.1	Recombinant Cid1 displays ambiguous substrate specificity <i>in vitro</i> .....	97
4.3.2	RNA uridylation is prominent in wildtype <i>S. pombe</i> and a <i>cid1</i> deletion strain.....	99
4.3.3	Deletion of the Dis3L2 exonuclease elicits changes in the transcriptome.....	100

4.3.4	Deletion of <i>dis3L2</i> confers resistance to hydroxyurea, whereas deletion of <i>cid1</i> increases sensitivity to protein misfolding stress.....	106
4.3.5	Deletion of <i>dis3L2</i> and <i>cid1</i> causes the accumulation of misfolded proteins.....	108
4.4	Discussion.....	109
4.4.1	mRNA uridylation does not exclusively depend on Cid1 .....	109
4.4.2	Mixed mRNA A/U tails .....	110
4.4.3	Uridylation-dependent RNA decay is linked to stress response and telomere maintenance .....	110
4.4.4	Dis3L2 depletion increases resistance to hydroxyurea-induced stress ...	112
4.5	Acknowledgements.....	114
4.6	References.....	114
Chapter 5	.....	118
5	Summary and Perspectives .....	118
5.1	The evolution of adenylyl- and uridylyltransferase nucleotide specificities .....	119
5.2	Regulation of Gld2 activity by post-translational phosphorylation.....	120
5.3	The role of the Cid1/Dis3L2 pathway in global RNA decay.....	121
5.4	Future Directions and Perspectives.....	122
5.4.1	The role of post-translational modifications in regulating Gld2.....	122
5.4.2	Regulating miR-122 through Gld2 in health and disease .....	124
5.4.3	Identifying new cellular roles of Gld2 .....	124
5.4.4	Impact of Cid1 and Dis3L2 on <i>Schizosaccharomyces pombe</i> RNA metabolism.....	125
5.5	Conclusion .....	127
5.6	References.....	127
Appendices	.....	131
Appendix A	.....	131

Appendix B .....	136
Curriculum Vitae .....	157

## List of Tables

Table 2.1: Nucleotide addition kinetics of Gld2. ....	46
Table 3.1: Activity and RNA-binding of wildtype (WT) and phosphomimetic Gld2 variants. ....	75
Table 4.1: Select genes up- or down-regulated in either <i>S. pombe</i> $\Delta cid1$ deletion strain or <i>S. pombe</i> $\Delta dis3L2$ deletion compared to wildtype <i>S. pombe</i> . ....	104
Table 4.2: Functional enrichments in genes >1.8-fold up-regulated in <i>S. pombe</i> $\Delta dis3L2$ compared to wildtype. ....	106

## List of Figures

Figure 1.1: Central Dogma of Molecular Biology.....	3
Figure 1.2: Transcription of mRNA.....	5
Figure 1.3: mRNA degradation pathways. ....	9
Figure 1.4: Maturation pathway of miRNAs. ....	12
Figure 1.5: miRNA modifications. ....	15
Figure 1.6: Domain organization of selected TENTs. ....	16
Figure 1.7: Overview of mRNA modifications by adenylation and uridylation in humans. .....	18
Figure 2.1: Proposed catalytic activities and domain organization of Gld2. ....	37
Figure 2.2: RNA substrates of Gld2. ....	42
Figure 2.3: <i>In vitro</i> transcription of pre-let-7 RNA. ....	43
Figure 2.4: Nucleotide substrates of Gld2. ....	44
Figure 2.5: Dependence of the reaction rate on nucleotide concentration. ....	45
Figure 2.6: Phylogeny of the Tutase/PAP superfamily. ....	48
Figure 2.7: Nucleotide preference is defined by a histidine residue. ....	50
Figure 2.8: Insertion of an active site histidine into Gld2 alters nucleotide specificity....	51
Figure 3.1: Pathways regulated by Gld2 and Gld2 domain architecture. ....	63
Figure 3.2: Multiple sequence alignment of mammalian Gld2 sequences. ....	69



Figure 3.3: Gld2 is phosphorylated at S62 when incubated with EGF-stimulated HEK 293 cell extract. ....	70
Figure 3.4: Phosphomimetic Gld2 variants modulate catalytic activity and RNA binding. ....	72
Figure 3.5: Akt1 and PKA phosphorylate Gld2 at S116. ....	77
Figure 3.6: Model of Akt1-mediated regulation of Gld2.....	84
Figure 4.1: Domain structure and amino acid composition of Cid1 and Dis3L2. ....	93
Figure 4.2: Cid1 displays a promiscuous substrate range <i>in vitro</i> . ....	98
Figure 4.3: RNA uridylation of diverse RNA transcripts is found in wildtype and <i>cid1</i> deletion strains. ....	100
Figure 4.4: Changes in relative abundance of mRNAs in WT, $\Delta cid1$ , and $\Delta dis3L2$ cells. ....	102
Figure 4.5: Genes differentially expressed in <i>S. pombe</i> deletion strains compared to wildtype.....	102
Figure 4.6: Northern blot and RT-qPCR showing differential expression of genes in wildtype <i>versus</i> $\Delta cid1$ or $\Delta dis3L2$ <i>S. pombe</i> . ....	103
Figure 4.7: Search tool for the retrieval of interacting genes/proteins (STRING) diagram of RNAs with altered expression levels in a <i>dis3L2</i> deletion strain compared to wildtype. ....	105
Figure 4.8: Growth assay of <i>S. pombe</i> WT, $\Delta dis3L2$ , and $\Delta cid1$ . ....	107
Figure 4.9: Sedimentation assay of aggregated proteins in WT, $\Delta cid1$ , and $\Delta dis3L2$ deletion strains. ....	108

## List of Appendices

### Appendix A

Table A1: Primers for cloning of Gld2 into pGEX-6P-2 and mutagenesis of Gld2. ....	131
Table A2: Kinases predicted to phosphorylate Gld2 at residues S62, S69, S95, S110 and S116. ....	131
Figure A1: Purified Gld2 constructs. ....	132
Figure A2: Catalytic activity and RNA binding of Gld2 phosphomimetic variants. ....	133
Figure A3: PKA and Akt1 produce phosphorylated Gld2. ....	134
Figure A4: PKA-mediated phosphorylation down-regulates Gld2 catalytic activity. ....	134
Figure A5: Gld2 S116A does not mimic wildtype activity and binding. ....	135

### Appendix B

Table B1: Oligonucleotide primers. ....	136
Table B2: Genes up- or down-regulated in either <i>S. pombe</i> $\Delta$ <i>cid1</i> deletion strain or <i>S. pombe</i> $\Delta$ <i>dis3L2</i> deletion compared to wildtype <i>S. pombe</i> . ....	136
Table B3: Genes up-regulated at least 1.8-fold in the <i>S. pombe</i> $\Delta$ <i>cid1</i> deletion strain compared to wildtype <i>S. pombe</i> . ....	148
Table B4: Genes down-regulated at least 1.8-fold in the <i>S. pombe</i> $\Delta$ <i>cid1</i> deletion strain compared to wildtype <i>S. pombe</i> . ....	148
Table B5: Genes down-regulated at least 1.8-fold in the <i>S. pombe</i> $\Delta$ <i>dis3L2</i> deletion strain compared to wildtype <i>S. pombe</i> . ....	149

Table B6: Genes up-regulated at least 1.8-fold in the <i>S. pombe</i> $\Delta dis3L2$ deletion strain compared to wildtype <i>S. pombe</i> .....	151
Figure B1: Colour map of gene expression changes between wildtype <i>S. pombe</i> and $\Delta cid1$ or $\Delta dis3L2$ strains.....	155
Figure B2: Search tool for the retrieval of interacting genes/proteins (STRING) diagram of RNAs with altered expression levels in <i>cid1</i> and <i>dis3L2</i> deletion strains compared to wildtype.....	156

## List of Abbreviations

<sup>32</sup> P	phosphorus-32
6-FAM	6-carboxyfluorescein
Abl	tyrosine-protein kinase Abl
Ac-pre-miRNA	Ago-cleaved pre-miRNAs
Ade (A)	adenine
AGC	protein kinase A, G, and C
Ago	Argonaute
Ago2	protein argonaute 2
<i>ahal</i>	chaperone activator Aha1
Akt1	RAC-alpha serine/threonine-protein kinase or protein kinase B
ALDEx2	ANOVA-like Differential Expression version 2
ANOVA	analysis of variance
Ala (A)	alanine
AMP	adenosine monophosphate
Arg (R)	arginine
Asn (N)	asparagine
Asp (D)	aspartic acid
ATP (A)	adenosine triphosphate
BIK	Bcl2-interacting killer
<i>bip1</i>	endoplasmic reticulum heat shock protein BiP1
BSA	bovine serum albumin
<i>C. elegans</i>	<i>Caenorhabditis elegans</i>
Caf1	CCR4-associated factor 1
CAT-1	cationic amino acid transporter 1
Ccr4	glucose-repressible alcohol dehydrogenase transcriptional effector
<i>cdc1</i>	DNA polymerase delta small subunit Cdc1
<i>cdc22</i>	ribonucleoside reductase large subunit Cdc22
<i>cdc37</i>	Hsp90 co-chaperone Cdc37
<i>cdc48</i>	AAA family ATPase involved in ubiquitin-mediated protein degradation Cdc48

CDK5	cyclin dependent kinase 5
cDNA	complementary DNA
Cid1	caffeine-induced death suppressor protein 1
Cid11	caffeine-induced death suppressor protein 11
Cid12	caffeine-induced death suppressor protein 12
Cid13	caffeine-induced death suppressor protein 13
Cid14	caffeine-induced death suppressor protein 14
Cid16	caffeine-induced death suppressor protein 16
CK1 $\alpha$	casein kinase 1 isoform alpha
CK2	casein kinase 2
CK2 $\alpha$	casein kinase 2 subunit alpha
clr	centered log-ratio
CPEB	cytoplasmic polyadenylation element binding protein
Cpsf	cleavage and polyadenylation stimulating factor
cRACE	circular rapid amplification of cDNA ends
CSD	cold shock domain
CTP (C)	cytidine triphosphate
<i>dak2</i>	dihydroxyacetone kinase Dak2
Dcp1/2	mRNA-decapping enzyme subunits 1 and 2
DGCR8	DiGeorge syndrome critical region in gene 8
Dis3L2	Dis3-like 3'-5' exoribonuclease 2
DMEM	Dulbecco's Modified Eagle's Medium
DNA	deoxyribonucleic acid
DTT	dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
<i>ec11</i>	extender of the chronological lifespan protein Ecl1
EDTA	ethylenediaminetetraacetic acid
EGF	epidermal growth factor
EGTA	ethylene glycol-bis( $\beta$ -aminoethyl ester)-N,N,N',N'-tetraacetic acid
eIF4	eukaryotic translation initiation factor 4

eIF4E	eukaryotic translation initiation factor 4E
eIF4G	eukaryotic translation initiation factor 4G
EMM	Edinburgh minimal media
ERK	extracellular signal-regulated kinase
Exp-5	exportin-5
FBS	fetal bovine serum
<i>gal1</i>	galactokinase Gal1
<i>gal7</i>	galactose-1-phosphate uridylyltransferase Gal7
<i>gal10</i>	UDP-glucose 4-epimerase/aldose 1-epimerase Gal10
GFP	green fluorescent protein
Gld2	germline-development 2
Gld3	defective in germline-development protein 3
Gld4	poly(A) RNA polymerase gld-4
<i>glo1</i>	glyoxalase I
Glu (E)	glutamic acid
GMP	guanosine monophosphate
GPCR	G-protein coupled receptor
GST	glutathione S-transferase
GTP (G)	guanosine triphosphate
<i>H. sapiens</i>	<i>Homo sapiens</i>
H <sub>2</sub> O <sub>2</sub>	hydrogen peroxide
HBV	Hepatitis B virus
HBx	Hepatitis B virus X-protein
HCC	hepatocellular carcinoma
HCV	Hepatitis C virus
HEK 293	human embryonic kidney 293 cells
HEK 293T	human embryonic kidney 293 cells with SV40 T-antigen
HeLa	Henrietta Lacks cells
HepB	Hepatitis B
HepC	Hepatitis C
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

hExo	human exonuclease 1
His (H)	histidine
HO-1	heme oxygenase-1
<i>hsp78</i>	mitochondrial heat shock protein Hsp78
<i>hsp90</i>	Hsp90 chaperone
<i>hsp104</i>	heat shock protein Hsp104
HU	hydroxyurea
IgG	immunoglobulin G
IPTG	isopropyl $\beta$ -D-1-thiogalactopyranoside
$k_{\text{cat}}$	turnover number
$K_{\text{d}}$	dissociation constant
$K_{\text{M}}$	Michaelis constant
kan	kanamycin
KCl	potassium chloride
KSRP	KH-type splicing regulatory protein
LB	Luria Broth
LC-MS	liquid chromatography-mass spectrometry
let-7	lethal-7
leu	leucine
Lin28A	Lin-28 homolog A
Lsm1-7	like Smith 1-7
Lys (K)	lysine
<i>M. musculus</i>	<i>Mus musculus</i>
MDA-MB-231	M.D. Anderson metastatic breast adenocarcinoma 231 cells
MgCl <sub>2</sub>	magnesium chloride
MGV	mean gray value
miR-122	microRNA 122
miR-215	microRNA 215
miR-564	microRNA 564
miRISC	miRNA bound RISC
miRNA	microRNA

MOPS	3-( <i>N</i> -morpholino)propanesulfonic acid
mRNA	messenger RNA
mRNP	messenger ribonucleoprotein
MS	mass spectrometry
MS/dd-MS2	mass spectrometry/data dependent-mass spectrometry
MT	mitochondrion
N	any nucleotide
<i>n</i>	sample size
Na <sub>2</sub> VO <sub>4</sub>	sodium orthovanadate
NaCl	sodium chloride
NCBI	National Center for Biotechnology Information
ncRNA	non-coding RNA
nt	nucleotide
NTR	nucleotidyltransferase
NTP	nucleoside triphosphate
OD <sub>600</sub>	optical density at 600 nm
orf	open reading frame
P bodies	processing bodies
<i>p</i> value	probability value
PAN	poly(A) nuclease
Pan2-Pan3	poly(A) specific ribonuclease subunits Pan2 and Pan3
PABP	poly(A) binding protein
PAP	poly(A) polymerase
PAPD1	poly(A) polymerase associated domain containing 1
PAPD4	poly(A) polymerase associated domain containing 4
Parn	poly(A) specific ribonuclease
PBS-T	1 x phosphate-buffered saline 1% Tween
PCR	polymerase chain reaction
PDK1	3-phosphoinositide-dependent protein kinase 1
pGld2	phosphorylated Gld2
<i>pex22</i>	peroxisomal membrane protein Pex22 (predicted)



<i>pgi1</i>	glucose-6-phosphate isomerase (predicted)
PGK1	phosphoglycerate kinase 1
PI3K	phosphoinositide 3-kinase
PKA	protein kinase A
<i>plg7</i>	phospholipase A2, PAF family homolog
PMSF	phenylmethylsulfonyl fluoride
Pneumo-G	pneumoviridae
Pol II	RNA polymerase II
Pol III	RNA polymerase III
PP2A	protein phosphatase 2A
ppAkt1	doubly phosphorylated Akt1
pre-let-7a	precursor lethal-7a
pre-miRNA	precursor microRNA
pre-mRNA	precursor mRNA
pri-miRNA	primary microRNA
PRM	parallel-reaction monitoring
PRR	proline-rich region
pSer	phosphoserine
<i>psi1</i>	DNAJ domain protein Psi1
PTM	post-translational modification
PVDF	polyvinylidene difluoride
QKI-7	Quaking
R	replicate
RISC	RNA-induced silencing complex
RNA	ribonucleic acid
RNase	ribonuclease
RNB	ribonuclease domain
<i>rpp0</i>	60S acidic ribosomal protein Rpp0 (predicted)
<i>rpt1</i>	19S proteasome regulatory subunit Rpt1 (predicted)
<i>rpt3</i>	19S proteasome regulatory subunit Rpt3 (predicted)
<i>rpt6</i>	19S proteasome regulatory subunit Rpt6 (predicted)

RRM	RNA recognition motif
RT	room temperature
RT-qPCR	quantitative reverse transcription polymerase chain reaction
rRNA	ribosomal RNA
S1	nonspecific RNA binding domain
S2 cells	Schneider 2 cells
<i>S. pombe</i>	<i>Schizosaccharomyces pombe</i>
S-M	Synthesis-Mitosis
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
Ser (S)	serine
snRNA	small nuclear RNA
snRNP	small nuclear ribonucleoprotein
<i>ssa1</i>	heat shock protein Ssa1 (predicted)
<i>ssa2</i>	heat shock protein Ssa2
<i>ssc1</i>	mitochondrial heat shock protein Hsp70
<i>sti1</i>	chaperone activator Sti1 (predicted)
STRING	Search Tool for the Retrieval of Interacting Genes/proteins
<i>suc22</i>	ribonucleotide reductase small subunit Suc22
Tap	nuclear RNA export factor 1
<i>tcg1</i>	single-stranded telomeric binding protein Tgc1
<i>tdh1</i>	glyceraldehyde-3-phosphate dehydrogenase Tdh1
TENT	terminal RNA nucleotidyltransferase
TENT2	terminal RNA nucleotidyltransferase 2
TENT3A	terminal RNA nucleotidyltransferase 3A
TENT3B	terminal RNA nucleotidyltransferase 3B
TEV	tobacco etch virus
<i>thf1</i>	C1-5,6,7,8-tetrahydrofolate (THF) synthase, trifunctional enzyme Thf1
Thr (T)	threonine
Tris-HCl	tris(hydroxymethyl)aminomethane hydrochloride

tRNA	transfer RNA
<i>trx1</i>	cytosolic thioredoxin Trx1
TTP (T)	thymidine triphosphate
TUT/Tutase	terminal uridylyltransferase
TUT1	terminal uridylyltransferase 1
TUT2	terminal uridylyltransferase 2
TUT4	terminal uridylyltransferase 4
TUT7	terminal uridylyltransferase 7
tWT	treated wildtype
U	uridine
U6 snRNA	U6 small nuclear RNA
<i>ubp15</i>	ubiquitin C-terminal hydrolase Ubp15
ura	uracil
UTP (U)	uridine triphosphate
UTR	untranslated region
$v_o$	initial velocity
$V_{\max}$	maximum velocity
Val (V)	valine
<i>wos2</i>	p23 homolog, predicted co-chaperone Wos2
WT	wildtype
<i>X. laevis</i>	<i>Xenopus laevis</i>
Xrn1	5'-3' exoribonuclease 1
YES	yeast extract with supplements
YPD	yeast extract peptone dextrose
Zcchc6	zinc finger CCHC-type containing 6
Zcchc11	zinc finger CCHC-type containing 11

# Chapter 1

## 1 Introduction

### 1.1 Central Dogma of Molecular Biology

In 1958 Francis Crick defined the Central Dogma of molecular biology as the flow of information from DNA to RNA to proteins<sup>1</sup> (Figure 1.1). Deoxyribonucleic acid (DNA) contains the genetic information in the form of a double-stranded helix<sup>2</sup>, which serves as a blueprint for life. The genome contains all the genetic information for an organism, encoded in a four-letter DNA code represented by the nucleotide bases adenine, thymine, cytosine, and guanine. Genes are sequences of DNA that are transcribed to ribonucleic acid (RNA) polymers, including messenger RNAs (mRNAs) and non-coding RNAs (ncRNAs). RNA polymerase recognizes and binds to specific DNA sequences and uses the DNA as a template to synthesize complementary RNA strands<sup>3</sup>. Of the different RNA polymerases, RNA polymerase II (Pol II) transcribes mRNAs. The transcriptome includes the collection of all mRNAs in the cell. These RNA messages are read or translated by the ribosome to synthesize proteins of specific sequences as defined in the mRNA<sup>3,4</sup>.

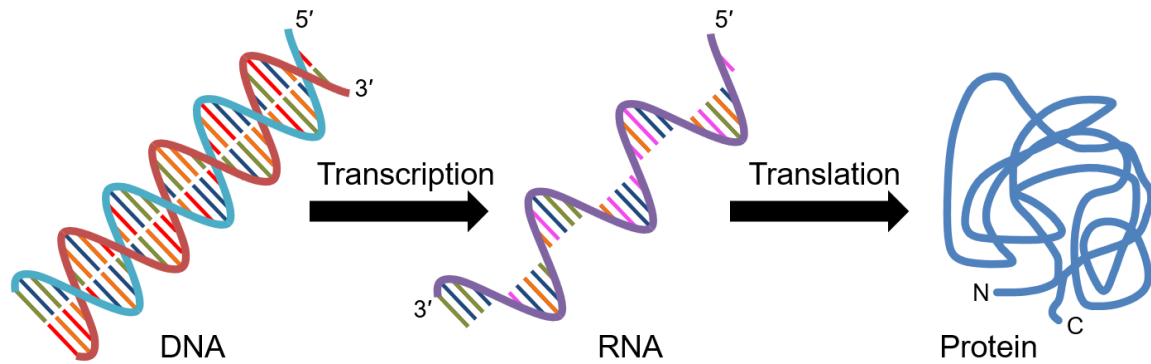
The mRNA carries the protein-encoding information in a four-letter code, similar to DNA. The thymine bases found in DNA are replaced by uracil in RNA transcripts, and the RNA bases are linked by a sugar phosphate backbone that contains ribose sugars instead of deoxyribose sugars found in DNA. The mRNA is recognized as a substrate by the ribosome in synthesizing all cellular proteins<sup>5</sup>. In protein synthesis, often referred to as mRNA translation, the mRNA is read in nucleotide (nt) triplets known as codons. RNAs encode up to 64 different codons, which either correspond to an amino acid or serve as a stop signal (stop codon) that terminate translation.

The genetic code is redundant and there can be more than one codon per amino acid; a total of 20 canonical amino acids are encoded by 61 codons. Three codons (UAG, UAA, UGA) serve as stop codons<sup>6</sup>. In some organisms, including *Escherichia coli* and

humans, UGA is recoded to the 21<sup>st</sup> amino acid selenocysteine in a small set of genes<sup>7-9</sup>. In a certain species of archaea and bacteria, UAG is reassigned to pyrrolysine, the 22<sup>nd</sup> genetically encoded amino acid<sup>7,10</sup>.

Small RNA molecules, transfer RNA (tRNA), serve as adaptors or decoders of the three-letter nucleic acid code into amino acids. tRNAs contain an anti-codon that is complementary to a codon or set of codons. Each tRNA species is charged with the cognate amino acid by the aminoacyl-tRNA synthetases<sup>11</sup>. tRNAs encode identity elements that allow aminoacyl-tRNA synthetases to specifically recognize tRNAs and ligate them with the correct amino acid<sup>6</sup>. Following elongation factor binding, the aminoacyl tRNA anti-codon base-pairs with the corresponding mRNA codon in the ribosome. The decoding process allows the correct amino acid, as specified in the mRNA, to be added to the growing protein chain. The ribosome, comprised of ribosomal RNAs (rRNAs) and protein subunits, catalyzes the peptidyl transfer reaction that adds the next amino acids to a growing peptide. The protein product is completed once a stop codon is reached and, following recognition by release factors, the ribosome dissociates and is prepared for the next round of translation<sup>12</sup>.

The primary structure of a protein is determined by its amino acid sequence. As the peptide grows, it forms a secondary structure. The secondary structure is influenced by the amino acid sequence and intramolecular interactions are formed between the amino acid residues and backbone atoms. At the secondary structure level, protein conformations include  $\alpha$ -helices,  $\beta$ -sheets, and random coils. Most proteins then fold into a ground state or low energy tertiary structure referred to as the native state. Quaternary structures may form when different proteins or subunit monomers that have already formed tertiary structures interact with each other to form higher order complexes<sup>13</sup>.



**Figure 1.1: Central Dogma of Molecular Biology.**

The schematic diagram depicts transcription and translation. For protein coding genes, their DNA sequences are transcribed to mRNAs. The translation machinery then decodes the mRNA into a specific protein sequence (N, N-terminal; C, C-terminal). The nucleotide bases are colour-coded: green, adenine; red, thymine; dark blue, guanine; orange, cytosine; pink, uracil.

## 1.2 Regulation of mRNA

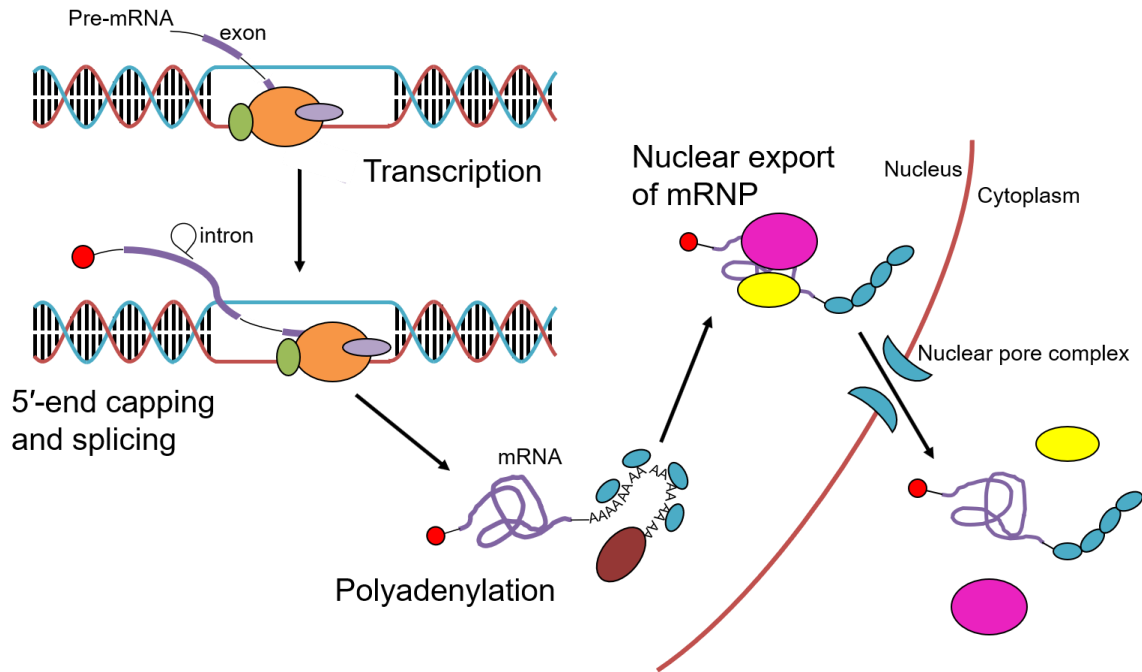
### 1.2.1 mRNA transcription in eukaryotes

Transcription of mRNA starts with Pol II and transcription factors binding to a DNA promoter and opening the double-stranded helix to form a transcription bubble<sup>3,4</sup> (Figure 1.2). During transcription elongation, Pol II moves along the DNA template and adds complementary RNA nucleotides to synthesize the growing mRNA transcript. Transcription ends when the complex reaches the transcription terminator and the resulting RNA is known as a precursor mRNA (pre-mRNA)<sup>14</sup>.

The pre-mRNA undergoes a series of modifications before becoming a mature mRNA. The processing starts as the pre-mRNA is formed with the addition of a 7-methylguanosine cap on the 5'-end<sup>15,16</sup>. RNA triphosphatase first hydrolyzes the 5'-triphosphate to a diphosphate; RNA guanylyltransferase then adds a guanosine monophosphate (GMP). Finally, RNA guanine-7 methyltransferase will methylate the guanosine base at position N7<sup>16</sup>.

The mRNA 5'-cap is important for many events including mRNA stability, processing, nuclear export, and translation<sup>16</sup>. The pre-mRNA also require splicing to remove the non-protein coding introns<sup>15,17</sup>. The spliceosome is a complex of five small nuclear ribonucleoproteins (snRNPs) that catalyzes the splicing event and requires specific sequences within the intron for efficient splicing<sup>17,18</sup>. Once the polyadenylation signal sequence and the downstream GU-rich motif are transcribed, the mRNA is cleaved between those two regions at the cleavage site and transcription is terminated<sup>19-21</sup>. The 3'-end of the transcript is polyadenylated by poly(A) polymerase (PAP) and the resulting poly(A) tail is bound by poly(A) binding proteins (PABPs)<sup>19-21</sup>.

For the mRNA to leave the nucleus, various mRNA-binding proteins come together to form a messenger ribonucleoprotein (mRNP) export complex. The complex is exported out of the nucleus by nuclear pore complexes and will undergo changes in the cytoplasm to release the mRNA and prevent the mRNA from returning to the nucleus<sup>22</sup>.



**Figure 1.2: Transcription of mRNA.**

The double-stranded DNA (blue and dark red strands) is bound by RNA polymerase II (orange oval, Pol II) and transcription factors such as transcription factor IIB (green oval) and transcription factor IID (purple oval) to form a transcription bubble. The pre-mRNA (purple and black line) is transcribed as Pol II moves along the DNA strand. The introns (black lines) are spliced out and a 7-methylguanosine cap (red circle) is added to the 5'-end as transcription occurs. Once transcription is terminated, a poly(A) tail is added by a poly(A) polymerase (dark red oval) to the 3'-end of the mRNA and bound by poly(A) binding proteins (light blue ovals). Various mRNA binding proteins such as the transcription-export complex (pink oval) and nuclear RNA export factor 1 and p15 heterodimer (Tap-p15, yellow oval) will bind the mRNA to form a messenger ribonucleoprotein (mRNP) export complex to export the mRNA out of the nucleus. Once in the cytoplasm, the mRNA binding proteins will undergo conformational changes to release the mRNA and this prevents the mRNA from returning to the nucleus.

### 1.2.2 Role of the poly(A) tail

The poly(A) tail plays an important role in the life cycle of mRNA. The poly(A) is added to the 3'-end of the mRNA during RNA processing after cleavage by nuclear PAPs and is important for events such as mRNA nuclear export, stability, and translation<sup>19–21</sup>. Transcripts lacking a poly(A) tail are confined in the nucleus<sup>23</sup>. More recent studies show that the poly(A) tail induces mRNA export by contributing to the length of the mRNA



transcript and the export can actually be inhibited if the poly(A) tail is too long<sup>24</sup>. The poly(A) tail is required to bind PABP and it is the PABPs that confer stability to the mRNA<sup>25,26</sup>. Indeed, a mRNA transcript without a poly(A) tail can be stabilized if PABP is tethered to the transcript<sup>25</sup>. The poly(A) tail is also involved in translation through PABP<sup>26–29</sup>. PABP binds eukaryotic translation initiation factor 4G (eIF4G), a component of the eukaryotic translation initiation factor 4 (eIF4) translation initiation complex, that then binds to the 5'-end cap-binding eukaryotic translation initiation factor 4E (eIF4E)<sup>27–30</sup>. This allows the mRNA to take on a circular form to promote translation and, possibly, the stability of the mRNA by preventing decapping at the 5'-end and deadenylation at the 3'-end<sup>26,28,29</sup>.

While initial polyadenylation occurs in the nucleus by canonical PAPs, the poly(A) tails of mRNAs can be further extended by non-canonical PAPs in the cytoplasm to facilitate translation<sup>31–34</sup>. Polyadenylation of mRNAs by mammalian homologs of the non-canonical PAP germline-development 2 (Gld2) has been shown to enhance mRNA translation in *Xenopus* oocytes<sup>31</sup>. Another study showed that dormant mRNAs in the cytoplasm have shortened poly(A) tails and require polyadenylation to be active for translation<sup>32,34</sup>. The extended poly(A) tail recruits additional PABP and the associated proteins required for translation<sup>26–29,34</sup>. The activation of dormant mRNAs is commonly observed during development with mRNAs involved in meiosis and mitosis<sup>34</sup>. In *Caenorhabditis elegans*, the absence of Gld2 led to abnormal germline cell growth<sup>35,36</sup>. The poly(A) tails of maternal mRNAs during oocyte maturation are extended to promote translation<sup>33</sup>, and the absence of the extended tails can be observed in mature oocytes and early-stage embryos when the non-canonical PAP is deleted<sup>37</sup>.

### 1.2.3 5'-3' and 3'-5' degradation of mRNAs

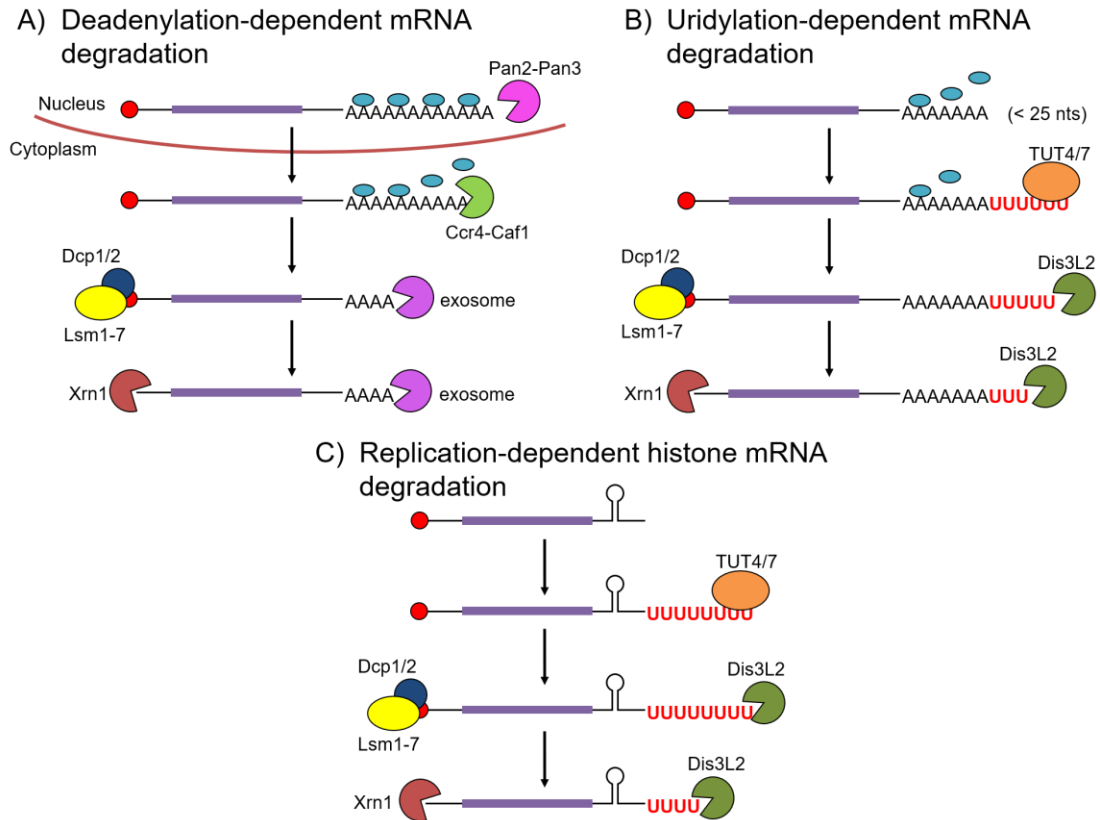
Degradation of mRNA can occur in both the 3'-5' and 5'-3' directions (Figure 1.3A). In eukaryotes, deadenylation-dependent degradation of the poly(A) tail is initially carried out by the poly(A) specific ribonuclease subunits Pan2 and Pan3 (Pan2-Pan3) of the poly(A) nuclease (PAN) deadenylase complex in the nucleus. The complex has been

observed to slightly trim the poly(A) tail to approximately 70-80 nts in *Saccharomyces cerevisiae* and 110 nts in mammalian cells before the mRNA transcript is exported out of the nucleus<sup>26,38-42</sup>. Most of the deadenylation is carried out in the cytoplasm by the glucose-repressible alcohol dehydrogenase transcriptional effector and CCR4-associated factor 1 (Ccr4-Caf1) complex<sup>43,44</sup>. This triggers the dissociation of PABP, leaving the 3'-end exposed for degradation by the exosome. Deadenylation, however, initially triggers the 5'-3' decay pathway. The loss of PABP will de-circularize the mRNA, exposing the 5'-end cap. This leads to decapping by the mRNA-decapping enzyme subunits 1 and 2 (Dcp1/2) and like Smith 1-7 (Lsm1-7) decapping complexes. Following decapping, the RNA is degraded by the 5'-3' exoribonuclease 1 (Xrn1)<sup>45</sup>.

An alternate deadenylation-independent RNA degradation pathway characterized by the addition of uridine residues to polyadenylated mRNAs was identified in *Schizosaccharomyces pombe*<sup>46</sup>. A similar, uridylation-dependent RNA degradation process has also been found in human cells<sup>46-50</sup> (Figure 1.3B). mRNAs with poly(A) tails less than 25 nts lose the protection of PABPs and are polyuridylated by human terminal uridylyltransferase 4 (TUT4, Zcchc11) and terminal uridylyltransferase 7 (TUT7, Zcchc6), leading to degradation by Dis3-like 3'-5' exonuclease 2 (Dis3L2)<sup>47</sup>. Dis3L2 is a 3'-5' cytoplasmic exonuclease, that specifically degrades 3'-uridylated RNA species, but is not part of the exosome<sup>51</sup>. Dis3L2 recognizes polyuridylated mRNAs<sup>51</sup> and miRNAs<sup>52</sup> and catalyzes their 3'-5' degradation. Mutations in Dis3L2 have been linked to Perlman syndrome and formation of Wilms' tumor in children<sup>53</sup>. Due to its ability to recognize and degrade any RNAs that are polyuridylated, Dis3L2 is believed to play a key role in maintaining cellular RNA homeostasis<sup>53</sup>. The poly(U) tails of mRNAs are also recognized by the Dcp1/2 and Lsm1-7 decapping complexes<sup>54</sup>, leading to 5'-3' degradation by Xrn1<sup>45</sup>. Interestingly, depletion of Lsm1 from the Lsm1-7 complex leads to an increase in mRNAs with oligo(U) tails, suggesting that mRNAs with 3'-uridine residues are first recognized by the decapping complexes and subjected to 5'-3' decay before 3'-5' degradation<sup>47</sup>.

This uridylation-dependent decay pathway is the major decay pathway for histone mRNAs (Figure 1.3C). These mRNAs encode a unique 3'-stem-loop structure, which is

uridylated at the end of S-phase to initiate degradation<sup>49</sup>. The terminal uridylyltransferase TUT4 is the key player in histone mRNA uridylation; knockdown of TUT4 reduces histone mRNA uridylation and increases overall histone mRNA abundance<sup>50</sup>. In addition, TUT7 was shown to uridylate mature histone mRNA. TUT7 interacts with the exonuclease 3'-hExo to maintain a length of 3 nts after the stem-loop structure and both enzymes take part in the initial degradation of the histone mRNA<sup>55</sup>. When the histone mRNA 3'-end stem-loop structure is polyuridylated, the RNA is subsequently degraded by both the 5'-3'<sup>56</sup> and Dis3L2 3'-5' degradation pathways<sup>49</sup>.



**Figure 1.3: mRNA degradation pathways.**

Three major RNA degradation pathways in eukaryotes. **A)** Deadenylation-dependent degradation of mRNAs begins with the Pan2-Pan3 deadenylase complex slightly trimming the poly(A) tail and continues in the cytoplasm with extensive deadenylation by Ccr4-Caf1. This causes PABPs (light blue ovals) to dissociate from the tail and will trigger 3'-5' degradation by the exosome as well as 5'-end decapping (5'-cap, red circle) by Dcp1/2 and Lsm1-7 and 5'-3' degradation by Xrn1. **B)** Uridylation-dependent degradation of mRNAs is initiated when poly(A) tails less than 25 nts lose the protection of PABPs and are polyuridylated by TUT4 or TUT7. The poly(U) tail is recognized by the U-specific exonuclease Dis3L2 for 3'-5' degradation. The poly(U) tail also triggers decapping by Dcp1/2 and Lsm1-7 and 5'-3' degradation by Xrn1. **C)** Replication-dependent histone mRNAs encode a 3'-end stem-loop structure instead of a poly(A) tail. The stem-loop acts as a *cis* element for mRNA degradation and is bound by stem-loop binding protein (SLBP) and other proteins (not shown in figure)<sup>49</sup>. The mRNAs are degraded through the uridylation-dependent degradation pathway.

## 1.3 miRNAs: Functions in Gene Expression

MicroRNAs (miRNAs) are small, single-stranded, regulatory RNAs that were discovered in the 1990s<sup>57,58</sup>. They are approximately 19-24 nts in length and take part in degrading mRNAs to suppress protein synthesis. miRNAs regulate gene expression by binding to complementary sequences in the target mRNA<sup>58</sup>. In animals, miRNAs regulate mRNAs through the RNA-induced silencing complex (RISC). As miRNAs in animals are typically not perfect complements to their mRNA targets, the mRNA is not degraded by endonucleolytic cleavage. The mRNA will be degraded and/or translationally suppressed through other methods (section 1.3.1)<sup>59,60</sup>. miRNAs play important roles in critical cellular pathways such as proliferation and apoptosis, ultimately affecting the well-being of the entire organism<sup>58,61</sup>. Correspondingly, de-regulation of miRNAs in humans has been shown to result in cardiovascular diseases<sup>62,63</sup> and other diseases such as cancer<sup>58,61,64,65</sup> and diabetes<sup>66</sup>.

### 1.3.1 Role of miRNAs in regulating gene expression

Many models have been put forth to explain translation repression<sup>5,67,68</sup>. Some models suggest that miRNA bound RISC (miRISC) inhibits translation initiation by competing with the 5'-cap binding protein eIF4E to bind the mRNA 5'-cap, preventing assembly of the 40S and 60S ribosomal subunits on the mRNA, or inhibiting formation of the translation initiation complex<sup>5,67,69-71</sup>. miRNAs also inhibit translation during elongation where miRISC causes early release of the ribosomes from the mRNA<sup>72</sup>. The miRISC has been suggested to prevent circularization of the mRNA, preventing translation from starting<sup>67,68</sup>. The miRISC can also sequester the mRNAs in processing bodies (P bodies), preventing translation of the mRNA<sup>67,73</sup>. The partial complementarity of the miRNA to its target mRNA directs the mRNA towards decay pathways instead of mRNA cleavage by RISC by recruiting the Ccr4-Caf1 complex for deadenylation<sup>60,67,74</sup>. Deadenylation will lead to deadenylation-dependent 3'-5' decay and 5'-3' decay through mRNA decapping as described in section 1.2.3.

### 1.3.2 miRNA maturation pathway

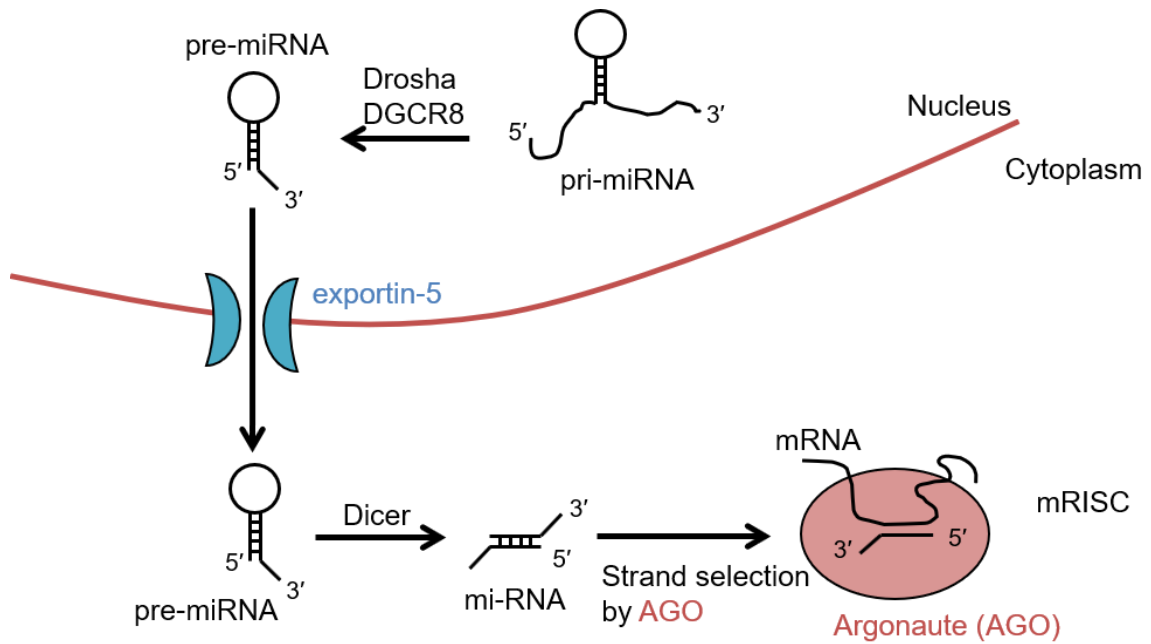
In animals, miRNA maturation takes place in the nucleus and the cytoplasm<sup>67,75</sup> (Figure 1.4). Most miRNAs are transcribed by Pol II as they are encoded within the sequences of coding and noncoding RNA transcripts<sup>67,75,76</sup>. miRNAs are mostly encoded within the introns but can be found in the exons<sup>67,75</sup>.

In addition, some miRNAs such as miR-23a/27a/24-2 have their own promoters that are recognized by Pol II<sup>76</sup>. A small subset of miRNAs with their own promoters, such as C19MC miRNAs and miR-886, are transcribed by RNA polymerase III (Pol III)<sup>77,78</sup>. The resulting RNA is known as a primary miRNA (pri-miRNA) and takes the form of a long RNA strand with a hairpin-loop structure. Each hairpin-loop structure corresponds to one miRNA and clusters of miRNA sequences will have multiple hairpin-loop structures on the RNA strand. Drosha, a class II RNase III enzyme, with the help of a cofactor (DiGeorge syndrome critical region in gene 8 (DGCR8) in humans) processes the pri-miRNA to the hairpin loop structure of about 60-70 nts<sup>79,80</sup>. Drosha recognizes a large loop ( $\geq 10$  nts) at the end of the hairpin structure as well as structural features of the hairpin stem<sup>81</sup>.

This processed form is known as the precursor miRNA (pre-miRNA). The pre-miRNA is transported out of the nucleus by exportin-5 (Exp-5), a nuclear transport receptor<sup>82,83</sup>. Once in the cytoplasm, the pre-miRNA is cleaved by the RNase III enzyme Dicer to form a double-stranded mature miRNA that is approximately 22 nts long<sup>67,75,84,85</sup>.

The double-stranded miRNA is then loaded onto the Argonaute (Ago) proteins to form the RISC. The miRNA will undergo strand selection where the passenger strand is degraded and the guide miRNA strand is left in the RISC<sup>86,87</sup>. Thermodynamic stability is an important factor in strand selection. The double-stranded miRNA is unwound and the miRNA strand with the weakest 5'-end binding is chosen as the guide strand. Thus, there is a bias to have an uracil residue on the 5'-end of the guide strand as this contributes to weaker binding. However, many other factors are involved in selecting the guide strand. Some examples are the orientation of the double-stranded duplex when it is loaded onto

RISC, the cell type, and proteins such as Ago, whose functions in the process remain unknown<sup>86</sup>. The RISC scans the mRNA pool to find the target mRNA through base-pairing interactions between the guide miRNA and mRNA 3'-untranslated region (UTR) to suppress translation of the mRNA target<sup>60,67</sup>.



**Figure 1.4: Maturation pathway of miRNAs.**

Schematic diagram depicting the miRNA maturation process. The primary miRNA (pri-miRNA) is transcribed from the DNA and processed by Drosha and DGCR8 to a single hairpin loop structure with a 2 nucleotide (nt) 3'-overhang, known as the precursor miRNA (pre-miRNA). The pre-miRNA is exported out of the nucleus by the nuclear transport receptor protein exportin-5 (blue) and undergoes processing by Dicer to form a double-stranded mature miRNA. The miRNA then undergoes strand selection by Argonaute (maroon) and incorporation into the RNA-induced silencing complex (RISC) where it will then bind its target mRNA to form the mRNA RISC (mRISC).

### 1.3.3 Regulation of miRNA through untemplated nucleotide addition

Recent high-throughput sequencing studies revealed the presence of untemplated nucleotide additions to the 3'-termini of nearly 40% of miRNAs<sup>88,89</sup>. As part of miRNA

maturation and degradation, untemplated uridine<sup>90,91</sup> and adenine<sup>92</sup> residues are added to pre-miRNAs and mature miRNAs. These additional nucleotides present an efficient means to control the level of active miRNA in a cell. Nucleotide additions to the 3'-end of miRNAs are widespread in eukaryotes and catalyzed by several terminal RNA nucleotidyltransferases (TENTs)<sup>48,90,93–102</sup>.

While single nucleotide additions play a crucial role in miRNA maturation, recent reports suggest that monoadenylation of miRNAs also leads to increased miRNA stability<sup>92,103</sup>. Monoadenylation of miRNAs are carried out by the same enzymes that polyadenylate mRNAs. The minimal TENT Gld2 is able to polyadenylate mRNAs in the cytoplasm as well as monoadenylate miRNAs<sup>54,92</sup>. Gld2, thus, promotes miRNA stabilization: Gld2-mediated monoadenylation stabilizes miRNA miR-122 transcripts in the liver<sup>104</sup> and human fibroblasts<sup>92</sup> and plays a role in the translational regulation of p53<sup>105,106</sup> (Figure 1.5B). In addition, Gld2 knockout mice show decreased miRNA monoadenylation<sup>107</sup>.

Uridylation of miRNAs can occur on pre-miRNAs and mature miRNAs. In the miRNA maturation pathway, a small group of pre-miRNAs known as Group II pre-miRNAs are processed to a single nucleotide overhang on the 3'-end and require the addition of a second nucleotide to generate the required 2 nt 3'-overhang<sup>108</sup>. *In vitro* uridylation assays and *in vivo* knockdowns showed that the humans TENTs TUT4, TUT7, and Gld2 act redundantly in monouridylating the Group II pre-let-7 pre-miRNA<sup>90</sup> (Figure 1.5A). This single nucleotide addition leads to the generation of a 2 nt 3'-overhang that would allow recognition and processing by Dicer, followed by the formation of the multiprotein RISC (Figure 1.5A), which leads to RNA silencing and cleavage by argonaute<sup>75,90,110,111</sup>.

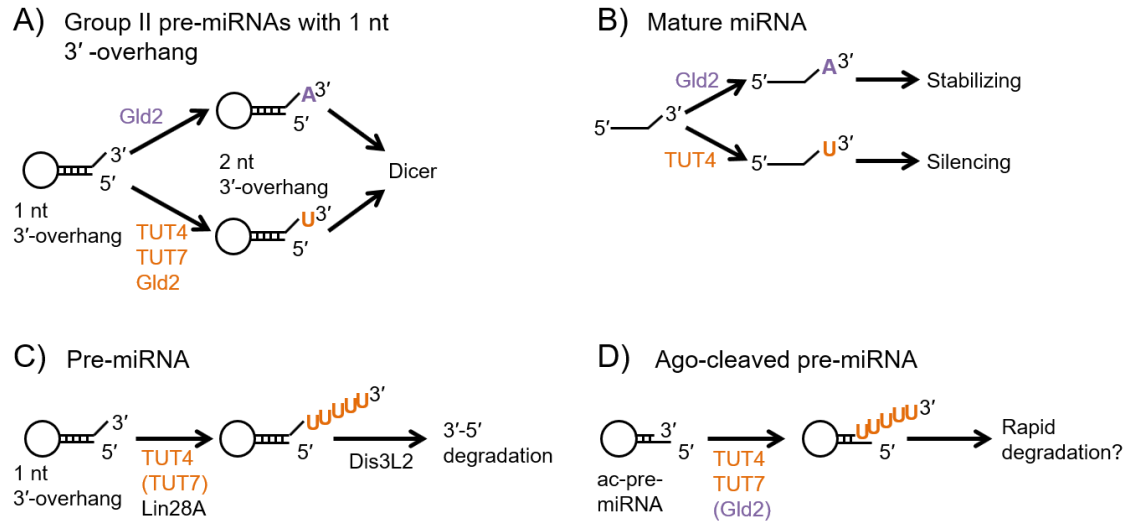
Polyuridylation, as opposed to monouridylation, of Group II pre-miRNAs leads to degradation by Dis3L2 (Figure 1.5C). In general, polyuridylation requires the presence of an accessory protein to enhance RNA binding. Lin-28 homolog A (Lin28A) is a RNA binding protein that recruits TUT4, and to a lesser extent TUT7, via its zinc knuckle domain to polyuridylate pre-let-7<sup>91,112,113</sup>. The full-length Lin28A protein recognizes the



let-7 stem-loop and a conserved GGAG motif near the 3'-end<sup>114</sup>. Lin28A suppresses the biogenesis of the tumor suppressor let-7 to prevent stem cell differentiation and protein levels slowly decrease during development to promote differentiation<sup>115</sup>. Lin28A is an oncogene as high levels correlate with a poor prognosis due to the suppression of mature let-7, leading differentiated cells to become pluripotent stem cells<sup>115,116</sup>. Lin28A expression is, however, tissue-specific after development and can be detected in tissues such as the placenta<sup>116</sup>.

A recent structure of Ago-cleaved pre-miRNAs (ac-pre-miRNAs) contains a 5'-overhang due to trimming of the 3'-end by Ago2 or an unknown nuclease<sup>117</sup> (Figure 1.5D). The overhang structures are thought to stem from imperfect pre-miRNA processing. The presence of the overhangs allows TUT4 and TUT7 binding in a Lin28A-independent manner, leading to polyuridylation and rapid RNA degradation<sup>117</sup>. In addition to pre-miRNA polyuridylation, which prevents miRNA maturation, the TENT terminal uridylyltransferase 1 (TUT1, Star-PAP) was shown to polyuridylate mature miRNAs, marking them for degradation<sup>118</sup>.

TENTs play an essential role in pre-miRNA maturation and are further involved in miRNA degradation. These mechanisms control the amount of miRNA transcripts in the cell, but do not directly influence their specific activity. However, a recent report suggests that TUT4 directly controls miR-26a activity<sup>93,94</sup>. The addition of a single 3'-uridine to mature miR-26a, a miRNA involved in cytokine expression, silences miR-26a activity without altering miRNA abundance<sup>94</sup> (Figure 1.5B). Deep sequencing of TUT4-depleted mouse livers revealed a decrease in the number of uridylated miRNAs but not a decrease in miRNA abundance<sup>93</sup>. TUT4 and TUT7 were shown to be redundant in monouridylating the 3'-end of mature miRNAs<sup>95</sup>. Depletion of one or both enzymes did not affect mature miRNA levels, suggesting that monouridylation leads to silencing and not degradation<sup>95</sup>. These uridylation events were observed for several distinct miRNA species such as let-7a, let-7g, and miR-10a, likely representing a general and direct mechanism to regulate miRNA activity<sup>93,95</sup>.



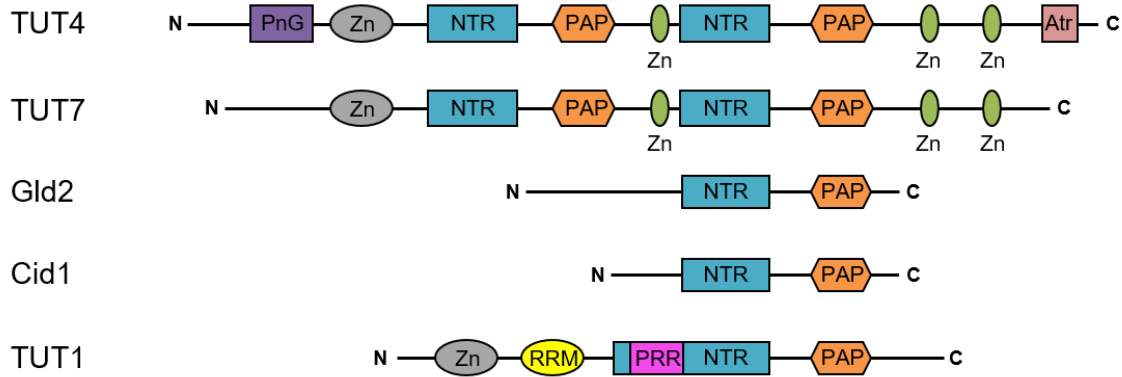
**Figure 1.5: miRNA modifications.**

The miRNA is regulated at various points during its maturation through untemplated nucleotide additions. **A)** Group II pre-miRNAs have a 1 nt 3'-overhang and can be monouridylated or monoadenylated to produce the 2 nt 3'-overhang recognized by Dicer. **B)** Mature miRNAs are regulated by Gld2 mediated 3' monoadenylation, enhancing their stability, or by TUT4 catalyzed 3' monouridylation, leading to silencing. **C)** Group II pre-miRNAs are polyuridylated by TUT4 or TUT7 in collaboration with Lin28A, leading to 3'-5' degradation by Dis3L2. **D)** Ago-cleaved pre-miRNAs (ac-pre-miRNAs) with a 5'-overhang are polyuridylated by TUT4, TUT7, or Gld2, possibly leading to rapid degradation.

## 1.4 Terminal RNA Nucleotidyltransferases

Many terminal RNA nucleotidyltransferase (TENT) homologs have a biochemically characterized biological function and are categorized as PAPs or terminal uridylyltransferases (Tutases), yet some homologs remain uncharacterized. Tutases have been identified as novel key players in mRNA turnover<sup>47,48,119–121</sup>, with additional roles in pre-mRNA<sup>122</sup>, non-adenylated histone mRNA<sup>49,50,55,56</sup>, and miRNA metabolism<sup>93–95,112,123</sup>. Humans encode three Tutases (TUT1<sup>124,125</sup>, TUT4<sup>47,50,90,93–95,126</sup>, TUT7<sup>47,55,90,95,117</sup>) and two non-canonical PAPs (Gld2<sup>92,104–106,127</sup>, PAPD1<sup>128,129</sup>). In fission yeast, several PAPs (Cid11<sup>130</sup>, Cid12<sup>131</sup>, Cid13<sup>132</sup>, Cid14<sup>133,134</sup>) and Tutases (Cid1<sup>135,136</sup>, Cid16<sup>137</sup>) have been identified. A variety of TENTs have been characterized, and we highlighted these discoveries in a recent review<sup>138</sup>. In the following sections, I

will focus on the TENTs relevant to this thesis: *S. pombe* TENT caffeine induced death protein 1 (Cid1) and the human TENTs TUT1, TUT4, TUT7, and Gld2.



**Figure 1.6: Domain organization of selected TENTs.**

TENTs share the major catalytic domains but differ in their RNA binding domains. Ntr, Nucleotidyltransferase domain; PAP, Poly(A) polymerase associated domain; Zn, Zinc finger; PnG, PneumoG domain; Atr, Atrophin-like domain; RRM, RNA recognition motif; PRR, proline-rich region.

#### 1.4.1 Adenylation vs. uridylation

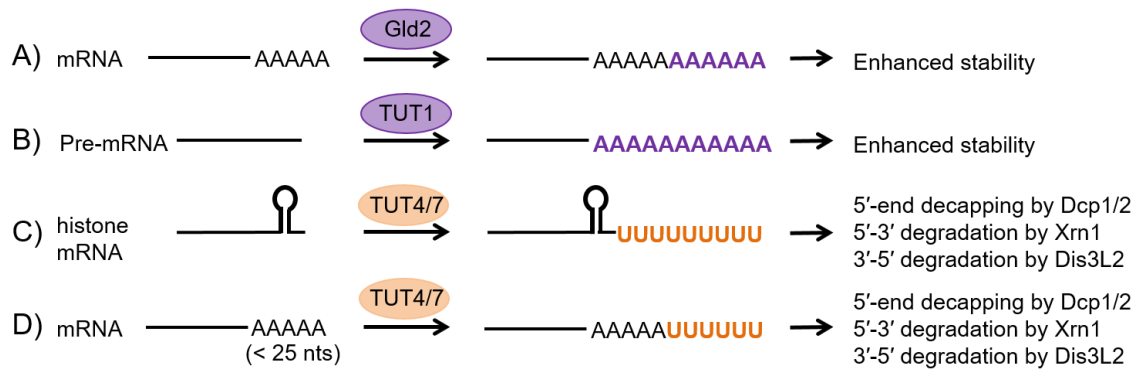
Enzymes with uridylyl- and adenylyltransferase activity are closely related in amino acid sequence and similar in their protein domain architecture and structure. Significant sequence similarity between PAPs and Tutases has obscured identification of sequence or structural features that differentiate a true adenylyltransferase from a true uridylyltransferase.

The biochemical characterization of the non-canonical PAP Cid1 from *S. pombe* revealed an unexpected *in vitro* and *in vivo* Tutase activity, introducing Cid1 as the first identified uridylyltransferase<sup>121</sup>. The catalytic core of Cid1 consists of a nucleotidyltransferase domain and PAP-associated domain, much like its adenylyltransferase counterpart Gld2<sup>109,136</sup>. This homologous catalytic core is also found in TUT4 and TUT7. All of these enzymes are members of the DNA polymerase  $\beta$ -like

superfamily (Figure 1.6). While Cid1, TUT4, and TUT7 have been shown to act as uridylyltransferases *in vivo*<sup>46,50,93–95,112,123,136,139</sup>, Gld2 has been shown to act as both an adenylyltransferase and uridylyltransferase<sup>31,92,104,105,109,140</sup>.

The crystal structure of U-specific Cid1 revealed the molecular basis of nucleotide discrimination: a single histidine residue in Tutases sterically hinders ATP from entering the active site (Figure 1.6)<sup>141</sup>. Some uridylyltransferases such as *Naegleria gruberi* Cid1 encode a phenylalanine instead of a histidine at this site, however both aromatic amino acids may serve to block ATP from the active site through steric hindrance<sup>109</sup>. Mutation of this histidine to an asparagine in Cid1 and to a leucine in *Xenopus laevis* TUT7 yielded an enzyme that lost its specificity for UTP, indicating that the histidine is crucial for UTP selectivity<sup>141,142</sup>.

It is important to note that adenylation and uridylation can have opposing effects on mRNA and miRNA stability (Figures 1.5 and 1.7). Generally, the addition of an adenine residue to mature miRNAs leads to stability<sup>104</sup> while an uridine residue leads to silencing<sup>93</sup>. Group II pre-miRNAs, however, require the addition of an uridine residue<sup>90</sup> on the 3'-end for proper maturation, while an oligo(U) tail, with the help of Lin28A, will lead to degradation by Dis3L2<sup>91,112</sup>. The addition of an oligo(U) tail to a shortened poly(A) tail (less than 25 nts) of mRNAs can lead to degradation by the 3'-5' exonuclease Dis3L2<sup>50</sup> and decapping of the 5'-end<sup>45</sup>. On the other hand, an oligo(A) tail will lengthen the poly(A) tail and extend the half-life of the mRNA<sup>31,96,143–145</sup>. In addition, histone mRNAs are subjected to 5'-3' and 3'-5' degradation when polyuridylated<sup>49</sup>. These opposing biological roles by very similar enzymes emphasize the need to define the biochemical activity of every TENT in order to correctly assign their biological function.



**Figure 1.7: 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.

## 1.4.2 Human uridylyltransferases

Like the yeast homolog Cid1, human TUT4 and TUT7 both encode a nucleotidyltransferase (NTR) domain and an adjacent PAP-associated domain<sup>141,146–150</sup> (Figure 1.6). While Cid1 lacks obvious RNA binding motifs, TUT4 and TUT7 encode several zinc finger domains thought to be responsible for substrate binding and substrate specificity. Both enzymes share a similar domain structure with three CCHC zinc fingers, an N-terminal C<sub>2</sub>H<sub>2</sub> zinc finger, and a C-terminal catalytic core consisting of the NTR domain and the PAP-associated domain (Figure 1.6). An additional copy of the catalytic core exists in the N-terminal region and is believed to be inactive as it lacks the aspartate triad in the NTR domain required for catalytic activity<sup>126</sup>. Due to the similarity in domain structure between TUT4 and TUT7, studies have shown that they may play redundant cellular roles<sup>47,90,112</sup>. TUT4 and TUT7 localize to the cytoplasm and can add single nucleotides to miRNAs (Figure 1.5A, B) or many uridine residues to histone mRNAs, miRNAs, and mRNAs (Figures 1.5C, D and 1.7C, D)<sup>50,93,94,112,123</sup>. The three CCHC-zinc finger domains flanking the active site are thought to assist in RNA binding<sup>112</sup> and the C<sub>2</sub>H<sub>2</sub> domain plays a specific role in the miRNA decay pathway, mediating the association of TUT4 to the RNA binding protein Lin28A<sup>117</sup> (Figure 1.6). Lin28A recruits

Tutases to certain substrate RNAs and is most likely required for processive uridylation<sup>112</sup>. Mutational analysis of the conserved cysteine residues in the C<sub>2</sub>H<sub>2</sub> zinc finger domain showed a decrease in Lin28A-enhanced uridylation activity<sup>112</sup>. Additional domains unique in TUT4 (Atrophin-like and Pneumo-G like domain, Figure 1.6) are not required for Tutase activity, but may be involved in substrate specificity and regulation of TUT4<sup>95</sup>. Taken together, these studies suggest polyuridylation in humans is coordinated by multiple Tutases that control mRNA degradation pathways.

Terminal uridylyltransferase 1 (TUT1, Star-PAP) is localized to the nuclear speckle and plays a role in cellular stress response, global regulation of miRNA abundance, and intron splicing<sup>88,118,124,125,151–153</sup>. TUT1 encodes an NTR domain and a PAP-associated domain but differs from the other TENTs with a proline-rich region (PRR) that is inserted in the NTR domain. TUT1 also encodes a N-terminal C<sub>2</sub>H<sub>2</sub> zinc finger and RNA recognition motif (RRM)<sup>154</sup> (Figure 1.6). TUT1 activity is associated with both adenylation-dependent pre-mRNA stabilization (Figure 1.7B) and uridylation-dependent RNA turnover. TUT1 restores the 3'-end uridyl residues of the U6 small nuclear RNA (snRNA) in the nucleus<sup>124,125,152</sup>.

TUT1 utilizes an unusual mechanism of 3'-cleavage and polyadenylation when modifying pre-mRNAs<sup>155</sup>. TUT1 is activated by phosphorylation which directs substrate specificity and enhances activity. The activated enzyme binds to a GC-rich region in the 3'-UTR of its pre-mRNA substrates via its RRM and C<sub>2</sub>H<sub>2</sub> zinc finger<sup>122,151,156</sup>. Upon binding, TUT1 directly interacts with the cleavage and polyadenylation stimulating factor (Cpsf) 160 and Cpsf 73 which subsequently recruit Cpsf 100 and Cpsf 30<sup>122</sup>. Cpsf 73 acts as an endonuclease and cleaves the pre-mRNA strand, followed by TUT1-catalyzed polyadenylation and subsequent stabilization of the mRNA<sup>122</sup>. Among the best-characterized pre-mRNA targets of TUT1 are coding sequences for heme oxygenase-1 (HO-1) and Bcl2-interacting killer (BIK). HO-1 and BIK are involved in oxidative stress and apoptotic responses, respectively<sup>122,151,155,157</sup>. Under oxidative stress, TUT1 is phosphorylated at serine and threonine residues in the PRR and targets HO-1<sup>120,122,155,158–161</sup>. Cellular stress such as viral infections and DNA damage will activate a different signaling pathway that results in phosphorylation of the PRR in TUT1, resulting in TUT1

targeting BIK<sup>151,162–164</sup>. TUT1 can also be phosphorylated at serine 6 in the N-terminal zinc finger domain by casein kinase 1 isoform  $\alpha$  (CK1 $\alpha$ ) independently of oxidative stress<sup>162</sup>.

### 1.4.3 Gld2, the minimal human nucleotidyltransferase

While single nucleotide additions play a crucial role in miRNA maturation, recent reports suggest that monoadenylation of miRNAs leads to increased miRNA stability<sup>92,103</sup>. The TENT homolog Gld2 was first described as a non-canonical PAP involved in *C. elegans* germline development, stabilizing mRNAs through polyadenylation<sup>36</sup> (Figure 1.7A). Gld2 may associate with RNA-binding proteins to target specific RNAs as the enzyme lacks identifiable RNA binding domains<sup>165</sup>. In *C. elegans*, efficient and processive mRNA polyadenylation requires the association of Gld2 with Gld3, a developmental regulator<sup>36</sup>. In humans, Gld2 has been shown to associate with the RNA binding proteins QKI-7 to polyadenylate mRNA targets<sup>166</sup>. Additionally, Gld2 promotes miRNA stabilization. Gld2-mediated monoadenylation stabilizes miRNA miR-122 transcripts in the liver<sup>104</sup> and human fibroblasts<sup>92</sup>, while it also plays a role in the translational regulation of p53<sup>105,106</sup> (Figure 1.5B). Gld2 knockout mice show decreased miRNA monoadenylation, suggesting a primary role of Gld2 in miRNA monoadenylation<sup>107</sup>. In liver cells, miR-122 abundance is likely controlled as a balance between the stabilizing effect of Gld2-catalyzed adenylation and the antagonistic, destabilizing effect of 3'-deadenylation by poly(A) specific ribonuclease (Parn)<sup>127</sup>.

Biochemical assays with recombinant human Gld2 discovered a previously unknown Gld2-mediated uridylation activity. Different reports documented Gld2 catalyzed adenylation and uridylation of miR-122<sup>92</sup> and monouridylation of pre-let-7<sup>90</sup> *in vitro*. As such, the biological role of Gld2 was unclear, as both uridylation and adenylation activity were shown, which have opposing effects on RNA stability.

#### 1.4.4 The founder uridylyltransferase Cid1

*S. pombe* Cid1 was initially identified as a protein required for S-M phase cell cycle checkpoint control<sup>135</sup> and was thought to exhibit PAP activity promoting mRNA stability<sup>165</sup>. Deletion of Cid1 prevents yeast cells from growing when exposed to the cell cycle check point inhibitors caffeine and hydroxyurea<sup>135</sup>. Overexpression of Cid1, however, allowed the cells to suppress the effects of hydroxyurea<sup>135</sup>.

When a deadenylation-independent pathway initiated by the addition of uridine residues to polyadenylated mRNAs was identified in *S. pombe*, it raised the possibility of Cid1 acting as a Tutase<sup>46</sup>. From a detailed biochemical and structural characterization<sup>121,141,146,147,165</sup>, it has now become clear that Cid1 is in fact a Tutase, catalyzing the polyuridylation of mRNAs *in vitro*, with predominantly monouridylated RNAs produced by Cid1 *in vivo*<sup>46,51</sup>. Upon uridylation, the Lsm1-7 decapping complex is recruited for mRNA 5'-3' degradation<sup>46</sup> and 3'-5' degradation is carried out by the exonuclease Dis3L2<sup>51,168–172</sup>. It is unknown how (U)-tail length is controlled in the cell, and whether these RNA species are exclusively uridylated by Cid1, as several additional TENT homologs are present in *S. pombe*<sup>109</sup>.

Cid1 is known as the minimal TENT, containing only the two domains necessary for catalytic activity, the NTR domain and PAP-associated domain. These are also the only two domains present in Cid1's adenylyltransferase counterpart Gld2. As both enzymes lack defined RNA binding domains, both enzymes display substrate promiscuity *in vitro*<sup>109,173</sup>. Substrate selectivity *in vivo* is thought to be achieved by interactions with RNA binding interacting proteins<sup>31,36,145,166,174</sup>.

The major difference between Cid1 and Gld2 lies in their nucleotide preference, Cid1 prefers UTP while Gld2 prefers ATP<sup>109,136</sup>. This is due to the presence of a critical histidine residue in the PAP-associated domain of Cid1 that sterically hinders ATP from entering the catalytic site<sup>141</sup>. Gld2 lacks this histidine residue and is thus able to accept ATP<sup>109</sup>. Interestingly, Cid1 is capable of adding a biotinylated ATP derivative, in the absence of UTP, to *in vitro* RNA transcripts under optimized conditions and is currently used for 3'-biotinylation of RNAs<sup>175</sup>.



## 1.5 Scope of Thesis

The regulation of RNAs through the addition of untemplated nucleotides to their 3'-ends has become an increasingly recognized regulatory modification. Adenylation and uridylation are associated with opposite biochemical outcomes with regard to RNA fate. Generally, adenylation stabilizes RNAs while uridylation silences miRNAs or marks RNAs for degradation. These additions play important roles in the regulation of mRNAs, either directly through nucleotide additions to the 3'-end of mRNAs or indirectly through miRNA regulation<sup>19,20,101,138,175</sup>. The enzymes that perform these tasks, TENTs, are thus of great interest. Therefore, I hypothesize that TENTs are required to maintain RNA homeostasis. In this thesis, the abilities of the TENTs Gld2 and Cid1 to catalyze nucleotide addition and regulate RNA stability were examined using precise biochemical assays.

Gld2 is a member of the non-canonical poly(A) polymerases, which include enzymes with varying nucleotide specificity, ranging from strictly ATP-adding to ambiguous to exclusively UTP-adding enzymes. Human Gld2 has been associated with transcript stabilizing miRNA monoadenylation and cytoplasmic mRNA polyadenylation<sup>92,104,166</sup>. More recent data revealed an unexpected miRNA uridylation activity, which promotes miRNA maturation<sup>90</sup>. These conflicting data raise the question of Gld2 nucleotide specificity. Chapter 2 biochemically characterizes human Gld2 and demonstrates that it is a *bona fide* adenylyltransferase *in vitro* with only weak activity toward other nucleotides. Despite its sequence similarity with uridylyltransferases (TUT4, TUT7), I show that Gld2 displays an 83-fold preference for ATP over UTP. I further show that Gld2 is promiscuous for its RNA substrate, with activity toward miRNA, pre-miRNA, and polyadenylated RNA substrates. *In vitro* Gld2 activity is restricted to adding single nucleotides<sup>109</sup> while processivity observed *in vivo* likely relies on additional RNA-binding proteins<sup>166</sup>. In a phylogenetic analysis of the PAP/Tutase superfamily, I further show that uridylyltransferase activity, which is derived from distinct adenylyltransferase ancestors, arose multiple times during evolution via insertion of an active site histidine residue. A corresponding histidine insertion into the Gld2 active site switches substrate specificity from ATP to UTP<sup>109</sup>.

Chapter 3 elucidates the first link between oncogenic kinase activity and the regulation of miRNA stability. The de-regulation of miRNAs is associated with multiple human diseases, yet cellular mechanisms governing miRNA abundance remain largely unknown<sup>58</sup>. While Gld2 activity was shown to stabilize miRNAs<sup>92,104</sup>, the regulation of Gld2 itself remained unclear. Human miR-122 is required for Hepatitis C proliferation and low miR-122 abundance is associated with hepatic cancer<sup>176</sup>. Gld2 catalyzes the post-transcriptional addition of a single adenine residue (A+1) to the 3'-end of miR-122, enhancing its stability<sup>92</sup>. I found that Gld2 activity is regulated by site-specific phosphorylation in its predicted disordered N-terminal domain. I identified two phosphorylation sites (S62, S110) where phosphomimetic substitutions increased Gld2 activity and one site (S116) that markedly reduced activity. Using mass spectrometry, I confirmed that HEK 293 cell extracts readily phosphorylate the N-terminus of Gld2 at S62. I also identified protein kinase A (PKA) and protein kinase B (Akt1) as kinases that site-specifically phosphorylate Gld2 at S116, abolishing Gld2-mediated nucleotide addition. The data demonstrate a novel phosphorylation-dependent mechanism to regulate Gld2 activity, revealing tumor suppressor miRNAs as a previously unknown target of Akt1-dependent signaling.

Uridylation-dependent RNA decay is a widespread eukaryotic pathway modulating RNA homeostasis<sup>99,168</sup>. Tutases add untemplated uridine residues to RNA 3'-ends, marking them for degradation by the U-specific exonuclease Dis3L2<sup>168</sup>. In *S. pombe*, Cid1 uridylates a variety of RNAs<sup>121,136</sup>. Chapter 4 investigates the prevalence and impact of uridylation-dependent RNA decay in *S. pombe* by transcriptionally profiling *cid1* and *dis3L2* deletion strains. I found that the exonuclease Dis3L2 represents a bottleneck in uridylation-dependent mRNA decay, whereas Cid1 plays a redundant role that can be complemented by other Tutases. Deletion of *dis3L2* elicits a cellular stress response, up-regulating transcription of genes involved in protein folding and protein degradation. Misfolded proteins accumulate in both deletion strains, yet only trigger a strong stress response in *dis3L2* deficient cells. While deletion of *cid1* increases sensitivity to protein misfolding stress, a *dis3L2* deletion showed no increased sensitivity or was even protective. Furthermore, uridylyl- and adenylyltransferases were shown to cooperate to generate a novel 5'-NxAUUAAAA-3' RNA motif on *dak2* mRNA. These studies

elucidate the role of uridylation-dependent RNA decay as part of a global mRNA surveillance, and perturbation of this pathway leads to the accumulation of misfolded proteins and elicits cellular stress responses.

This thesis presents work characterizing the adenylyltransferase Gld2 and identifies site-specific serine phosphorylation as a means of TENT regulation. In addition, the cellular impact of the founder uridylyltransferase Cid1 on global RNA uridylation was explored. Altogether, the work takes a closer look at the regulation and impact of the minimal TENTs Gld2 and Cid1 that can be used as a starting point in the study of other TENTs.

## 1.6 References

1. Crick, F. H. On protein synthesis. *Symp. Soc. Exp. Biol.* **12**, 138–163 (1958).
2. Travers, A. & Muskhelishvili, G. DNA structure and function. *FEBS J.* **282**, 2279–2295 (2015).
3. Hahn, S. Structure and mechanism of the RNA Polymerase II transcription machinery. *Nat. Struct. Mol. Biol.* **11**, 394–403 (2004).
4. Sainsbury, S., Bernecky, C. & Cramer, P. Structural basis of transcription initiation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.* **16**, 129–143 (2015).
5. Sonenberg, N. & Hinnebusch, A. G. Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. *Cell* **136**, 731–745 (2009).
6. Lant, J. T., Berg, M. D., Heinemann, I. U., Brandl, C. J. & O'Donoghue, P. Pathways to disease from natural variations in human cytoplasmic tRNAs. *J. Biol. Chem.* **294**, 5294–5308 (2019).
7. Rother, M. & Krzycki, J. A. Selenocysteine, pyrrolysine, and the unique energy metabolism of methanogenic archaea. *Archaea Vanc. BC* **2010**, (2010).
8. Turanov, A. A. *et al.* Biosynthesis of selenocysteine, the 21st amino acid in the genetic code, and a novel pathway for cysteine biosynthesis. *Adv. Nutr. Bethesda Md* **2**, 122–128 (2011).
9. Schmidt, R. L. & Simonović, M. Synthesis and decoding of selenocysteine and human health. *Croat. Med. J.* **53**, 535–550 (2012).
10. Gaston, M. A., Jiang, R. & Krzycki, J. A. Functional context, biosynthesis, and genetic encoding of pyrrolysine. *Curr. Opin. Microbiol.* **14**, 342–349 (2011).
11. Pang, Y. L. J., Poruri, K. & Martinis, S. A. tRNA synthetase: tRNA Aminoacylation and beyond. *Wiley Interdiscip. Rev. RNA* **5**, 461–480 (2014).
12. Dever, T. E. & Green, R. The Elongation, Termination, and Recycling Phases of Translation in Eukaryotes. *Cold Spring Harb. Perspect. Biol.* **4**, (2012).
13. Rehman, I. & Botelho, S. Biochemistry, Secondary Protein Structure. in *StatPearls* (StatPearls Publishing, 2019).

14. Sims, R. J., Belotserkovskaya, R. & Reinberg, D. Elongation by RNA polymerase II: the short and long of it. *Genes Dev.* **18**, 2437–2468 (2004).
15. Bentley, D. L. Coupling mRNA processing with transcription in time and space. *Nat. Rev. Genet.* **15**, 163–175 (2014).
16. Cowling, V. H. Regulation of mRNA cap methylation. *Biochem. J.* **425**, 295–302 (2010).
17. Pandya-Jones, A. Pre-mRNA splicing during transcription in the mammalian system. *Wiley Interdiscip. Rev. RNA* **2**, 700–717 (2011).
18. Wahl, M. C., Will, C. L. & Lührmann, R. The spliceosome: design principles of a dynamic RNP machine. *Cell* **136**, 701–718 (2009).
19. Proudfoot, N. J. Ending the message: poly(A) signals then and now. *Genes Dev.* **25**, 1770–1782 (2011).
20. Colgan, D. F. & Manley, J. L. Mechanism and regulation of mRNA polyadenylation. *Genes Dev.* **11**, 2755–2766 (1997).
21. Mandel, C. R., Bai, Y. & Tong, L. Protein factors in pre-mRNA 3'-end processing. *Cell. Mol. Life Sci. CMLS* **65**, 1099–1122 (2008).
22. Katahira, J. Nuclear Export of Messenger RNA. *Genes* **6**, 163–184 (2015).
23. Huang, Y. & Carmichael, G. C. Role of polyadenylation in nucleocytoplasmic transport of mRNA. *Mol. Cell. Biol.* **16**, 1534–1542 (1996).
24. Fuke, H. & Ohno, M. Role of poly (A) tail as an identity element for mRNA nuclear export. *Nucleic Acids Res.* **36**, 1037–1049 (2008).
25. Collier, J. M., Gray, N. K. & Wickens, M. P. mRNA stabilization by poly(A) binding protein is independent of poly(A) and requires translation. *Genes Dev.* **12**, 3226–3235 (1998).
26. Wilusz, C. J., Wormington, M. & Peltz, S. W. The cap-to-tail guide to mRNA turnover. *Nat. Rev. Mol. Cell Biol.* **2**, 237–246 (2001).
27. Imataka, H., Gradi, A. & Sonenberg, N. A newly identified N-terminal amino acid sequence of human eIF4G binds poly(A)-binding protein and functions in poly(A)-dependent translation. *EMBO J.* **17**, 7480–7489 (1998).
28. Kahvejian, A., Svitkin, Y. V., Sukarieh, R., M'Boutchou, M.-N. & Sonenberg, N. Mammalian poly(A)-binding protein is a eukaryotic translation initiation factor, which acts via multiple mechanisms. *Genes Dev.* **19**, 104–113 (2005).
29. Wells, S. E., Hillner, P. E., Vale, R. D. & Sachs, A. B. Circularization of mRNA by eukaryotic translation initiation factors. *Mol. Cell* **2**, 135–140 (1998).
30. Rhoads, R. E. eIF4E: New Family Members, New Binding Partners, New Roles. *J. Biol. Chem.* **284**, 16711–16715 (2009).
31. Kwak, J. E., Wang, L., Ballantyne, S., Kimble, J. & Wickens, M. Mammalian GLD-2 homologs are poly(A) polymerases. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 4407–4412 (2004).
32. Mendez, R. & Richter, J. D. Translational control by CPEB: a means to the end. *Nat. Rev. Mol. Cell Biol.* **2**, 521–529 (2001).
33. Tadros, W. & Lipshitz, H. D. Setting the stage for development: mRNA translation and stability during oocyte maturation and egg activation in *Drosophila*. *Dev. Dyn. Off. Publ. Am. Assoc. Anat.* **232**, 593–608 (2005).
34. Charlesworth, A., Meijer, H. A. & de Moor, C. H. Specificity factors in cytoplasmic polyadenylation. *Wiley Interdiscip. Rev. RNA* **4**, 437–461 (2013).

35. Kadyk, L. C. & Kimble, J. Genetic regulation of entry into meiosis in *Caenorhabditis elegans*. *Dev. Camb. Engl.* **125**, 1803–1813 (1998).
36. Wang, L., Eckmann, C. R., Kadyk, L. C., Wickens, M. & Kimble, J. A regulatory cytoplasmic poly(A) polymerase in *Caenorhabditis elegans*. *Nature* **419**, 312–316 (2002).
37. Cui, J., Sartain, C. V., Pleiss, J. A. & Wolfner, M. F. Cytoplasmic polyadenylation is a major mRNA regulator during oogenesis and egg activation in *Drosophila*. *Dev. Biol.* **383**, 121–131 (2013).
38. Yamashita, A. *et al.* Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. *Nat. Struct. Mol. Biol.* **12**, 1054–1063 (2005).
39. Brown, C. E. & Sachs, A. B. Poly(A) tail length control in *Saccharomyces cerevisiae* occurs by message-specific deadenylation. *Mol. Cell. Biol.* **18**, 6548–6559 (1998).
40. Brown, C. E., Tarun, S. Z., Boeck, R. & Sachs, A. B. PAN3 encodes a subunit of the Pab1p-dependent poly(A) nuclease in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **16**, 5744–5753 (1996).
41. Chen, C.-Y. A. & Shyu, A.-B. Mechanisms of deadenylation-dependent decay. *Wiley Interdiscip. Rev. RNA* **2**, 167–183 (2011).
42. Garneau, N. L., Wilusz, J. & Wilusz, C. J. The highways and byways of mRNA decay. *Nat. Rev. Mol. Cell Biol.* **8**, 113–126 (2007).
43. Tucker, M. *et al.* The transcription factor associated Ccr4 and Caf1 proteins are components of the major cytoplasmic mRNA deadenylase in *Saccharomyces cerevisiae*. *Cell* **104**, 377–386 (2001).
44. Maryati, M., Airhihen, B. & Winkler, G. S. The enzyme activities of Caf1 and Ccr4 are both required for deadenylation by the human Ccr4-Not nuclease module. *Biochem. J.* **469**, 169–176 (2015).
45. Collier, J. & Parker, R. Eukaryotic mRNA decapping. *Annu. Rev. Biochem.* **73**, 861–890 (2004).
46. 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).
47. Lim, J. *et al.* Uridylation by TUT4 and TUT7 Marks mRNA for Degradation. *Cell* **159**, 1365–1376 (2014).
48. Sement, F. M. *et al.* Uridylation prevents 3' trimming of oligoadenylated mRNAs. *Nucleic Acids Res.* **41**, 7115–7127 (2013).
49. 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).
50. Schmidt, M.-J., West, S. & Norbury, C. J. The human cytoplasmic RNA terminal U-transferase ZCCHC11 targets histone mRNAs for degradation. *RNA N. Y. N* **17**, 39–44 (2011).
51. Malecki, M. *et al.* The exoribonuclease Dis3L2 defines a novel eukaryotic RNA degradation pathway. *EMBO J.* **32**, 1842–1854 (2013).
52. Faehnle, C. R., Walleshauser, J. & Joshua-Tor, L. Mechanism of Dis3L2 substrate recognition in the Lin28-let-7 pathway. *Nature* **514**, 252–256 (2014).
53. Luan, S., Luo, J., Liu, H. & Li, Z. Regulation of RNA decay and cellular function by 3'-5' exoribonuclease DIS3L2. *RNA Biol.* **16**, 160–165 (2019).

54. Chowdhury, A., Mukhopadhyay, J. & Tharun, S. The decapping activator Lsm1p-7p-Pat1p complex has the intrinsic ability to distinguish between oligoadenylated and polyadenylated RNAs. *RNA* **13**, 998–1016 (2007).
55. Lackey, P. E., Welch, J. D. & Marzluff, W. F. TUT7 catalyzes the uridylation of the 3' end for rapid degradation of histone mRNA. *RNA N. Y. N* **22**, 1673–1688 (2016).
56. Su, W. *et al.* mRNAs containing the histone 3' stem-loop are degraded primarily by decapping mediated by oligouridylation of the 3' end. *RNA N. Y. N* **19**, 1–16 (2013).
57. Lee, R. C., Feinbaum, R. L. & Ambros, V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* **75**, 843–854 (1993).
58. Bhaskaran, M. & Mohan, M. MicroRNAs: history, biogenesis, and their evolving role in animal development and disease. *Vet. Pathol.* **51**, 759–774 (2014).
59. Ameres, S. L. & Zamore, P. D. Diversifying microRNA sequence and function. *Nat. Rev. Mol. Cell Biol.* **14**, 475–488 (2013).
60. Iwakawa, H.-O. & Tomari, Y. The Functions of MicroRNAs: mRNA Decay and Translational Repression. *Trends Cell Biol.* **25**, 651–665 (2015).
61. Kloosterman, W. P. & Plasterk, R. H. A. The Diverse Functions of MicroRNAs in Animal Development and Disease. *Dev. Cell* **11**, 441–450 (2006).
62. Small, E. M. & Olson, E. N. Pervasive roles of microRNAs in cardiovascular biology. *Nature* **469**, 336–342 (2011).
63. Zhou, S.-S. *et al.* miRNAs in cardiovascular diseases: potential biomarkers, therapeutic targets and challenges. *Acta Pharmacol. Sin.* **39**, 1073–1084 (2018).
64. Zhu, X.-M., Wu, L.-J., Xu, J., Yang, R. & Wu, F.-S. Let-7c microRNA expression and clinical significance in hepatocellular carcinoma. *J. Int. Med. Res.* **39**, 2323–2329 (2011).
65. Cammarata, G. *et al.* Differential expression of specific microRNA and their targets in acute myeloid leukemia. *Am. J. Hematol.* **85**, 331–339 (2010).
66. Feng, J., Xing, W. & Xie, L. Regulatory Roles of MicroRNAs in Diabetes. *Int. J. Mol. Sci.* **17**, (2016).
67. Wahid, F., Shehzad, A., Khan, T. & Kim, Y. Y. MicroRNAs: synthesis, mechanism, function, and recent clinical trials. *Biochim. Biophys. Acta* **1803**, 1231–1243 (2010).
68. Valinezhad Orang, A., Safaralizadeh, R. & Kazemzadeh-Bavili, M. Mechanisms of miRNA-Mediated Gene Regulation from Common Downregulation to mRNA-Specific Upregulation. *Int. J. Genomics* **2014**, 970607 (2014).
69. Wang, B., Yanez, A. & Novina, C. D. MicroRNA-repressed mRNAs contain 40S but not 60S components. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 5343–5348 (2008).
70. Fukaya, T., Iwakawa, H.-O. & Tomari, Y. MicroRNAs block assembly of eIF4F translation initiation complex in *Drosophila*. *Mol. Cell* **56**, 67–78 (2014).
71. Chen, S. & Gao, G. MicroRNAs recruit eIF4E2 to repress translation of target mRNAs. *Protein Cell* **8**, 750–761 (2017).
72. Petersen, C. P., Bordeleau, M.-E., Pelletier, J. & Sharp, P. A. Short RNAs repress translation after initiation in mammalian cells. *Mol. Cell* **21**, 533–542 (2006).
73. Parker, R. & Sheth, U. P bodies and the control of mRNA translation and degradation. *Mol. Cell* **25**, 635–646 (2007).
74. Niinuma, S., Fukaya, T. & Tomari, Y. CCR4 and CAF1 deadenylases have an intrinsic activity to remove the post-poly(A) sequence. *RNA N. Y. N* **22**, 1550–1559 (2016).

75. Ha, M. & Kim, V. N. Regulation of microRNA biogenesis. *Nat. Rev. Mol. Cell Biol.* **15**, 509–524 (2014).
76. Lee, Y. *et al.* MicroRNA genes are transcribed by RNA polymerase II. *EMBO J.* **23**, 4051–4060 (2004).
77. Borchert, G. M., Lanier, W. & Davidson, B. L. RNA polymerase III transcribes human microRNAs. *Nat. Struct. Mol. Biol.* **13**, 1097–1101 (2006).
78. Canella, D., Praz, V., Reina, J. H., Cousin, P. & Hernandez, N. Defining the RNA polymerase III transcriptome: Genome-wide localization of the RNA polymerase III transcription machinery in human cells. *Genome Res.* **20**, 710–721 (2010).
79. Lee, Y. *et al.* The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**, 415–419 (2003).
80. Han, J. *et al.* The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev.* **18**, 3016–3027 (2004).
81. Zeng, Y., Yi, R. & Cullen, B. R. Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha. *EMBO J.* **24**, 138–148 (2005).
82. Yi, R., Qin, Y., Macara, I. G. & Cullen, B. R. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev.* **17**, 3011–3016 (2003).
83. Bohnsack, M. T., Czaplinski, K. & Gorlich, D. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA N. Y. N* **10**, 185–191 (2004).
84. Hutvagner, G. *et al.* A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* **293**, 834–838 (2001).
85. Ketting, R. F. *et al.* Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev.* **15**, 2654–2659 (2001).
86. Meijer, H. A., Smith, E. M. & Bushell, M. Regulation of miRNA strand selection: follow the leader? *Biochem. Soc. Trans.* **42**, 1135–1140 (2014).
87. Gregory, R. I., Chendrimada, T. P., Cooch, N. & Shiekhattar, R. Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell* **123**, 631–640 (2005).
88. Burroughs, A. M. *et al.* A comprehensive survey of 3' animal miRNA modification events and a possible role for 3' adenylation in modulating miRNA targeting effectiveness. *Genome Res.* **20**, 1398–1410 (2010).
89. Wyman, S. K. *et al.* Post-transcriptional generation of miRNA variants by multiple nucleotidyl transferases contributes to miRNA transcriptome complexity. *Genome Res.* **21**, 1450–1461 (2011).
90. 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).
91. Heo, I. *et al.* TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. *Cell* **138**, 696–708 (2009).
92. 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).
93. Jones, M. R. *et al.* Zcchc11 uridylates mature miRNAs to enhance neonatal IGF-1 expression, growth, and survival. *PLoS Genet.* **8**, e1003105 (2012).
94. Jones, M. R. *et al.* Zcchc11-dependent uridylation of microRNA directs cytokine expression. *Nat. Cell Biol.* **11**, 1157–1163 (2009).

95. Thornton, J. E. *et al.* Selective microRNA uridylation by Zcchc6 (TUT7) and Zcchc11 (TUT4). *Nucleic Acids Res.* **42**, 11777–11791 (2014).
96. Kwak, J. E. & Wickens, M. A family of poly(U) polymerases. *RNA* **13**, 860–867 (2007).
97. Norbury, C. J. Cytoplasmic RNA: a case of the tail wagging the dog. *Nat. Rev. Mol. Cell Biol.* **14**, 643–653 (2013).
98. Aphasizhev, R. RNA uridylyltransferases. *Cell. Mol. Life Sci. CMLS* **62**, 2194–2203 (2005).
99. Norbury, C. J. 3' Uridylation and the regulation of RNA function in the cytoplasm. *Biochem. Soc. Trans.* **38**, 1150–1153 (2010).
100. Ren, G., Chen, X. & Yu, B. Uridylation of miRNAs by HEN1 SUPPRESSOR1 in Arabidopsis. *Curr. Biol.* **22**, 695–700 (2012).
101. Scott, D. D. & Norbury, C. J. RNA decay via 3' uridylation. *Biochim. Biophys. Acta BBA - Gene Regul. Mech.* **1829**, 654–665 (2013).
102. Perumal, K. & Reddy, R. The 3' end formation in small RNAs. *Gene Expr.* **10**, 59–78 (2002).
103. Peng, F. *et al.* HBx down-regulated Gld2 plays a critical role in HBV-related dysregulation of miR-122. *PloS One* **9**, e92998 (2014).
104. 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).
105. Glahder, J. A. & Norrild, B. Involvement of hGLD-2 in cytoplasmic polyadenylation of human p53 mRNA. *APMIS* **119**, 769–775 (2011).
106. 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).
107. Mansur, F. *et al.* Gld2-catalyzed 3' monoadenylation of miRNAs in the hippocampus has no detectable effect on their stability or on animal behavior. *RNA* **22**, 1492–1499 (2016).
108. Yates, L. A., Norbury, C. J. & Gilbert, R. J. C. The long and short of microRNA. *Cell* **153**, 516–519 (2013).
109. Chung, C. Z., Jo, D. H. S. & Heinemann, I. U. Nucleotide specificity of the human terminal nucleotidyltransferase Gld2 (TUT2). *RNA N. Y. N* **22**, 1239–1249 (2016).
110. Lin, S. & Gregory, R. I. Identification of small molecule inhibitors of Zcchc11 TUTase activity. *RNA Biol.* **12**, 792–800 (2015).
111. Liu, X. *et al.* A MicroRNA Precursor Surveillance System in Quality Control of MicroRNA Synthesis. *Mol. Cell* **55**, 868–879 (2014).
112. 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 N. Y. N* **18**, 1875–1885 (2012).
113. Heo, I. *et al.* Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA. *Mol. Cell* **32**, 276–284 (2008).
114. Nam, Y., Chen, C., Gregory, R. I., Chou, J. J. & Sliz, P. Molecular basis for interaction of let-7 microRNAs with Lin28. *Cell* **147**, 1080–1091 (2011).
115. Thornton, J. E. & Gregory, R. I. How does Lin28 let-7 control development and disease? *Trends Cell Biol.* **22**, 474–482 (2012).



116. Shyh-Chang, N. & Daley, G. Q. Lin28: primal regulator of growth and metabolism in stem cells. *Cell Stem Cell* **12**, 395–406 (2013).
117. Kim, B. *et al.* TUT7 controls the fate of precursor microRNAs by using three different uridylation mechanisms. *EMBO J.* **34**, 1801–1815 (2015).
118. 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).
119. Guhaniyogi, J. & Brewer, G. Regulation of mRNA stability in mammalian cells. *Gene* **265**, 11–23 (2001).
120. 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).
121. Rissland, O. S. & Norbury, C. J. The Cid1 poly(U) polymerase. *Biochim. Biophys. Acta* **1779**, 286–294 (2008).
122. 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).
123. 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).
124. Trippe, R., Sandrock, B. & Benecke, B. J. A highly specific terminal uridylyl transferase modifies the 3'-end of U6 small nuclear RNA. *Nucleic Acids Res.* **26**, 3119–3126 (1998).
125. Trippe, R., Richly, H. & Benecke, B.-J. Biochemical characterization of a U6 small nuclear RNA-specific terminal uridylyltransferase. *Eur. J. Biochem.* **270**, 971–980 (2003).
126. Minoda, Y. *et al.* A novel Zinc finger protein, ZCCHC11, interacts with TIFA and modulates TLR signaling. *Biochem. Biophys. Res. Commun.* **344**, 1023–1030 (2006).
127. Katoh, T., Hojo, H. & Suzuki, T. Destabilization of microRNAs in human cells by 3' deadenylation mediated by PARN and CUGBP1. *Nucleic Acids Res.* **43**, 7521–7534 (2015).
128. Nagaike, T., Suzuki, T., Katoh, T. & Ueda, T. Human mitochondrial mRNAs are stabilized with polyadenylation regulated by mitochondria-specific poly(A) polymerase and polynucleotide phosphorylase. *J. Biol. Chem.* **280**, 19721–19727 (2005).
129. Bai, Y., Srivastava, S. K., Chang, J. H., Manley, J. L. & Tong, L. Crystal structure of human PAPD1, a noncanonical poly(A) polymerase. *Mol. Cell* **41**, 311–320 (2011).
130. 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. U. S. A.* **99**, 12079–12084 (2002).
131. Win, T. Z., Stevenson, A. L. & Wang, S.-W. Fission yeast Cid12 has dual functions in chromosome segregation and checkpoint control. *Mol. Cell. Biol.* **26**, 4435–4447 (2006).
132. Saitoh, S. *et al.* Cid13 is a cytoplasmic poly(A) polymerase that regulates ribonucleotide reductase mRNA. *Cell* **109**, 563–573 (2002).
133. Keller, C., Woolcock, K., Hess, D. & Bühler, M. Proteomic and functional analysis of the noncanonical poly(A) polymerase Cid14. *RNA N. Y. N* **16**, 1124–1129 (2010).

134. Win, T. Z. *et al.* Requirement of Fission Yeast Cid14 in Polyadenylation of rRNAs. *Mol. Cell. Biol.* **26**, 1710–1721 (2006).
135. Wang, S.-W., Toda, T., MacCallum, R., Harris, A. L. & Norbury, C. Cid1, a Fission Yeast Protein Required for S-M Checkpoint Control when DNA Polymerase  $\delta$  or  $\epsilon$  Is Inactivated. *Mol. Cell. Biol.* **20**, 3234 (2000).
136. Rissland, O. S., Mikulasova, A. & Norbury, C. J. Efficient RNA polyuridylation by noncanonical poly(A) polymerases. *Mol. Cell. Biol.* **27**, 3612–3624 (2007).
137. Pisacane, P. & Halic, M. Tailing and degradation of Argonaute-bound small RNAs protect the genome from uncontrolled RNAi. *Nat. Commun.* **8**, 15332 (2017).
138. 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* **1861**, 2971–2979 (2017).
139. Blahna, M. T., Jones, M. R., Quinton, L. J., Matsuura, K. Y. & Mizgerd, J. P. Terminal uridylyltransferase enzyme Zcchc11 promotes cell proliferation independent of its uridylyltransferase activity. *J. Biol. Chem.* **286**, 42381–42389 (2011).
140. Kwak, J. E. *et al.* GLD2 poly(A) polymerase is required for long-term memory. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 14644–14649 (2008).
141. 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).
142. Lapointe, C. P. & Wickens, M. The nucleic acid-binding domain and translational repression activity of a *Xenopus* terminal uridylyl transferase. *J. Biol. Chem.* **288**, 20723–20733 (2013).
143. 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).
144. Rouhana, L. *et al.* Vertebrate GLD2 poly(A) polymerases in the germline and the brain. *RNA N. Y. N* **11**, 1117–1130 (2005).
145. Kim, J. H. & Richter, J. D. Opposing polymerase-deadenylase activities regulate cytoplasmic polyadenylation. *Mol. Cell* **24**, 173–183 (2006).
146. Munoz-Tello, P., Gabus, C. & Thore, S. Functional implications from the Cid1 poly(U) polymerase crystal structure. *Struct. Lond. Engl.* **1993** **20**, 977–986 (2012).
147. Munoz-Tello, P., Gabus, C. & Thore, S. A critical switch in the enzymatic properties of the Cid1 protein deciphered from its product-bound crystal structure. *Nucleic Acids Res.* **42**, 3372–3380 (2014).
148. Deng, J., Ernst, N. L., Turley, S., Stuart, K. D. & Hol, W. G. Structural basis for UTP specificity of RNA editing TUTases from *Trypanosoma brucei*. *EMBO J.* **24**, 4007–4017 (2005).
149. Stagno, J., Aphasizheva, I., Bruystens, J., Luecke, H. & Aphasizhev, R. Structure of the Mitochondrial Editosome-like Complex Associated TUTase 1 Reveals Divergent Mechanisms of UTP Selection and Domain Organization. *J. Mol. Biol.* **399**, 464–475 (2010).
150. 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).
151. Li, W. *et al.* Star-PAP control of BIK expression and apoptosis is regulated by nuclear PIPKI $\alpha$  and PKC $\delta$  signaling. *Mol. Cell* **45**, 25–37 (2012).

152. Trippe, R. *et al.* Identification, cloning, and functional analysis of the human U6 snRNA-specific terminal uridylyl transferase. *RNA N. Y. N* **12**, 1494–1504 (2006).
153. Kuhn, A. N. & Brow, D. A. Suppressors of a cold-sensitive mutation in yeast U4 RNA define five domains in the splicing factor Prp8 that influence spliceosome activation. *Genetics* **155**, 1667–1682 (2000).
154. Martin, G. & Keller, W. RNA-specific ribonucleotidyl transferases. *RNA N. Y. N* **13**, 1834–1849 (2007).
155. Mellman, D. L. *et al.* A PtdIns4,5P2-regulated nuclear poly(A) polymerase controls expression of select mRNAs. *Nature* **451**, 1013–1017 (2008).
156. Ray, D. *et al.* A compendium of RNA-binding motifs for decoding gene regulation. *Nature* **499**, 172–177 (2013).
157. Maines, M. D. & Gibbs, P. E. M. 30 some years of heme oxygenase: from a ‘molecular wrecking ball’ to a ‘mesmerizing’ trigger of cellular events. *Biochem. Biophys. Res. Commun.* **338**, 568–577 (2005).
158. Gonzales, M. L., Mellman, D. L. & Anderson, R. A. CKI $\alpha$  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).
159. Laishram, R. S., Barlow, C. A. & Anderson, R. A. CKI isoforms  $\alpha$  and  $\epsilon$  regulate Star-PAP target messages by controlling Star-PAP poly(A) polymerase activity and phosphoinositide stimulation. *Nucleic Acids Res.* **39**, 7961–7973 (2011).
160. Okada, M., Jang, S.-W. & Ye, K. Akt phosphorylation and nuclear phosphoinositide association mediate mRNA export and cell proliferation activities by ALY. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 8649–8654 (2008).
161. Gross, S. D. & Anderson, R. A. Casein kinase I: spatial organization and positioning of a multifunctional protein kinase family. *Cell. Signal.* **10**, 699–711 (1998).
162. Mohan, N., AP, S., Francis, N., Anderson, R. & Laishram, R. S. Phosphorylation regulates the Star-PAP-PIPKI $\alpha$  interaction and directs specificity toward mRNA targets. *Nucleic Acids Res.* **43**, 7005–7020 (2015).
163. Humphries, M. J. *et al.* Suppression of Apoptosis in the Protein Kinase C $\delta$  Null Mouse in Vivo. *J. Biol. Chem.* **281**, 9728–9737 (2006).
164. Steinberg, S. F. Distinctive activation mechanisms and functions for protein kinase Cdelta. *Biochem. J.* **384**, 449–459 (2004).
165. Stevenson, A. L. & Norbury, C. J. The Cid1 family of non-canonical poly(A) polymerases. *Yeast Chichester Engl.* **23**, 991–1000 (2006).
166. Yamagishi, R., Tsusaka, T., Mitsunaga, H., Maehata, T. & Hoshino, S. The STAR protein QKI-7 recruits PAPD4 to regulate post-transcriptional polyadenylation of target mRNAs. *Nucleic Acids Res.* **44**, 2475–2490 (2016).
167. Traut, T. W. Physiological concentrations of purines and pyrimidines. *Mol. Cell. Biochem.* **140**, 1–22 (1994).
168. Ustianenko, D. *et al.* TUT-DIS3L2 is a mammalian surveillance pathway for aberrant structured non-coding RNAs. *EMBO J.* **35**, 2179–2191 (2016).
169. Pirouz, M., Du, P., Munafò, M. & Gregory, R. I. Dis3l2-Mediated Decay Is a Quality Control Pathway for Noncoding RNAs. *Cell Rep.* **16**, 1861–1873 (2016).
170. Gallouzi, I. E. & Wilusz, J. A DIStinctively novel exoribonuclease that really likes U. *EMBO J.* **32**, 1799–1801 (2013).

171. Lv, H. *et al.* Structural analysis of Dis3l2, an exosome-independent exonuclease from *Schizosaccharomyces pombe*. *Acta Crystallogr. D Biol. Crystallogr.* **71**, 1284–1294 (2015).
172. Robinson, S. R., Oliver, A. W., Chevassut, T. J. & Newbury, S. F. The 3' to 5' Exoribonuclease DIS3: From Structure and Mechanisms to Biological Functions and Role in Human Disease. *Biomolecules* **5**, 1515–1539 (2015).
173. Chung, C. Z. *et al.* RNA surveillance by uridylation dependent RNA decay in *Schizosaccharomyces pombe*. *Nucleic Acids Res.* (2019). doi:10.1093/nar/gkz043
174. Yates, L. A. *et al.* Structural plasticity of Cid1 provides a basis for its distributive RNA terminal uridylyl transferase activity. *Nucleic Acids Res.* **43**, 2968–2979 (2015).
175. Moritz, B. & Wahle, E. Simple methods for the 3' biotinylation of RNA. *RNA N. Y. N* **20**, 421–427 (2014).
176. Menezes, M. R., Balzeau, J. & Hagan, J. P. 3' RNA Uridylation in Epitranscriptomics, Gene Regulation, and Disease. *Front. Mol. Biosci.* **5**, 61 (2018).
177. Bandiera, S., Pfeffer, S., Baumert, T. F. & Zeisel, M. B. miR-122--a key factor and therapeutic target in liver disease. *J. Hepatol.* **62**, 448–457 (2015).

## Chapter 2

# 2 Nucleotide specificity of the human terminal nucleotidyltransferase Gld2 (TUT2)

## 2.1 Introduction

Since microRNA (miRNA) discovery in the early 1990s<sup>1,2</sup>, it has become evident that post-transcriptional gene regulation by miRNAs is involved in most biological processes<sup>3-7</sup>. The latest miRBase release contains 24,521 experimentally validated miRNA genes from 206 species<sup>8</sup>. Dysfunctional miRNA expression, processing, and degradation have been found in breast cancer<sup>9</sup>, acute myeloid leukemia<sup>10</sup>, ovarian cancer<sup>11</sup>, and hepatocellular carcinoma<sup>12</sup>. Deregulated miRNA processing also contributes to other major diseases such as Hepatitis C<sup>13</sup> and cardiovascular diseases<sup>14</sup>. Because miRNAs regulate genes that change cellular fate, miRNAs and proteins involved in miRNA regulation are promising next-generation cancer therapeutic targets. In addition, specific components of the RNA processing machinery are currently used as biomarkers for cancer detection<sup>15-19</sup>.

The generation of miRNAs is a multistage process and translational inhibition by miRNAs is achieved through base pairing with the 3'-UTR of the respective target mRNA, leading to mRNA decay or silencing<sup>20</sup>. Recent high-throughput sequencing studies revealed the presence of untemplated nucleotide additions to the 3' termini of nearly 40% of miRNAs<sup>21,22</sup>. Of these, depending on the miRNA species, ~50% displayed an extra adenine, 25% contained a single additional uridine, and the remaining 25% contained multiple nucleotides appended to the 3'-termini<sup>22</sup>. During miRNA maturation and degradation, untemplated uridine<sup>23,24</sup> and adenine<sup>25</sup> residues are added to pre-miRNAs and mature miRNAs. These untemplated nucleotide additions are an efficient means to control the levels of active miRNAs in the cell. The seemingly innocuous addition of a single nucleotide can initiate miRNA maturation, stabilization, or convert an active miRNA to an inactive form<sup>26</sup>. While multiple adenine residues are added to mRNA for stabilization, extending the transcript life-span<sup>27</sup>, multiple uridine residues mark both miRNA and mRNA for degradation<sup>28-30</sup>. The presence of untemplated nucleotides on a

variety of miRNAs is now well known; however, the corresponding enzymes have only recently become the focus of biochemical characterization.

Nucleotidyltransferases such as the human terminal uridylyltransferases (TUTs) TUT4 (Zcchc11, TENT3A), TUT7 (Zcchc6, TENT3B) and the minimal homolog Gld2 (TUT2, PAPD4, TENT2) have been shown to play fundamental roles in the regulation and maturation of miRNAs let-7 and mir-122. Tutase homologs are part of the nucleotidyltransferase superfamily of enzymes and were initially identified as adenylyltransferases associated with miRNA and mRNA adenylation. Gld2 was first described as a cytoplasmic non-canonical poly(A) polymerase involved in *Caenorhabditis elegans* germline development. *C. elegans* Gld2 displayed very little activity on its own however, and relies on an additional protein, Gld3, to promote adenylation<sup>31,32</sup>. In *Drosophila*, specific depletion of the Gld2 homolog WISPY connected its function with mRNA polyadenylation required for oocyte to egg activation<sup>33</sup>, as well as long term memory<sup>34</sup> (Figure 2.1A). Recent data showed that in addition to mRNA adenylation, WISPY adenylates miRNAs in S2 cells leading to a reduction of miRNA levels<sup>35</sup>. This observation that adenylation reduces miRNA levels<sup>35</sup> contradicts reports that monoadenylation stabilizes naturally unstable miRNAs<sup>25</sup> and demonstrates the complexity of untemplated nucleotide additions.

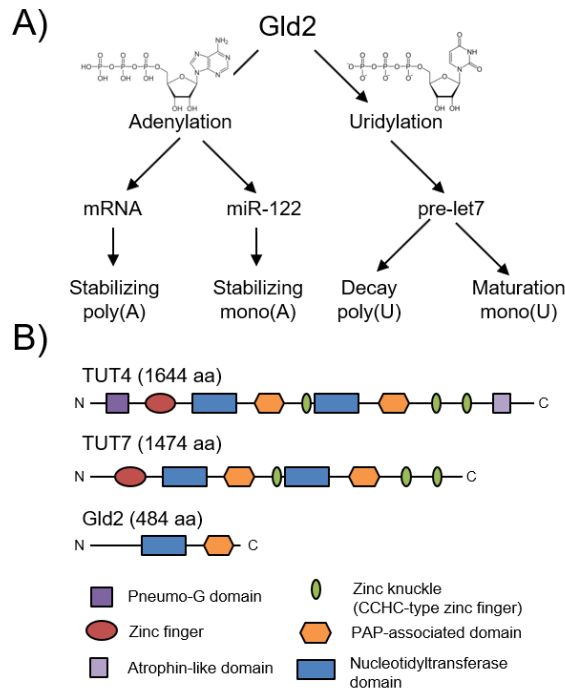
In humans, Gld2 has been equally associated with miRNA and mRNA adenylation as well as miRNA uridylation. The first reports associated human Gld2 with mRNA polyadenylation<sup>36</sup> by monitoring translation of a reporter mRNA tethered to human Gld2 and injected into *Xenopus* oocytes. *In vivo*, RNA polyadenylation requires either artificial RNA tethering<sup>36</sup> or accessory RNA binding proteins such as the cytoplasmic polyadenylation element binding protein (CPEB) in *Xenopus*<sup>37,38</sup> and Gld3 in *C. elegans*<sup>31</sup>. Further studies showed that Gld2 mediated monoadenylation stabilizes miR-122 transcripts in human fibroblasts<sup>25</sup> and plays a role in translational regulation of p53<sup>39,40</sup>. Monoadenylation is, in contrast to polyadenylation not entirely dependent on RNA binding proteins, as purified Gld2 from human cells displayed catalytic activity *in vitro*. With the discovery of the poly(U) polymerase activity of enzymes previously thought to be poly(A) polymerases, specifically the human Gld2 homologs TUT4 and

TUT7<sup>41</sup>, most recent research has uncovered a previously unknown Gld2 mediated uridylation activity. D'Ambrogio and colleagues demonstrated for the first time that human Gld2 is the enzyme responsible for monoadenylation and subsequent stabilization of miRNA-122, but they also reported a weaker uridylation activity<sup>25</sup>. Gld2 has further been shown to catalyze the monouridylation of pre-microRNA let-7a, which is crucial for its maturation<sup>24</sup>. Flag-tagged human Gld2 purified from HEK 293T cells adds a single uridine to pre-let-7a but also displayed catalytic activity adding GTP and ATP, but not CTP *in vitro*<sup>24</sup>. Interestingly, Gld2-mediated polyuridylation has been observed on pre-let-7a overhang variants<sup>42</sup> in the absence of accessory proteins. Further evidence linking Gld2 to pre-microRNA uridylation stems from knockdown assays, showing that TUT4, TUT7 and Gld2 redundantly control pre-let-7 maturation and are required for let-7 biogenesis<sup>24</sup>. Gld2 can thus function as either a Poly(A)-Polymerase (PAP) or a TUT *in vitro*.

Gld2 is composed of two major domains, a PAP associated domain and a nucleotidyltransferase (NTR) domain (Figure 2.1B). Its closest human homologs, TUT4 and TUT7 are comprised of the same domains but feature additional RNA-binding motifs, such as Zinc-finger domains. TUT4 and TUT7 have been characterized *in vivo* and *in vitro* as true uridylyltransferases and are involved in multiple processes including miRNA and mRNA uridylation. For example, uridylation of the let-7a precursor by TUT4 can drive processing by Dicer or mark the precursor miRNA for degradation, thus directly controlling let-7a levels in the cell<sup>23,26,30,43</sup>. Gld2 has been proposed to carry out a similar function during miRNA maturation<sup>24,42</sup>. While the role of TUT4 and TUT7 in these processes is becoming increasingly clear, the catalytic activity and biological role of the minimal nucleotidyltransferase Gld2 is uncertain. Evidence for both uridylation and adenylation activity of the human enzyme has been shown in *in vivo* and *in vitro* experiments, but a conclusive investigation of Gld2 nucleotide preference is lacking.

We here present a biochemical characterization of Gld2, identifying it as a *bona fide* adenylyltransferase *in vitro* with only weak activity towards UTP and GTP. Conversely, Gld2 displays a wide target RNA substrate range, adenylating multiple RNAs *in vitro*. The data suggest that Gld2 RNA substrate selectivity may require association with other

protein factors in the cell. A detailed phylogenetic analysis shows that uridylyl- and adenylyltransferases are closely related, and that uridylyltransferase activity arose independently multiple times during evolution.



**Figure 2.1: Proposed catalytic activities and domain organization of Gld2.**

**A)** Gld2 has been implicated in nucleotide addition to multiple RNA substrates in different pathways. **B)** Domain organization of Gld2 and its homologs TUT4 and TUT7. Dark purple: Pneumo-G domain; Red: C<sub>2</sub>H<sub>2</sub> Zinc finger domain; Blue: Nucleotidyltransferase domain; Orange: Poly(A) polymerase associated domain; Green: CCHC type Zinc finger domain; Light purple: Atrophin-like domain.

## 2.2 Materials and Methods

### 2.2.1 Gld2 cloning and site-directed mutagenesis

Total RNA was extracted from HeLa cells using the GeneJET RNA purification Kit (Thermo Scientific) and reverse transcription was performed with SuperScript II Reverse Transcriptase (Invitrogen) using an oligo(dT)<sub>16</sub> primer. PCR was carried out on the



cDNA with gene specific primers (Gld2EcoR1for 5'-GAATTCGATGTTCCCAAACCTC AATTTTGGG-3' and Gld2XhoIrev 5'-CTCGAGTCTTTTCAGGA-CAGCAGCTC-3'). The cDNA was digested with *EcoRI/XhoI* and ligated into pET20b. Quickchange Site directed mutagenesis (Agilent) was employed according to manufacturer's instructions to generate a Gld2-His insertion variant, using primers Tut2Hisfor 5'-GAACCTTTTGATGGAACA CATAATACAGCCAGAGCAGTGC-3' and Tut2Hisrev 5'-GCACTGCTCTGGCTGTA TTATGTGTTCCATCAAAAGGTTC-3'. The construct and mutation were verified by DNA sequencing (Genewiz).

### 2.2.2 Gld2 expression and purification

*Escherichia coli* BL21 (DE3) RIL cells (Agilent) were transformed with pET20b-Gld2 and grown in LB medium containing ampicillin (100 µg/mL) and chloramphenicol (34 µg/mL) at 37°C until OD<sub>600</sub> = 0.6. The temperature was lowered to 20°C and isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 250 µM to induce protein expression. Cells were harvested after 19 hours by centrifugation and suspended in Buffer A (10 mM Tris-HCl, pH 8.0, 200 mM KCl, 5 mM MgCl<sub>2</sub>). Cells were broken by the addition of lysozyme followed by sonication on ice. The cell lysate was centrifuged for 45 minutes at 15,000g and 4°C. Cell free extract was loaded onto HisPur Ni-NTA resin (Thermo Scientific) equilibrated with Buffer A. The resin was washed with Buffer B (10 mM Tris-HCl, pH 8.0, 200 mM KCl, 5 mM MgCl<sub>2</sub>, 10 mM imidazole) and proteins were eluted with Buffer C (10 mM Tris-HCl [pH 8.0], 200 mM KCl, 5 mM MgCl<sub>2</sub>, 250 mM imidazole). The elution fractions containing Gld2 were dialyzed against Buffer D (10 mM Tris-HCl, pH 8.0, 200 mM KCl, 5 mM MgCl<sub>2</sub>, 10% glycerol) overnight at 4°C with gentle mixing and stored at -80°C. Gld2 protein concentration was determined by a Bradford test.

### 2.2.3 Size exclusion chromatography

A 200 µL sample containing Gld2 in Buffer A was passed through an ENrich SEC 650 high-resolution size exclusion column (Bio-Rad) equilibrated with Buffer A. The flow rate was 0.75 mL/min and 1 mL fractions were collected upon injection of the sample. The fractions that corresponded to the peaks on the absorbance graph were

subjected to ammonium sulfate protein precipitation and analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) on a 15% polyacrylamide gel. For the purposes of calibration, vitamin B12 ( $M_r \approx 1.3$  kDa), bovine insulin ( $M_r \approx 6$  kDa), cytochrome c ( $M_r \approx 13.6$  kDa) bovine carbonic anhydrase ( $M_r \approx 30$  kDa), ovalbumin ( $M_r \approx 43$  kDa), BSA ( $M_r \approx 67$  kDa), ferritin ( $M_r \approx 440$  kDa) and thyroglobulin ( $M_r \approx 669$  kDa) were used as marker components and chromatographed under identical conditions.

#### 2.2.4 RNA substrate preparation

The following RNA substrates were purchased from Sigma Aldrich: monophosphorylated human let-7a-5p (5' (p)-UGAGGUAGUAGGUUGUAUAGUU-3'), unphosphorylated human let-7a-0p (5'-UGAGGUAGUAGGUUGUAUAGUU-3'), diphosphorylated human let-7a-5p-2p (5' (pp)-UGAGGUAGUAGGUUGUAUAGUU-3'), triphosphorylated human let-7a-5p-3p (5' (ppp)-UGAGGUAGUAGGUUGUAUAGUU-3'), the poly(A) tail mimic 15A RNA (5' (p)-AAAAAAAAAAAAAAAA-3'), human microRNA miR-122-5p (5' (p)-UGGAGUGUGACAAUGGUGUUUG-3'), and total *E. coli* tRNA. The coding sequence for a ribozyme-pre-let-7a with a T7 Promoter sequence was cloned into pUC19 for *in vitro* transcription with T7 RNA polymerase using the following primers (pre-let7 5'-CTAGATAATACGACTCACTATAGGGAGACTACTA CCTCACTGATGAGTCCGTGAGGACGAAACGGTACCCGGTATAGGTTGTATAG TTTTAGGGTCACACCCA ACTGGGAGATAACTATAACAATCTACTGTCTTTTCGAA -3' and pre-let-7rev 5'-ATCCTTCGAAAGACAGTAGATTGTATAGTTATCTCCCA G TGGTGGGTGTGACCCTAAACTATAACAACCTACTACGGGTACCGTTTCGTCCT CACGGACTCATCAGTGAGGTAGTAGTCTCCCTATAGTGAGTCGTATTA-3').

Primers were phosphorylated, annealed and cloned into pUC19 using *XbaI/BamHI*. For *in vitro* transcription, the DNA template was amplified using primers Theforward (5'-GTTGGGAAGGGCGATCGGTG-3') and let-7PCRrev (5'-GAAAGACAGTAGATTGT ATAG-3'). The PCR product was purified by phenol chloroform extraction, and RNA was transcribed with T7 RNA polymerase as described previously<sup>44</sup>. Upon transcription, the ribozyme constructs auto-cleaves into ribozyme and pre-let-7a. The transcripts were separated by electrophoresis on a 12% denaturing polyacrylamide gel, the band

corresponding to self-cleaved pre-let-7a was excised from the gel and eluted as described previously<sup>45</sup>.

### 2.2.5 Determination of enzymatic activity and substrate range

10  $\mu$ L reactions were carried out containing 100 nM Gld2 in Buffer D and 1  $\mu$ M of the respective RNA substrates. Dithiothreitol (DTT) and  $MgCl_2$  were added for a final concentration of 1 mM and 3.2 mM, respectively. [ $\alpha$ -<sup>32</sup>P]-UTP or [ $\alpha$ -<sup>32</sup>P]-ATP (Perkin Elmer) were used as indicated at a final concentration of 0.33  $\mu$ M. All reactions were incubated at 37°C for 20 minutes and stopped by the addition of 2 x RNA loading dye (95% v/v formamide, 0.1% w/v xylene xyanol, 0.1% w/w bromophenol blue, 10 mM ethylenediaminetetraacetic acid (EDTA)). Reactions were analyzed via gel electrophoresis on a 12% denaturing polyacrylamide gel and visualized with a phosphorimager (Storm 860 Molecular Imager). The radiolabelled RNA Decade marker (Ambion) was used as reference.

### 2.2.6 Determination of enzyme kinetics

20  $\mu$ L reactions were carried out containing 100 nM Gld2 in Buffer D and 2  $\mu$ M of let-7a. DTT and  $MgCl_2$  were added to a final concentration of 1 mM and 3.2 mM, respectively. Separate reactions contained one of the four NTPs at various concentrations. Three technical replicates were performed for each nucleotide concentration. Higher nucleotide concentrations were achieved by using a mixture of [ $\alpha$ -<sup>32</sup>P]-labelled and unlabelled nucleotides. Nucleotide concentrations were adjusted so that in an 8-minute time course the reaction progressed linearly and < 10% of the total substrate RNA was converted to product. The following nucleotide concentrations were used: ATP 0 - 15  $\mu$ M (0  $\mu$ M, 0.01  $\mu$ M, 0.1  $\mu$ M, 1  $\mu$ M, 10  $\mu$ M, 15  $\mu$ M); UTP 0 - 1000  $\mu$ M (0  $\mu$ M, 0.01  $\mu$ M, 0.1  $\mu$ M, 1  $\mu$ M, 10  $\mu$ M, 50  $\mu$ M, 100  $\mu$ M, 300  $\mu$ M, 1000  $\mu$ M); GTP 0 - 1300  $\mu$ M (0  $\mu$ M, 1  $\mu$ M, 10  $\mu$ M, 50  $\mu$ M, 100  $\mu$ M, 300  $\mu$ M, 1000  $\mu$ M, 1300  $\mu$ M); CTP 0 - 1500  $\mu$ M (0  $\mu$ M, 1  $\mu$ M, 10  $\mu$ M, 100  $\mu$ M, 600  $\mu$ M, 1000  $\mu$ M, 1200  $\mu$ M, 1500  $\mu$ M). All reactions were incubated at 37°C for 30 seconds before Gld2 was added (at t = 0) and 5  $\mu$ L samples were then taken out at 2, 4, 6, and 8 minutes. The reactions were stopped with the addition of 2 x RNA loading dye. Reactions were analyzed via gel electrophoresis on a

12% denaturing polyacrylamide gel and visualized by phosphorimaging overnight. To quantify product formation, a strip of Whatman filter paper was spotted with different known concentrations for each [ $\alpha$ - $^{32}\text{P}$ ]-labelled nucleotide and imaged on the same phosphorimaging screen. Spot intensity was quantified using ImageJ. Kinetic constants were derived from plotting initial velocity ( $v_o$ ) against nucleotide concentration. Kinetics were fitted to the standard Michaelis–Menten curve using Kaleida Graph 3.1 (Synergy Software) and SigmaPlot (Systat Software). Error bars represent 1 Standard Deviation from 3 replicates.

### 2.2.7 Phylogenetic analysis

Sequences were downloaded from the National Center for Biotechnology Information (NCBI). Sequence alignment and alignment editing was carried out using Muscle<sup>46</sup> and the Multiseq alignment editor from VMD 1.8.7<sup>47</sup>. A maximum likelihood phylogeny for Gld2 sequences was determined using PhyML<sup>48</sup>. The starting tree was generated with BioNJ, and the tree space was searched with the SPR followed by the NNI algorithm to find the best tree. The JTT+ $\Gamma$  model with 4 rate categories was applied. Likelihood parameters were initially estimated from the alignment, Shimodaira–Hasegawa bootstrap values were computed as implemented in PhyML.

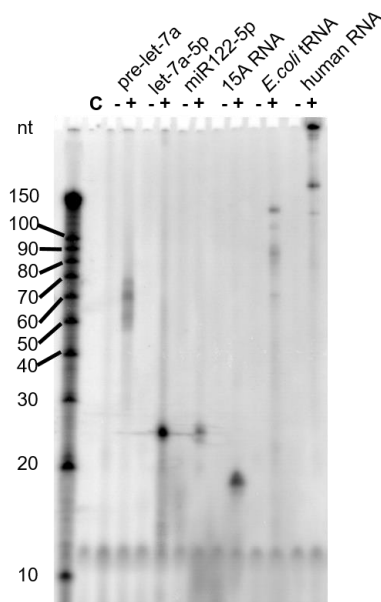
## 2.3 Results

### 2.3.1 Gld2 displays RNA substrate promiscuity

Gld2 has been implicated in multiple pathways of mRNA and miRNA regulation. To test whether Gld2 displays enzymatic activity *in vitro*, we recombinantly expressed and purified full length human Gld2 from *E. coli*. Gld2 eluted from a gel filtration column as a monomer. To evaluate Gld2 activity, we first assessed its substrate range. Previously suggested roles for Gld2 activity are in miRNA regulation, including miR-122 adenylation<sup>25</sup> and the uridylation of pre-let-7a and let-7a<sup>24,38</sup>. Another study linked Gld2 to mRNA adenylation<sup>49</sup>. We therefore included miR-122, pre-let-7a, let-7a, total human mRNA and a poly(A) tail mimic comprised of 15 adenines (15A) in our initial assays and

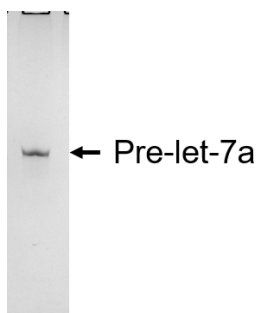
total *E. coli* tRNA as a control.

Surprisingly, Gld2 displayed adenylation activity on all substrates *in vitro* (Figure 2.2). For miRNAs and the 15A RNA a single band was observed, corresponding to monoadenylation of the substrate RNAs. For the premature miRNA substrate, pre-let-7a, a band around 70 bases was observed, indicating single nucleotide addition. While the *in vitro* transcribed pre-let-7a was purified as a single transcript (Figure 2.3), additional bands of lower molecular weight were observed. These bands are likely due to the adenylation of partially degraded substrate RNA or degradation during the enzymatic reaction. For both total tRNA and total mRNA multiple bands were observed, but due to heterogeneity of the substrates we are unable to confirm that these are single nucleotide additions. Nevertheless, no clusters of ladder-like nucleotide additions were observed in these cases, indicating that the heterogenic substrates were likely monoadenylated.



**Figure 2.2: RNA substrates of Gld2.**

Gld2 was incubated with different RNA substrates and [ $\alpha$ - $P^{32}$ ]-ATP as indicated. Formation of [ $\alpha$ - $P^{32}$ ]-labelled RNA products was monitored by electrophoretic separation and subsequent phosphorimaging. Gld2 catalyzed [ $\alpha$ - $P^{32}$ ]-ATP addition to RNA substrates pre-let-7a (72 nts), mature human miRNAs let-7a-5p (22 nts), miR122 (22 nts), an oligo (A) tail mimic 15A (15 nts), and total *E. coli* tRNA and total human mRNA. A representative gel is shown. C is no enzyme control.

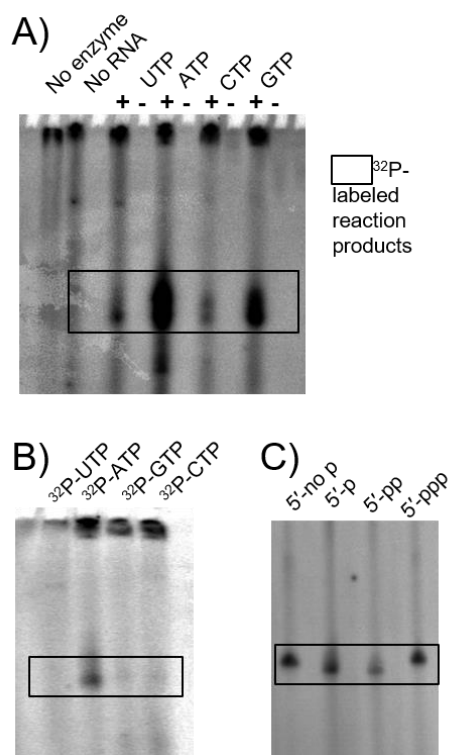


**Figure 2.3: *In vitro* transcription of pre-let-7 RNA.**

Pre-let-7a was transcribed as a ribozyme-fusion RNA, which self-cleaves after transcription. The band corresponding to pre-let-7a (72 nts) was excised from the gel, purified and a single band was observed in a denaturing polyacrylamide gel electrophoresis after purification.

### 2.3.2 Nucleotide preference of recombinant Gld2

We determined the specificity of Gld2 for all of the four nucleotides ATP, CTP, GTP and UTP. In enzyme assays containing a single nucleotide species, pre-let-7a, and Gld2, the enzyme was active with all NTPs (Figure 2.4A). Gld2 can accommodate each NTP in the active site and catalyze their 3' addition to pre-let-7a. In a competition assay, which included all four nucleotides in equimolar concentrations with only one [ $\alpha$ - $^{32}$ P]-labelled nucleotide, only ATP was added to pre-le-7a (Figure 2.4B). Thus, while Gld2 displays relaxed specificity towards RNA substrates *in vitro*, it exhibits a clear preference for ATP in the presence of all four NTPs. We further investigated whether the 5'-end of the RNA substrates influenced substrate recognition. No difference in enzyme activity was seen when differentially 5'-phosphorylated RNA substrates were assayed with [ $\alpha$ - $^{32}$ P]-ATP (Figure 2.4C).



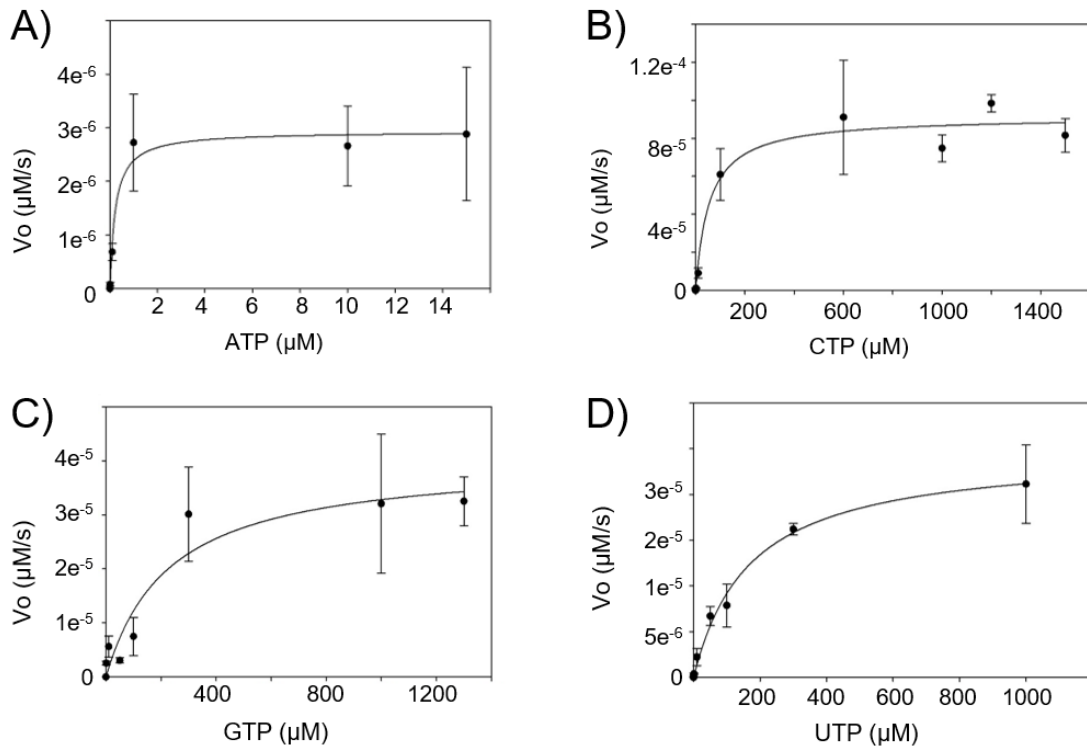
**Figure 2.4: Nucleotide substrates of Gld2.**

Gld2 was incubated with varying nucleotides and the precursor miRNA let-7a or mature miRNA let-7a-5p. **A)** Addition of single [ $\alpha$ - $^{32}$ P]-labelled nucleotides to pre-let-7a and Gld2 as indicated. **B)** Competitive nucleotide addition: In a competition assay, Gld2 was incubated with pre-let-7a in the presence of all four unlabelled nucleotides in equimolar amounts with a portion of the indicated nucleotide in an [ $\alpha$ - $^{32}$ P]-labelled form. **C)** 5' end phosphorylation: Gld2 activity on let-7a-5p substrates with differentially phosphorylated 5' ends were assayed with [ $\alpha$ - $^{32}$ P]-ATP. Representative gels are shown. 5'-no p = unphosphorylated, 5'-p = monophosphate, 5'-pp = diphosphate, 5'-ppp = triphosphate.

### 2.3.3 Gld2 is an adenylyltransferase

Our initial experiments showed that Gld2 catalyzes the addition of all four nucleotides to pre-let-7a (Figure 2.4A). In the subsequent competition assay, a clear preference for ATP was observed (Figure 2.4B). This led us to further investigate the nucleotide specificity of Gld2, and we performed a detailed kinetic characterization of Gld2 with all four nucleotides with let-7a (Figure 2.5, Table 2.1). Our kinetic analysis confirmed that Gld2 displays the most affinity for ATP ( $K_M \sim 0.2 \mu\text{M}$ ) and is most

efficient with ATP. Significantly increased  $K_M$  values for the other nucleotides, suggests far weaker binding affinity. Gld2 showed increases in  $K_M$  of  $> 700$ -fold for UTP, 240-fold for CTP, and  $> 1000$ -fold for GTP. The relative catalytic efficiency for the reactions indicates an 83-fold preference of ATP over UTP and 71-fold over GTP. Overall catalytic efficiency is greatest for ATP ( $k_{cat}/K_M = 12.8 \times 10^{-5} \mu\text{M}^{-1}\text{s}^{-1}$ ) with 12.9% relative efficiency for CTP ( $k_{cat}/K_M = 1.66 \times 10^{-5} \mu\text{M}^{-1}\text{s}^{-1}$ ), and 1.2% for UTP ( $k_{cat}/K_M = 0.15 \times 10^{-5} \mu\text{M}^{-1}\text{s}^{-1}$ ) and 1.4% for GTP ( $k_{cat}/K_M = 0.18 \times 10^{-5} \mu\text{M}^{-1}\text{s}^{-1}$ ). Taken together, these data indicate that Gld2 NTP specificity is determined by productive binding of Gld2 to the respective nucleotide. While ATP is preferred and outcompetes all other NTPs, no nucleotide is specifically excluded from the active site.



**Figure 2.5: Dependence of the reaction rate on nucleotide concentration.**

The plot shows the initial velocity of the enzyme reaction plotted against concentration of **A) ATP**, **B) CTP**, **C) GTP** and **D) UTP**. Three technical replicates were performed for each nucleotide concentration. Error bars show one standard deviation.



**Table 2.1: Nucleotide addition kinetics of Gld2.**

Relative catalytic efficiency is the relative percentage in  $k_{cat}/K_M$  that is calculated as the ratio of  $k_{cat}/K_M$  for the nucleotide listed in the far-left column over the  $k_{cat}/K_M$  for ATP. Physiological nucleotide conditions are derived from Traut *et al.*, 1994. Standard deviations are reported. Reaction conditions are given in Materials and Methods.

	$V_{max}$ ( $\mu\text{M/s}$ )	$K_M$ ( $\mu\text{M}$ )	$k_{cat}$ ( $\text{s}^{-1}$ )	$k_{cat}/K_M$ ( $\mu\text{M}^{-1}\text{s}^{-1}$ )
<b>ATP</b>	$(2.93 \pm 0.16) \times 10^{-6}$	$0.229 \pm 0.077$	$(2.93 \pm 0.16) \times 10^{-5}$	$12.8 \times 10^{-5}$
<b>UTP</b>	$(24.8 \pm 1.17) \times 10^{-6}$	$169 \pm 23.3$	$(24.8 \pm 1.17) \times 10^{-5}$	$0.15 \times 10^{-5}$
<b>CTP</b>	$(91.4 \pm 5.16) \times 10^{-6}$	$55.0 \pm 21.9$	$(91.4 \pm 5.16) \times 10^{-5}$	$1.66 \times 10^{-5}$
<b>GTP</b>	$(40.3 \pm 5.97) \times 10^{-6}$	$230 \pm 112$	$(40.3 \pm 5.97) \times 10^{-5}$	$0.18 \times 10^{-5}$
	Relative efficiency $k_{cat}/K_M$		Physiological concentrations	
<b>ATP</b>	100%		2102 $\mu\text{M}$	
<b>UTP</b>	1.2%		253 $\mu\text{M}$	
<b>CTP</b>	12.9%		91 $\mu\text{M}$	
<b>GTP</b>	1.4%		305 $\mu\text{M}$	

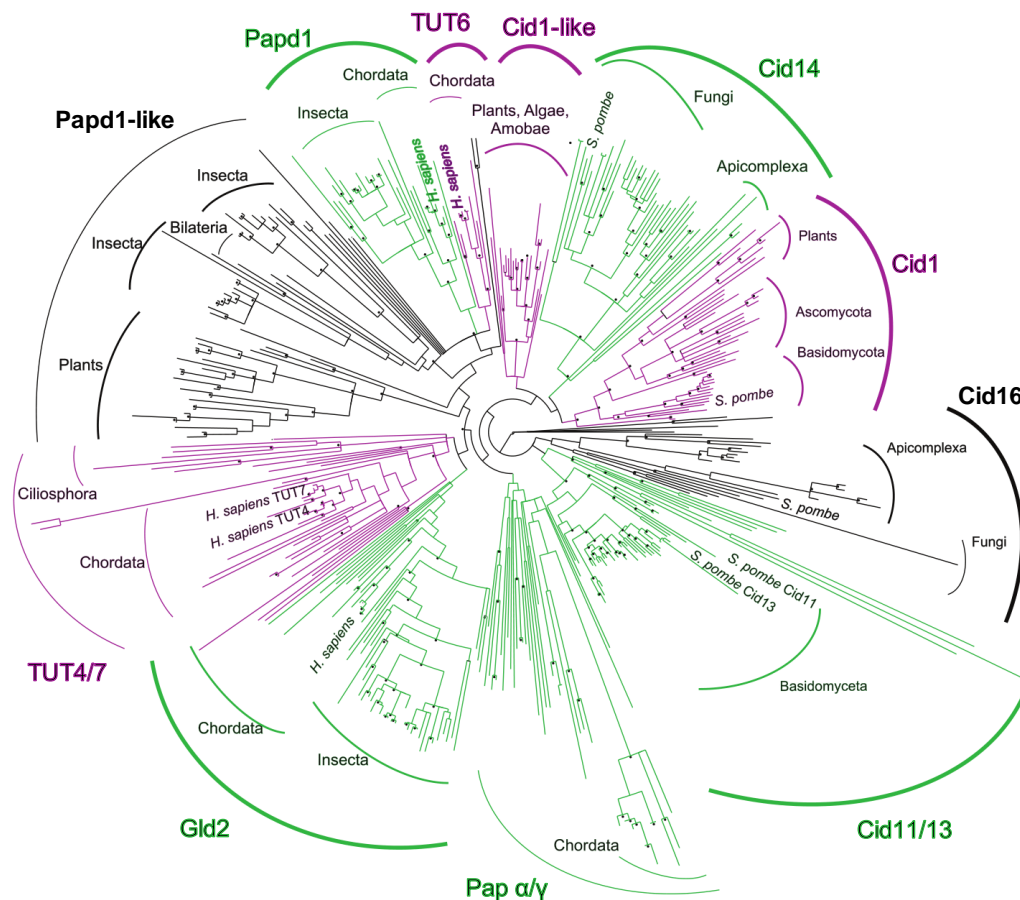
### 2.3.4 Nucleotidyltransferase specificity arose multiple times during evolution

Gld2 is a member of the non-canonical poly(A) polymerases, a diverse group of enzymes with varying RNA and nucleotide preferences. It includes enzymes with nucleotide specificity ranging from strictly ATP to ambiguous ATP or UTP, to exclusive UTP-adding enzymes. While some enzymes have a distinct substrate preference, such as the U6 snRNA uridylating enzyme TUT6<sup>28,50</sup>, which is both UTP and U6 snRNA specific, other enzymes are more promiscuous in their substrate specificity. TUT4 and TUT7, for example, have been shown to uridylate miRNAs<sup>24,36</sup>, as well as histone mRNA<sup>51</sup> and cytoplasmic mRNA<sup>30</sup>. The *Schizosaccharomyces pombe* homolog Cid1

protein was initially thought to act as an mRNA adenylylating enzyme but was later characterized as a uridylyltransferase with 1% residual adenylation activity<sup>52</sup>.

Gld2 was initially annotated as an adenylyltransferase, but recent *in vivo*<sup>24</sup> and *in vitro*<sup>24,25,38</sup> evidence and its sequence similarity to confirmed TUTs (30% and 32% amino acid sequence identity with TUT4 and TUT7, respectively) suggests a possible uridylyltransferase activity. The similar domain structure of uridylyl- and adenylyltransferases, as well as a high sequence similarity, suggests that these enzymes evolved from a common ancestor. To elucidate the phylogenetic background evolution of these enzymes, we performed a detailed phylogenetic analysis of the enzyme superfamily to trace the evolutionary origins of NTP specificity. The phylogeny of the nucleotidyltransferase family includes over 400 sequences (Figure 2.6).

Our analysis shows that nucleotidyltransferases display a classical star phylogeny, with several distinct subgroups. Interestingly, uridylyl- and adenylyltransferases do not form two separate clades, but rather Tutases emerge from distinct groups dominated by PAPs. Uridylyltransferases are derived from distinct subfamilies of adenylyltransferases, and Tutase activity, thus, evolved multiple times independently. One group shows the U6 snRNA uridylylating enzyme TUT6<sup>50</sup> evolved from a parent clade composed of known and putative adenylyltransferases. The nucleotide specificity of TUT6 is, however, not restricted to snRNA uridylation, as it was additionally found to adenylate selected mRNAs<sup>53,54</sup>. TUT6 is most closely related to the non-canonical mitochondrial poly(A) polymerase PAPD1, which mediates RNA decay by polyadenylation<sup>55</sup>. A second distinct subgroup includes the uridylyltransferase Cid1, which initiates RNA decay by uridylation<sup>29</sup>, and Cid14, a poly(A) polymerase that acts on rRNAs<sup>56</sup> and telomeres<sup>57</sup>.



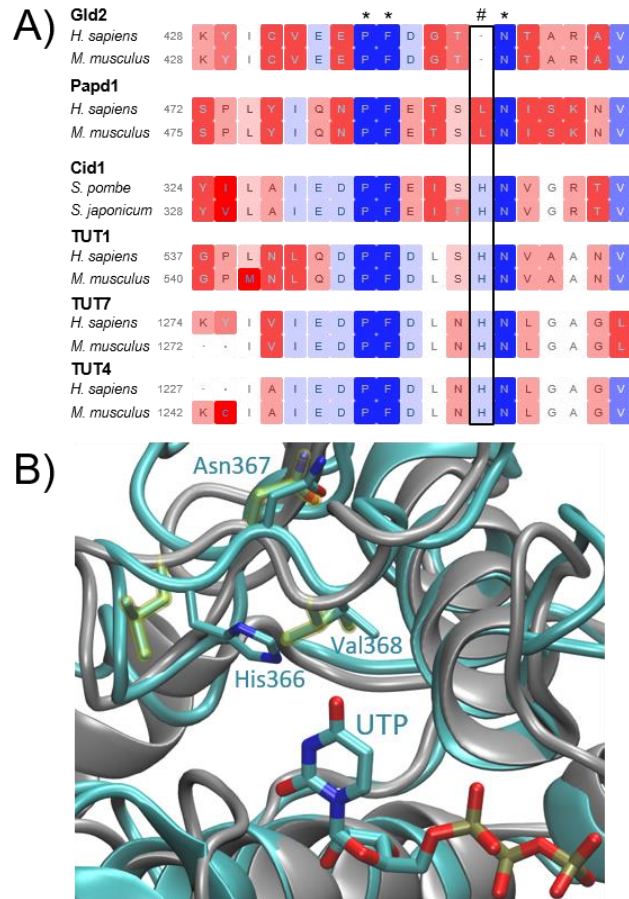
**Figure 2.6: Phylogeny of the Tutase/PAP superfamily.**

Gld2 is most closely related to the genuine uridylyltransferases TUT4/TUT7. Known enzymatic activities are colour-coded: purple, UTP preference; green, ATP preference; black, unknown. Bootstrap values over 90% are denoted with a star. Sequence data were downloaded from the Integrated Microbial Genomes database. The tree was calculated with PHYML using a BioNJ starting tree and SPR tree search followed by NNI branch swapping to optimize the tree. Bootstrap values were computed according to the Shimodaira–Hasegawa re-estimation of log-likelihood test implemented in PHYML.

### 2.3.5 Insertion of a histidine residue confers UTP specificity

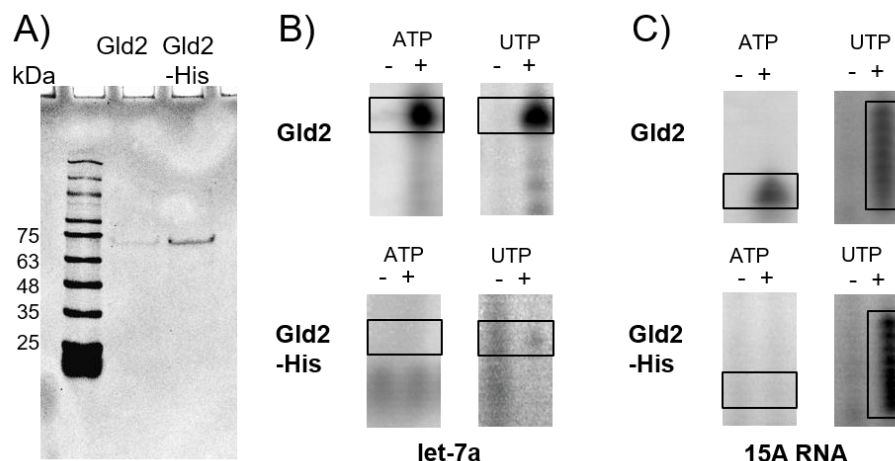
Gld2 displays clear nucleotide specificity for ATP, as demonstrated by kinetic analysis of Gld2 activity on all four nucleotides and in competition experiments. Our phylogenetic analysis of the nucleotidyltransferase family suggests that uridylyltransferase activity evolved from adenylyltransferases multiple times during

evolution (Figure 2.6). Previous reports showed that the mutation of the *S. pombe* Cid1 uridylyltransferase active site histidine (H336, Figure 2.7) to asparagine broadens its substrate specificity to include ATP<sup>58</sup>. In Gld2, the corresponding amino acid is lacking (Figure 2.7A). Consequently, we tested whether an insertion of a histidine residue at the position homologous to Cid1 position H336 (between Gld2 amino acids T439 and N440 (Figure 2.7B)) confers UTP specificity over ATP. Purified recombinant Gld2 and Gld2-His (Figure 2.8A) were tested with RNA substrates let-7a (Figure 2.8B) and 15A (Figure 2.8C) and radiolabelled ATP or UTP. As expected, Gld2 displayed activity with ATP and UTP for both substrates in a non-competitive assay (Figures 2.8B panel 1, and 2.8C panel 1 and Figure 2.4A). The Gld2-His insertion variant, however, displayed significantly decreased activity with ATP. For let-7a, no activity was observed with ATP, while a band is visible when incubated with UTP. Similarly, we observed no band for the poly(A) tail mimic 15A when incubated with ATP, while incubation with UTP lead to a ladder-like addition of nucleotides. In all cases, unlike UTP, ATP was excluded as a substrate. Thus, the insertion of a histidine residue homologous to Cid1 H336 conferred UTP selectivity in Gld2.



**Figure 2.7: Nucleotide preference is defined by a histidine residue.**

**A)** Excerpt from a multiple sequence alignment of 440 nucleotidyltransferase sequences. After the alignment sequences were grouped into phylogenetic subgroups and two representative sequences are depicted. Sequence similarities within each group were colour-coded. Confirmed nucleotidyltransferases Cid1, TUT4, TUT6 and TUT7 and their homologs contain a histidine (#), which confers uridyl selectivity. Adenylyltransferases either substitute histidine with a smaller amino acid (leucine in PAPD1) or entirely lack this residue (Gld2). **B)** Excerpt of a structural alignment between Cid1 (pdb 4FHP) and PAPD1 (pdb 3PQ1). The depicted ribbon diagram shows a structural superposition of human mitochondrial PAPD1 (grey) and Cid1 (cyan). Asn367 of Cid1 aligns with the homologous Asn273 of PAPD1. Cid1 His336 points into the active site and makes contact with UTP.



**Figure 2.8: Insertion of an active site histidine into Gld2 alters nucleotide specificity.**

**A)** Gld2 and Gld2-His purification. Recombinant His-tagged Gld2 and the Gld2-His mutant were purified via Ni-NTA chromatography to apparent homogeneity as judged by SDS-PAGE. **B)** and **C):** Nucleotide preference of Gld2 and Gld2-His. Recombinant enzymes were incubated with and without RNA substrates and radiolabelled ATP or UTP. Formation of [ $\alpha$ - $^{32}$ P]-labelled RNA products was monitored by electrophoretic separation and subsequent phosphorimaging. RNA substrates were **B)** let-7a and **C)** 15A. Representative gels are shown.

## 2.4 Discussion

### 2.4.1 Gld2 is an adenylyltransferase

Previous studies presented evidence associating Gld2 with monoadenylation<sup>25</sup> or monouridylation<sup>24,25,38</sup> of miRNAs in humans. While monoadenylation confers increased miRNA stability, monouridylation is a required step in biogenesis of Group II miRNAs. During maturation, Group II miRNAs are processed into pre-miRNAs with a single nucleotide 3'-overhang. TUT4, TUT7, and potentially Gld2 add an essential uridyl residue to 1 nt 3'-overhang pre-miRNA to yield a 2 nucleotide 3'-overhang, which is a prerequisite for processing by Dicer<sup>24,42</sup>. In *Xenopus laevis* and *Caenorhabditis elegans*, Gld2 is required for polyadenylation of specific mRNAs, aided by RNA binding proteins such as CPEB and Gld3, respectively<sup>31,38</sup>.

To elucidate the nucleotide preference of human Gld2, we carried out a detailed enzyme kinetic analysis. Using purified enzyme, we demonstrated that Gld2 is an adenylyltransferase preferentially adding single nucleotides to small RNAs. The catalytic efficiency of Gld2 is reduced 83-fold for UTP in comparison to ATP. Conversely, the uridylyltransferase Cid1 displays a 100-fold higher specificity for UTP over ATP<sup>58,59</sup>. Our data shows a clear preference of ATP over all other nucleotides. The rate-limiting step is most likely the NTP binding event. The  $K_M$  for ATP is 0.23  $\mu\text{M}$ , which is about 10,000-fold lower than the cellular ATP concentration of 2.1 mM<sup>60</sup>. Thus, Gld2 encounters a vast excess of ATP in the cell and will attain maximal substrate turnover. Cellular UTP, GTP, and CTP concentrations are overall lower than ATP concentrations<sup>60</sup>. In addition, we measured  $K_M$  values in the 100  $\mu\text{M}$  range for the other NTPs. Physiological concentrations for UTP and GTP are between 250-300  $\mu\text{M}$ <sup>60</sup>, which is between 1 and 6-fold higher than the respective  $K_M$ . The overall higher  $K_M$  of Gld2 for UTP, CTP, and GTP combined with a much higher cellular ATP concentration further shifts enzymatic activity towards adenylation. Interestingly, at a CTP concentration 1000-fold in excess of the physiological concentration the catalytic turnover for CTP is 30-fold higher than for ATP. Once NTP binding occurs, the other NTPs are ligated to the substrate RNA more rapidly than ATP (Table 2.1). Nevertheless, the low affinities of Gld2 for nucleotides other than ATP and the fact the other NTPs fail to outcompete ATP (Figure 2.4B) indicate Gld2 activity is shifted to adenylation under physiological conditions.

Our data indicate that Gld2 may not have evolved to function exclusively with ATP, as we could show that Gld2 is active with UTP as well as GTP and CTP, albeit with low efficiency. While we cannot exclude that post-translational modifications or interactions with other proteins may influence nucleotide specificity, our data suggests that the observed NTP specificity is only in part determined by binding constants. Gld2 nucleotide specificity alone provides an 83-fold preference for ATP over UTP, which is combined with a cellular environment that has 10-fold excess of ATP compared to the other NTPs. Gld2 substrate preference and the cellular nucleotide concentrations together increase Gld2 selectivity towards ATP to >800-fold, enhancing the enzyme's specificity without evolving a precisely selective adenylyltransferase.

#### 2.4.2 Gld2 monoadenylates small RNA substrates

In *in vitro* activity assays, Gld2 monoadenylates a variety of small RNAs, and does not specifically discriminate between different substrate RNAs. Our data shows that Gld2 displays activity on tRNAs, miRNAs, pre-miRNAs and mRNA alike, with a slight preference for miRNAs (Figure 2.2). This agrees with earlier studies of *X. laevis* and *C. elegans* Gld2 homologs, which showed a role for Gld2 in miRNA metabolism<sup>32</sup>. The RNA substrate promiscuity that we observed with human Gld2 may be related to the fact that this minimal adenylyltransferase lacks RNA binding domains, which are thought to confer substrate specificity in Gld2's closest human homologs, TUT4 and TUT7 (Figure 2.1). Interestingly, we found no evidence of polyadenylation activity on any of the employed substrates. For processive polyadenylation, human Gld2 most likely requires the assistance of RNA proteins *in vivo*, which may confer specificity or activate elongation. Intriguingly, an extended incubation of Gld2 or Gld2-His with 15A RNA and UTP lead to a ladder-like addition of nucleotides (Figure 2.8C). While UTP is not the natural Gld2 substrate, it is possible that polyuridylation, in contrast to polyadenylation, does not require accessory proteins.

Several studies from non-human Gld2 homologs show that in the presence of RNA binding proteins such as CPEB<sup>38</sup> and Gld3<sup>31</sup> in *X. laevis* and *C. elegans*, respectively, Gld2 can processively add multiple adenine residues. To date, no such interaction has been shown for human Gld2, but the fact that Gld2 alone hardly discriminates between several RNA substrates suggests the regulation of substrate specificity by additional RNA binding proteins. While humans lack a Gld3 homolog, several CPEB homologs are encoded in the genome. In *X. laevis*, an additional poly(A) polymerase Gld4 enzyme polyadenylates p53 mRNA in a CPEB dependent manner. In this case, Gld2 is not associated with CPEB directly, but regulates its expression via miR-122 adenylation<sup>39</sup>. In *C. elegans* the RNA binding protein Gld3 stimulates Gld2 catalyzed polyadenylation by increasing its affinity to the substrate RNA<sup>32</sup>. Similarly, TUT7 can be triggered to polyuridylate RNA substrates in association with the RNA binding protein Lin28A<sup>23,61</sup>. Potential Gld2 associated proteins, however, remain to be identified.



### 2.4.3 Convergent evolution of Tutase activity by histidine insertion in the PAP active site

Recent biochemical and structural data on fission yeast Cid1<sup>58,59</sup> and *X. laevis* TUT7<sup>62</sup> suggested that the nucleotide preference of nucleotidyltransferases is determined by a single histidine near the active site (Figure 2.7A and B). In Cid1, histidine 336, which is located on a flexible loop near the catalytic site, sterically excludes ATP from the active site. Prior experiments have shown that mutation of this histidine to a smaller amino acid broadens nucleotide specificity in uridylyltransferases to include ATP, concluding that an asparagine to histidine mutation confers ATP specificity<sup>58,59,62</sup>. Our multiple sequence alignment (Figures 2.7B) and structural superposition of Cid1 and human mitochondrial PAPD1 (Figure 2.7B) show that the Asn337 in Cid1 is homologous to Asn273 residue of PAPD1. This Asn is strictly conserved in all members of the nucleotidyltransferase superfamily. The multiple sequence alignment clearly shows that Cid1 His336 is an inserted residue relative to the PAPD1 homolog (Figure 2.7A). His336 points directly into the active site in the structural superposition, making contact with the UTP substrate, while the Asn337/Asn273 residue is oriented away from the active site (Figure 2.7B). The Asn conformation is structurally conserved between Tutases and PAPs. Consequently, UTP selectivity of Gld2, which is phylogenetically derived from TUT4/7 (Figure 2.6), is conferred by a histidine insertion, rather than an Asn to His mutation as previously suggested<sup>58,59,62</sup>.

The here presented phylogenetic analysis shows that the presence/absence of the TUT-specific histidine is consistent within all nucleotidyltransferase groups and coincides with the divergence of uridylyltransferases from adenylyltransferases. While previous studies alluded to a point mutation from histidine (denoted with # in Figure 2.7B) to asparagine<sup>58</sup>, conferring ATP over UTP selectivity, our sequence alignment clearly shows that the respective asparagine (denoted with a \* in Figure 2.7B) is part of a highly conserved motif found in all nucleotidyltransferases. Thus, a histidine insertion rather than a mutation confers uridine specificity. In agreement with our biochemical data, the respective residue is absent from the Gld2 amino acid sequence, which vacates the active site for the larger ATP. Similarly, Cid14 is also lacking the respective histidine

residue, conferring a preference of ATP over UTP. Consequently, a histidine insertion into Gld2 switched the nucleotide preference from ATP to UTP, excluding the larger ATP from the active site. Gld2-His shows no activity towards ATP but is active on multiple substrates with UTP (Figure 2.8B and C). These data clearly show that a single amino acid insertion can change nucleotide specificity in nucleotidyltransferases.

The human uridylyltransferases TUT4 and TUT7 are closely grouped within the phylogeny and are likely the result of recent gene duplication in the chordata lineage. Consistent with their amino acid sequence similarity, Gld2 is most closely related to the TUT4/7 phylogenetic group, although Gld2 has the conserved Tutase histidine deleted. In adenylyltransferases, such as human PAPD1 and its homologs, leucine replaces histidine. How a leucine insertion impacts enzyme activity and substrate specificity remains to be elucidated. Considering that uridylyltransferase activity has evolved multiple times, the flexible loop which harbors histidine 336 in Cid1 can be denoted as a preferred spot for mutations and insertions, allowing for facile alterations in substrate specificity. Once RNA specificity has been determined, the ability to change nucleotide specificity through this preferred spot would allow the cell to easily obtain different enzymes that differentially modify the same RNAs.

The phylogeny, supported by our mutational analysis of Gld2, shows that uridylyltransferase activity diverged from adenylyltransferase activity multiple times during evolution and prior to the split of slime molds and bilateria. Interestingly, non-canonical nucleotidyltransferases are very prevalent in Fungi, which include several nucleotidyltransferase clades (e.g., Cid16, Cid11/13) not found in other organisms. These homologs appear to result from initial gene duplication, giving rise to Cid16 and Cid11/13 groups, with a more recent duplication leading to the divergence of Cid11 and Cid13. The nuclear poly(A) polymerase Cid11<sup>63</sup> and the Cid13 homolog encode an Arg, while Cid16 homologs a Lys in position His336 of Cid1. The biological function of these enzymes is not entirely clear, but the small positively charged side chain suggests a role in RNA adenylation, as ATP could still be spatially accommodated in the active site. One nucleotidyltransferase clade of entirely unknown function contains PAPD1-like proteins with homologs in *Drosophilidae* and plants. No data are available on nucleotide

specificity or biological function of these proteins. This group is diverse in its active site constituents. Interestingly, members of the *Drosophilidae* encode two PAPD1-like homologs. One homolog encodes an arginine, and one a histidine, suggesting a recent gene duplication to allow for both uridylyltransferase and adenylyltransferase activity. These homologs could potentially share a specific RNA substrate range or interacting protein partners, while differing in nucleotide preference, thus fulfilling distinct biological functions.

## 2.5 Acknowledgements

This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada to IUH (RGPIN 04776-2014) and a Seed Grant for CIHR success from The University of Western Ontario. We thank Julia Jaramillo, Mathura Vithyananthan, Riddhi Desai, and Kunmo Kim for their support, Susanna George, Nileeka Balasuriya, and Yumin Bi for their advice, and Patrick O'Donoghue and Murray Junop for critical discussions.

## 2.6 References

1. Arasu, P., Wightman, B. & Ruvkun, G. Temporal regulation of lin-14 by the antagonistic action of two other heterochronic genes, lin-4 and lin-28. *Genes Dev.* **5**, 1825–1833 (1991).
2. Lee, R. C., Feinbaum, R. L. & Ambros, V. The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75**, 843–854 (1993).
3. Grosshans, H. & Slack, F. J. Micro-RNAs: small is plentiful. *J. Cell Biol.* **156**, 17–21 (2002).
4. Friedman, R. C., Farh, K. K.-H., Burge, C. B. & Bartel, D. P. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* **19**, 92–105 (2009).
5. Ameres, S. L. & Zamore, P. D. Diversifying microRNA sequence and function. *Nat. Rev. Mol. Cell Biol.* **14**, 475–488 (2013).
6. Sun, K. & Lai, E. C. Adult-specific functions of animal microRNAs. *Nat. Rev. Genet.* **14**, 535–548 (2013).
7. Takahashi, R.-U., Miyazaki, H. & Ochiya, T. The role of microRNAs in the regulation of cancer stem cells. *Front. Genet.* **4**, 295 (2014).

8. Kozomara, A. & Griffiths-Jones, S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.* **42**, D68–73 (2014).
9. Cammarata, G. *et al.* Differential expression of specific microRNA and their targets in acute myeloid leukemia. *Am. J. Hematol.* **85**, 331–339 (2010).
10. Kobayashi, M. *et al.* Ovarian cancer cell invasiveness is associated with discordant exosomal sequestration of Let-7 miRNA and miR-200. *J. Transl. Med.* **12**, 4 (2014).
11. Mulrane, L., McGee, S. F., Gallagher, W. M. & O'Connor, D. P. miRNA dysregulation in breast cancer. *Cancer Res.* **73**, 6554–6562 (2013).
12. Zhu, X.-M., Wu, L.-J., Xu, J., Yang, R. & Wu, F.-S. Let-7c microRNA expression and clinical significance in hepatocellular carcinoma. *J. Int. Med. Res.* **39**, 2323–2329 (2011).
13. Van der Ree, M. H. *et al.* Miravirsin dosing in chronic hepatitis C patients results in decreased microRNA-122 levels without affecting other microRNAs in plasma. *Aliment. Pharmacol. Ther.* **43**, 102–113 (2016).
14. Small, E. M. & Olson, E. N. Pervasive roles of microRNAs in cardiovascular biology. *Nature* **469**, 336–342 (2011).
15. Barh, D., Malhotra, R., Ravi, B. & Sindhurani, P. MicroRNA let-7: an emerging next-generation cancer therapeutic. *Curr. Oncol. Tor. Ont* **17**, 70–80 (2010).
16. Di Leva, G. & Croce, C. M. miRNA profiling of cancer. *Curr. Opin. Genet. Dev.* **23**, 3–11 (2013).
17. Fendler, A. & Jung, K. MicroRNAs as new diagnostic and prognostic biomarkers in urological tumors. *Crit. Rev. Oncog.* **18**, 289–302 (2013).
18. Zheng, H., Liu, J.-Y., Song, F.-J. & Chen, K.-X. Advances in circulating microRNAs as diagnostic and prognostic markers for ovarian cancer. *Cancer Biol. Med.* **10**, 123–130 (2013).
19. Moitra, K. Overcoming Multidrug Resistance in Cancer Stem Cells. *BioMed Res. Int.* **2015**, 635745 (2015).
20. Yates, L. A., Norbury, C. J. & Gilbert, R. J. C. The long and short of microRNA. *Cell* **153**, 516–519 (2013).
21. Burroughs, A. M. *et al.* A comprehensive survey of 3' animal miRNA modification events and a possible role for 3' adenylation in modulating miRNA targeting effectiveness. *Genome Res.* **20**, 1398–1410 (2010).
22. Wyman, S. K. *et al.* Post-transcriptional generation of miRNA variants by multiple nucleotidyl transferases contributes to miRNA transcriptome complexity. *Genome Res.* **21**, 1450–1461 (2011).
23. Heo, I. *et al.* TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. *Cell* **138**, 696–708 (2009).
24. 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).
25. 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).
26. Thornton, J. E. *et al.* Selective microRNA uridylation by Zcchc6 (TUT7) and Zcchc11 (TUT4). *Nucleic Acids Res.* **42**, 11777–11791 (2014).
27. Norbury, C. J. Cytoplasmic RNA: a case of the tail wagging the dog. *Nat. Rev. Mol. Cell Biol.* **14**, 643–653 (2013).

28. 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).
29. 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).
30. Lim, J. *et al.* Uridylation by TUT4 and TUT7 Marks mRNA for Degradation. *Cell* **159**, 1365–1376 (2014).
31. Wang, L., Eckmann, C. R., Kadyk, L. C., Wickens, M. & Kimble, J. A regulatory cytoplasmic poly(A) polymerase in *Caenorhabditis elegans*. *Nature* **419**, 312–316 (2002).
32. Kwak, J. E., Wang, L., Ballantyne, S., Kimble, J. & Wickens, M. Mammalian GLD-2 homologs are poly(A) polymerases. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 4407–4412 (2004).
33. Cui, J., Sartain, C. V., Pleiss, J. A. & Wolfner, M. F. Cytoplasmic polyadenylation is a major mRNA regulator during oogenesis and egg activation in *Drosophila*. *Dev. Biol.* **383**, 121–131 (2013).
34. Kwak, J. E. *et al.* GLD2 poly(A) polymerase is required for long-term memory. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 14644–14649 (2008).
35. Lee, M. *et al.* Adenylation of maternally inherited microRNAs by Wispy. *Mol. Cell* **56**, 696–707 (2014).
36. Kwak, J. E. & Wickens, M. A family of poly(U) polymerases. *RNA* **13**, 860–867 (2007).
37. 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).
38. Kim, J. H. & Richter, J. D. Opposing polymerase-deadenylase activities regulate cytoplasmic polyadenylation. *Mol. Cell* **24**, 173–183 (2006).
39. 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).
40. Glahder, J. A. & Norrild, B. Involvement of hGLD-2 in cytoplasmic polyadenylation of human p53 mRNA. *APMIS* **119**, 769–775 (2011).
41. Rissland, O. S., Mikulasova, A. & Norbury, C. J. Efficient RNA polyuridylation by noncanonical poly(A) polymerases. *Mol. Cell. Biol.* **27**, 3612–3624 (2007).
42. Kim, B. *et al.* TUT7 controls the fate of precursor microRNAs by using three different uridylation mechanisms. *EMBO J.* **34**, 1801–1815 (2015).
43. 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 N. Y. N* **18**, 1875–1885 (2012).
44. Milligan, J. F., Groebe, D. R., Witherell, G. W. & Uhlenbeck, O. C. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res.* **15**, 8783–8798 (1987).
45. Köhrer, C. & RajBhandary, U. L. The many applications of acid urea polyacrylamide gel electrophoresis to studies of tRNAs. *Methods San Diego Calif* **44**, 129–138 (2008).
46. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).

47. Roberts, E., Eargle, J., Wright, D. & Luthey-Schulten, Z. MultiSeq: unifying sequence and structure data for evolutionary analysis. *BMC Bioinformatics* **7**, 382 (2006).
48. Guindon, S. & Gascuel, O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**, 696–704 (2003).
49. Sartain, C. V., Cui, J., Meisel, R. P. & Wolfner, M. F. The poly(A) polymerase GLD2 is required for spermatogenesis in *Drosophila melanogaster*. *Dev. Camb. Engl.* **138**, 1619–1629 (2011).
50. Trippe, R. *et al.* Identification, cloning, and functional analysis of the human U6 snRNA-specific terminal uridylyl transferase. *RNA N. Y. N* **12**, 1494–1504 (2006).
51. Schmidt, M.-J., West, S. & Norbury, C. J. The human cytoplasmic RNA terminal U-transferase ZCCHC11 targets histone mRNAs for degradation. *RNA N. Y. N* **17**, 39–44 (2011).
52. Rissland, O. S. & Norbury, C. J. The Cid1 poly(U) polymerase. *Biochim. Biophys. Acta* **1779**, 286–294 (2008).
53. Mellman, D. L. *et al.* A PtdIns4,5P2-regulated nuclear poly(A) polymerase controls expression of select mRNAs. *Nature* **451**, 1013–1017 (2008).
54. Li, W. *et al.* Star-PAP control of BIK expression and apoptosis is regulated by nuclear PIPKI $\alpha$  and PKC $\delta$  signaling. *Mol. Cell* **45**, 25–37 (2012).
55. 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).
56. Win, T. Z. *et al.* Requirement of Fission Yeast Cid14 in Polyadenylation of rRNAs. *Mol. Cell. Biol.* **26**, 1710–1721 (2006).
57. Bah, A., Wischnewski, H., Shchepachev, V. & Azzalin, C. M. The telomeric transcriptome of *Schizosaccharomyces pombe*. *Nucleic Acids Res.* **40**, 2995–3005 (2012).
58. 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).
59. Yates, L. A. *et al.* Structural Basis for the Activity of a Cytoplasmic RNA Terminal U-transferase. *Nat. Struct. Mol. Biol.* **19**, 782–787 (2012).
60. Traut, T. W. Physiological concentrations of purines and pyrimidines. *Mol. Cell. Biochem.* **140**, 1–22 (1994).
61. 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).
62. Lapointe, C. P. & Wickens, M. The nucleic acid-binding domain and translational repression activity of a *Xenopus* terminal uridylyl transferase. *J. Biol. Chem.* **288**, 20723–20733 (2013).
63. Saitoh, S. *et al.* Cid13 is a cytoplasmic poly(A) polymerase that regulates ribonucleotide reductase mRNA. *Cell* **109**, 563–573 (2002).

## Chapter 3

### 3 Gld2 activity is regulated by phosphorylation in the N-terminal domain

#### 3.1 Introduction

MicroRNAs (miRNAs) are critical regulators of gene expression that are essential to human life, normal cellular function, and development. De-regulation of miRNAs is, perhaps not surprisingly, associated with a number of human diseases<sup>1</sup>. MiRNAs regulate the expression of many genes, including oncogenes, by complementary base pairing with the 3'-untranslated regions (UTRs) of mRNAs, which normally inhibits protein synthesis<sup>1</sup>.

MiRNAs themselves are post-transcriptionally regulated by the addition of single or multiple adenine (A) or uridine (U) residues to their 3'-ends. This untemplated RNA editing is now recognized as an important mechanism regulating cellular miRNA homeostasis<sup>2,3</sup>. The addition of a single A to the 3'-end on certain miRNAs leads to increased stability<sup>4</sup>. Conversely, the activity of mature miRNAs is reduced by the addition of a single 3'-U residue<sup>5,6</sup>. The addition of multiple U residues to precursor miRNAs (pre-miRNAs) triggers subsequent degradation by the U-specific exonuclease Dis3L2 (Dis3-like 3'-5' exonuclease 2)<sup>3</sup>. Although uridylation is commonly associated with silencing and degradation of RNAs, monouridylation of Group II pre-miRNAs lacking a critical 3'-end overhang nucleotide is required for miRNA maturation and processing by Dicer<sup>7</sup>. Cellular mechanisms that regulate miRNAs through 3'-terminal nucleotide additions are of fundamental relevance to the molecular basis of diseases characterized by de-regulated miRNA metabolism<sup>3,8</sup>.

A diverse family of terminal RNA nucleotidyltransferases (TENTs) catalyzes 3'-A and U additions to RNAs in human cells. The nucleotidyltransferase Gld2 (germline-development 2, TENT2) was first identified as a regulator of meiosis in *Caenorhabditis elegans*<sup>9</sup> and was later shown to extend the poly(A) tails of mRNAs (Figure 3.1A), leading to enhanced mRNA stability and increased abundance of the encoded protein<sup>10</sup>.

In humans, Gld2 stabilizes miR-122 in the liver and fibroblasts through monoadenylation<sup>4,11</sup> and mRNAs via polyadenylation<sup>12</sup> (Figure 3.1A).

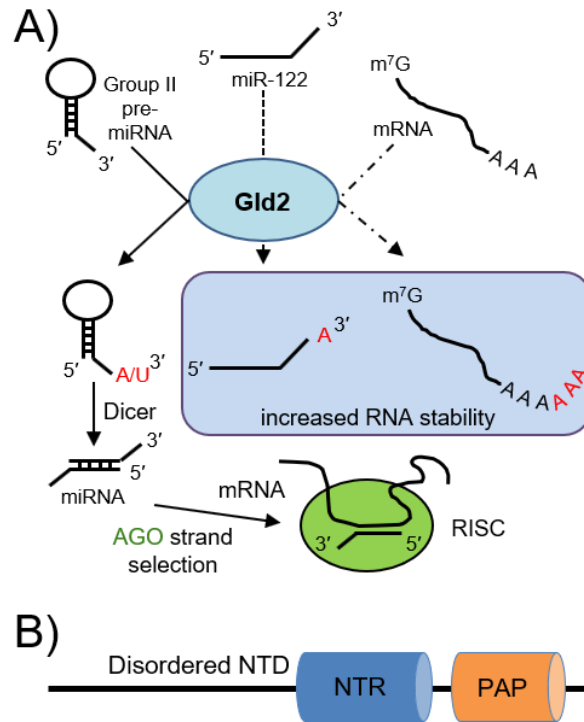
Gld2 is thought to be part of a larger protein complex involved in RNA modification and germ cell formation<sup>13</sup>. Although some reports<sup>7</sup> suggested that Gld2 may function as a uridylyltransferase, we recently characterized human Gld2 as a *bona fide* adenylyltransferase *in vitro*<sup>14</sup>. Our data confirmed a basal activity of Gld2 with UTP, but the 80-fold higher catalytic efficiency for ATP makes the enzyme strongly selective for A additions<sup>14</sup>. Gld2 encodes a nucleotidyltransferase domain and a poly(A) polymerase-associated domain that are required for catalytic activity as well as a predicted disordered N-terminal domain of unknown function<sup>10</sup> (Figure 3.1B), yet lacks identifiable RNA binding motifs. The crystal structure of a truncated *C. elegans* Gld2 in complex with the interacting protein Gld3 shows that the two essential Gld2 catalytic domains share the same fold as other nucleotidyltransferases<sup>15</sup>.

Cellular mechanisms that regulate miRNAs through 3'-end nucleotide additions are of fundamental relevance to the molecular basis of diseases characterized by de-regulated miRNA metabolism<sup>3,8</sup>. Gld2 and its substrate miR-122 play a role in Hepatitis C virus (HCV) infection and in hepatic cancer<sup>16</sup>. MiR-122 is one of the most abundant miRNAs in the liver, with an essential role in maintaining liver homeostasis and differentiation<sup>16</sup>. During HCV infection, miR-122 binds to two sites in the viral 5'-UTR of the Hepatitis C viral RNA and is required for HCV infection<sup>16,17</sup>. The miR-122 interaction with the 5'-UTR enhances viral replication by increasing the formation of ribosome complexes to increase viral protein production. The binding of miR-122 to protein argonaute-2 (Ago2) in the RNA-induced silencing complex (RISC) also protects viral RNA from exonucleases<sup>16</sup>. Interestingly, the HCV core protein was shown to bind to Gld2 in the cytoplasm and inhibit its nucleotide addition activity. The subsequent reduction in miR-122 abundance allows HCV to maintain low levels of viral protein production to facilitate continuous viral replication and infection of host cells<sup>18</sup>. Consequently, inhibition of Gld2 by the HCV core protein decreases miR-122 stability and abundance. Low miR-122 levels, in turn, are associated with hepatic cancer, linking HCV infection to the development of hepatocellular carcinoma (HCC)<sup>18,19</sup>. Hepatitis B virus X-protein (HBx)



was also shown to reduce Gld2 protein levels and cause an increase in cationic amino acid transporter 1 (CAT-1), a target of miR-122<sup>20-22</sup>. CAT-1 is involved in the tumorigenesis of the Hepatitis B virus (HBV)<sup>20</sup>. Miravirsen, an anti-miR-122 oligonucleotide, is in Phase II trials to treat Hepatitis C and has been shown to decrease levels of miR-122 for a prolonged period of time, resulting in decreased HCV RNA levels in patients<sup>23-25</sup>. As high levels of miR-122 have been observed in colorectal liver metastasis, Miravirsen has been suggested as a potential anti-cancer drug as well<sup>26</sup>.

While it is clear that Gld2 plays a role in promoting miRNA stability<sup>3,14,16</sup>, cellular mechanisms that regulate Gld2 activity were previously unknown. In HCC cells, miR-122 is destabilized despite no observed changes in Gld2 protein levels<sup>18,19</sup>. These data suggest the existence of a clinically relevant mechanism that regulates Gld2 activity via post-translational modification. We demonstrate that Gld2 activity is indeed regulated by phosphorylation. We found that Gld2 is phosphorylated at specific serine residues in the predicted disordered N-terminal domain *in vivo*, which dramatically impact catalytic activity and substrate specificity. We found protein kinases A (PKA) and B (Akt1) site-specifically phosphorylate Gld2 at S116, which abolishes 3'-nucleotide addition activity. The data reveal tumor suppressor miRNAs as a previously unrecognized target of oncogenic protein kinases.



**Figure 3.1: Pathways regulated by Gld2 and Gld2 domain architecture.**

**A)** Known functions of Gld2. Gld2 stabilizes mature miRNA and mRNA through monoadenylation or polyadenylation of the 3'-end. Mononucleotide addition of Group II pre-miRNAs on the 3'-end by Gld2 allows recognition by Dicer to be processed to mature miRNAs. This is followed by strand selection by Argonaute (AGO) and incorporation into the RNA-induced silencing complex (RISC). The different pathways are represented by solid or dashed lines. **B)** Schematic of Gld2 showing the nucleotidyltransferase domain (NTR) and poly(A) polymerase-like domain (PAP).

## 3.2 Materials and Methods

### 3.2.1 Multiple sequence alignment

Alignments was performed as previously described in Chapter 2<sup>14</sup>. Briefly, 250 mammalian Gld2 sequences were downloaded from NCBI. Sequence alignment and alignment editing were performed with Muscle<sup>27</sup>, MultiSeq from VMD 1.8.7<sup>28</sup>, and Wasabi<sup>29</sup>.

### 3.2.2 Plasmids

*Homo sapiens* Gld2 was codon-optimized (Genewiz, South Plainfield, NJ, USA) for expression in *Escherichia coli*. The gene was cloned into pGEX-6P-2 with an N-terminal TEV cleavage site using *Bam*HI and *Xho*I restriction sites. Mutants were generated through site-directed mutagenesis<sup>30</sup>. All primers are listed in Appendix A Table A1. Successful cloning was verified by DNA sequencing at the London Regional Genomics Centre, London, ON, Canada. Cloning of Akt1 and PDK1 and Akt1 production and purification were previously described<sup>31</sup>.

### 3.2.3 Gld2 protein production and purification

Wildtype Gld2, glutamic acid mutants, and alanine mutants were transformed into *E. coli* BL21 (DE3) cells (Invitrogen, Carlsbad, CA, USA) and grown to an OD<sub>600</sub> of 0.6 at 37°C. Protein production was induced by 500 µM isopropyl β-D-1-thiogalactopyranoside (IPTG) and grown at 16°C for 18 hrs. Cells were harvested in GST wash buffer (50 mM Tris-HCl [pH 8.0], 300 mM NaCl, 5 mM MgCl<sub>2</sub>) and ethylenediaminetetraacetic acid (EDTA)-free mini protease inhibitor cocktail (Roche, Basel, Switzerland) before cell lysis with a French pressure cell press. The cell lysate was centrifuged at 64,000 x g for 1 hr at 4°C and the supernatant was loaded onto a GSTrap Fast Flow 5mL (GE Healthcare). Protein purification was automated on the ÄKTA Pure (GE Healthcare). Protein was eluted with an increasing gradient of GST elution buffer (50 mM Tris-HCl [pH 8.0], 300 mM NaCl, 5 mM MgCl<sub>2</sub>, 10 mM reduced glutathione). Eluted fractions were pooled, concentrated, and dialyzed against storage buffer (50 mM Tris-HCl [pH 8.0], 200 mM NaCl, 5 mM MgCl<sub>2</sub>, 10% glycerol) at 4°C overnight. The proteins were aliquoted and stored at -80°C until further use.

### 3.2.4 Western blotting

Purified enzyme samples were combined with 3 x sodium dodecyl sulfate (SDS) loading dye (188 mM Tris-HCl [pH 6.8], 3% SDS (w/v), 30% glycerol, 0.01% bromophenol blue, 300 mM Dithiothreitol (DTT)) and electrophoresed on two identical 10% polyacrylamide SDS gels. One gel was stained with Coomassie Brilliant Blue G-250 and imaged on the ChemiDoc MP Imaging System (BioRad). The other gel was

transferred to a polyvinylidene difluoride (PVDF) membrane using the Trans-Blot Turbo Transfer System (BioRad). The membrane was blocked in 3% bovine serum albumin (BSA), 1 x phosphate-buffered saline 1% Tween (PBS-T) for 2 hrs at room temperature (RT) and incubated with anti-Gld2 (PA5-25015, ThermoFisher Scientific) in 3% BSA, 1 x PBS-T (1:1000) overnight at 4°C. The membrane was washed 3 x 10 min in 1% BSA, 1 x PBS-T at RT, incubated with IRDye 800CW goat anti-rabbit IgG (926-32211, LI-COR) in 1% BSA, 1 x PBS-T (1:5000) for 2 hrs at RT, washed 3 x 10 min in 1 x PBS-T, and 10 min in 1 x PBS. The membrane was imaged on the Odyssey Classic (LI-COR).

### 3.2.5 Nucleotide addition assay

Gld2-catalyzed reactions and product quantification were carried out as described previously in Chapter 2. Briefly, reactions contained 1  $\mu$ M ATP (0.835  $\mu$ M unlabelled ATP and 0.165  $\mu$ M [ $\alpha$ -<sup>32</sup>P]-ATP (Perkin Elmer)), 2  $\mu$ M 5p-miR-122 (22 nt) or 15A RNA (15 nt) (SigmaAldrich), and 100 nM Gld2. Each Gld2 enzyme was incubated with 1  $\mu$ M unlabelled and [ $\alpha$ -<sup>32</sup>P]-labelled ATP and 2  $\mu$ M RNA substrate. Reactions were incubated at 37°C and samples were taken every 2 minutes and stopped with the addition of 2 x RNA loading dye. Reactions were analyzed via electrophoretic separation and subsequent phosphorimaging on a Storm 860 Molecular Imager. Product formation was quantified by spotting a range of known concentrations of [ $\alpha$ -<sup>32</sup>P]-ATP onto a strip of Whatman filter paper that was imaged on the same phosphorimaging screen as the gel. Specific activity was calculated from the linear slope of the curve using Microsoft Excel and the standard errors and deviations were obtained from triplicate reactions. SigmaPlot (Systat Software) was used to determine statistical significance (*p* values) for changes in Gld2 activity.

### 3.2.6 Fluorescence anisotropy

Dissociation constants ( $K_d$ ) of Gld2 substrates miR-122 and 15A were determined in 100  $\mu$ L reactions in black plates containing 3.2 mM MgCl<sub>2</sub>, 1 mM DTT, 20 nM 5'-end labelled miR-122 (22 nt) or 15A RNA (15 nt), and 0-100 nM Gld2 enzyme incubated for 20 minutes at room temperature in the dark. The RNAs were labelled on the 5'-end with 6-carboxyfluorescein (6-FAM) (SigmaAldrich) and the following enzyme concentrations

were used (nM): 0, 0.5, 1, 2.5, 5, 7.5, 10, 12.5, 15, 17.5, 20, 30, 40, 50, 60, 80, 100. Fluorescence polarization was measured on a Victor<sup>3</sup>V (PerkinElmer) with an excitation of 492 nm and emission of 535/20 nm. Readings were subtracted from the no enzyme control and three technical replicates were performed for each RNA and enzyme. SigmaPlot (Systat Software) was used to generate the plots, determine the  $K_d$ , and calculate the standard errors and  $p$  values.

### 3.2.7 Identification of potential kinases

Two online tools were used to generate a list of potential Gld2 kinases. PhosphoMotif Finder identifies putative kinase binding sequences in a query sequence based on the binding motifs of kinases as well as their substrate sequences identified in the literature<sup>32</sup>. GPS 3.0 predicts kinase phosphorylation sites in the query sequence using a computational prediction program<sup>33</sup>.

### 3.2.8 Dot plot kinase activity assays

Kinase assays were performed as previously described<sup>31</sup>. Briefly, reactions containing a kinase, wildtype Gld2, and [ $\gamma$ -<sup>32</sup>P]-ATP (Perkin Elmer) were carried out at 37°C. Samples were taken at various timepoints and stopped by spotting on P81 paper. The P81 paper was washed, air-dried, exposed to a phosphor screen, and visualized with a phosphorimager (Storm 860 Molecular Imager).

### 3.2.9 Kinase activity assays using SDS gels

In the following assays, Gld2 was tested as a protein substrate for several human kinases. Kinase assays were performed in 60  $\mu$ L reactions containing 900 nM Gld2, kinase buffer (20 mM MOPS [pH 7.0], 25 mM  $\beta$ -glyceraldehydephosphate, 25 mM MgCl<sub>2</sub>, 5 mM EGTA [pH 8.0], 1 mM Na<sub>2</sub>VO<sub>4</sub>, 0.1 mM ATP, 13.2 nM [ $\gamma$ -<sup>32</sup>P]-ATP (Perkin Elmer)). Reactions were initiated with the addition of the specified kinase. Since kinases varied in activity, final concentrations were adjusted according to published values using 25 nM CK2 $\alpha$ , 1.43 nM CK2 holoenzyme, 33 nM fully activated ppAkt1<sup>T308,S473</sup>, 30 nM Abl, 0.45 nM PKA, or 12 nM CDK5 in a kinase dilution buffer (0.1 mg/mL BSA, 5 mM MOPS [pH 7.2], 25 mM  $\beta$ -glyceraldehydephosphate, 5 mM

MgCl<sub>2</sub>, 1 mM EGTA [pH 8.0], 1 mM Na<sub>2</sub>VO<sub>4</sub>, 100 mM NaCl). Reactions were incubated at 37 °C for 15 minutes on a microcentrifuge shaker. Samples (20 µL) were taken every 5 minutes and the reaction was stopped with the addition of 2 x SDS loading dye. Purified recombinant kinases CK2α, CK2 holoenzyme, Abl, PKA, and CDK5 were a generous gift from Dr. David W. Litchfield (The University of Western Ontario, Canada). Reaction products at each time point were separated on a 10% polyacrylamide SDS gel. The gel was exposed to a storage phosphor screen overnight at -80°C and visualized with a phosphorimager (Storm 860 Molecular Imager). Kinase assays and quantification were previously described<sup>31,34</sup>.

### 3.2.10 Isolation of phosphorylated Gld2 for downstream assays

For mass spectrometry and downstream Gld2 activity assays, large-scale kinase reactions were performed as above with 0.1 mM unlabelled ATP and 80 nM PKA for 15 minutes. To isolate the resulting phosphorylated Gld2 (pGld2), the reactions were loaded onto GST SpinTrap columns (GE Healthcare) and pGld2 was eluted with GST elution buffer. The isolated pGld2 was used immediately for downstream assays. The large-scale kinase reaction was repeated with a control lacking the kinase.

### 3.2.11 Phosphorylation of Gld2 using HEK 293 cell extract

HEK 293 cells in 100 mm plates were grown to approximately 90% confluency in Dulbecco's Modified Eagle's Medium (DMEM) (319-005-3L, Wisent) with 10% fetal bovine serum (FBS) (098150, Lot #185700, Wisent) and 1% penicillin-streptomycin (450-201-EL, Wisent). Epidermal growth factor (EGF) was added to each plate to a final concentration of 50 ng/mL and incubated at 37°C for 1 hr. Cells were harvested and resuspended in 5 x kinase buffer with 5 mM phenylmethylsulfonyl fluoride (PMSF). Cells were broken with a Q125 Sonicator (Qsonica) six times at 20% amplitude and 1 sec on, 1 sec off. Using the cell extract as a source of active kinases, we performed a kinase reaction. This was repeated with unstimulated HEK 293 cell extract. Gld2 was isolated using a GST Spintrap column and possible Gld2 phosphorylation sites were analyzed by mass spectrometry.

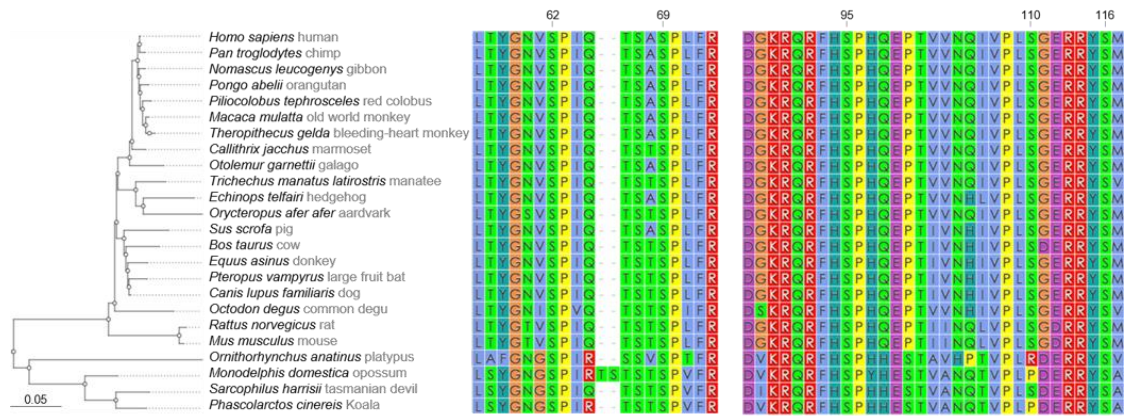
### 3.2.12 Mass spectrometry

Mass spectrometry (MS) analysis of tryptic digested peptides was carried out on the EasyLC1000-QExactive tandem LC-MS system (ThermoFisher Scientific). A full scanning (full MS/dd-MS2 TopN, data dependent acquisition mode in a Q-Exactive) was performed to obtain an overview of all possible protein modifications within Gld2. Parallel-reaction monitoring (PRM) was then carried out to further verify the phosphorylation at S62 or S116 in Gld2. We analyzed Gld2 and Gld2 phosphorylated by purified recombinant kinases and by HEK 293 cell extract. Gld2 or pGld2 was precipitated in ice-cold acetone/ethanol/acetic acid (50/50/0.1 v/v/v). The protein precipitate was re-suspended in 8 M urea then reduced in 5 mM DTT at 37°C for 1 hr and alkylated in 14 mM iodoacetamide in darkness at RT for 1 hr. Unreacted iodoacetamide was neutralized by adding 5 mM DTT and final protein concentration was determined by Bradford assay. Trypsin digestion was performed at 37°C overnight with a protein:trypsin ratio of 20:1 w/w. The digest was desalted in a C18 column (Phenomenex, Torrance, CA, USA) according to the manufacturer's protocol and re-suspended in MS-grade water for MS injection. Data were analyzed using Skyline software<sup>35</sup>.

## 3.3 Results

Gld2 plays an important role in miRNA stability, but the regulation of Gld2 activity or substrate specificity is unknown. Studies in other nucleotide polymerases<sup>36,37</sup> found that phosphorylation of serine and threonine residues can increase activity or processivity. For example, serine phosphorylation in the disordered C-terminal domain of RNA polymerase II is required for transcription initiation and elongation<sup>36</sup>. Phosphorylation of the terminal uridylyltransferase TUT1 at S6 is required for TUT1 nuclear retention and regulation of specific mRNAs<sup>37</sup>. Multiple independent proteome-level mass spectrometry studies of human cells revealed phosphorylated residues in Gld2<sup>38-40</sup>, including five conserved serine residues (S62, S69, S95, S110, S116) in the predicted disordered N-terminal domain (Figure 3.2). Despite these observations, the

putative Gld2 kinase(s) and the impact of phosphorylation on Gld2 activity was unknown.



**Figure 3.2: Multiple sequence alignment of mammalian Gld2 sequences.**

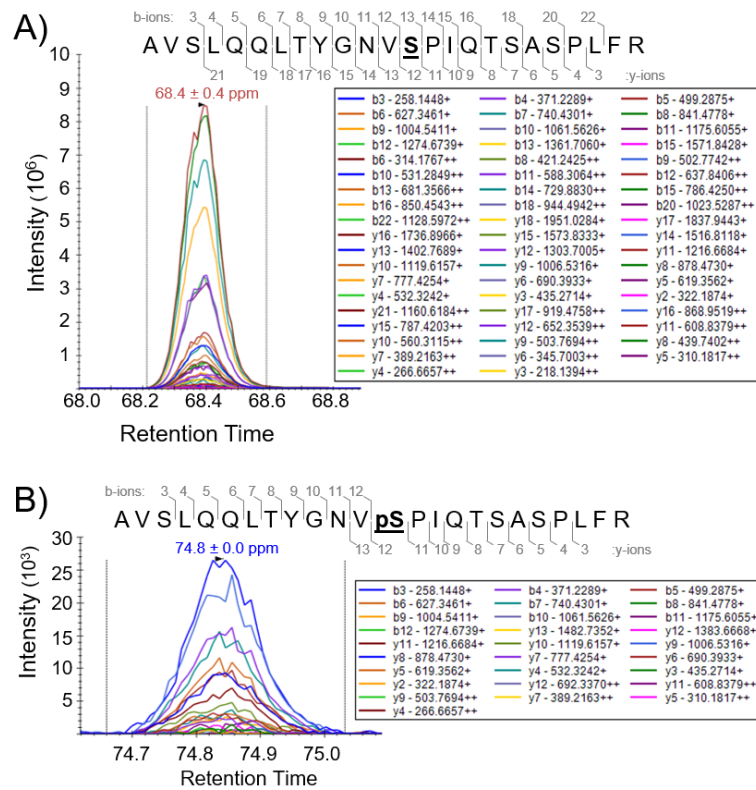
Sequences were downloaded from NCBI and the alignment and editing were performed with Muscle<sup>27</sup>, MultiSeq from VMD 1.8.7<sup>28</sup>, and Wasabi<sup>29</sup>. Numbers above the alignment indicate the position in *H. sapiens* Gld2.

### 3.3.1 Phosphorylation of the Gld2 N-terminal domain by HEK 293 cells

Although the above studies suggest the existence of a Gld2 kinase in human cells, we analyzed Gld2 following incubation with HEK 293 cell lysates to confirm phosphorylation activity towards Gld2. The HEK 293 cells were stimulated with epidermal growth factor (EGF) to activate cellular kinases. Purified Gld2 (Appendix A Figure A1) was incubated with cell extracts from HEK 293 cells after EGF stimulation of signaling pathways. The phosphorylation status of Gld2 was subsequently analyzed by mass spectrometry and we unambiguously identified phosphorylation at S62 in the sample incubated with EGF-stimulated cell lysate (Figure 3.3B). We did not identify pS62 in unstimulated cells or in recombinantly produced Gld2. As Gld2 has been previously shown to be involved in miRNA metabolism<sup>4,7</sup>, this indicates the existence of physiologically relevant signaling pathways connecting EGF-stimulated protein kinases to miRNA metabolism via phosphorylation-dependent regulation of Gld2. EGF activates many cellular pathways involved in regulating growth, proliferation, differentiation, and survival. EGF binds to receptor tyrosine kinases, leading to their activation. This in turn



activates cascades of cellular kinases. Within 10 minutes of EGF stimulation, mammalian cells display 100s of new phosphorylation events<sup>41</sup>. Data from this study also indicate that following EGF stimulation 30 tyrosine and more than 100 serine/threonine kinases are activated by phosphorylation, which may be responsible for S62 phosphorylation. Any of these serine/threonine kinases are potential candidates for catalyzing S62 phosphorylation.



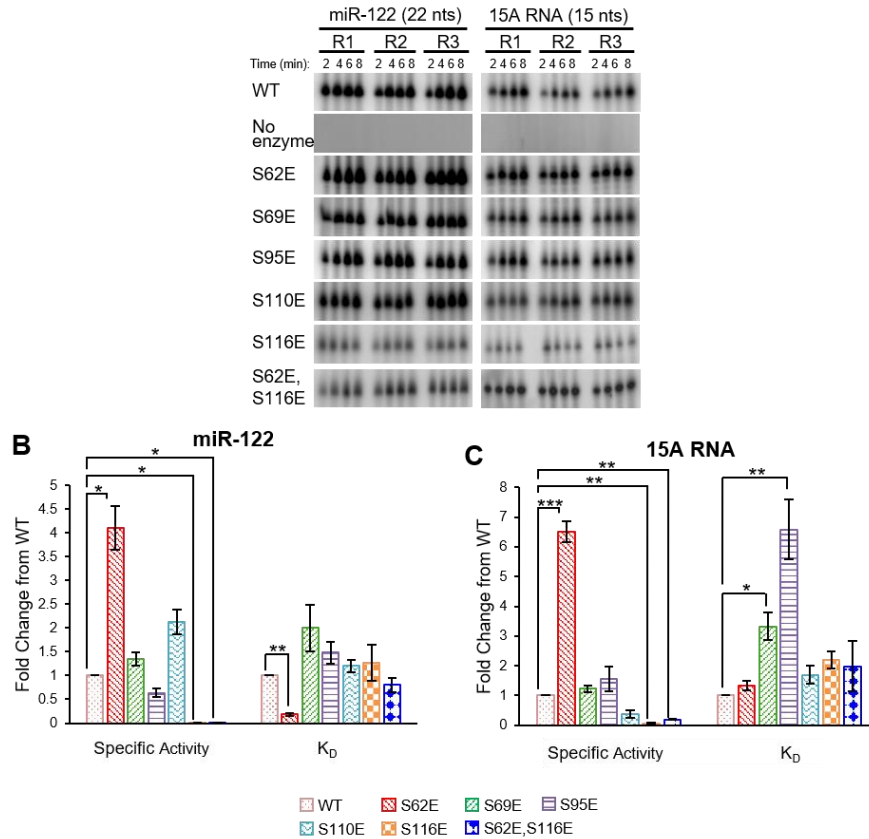
**Figure 3.3: Gld2 is phosphorylated at S62 when incubated with EGF-stimulated HEK 293 cell extract.**

Mass spectra of Gld2 after incubation with cell extracts from EGF stimulated HEK 293 cells showing **A)** unphosphorylated S62 and **B)** phosphorylation at S62 (bolded and underlined). The intensity for y and b-ions resulting from fragmentation of the peptide containing S62 is shown; these intensities are overlaid on the retention time position of the full peptide mass. M/z values for each y and b ion are shown. Gld2 was isolated using a GST Spintrap column before mass spectrometry. Trypsin was used to generate the Gld2 peptides and the ions from the peptides are shown in the mass spectra. A full scanning was performed to obtain all possible modifications and was followed by parallel reaction monitoring (PRM) to verify the modification at S62.

### 3.3.2 Gld2 N-terminal domain phosphomimetic variants regulate catalytic activity

To rapidly assess the effects of phosphorylation at positions in the Gld2 N-terminal domain, human Gld2 variants were produced in *Escherichia coli* with respective serine phosphorylation sites mutated to the phosphomimetic glutamic acid<sup>30</sup>. Wild type and phosphomimetic Gld2 variants were produced and purified to homogeneity (Appendix A Figure A1). Kinetic parameters for specific activity and binding affinity were determined for wildtype Gld2 and phosphomimetic variants with glutamate substitutions at the phosphorylation sites S62, S69, S95, S110, and S116

Nucleotide addition activity was measured by incubating each enzyme variant with an RNA substrate and [ $\alpha$ -<sup>32</sup>P]-ATP (Figures 3.4, and A2A, A2B). Enzymatic and binding assays were conducted with RNA substrates miR-122 and a mRNA poly(A) tail mimic of 15 adenine residues (15A). MiR-122 and 15A RNA were used based on previous studies demonstrating their competence as Gld2 substrates<sup>4,14</sup>.



**Figure 3.4: Phosphomimetic Gld2 variants modulate catalytic activity and RNA binding.**

**A)** Activity assay gels of wildtype and mutant Gld2. Wildtype (WT) Gld2 and glutamic acid mutants were incubated with  $[\alpha\text{-}^{32}\text{P}]\text{-ATP}$  and miR-122 (22 nts) or 15A RNA (15 nts) at 37°C with samples taken every 2 minutes for 8 minutes. Reactions were repeated in triplicate (R1-3) and analyzed via gel electrophoresis and phosphorimaging. A no enzyme control was performed for each RNA substrate in triplicate and the average of the no enzyme triplicates for each RNA was calculated for the 0-minute timepoint. Reaction products were quantified by exposing a Whatman filter strip dotted with different known concentrations of  $[\alpha\text{-}^{32}\text{P}]\text{-ATP}$  to the same phosphorscreen as the gel. R, replicate. **B)** and **C)** Bar graphs showing the fold change in specific activity at 1  $\mu\text{M}$  ATP calculated from the activity assays and binding affinity ( $K_d$ ) between wildtype Gld2 and Gld2 glutamic acid mutants with **B)** miR-122 (22 nts) or **C)** oligo(A) tail mimic 15A RNA (15 nts). Specific activity is the activity of an enzyme per milligram of purified enzyme and was calculated from the linear slope of the curve. Fluorescence anisotropy was used to determine the  $K_d$ . Each Gld2 enzyme was incubated with a RNA substrate fluorescently labelled on the 5'-end with 6-FAM and incubated at room temperature for 20 minutes. Fluorescence polarization was measured at Ex. 492nm and Em. 535/20 nm and the  $K_d$  was calculated using SigmaPlot. Error bars represent one standard error calculated from triplicate reactions. Significant changes calculated using a two-tailed t-test are indicated by asterisks.  $p \leq 0.05$  (\*);  $p \leq 0.01$  (\*\*);  $p \leq 0.001$  (\*\*\*). Fold changes were calculated using data from Table 3.1 and Appendix A Figure A2.

Depending on the residue location, phosphomimetic substitutions had distinct effects on enzyme activity (Table 3.1, Figure 3.4). S62E markedly increased activity with both RNAs compared to wildtype Gld2, indicating an overall activating effect. In contrast, a S116E mutation severely decreased Gld2 activity with miR-122 and 15A RNA. Interestingly, S110E increased activity for miR-122 but decreased activity for 15A RNA. Gld2 S69E showed no significant changes in activity for either RNA while S95E was 1.6-fold more active with 15A RNA.

Based on a comparison of the specific activities (Figure 3.4, Table 3.1), only S62E and S116E displayed statistically significant changes in activity with both RNA substrates. S62E enhanced the nucleotide addition activity by ~5-fold. S116E exhibited the opposite effect, decreasing Gld2 activity by 111-fold with miR-122 and 16-fold with 15A RNA. As the S62E and S116E mutants displayed opposite effects on Gld2 activity, a double mutant (S62E/S116E) was generated to investigate cumulative effects. Interestingly, the inhibitory effect of the S116E mutation overpowered the activating effect of the S62E mutation and decreased the activity of Gld2 74-fold with miR-122 and 5-fold with 15A RNA compared to the wildtype enzyme. The double mutant counteracted the silencing effect of S116E alone by 3.1-fold with 15A RNA and 1.5-fold with miR-122.

The ability of the mutants to alter the nucleotide addition activity varied between RNA substrates. The molecular basis for the higher specific activity of Gld2 with miR-122 compared to 15A RNA remains to be elucidated to discern whether Gld2 recognizes a specific RNA sequence and/or discriminates substrates based on the RNA length.

### 3.3.3 Gld2 phosphomimetic substitutions impact RNA substrate affinity

As changes in catalytic activity were RNA-dependent, the RNA binding affinities of all Gld2 phosphomimetic variants were quantified using fluorescence anisotropy (Figures A2C, A2D). The binding affinities ( $K_d$ ) for all enzyme variants were in the nanomolar range (Figure 3.4, Table 3.1). Changes in RNA binding affinity were substrate dependent.

The binding affinity to miR-122 was unchanged for most mutants, except for S62E, which showed a 5.5-fold increase in RNA binding compared to wildtype Gld2 with miR-122. The same mutant showed no change in binding to a 15A substrate. For the 15A RNA, two mutants (S69E and S95E) showed a decrease in affinity, while all other mutants showed no change. With a 6.6-fold reduced  $K_d$  compared to wildtype, Gld2 S95E showed the most dramatic impact on 15A RNA binding; S69E was 3.3-fold decreased in binding affinity to 15A.

Overall, our phosphomimetic analysis suggests that each phosphorylation site has a distinct role in regulating Gld2 activity or substrate selectivity. We found that S62E increases activity with either no change (15A) or with increased binding affinity (miR-122), which may favor miRNA stabilization over mRNA. S69E and S95E caused no significant change in activity, but decreased affinity towards 15A RNA, indicating a reduced preference for mRNA adenylation. S110E appears to have an insignificant impact on activity and binding on the tested substrates. Finally, S116E and the S62E/S116E double mutant markedly reduced activity without significantly impacting RNA binding. The data indicate that Gld2 S116E and the double mutant are able to bind to the RNA target, but perhaps not in a catalytically competent conformation.

**Table 3.1: Activity and RNA-binding of wildtype (WT) and phosphomimetic Gld2 variants.**

	miR-122		15A RNA	
	Specific activity ( $\mu\text{mol}/\text{min}/\text{mg}$ )	$K_d$ (nM)	Specific activity ( $\mu\text{mol}/\text{min}/\text{mg}$ )	$K_d$ (nM)
<b>WT</b>	$3452 \pm 182$	$15 \pm 3$	$448 \pm 10$	$1.7 \pm 0.3$
<b>S62E</b>	$14186 \pm 774$	$2.7 \pm 0.8$	$2919 \pm 239$	$2.3 \pm 0.3$
<b>S69E</b>	$4670 \pm 272$	$30 \pm 9$	$555 \pm 22$	$5.8 \pm 1.1$
<b>S95E</b>	$2191 \pm 78$	$22 \pm 6$	$698 \pm 30$	$11 \pm 2$
<b>S110E</b>	$7340 \pm 433$	$18 \pm 5$	$171 \pm 7$	$2.9 \pm 0.6$
<b>S116E</b>	$31 \pm 5$	$19 \pm 5$	$28 \pm 0.05$	$3.8 \pm 0.8$
<b>S62E/S116E</b>	$47 \pm 2$	$12 \pm 3$	$89 \pm 6$	$3.4 \pm 0.6$
<b>S116A</b>	$150 \pm 0.05$	$35 \pm 7$	$38 \pm 16$	$7.4 \pm 1.8$

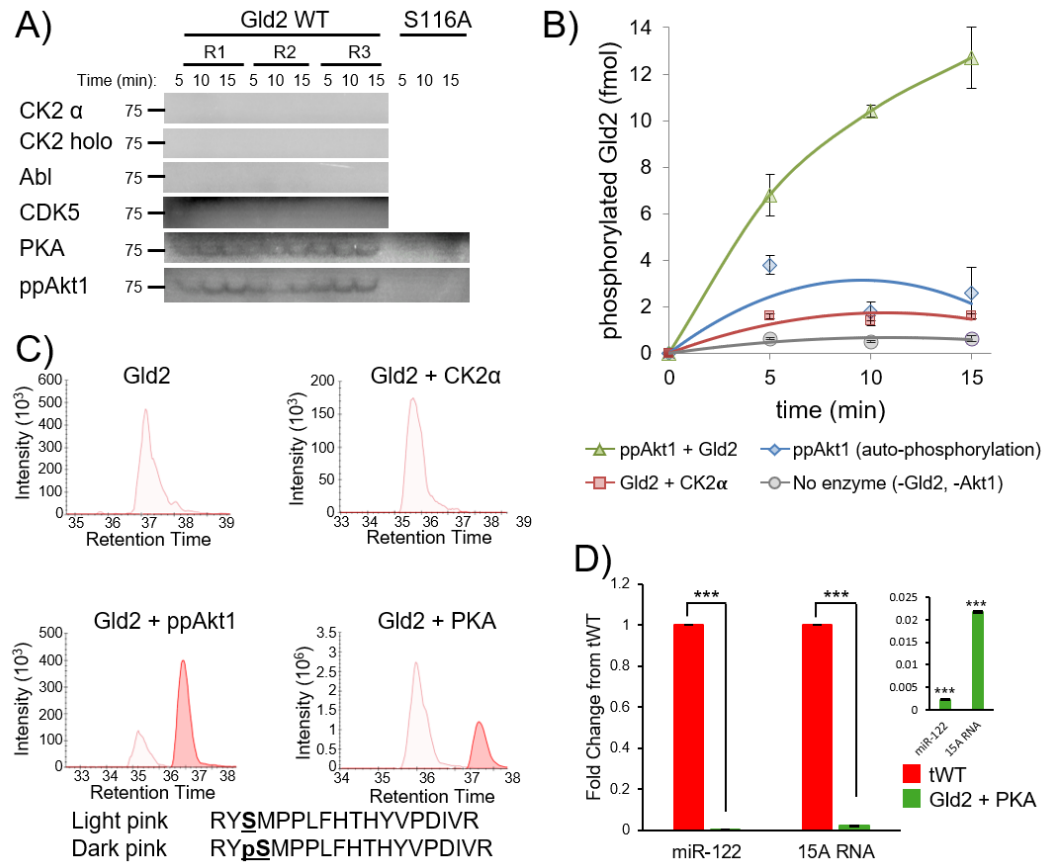
Standard error is reported; specific activities at 1  $\mu\text{M}$  ATP.

### 3.3.4 PKA and Akt1 site-specifically phosphorylate Gld2 at S116

As the phosphomimetic mutants displayed significant changes in activity and RNA substrate binding compared to wildtype Gld2, we next identified kinases that phosphorylate the Gld2 N-terminus. PhosphoMotif Finder<sup>42</sup> and GPS 3.0<sup>33</sup> were used to generate a list of potential kinases (Table A2). Using the kinase assay detailed in Materials and Methods, wildtype Gld2 was incubated with recombinant and active human kinases (CK2 $\alpha$ , CK2 holoenzyme, CDK5, PKA, and Akt1) predicted to have a recognition motif in Gld2 (Figure 3.5A). The kinase Abl, which was not identified as a potential Gld2 kinase, was used as a negative control. We used our recently developed approach<sup>31</sup> combining genetic code expansion with *in vivo* enzymatic phosphorylation to prepare fully activated and purified recombinant Akt1 with programmed phosphorylation at both activating sites (ppAkt1<sup>T308,S473</sup>). This method involves protein production in an *E. coli* strain that co-expresses the kinase PDK1 to phosphorylate Akt1 at T308. The strain

also genetically encodes phosphoserine (pSer) at UAG codons. The serine codon at position 473 was replaced with a UAG codon to direct pSer incorporation in Akt1.

Following incubation of Gld2 and [ $\gamma$ - $^{32}$ P]-ATP with each kinase in separate reactions, the radio-labelled phosphorylated Gld2 (pGld2) product was only observed when Gld2 was incubated with PKA or ppAkt1<sup>T308,S473</sup> (Figure A3, Figure 3.5A). Quantification of the Akt1-dependent reaction showed a rapid increase in phosphorylated Gld2 over a 15 min time course (Figure 3.5B). Independent pGld2 preparations resulting from incubation with PKA or ppAkt1<sup>T308,S473</sup> were analyzed by mass spectrometry to determine the site(s) of phosphorylation. Both unphosphorylated Gld2 and Gld2 incubated with CK2 $\alpha$ , which was inactive in phosphorylating Gld2, were also analyzed by mass spectrometry as controls (Figure 3.5C). S116 was unambiguously identified as the site of specific phosphorylation by Akt1 and PKA. Phosphorylation at S116 was not observed in unphosphorylated Gld2 or in the preparation incubated with CK2 $\alpha$ . Next, S116 was mutated to an alanine residue (Figure A1) to determine if S116 represents the sole Akt1 and PKA phosphorylation site on Gld2. The kinase assay was repeated with Gld2 S116A incubated with PKA or ppAkt1<sup>T308,S473</sup> (Figure 3.5A). For both kinases, no phosphorylated Gld2 S116A product was observed, confirming that PKA and Akt1 phosphorylate Gld2 specifically and exclusively at S116.



**Figure 3.5: Akt1 and PKA phosphorylate Gld2 at S116.**

**A)** Gld2 or Gld2 S116A were incubated with [ $\gamma$ - $^{32}$ P]-ATP and the indicated kinases. Formation of phosphorylated Gld2 (75 kDa) was monitored by electrophoretic separation and subsequent phosphorimaging. R, replicate. **B)** Quantification of phosphorylated product formation from a kinase reaction over 15 minutes. **C)** Mass spectra of unphosphorylated Gld2 or Gld2 phosphorylated by CK2 $\alpha$ , ppAkt1<sup>T308,S473</sup> (ppAkt1), or PKA. Unphosphorylated peptide is indicated by the light pink peak and the phosphorylated peptide by the dark pink peak. Position 116 is bolded and underlined. **D)** Bar graphs showing the fold change in specific activity at 1  $\mu$ M ATP between treated WT (tWT) and Gld2 phosphorylated by PKA with miR-122 (22nt) or oligo(A) tail mimic 15A RNA (15nt). The inset shows the fold change in activity of Gld2 phosphorylated by PKA compared to tWT. Error bars show one standard error calculated from triplicate reactions. Significant changes calculated using a two-tailed t-test are indicated by asterisks. p  $\leq$  0.001 (\*\*\*).



### 3.3.5 Phosphorylation of Gld2 at S116 abolishes nucleotide addition activity

The phosphomimetic mutant (S116E) was not competent in nucleotide addition, yet the variant retained RNA binding affinity (Figure 3.4, Table 3.1). As Akt1 and PKA both phosphorylate S116, we produced pGld2<sup>S116</sup> following incubation with PKA as noted above to investigate the nucleotide addition activity of Gld2 with phosphate on S116. We performed nucleotide addition activity assays with pGld2<sup>S116</sup> and with an unphosphorylated Gld2 control without PKA addition (treated wildtype, tWT).

PKA-dependent phosphorylation of Gld2 decreased nucleotide addition activity by two orders of magnitude (Figures 3.5D, A4). Significant reductions in activity were observed for both 15A RNA (~45-fold) and miR-122 (~400-fold) substrates. Although both pGld2<sup>S116</sup> and the S116E phosphomimetic variant reduced nucleotide addition activity, as we anticipated, the phosphate at S116 had a significantly stronger inhibitory effect compared to acidic amino acid substitutions (~3-fold). This observation is even more striking in light of the fact that our pGld2<sup>S116</sup> preparations are only partially phosphorylated (Figure 3.5C), suggesting phosphorylation of the Gld2 N-terminal domain is a potent mechanism for the cell to control nucleotide addition activity.

Nucleotide addition activity assays and binding assays were also performed for the S116A mutant (Table 3.1, Figure A5). Alanine substitutions are often used as phospho-ablated enzyme models. Although the mutant was expected to act similarly to the wildtype enzyme, the alanine substitution in fact reduced the activity by 23-fold with miR-122. As described above, the S116E mutation reduced activity by 111-fold with miR-122, indicating that a serine residue is crucial at this position and cannot be replaced by alanine. Thus, S116A is not an appropriate model for an unphosphorylated Gld2 and indicates that a serine in position S116 is required for activity. We previously showed that alanine is not necessarily a good model for a non-phosphorylatable residue<sup>31</sup>. Overall, these data indicate that ppAkt1<sup>T308,S473</sup> dependent phosphorylation of Gld2 inactivates Gld2. Our data reveals a novel molecular pathway linking Akt1 activity to miR-122 stability and activity *in vivo*.

## 3.4 Discussion

### 3.4.1 Gld2 activity is regulated by phosphorylation

Gld2 is a key regulator in the stabilization and maturation of tumor suppressors miR-122 and let-7<sup>4,7,11</sup>. Cellular mechanisms that regulate Gld2 mediated nucleotide addition were previously unknown. Data from HCC cells<sup>18</sup> and proteomic analysis<sup>38,40</sup> implicated post-translational modification as a potential mechanism regulating Gld2 activity. Here, we presented the first evidence that serine phosphorylation of Gld2 has a profound impact on catalytic activity and RNA binding.

We identified phosphorylation sites in the N-terminal domain of Gld2 that positively or negatively regulate nucleotide addition activity. Our findings are reminiscent of regulation identified in other polymerases. Phosphorylation of specific serine residues in the C-terminal domain of RNA polymerase II and in the terminal uridylyltransferase TUT1 regulate their activity and substrate recognition<sup>36,37</sup>. For RNA polymerase II, serine phosphorylation in the disordered C-terminal domain is required for promoter clearance<sup>36</sup>. Phosphorylation of the uridylyltransferase TUT1 at position S6 plays a role in its regulation of specific mRNAs and in its nuclear retention, possibly by facilitating interactions between TUT1 and nuclear proteins<sup>37</sup>. Similarly, we found that both phosphomimetic mutations to acidic residues or true phosphorylation of Gld2 at different sites in the N-terminal domain substantially altered substrate specificity, enhanced, or abolished enzyme activity. Although these phosphorylation sites are conserved among mammalian Gld2 proteins, they are not conserved in the human terminal uridylyltransferases TUT4 and TUT7, as these enzymes lack the large disordered N-terminal region.

Undoubtedly, human cells possess a robust Gld2 phosphorylation activity (Figure 3.3). We found that lysates from EGF-stimulated HEK 293 cells were active in phosphorylating Gld2 at S62. In this particular experiment, we were not able to identify additional phosphorylation sites in Gld2, suggesting that additional Gld2 kinases may not be activated or sufficiently active in our experimental conditions. While Akt1 and PKA are expressed in HEK 293 cells, Akt1 activity in these cells even upon EGF stimulation is

low<sup>43</sup> and likely not sufficient to yield quantitative phosphorylation of Gld2 S116 required for mass spectrometry. Alternatively, Gld2 phosphatases may be active at sites other than Ser62. The data, nevertheless, show that human cells are competent in phosphorylation of Gld2 in its N-terminal domain.

Our experiments with phosphomimetic mutants indicate that phosphorylation at S62 significantly increases Gld2 activity with miR-122 and the 15A RNA. Although Gld2 S62E showed a significant increase in activity with both RNAs, increased binding affinity was only observed with miR-122. It is interesting to note that wildtype Gld2 is 9-fold more active with miR-122 than with the poly(A) tail mimic. Even the increase in activity with the poly(A) tail mimic by S62E does not reach the level of wildtype activity with miR-122. In contrast, experiments with Gld2 S116E and pGld2<sup>S116</sup> show that phosphorylation at this site abolishes nucleotide addition activity. Although enzyme activity is more than 100-fold reduced with miR-122, Gld2 S116E and pGld2<sup>S116</sup> still retained very low levels of enzymatic activity. In comparison, a mutation in the active site, D215A, completely abolished enzyme activity on miRNA substrates<sup>4,18</sup>. D215A is part of the conserved catalytic triad responsible for activity<sup>3,44,45</sup>, while S116 is found in the predicted disordered N-terminal domain, which was previously shown to be dispensable for catalytic activity in the related uridylyltransferase Cid1<sup>46-48</sup>, but we here show that its function lies in the regulation of enzyme activity. Using enzymatic assays and mass spectrometry, we identified and validated Akt1 and PKA as kinases with site-specific phosphorylation activity at S116 in Gld2. Although acidic amino acids are not always able to mimic the functional impact of phosphate<sup>31</sup>, the glutamate variant showed reduced activity similarly to the pGld2<sup>S116</sup> enzyme. Phosphorylation at S116, however, led to a significantly greater reduction in Gld2 activity, which was two orders of magnitude below the activity of the unphosphorylated enzyme. Thus, phosphorylation at S116 effectively controls Gld2 activity.

During HCV infection, the Hepatitis C core protein binds to Gld2 to inhibit its adenylation activity. Core protein binding is somewhat inefficient, with 13% binding at a 1:1 ratio. Nonetheless, this inhibition leads to a reduction of cellular miR-122 levels by 30%<sup>18</sup>. Even partial phosphorylation of cellular Gld2 at S116 is expected to have a

similar or greater effect on miR-122 levels, making post-translational modification of Gld2 an efficient means to control cellular miRNA levels.

Although phosphorylation of S116 inhibits Gld2 catalytic activity, RNA binding affinity was unperturbed, suggesting Gld2 phosphorylated at S116 binds the RNA substrate in a non-productive conformation. The crystal structures of related nucleotidyltransferases from yeast<sup>49</sup> and *C. elegans*<sup>15,50</sup> as well as the human terminal uridylyltransferases TUT1 and TUT7<sup>51,52</sup> revealed a conserved positively charged surface that may facilitate RNA binding. Despite these efforts, no structural information is available on the N-terminal domain of Gld2. Our data suggest that Gld2 can assume different RNA binding modes. The wildtype Gld2 enzyme binds RNA with high affinity in a catalytically competent mode, which is perhaps stabilized yet further in the activating mutants S62E and S110E. Conversely, Gld2 variants with phosphomimetic substitution or phosphorylation at S116 appear to bind the RNA substrate in a non-catalytic mode that interferes with nucleotide addition. This is not unique to Gld2, as other cases of catalytically incompetent enzymes have been described in the literature. A small deletion in RNA polymerase (RNAP) leads to a catalytically incompetent RNAP, that remains bound to the promoter complex<sup>53</sup>. It is also well known that phosphorylation of the C-terminal domain of RNAP II is required for promoter clearance but not its activity<sup>36,54</sup>. While the structural basis for phosphorylation-dependent modulation of Gld2 activity is not yet defined, our experiments suggest an allosteric mechanism. We identified Gld2 variants that impact activity independently of RNA binding, indicating that allosteric or conformational changes in the Gld2 RNA complex may play an important role in Gld2 catalyzed nucleotide addition.

Furthermore, in the cell, phosphorylation of Gld2 may affect interactions with proteins that are implicated in RNA substrate selectivity<sup>12,55–58</sup>. In humans, Gld2 was recently shown to interact with the RNA-binding protein Quaking (QKI-7) to facilitate polyadenylation of target mRNAs. QKI-7 was shown to bind between residues 1-141 on Gld2<sup>58</sup>, which correspond to the predicted disordered N-terminal domain, including all of the phosphorylation sites investigated here. Future efforts will determine the effect of

phosphorylation on the ability of RNA binding proteins to interact with and regulate Gld2 activity and substrate specificity.

### 3.4.2 Oncogenic protein kinases signal to miRNA regulation

We identified Gld2 as a previously unknown substrate of two oncogenic kinases. PKA and Akt1 belong to the evolutionarily conserved AGC family of protein kinases that are activated upon stimulation with growth factors such as EGF<sup>59</sup>. PKA is the key kinase in the cyclic AMP signaling pathway that is activated upon hormone binding to G-protein coupled receptors (GPCR). PKA activation has been shown to modulate the expression of miR-122 through the extracellular signal-regulated kinase (ERK)<sup>60</sup> and miRNA let-7b levels through protein phosphatase 2A (PP2A) activation but the underlying pathway remained unclear<sup>61</sup>.

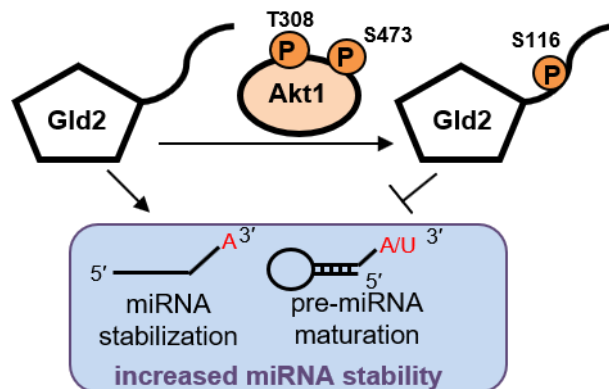
Akt1 is a central hub of the PI3K/Akt signaling pathway, which is the most commonly activated signal transduction pathway in human cancers<sup>62</sup>. Active Akt1 signals for cell survival and proliferation while also inhibiting apoptosis<sup>59</sup>. Akt1 activity is dependent on phosphorylation at two key regulatory sites (T308, S473). Over-active and hyper-phosphorylated Akt1 is a hallmark of diverse human malignancies<sup>62,63</sup>, while the unmodified Akt1 protein is inactive and rapidly degraded in cells<sup>64</sup>. Several reports show that Akt1 expression is regulated by miRNAs. The miRNAs miR-564 and miR-215 directly negatively regulate Akt1 mRNA stability<sup>65,66</sup>, while miRNA let-7 inhibits cyclin D1 expression, leading to reduced Akt1 phosphorylation at S473<sup>67</sup>. Conversely, miRNA-122 overexpression is associated with increased Akt1 phosphorylation in T-cell lymphoma and renal cell carcinoma cells<sup>68</sup>, enhancing cancer progression. These reports correlate PKA or Akt1 activity with miRNA levels yet fail to identify the factor linking PKA or Akt1 activity to miRNA levels, exposing gaps in our understanding of the signal transduction network. In addition, neither PKA nor Akt1 were previously shown to directly phosphorylate any regulator of RNA.

Although miRNAs are known to regulate Akt1, the ability of Akt1 activity to regulate an enzyme involved in the regulation of miRNA metabolism has not been documented previously. The Akt isozyme Akt2 is known to phosphorylate the single-stranded RNA binding protein KH-type splicing regulatory protein (KSRP), allowing it to switch its RNA preference from mRNA to primary miRNA (pri-miRNA) and facilitate Drosha processing to pre-miRNA<sup>69</sup>. Although KSRP binds RNA, no enzyme that directly regulates RNA stability has been previously identified as a substrate of Akt1. Using precise biochemical experiments, we found that both Akt1 and PKA site-specifically phosphorylate Gld2 at S116, abolishing Gld2 activity with miRNA and mRNA substrates.

### 3.4.3 Relevance of Akt1-dependent regulation of miRNAs to disease

Hyperactivity of Akt1 and PKA is common in many cancers<sup>70,71</sup>. Our data suggest that phosphorylation of Gld2 by these kinases would further promote carcinogenesis by destabilizing tumor suppressor miRNAs, thus, further inducing tumorigenesis. Abolishing Gld2 activity leads to a decrease in levels of tumor suppressor miRNAs including miR-122 and let-7<sup>7,16</sup> (Figure 3.6). Decreased miR-122 and let-7 levels and activity enable over- or un-regulated expression of their target genes, including oncogenes with roles in cell growth, metastasis, and apoptosis<sup>16,72</sup>.

In a related disease context, Gld2 is down-regulated in Hepatitis B and inhibited in Hepatitis C infections<sup>16,18</sup>. Both HBV and HCV are contributing factors in the development of HCC and other liver diseases due to the dysregulation in miR-122 levels and the resulting expression of miR-122 regulated oncogenes<sup>16,19</sup>. The extent of Gld2 catalyzed mRNA adenylation, and the effect of specific phosphorylations on the transcriptome and miRnome remain to be investigated and may reveal additional contributions of Gld2 regulation to pathogenesis.



**Figure 3.6: Model of Akt1-mediated regulation of Gld2.**

Gld2 monoadenylates miRNAs to increase stability and adds a single nucleotide to pre-miRNAs to enable recognition by Dicer and miRNA maturation. Decreased miRNA levels (e.g. let-7b<sup>61</sup>) are associated with increased Akt1 phosphorylation status and activity. Thus, fully activated Akt1 (ppAkt1<sup>T308,S473</sup>) phosphorylates Gld2 at S116 and silences the stabilizing/maturing effect of Gld2 on miRNA. Through phosphorylation of Gld2, Akt1 activity is expected to reduce miRNA levels. Phosphorylation is indicated by the orange circles.

#### 3.4.4 Conclusion

While hundreds of Akt1 substrates have been validated and/or predicted in human cells<sup>73</sup>, miRNA editing enzymes were not previously known to be part of the Akt1 signaling network. Similarly, it was unclear how Gld2 activity may be regulated to respond to external stimuli and signaling pathways, in turn controlling miRNAs and an even larger number of downstream mRNA substrates. We here revealed the first link between the activity of oncogenic kinases Akt1 and PKA and the regulation of Gld2. We found that HEK 293 cells contain N-terminal Gld2 kinase activity and that phosphorylation sites in the N-terminal domain of Gld2 can either positively or negatively regulate nucleotide addition activity. We identified Gld2 as a *bona fide* substrate of PKA and Akt1 and the site-specific phosphorylation catalyzed by either kinase at Gld2 S116 abolishes nucleotide addition activity. While the overall impact of Akt1/PKA signaling on miRNA metabolism remains to be investigated in a cellular context, these data significantly enhance our knowledge on miRNA regulation and reveal a

previously unrecognized link between oncogenic signal transduction and the regulation of tumor suppressor miRNAs.

### 3.5 Acknowledgements

This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada (RGPIN 04776-2014 to I.U.H.; RGPIN 04282-2014 to P.O.), the J. P. Bickell Foundation to I.U.H., the Canada Foundation for Innovation (229917 to P.O.), the Ontario Research Fund (229917 to P.O.), the Canada Research Chairs Program (950-229917 to P.O.), the Canadian Cancer Society Research Institute Innovation grant (704324 to P.O.), an Ontario Graduate Scholarship to C.Z.C., an Alexander Graham Bell Canada Graduate Scholarship-Doctoral from the Natural Sciences and Engineering Research Council of Canada to C.Z.C., a Queen Elizabeth II Scholarship in Science and Technology to N.B., and a scholarship from the Saudi Arabian Cultural Bureau to E.M. We are thankful to Dr. David W. Litchfield, Dr. Laszlo Gyenis, and Teresa Nuñez de Villavicencio-Díaz for their advice on kinases and their gift of CK2 $\alpha$ , CK2 holoenzyme, Abl, PKA, and CDK5 kinases and to Matthew Turk and Anish Engineer for advice and discussion.

### 3.6 References

1. Paul, P. *et al.* Interplay between miRNAs and human diseases. *J. Cell. Physiol.* **233**, 2007–2018 (2018).
2. De Almeida, C., Scheer, H., Zuber, H. & Gagliardi, D. RNA uridylation: a key posttranscriptional modification shaping the coding and noncoding transcriptome. *Wiley Interdiscip. Rev. RNA* **9**, (2018).
3. 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* **1861**, 2971–2979 (2017).
4. 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).
5. Jones, M. R. *et al.* Zcchc11 uridylates mature miRNAs to enhance neonatal IGF-1 expression, growth, and survival. *PLoS Genet.* **8**, e1003105 (2012).



6. Jones, M. R. *et al.* Zcchc11-dependent uridylation of microRNA directs cytokine expression. *Nat. Cell Biol.* **11**, 1157–1163 (2009).
7. 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).
8. Winter, J., Jung, S., Keller, S., Gregory, R. I. & Diederichs, S. Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nat. Cell Biol.* **11**, 228–234 (2009).
9. Kadyk, L. C. & Kimble, J. Genetic regulation of entry into meiosis in *Caenorhabditis elegans*. *Dev. Camb. Engl.* **125**, 1803–1813 (1998).
10. Wang, L., Eckmann, C. R., Kadyk, L. C., Wickens, M. & Kimble, J. A regulatory cytoplasmic poly(A) polymerase in *Caenorhabditis elegans*. *Nature* **419**, 312–316 (2002).
11. 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).
12. Glahder, J. A. & Norrild, B. Involvement of hGLD-2 in cytoplasmic polyadenylation of human p53 mRNA. *APMIS* **119**, 769–775 (2011).
13. Nousch, M., Minasaki, R. & Eckmann, C. R. Polyadenylation is the key aspect of GLD-2 function in *C. elegans*. *RNA N. Y. N* **23**, 1180–1187 (2017).
14. Chung, C. Z., Jo, D. H. S. & Heinemann, I. U. Nucleotide specificity of the human terminal nucleotidyltransferase Gld2 (TUT2). *RNA N. Y. N* **22**, 1239–1249 (2016).
15. 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. U. S. A.* **112**, 8614–8619 (2015).
16. Bandiera, S., Pfeffer, S., Baumert, T. F. & Zeisel, M. B. miR-122--a key factor and therapeutic target in liver disease. *J. Hepatol.* **62**, 448–457 (2015).
17. Masaki, T. *et al.* miR-122 stimulates hepatitis C virus RNA synthesis by altering the balance of viral RNAs engaged in replication versus translation. *Cell Host Microbe* **17**, 217–228 (2015).
18. Kim, G.-W. *et al.* Hepatitis C Virus Core Protein Promotes miR-122 Destabilization by Inhibiting GLD-2. *PLoS Pathog.* **12**, e1005714 (2016).
19. Moriya, K. *et al.* The core protein of hepatitis C virus induces hepatocellular carcinoma in transgenic mice. *Nat. Med.* **4**, 1065–1067 (1998).
20. Dai, R. *et al.* Hepatitis B virus X protein-induced upregulation of CAT-1 stimulates proliferation and inhibits apoptosis in hepatocellular carcinoma cells. *Oncotarget* (2017). doi:10.18632/oncotarget.17631
21. Chang, J. *et al.* miR-122, a mammalian liver-specific microRNA, is processed from hcr mRNA and may downregulate the high affinity cationic amino acid transporter CAT-1. *RNA Biol.* **1**, 106–113 (2004).
22. Peng, F. *et al.* HBx down-regulated Gld2 plays a critical role in HBV-related dysregulation of miR-122. *PloS One* **9**, e92998 (2014).
23. Gebert, L. F. R. *et al.* Miravirsin (SPC3649) can inhibit the biogenesis of miR-122. *Nucleic Acids Res.* **42**, 609–621 (2014).
24. Van der Ree, M. H. *et al.* Miravirsin dosing in chronic hepatitis C patients results in decreased microRNA-122 levels without affecting other microRNAs in plasma. *Aliment. Pharmacol. Ther.* **43**, 102–113 (2016).

25. Janssen, H. L. A. *et al.* Treatment of HCV infection by targeting microRNA. *N. Engl. J. Med.* **368**, 1685–1694 (2013).
26. Iino, I. *et al.* Effect of miR-122 and its target gene cationic amino acid transporter 1 on colorectal liver metastasis. *Cancer Sci.* **104**, 624–630 (2013).
27. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
28. Roberts, E., Eargle, J., Wright, D. & Luthey-Schulten, Z. MultiSeq: unifying sequence and structure data for evolutionary analysis. *BMC Bioinformatics* **7**, 382 (2006).
29. Veidenberg, A., Medlar, A. & Löytynoja, A. Wasabi: An Integrated Platform for Evolutionary Sequence Analysis and Data Visualization. *Mol. Biol. Evol.* **33**, 1126–1130 (2016).
30. 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).
31. Balasuriya, N. *et al.* Genetic code expansion and live cell imaging reveal that Thr308 phosphorylation is irreplaceable and sufficient for Akt1 activity. *J. Biol. Chem.* (2018). doi:10.1074/jbc.RA118.002357
32. Amanchy, R. *et al.* A curated compendium of phosphorylation motifs. *Nat. Biotechnol.* **25**, 285–286 (2007).
33. Xue, Y. *et al.* GPS 2.0, a Tool to Predict Kinase-specific Phosphorylation Sites in Hierarchy. *Mol. Cell. Proteomics MCP* **7**, 1598–1608 (2008).
34. Balasuriya, N., McKenna, M., Liu, X., Li, S. & O'Donoghue, P. Phosphorylation-Dependent Inhibition of Akt1. *Genes* **9**, 450 (2018).
35. MacLean, B. *et al.* Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinforma. Oxf. Engl.* **26**, 966–968 (2010).
36. Hirose, Y. & Ohkuma, Y. Phosphorylation of the C-terminal domain of RNA polymerase II plays central roles in the integrated events of eucaryotic gene expression. *J. Biochem. (Tokyo)* **141**, 601–608 (2007).
37. Mohan, N., AP, S., Francis, N., Anderson, R. & Laishram, R. S. Phosphorylation regulates the Star-PAP-PIPK1 $\alpha$  interaction and directs specificity toward mRNA targets. *Nucleic Acids Res.* **43**, 7005–7020 (2015).
38. Mertins, P. *et al.* Proteogenomics connects somatic mutations to signalling in breast cancer. *Nature* **534**, 55–62 (2016).
39. Mertins, P. *et al.* Ischemia in tumors induces early and sustained phosphorylation changes in stress kinase pathways but does not affect global protein levels. *Mol. Cell. Proteomics MCP* **13**, 1690–1704 (2014).
40. Kettenbach, A. N. *et al.* Quantitative phosphoproteomics identifies substrates and functional modules of Aurora and Polo-like kinase activities in mitotic cells. *Sci. Signal.* **4**, rs5 (2011).
41. Francavilla, C. *et al.* Multilayered proteomics reveals molecular switches dictating ligand-dependent EGFR trafficking. *Nat. Struct. Mol. Biol.* **23**, 608–618 (2016).
42. Amanchy, R. *et al.* A curated compendium of phosphorylation motifs. *Nat. Biotechnol.* **25**, 285–286 (2007).
43. Lee, R. S. *et al.* Relative Expression Levels Rather Than Specific Activity Plays the Major Role in Determining In Vivo AKT Isoform Substrate Specificity. *Enzyme Res.* **2011**, 720985 (2011).

44. Schmidt, M.-J. & Norbury, C. J. Polyadenylation and beyond: emerging roles for noncanonical poly(A) polymerases. *Wiley Interdiscip. Rev. RNA* **1**, 142–151 (2010).
45. Munoz-Tello, P., Gabus, C. & Thore, S. Functional implications from the Cid1 poly(U) polymerase crystal structure. *Struct. Lond. Engl.* **20**, 977–986 (2012).
46. Chung, C. Z. *et al.* RNA surveillance by uridylation dependent RNA decay in *Schizosaccharomyces pombe*. *Nucleic Acids Res.* (2019). doi:10.1093/nar/gkz043
47. Rissland, O. S., Mikulasova, A. & Norbury, C. J. Efficient RNA polyuridylation by noncanonical poly(A) polymerases. *Mol. Cell. Biol.* **27**, 3612–3624 (2007).
48. Yates, L. A. *et al.* Structural plasticity of Cid1 provides a basis for its distributive RNA terminal uridylyl transferase activity. *Nucleic Acids Res.* **43**, 2968–2979 (2015).
49. 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).
50. Nakel, K. *et al.* Structural basis for the antagonistic roles of RNP-8 and GLD-3 in GLD-2 poly(A)-polymerase activity. *RNA N. Y. N* **22**, 1139–1145 (2016).
51. Faehnle, C. R., Walleshauser, J. & Joshua-Tor, L. Multi-domain utilization by TUT4 and TUT7 in control of let-7 biogenesis. *Nat. Struct. Mol. Biol.* **24**, 658–665 (2017).
52. Yamashita, S., Takagi, Y., Nagaike, T. & Tomita, K. Crystal structures of U6 snRNA-specific terminal uridylyltransferase. *Nat. Commun.* **8**, 15788 (2017).
53. Maitra, A., Shulgina, I. & Hernandez, V. J. Conversion of active promoter-RNA polymerase complexes into inactive promoter bound complexes in *E. coli* by the transcription effector, ppGpp. *Mol. Cell* **17**, 817–829 (2005).
54. Phatnani, H. P. & Greenleaf, A. L. Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev.* **20**, 2922–2936 (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. Rouhana, L. *et al.* Vertebrate GLD2 poly(A) polymerases in the germline and the brain. *RNA N. Y. N* **11**, 1117–1130 (2005).
58. Yamagishi, R., Tsusaka, T., Mitsunaga, H., Maehata, T. & Hoshino, S. The STAR protein QKI-7 recruits PAPD4 to regulate post-transcriptional polyadenylation of target mRNAs. *Nucleic Acids Res.* **44**, 2475–2490 (2016).
59. Pearce, L. R., Komander, D. & Alessi, D. R. The nuts and bolts of AGC protein kinases. *Nat. Rev. Mol. Cell Biol.* **11**, 9–22 (2010).
60. Menon, B., Sinden, J., Franzo-Romain, M., Botta, R. B. & Menon, K. M. J. Regulation of LH receptor mRNA binding protein by miR-122 in rat ovaries. *Endocrinology* **154**, 4826–4834 (2013).
61. Yamada, S. *et al.* Epigallocatechin-3-O-gallate up-regulates microRNA-let-7b expression by activating 67-kDa laminin receptor signaling in melanoma cells. *Sci. Rep.* **6**, 19225 (2016).
62. Agarwal, E., Brattain, M. G. & Chowdhury, S. Cell Survival and Metastasis Regulation by Akt Signalling in Colorectal Cancer. *Cell. Signal.* **25**, 1711–1719 (2013).

63. Spencer, A. *et al.* The novel AKT inhibitor afuresertib shows favorable safety, pharmacokinetics, and clinical activity in multiple myeloma. *Blood* **124**, 2190–2195 (2014).
64. Oh, W. J. *et al.* mTORC2 can associate with ribosomes to promote cotranslational phosphorylation and stability of nascent Akt polypeptide. *EMBO J.* **29**, 3939–3951 (2010).
65. Yao, J., Zhang, P., Li, J. & Xu, W. MicroRNA-215 acts as a tumor suppressor in breast cancer by targeting AKT serine/threonine kinase 1. *Oncol. Lett.* **14**, 1097–1104 (2017).
66. Ru, N. *et al.* MiR-564 is down-regulated in osteosarcoma and inhibits the proliferation of osteosarcoma cells via targeting Akt. *Gene* **645**, 163–169 (2018).
67. Sun, H., Ding, C., Zhang, H. & Gao, J. Let-7 miRNAs sensitize breast cancer stem cells to radiation-induced repression through inhibition of the cyclin D1/Akt1/Wnt1 signaling pathway. *Mol. Med. Rep.* **14**, 3285–3292 (2016).
68. Lian, J.-H., Wang, W.-H., Wang, J.-Q., Zhang, Y.-H. & Li, Y. MicroRNA-122 promotes proliferation, invasion and migration of renal cell carcinoma cells through the PI3K/Akt signaling pathway. *Asian Pac. J. Cancer Prev. APJCP* **14**, 5017–5021 (2013).
69. Blahna, M. T. & Hata, A. Regulation of miRNA biogenesis as an integrated component of growth factor signaling. *Curr. Opin. Cell Biol.* **25**, 233–240 (2013).
70. Manning, B. D. & Toker, A. AKT/PKB Signaling: Navigating the Network. *Cell* **169**, 381–405 (2017).
71. Caretta, A. & Mucignat-Caretta, C. Protein Kinase A in Cancer. *Cancers* **3**, 913–926 (2011).
72. Johnson, S. M. *et al.* RAS is regulated by the let-7 microRNA family. *Cell* **120**, 635–647 (2005).
73. Manning, B. D. & Cantley, L. C. AKT/PKB signaling: navigating downstream. *Cell* **129**, 1261–1274 (2007).

## Chapter 4

### 4 RNA surveillance by uridylation-dependent RNA decay in *Schizosaccharomyces pombe*

#### 4.1 Introduction

RNA synthesis and degradation are regulated through a variety of mechanisms that amend the transcriptome to match cellular needs throughout the cell cycle and adaptation to environmental changes<sup>1</sup>. Messenger RNA (mRNA) degradation can proceed by two general pathways, in either a 5'-3' or 3'-5' direction, catalyzed by exonucleases or the exosome complex, respectively. These canonical RNA degradation processes usually commence with an initial deadenylation step, followed by decapping by Dcp1/2 and the Lsm1–7 complex. Decapped mRNA is subsequently accessible to 5'-3' decay catalyzed by the exonuclease Xrn1, while exosome-catalyzed 3'-5' degradation does not require decapping<sup>2</sup>. Recently, a second deadenylation-independent pathway of mRNA decay was discovered and appears to be conserved in many eukaryotes. Here, uridylation of polyadenylated mRNAs recruits the Lsm1–7 complex and subsequently leads to mRNA degradation by designated exonucleases<sup>2</sup>. This template-independent addition of nucleotides is catalyzed by terminal RNA nucleotidyltransferases (TENTs), a subfamily of the polymerase  $\beta$  superfamily of nucleotidyltransferases<sup>3</sup>. TENTs add ribonucleoside monophosphates to an RNA substrate through a catalytic process involving two metal ion cofactors<sup>3</sup>. Of note, non-templated 3'-end uridylation of a variety of RNA species plays key roles in eukaryotic RNA processing pathways including mRNA and pre-miRNA degradation, pre-miRNA maturation, and miRNA silencing<sup>4–6</sup>. RNA uridylation is catalyzed by terminal uridylyltransferases (Tutases), and polyuridylylated RNAs are subsequently degraded by the U-specific exonuclease Dis3L2<sup>6–8</sup>. While uridylation and deadenylation-dependent RNA decay show some redundancy, uridylation is conserved in many different species indicating that it is important for RNA turnover<sup>9–11</sup>.

Fission yeast Cid1 (caffeine-induced death suppressor protein 1) was first discovered in a genetic screen identifying components of the S-M cell cycle checkpoint

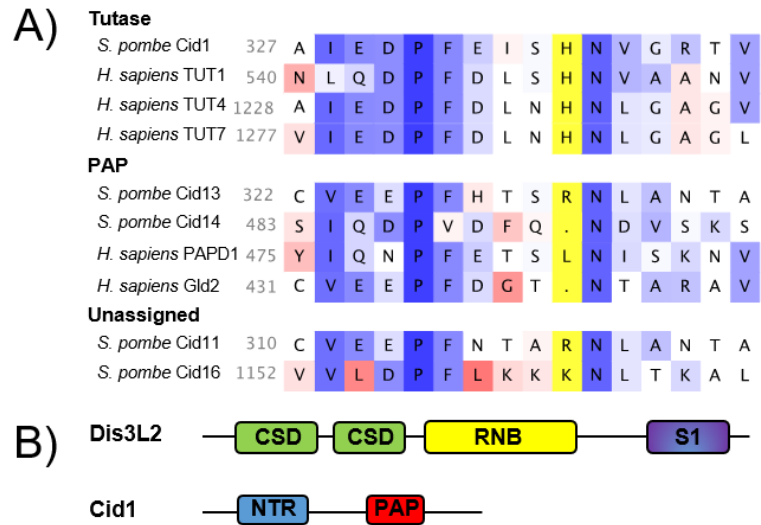
in *Schizosaccharomyces pombe*<sup>12</sup>. Although *S. pombe*  $\Delta cid1$  strains are viable, they are sensitive to a combination of hydroxyurea, a ribonucleotide reductase inhibitor, and caffeine, which overrides the S-M checkpoint and induces mitosis. Overexpression of Cid1 confers resistance to this combination of stressors<sup>12</sup>. Cid1 was originally thought to be a poly(A) polymerase due to its significant *in vitro* poly(A) polymerase activity<sup>13</sup>, but recent evidence characterized it as an efficient TUTase *in vitro* and *in vivo*<sup>14–16</sup>. Cid1 encodes a catalytic nucleotidyltransferase motif and a poly(A) polymerase-associated motif<sup>17</sup>, but lacks an identifiable RNA recognition motif. Interestingly, nucleotide specificity appears to have evolved after RNA specificity, with adenylyltransferases and uridylyltransferases playing opposing roles in promoting RNA stability or degradation in eukaryotes, respectively<sup>18</sup>. Nucleotide specificity depends on a critical histidine residue (H336), which is responsible for UTP over ATP preference<sup>19,20</sup> (Figure 4.1A). A H336N mutation in Cid1 converts the enzyme to an adenylyltransferase<sup>16,20</sup>, whereas a histidine insertion in its human adenylyltransferase counterpart Gld2 confers UTP specificity<sup>18</sup>.

One of the first Cid1 RNA substrates to be identified was *actin1* mRNA, which was shown to be uridylated upon S-phase arrest in a Cid1-dependent manner<sup>15</sup>. In *S. pombe*, RNA uridylation mediates mRNA turnover: Cid1 uridylates polyadenylated mRNAs to trigger Lsm1–7-mediated decapping of the RNA 5'-end and subsequent degradation by the U-specific exonuclease Dis3L2<sup>7,10</sup>. Biochemical and structural investigations revealed that despite the absence of a specific RNA recognition motif (Figure 4.1B), Cid1 is capable of binding and uridylating RNAs in a sequence-independent manner<sup>14</sup>. Due to its substrate promiscuity, Cid1 is thought to participate in a widespread mechanism of mRNA decay in *S. pombe*<sup>11,17,19,21,22</sup>, and substrate specificity and selectivity may require accessory proteins, in analogy to the human homologs, Tutases TUT4 and TUT7 and the adenylyltransferase Gld2<sup>18,23–26</sup>.

Following uridylation, RNAs are quickly degraded by the U-specific 3'-5' exonuclease Dis3L2<sup>6–8,27–29</sup>. Recent studies revealed that Dis3L2-catalyzed exonucleolytic RNA degradation constitutes an alternative pathway for RNA decay, independent of exosome and Xrn1-catalyzed decay pathways<sup>7</sup>. In *S. pombe*, Dis3L2 localizes to the cytoplasm and does not associate with the exosome but interacts with

components of the cytoplasmic mRNA degradation pathway. While a recent study reported no significant changes in mRNA accumulation in a *dis3L2* deletion strain, uridylated mRNAs were found elevated in a *dis3L2* and *lsm1* double mutant strain, and recombinant Dis3L2 degraded uridylated RNA transcripts *in vitro*<sup>7</sup>. In humans, Dis3L2 is involved in the degradation of uridylated mRNA and miRNA transcripts<sup>6,7,30–32</sup>. Mutations in Dis3L2 in humans are associated with the Perlman syndrome of fetal overgrowth, likely due to its role in the degradation of miRNAs and pre-miRNAs of the let-7 family<sup>33</sup>. Dis3L2 displays a typical RNase II-like protein domain organization, and encodes two cold shock domains (CSDs), an exonucleolytic ribonuclease domain (RNB), and a nonspecific RNA binding domain (S1) (Figure 4.1B). Structural analysis of Dis3L2 showed that in the absence of RNA, the enzyme displays an open conformation<sup>28</sup> and RNA binding induces a closed conformation, where three RNA binding domains form a funnel to position the RNA substrate for exonucleolytic degradation<sup>32</sup>.

In *S. pombe*, the Cid1/Dis3L2 RNA degradation pathway constitutes one of three mRNA surveillance pathways. While the individual proteins, Cid1 and Dis3L2, are now recognized and biochemically and structurally characterized, it is unclear whether the three RNA decay pathways (Xrn1, exosome, and Cid1/Dis3L2) are dedicated to specific substrate RNAs or act as three global albeit independent decay mechanisms. In this study, we capture the extent of the Cid1/Dis3L2 mediated RNA degradation and find that depletion of uridylation-dependent RNA decay causes the accumulation of misfolded proteins and an increase in abundance of mRNAs involved in the stress response. Using deep sequencing, we find that while Cid1 depletion has little impact on mRNA homeostasis, Dis3L2 represents a bottleneck in uridylation-dependent RNA decay and its depletion leads to an increase in mRNAs involved in protein folding and degradation pathways. We conclude that perturbation of uridylation-dependent RNA decay elicits a stress response, likely due to the accumulation of misfolded proteins.



**Figure 4.1: Domain structure and amino acid composition of Cid1 and Dis3L2.**

**A)** Amino acid sequence alignment adapted from<sup>18</sup>. Enzymes known to exercise Tutase activity encode a histidine residue (His336 in Cid1, highlighted in yellow), that sterically hinders the larger ATP from entering the active site. Adenylyltransferases (PAPs) do not encode the respective histidine residue. Nucleotide preference for *S. pombe* Cid11 and Cid16 is undetermined, though Cid16 likely prefers UTP. **B)** Dis3L2 displays a typical RNase II domain organization, encoding two cold shock domains (CSD), an exonucleolytic ribonuclease domain (RNB), and a nonspecific RNA binding domain (S1). Cid1 is composed of a nucleotidyltransferase domain (NTR) and a poly(A)polymerase domain (PAP).

## 4.2 Materials and Methods

### 4.2.1 Yeast strains and growth conditions

*S. pombe* strains were obtained from Bioneer (Alameda, CA, USA): Wildtype (BG\_0000H6, ade6-M210 ura4-D18 leu1-32);  $\Delta$ dis3L2 (BG\_H0669; orf $\Delta$  SPAC2C4.07c: kanMX4/ORF ade6-M210 ura4-D18 leu1-32) and  $\Delta$ cid1 (BG\_H0513; orf $\Delta$  SPAC19D5.03: kanMX4/ORF ade6-M210 ura4-D18 leu1-32). Liquid cultures were grown at 30°C in YPD supplemented with adenine to an optical density of OD<sub>600</sub> = 0.5. For spotting assays, yeast were grown on Edinburgh minimal media (EMM) (3 g/l potassium hydrogen phthalate, 2.2 g/l sodium phosphate dibasic, 5 g/l ammonium



chloride, 20 g/l dextrose, 2.1 g/l mineral salts, 0.02 g/l vitamins, 3 mg/l trace elements) supplemented with 12 g/l L-leucine (leu), 2 g/l uracil (ura), 2 g/l adenine (ade).

#### 4.2.2 Spotting assays

*S. pombe* cells were inoculated in 4 mL EMM-URA-LEU-ADE liquid media overnight in a 30°C incubator shaker. 100 µl of cells were diluted 1:10 in ddH<sub>2</sub>O to measure the OD<sub>600</sub> to determine cell density. Cells were standardized to OD<sub>600</sub> = 1 in the first row of wells on a 96-well plate. A 1:5 serial dilution of cells was performed in the subsequent five rows of wells. Cells were spotted on YES, and EMM-uracil-leucine-adenine media agar plates with or without 100 µM H<sub>2</sub>O<sub>2</sub>, 5 mM caffeine, 2 mM caffeine, 5 mM hydroxyurea, 2 mM hydroxyurea, and 2.5 mM caffeine + 10 mM hydroxyurea. Plates were incubated in a 30°C incubator. Photographs of plates were taken on different days to document growth. Spotting assays were photographed, and the image was modified to black and white with the background blackened out and the yeast colonies being white. The circular selection tool on ImageJ was used to select an equal area of colonies and the mean gray value (MGV) was measured for density of cell growth. The blackened plate background gave a MGV of 0 and complete colony growth gave a value of 255. Wildtype values were normalized to 1 and the growth of deletion strains were normalized against wildtype to give a fraction of 1. Unpaired *t*-test was employed to infer statistical significance between wildtype and each deletion strain at 95% confidence interval.

#### 4.2.3 Cid1 cloning, purification, and activity assays

Total RNA was extracted from *S. pombe* cells using MasterPure Complete DNA and RNA Purification kit (Epicentre) and reverse transcribed with SuperScript II Reverse Transcriptase (Invitrogen) and random hexamer primers. The resulting cDNA was amplified by PCR using gene specific primers (Cid1F 5'-AAGCTTATGAACATTTCTTCTGCACAATTTATTCCTGGTGT-3' and Cid1R 5'-CTCGAGCTCAGAATTGTCACCATCGGTTTCATTC-3') and inserted into a pET-20b(+) expression vector with HindIII and XhoI restriction sites. The construct was confirmed by DNA sequencing (London Regional Genomics Centre). *Escherichia coli* BL21 Codon Plus cells were transformed

with pET-20b(+) encoding His-tagged *cid1* and grown in LB media with ampicillin (100 µg/ml) and chloramphenicol (34 µg/ml) at 37°C to an OD<sub>600</sub> of 0.6. Protein expression was induced by the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM and the culture was grown overnight at 18°C. Cells were harvested and resuspended in Buffer A (50 mM HEPES, pH 6.8, 100 mM NaCl, 2 mM 2-mercaptoethanol) supplemented with 1 mM phenylmethylsulfonyl fluoride (PMSF) and 0.25 mg/ml lysozyme and lysed with a French Pressure Cell. Following 1 hour of centrifugation (41 000 rpm at 4°C), cell free extract was loaded onto a gravity column containing HisPur Ni-NTA resin (Thermo Scientific) pre-equilibrated in Buffer A. The resin was washed with Buffer B (50 mM HEPES, pH 6.8, 100 mM NaCl, 2 mM 2-mercaptoethanol, 50 mM imidazole) and Cid1 was eluted with Buffer C (50 mM HEPES, pH 6.8, 100 mM NaCl, 100 mM imidazole). After concentrating the eluted protein, remaining contaminants were removed by size exclusion chromatography using a Superdex™ 200 Increase 5/150 GL column and Buffer A. Protein concentration was determined using a Bradford assay and purified proteins were stored at –80°C with 10% glycerol. Enzyme activity assays were carried out as described previously<sup>18</sup>.

#### 4.2.4 Circular rapid amplification of cDNA ends (cRACE) and Northern blotting

RNA was isolated using the Masterpure RNA purification kit (Epicentre), dephosphorylated with shrimp alkaline phosphatase (NEB), and phenol chloroform extracted. For decapping, RNA was incubated with tobacco acid pyrophosphatase (NEB) and circularized with T4 RNA ligase (NEB). Reverse transcription was carried out using Superscript Reverse Transcriptase (ThermoFisher) and random hexamer primers. Gene specific amplification was carried out using primers listed in Appendix B Table B1 and PCR products cloned into pCR3.1- TopoTA vector (ThermoFisher) and sequenced at the London Regional Genomics Centre. Northern blots were performed as previously described<sup>34</sup> using 5 or 10 µg of total RNA and 5'-<sup>32</sup>P-labelled gene specific probes amplified from genomic DNA using primers listed in Appendix B Table B1.

#### 4.2.5 RT-qPCR

Reverse transcription and quantitative PCR (RT-qPCR) were performed as described<sup>35</sup>. Briefly, total RNA was extracted using the Masterpure RNA purification kit (Epicentre) and reverse transcribed using the High-Capacity cDNA Reverse Transcription kit (Applied Biosystems). PowerUp SYBR Green Master Mix (Applied Biosystems) was used for the qPCR and amplification was performed on the ViiA 7 Real-Time PCR System. Six biological replicates of each *S. pombe* strain (wildtype,  $\Delta dis3L2$ , and  $\Delta cid1$ ) were analyzed in technical triplicates. All primers are listed in Table B1.

#### 4.2.6 Yeast sedimentation assay and Western blot

Wildtype,  $\Delta dis3L2$ , and  $\Delta cid1$  *S. pombe* strains were streaked onto YPD agar plates supplemented with adenine, uracil, and leucine and incubated at 30°C. Three biological replicates for each strain were grown in 5 mL YPD media supplemented with adenine, uracil, and leucine overnight at 30°C. The OD<sub>600</sub> of all cultures was measured the next morning before the cells were centrifuged at 4000 × g for 5 min at 4°C. The sedimentation assay was adapted from<sup>36</sup>. Briefly, cell pellets were lysed in 200 µl lysis buffer (100 mM Tris-HCl pH 7.5, 200 mM NaCl, 1 mM ethylenediaminetetraacetic acid (EDTA), 1 mM dithiothreitol (DTT), 5% glycerol, 0.5% Triton-X, 2 mM phenylmethylsulfonyl fluoride (PMSF)) supplemented with protease inhibitors (Roche, 04 693 159 001) using glass beads on a Disruptor Genie. Cells were disrupted with six 30 s bursts followed by 30 s on ice between each burst. The lysates were separated from the glass beads and 50 µl was added to 50 µl SUMEB (1% sodium dodecyl sulfate (SDS), 8 M urea, 10 mM 3-(*N*-morpholino)propanesulfonic acid (MOPS), 10 mM EDTA, 0.01% bromophenol blue). The remaining lysates were centrifuged at 500 × g for 15 mins at 4°C and 100 µl of the supernatant was added to 100 µl of SUMEB. The pellets were resuspended in 100 µl lysis buffer with no PMSF and 100 µl SUMEB. Samples were analyzed on a 10% SDS gel and total protein in each lane was quantified using the Bio-Rad ChemiDoc MP and Image Lab software. Total lysates were blotted for PGK1 (ThermoFisher Scientific, 459250) as loading control.

#### 4.2.7 RNA sequencing and data analysis

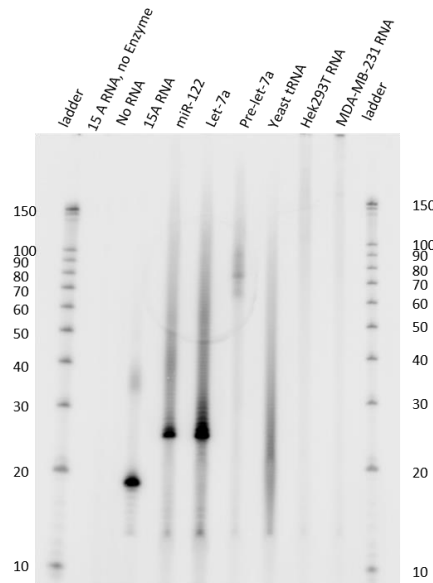
Three biological and three technical replicates for each *S. pombe* strain (WT,  $\Delta$ dis3L2 and  $\Delta$ cid1) were grown in YPD media to an OD<sub>600</sub> of 0.6–0.65. RNA was isolated using the Masterpure RNA purification kit (Epicentre). Ribosomal RNAs were depleted using the Ribo zero RNA kit and the RNA library was generated with the NEBNext® Ultra™ Directional RNA Library Prep Kit for Illumina. Samples were analyzed on the MiSeq sequencing v2. Sequencing reads were mapped to the *S. pombe* mitochondrion (MT) genome (NC\_001326.1) using the CLC Genomics Workbench and changes in gene expression were analyzed with the ANOVA-like Differential Expression (ALDEx2) tool<sup>37</sup> and an effect size cut-off of 1.5. The relative expression (abundance) of each gene within a sample was calculated as the median centered log-ratio (clr) from 1000 Monte Carlo Dirichlet instances. Genes were considered to be differentially expressed if the ALDEx2 effect size was greater than 1.5 (i.e. the difference in abundance between two strains was at least 1.5-fold greater than the difference between biological replicates). Genes differentially up- or down-regulated by at least 1.5-fold were analyzed for enrichments in specific pathways using the Search Tool for the Retrieval of Interacting Genes/proteins (STRING)<sup>38</sup>.

### 4.3 Results

#### 4.3.1 Recombinant Cid1 displays ambiguous substrate specificity *in vitro*

Previous studies have shown that *S. pombe* Cid1 uridylates a variety of RNA substrates *in vivo*<sup>15</sup>, which may subsequently be degraded by the exonuclease Dis3L2<sup>7</sup>. To assess whether Cid1 uridylation is ambiguous or dedicated to specific substrates, we produced and purified full length Cid1 and assessed its substrate specificity on several RNA substrates *in vitro*. We found that Cid1 uridylates a poly(A) tail mimic (15A), tRNA, total mRNA, pre-miRNA and miRNAs equally (Figure 4.2). It appears that Cid1 uridylates RNA substrates regardless of their secondary structure (pre-miRNA hairpin structure, tRNA structure), or sequence (miRNAs and poly(A) RNA). Total yeast tRNA

and total human RNA preparations from cell lines HEK 293T and MDA-MB-231 are heterogeneous mixtures by nature and the frequency of poly- *versus* monouridylation cannot be assessed. In these cases, the observed products are presented in a smear, consistent with either the uridylation of RNA substrates of varying length or also a mixture of poly- and monouridylation events. Substrates 15A RNA, miRNAs and pre-miRNA were either purchased oligoribonucleotides or products of *in vitro* transcription (pre-let-7a)<sup>18</sup>. For these homogenous RNA substrates, the predominant product is consistent in length with a monouridylated RNA product. In the case of the 15A nucleotide poly(A) tail mimic 15A RNA, a second band at ~35 nucleotides is detectable, indicating the addition of roughly 20 uridines *in vitro*. For the microRNA substrates miR-122 and let-7a, polyuridylation can be observed in the form of a ladder-like pattern, as a result of multiple nucleotide additions with variable product lengths. Thus, Cid1 displays no substrate preference *in vitro* and can act as both a distributive or processive Tutase.

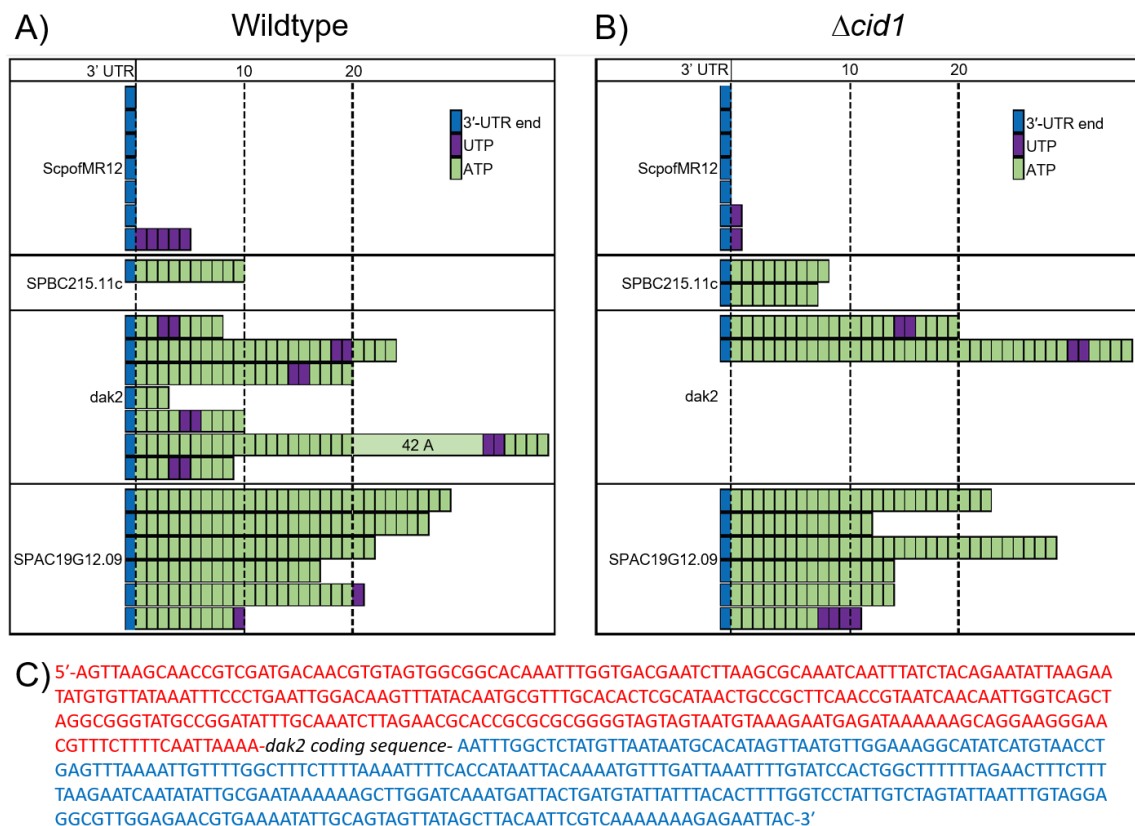


**Figure 4.2: Cid1 displays a promiscuous substrate range *in vitro*.**

Cid1 was incubated with different RNA substrates and [ $\alpha$ -<sup>32</sup>P]-UTP as indicated. Formation of [ $\alpha$ -<sup>32</sup>P]-labelled RNA products was monitored by electrophoretic separation and subsequent phosphorimaging. Cid1 catalyzed [ $\alpha$ -<sup>32</sup>P]-UTP addition to RNA substrates pre-let-7a (72 nts), mature human miRNAs let-7a-5p (22 nts) and miR122 (22 nts), an oligo(A) tail mimic 15A (15 nts), total yeast tRNA, and total RNA isolated from HEK 293T or MDA-MB-321 cells. Radiolabelled RNA Decade marker is used for reference.

#### 4.3.2 RNA uridylation is prominent in wildtype *S. pombe* and a *cid1* deletion strain

To further investigate the function and substrate range of Cid1 *in vivo*, we purified mRNA from wildtype and  $\Delta cid1$  *S. pombe* strains and amplified several mRNA species to sequence their 3'-end by cRACE. Surprisingly, uridylation of RNAs was found in both the wildtype and the deletion strain. The small subunit ribosomal RNA, *ScpofMR12* is, as typical for ribosomal RNA, not adenylated, and most samples also did not contain additional uridyl residues at the 3'-end (Figure 4.3). In *S. pombe* wildtype cells, one sequence was retrieved with multiple uridines added to the RNA 3'-end (Figure 4.3A) and in the *cid1* deletion strain, monouridylated RNA was found (Figure 4.3B). Since uridylation is rare and only few sequences were retrieved, no conclusions as to the general uridylation pattern can be drawn from this data. *SPBC215.11c*, a protein coding RNA, was polyadenylated in both strains, with no 3'-end uridylated RNAs recovered. Interestingly, a uridylation/adenylation pattern was found in several samples of *dak2*, a protein coding RNA, where a poly(A) tail of differing length was interceded by two uridyl residues, followed by an additional four adenines (5'-N<sub>x</sub>AUUAAAA-3'). This pattern was found only in *dak2* RNA, in eight out of nine sequenced RNA samples, and is not derived from the *dak2* 5'- or 3'-UTRs (Figure 4.3C). For another protein coding mRNA, *SPAC19G12.09*, samples encoding poly(A) tails without uridines were most prevalent. One polyuridylated sample was recovered from the *cid1* deletion strain and two monouridylated RNAs from wildtype *S. pombe* RNA. Again, due to sample size, no conclusions can be drawn whether this represents a general uridylation pattern. Though it appears that uridylation is slightly less prevalent in the *cid1* deletion strain, the small sample size and methodology of cRACE does not allow for a quantification of uridylation, but rather the qualitative observation that mono- and polyuridylation occurs in both wildtype and *cid1* deletion strains, and that *dak2* RNA is prone to an unusual pattern of RNA uridylation and adenylation.



**Figure 4.3: RNA uridylation of diverse RNA transcripts is found in wildtype and *cid1* deletion strains.**

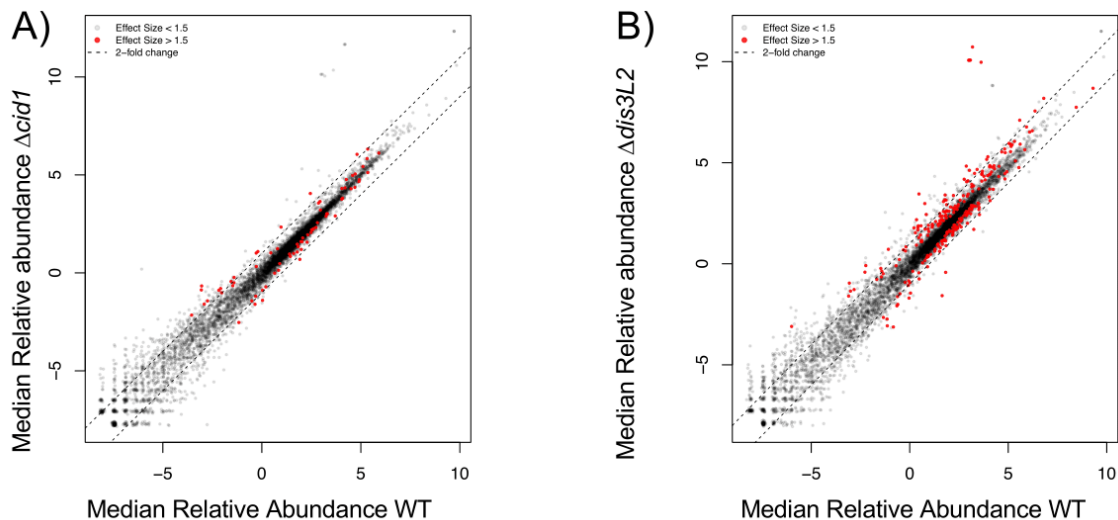
RNA was extracted from wildtype and *cid1* deletion strain, and 3'-ends of selected mRNAs were analyzed by cRACE. In both **A)** wildtype and **B)**  $\Delta cid1$  *S. pombe* cells transcripts containing terminal uridyl residues, and residues incorporated into the poly(A) tail were detected. A slight decrease in uridylated transcripts was detected upon *cid1* deletion. **(C)** Sequence of the *dak2* mRNA 3'-UTR (red) and 5'-UTR (blue). The protein coding sequence was omitted and is indicated as *dak2* coding sequence.

#### 4.3.3 Deletion of the Dis3L2 exonuclease elicits changes in the transcriptome

To further probe the prevalence of uridylation-dependent RNA decay in *S. pombe*, we isolated total RNA from wildtype,  $\Delta cid1$ , and  $\Delta dis3L2$  strains, depleted ribosomal RNA and analyzed the RNA content using deep sequencing. Reads were mapped to the *S. pombe* genome, and differentially expressed genes (effect size >1.5) with more than a 1.5-fold change in expression between wildtype and  $\Delta cid1$  or wildtype and  $\Delta dis3L2$  were

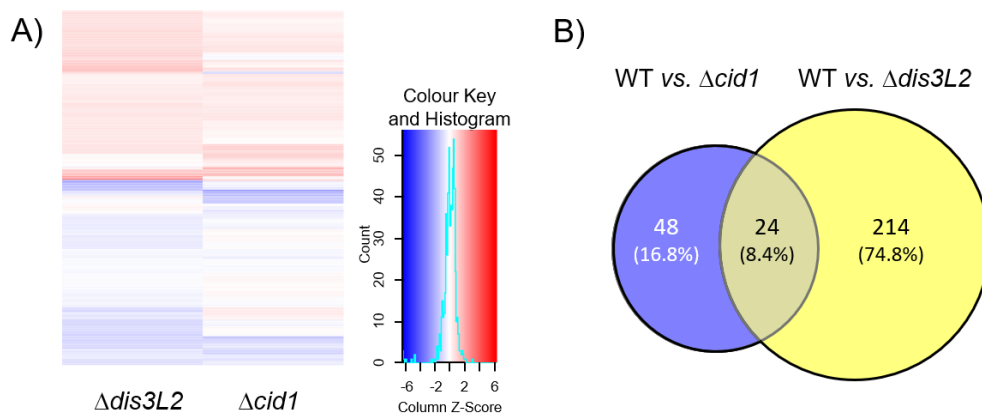
considered for the data analysis (Figure 4.4A and B, Appendix B Tables B2-B6). Gene expression changes in  $\Delta cid1$  and  $\Delta dis3L2$  deletion strains followed a similar trend, as outlined in Figure 4.5A, B and Appendix B Figure B1 and Tables B2-B6. Overall, 72 genes were found differentially expressed >1.5-fold between wildtype and  $\Delta cid1$ , and 214 genes were differentially expressed >1.5-fold between wildtype and  $\Delta dis3L2$  (Table B2). 24 of the genes were differentially regulated more than 1.5-fold in both deletion strains (Figure 4.5B). While changes in the transcriptome of >1.5-fold were more noticeable in  $\Delta dis3L2$ , many of the same genes were similarly up- or down-regulated in  $\Delta cid1$  (Figure 4.5A), albeit to a lesser extent. To verify the results obtained by RNA sequencing, we performed Northern Blotting on several RNAs found to be differentially expressed in the deletion strains. *ecf1* was shown to be 1.4-fold up-regulated in the *cid1* deletion strain, which was confirmed by Northern Blot (Figure 4.6A), and little to no expression changes were seen in *spac19g12.09*, *spac27e2.11c*, *thf1* and *tdh1*, which confirms our sequencing results (Table 4.1). We further used RT-qPCR to confirm our sequencing results (Figure 4.6B), and all data confirmed the data observed in our Next Generation Sequencing data. The qPCR data confirmed no significant change in the expression of *pex22* (Wt/dis3L2: qPCR 0.73-fold change, Sequencing 1.0-fold change). Four genes, *hsp104* (Wt/dis3L2: qPCR 2.8-fold change, Sequencing 3.9-fold change), *hsp78* (Wt/dis3L2: qPCR 1.8-fold change, Sequencing 2.7-fold change), *ssa2* (Wt/dis3L2: qPCR 1.8-fold change, Sequencing 2.9-fold change) and *tcg1* (Wt/dis3L2: qPCR 1.24-fold change, Sequencing 2.3-fold change) were more abundant in  $\Delta dis3L2$ , confirming our sequencing results (Table 4.1). As observed in our next generation sequencing data, the changes in mRNA abundance as measured by RT-qPCR are less pronounced in the *cid1* deletion strain than in the *dis3L2* deletion strain (Figure 4.6B).





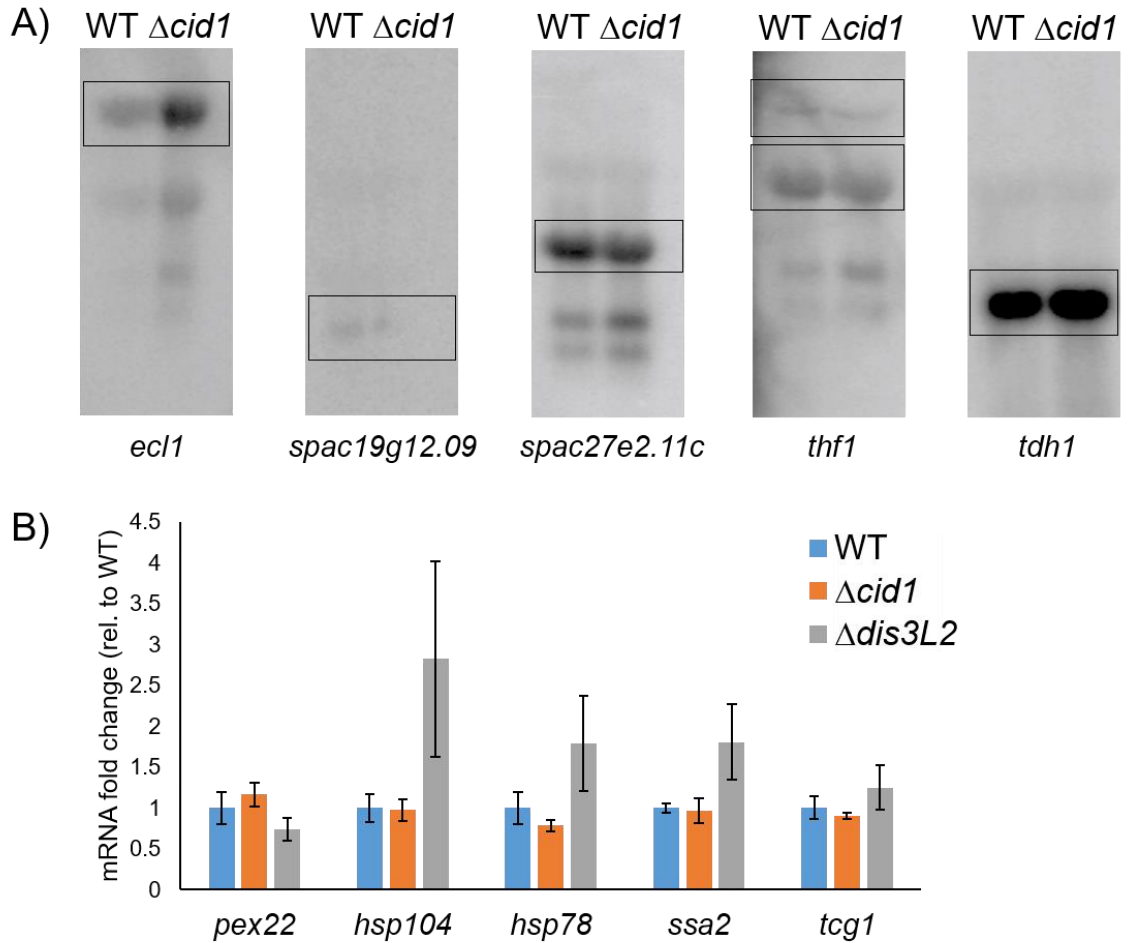
**Figure 4.4: Changes in relative abundance of mRNAs in WT,  $\Delta cid1$ , and  $\Delta dis3L2$  cells.**

Expression plot comparing relative abundance ( $\log_2$  centered log ratio, clr) of transcripts from a WT (x-axis) and **A)**  $\Delta cid1$  or **B)**  $\Delta dis3L2$  strain (y-axis). Differentially expressed genes (ALDEx2 effect size  $> 1.5$ ) are indicated in red, and dotted lines indicated a 2-fold change in expression from the line of best fit for the data (Pearson's  $r = 0.9805$ ).



**Figure 4.5: Genes differentially expressed in *S. pombe* deletion strains compared to wildtype.**

**A)** Heat map showing fold-change for significantly different genes with hits for both WT *S. pombe* and  $\Delta cid1$  and WT and  $\Delta dis3L2$  strains. A histogram is included in the colour key to show the gene fold-change distribution. **B)** Venn diagram of genes differentially expressed in  $\Delta cid1$  and  $\Delta dis3L2$  deletion strains versus wildtype.



**Figure 4.6: Northern blot and RT-qPCR showing differential expression of genes in wildtype versus  $\Delta cid1$  or  $\Delta dis3L2$  *S. pombe*.**

**A)** Northern blots of total RNA extracted from WT and  $\Delta cid1$  *S. pombe* was run on a 1% agarose gel in 1 x MOPS and capillary blotted onto a Nylon membrane overnight at 4°C. RNA was UV-crosslinked to the membrane and probed with gene-specific oligonucleotides labelled on the 5'-end with  $^{32}\text{P}$ . The blot was exposed to a phosphorimaging screen for two days at -80°C. *tdh1* (GAPDH) was used as loading control. Expected sizes are as following: *ecf1* with UTRs, 3597 nts; *spac19G12.09* with UTRs, 1213 nts; *spac27E2.11c* with UTRs, 2140 nts; *thf1* with UTRs, 3449 nts, with UTRs and introns, 3777 nts; *tdh1* with UTRs, 1518 nts. **B)** RT-qPCR was performed on WT,  $\Delta cid1$ , or  $\Delta dis3L2$  *S. pombe* cells to assess the gene expression changes in the absence of *cid1* or *dis3L2*. Cultures were grown to exponential phase ( $\text{OD}_{600} = 0.6$ ), harvested, and RNA isolated. Expression of each gene was normalized to the WT strain grown to early-exponential phase. Internal control used was *rpp0*. Error bars show the standard error on the mean ( $n = 6$ ).

**Table 4.1: Select genes up- or down-regulated in either *S. pombe*  $\Delta cid1$  deletion strain or *S. pombe*  $\Delta dis3L2$  deletion compared to wildtype *S. pombe*.**

Genes targeted in the Northern blot or RT-qPCR are listed.

		Fold change in $\Delta cid1$	Effect change in $\Delta cid1$	Fold change in $\Delta dis3L2$	Effect change in $\Delta dis3L2$
<i>SPAC19D5.02c</i>	peroxisomal membrane protein Pex22 (predicted)	2.56	31.96	1.01	1.03
<i>SPAC27E2.11C</i>	<i>Schizosaccharomyces</i> specific protein	2.00	3.49	1.82	3.57
<i>SPBC16D10.08c</i>	heat shock protein Hsp104 mitochondrial heatshock protein	1.60	2.85	3.96	5.18
<i>SPBC4F6.17c</i>	Hsp78 (predicted)	1.21	2.01	2.69	10.07
<i>SPBC839.16</i>	C1-5,6,7,8-tetrahydrofolate (THF) synthase, trifunctional enzyme Thf1	0.80	0.25	0.90	0.48
<i>ssa2</i>	heat shock protein Ssa2	1.63	2.13	2.88	4.96
<i>tcg1</i>	single-stranded telomeric binding protein Tgc1	1.30	2.34	1.76	3.60
<i>tdh1</i>	glyceraldehyde-3-phosphate dehydrogenase Tdh1	1.57	1.68	2.62	2.88

For genes >1.8-fold up- or down-regulated, we performed STRING analysis for enrichment of specific pathways<sup>38</sup> (Figure 4.7). In the  $\Delta dis3L2$  strain, we found significant enrichment in genes up-regulated in protein folding and protein degradation pathways (false discovery rate <0.001, Table 4.2). Differential expression for genes involved in stress response, especially heat shock proteins, chaperones, and protein degradation were most prominent, but enrichment was also observed in sugar and nucleotide metabolism (false discovery rate <0.01, Table 4.2), specifically in galactose metabolism (false discovery rate <0.02, Table 4.2). No significant enrichment was found for genes up- or down-regulated in the *cid1* deletion strain (Appendix B Figure B2A and B). Similarly, few genes were down-regulated in  $\Delta dis3L2$ , and the gene products did not show enrichment in specific pathways according to our STRING analysis (Appendix B Figure B2C).



**Table 4.2: Functional enrichments in genes >1.8-fold up-regulated in *S. pombe*  $\Delta$ *dis3L2* compared to wildtype.**

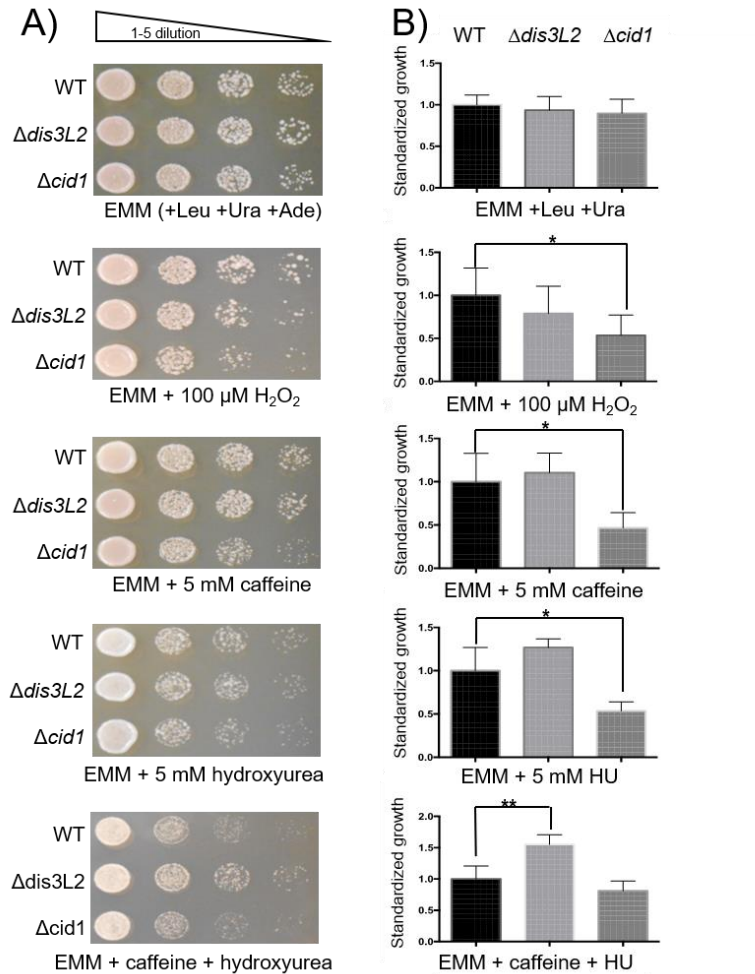
Gene ontology processes enriched with a false discovery rate <0.01 are listed.

Gene ontology biological process	observed gene count	false discovery rate	matching proteins in <i>S. pombe</i> network
protein folding	12	0.000414	<i>SPBC1711.08.1, bip1, cdc37, hsp104, hsp78, hsp90, psi1, ssa1, ssa2, ssc1, stil, trx1</i>
galactose catabolic process	4	0.0021	<i>SPBC32F12.10.1, gal1, gal10, gal7</i>
single-organism catabolic process	15	0.00848	<i>SPAC26F1.07.1, SPBC32F12.10.1, SPBC3B9.01, SPCC5E4.05c.1, bip1, cdc48, gal1, gal10, gal7, glo1, pgil, plg7, rpt1, rpt3, trx1</i>
monosaccharide catabolic process	6	0.00848	<i>SPAC26F1.07.1, SPBC32F12.10.1, gal1, gal10, gal7, pgil</i>

#### 4.3.4 Deletion of *dis3L2* confers resistance to hydroxyurea, whereas deletion of *cid1* increases sensitivity to protein misfolding stress

Since our sequencing analysis revealed major changes to the transcriptome of stress related genes in the *dis3L2* deletion strain, we assessed phenotypic effects on *S. pombe* in response to chemical stress (Figure 4.8). Cid1 was first identified as a protein involved in S-M checkpoint control and *cid1* deletion strains were found to be sensitive to caffeine stress<sup>12</sup>. For phenotypic analysis, cells were grown on EMM containing 10 mM caffeine, 10 mM hydroxyurea (HU), or a combination of 2.5 mM caffeine and 10 mM HU as described previously<sup>12</sup>. Our data shows that the *cid1* deletion strain is sensitive to caffeine and HU (Figure 4.8). Since Cid1 and Dis3L2 function in the same pathway, we next tested the impact of these chemical stressors on the *dis3L2* deletion strain. Surprisingly,  $\Delta$ *dis3L2* cells were more resistant to HU and caffeine, and a combination of caffeine and HU than wildtype cells (Figure 4.8). In addition, we tested the sensitivity of strains

bearing a deletion of *dis3L2* or *cid1* to conditions of protein misfolding stress for growth in media containing hydrogen peroxide, which causes oxidative damage. Deletion of *dis3L2* did not cause significant changes in sensitivity to hydrogen peroxide compared to wildtype (Figure 4.8). By contrast, deletion of *cid1* resulted in increased sensitivity to oxidative stress.

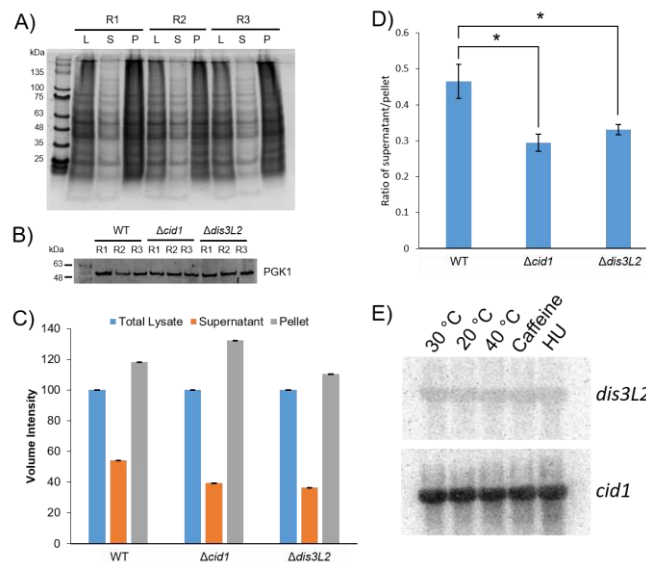


**Figure 4.8: Growth assay of *S. pombe* WT,  $\Delta dis3L2$ , and  $\Delta cid1$ .**

Cells were grown on EMM with required nutrients and with or without drug treatment at 30°C. **A)** Cells were grown overnight in EMM media with Ura, Leu, and Ade and then spotted on media plates. **B)** Quantification of standardized growth differences between wildtype and deletion strains. WT growth was normalized to 1 and compared against deletion strains. The  $p$  values were calculated using a two-tailed  $t$ -test. A  $p$  value  $< 0.05$  is indicated by an asterisk (\*) and  $< 0.01$  by (\*\*).

#### 4.3.5 Deletion of *dis3L2* and *cid1* causes the accumulation of misfolded proteins

To test whether the up-regulation of stress response genes in the deletion strains was a transcriptional response to cellular stress, we tested WT and deletion strains for the accumulation of misfolded proteins in the cell. Indeed, the overall protein abundance in the insoluble protein fraction was significantly higher in both deletion strains (Figure 4.9A, C, and D). PGK1 (Phosphoglycerate kinase 1) was blotted for in total cell lysates as a loading control (Figure 4.9B). To investigate whether *dis3L2* or *cid1* expression is up-regulated under stress conditions, we performed a Northern blot on RNAs extracted from WT cells grown under heat, cold, caffeine and hydroxyurea stress (Figure 4.9E). No changes in abundance were observed.



**Figure 4.9: Sedimentation assay of aggregated proteins in WT,  $\Delta cid1$ , and  $\Delta dis3L2$  deletion strains.**

**A)** Representative SDS gel of a sedimentation assay showing total, soluble, and insoluble protein of WT *S. pombe* cells in triplicate. **B)** Western Blot showing similar amount of PGK1 of total cell lysate of WT,  $\Delta cid1$ , and  $\Delta dis3L2$  deletion strains were used for sedimentation assays. R, replicate. **C)** Quantification of total amount of protein in the supernatant and pellet fractions of three biological replicates from the wildtype,  $\Delta cid1$ , and  $\Delta dis3L2$  strains. **D)** The ratio between the supernatant and pellet for each strain was calculated and plotted on a bar graph. The *p* values were calculated using a two-tailed *t*-test (0.03 between WT and  $\Delta cid1$ , 0.05 between WT and  $\Delta dis3L2$ ). A *p* value <0.05 is indicated by an asterisk (\*). **E)** Northern blot showing no change in expression of *cid1* or *dis3L2* RNA in response to different growth conditions.



## 4.4 Discussion

### 4.4.1 mRNA uridylation does not exclusively depend on Cid1

Uridylation-dependent RNA decay is now well established as an alternative RNA degradation pathway<sup>9</sup>. Despite growing knowledge on the biochemical and structural properties of the responsible enzymes Cid1 and Dis3L2 in *S. pombe*, little is known about the prevalence and substrate-specificity of uridylation-dependent mRNA decay. A truncated Cid1, lacking amino acids 1–31 of the N-terminal domain was previously shown to be highly processive *in vitro*<sup>15,21</sup>, yet only a few U residues are added *in vivo* (Figure 4.3 and <sup>15</sup>). By contrast, our full-length protein is significantly less processive and its uridylation activity is restricted to few residues (Figure 4.2). It is therefore possible that the N-terminal domain of Cid1 serves as an auto-inhibitory domain to prevent excessive RNA uridylation.

In the cellular context, our data shows that mRNA uridylation in *S. pombe* is not exclusive to the founder Tutase Cid1 but is likely also executed by a partially redundant Tutase. We found evidence of mRNA uridylation in a *cid1* deletion strain (Figure 4.3), indicating the activity of an alternate Tutase in this pathway. In addition, less drastic changes were observed in the transcriptome upon *cid1* deletion (Figures 4.4, 4.5 and Appendix B Figure B2). It remains to be elucidated which TENTs are responsible for RNA uridylation in the absence of Cid1. *S. pombe* encodes several TENT homologs besides Cid1, namely Cid11, Cid12, Cid13, Cid14, and Cid16. Interestingly, transcriptional levels of the TENT Cid14 are 2-fold down-regulated in the *cid1* deletion strain. Cid14 is thought to be a poly(A) polymerase and lacks the histidine residue specifying uridine over adenine specificity (Figure 4.1A). Furthermore, Cid14 has been shown to act as a poly(A) polymerase in ribosomal RNA processing<sup>39</sup>, and is thus unlikely to also act as a Tutase. Similarly, Cid11, Cid12 and Cid13 are thought to be adenylyltransferases rather than Tutases<sup>9,40</sup>. The most likely candidate for mRNA uridylation appears to be Cid16, an enzyme previously shown to act as a Tutase on small RNAs, targeting Argonaute-bound RNAs to promote their degradation<sup>41</sup>. In a recent study, spatially separated activities of the nuclear adenylyltransferase Cid14 and the cytoplasmic Tutase Cid16 were shown to regulate small RNA stability<sup>41</sup>. While Cid1



does not compensate for small RNA uridylation in the absence of Cid16<sup>41</sup>, it is possible that Cid16 uridylates mRNAs in the absence of Cid1, or generally functions as a redundant Tutase acting on mRNA.

#### 4.4.2 Mixed mRNA A/U tails

While adenylyltransferases and Tutases have previously been shown to act on the same set of RNAs on separate instances<sup>41</sup>, we show here a previously unreported combined 3' A/U modification in *S. pombe*. The observed combination of A/U addition suggests that in contrast to the uridylation and adenylation of sRNAs catalyzed by the spatially separate Cid14/Cid16, the *dak2* mRNA 5'-N<sub>x</sub>AUUAAAA-3' tail most likely occurs in the cytoplasm as a cooperative effort between uridylyl- and adenylyltransferases. The Tutases Cid1 and Cid16, as well as the adenylyltransferases Cid11 and Cid13, localize to the cytoplasm<sup>42</sup> and could interact to form this unusual 3'-end RNA motif. Whether the 5'-N<sub>x</sub>AUUAAAA-3' motif serves as a specific signal for downstream RNA processing, RNA localization, or degradation remains to be determined. We observed the 5'-N<sub>x</sub>AUUAAAA-3' motif on *dak2* mRNA, which encodes a dihydroxyacetone kinase. Whether the 5'-N<sub>x</sub>AUUAAAA-3' motif is unique to this mRNA species is speculative, as we only tested a small sample number of mRNAs. Excitingly, guanylyl- and adenylyltransferases were shown to generate mixed-tails in human cells, and mixed tails were shown to shield mRNA from rapid deadenylation<sup>43</sup>. Future research will have to explore the biological function and prevalence of combined A/U tailored transcripts in *S. pombe in vivo*.

#### 4.4.3 Uridylation-dependent RNA decay is linked to stress response and telomere maintenance

While Cid1 is not essential for the uridylation-dependent RNA decay pathway under non-stress conditions, Dis3L2 appears to represent more of a bottleneck in RNA degradation. Our data shows that the deletion of *dis3L2* leads to an accumulation of

transcripts predominantly in protein folding and degradation pathways, as well as sugar catabolic processes. While Cid1 does not contain specific RNA recognition motifs, such as Zinc fingers, its crystal structure revealed a positively charged surface area thought to facilitate general, sequence-independent RNA binding<sup>19,20</sup>. It has been suggested that RNA binding proteins interact with Cid1 to direct the TENT to substrate RNAs, but these proteins remain to be identified<sup>21</sup>.

As previously described for chemical stress, the transcriptome of *S. pombe* undergoes global changes in response to stress conditions<sup>44</sup> (Appendix B Tables B2-B6). We found most significant changes in the transcriptome in genes related to protein folding, such as heat shock protein genes *hsp90* and *hsp70*, co-chaperones *cdc37* and *wos2*, protein remodeling factor *hsp104*, and chaperone activators *aha1* and *sti1* with an effect change between 2- and 10-fold in the *dis3L2* deletion strain (Appendix B Tables B2, B5, B6). Changes to the  $\Delta$ *cid1* transcriptome generally followed the same trend, albeit to a lesser extent, indicating that a partially redundant Tase reduces the effect of the *cid1* deletion (Appendix B Tables B2, B3, B4). Furthermore, protein catabolic processes and protein degradation pathways, including AAA-type ATPase *cdc48*, and 19S proteasome regulatory subunits *rpt1*, *rpt3* and *rpt6*, and the ubiquitin C-terminal hydrolase *ubp15*, were up-regulated between 2- and 5.4-fold in  $\Delta$ *dis3L2*, and up to 2.2-fold in  $\Delta$ *cid1*. Other accumulated transcripts include RNAs of genes from a variety of metabolic pathways, transcriptional regulators, cell cycle, and the cytoskeleton (Figure 4.7).

Galactose metabolism genes *gal1*, *gal7*, and *gal10* are required for using galactose as a carbon or energy source and are usually repressed in wildtype fission yeast in the absence of galactose<sup>45</sup>. We found these genes significantly up-regulated in the  $\Delta$ *dis3L2* strain with effect changes between 3.5 and 8.5 (Appendix B Tables B2, B5, B6). Expression of *gal* genes is regulated by telomeric silencing<sup>45</sup>. We observed an increased expression of genes involved in telomere organization in the *dis3L2* deletion strain, namely DNA replication factor A subunit Ssb1 (effect change = 6.5), and single-stranded telomeric binding protein Tgc1 (effect change = 3.6). Overexpression of proteins involved in telomere organization may be compensatory effects in an effort to counteract a disruption of telomere organization, which is evident from the de-repression

of *gal* genes. Cid1 was initially characterized as a protein involved in S-M checkpoint control, and it may well be that disruption of uridylation-dependent RNA decay interferes with telomere maintenance and repair.

Previous studies in frog oocytes show that uridylation-dependent decay is crucial for clearance of the maternal transcriptome, indicating that this decay pathway plays a role in the degradation of transcripts that are no longer required<sup>46</sup>. Similarly, in humans, Dis3L2 depletion is associated with the accumulation of damaged RNA transcripts<sup>6</sup>. Furthermore, Dis3L2 depletion inhibits global apoptotic mRNA decay and cell death<sup>47</sup>. These and other studies indicate that uridylation-dependent RNA decay is part of a global mRNA surveillance, aiding in the clearance of unneeded or damaged RNAs. Our data show that perturbation in uridylation-dependent RNA decay elicits a stress response as evidenced by increased abundance of transcripts enriched in protein folding and degradation pathways. It is possible that some of the other transcriptional changes beyond stress response result in a higher expression rate of those genes. We hypothesize that excess protein production may overwhelm the cellular protein quality control, leading to the accumulation of unfolded proteins and consequently elicit the unfolded protein response. In addition, since the TUT/Dis3L2 pathway in humans functions in damaged transcript decay<sup>6</sup>, these transcripts may also be accumulating in *S. pombe*, but not be detected as a significant change in abundance rates. Stress response genes are likely up-regulated as a response to accumulated proteins in the deletion strains (Figure 4.9).

#### 4.4.4 Dis3L2 depletion increases resistance to hydroxyurea-induced stress

Taking into account that both Cid1 and Dis3L2 are dispensable under normal growth conditions<sup>7,9,12</sup>, it is likely that uridylation-dependent RNA decay targets damaged or incomplete RNA transcripts as part of a stress response, similar to the Cid14/16 small RNA surveillance pathway<sup>41</sup>. Cid1 was first described as a protein involved in the S-M cell cycle checkpoint<sup>12</sup>. A *cid1* deletion strain displays a growth retarded phenotype under stress conditions when exposed to caffeine, HU, or a combination of caffeine and

HU (Figure 4.8) and overexpression of *Cid1* increases resistance to this stress-inducing combination<sup>12,48</sup>. Our data shows that a *dis3L2* deletion on the other hand increases viability when exposed to HU and caffeine (Figure 4.8). Both HU and caffeine are known for their interference with the cell cycle<sup>49</sup>. HU inhibits the enzyme ribonucleotide reductase, which is essential for DNA synthesis, and its depletion impairs DNA replication and subsequently arrests cells in S phase<sup>50</sup>. Interestingly, the RNAs up-regulated in the *dis3L2* deletion strain included ribonucleotide reductase small subunit *suc22* (1.9-fold) and ribonucleoside reductase large subunit *cdc22* (3.4-fold), which were previously shown to be up-regulated in response to HU<sup>51</sup>. The increased expression of ribonucleotide reductase subunits, in combination with an already activated stress response may give *Dis3L2* depleted cells the growth advantage over wildtype cells. While overexpression or depletion of enzymes in the uridylation-dependent RNA decay pathway can alter cell viability, we found no evidence that expression of *cid1* or *dis3L2* is altered at the transcriptional level (Figure 4.9E). Both *cid1* and *dis3L2* show similar abundance independent of growth temperature or chemical stress (Figure 4.8). However, protein production may instead serve as the point of control on the translational level, or enzyme activity could be modulated by post-translational modification, as described for other nucleotidyltransferases<sup>52,53</sup>. It is plausible that enzymes involved in RNA uridylation and decay are constitutively active and degrade damaged RNA transcripts on demand.

We also tested the effect of the deletion of *dis3L2* or *cid1* on protein quality control. When grown on hydrogen peroxide to induce protein misfolding, the *cid1* deletion, but not the *dis3L2* deletion strain showed a significant increase in sensitivity (Figure 4.8). It is possible that the up-regulation of the cellular stress responses (e.g. the heat shock response) that we observed in the *dis3L2* deletion exerts a protective function compared to the *cid1* deletion strain. We further found that both deletion strains accumulated higher proportions of misfolded and insoluble proteins compared to wildtype (Figure 4.9). These results indicate that the defects in RNA processing in these strains leads to an accumulation of misfolded protein, which in turn appears to induce the expression of many protein quality control genes possibly by activation of the heat shock response.

We conclude that uridylation-dependent RNA decay is part of an RNA surveillance system, and RNA transcripts are not efficiently disposed of in the absence of Dis3L2 or to a lesser extent Cid1. Translation of these potentially damaged or unwanted RNA transcripts leads to the accumulation of misfolded proteins, eliciting the cellular stress response and the increased expression of chaperones and enzymes involved in protein degradation.

## 4.5 Acknowledgements

This work was supported by grants from the J.P. Bickell Foundation and Natural Sciences and Engineering Research Council of Canada (RGPIN 04776-2014) to I.U.H., Ontario Graduate Scholarship and NSERC Canada Graduate Scholarship to C.Z.C., an Alexander Graham Bell Canada Graduate Scholarship-Doctoral from the Natural Sciences and Engineering Research Council of Canada to C.Z.C., and a NSERC Canada Graduate Scholarship to M.J.E. Funding for open access charge: The Natural Sciences and Engineering Research Council of Canada. We thank Nileeka Balasuriya and Susanna George for their advice, and Patrick O'Donoghue and Gregory Gloor for critical discussions.

## 4.6 References

1. Cheadle, C. *et al.* Stability regulation of mRNA and the control of gene expression. *Ann. N. Y. Acad. Sci.* **1058**, 196–204 (2005).
2. Norbury, C. J. Cytoplasmic RNA: a case of the tail wagging the dog. *Nat. Rev. Mol. Cell Biol.* **14**, 643–653 (2013).
3. Martin, G. & Keller, W. RNA-specific ribonucleotidyl transferases. *RNA N. Y. N* **13**, 1834–1849 (2007).
4. Lin, C.-J. *et al.* Characterization of a TUTase/RNase complex required for *Drosophila* gametogenesis. *RNA N. Y. N* **23**, 284–296 (2017).
5. Norbury, C. J. 3' Uridylation and the regulation of RNA function in the cytoplasm. *Biochem. Soc. Trans.* **38**, 1150–1153 (2010).
6. Ustianenko, D. *et al.* TUT-DIS3L2 is a mammalian surveillance pathway for aberrant structured non-coding RNAs. *EMBO J.* **35**, 2179–2191 (2016).

7. Malecki, M. *et al.* The exoribonuclease Dis3L2 defines a novel eukaryotic RNA degradation pathway. *EMBO J.* **32**, 1842–1854 (2013).
8. Pirouz, M., Du, P., Munafò, M. & Gregory, R. I. Dis3L2-Mediated Decay Is a Quality Control Pathway for Noncoding RNAs. *Cell Rep.* **16**, 1861–1873 (2016).
9. 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* **1861**, 2971–2979 (2017).
10. Kwak, J. E. & Wickens, M. A family of poly(U) polymerases. *RNA* **13**, 860–867 (2007).
11. 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).
12. Wang, S.-W., Toda, T., MacCallum, R., Harris, A. L. & Norbury, C. Cid1, a Fission Yeast Protein Required for S-M Checkpoint Control when DNA Polymerase  $\delta$  or  $\epsilon$  Is Inactivated. *Mol. Cell. Biol.* **20**, 3234 (2000).
13. 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. U. S. A.* **99**, 12079–12084 (2002).
14. Rissland, O. S. & Norbury, C. J. The Cid1 poly(U) polymerase. *Biochim. Biophys. Acta* **1779**, 286–294 (2008).
15. Rissland, O. S., Mikulasova, A. & Norbury, C. J. Efficient RNA polyuridylation by noncanonical poly(A) polymerases. *Mol. Cell. Biol.* **27**, 3612–3624 (2007).
16. Munoz-Tello, P., Gabus, C. & Thore, S. A critical switch in the enzymatic properties of the Cid1 protein deciphered from its product-bound crystal structure. *Nucleic Acids Res.* **42**, 3372–3380 (2014).
17. Stevenson, A. L. & Norbury, C. J. The Cid1 family of non-canonical poly(A) polymerases. *Yeast Chichester Engl.* **23**, 991–1000 (2006).
18. Chung, C. Z., Jo, D. H. S. & Heinemann, I. U. Nucleotide specificity of the human terminal nucleotidyltransferase Gld2 (TUT2). *RNA N. Y. N* **22**, 1239–1249 (2016).
19. Munoz-Tello, P., Gabus, C. & Thore, S. Functional implications from the Cid1 poly(U) polymerase crystal structure. *Struct. Lond. Engl. 1993* **20**, 977–986 (2012).
20. 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).
21. Yates, L. A. *et al.* Structural plasticity of Cid1 provides a basis for its distributive RNA terminal uridylyl transferase activity. *Nucleic Acids Res.* **43**, 2968–2979 (2015).
22. Yates, L. A. *et al.* Improved crystallization and diffraction of caffeine-induced death suppressor protein 1 (Cid1). *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **71**, 346–353 (2015).
23. Aphasizhev, R. RNA uridylyltransferases. *Cell. Mol. Life Sci. CMLS* **62**, 2194–2203 (2005).
24. 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).
25. Lim, J. *et al.* Uridylation by TUT4 and TUT7 Marks mRNA for Degradation. *Cell* **159**, 1365–1376 (2014).

26. 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 N. Y. N* **18**, 1875–1885 (2012).
27. Gallouzi, I. E. & Wilusz, J. A DIStinctively novel exoribonuclease that really likes U. *EMBO J.* **32**, 1799–1801 (2013).
28. Lv, H. *et al.* Structural analysis of Dis3l2, an exosome-independent exonuclease from *Schizosaccharomyces pombe*. *Acta Crystallogr. D Biol. Crystallogr.* **71**, 1284–1294 (2015).
29. Robinson, S. R., Oliver, A. W., Chevassut, T. J. & Newbury, S. F. The 3' to 5' Exoribonuclease DIS3: From Structure and Mechanisms to Biological Functions and Role in Human Disease. *Biomolecules* **5**, 1515–1539 (2015).
30. Towler, B. P., Jones, C. I., Harper, K. L., Waldron, J. A. & Newbury, S. F. A novel role for the 3'-5' exoribonuclease Dis3L2 in controlling cell proliferation and tissue growth. *RNA Biol.* **13**, 1286–1299 (2016).
31. Lubas, M. *et al.* Exonuclease hDIS3L2 specifies an exosome-independent 3'-5' degradation pathway of human cytoplasmic mRNA. *EMBO J.* **32**, 1855–1868 (2013).
32. Faehnle, C. R., Walleshauser, J. & Joshua-Tor, L. Mechanism of Dis3l2 substrate recognition in the Lin28-let-7 pathway. *Nature* **514**, 252–256 (2014).
33. Chang, H.-M., Triboulet, R., Thornton, J. E. & Gregory, R. I. A role for the Perlman syndrome exonuclease Dis3l2 in the Lin28-let-7 pathway. *Nature* **497**, 244–248 (2013).
34. Ellis, M. J., Trussler, R. S. & Haniford, D. B. Hfq binds directly to the ribosome-binding site of IS10 transposase mRNA to inhibit translation. *Mol. Microbiol.* **96**, 633–650 (2015).
35. Ellis, M. J., Trussler, R. S., Charles, O. & Haniford, D. B. A transposon-derived small RNA regulates gene expression in *Salmonella Typhimurium*. *Nucleic Acids Res.* **45**, 5470–5486 (2017).
36. Shiber, A., Breuer, W., Brandeis, M. & Ravid, T. Ubiquitin conjugation triggers misfolded protein sequestration into quality control foci when Hsp70 chaperone levels are limiting. *Mol. Biol. Cell* **24**, 2076–2087 (2013).
37. Fernandes, A. D., Macklaim, J. M., Linn, T. G., Reid, G. & Gloor, G. B. ANOVA-like differential expression (ALDEx) analysis for mixed population RNA-Seq. *PloS One* **8**, e67019 (2013).
38. Szklarczyk, D. *et al.* The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res.* **39**, D561–568 (2011).
39. Keller, C., Woolcock, K., Hess, D. & Bühler, M. Proteomic and functional analysis of the noncanonical poly(A) polymerase Cid14. *RNA N. Y. N* **16**, 1124–1129 (2010).
40. Saitoh, S. *et al.* Cid13 is a cytoplasmic poly(A) polymerase that regulates ribonucleotide reductase mRNA. *Cell* **109**, 563–573 (2002).
41. Pisacane, P. & Halic, M. Tailing and degradation of Argonaute-bound small RNAs protect the genome from uncontrolled RNAi. *Nat. Commun.* **8**, 15332 (2017).
42. Matsuyama, A. *et al.* ORFeome cloning and global analysis of protein localization in the fission yeast *Schizosaccharomyces pombe*. *Nat. Biotechnol.* **24**, 841–847 (2006).
43. Lim, J. *et al.* Mixed tailing by TENT4A and TENT4B shields mRNA from rapid deadenylation. *Science* **361**, 701–704 (2018).

44. Biswas, P. & Ghosh, S. Global transcriptomic profiling of *Schizosaccharomyces pombe* in response to nitrosative stress. *Gene* **558**, 241–253 (2015).
45. Matsuzawa, T. *et al.* New insights into galactose metabolism by *Schizosaccharomyces pombe*: isolation and characterization of a galactose-assimilating mutant. *J. Biosci. Bioeng.* **111**, 158–166 (2011).
46. Chang, H. *et al.* Terminal Uridyltransferases Execute Programmed Clearance of Maternal Transcriptome in Vertebrate Embryos. *Mol. Cell* **70**, 72–82.e7 (2018).
47. Thomas, M. P. *et al.* Apoptosis Triggers Specific, Rapid, and Global mRNA Decay with 3' Uridylated Intermediates Degraded by DIS3L2. *Cell Rep.* **11**, 1079–1089 (2015).
48. Wang, S. W., Norbury, C., Harris, A. L. & Toda, T. Caffeine can override the S-M checkpoint in fission yeast. *J. Cell Sci.* **112** ( Pt 6), 927–937 (1999).
49. Moser, B. A., Brondello, J. M., Baber-Furnari, B. & Russell, P. Mechanism of caffeine-induced checkpoint override in fission yeast. *Mol. Cell. Biol.* **20**, 4288–4294 (2000).
50. Koç, A., Wheeler, L. J., Mathews, C. K. & Merrill, G. F. Hydroxyurea arrests DNA replication by a mechanism that preserves basal dNTP pools. *J. Biol. Chem.* **279**, 223–230 (2004).
51. Fernandez Sarabia, M. J., McInerney, C., Harris, P., Gordon, C. & Fantes, P. The cell cycle genes *cdc22+* and *suc22+* of the fission yeast *Schizosaccharomyces pombe* encode the large and small subunits of ribonucleotide reductase. *Mol. Gen. Genet. MGG* **238**, 241–251 (1993).
52. Mohan, N., AP, S., Francis, N., Anderson, R. & Laishram, R. S. Phosphorylation regulates the Star-PAP-PIPK1 $\alpha$  interaction and directs specificity toward mRNA targets. *Nucleic Acids Res.* **43**, 7005–7020 (2015).
53. Phatnani, H. P. & Greenleaf, A. L. Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev.* **20**, 2922–2936 (2006).



## Chapter 5

### 5 Summary and Perspectives

RNAs are a fundamental part of cellular homeostasis and are regulated throughout their life cycle<sup>1-11</sup>. The addition of non-templated 3'-end nucleotides plays a key role in RNA regulation<sup>12</sup>. The enzymes that these reactions are known as terminal RNA nucleotidyltransferases (TENTs) and are divided into two functional groups, adenylyltransferases and uridylyltransferases, based on their nucleotide preference.

Adenylation is associated with stabilization while uridylation leads to silencing and degradation of RNAs, with the exception of a role for uridylation in Group II pre-miRNA maturation<sup>7,12</sup>. The first uridylyltransferase discovered was Cid1 in *Schizosaccharomyces pombe*. Cid1 was initially assumed to be an adenylyltransferase shown to play a role in the S-M cell cycle checkpoint<sup>13</sup>. The unexpected uridylation activity was later identified through *in vitro* assays with recombinant Cid1 and *in vivo* assays in *S. pombe*<sup>14-16</sup>. This uridylation activity is thought to be part of a uridylation-dependent RNA decay pathway where polyuridylated RNAs are recognized and degraded by the U-specific 3'-5' exonuclease Dis3L2<sup>17</sup>. The discovery of Cid1 as an uridylyltransferase sparked new investigations resulting in many adenylyltransferases to be re-classified as uridylyltransferases<sup>10-12,18-20</sup>. One human homolog of Cid1 is Gld2, which has been associated with both adenylation and uridylation in different reports<sup>7,9,21,22</sup>.

Both Cid1 and Gld2 encode two domains necessary for catalytic activity but lack recognizable RNA binding domains<sup>23</sup>. Therefore, their catalytic activity and RNA specificity are thought to be regulated by interacting proteins. In humans, some of those interacting proteins were identified and characterized, but remain uncharacterized in *S. pombe*<sup>22,24,25</sup>. The fact that Cid1 and Gld2 likely require accessory proteins to determine their substrate specificity makes it challenging to identify what RNAs are targeted by these enzymes. In addition, their similar domain structure and high amino acid conservation, but seemingly different nucleotide preferences, suggest that uridylyltransferases and adenylyltransferases are closely related despite having opposing

roles. The work in this thesis provides insights into the regulation and cellular impact of TENT-catalyzed nucleotide additions to RNAs.

## 5.1 The evolution of adenylyl- and uridylyltransferase nucleotide specificities

In previous studies, Gld2 has been reported to both adenylate and uridylate the 3'-end of RNAs<sup>7,9,21,22,25,26</sup>. As adenylation (stabilizing) and uridylation (silencing and/or degradation) have opposing effects on RNA stability<sup>12</sup>, the cellular roles of Gld2 were not clear. Through extensive enzyme kinetic analysis, I showed that Gld2 is a true adenylyltransferase with an 83-fold preference for ATP over UTP. As the cellular concentrations of ATP vastly exceeds that of UTP<sup>27</sup>, it is most likely that Gld2 acts as an adenylyltransferase *in vivo* to stabilize RNAs.

The high amino acid conservation and structural similarity between adenylyltransferases and uridylyltransferases suggests these enzymes are homologous and evolved from a common ancestor. A detailed phylogenetic analysis showed that uridylyl- and adenylyltransferases do not separate into two distinct groups according to nucleotide specificity. Rather the phylogeny (see Figure 2.6 in Chapter 2) showed that nucleotide specificity evolved multiple times, suggesting a simple mechanism governing ATP *versus* UTP selectivity in the TENT family. My experiments revealed that UTP specificity is accomplished through insertion of a crucial histidine residue in the active site. This histidine is absent in adenylyltransferases and present in uridylyltransferases, where it sterically blocks the larger ATP from the active site. Insertion of the histidine into the coding sequence of Gld2 did indeed switch the nucleotide preference from ATP to UTP. This simple insertion/deletion of a histidine residue could explain how uridylyl- and adenylyltransferases appeared multiple times throughout evolution of the TENT family.

## 5.2 Regulation of Gld2 activity by post-translational phosphorylation

The activity and RNA specificity of Gld2 is regulated through interactions with other proteins such as the Hepatitis C (HepC) core protein<sup>28</sup> and the RNA binding proteins QKI-7<sup>24</sup> and CPEB<sup>25,29,30</sup>. No other modes of regulation, such as regulation of protein stability, have been described in the literature. Large scale proteomic studies using mass spectrometry revealed five phosphorylated serine residues in the predicted disordered N-terminal domain of Gld2<sup>31–33</sup>, suggesting the possibility of Gld2 regulation through phosphorylation. I showed that phosphomimetic substitutions at each of the five serine positions lead to changes in Gld2 activity and RNA binding that were dependent on the position as well as the RNA substrate. Two sites of interest were identified where one site (S62) increased enzymatic activity by ~5-fold while another site (S116) severely inhibited Gld2's adenylation activity up to 111-fold. When both sites were combined, the inhibiting ability of S116E overwhelmed the activating function of S62E. This was the first evidence that phosphorylation at these sites may regulate Gld2 activity.

Using kinase activity assays and mass spectrometry, I identified the first known kinases of Gld2: protein kinase A (PKA) and protein kinase B (Akt1). Both kinases were found to specifically and exclusively phosphorylate Gld2 at S116 and reduced Gld2 activity to an even greater extent than the phosphomimetic. The data suggest that these kinases may play a role in inactivating Gld2. Gld2 promotes the stability and maturation of the miRNAs miR-122 and let-7a, respectively<sup>7,9,21</sup>. Both miRNAs function as tumor suppressors and reduced levels of these miRNAs lead to dysregulation of their target genes, some of which are involved in cell growth, metastasis, and apoptosis<sup>34,35</sup>. As Akt1 and PKA are commonly over-activated in many cancers<sup>36,37</sup>, their ability to inhibit Gld2 may promote tumorigenesis by decreasing the levels of miRNA tumor suppressors. Thus, the data reveal a possible novel link between oncogenic kinases signaling and miRNA regulation.

### 5.3 The role of the Cid1/Dis3L2 pathway in global RNA decay

The Cid1/Dis3L2 RNA decay pathway in *S. pombe* plays a major role in mRNA homeostasis<sup>17,19,38–42</sup>. However, the contribution of the Cid1/Dis3L2 pathway to total RNA decay is unknown. Deletion of *cid1* did not display any major changes in the uridylation status compared to wildtype *S. pombe*. However, an interesting uridylation/adenylation pattern (5'-N<sub>x</sub>AUUAAAA-3') was found on the 3'-ends of *dak2*, a protein coding RNA. Such mixed RNA ends were previously shown to alter RNA stability in human cells<sup>43</sup>. The role of the mixed RNA tail in *S. pombe* remains to be investigated. Deep sequencing revealed that deleting *cid1* or *dis3L2* elicited similar changes to the transcriptome. However, the changes in gene expression were much greater in the *dis3L2* deletion strain. Genes involved in protein folding and degradation were up-regulated with the greatest change, indicating an up-regulated stress response. The transcriptome of both deletion strains showed similar accumulation of transcripts related to stress response, yet larger and more significant changes were observed in the *dis3L2* deletion strain. The greater transcriptional response as well as the presence of somewhat redundant uridylyltransferase activity in the *cid1* deletion strain indicates that Dis3L2 represents a bottleneck in the RNA decay pathway.

The data suggests that the increase of mRNA transcripts of stress response genes could be due to the accumulation of misfolded proteins. As the Cid1/Dis3L2 RNA decay pathway was compromised in the deletion strains, we expect accumulation of RNA transcripts that would usually be decayed in the Cid1/Dis3L2 pathway. These transcripts could be translated into proteins, resulting in a dysregulated proteome, which may then overwhelm the protein control machinery. The accumulation of misfolded proteins, which we observed experimentally, may trigger the cellular stress response. Our data suggests that the uridylation-dependent RNA decay pathway mediated by Cid1 and Dis3L2 plays an important role in maintaining mRNA and proteome homeostasis.

## 5.4 Future Directions and Perspectives

### 5.4.1 The role of post-translational modifications in regulating Gld2

Post-translational modifications (PTMs) are an efficient means for the cell to regulate proteins<sup>44</sup>. In Chapter 3, phosphomimetic substitutions at specific serine residues in the predicted disordered N-terminal domain of Gld2 demonstrated that the activity and RNA binding of Gld2 is affected and that two different oncogenic kinases were able to phosphorylate and inhibit Gld2 activity through one site (S116). Confirming these results in human cell lines and animal models would be the next steps in connecting the kinases to miRNA regulation. As PKA and Akt1 are hyperactive in many cancers<sup>36,37</sup>, mass spectrometry of tumor samples that show hyperactivity of either kinase would be a valuable tool to determine if Gld2 pS116 is present in those samples. In addition, determining additional upstream kinases for Gld2 would reveal different pathways that regulate Gld2 activity.

The majority of PTMs were shown to be in the predicted disordered N-terminal domain of Gld2<sup>45</sup>. Disordered regions are known as protein docking areas due to their ability to adopt different conformations<sup>46</sup>. In fact, the RNA binding protein QKI-7 was shown to bind to the predicted disordered N-terminal domain of Gld2<sup>24</sup>. As Gld2 lacks known RNA binding motifs, it seems the enzyme depends on RNA binding proteins to promote recognition with specific RNA substrates and regulate Gld2 activity<sup>12,22–25</sup>. It is possible that these interacting proteins bind the N-terminal domain of Gld2 and PTMs in this region may regulate Gld2 binding to specific proteins at specific times.

Although PTMs may affect the binding of interacting proteins, phosphorylation of five serine sites were shown to directly affect Gld2 activity and RNA binding in Chapter 3. It would be interesting to extend the study to include different RNAs such as a wider range of miRNA forms, poly(A) tail mimics of different lengths, and poly(A) tail mimics that include different sequence motifs. In addition, testing the effect that serine phosphorylation may have on nucleotide specificity may be worthwhile as Gld2 can accept and use ATP and UTP<sup>7,9,21,47</sup>. As Gld2 is currently viewed as an enzyme involved in promoting RNA stability through adenylation, this would significantly change the

biological function of Gld2 as a RNA regulating enzyme. The ability to switch between either nucleotide based on the phosphorylation status would allow Gld2 to take part in stabilizing (adenylation) and silencing/degrading (uridylation) RNAs, two opposing roles thought to be carried out by different enzymes.

The Gld2 N-terminal domain constitutes approximately one third of the protein and it is possible that the domain can act as a regulatory element for the two catalytic domains. This was observed for Cid1, where the full-length enzyme was less processive than an enzyme lacking the first 31 amino acids<sup>15,48</sup>. I have unpublished data revealing that Gld2 lacking the N-terminal domain is approximately 20-fold more active than wildtype Gld2. Besides the five serine positions that were studied, other phosphorylation sites throughout Gld2 and modifications such as methylation and ubiquitination were also reported in mass proteomic studies of various samples<sup>31,32,49–53</sup>. Future studies will probe the function of these other modifications.

In a disease context, determining if there are different PTMs on Gld2 in diseases such as cancer or cardiovascular diseases would be of great interest as a diagnostic tool. The HepC core protein has been shown to inhibit Gld2 activity without affecting protein expression levels<sup>28</sup>, leading to decreased miR-122 levels. The reported PTMs of Gld2<sup>31,32,45,49–53</sup> were identified in a variety of tumors, tissues, and immortalized human cells, making it difficult to determine what modifications are a result of healthy conditions or cellular dysregulation. As miR-122 is a known substrate of Gld2<sup>9,21</sup> and miR-122 is highly expressed in the liver<sup>54</sup>, immunoprecipitation and mass spectrometry of Gld2 from healthy livers of animal models such as mice would reveal PTMs on Gld2 in healthy tissue. These can be compared to PTMs reported in the literature<sup>31,32,45,49–53</sup> or disease mouse models. Immunoprecipitation and mass spectrometry of Gld2 from a variety of liver cells lines, such as the HepG2 liver hepatocellular carcinoma cell line and the PLC/PRF/5 hepatoma cell line, would also give an overview of PTMs that are associated with specific conditions.

#### 5.4.2 Regulating miR-122 through Gld2 in health and disease

The miRNA miR-122 is the most abundant miRNA in the liver, comprising 70% of the total miRNA population in human livers<sup>54</sup>. It is a naturally unstable miRNA that requires the post-transcriptional addition of a 3'-end adenine residue to enhance its stability<sup>9,21</sup>. Gld2 was identified as the nucleotidyltransferase responsible for stabilizing miR-122<sup>9,21</sup>. Not only does miR-122 play important roles in liver development and homeostasis, but also in viral infections and liver diseases<sup>34</sup>. The levels of miR-122 in viral infections vary. HepC infected livers have moderate levels of miR-122 as the virus requires miR-122 for replication. On the other hand, Hepatitis B (HepB) infected livers have non-detectable miR-122 levels. HepB actively reduces miR-122 as miR-122 was shown to exert an inhibitory effect on the replicative ability of the virus<sup>28,34,55</sup>. Both HepC and HepB regulate miR-122 stability. The HepC core protein binds to Gld2 and inhibits its activity<sup>28</sup> while the HepB virus X protein (HBx) impairs Gld2 gene expression<sup>55</sup>. Although miR-122 has been a target for HepC therapeutic intervention, patients require constant dosages to reduce miR-122 levels due to the instability of the therapeutic RNAs<sup>56-58</sup>. Thus, inhibiting Gld2 with a small molecule is a potentially more desirable route to deplete miR-122 levels.

Investigating how Gld2 is regulated during viral infection or in cancers may reveal specific post-translational modifications that are not present in healthy cells. The resulting data would enable downstream development of drugs, such as antibodies, against those different modifications or protein conformations. However, inhibiting Gld2 to reduce miR-122 levels would be counterproductive in treating diseases such as HepB that require therapies to increase miR-122 levels. Therefore, drugs that could lock Gld2 in a catalytically-active conformation or inhibitors of kinases that inactivate Gld2, such as Akt1 and PKA, would be preferred in this situation.

#### 5.4.3 Identifying new cellular roles of Gld2

In studies, *C. elegans* Gld2 has been shown to play an important role in development<sup>26,59,60</sup>. Here, Gld2 is involved in extending the poly(A) tails of genes

associated with meiosis to increase protein expression and promote entry into meiosis<sup>26,59,60</sup>. In humans, Gld2 plays a role in controlling mRNA stability, in addition to miRNA maturation and stability<sup>7,9,12,21,61</sup>. However, whether Gld2 has the same function in different tissues and if Gld2 regulation is tissue-specific remains to be answered.

As Gld2 lacks recognizable RNA binding domains, studies have shown that it interacts with RNA binding proteins to recognize specific substrates<sup>22,24,25</sup>. In addition, these RNA binding proteins may also modulate Gld2 processivity by allowing Gld2 to stay in contact with the RNA longer<sup>22,24,25</sup>. Thus, Gld2 could be regulated through RNA binding proteins that are differentially expressed in different cell types or tissues. As cell lines are commonly used to study proteins in a cellular context, co-immunoprecipitation assays can be used to identify interacting proteins that can then be verified through biochemical approaches and further *in vivo* experiments such as protein expression knockdowns with siRNAs and co-localization studies.

Although the studies in Chapter 3 showed that phosphomimetics at specific serine residues on Gld2 changed the activity and RNA binding, the experiments were performed *in vitro* and further studies in cells and ultimately animal models should be pursued. As Gld2 has been shown to stabilize the liver-specific miRNA miR-122<sup>9,21</sup>, liver cell lines are the standard used by the field to study miR-122. The phosphomimetic variants and a non-phosphorylatable variant as a control could be expressed through transient transfections and their activity can be measured through miR-122 levels either indirectly using a GFP reporter construct developed in the Heinemann lab<sup>62</sup> or directly using miRNA RT-qPCR. Since we found glutamate substitutions provided an accurate mimic of phosphorylation at Ser116, a systems level analysis of the impact of phosphomimetic variants on mRNA levels can also be studied through deep sequencing.

#### 5.4.4 Impact of Cid1 and Dis3L2 on *Schizosaccharomyces pombe* RNA metabolism

Deep sequencing of *S. pombe cid1* and *dis3L2* deletion strains revealed changes across the transcriptome. Although both deletion strains showed similar changes in RNA



expression, the *dis3L2* deletion strain showed gene expression changes of far greater magnitude. This is potentially due to a partially redundant uridylyltransferase such as Cid16 in *S. pombe* that can take on Cid1's role. Deep sequencing of a *cid1/dis3L2* double deletion strain would provide additional information on what genes are specifically affected through the Cid1/Dis3L2 decay pathway. Other deletion strains to study would be a *cid16* single deletion strain and double deletion strains of *cid1/cid16* and *cid16/dis3L2*.

Select genes from different pathways that were differentially expressed were further verified in Chapter 4 but many others involved in pathways such as telomere and mitochondria organization were identified in the analysis. Genes that were affected in both deletion strains, such as ubiquitin-protein ligase E3 Dbl5 and acetyl-CoA C-acetyltransferase Erg10, would be a good starting point for further studies. Chapter 4 discussed how the *dis3L2* deletion disrupted telomere organization which led to an up-regulation in galactose metabolism genes and telomere organization genes. As many genes were up- or down-regulated, it is possible that not all of them are a direct result from the *cid1* or *dis3L2* deletion and are a downstream effect of other dysregulated genes. A possible experiment to determine what RNAs Cid1 uridylate would be to crosslink Cid1 with interacting RNA in *S. pombe* cells, pull-down Cid1 through immunoprecipitation, and identify the RNAs through RNA sequencing.

Terminal nucleotidyltransferases that exhibit different nucleotide preferences have been reported to act on the same RNA substrates at different times<sup>63</sup>. In *S. pombe*, a unique pattern was found at the end of the *dak2* mRNA tail where 2 uridine residues were added before the terminal four adenine residues. Such U/A mixed 3'-terminal RNA tails were also observed in human cells where uridine residues were found on short poly(A) tails (<25 nucleotides) and guanine residues on long tails (>40 nucleotides)<sup>64</sup>. Polyuridylation at the 3'-end of mRNAs is commonly associated with degradation<sup>10,18</sup> but the addition of single guanine residues within the poly(A) tail was shown to prevent rapid deadenylation<sup>43</sup>. As the U/A pattern was observed on short and long poly(A) tails and consists of only 2 uridine residues, the poly(A) tail interspaced with two uridine residues potentially have the same protective function in *S. pombe* as the single guanine residues

do in human cells<sup>43</sup>. This can be studied by determining the half-lives of RNAs with and without the mixed tails transfected into human cells. As the guanylation was found on multiple mRNA tails<sup>64</sup>, determining if the uridine/adenine pattern also exists on other *S. pombe* mRNA tails would indicate that mixed tails is an evolutionarily conserved mechanism to prevent rapid deadenylation of RNAs.

## 5.5 Conclusion

RNAs play an essential role in many cellular processes and RNA homeostasis is thus highly regulated. The discovery of RNA regulation through non-templated 3'-end nucleotide additions by terminal nucleotidyltransferases added an additional layer of complexity to the network regulating RNA metabolism. The work presented in this thesis aimed to elucidate the impact of these nucleotide additions on RNA homeostasis and the regulatory mechanisms controlling terminal nucleotidyltransferases. The included data broadens our understanding of the enzymes and pathways regulating cellular RNAs.

## 5.6 References

1. Cowling, V. H. Regulation of mRNA cap methylation. *Biochem. J.* **425**, 295–302 (2010).
2. Proudfoot, N. J. Ending the message: poly(A) signals then and now. *Genes Dev.* **25**, 1770–1782 (2011).
3. Colgan, D. F. & Manley, J. L. Mechanism and regulation of mRNA polyadenylation. *Genes Dev.* **11**, 2755–2766 (1997).
4. Mandel, C. R., Bai, Y. & Tong, L. Protein factors in pre-mRNA 3'-end processing. *Cell. Mol. Life Sci. CMLS* **65**, 1099–1122 (2008).
5. Ameres, S. L. & Zamore, P. D. Diversifying microRNA sequence and function. *Nat. Rev. Mol. Cell Biol.* **14**, 475–488 (2013).
6. Iwakawa, H.-O. & Tomari, Y. The Functions of MicroRNAs: mRNA Decay and Translational Repression. *Trends Cell Biol.* **25**, 651–665 (2015).
7. 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).
8. Heo, I. *et al.* TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. *Cell* **138**, 696–708 (2009).
9. 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).

10. Schmidt, M.-J., West, S. & Norbury, C. J. The human cytoplasmic RNA terminal U-transferase ZCCHC11 targets histone mRNAs for degradation. *RNA N. Y. N* **17**, 39–44 (2011).
11. 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).
12. 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* **1861**, 2971–2979 (2017).
13. Wang, S.-W., Toda, T., MacCallum, R., Harris, A. L. & Norbury, C. Cid1, a Fission Yeast Protein Required for S-M Checkpoint Control when DNA Polymerase  $\delta$  or  $\epsilon$  Is Inactivated. *Mol. Cell. Biol.* **20**, 3234 (2000).
14. Rissland, O. S. & Norbury, C. J. The Cid1 poly(U) polymerase. *Biochim. Biophys. Acta* **1779**, 286–294 (2008).
15. Rissland, O. S., Mikulasova, A. & Norbury, C. J. Efficient RNA polyuridylation by noncanonical poly(A) polymerases. *Mol. Cell. Biol.* **27**, 3612–3624 (2007).
16. Munoz-Tello, P., Gabus, C. & Thore, S. A critical switch in the enzymatic properties of the Cid1 protein deciphered from its product-bound crystal structure. *Nucleic Acids Res.* **42**, 3372–3380 (2014).
17. Malecki, M. *et al.* The exoribonuclease Dis3L2 defines a novel eukaryotic RNA degradation pathway. *EMBO J.* **32**, 1842–1854 (2013).
18. Lim, J. *et al.* Uridylation by TUT4 and TUT7 Marks mRNA for Degradation. *Cell* **159**, 1365–1376 (2014).
19. 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).
20. Sement, F. M. *et al.* Uridylation prevents 3' trimming of oligoadenylated mRNAs. *Nucleic Acids Res.* **41**, 7115–7127 (2013).
21. 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).
22. Glahder, J. A. & Norrild, B. Involvement of hGLD-2 in cytoplasmic polyadenylation of human p53 mRNA. *APMIS* **119**, 769–775 (2011).
23. Martin, G. & Keller, W. RNA-specific ribonucleotidyl transferases. *RNA N. Y. N* **13**, 1834–1849 (2007).
24. Yamagishi, R., Tsusaka, T., Mitsunaga, H., Maehata, T. & Hoshino, S. The STAR protein QKI-7 recruits PAPD4 to regulate post-transcriptional polyadenylation of target mRNAs. *Nucleic Acids Res.* **44**, 2475–2490 (2016).
25. 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).
26. Wang, L., Eckmann, C. R., Kadyk, L. C., Wickens, M. & Kimble, J. A regulatory cytoplasmic poly(A) polymerase in *Caenorhabditis elegans*. *Nature* **419**, 312–316 (2002).
27. Traut, T. W. Physiological concentrations of purines and pyrimidines. *Mol. Cell. Biochem.* **140**, 1–22 (1994).

28. Kim, G.-W. *et al.* Hepatitis C Virus Core Protein Promotes miR-122 Destabilization by Inhibiting GLD-2. *PLoS Pathog.* **12**, e1005714 (2016).
29. 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).
30. Kim, J. H. & Richter, J. D. Opposing polymerase-deadenylase activities regulate cytoplasmic polyadenylation. *Mol. Cell* **24**, 173–183 (2006).
31. Mertins, P. *et al.* Proteogenomics connects somatic mutations to signalling in breast cancer. *Nature* **534**, 55–62 (2016).
32. Mertins, P. *et al.* Ischemia in tumors induces early and sustained phosphorylation changes in stress kinase pathways but does not affect global protein levels. *Mol. Cell. Proteomics MCP* **13**, 1690–1704 (2014).
33. Kettenbach, A. N. *et al.* Quantitative phosphoproteomics identifies substrates and functional modules of Aurora and Polo-like kinase activities in mitotic cells. *Sci. Signal.* **4**, rs5 (2011).
34. Bandiera, S., Pfeffer, S., Baumert, T. F. & Zeisel, M. B. miR-122--a key factor and therapeutic target in liver disease. *J. Hepatol.* **62**, 448–457 (2015).
35. Johnson, S. M. *et al.* RAS is regulated by the let-7 microRNA family. *Cell* **120**, 635–647 (2005).
36. Manning, B. D. & Toker, A. AKT/PKB Signaling: Navigating the Network. *Cell* **169**, 381–405 (2017).
37. Caretta, A. & Mucignat-Caretta, C. Protein Kinase A in Cancer. *Cancers* **3**, 913–926 (2011).
38. Kwak, J. E. & Wickens, M. A family of poly(U) polymerases. *RNA* **13**, 860–867 (2007).
39. Stevenson, A. L. & Norbury, C. J. The Cid1 family of non-canonical poly(A) polymerases. *Yeast Chichester Engl.* **23**, 991–1000 (2006).
40. Munoz-Tello, P., Gabus, C. & Thore, S. Functional implications from the Cid1 poly(U) polymerase crystal structure. *Struct. Lond. Engl. 1993* **20**, 977–986 (2012).
41. Yates, L. A. *et al.* Structural plasticity of Cid1 provides a basis for its distributive RNA terminal uridylyl transferase activity. *Nucleic Acids Res.* **43**, 2968–2979 (2015).
42. Yates, L. A. *et al.* Improved crystallization and diffraction of caffeine-induced death suppressor protein 1 (Cid1). *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **71**, 346–353 (2015).
43. Lim, J. *et al.* Mixed tailing by TENT4A and TENT4B shields mRNA from rapid deadenylation. *Science* **361**, 701–704 (2018).
44. Karve, T. M. & Cheema, A. K. Small changes huge impact: the role of protein posttranslational modifications in cellular homeostasis and disease. *J. Amino Acids* **2011**, 207691 (2011).
45. Hornbeck, P. V. *et al.* PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res.* **43**, D512–520 (2015).
46. Mollica, L. *et al.* Binding Mechanisms of Intrinsically Disordered Proteins: Theory, Simulation, and Experiment. *Front. Mol. Biosci.* **3**, 52 (2016).
47. Chung, C. Z., Jo, D. H. S. & Heinemann, I. U. Nucleotide specificity of the human terminal nucleotidyltransferase Gld2 (TUT2). *RNA N. Y. N* **22**, 1239–1249 (2016).
48. Chung, C. Z. *et al.* RNA surveillance by uridylation dependent RNA decay in *Schizosaccharomyces pombe*. *Nucleic Acids Res.* (2019). doi:10.1093/nar/gkz043

49. Gu, T.-L. *et al.* Identification of activated Tnk1 kinase in Hodgkin's lymphoma. *Leukemia* **24**, 861–865 (2010).
50. Geoghegan, V., Guo, A., Trudgian, D., Thomas, B. & Acuto, O. Comprehensive identification of arginine methylation in primary T cells reveals regulatory roles in cell signalling. *Nat. Commun.* **6**, 6758 (2015).
51. Akimov, V. *et al.* UbiSite approach for comprehensive mapping of lysine and N-terminal ubiquitination sites. *Nat. Struct. Mol. Biol.* **25**, 631–640 (2018).
52. Mertins, P. *et al.* Integrated proteomic analysis of post-translational modifications by serial enrichment. *Nat. Methods* **10**, 634–637 (2013).
53. Larsen, S. C. *et al.* Proteome-wide analysis of arginine monomethylation reveals widespread occurrence in human cells. *Sci. Signal.* **9**, rs9 (2016).
54. Girard, M., Jacquemin, E., Munnich, A., Lyonnet, S. & Henrion-Caude, A. miR-122, a paradigm for the role of microRNAs in the liver. *J. Hepatol.* **48**, 648–656 (2008).
55. Peng, F. *et al.* HBx down-regulated Gld2 plays a critical role in HBV-related dysregulation of miR-122. *PloS One* **9**, e92998 (2014).
56. Janssen, H. L. A. *et al.* Treatment of HCV infection by targeting microRNA. *N. Engl. J. Med.* **368**, 1685–1694 (2013).
57. Van der Ree, M. H. *et al.* Miravirsin dosing in chronic hepatitis C patients results in decreased microRNA-122 levels without affecting other microRNAs in plasma. *Aliment. Pharmacol. Ther.* **43**, 102–113 (2016).
58. Gebert, L. F. R. *et al.* Miravirsin (SPC3649) can inhibit the biogenesis of miR-122. *Nucleic Acids Res.* **42**, 609–621 (2014).
59. Kadyk, L. C. & Kimble, J. Genetic regulation of entry into meiosis in *Caenorhabditis elegans*. *Dev. Camb. Engl.* **125**, 1803–1813 (1998).
60. Mohammad, A. *et al.* Initiation of Meiotic Development Is Controlled by Three Posttranscriptional Pathways in *Caenorhabditis elegans*. *Genetics* (2018). doi:10.1534/genetics.118.300985
61. Nousch, M., Minasaki, R. & Eckmann, C. R. Polyadenylation is the key aspect of GLD-2 function in *C. elegans*. *RNA N. Y. N* **23**, 1180–1187 (2017).
62. Turk, M. A. *et al.* MiRAR-miRNA Activity Reporter for Living Cells. *Genes* **9**, (2018).
63. Pisacane, P. & Halic, M. Tailing and degradation of Argonaute-bound small RNAs protect the genome from uncontrolled RNAi. *Nat. Commun.* **8**, 15332 (2017).
64. 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).

## Appendices

### Appendix A

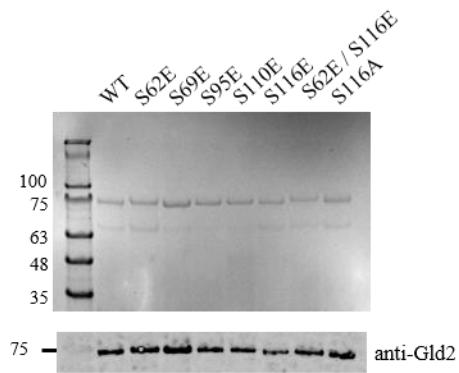
**Table A1: Primers for cloning of Gld2 into pGEX-6P-2 and mutagenesis of Gld2.**

Primer Name	Primer Sequence
Gld2BamHIfor	5'-ATGCGGATCCGAAAATCTGTACTTC-3'
Gld2XhoIrev	5'-TAATCTCGAGTTAACGTTTTTAACACGG-3'
Gld2Glu62for	5'-CATACGGTAACGTGGAACCGATCCAGACCAGCGC-3'
Gld2Glu62rev	5'-GCGCTGGTCTGGATCGGTTCCACGTTACCGTATG-3'
Gld2Glu69for	5'-GATCCAGACCAGCGCCGAACCTCTGTTCCGTGG-3'
Gld2Glu69rev	5'-CCACGGAACAGAGGTTTCGGCGCTGGTCTGGATC-3'
Gld2Glu95for	5'-CGCCAGCGTTTTTCATGAACCGCACCAAGAACCG-3'
Gld2Glu95rev	5'-CGGTTCTTGGTGCGGTTTCATGAAAACGCTGGCG-3'
Gld2Glu110for	5'-GAACCAGATTGTGCCGTTAGAAGGTGAACGTCGCTATAGC-3'
Gld2Glu110rev	5'-GCTATAGCGACGTTACCTTCTAACGGCACAATCTGGTTC-3'
Gld2Glu116for	5'-CGGTGAACGTCGCTATGAAATGCCTCCGCTGTTTC-3'
Gld2Glu116rev	5'-GAAACAGCGGAGGCATTTTCATAGCGACGTTACCG-3'
Gld2Ala116for	5'-CGGTGAACGTCGCTATGCCATGCCTCCGCTGTTTC-3'
Gld2Ala116rev	5'-GAAACAGCGGAGGCATGGCATAAGCGACGTTACCG-3'

**Table A2: Kinases predicted to phosphorylate Gld2 at residues S62, S69, S95, S110 and S116.**

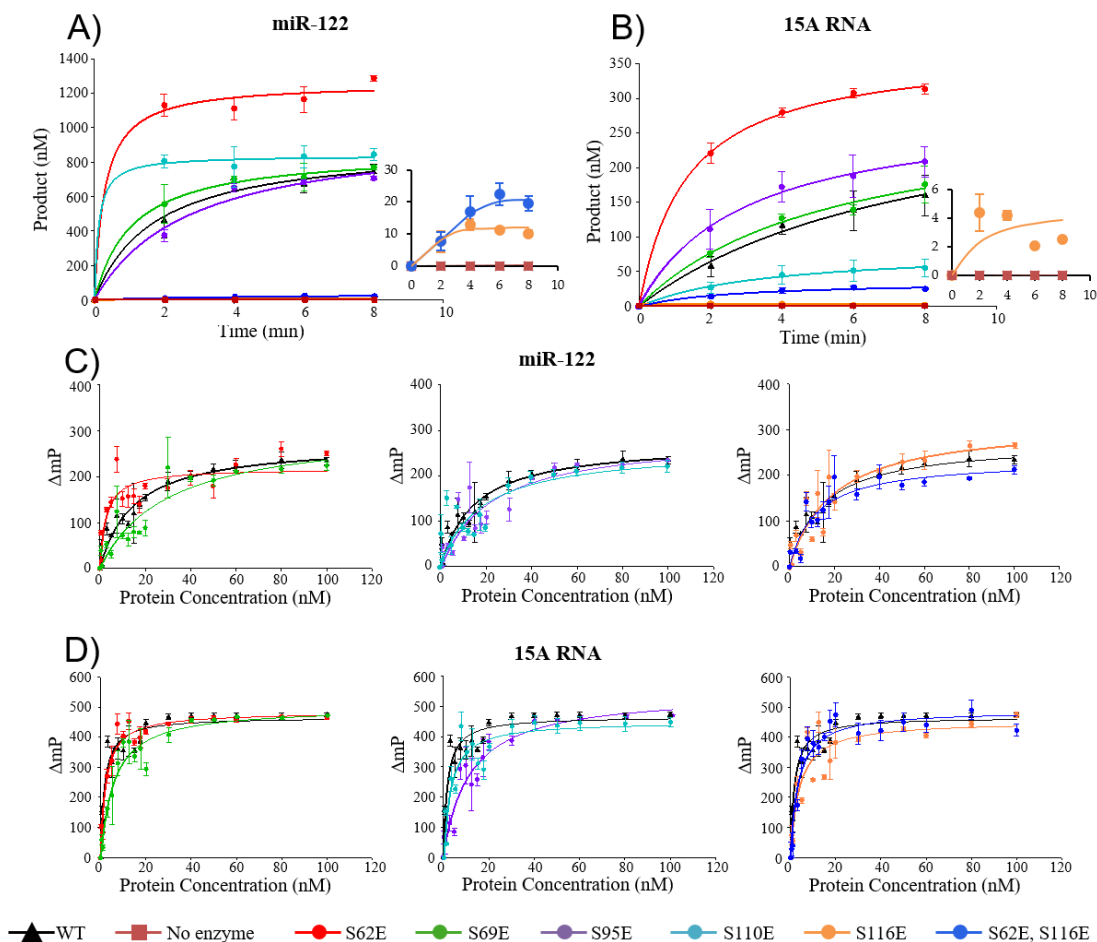
Position in Gld2	Gld2 Sequence*	Predicted Kinases
62	NV <u><b>S</b></u> P	GSK-3, ERK1, ERK2, CDK5
62	<u><b>S</b></u> PIQT	MAPKAPK2
62	QLTYGNV <u><b>S</b></u> PIQTSAS	PKA, PIKK, CDK, MAPK, ERK1, ERK2
69	SA <u><b>S</b></u> P	GSK-3, ERK1, ERK2, CDK5
69	SPIQTSAS <u><b>S</b></u> PLFRGRK	PKA, PKC
95	RQRF <u><b>S</b></u>	Akt2, MAPKAPK
95	RF <u><b>S</b></u>	PKA, PKC
95	FH <u><b>S</b></u> P	GSK-3, ERK1, ERK2, CDK5
95	GKRQRFH <u><b>S</b></u> PHQEPTV	Akt, Akt2, AGC, CDK5
110	<u><b>S</b></u> GE	CK2
116	LSGERRY <u><b>S</b></u> MPPLFHT	PKA, PKC, Akt, AGC, CAMK

\*Serine residue predicted to be phosphorylated is bolded and underlined



**Figure A1: Purified Gld2 constructs.**

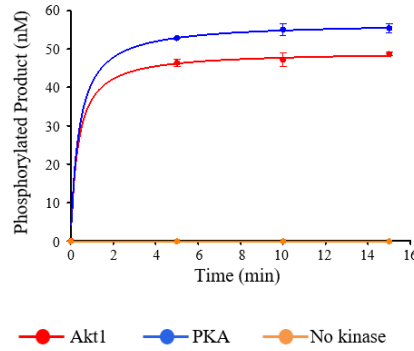
Purified GST-Gld2 constructs were visualized on a 10% polyacrylamide SDS gel stained with Coomassie Brilliant Blue G-250 and Western blotted using an anti-Gld2 antibody.



**Figure A2: Catalytic activity and RNA binding of Gld2 phosphomimetic variants.**

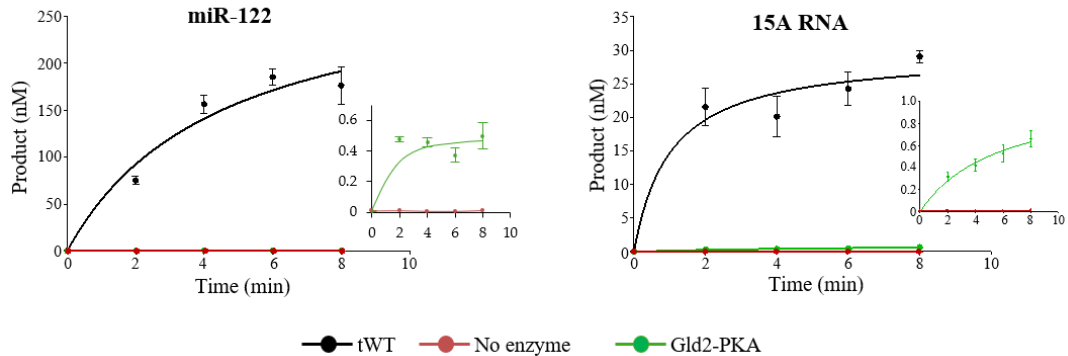
The activity plots show the amount of product formed at 1  $\mu$ M ATP plotted over time with **A)** miR-122 (22 nts) or **B)** oligo(A) tail mimic 15A RNA (15 nts). Insets of Gld2 S116E and the no enzyme control are for better visualization of the data. Wildtype Gld2 (WT) and no enzyme reactions are plotted on all graphs for comparison and the insets show the low activity Gld2 variants. Error bars represent the standard error. The binding assay plots show the binding of wildtype Gld2 and Gld2 mutants to **C)** miR-122 or **D)** oligo(A) tail mimic 15A RNA. WT reaction is plotted on all graphs for comparison. Gld2 mutants are graphed separately for better visualization of the data. Each Gld2 variant was incubated with 1  $\mu$ M unlabelled and [ $\alpha$ - $^{32}$ P]-labelled ATP and 2  $\mu$ M RNA substrate. Samples were taken every 2 minutes and stopped with the addition of 2 x RNA loading dye. Reactions were analyzed via electrophoretic separation and subsequent phosphorimaging. Specific activity was calculated from the linear slope of the curve. Fluorescence anisotropy was used to determine the  $K_d$ . Each Gld2 enzyme was incubated with a RNA substrate fluorescently labelled on the 5'-end with 6-FAM and incubated at room temperature for 20 minutes. Fluorescence polarization was measured at Ex. 492nm and Em. 535/20 nm and the  $K_d$  was calculated using SigmaPlot. Error bars represent the standard error from triplicate reactions.  $\Delta$ mP, change in fluorescence polarization in millipolarization units.





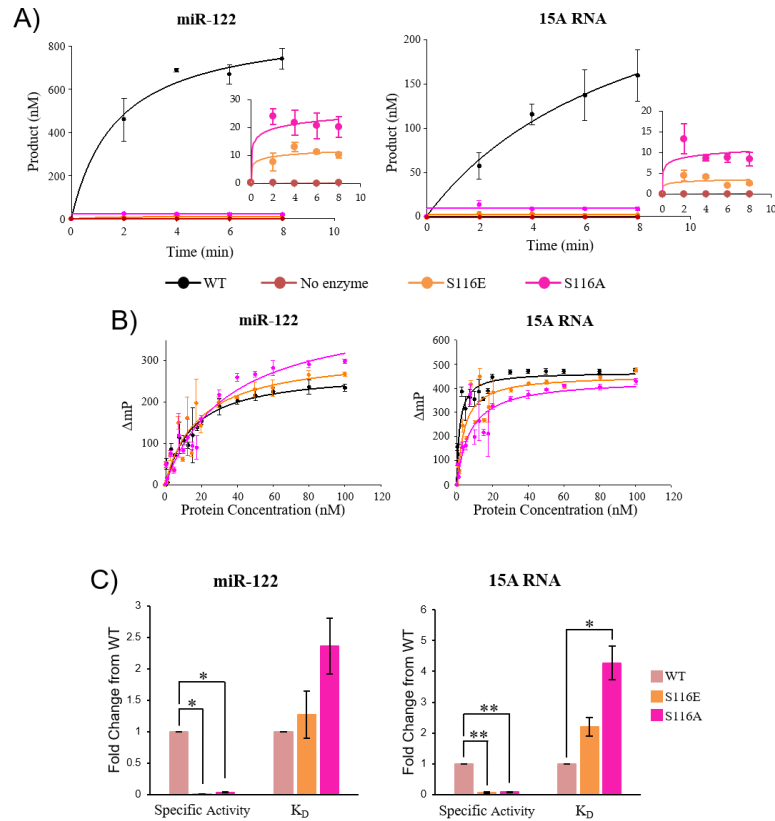
**Figure A3: PKA and Akt1 produce phosphorylated Gld2.**

The activity plot shows the amount of phosphorylated Gld2 (nM) produced over time when incubated with PKA or Akt1. A no kinase control is also plotted. Gld2 WT and a kinase were incubated with  $[\gamma\text{-}^{32}\text{P}]\text{-ATP}$  and samples were taken every 5 minutes and stopped with the addition of 2 x SDS loading dye. Reactions were analyzed on a 10% polyacrylamide SDS gel and subsequent phosphorimaging and product quantification. Error bars are one standard error calculated from triplicate reactions.



**Figure A4: PKA-mediated phosphorylation down-regulates Gld2 catalytic activity.**

The activity plots show the amount of product formed at 1  $\mu\text{M}$  ATP plotted against time with miR-122 (22 nts) or oligo(A) tail mimic 15A RNA (15 nts). Gld2 was incubated with PKA and purified before activity assay. Phosphorylated Gld2 was incubated with 1  $\mu\text{M}$  unlabelled and  $[\alpha\text{-}^{32}\text{P}]\text{-ATP}$  and 2  $\mu\text{M}$  RNA substrate. Samples were taken every 2 minutes and stopped with the addition of 2 x RNA loading dye. Reactions were analyzed via electrophoretic separation and subsequent phosphorimaging. Wildtype Gld2 treated in the same way without kinase (treated wildtype, tWT) and no enzyme reactions are plotted on all graphs for comparison. Insets of Gld2 phosphorylated by PKA and the no enzyme control are for better visualization of the data. Error bars are one standard error calculated from triplicate reactions.



**Figure A5: Gld2 S116A does not mimic wildtype activity and binding.**

**A)** The activity plots show the amount of product formed at 1  $\mu$ M ATP plotted over time with miR-122 (22 nts) or oligo(A) tail mimic 15A RNA (15 nts). Insets of Gld2 S116E, Gld2 S116A, and the no enzyme control are for better visualization of the data. Wildtype Gld2 (WT) and no enzyme reactions are plotted on all graphs for comparison. Each Gld2 variant was incubated with 1  $\mu$ M unlabelled and [ $\alpha$ - $^{32}$ P]-labelled ATP and 2  $\mu$ M RNA substrate. Samples were taken every 2 minutes and stopped with the addition of 2 x RNA loading dye. Reactions were analyzed via electrophoretic separation and subsequent phosphorimaging. Specific activity was calculated from the linear slope of the curve. Error bars represent the standard error. **B)** The binding assay plots generated through fluorescence anisotropy show the binding of wildtype Gld2 and Gld2 S116E and S116A mutants to miR-122 or oligo(A) tail mimic 15A RNA. WT reaction is plotted on all graphs for comparison. Each Gld2 enzyme was incubated with a RNA substrate fluorescently labelled on the 5'-end with 6-FAM and incubated at room temperature for 20 minutes. Fluorescence polarization was measured at Ex. 492nm and Em. 535/20 nm and the binding affinity ( $K_d$ ) was calculated using SigmaPlot. Error bars represent the standard error from triplicate reactions. **C)** Bar graphs showing the fold change in specific activity at 1  $\mu$ M ATP and  $K_d$  between wildtype Gld2, S116E, and S116A with miR-122 or oligo(A) tail mimic 15A RNA calculated from **A)** and **B)**.  $\Delta$ mP, change in fluorescence polarization in millipolarization units.

## Appendix B

**Table B1: Oligonucleotide primers.**

Target	Method	Forward Primer 5'-3'	Reverse Primer 5'-3'
Thf1	Northern Blot	CGAAGGTCAACCCTTGTTTACCG	CCAACATCAATAGCGACAACG
ECI1		GGAGACATACATTTACAAAGCG	GAGCAGTCATGATTTCTTCCTGTC
SPAC19G12.09		GGTACAGCTTTGTTTAAGAAAG	GGAATAGGACTATGCAAAAGG
Tdh1		GGTGCTGACTACGTTATCGAG	CTTGGAGGGACCGTCAACG
SPAC27E2.11c		GCCATGATTGCTGTCGCTTGCCTGTC	CGACACCACACAGAATAATATAGTTGAATGG
Cid1		TCACGTTATCAAGCCTCCCG	AAGCGGCCATAAATTCCCCTC
Dis3L2		AGGAGTCATCGGGAGCAACT	AGACTGGCACCATTACGCTC
ScpofMr12	cRACE	GAAGGAGGAATTGCGAG	GATTACGATTTGAGCTTG
SPBC215.11c		CAGTCCGGTTATGCTACC	GCAAGCCTCTTTGTC
dak2		GGCATATCATGTAACTTG	GTTTAGTAGAGGGAGAAGC
SPAC19G12.09		GCGTTATACCTATCACTAC	CAACATGCTTATCGCTGC
Pex22	RT-qPCR	GTCTTTTCCTCAGGTTTCGGACT	TCCGACTCAGAAAGTGCTGT
Hsp104		CTTCGTCCTTCTCACGCTCT	TGGTTACTGCTGCCCATCTC
Hsp78		GCTCCTTCCAGGTCCTCAG	ATCCTGCGTTAGTTGGTCCG
Ssa2		GGTGACGCTGCTAAGAACCA	CGGGGGTAAAGGTCTTGGTC
Tcg1		CCGCTGAGGAACTGTAACC	CCGTTTCAAACAATGAACGGATT
Rpp0		GTTACCGGCAGGGACAAAGA	GATTCGTCGTGCTATGCGTG

**Table B2: Genes up- or down-regulated in either *S. pombe*  $\Delta cid1$  deletion strain or *S. pombe*  $\Delta dis3L2$  deletion compared to wildtype *S. pombe*.**

		Fold change in $\Delta cid1$	Effect change in $\Delta cid1$	Fold change in $\Delta dis3L2$	Effect change in $\Delta dis3L2$
<b>Actin cytoskeleton organization</b>					
<i>adf1</i>	actin depolymerizing factor, cofilin	1.21	2.05	1.34	3.80
<i>arp2</i>	ARP2/3 actin-organizing complex subunit Arp2	0.98	0.89	1.27	5.65
<i>arp3</i>	Arp2/3 protein complex, actin-like protein subunit Arp3	1.49	2.40	1.83	3.76
<i>myo1</i>	myosin type I	0.98	0.89	1.33	6.11
<i>SPAC637.13c</i>	cytoskeletal signaling protein Slm1 (predicted)	0.98	0.95	1.54	3.46
<b>Antisense RNA</b>		<b>Fold</b>	<b>Effect</b>	<b>Fold</b>	<b>Effect</b>

		change in $\Delta cid1$	change in $\Delta cid1$	change in $\Delta dis3L2$	change in $\Delta dis3L2$
<i>SPNCRNA.1036</i>	antisense RNA (predicted)	1.37	2.59	2.29	14.8
<i>SPNCRNA.1132</i>	antisense RNA (predicted)	3.02	2.91	1.65	1.56
<i>SPNCRNA.1138</i>	antisense RNA (predicted)	0.86	0.86	0.19	0.16
<i>SPNCRNA.1170</i>	antisense RNA (predicted)	2.49	3.06	1.24	1.24
<i>SPNCRNA.1204</i>	antisense RNA (predicted)	0.88	0.87	0.38	0.22
<i>SPNCRNA.1212</i>	antisense RNA (predicted)	0.39	0.26	1.00	1.00
<i>SPNCRNA.1235</i>	antisense RNA (predicted)	1.10	1.34	1.82	2.89
<i>SPNCRNA.1447</i>	antisense RNA (predicted)	1.14	1.58	1.72	8.58
<i>SPNCRNA.1451</i>	antisense RNA (predicted)	0.87	0.68	0.61	0.13
<i>SPNCRNA.1467</i>	antisense RNA (predicted), possible alternative UTR	0.94	0.78	0.58	0.14
<i>SPNCRNA.1548</i>	antisense RNA (predicted)	0.39	0.25	0.27	0.16
<i>SPNCRNA.1563</i>	antisense RNA (predicted)	2.01	1.61	2.85	2.89
<i>SPNCRNA.1626</i>	antisense RNA (predicted)	0.88	0.76	2.66	7.15
<i>SPNCRNA.1665</i>	antisense RNA (predicted)	1.27	2.07	1.42	2.83
<i>SPNCRNA.579</i>	antisense RNA (predicted)	0.79	0.74	0.51	0.23
<i>SPNCRNA.606</i>	antisense RNA (predicted)	0.68	0.55	0.33	0.25
<i>SPNCRNA.636</i>	antisense RNA (predicted)	1.50	4.30	1.73	2.95
<i>SPNCRNA.706</i>	antisense RNA (predicted)	1.26	1.15	4.98	3.15
<i>SPNCRNA.857</i>	antisense RNA (predicted)	1.00	1.00	1.79	3.07
<i>SPNCRNA.886</i>	antisense RNA (predicted)	0.78	0.64	0.48	0.34
<i>SPNCRNA.949</i>	antisense RNA (predicted)	1.20	1.46	0.67	0.30
<i>SPNCRNA.967</i>	antisense RNA (predicted)	0.59	0.34	0.69	0.38
		Fold change in $\Delta cid1$	Effect change in $\Delta cid1$	Fold change in $\Delta dis3L2$	Effect change in $\Delta dis3L2$
<b>Carbohydrate metabolic process</b>					
<i>adh4</i>	alcohol dehydrogenase Adh4	0.90	0.73	0.46	0.16
<i>dak1</i>	dihydroxyacetone kinase Dak1	1.02	1.05	1.81	3.41
<i>gal1</i>	galactokinase Gal1	1.35	1.65	3.78	8.51
<i>gal10</i>	UDP-glucose 4-epimerase/aldose 1- epimerase Gal10	0.83	0.82	2.63	6.07
<i>gal7</i>	galactose-1-phosphate uridylyltransferase Gal7	0.84	0.78	1.67	3.52
<i>gpd2</i>	glycerol-3-phosphate dehydrogenase Gpd2	0.93	0.69	0.63	0.12
<i>mal1</i>	maltase alpha-glucosidase Mal1	0.74	0.34	0.88	0.73
<i>pgi1</i>	glucose-6-phosphate isomerase (predicted)	1.16	2.49	1.87	4.75
<i>SPAC26F1.07</i>	glucose 1-dehydrogenase (NADP+) (predicted)	0.87	0.78	2.36	4.63
<i>SPACUNK4.16c</i>	alpha,alpha-trehalose-phosphate synthase (predicted)	0.99	0.99	2.45	5.29
<i>SPBC1683.04</i>	glycosyl hydrolase family 3 (predicted)	0.77	0.26	1.17	1.59
<i>SPBC2G2.17c</i>	beta-glucosidase Psu2 (predicted)	0.53	0.54	0.33	0.28
<i>SPBC32F12.10</i>	phosphoglucomutase (predicted)	1.11	1.44	1.93	3.75
<i>SPCC306.06c</i>	ER membrane protein, BIG1 family	0.80	0.32	0.81	0.23

	(predicted)				
<i>tdh1</i>	glyceraldehyde-3-phosphate dehydrogenase Tdh1	1.57	1.68	2.62	2.88
<i>tps1</i>	alpha,alpha-trehalose-phosphate synthase [UDP-forming]	1.18	1.28	2.75	4.08
<b>Carbohydrate derived metabolic process</b>		<b>Fold change in <math>\Delta cid1</math></b>	<b>Effect change in <math>\Delta cid1</math></b>	<b>Fold change in <math>\Delta dis3L2</math></b>	<b>Effect change in <math>\Delta dis3L2</math></b>
<i>SPCC1322.04</i>	UTP-glucose-1-phosphate uridylyltransferase Fyu1	1.19	1.38	2.25	4.67
<b>Cell adhesion</b>		<b>Fold change in <math>\Delta cid1</math></b>	<b>Effect change in <math>\Delta cid1</math></b>	<b>Fold change in <math>\Delta dis3L2</math></b>	<b>Effect change in <math>\Delta dis3L2</math></b>
<i>SPBPJ4664.02</i>	cell surface glycoprotein, flocculin, related to Gsf2	0.63	0.34	1.07	1.24
<b>Cell wall organization or biogenesis</b>		<b>Fold change in <math>\Delta cid1</math></b>	<b>Effect change in <math>\Delta cid1</math></b>	<b>Fold change in <math>\Delta dis3L2</math></b>	<b>Effect change in <math>\Delta dis3L2</math></b>
<i>omh5</i>	alpha-1,2-mannosyltransferase Omh5 (predicted)	1.36	3.14	1.04	1.15
<b>Cellular amino acid metabolic process</b>		<b>Fold change in <math>\Delta cid1</math></b>	<b>Effect change in <math>\Delta cid1</math></b>	<b>Fold change in <math>\Delta dis3L2</math></b>	<b>Effect change in <math>\Delta dis3L2</math></b>
<i>arg1</i>	acetylornithine aminotransferase	0.86	0.29	0.82	0.20
<i>arg11</i>	N-acetyl-gamma-glutamyl-phosphate reductase/acetylglutamate kinase	0.86	0.32	0.86	0.29
<i>gdh1</i>	NADP-specific glutamate dehydrogenase Gdh1 (predicted)	1.28	3.27	1.35	7.05
<i>leu2</i>	3-isopropylmalate dehydratase Leu2	0.94	0.61	0.61	0.10
<i>SPBC19F5.04</i>	aspartate kinase (predicted)	0.89	0.35	0.77	0.10
<i>SPBC776.03</i>	homoserine dehydrogenase (predicted)	1.09	1.46	1.28	3.01
<i>SPBPB2B2.05</i>	peptidase family C26 protein	1.57	1.54	3.80	5.73
<i>SPCC70.03c</i>	proline dehydrogenase Put1 (predicted)	0.84	0.24	0.90	0.36
<i>trx1</i>	cytosolic thioredoxin Trx1	1.02	1.10	1.51	6.99
<b>Chromatin organization</b>		<b>Fold change in <math>\Delta cid1</math></b>	<b>Effect change in <math>\Delta cid1</math></b>	<b>Fold change in <math>\Delta dis3L2</math></b>	<b>Effect change in <math>\Delta dis3L2</math></b>
<i>ams2</i>	cell cycle regulated GATA-type transcription factor Ams2	1.20	1.62	1.54	4.09
<i>asa1</i>	Astra associated protein 1 Asa1	0.78	0.50	0.64	0.29
<i>SPBC582.04c</i>	RNAi protein, Dsh1	0.72	0.29	0.72	0.28
<i>swc3</i>	Swr1 complex subunit Swc3	1.12	1.28	1.52	2.93
<i>swd3</i>	WD repeat protein Swd3	0.81	0.42	0.73	0.26
<i>tel2</i>	Tel2/Rad-5/Clk-2 family protein Tel2	0.72	0.32	0.86	0.58
<b>Cofactor metabolic process</b>		<b>Fold change in <math>\Delta cid1</math></b>	<b>Effect change in <math>\Delta cid1</math></b>	<b>Fold change in <math>\Delta dis3L2</math></b>	<b>Effect change in <math>\Delta dis3L2</math></b>
<i>coq3</i>	hexaprenyldihydroxybenzoate methyltransferase Coq3	1.44	1.73	2.15	3.15

<i>pdx1</i>	pyruvate dehydrogenase protein x component, Pdx1 (predicted)	0.95	0.81	0.69	0.29
<i>SPAC806.06c</i>	nicotinamide mononucleotide (NMN) adenylyltransferase (predicted)	2.21	4.15	2.02	6.54
<i>SPBC1709.19c</i>	mitochondrial iron-sulfur cluster assembly protein Nfu1 (predicted)	0.82	0.48	0.70	0.23
<i>SPBC4B4.01c</i>	fumble family pantothenate kinase (predicted)	1.39	2.38	1.44	2.85
<i>SPBC947.15c</i>	mitochondrial NADH dehydrogenase (ubiquinone) Nde1 (predicted)	1.14	1.56	1.71	10.56
<b>Cytoplasmic translation</b>		<b>Fold change in <math>\Delta cid1</math></b>	<b>Effect change in <math>\Delta cid1</math></b>	<b>Fold change in <math>\Delta dis3L2</math></b>	<b>Effect change in <math>\Delta dis3L2</math></b>
<i>dph2</i>	diphthamide biosynthesis protein (predicted)	0.91	0.78	0.69	0.28
<i>hri2</i>	eIF2 alpha kinase Hri1	0.96	0.87	1.42	4.06
<i>psi1</i>	DNAJ domain protein Psi1	1.07	1.24	2.60	7.32
<i>SPBC3B9.01</i>	Hsp70 nucleotide exchange factor Fes1 (predicted)	1.33	2.59	2.74	4.35
<i>tif212</i>	translation initiation factor eIF2 beta subunit (predicted)	1.31	3.53	0.99	0.94
<b>Detoxification</b>		<b>Fold change in <math>\Delta cid1</math></b>	<b>Effect change in <math>\Delta cid1</math></b>	<b>Fold change in <math>\Delta dis3L2</math></b>	<b>Effect change in <math>\Delta dis3L2</math></b>
<i>glo1</i>	glyoxalase I	1.27	2.44	1.83	4.86
<i>SPAC869.02c</i>	nitric oxide dioxygenase Yhb1	0.55	0.11	1.12	1.37
<b>DNA-templated transcription</b>		<b>Fold change in <math>\Delta cid1</math></b>	<b>Effect change in <math>\Delta cid1</math></b>	<b>Fold change in <math>\Delta dis3L2</math></b>	<b>Effect change in <math>\Delta dis3L2</math></b>
<i>nut2</i>	mediator complex subunit Med10	0.84	0.58	0.69	0.33
<i>rad24</i>	14-3-3 protein Rad24	1.19	1.68	1.49	3.55
<i>rpc53</i>	DNA-directed RNA polymerase III complex subunit Rpc53 (predicted)	0.74	0.50	0.70	0.33
<i>rrn3</i>	ribosomal DNA (rDNA)				
	transcription factor Rrn3	0.90	0.66	0.66	0.17
<i>SPBC2G5.02c</i>	CK2 family regulatory subunit Ckb2 (predicted)	0.91	0.60	0.76	0.22
<i>SPBC83.17</i>	transcriptional coactivator, multiprotein bridging factor Mbf1 (predicted)	1.28	2.34	1.51	3.78
<i>SPCC548.05c</i>	ubiquitin-protein ligase E3 Dbl5	0.36	0.27	0.37	0.26
<b>DNA recombination</b>		<b>Fold change in <math>\Delta cid1</math></b>	<b>Effect change in <math>\Delta cid1</math></b>	<b>Fold change in <math>\Delta dis3L2</math></b>	<b>Effect change in <math>\Delta dis3L2</math></b>
<i>mcp7</i>	meiosis specific coiled-coil protein Mcp7	0.83	0.60	0.64	0.19
<i>rec10</i>	meiotic recombination protein Rec10	0.54	0.32	0.87	0.75
<i>Tf2-1</i> (NC_003424 1465326.14702 52)	retrotransposable element/transposon Tf2-type	2.58	2.90	1.70	1.60

		Fold change in $\Delta cid1$	Effect change in $\Delta cid1$	Fold change in $\Delta dis3L2$	Effect change in $\Delta dis3L2$
<b>DNA repair</b>					
<i>lub1</i>	WD repeat protein Lub1	0.98	0.91	1.47	2.97
<i>msh6</i>	MutS protein homolog	1.18	1.61	1.49	5.18
<i>nse4</i>	Smc5-6 complex non-SMC delta-kleisin subunit Nse4	0.90	0.70	0.71	0.35
<i>nth1</i>	DNA endonuclease III	0.89	0.71	0.70	0.34
<i>pnk1</i>	DNA kinase/phosphatase Pnk1	0.95	0.85	0.76	0.34
<i>pso2</i>	DNA 5' exonuclease (predicted)	0.70	0.39	0.67	0.26
<i>rik1</i>	silencing protein Rik1	5.07	3.52	3.65	2.61
<i>SPBC23E6.02</i>	ATP-dependent DNA helicase, ubiquitin-protein ligase E3 (predicted)	2.45	4.46	1.39	1.69
<i>tdp1</i>	tyrosyl-DNA phosphodiesterase Tdp1	0.74	0.34	0.93	0.83
<i>tra2</i>	NuA4 complex phosphatidylinositol pseudokinase complex subunit Tra2	0.80	0.34	0.88	0.64
<b>DNA replication</b>					
<i>cdc23</i>	MCM-associated protein Mcm10	0.52	0.33	0.50	0.26
<i>cdt2</i>	WD repeat protein Cdt2	1.38	1.97	1.78	4.01
<i>dna2</i>	DNA replication endonuclease-helicase Dna2	0.81	0.39	0.80	0.25
<b>Establishment or maintenance of cell polarity</b>					
<i>paa1</i>	protein phosphatase regulatory subunit Paa1	1.19	2.41	1.38	4.45
<i>tea3</i>	cell end marker Tea3	0.84	0.38	0.77	0.19
<b>Generation of precursor metabolites and energy</b>					
<i>cob</i>	cytochrome b, Cob1 (predicted)	0.89	0.61	0.65	0.25
<i>cox1</i>	cytochrome c oxidase 1 (predicted)	0.97	0.92	0.66	0.35
<i>idp1</i>	isocitrate dehydrogenase Idp1 (predicted)	0.81	0.64	1.35	4.33
<i>SPBC660.16</i>	phosphogluconate dehydrogenase, decarboxylating	1.15	1.98	1.80	4.88
<i>SPCC1620.08</i>	succinate-CoA ligase beta subunit Lsc2 (predicted)	0.82	0.32	1.15	1.81
<b>Intergenic RNA</b>					
<i>SPNCRNA.1115</i>	intergenic RNA (predicted), possible alternative UTR	1.12	1.40	1.37	3.04
<i>SPNCRNA.1164</i>	intergenic RNA (predicted)	1.76	3.34	1.13	1.24
<i>SPNCRNA.1297</i>	intergenic RNA (predicted), possible alternative UTR	0.80	0.67	0.31	0.16
<i>SPNCRNA.1325</i>	intergenic RNA (predicted), possible alternative UTR	2.28	2.05	3.05	2.93

<i>SPNCRNA.1474</i>	intergenic RNA (predicted)	0.50	0.60	0.40	0.35
<i>SPNCRNA.1657</i>	intergenic RNA (predicted), possible alternative UTR	2.14	2.96	0.62	0.54
<i>SPNCRNA.1673</i>	intergenic RNA (predicted)	2.05	3.19	0.84	0.83
<i>SPNCRNA.671</i>	intergenic RNA (predicted)	1.00	1.00	0.53	0.29
<i>SPNCRNA.672</i>	intergenic RNA (predicted)	2.24	3.23	1.26	1.52
<i>SPNCRNA.737</i>	intergenic RNA (predicted)	1.14	1.08	0.49	0.34
<i>SPNCRNA.781</i>	intergenic RNA (predicted)	1.63	2.69	0.48	0.30
<i>SPNCRNA.877</i>	intergenic RNA (predicted)	2.05	1.94	2.98	3.26
<i>SPNCRNA.935</i>	intergenic RNA (predicted), possible alternative UTR	1.06	1.16	0.70	0.19
<i>SPNCRNA.968</i>	intergenic RNA (predicted), possible alternative UTR	1.13	1.28	0.62	0.35
<b>Lipid metabolic process</b>		<b>Fold change in <i>Δcid1</i></b>	<b>Effect change in <i>Δcid1</i></b>	<b>Fold change in <i>Δdis3L2</i></b>	<b>Effect change in <i>Δdis3L2</i></b>
<i>aim22</i>	lipoate-protein ligase A (predicted)	0.74	0.34	0.72	0.31
<i>erg10</i>	acetyl-CoA C-acetyltransferase Erg10 (predicted)	1.39	3.93	1.47	4.34
<i>plg7</i>	phospholipase A2, PAF family homolog	1.02	1.10	1.36	3.89
<i>SPAC4A8.10</i>	acylglycerol lipase (predicted)	0.77	0.43	0.64	0.31
<i>SPAC977.09c</i>	phospholipase (predicted)	1.14	1.33	2.92	10.71
<i>SPBC36.10</i>	mitochondrial intermembrane space protein; involved in phospholipid metabolism Ups2 (predicted)	0.97	0.93	0.68	0.35
<i>SPCC5E4.05c</i>	mitochondrial acylglycerol lipase Mgl1 (predicted)	1.28	2.36	1.48	3.69
<b>Microtubule cytoskeleton organization</b>		<b>Fold change in <i>Δcid1</i></b>	<b>Effect change in <i>Δcid1</i></b>	<b>Fold change in <i>Δdis3L2</i></b>	<b>Effect change in <i>Δdis3L2</i></b>
<i>sad1</i>	spindle pole body SUN domain protein Sad1	1.07	1.26	1.32	3.30
<b>Mitochondrion organization</b>		<b>Fold change in <i>Δcid1</i></b>	<b>Effect change in <i>Δcid1</i></b>	<b>Fold change in <i>Δdis3L2</i></b>	<b>Effect change in <i>Δdis3L2</i></b>
<i>exo5</i>	mitochondrial single stranded DNA specific 5'-3' exodeoxyribonuclease Exo5 (predicted)	0.54	0.33	0.71	0.50
<i>mdj1</i>	mitochondrial DNAJ domain protein Mdj1 (predicted)	1.11	1.32	1.62	3.01
<i>phb1</i>	prohibitin Phb1 (predicted)	0.97	0.89	1.23	2.94
<i>SPCC1442.05c</i>	MICOS complex subunit Mic26/27 (predicted)	1.25	1.81	1.40	3.72
<i>SPCC4B3.03c</i>	CBS domain protein implicated in magnesium homeostasis (predicted)	0.85	0.46	0.77	0.33
<b>Mitotic cytokinesis</b>		<b>Fold change in <i>Δcid1</i></b>	<b>Effect change in <i>Δcid1</i></b>	<b>Fold change in <i>Δdis3L2</i></b>	<b>Effect change in <i>Δdis3L2</i></b>
<i>fim1</i>	fimbrin	1.03	1.15	1.62	4.60
<i>mid2</i>	medial ring protein Mid2	1.95	2.96	2.05	4.30
<i>rng2</i>	IQGAP	0.75	0.49	0.80	0.33



		Fold change in <i>Δcid1</i>	Effect change in <i>Δcid1</i>	Fold change in <i>Δdis3L2</i>	Effect change in <i>Δdis3L2</i>
<b>Mitotic sister chromatid segregation</b>					
<i>cnp3</i>	kinetochore protein, CENP-C ortholog Cnp3	0.89	0.71	0.57	0.28
<i>rpn2</i>	19S proteasome regulatory subunit Rpn2 (predicted)	1.04	1.34	1.31	5.02
<i>rpn7</i>	19S proteasome regulatory subunit Rpn7	1.01	1.06	1.25	3.78
<i>SPBC2G2.14</i>	mitotic centromere-SPB clustering protein Csi1	0.88	0.72	0.79	0.26
<i>ssl3</i>	cohesin loading factor Ssl3	0.93	0.71	0.79	0.34
<b>mRNA metabolic process</b>		Fold change in <i>Δcid1</i>	Effect change in <i>Δcid1</i>	Fold change in <i>Δdis3L2</i>	Effect change in <i>Δdis3L2</i>
<i>cdc28</i>	ATP-dependent RNA helicase Cdc28	1.67	1.91	1.83	2.85
<i>cft2</i>	cleavage factor two Cft2/polyadenylation factor CPSF- 73 (predicted)	1.32	2.54	1.75	5.59
<i>cwf21</i>	complexed with Cdc5 protein Cwf21	0.88	0.64	0.72	0.35
<i>edc3</i>	enhancer of mRNA decapping Edc3	0.74	0.25	0.98	0.93
<i>frg1</i>	FRG1 family protein, involved in mRNA splicing (predicted)	0.89	0.68	0.51	0.08
<i>mug161</i>	CwfJ family protein, splicing factor (predicted)	1.05	1.13	1.80	4.33
<i>prp10</i>	U2 snRNP-associated protein Sap155	0.82	0.35	0.92	0.61
<i>SPAC2C4.07c</i>	Dis3L2	0.87	0.75	0.09	0.04
<i>SPBC56F2.08c</i>	pumilio family RNA-binding protein Puf1 (predicted)	0.92	0.76	0.61	0.13
<b>Nucleobase-containing small molecule metabolic process</b>		Fold change in <i>Δcid1</i>	Effect change in <i>Δcid1</i>	Fold change in <i>Δdis3L2</i>	Effect change in <i>Δdis3L2</i>
<i>ade3</i>	phosphoribosylformylglycinamidine synthase Ade3	0.90	0.47	0.87	0.32
<i>adk1</i>	adenylate kinase activity	2.53	3.19	2.73	2.73
<i>cdc22</i>	ribonucleoside reductase large subunit Cdc22	1.19	1.89	1.68	3.29
<i>SPACUNK4.15</i>	2',3'-cyclic-nucleotide 3'- phosphodiesterase (predicted)	1.13	1.53	1.47	3.65
<i>SPBC839.16</i>	C1-5,6,7,8-tetrahydrofolate (THF) synthase, trifunctional enzyme Thf1	0.80	0.25	0.90	0.48
<i>suc22</i>	ribonucleotide reductase small subunit Suc22	1.10	3.72	1.12	1.87
<b>Nucleocytoplasmic transport</b>		Fold change in <i>Δcid1</i>	Effect change in <i>Δcid1</i>	Fold change in <i>Δdis3L2</i>	Effect change in <i>Δdis3L2</i>
<i>SPCC550.15c</i>	ribosome biogenesis protein Rei1 (predicted)	0.93	0.71	0.80	0.12
<b>Other</b>		Fold change in <i>Δcid1</i>	Effect change in <i>Δcid1</i>	Fold change in <i>Δdis3L2</i>	Effect change in <i>Δdis3L2</i>

<i>abp2</i>	ARS binding protein Abp2 centaurin ADOP ribosylation factor GTPase activating protein family	0.97	0.95	0.54	0.30
<i>cnt6</i>	(predicted) <i>Schizosaccharomyces</i> specific	0.76	0.31	0.76	0.25
<i>meu31</i>	protein Meu31 <i>Schizosaccharomyces</i> specific	1.08	1.12	0.39	0.31
<i>mug74</i>	protein Mug74 mitochondrial PPR repeat protein	4.02	3.53	1.91	2.14
<i>ppr3</i>	Ppr3 mitochondrial PPR repeat protein	0.82	0.49	0.74	0.34
<i>ppr5</i>	Ppr5 <i>Schizosaccharomyces</i> specific	0.76	0.46	0.70	0.32
<i>slt1</i>	protein Slt1 mitochondrial thioredoxin family	1.05	1.15	2.47	4.87
<i>SPAC11E3.12</i>	protein MTC tricarboxylate transmembrane	1.27	1.87	1.42	4.05
<i>SPAC17G6.15c</i>	transporter Fsf1 (predicted)	0.99	0.96	0.77	0.35
<i>SPAC22A12.06c</i>	serine hydrolase-like BSD domain protein, unknown	0.69	0.61	0.36	0.16
<i>SPAC22A12.14c</i>	biological role short chain dehydrogenase	1.39	2.02	1.63	3.71
<i>SPAC22A12.17c</i>	(predicted)	1.25	3.23	1.75	12.70
<i>SPAC27D7.09c</i>	But2 family protein <i>Schizosaccharomyces</i> specific	1.24	1.18	5.12	3.15
<i>SPAC27E2.11c</i>	protein	2.00	3.49	1.82	3.57
<i>SPAC2E1P3.05c</i>	fungal cellulose binding domain protein	0.90	0.68	1.85	8.22
<i>SPAC30C2.03</i>	<i>Schizosaccharomyces</i> specific protein	1.25	1.18	2.80	4.45
<i>SPAC513.02</i>	phosphoglycerate mutase family flavonol reductase/cinnamoyl-CoA	0.58	0.69	5.34	3.46
<i>SPAC513.07</i>	reductase family <i>S. pombe</i> specific 5Tm protein	1.00	1.00	1.66	3.13
<i>SPAC750.05c</i>	family <i>S. pombe</i> specific DUF999 protein	1.25	1.23	4.26	3.91
<i>SPAC750.06c</i>	family 4	1.56	1.59	2.44	3.77
<i>SPAC7D4.05</i>	hydrolase (predicted)	0.88	0.76	0.68	0.27
<i>SPAC977.15</i>	dienelactone hydrolase family	0.78	0.68	1.83	3.03
<i>SPAPB18E9.05c</i>	<i>Schizosaccharomyces pombe</i> specific protein	0.61	0.34	1.11	1.33
<i>SPAPB2B4.07</i>	ubiquitin family protein, human UBTD1 homolog	0.73	0.43	0.60	0.31
<i>SPBC216.01c</i>	protein phosphatase PP4 complex regulatory subunit 3 Psy2 (predicted)	0.85	0.35	0.81	0.34
<i>SPBC21B10.08c</i>	antibiotic biosynthesis monooxygenase-like domain (predicted)	1.11	1.59	1.86	4.38
<i>SPBC21C3.19</i>	SBDS family protein Rtc3 (predicted)	1.02	1.04	1.87	3.13
<i>SPBC29A10.17</i>	<i>Schizosaccharomyces</i> specific protein	0.77	0.27	0.82	0.37
<i>SPBC30D10.14</i>	dienelactone hydrolase family	1.09	1.20	2.71	4.76

	(predicted)				
<i>SPBC3B8.06</i>	conserved fungal protein	1.17	1.78	1.33	2.91
<i>SPBC651.04</i>	<i>Schizosaccharomyces</i> specific protein	0.99	0.97	1.50	4.80
<i>SPBC8E4.05c</i>	fumarate lyase superfamily, bacterial 3-carboxy-cis,cis-muconate cycloisomerase related	1.02	1.02	1.86	3.98
<i>SPBP8B7.32</i>	dubious	1.22	1.45	0.54	0.20
<i>SPBPB21E7.08</i>	Unassigned	1.07	1.08	2.95	4.70
<i>SPBPB2B2.08</i>	conserved fungal protein	1.42	1.10	7.42	3.20
<i>SPCC1235.01</i>	<i>Schizosaccharomyces</i> specific protein	0.94	0.82	1.94	3.70
<i>SPCC1322.09</i>	conserved fungal protein	1.06	1.30	1.36	3.77
<i>SPCC1682.08c</i>	pumilio family RNA-binding protein Mcp2	0.83	0.54	0.64	0.33
<i>SPCC569.03</i>	mug2/mug135/meu2 family short chain dehydrogenase	0.95	0.86	1.87	5.54
<i>SPCC663.09c</i>	(predicted)	1.06	1.26	1.67	4.04
<i>SPNCRNA.1436</i>	box H/ACA small nucleolar RNA snR95	1.16	2.92	1.02	1.10
<i>vps901</i>	guanyl-nucleotide exchange factor Vps902	0.99	0.93	0.85	0.28
<b>Protein catabolic process</b>		<b>Fold change in <math>\Delta cid1</math></b>	<b>Effect change in <math>\Delta cid1</math></b>	<b>Fold change in <math>\Delta dis3L2</math></b>	<b>Effect change in <math>\Delta dis3L2</math></b>
<i>bip1</i>	ER heat shock protein BiP	1.10	2.21	1.40	5.70
<i>cdc48</i>	AAA family ATPase involved in ubiquitin-mediated protein degradation Cdc48	1.10	2.11	1.81	5.03
<i>fub2</i>	PI31 proteasome regulator Fub2 (predicted)	0.86	0.47	0.78	0.24
<i>nta1</i>	protein N-terminal amidase Nta1 (predicted)	0.34	0.25	0.83	0.75
<i>rpt1</i>	19S proteasome regulatory subunit Rpt1 (predicted)	1.15	1.99	1.34	3.78
<i>rpt3</i>	19S proteasome regulatory subunit Rpt3 (predicted)	1.14	1.61	1.42	3.89
<i>rpt6</i>	19S proteasome regulatory subunit Rpt6 (predicted)	1.00	0.97	1.23	3.37
<i>ubp15</i>	ubiquitin C-terminal hydrolase Ubp15	1.08	1.19	1.62	3.24
<b>Protein complex assembly</b>		<b>Fold change in <math>\Delta cid1</math></b>	<b>Effect change in <math>\Delta cid1</math></b>	<b>Fold change in <math>\Delta dis3L2</math></b>	<b>Effect change in <math>\Delta dis3L2</math></b>
<i>SPCC18.17c</i>	proteasome assembly chaperone (predicted)	0.77	0.36	0.66	0.24
<b>Protein folding</b>		<b>Fold change in <math>\Delta cid1</math></b>	<b>Effect change in <math>\Delta cid1</math></b>	<b>Fold change in <math>\Delta dis3L2</math></b>	<b>Effect change in <math>\Delta dis3L2</math></b>
<i>cct5</i>	chaperonin-containing T-complex epsilon subunit Cct5	1.50	1.76	2.13	3.20
<i>cdc37</i>	Hsp90 co-chaperone Cdc37	0.95	0.95	2.04	3.73
<i>cnx1</i>	calnexin Cnx1	1.18	1.85	1.58	3.28

<i>hsp90</i>	Hsp90 chaperone mitochondrial heat shock protein	1.64	2.58	2.29	6.48
<i>mcp60</i>	Hsp60/Mcp60	1.40	2.02	1.67	3.17
<i>SPBC16D10.08c</i>	heat shock protein Hsp104	1.60	2.85	3.96	5.18
<i>SPBC1711.08</i>	chaperone activator Aha1 mitochondrial heatshock protein	1.63	3.50	2.20	7.46
<i>SPBC4F6.17c</i>	Hsp78 (predicted)	1.21	2.01	2.69	10.07
<i>ssa1</i>	heat shock protein Ssa1 (predicted)	1.36	1.84	3.08	5.16
<i>ssa2</i>	heat shock protein Ssa2 mitochondrial heat shock protein	1.63	2.13	2.88	4.96
<i>ssc1</i>	Hsp70	1.38	2.34	1.88	5.15
<i>sti1</i>	chaperone activator Sti1 (predicted) p23 homolog, predicted co- chaperone	1.46	2.51	2.78	5.34
<i>wos2</i>	Wos2	1.22	1.62	1.74	3.15
<b>Protein glycosylation</b>		<b>Fold change in <math>\Delta cid1</math></b>	<b>Effect change in <math>\Delta cid1</math></b>	<b>Fold change in <math>\Delta dis3L2</math></b>	<b>Effect change in <math>\Delta dis3L2</math></b>
<i>gma12</i>	alpha-1,2-galactosyltransferase Gma12	1.05	1.16	0.78	0.30
<b>Protein maturation</b>		<b>Fold change in <math>\Delta cid1</math></b>	<b>Effect change in <math>\Delta cid1</math></b>	<b>Fold change in <math>\Delta dis3L2</math></b>	<b>Effect change in <math>\Delta dis3L2</math></b>
<i>grx5</i>	monothiol glutaredoxin Grx5 BRO1 domain protein Rim20 (predicted)	0.85	0.57	0.73	0.32
<i>SPAC2G11.05c</i>	(predicted)	0.80	0.68	0.64	0.35
<b>Protein modification by small protein conjugation or removal</b>		<b>Fold change in <math>\Delta cid1</math></b>	<b>Effect change in <math>\Delta cid1</math></b>	<b>Fold change in <math>\Delta dis3L2</math></b>	<b>Effect change in <math>\Delta dis3L2</math></b>
<i>ubc4</i>	ubiquitin conjugating enzyme E2 Ubc4/UbcP1	1.44	2.98	1.14	1.39
<i>ulp2</i>	SUMO deconjugating cysteine peptidase Ulp2 (predicted)	0.80	0.34	0.74	0.23
<b>Protein targeting</b>		<b>Fold change in <math>\Delta cid1</math></b>	<b>Effect change in <math>\Delta cid1</math></b>	<b>Fold change in <math>\Delta dis3L2</math></b>	<b>Effect change in <math>\Delta dis3L2</math></b>
<i>kap109</i>	karyopherin Kap109 mitochondrial processing peptidase (MPP) complex beta subunit Mas1 (predicted)	0.98	0.94	1.30	3.38
<i>qcr1</i>	peroxisomal membrane protein Pex22 (predicted)	0.85	0.25	1.08	1.39
<i>SPAC19D5.02c</i>	TIM23 translocase complex subunit Tim21 (predicted)	2.56	31.96	1.01	1.03
<i>tim21</i>	Tim21 (predicted)	1.19	1.96	1.39	6.63
<i>tlh2</i>	RecQ type DNA helicase Tlh1	1.18	1.16	0.21	0.29
<b>Regulation of mitotic cell cycle phase transition</b>		<b>Fold change in <math>\Delta cid1</math></b>	<b>Effect change in <math>\Delta cid1</math></b>	<b>Fold change in <math>\Delta dis3L2</math></b>	<b>Effect change in <math>\Delta dis3L2</math></b>
<i>wee1</i>	M phase inhibitor protein kinase Wee1	0.60	0.35	0.74	0.47
<b>Ribosome biogenesis</b>		<b>Fold change in <math>\Delta cid1</math></b>	<b>Effect change in <math>\Delta cid1</math></b>	<b>Fold change in <math>\Delta dis3L2</math></b>	<b>Effect change in <math>\Delta dis3L2</math></b>

<i>fib1</i>	fibrillarin, rRNA methyltransferase mitochondrial 2' O-ribose	2.70	2.59	2.40	3.06
<i>mrm2</i>	methyltransferase Mrm2 (predicted)	2.48	3.22	1.30	1.29
<i>SPAPB8E5.07c</i>	rRNA processing protein Rrp12 (predicted)	0.92	0.56	0.66	0.08
<i>SPBC13G1.09</i>	bystin family U3 and U14 snoRNA associated protein Enp1 (predicted)	0.88	0.41	0.65	0.05
<i>SPBP8B7.10c</i>	U3 snoRNP-associated protein Utp16 (predicted)	0.86	0.71	0.61	0.33
<i>SPCC830.09c</i>	RNase P and RNase MRP subunit (predicted)	1.00	0.99	0.75	0.27
<i>SPCP1E11.08</i>	ribosome biogenesis protein Nsa2 (predicted)	1.06	1.32	0.73	0.33
<i>utp20</i>	U3 snoRNP protein Utp20 (predicted)	0.89	0.40	0.78	0.11
<b>Signaling</b>		<b>Fold change in <math>\Delta cid1</math></b>	<b>Effect change in <math>\Delta cid1</math></b>	<b>Fold change in <math>\Delta dis3L2</math></b>	<b>Effect change in <math>\Delta dis3L2</math></b>
<i>efc25</i>	Ras1 guanyl-nucleotide exchange factor Efc25	0.81	0.83	0.64	0.34
<i>gap1</i>	GTPase activating protein Gap1 neuronal calcium sensor related	0.90	0.75	0.61	0.12
<i>ncs1</i>	protein Ncs1	0.59	0.20	0.67	0.35
<i>ptc2</i>	protein phosphatase 2C Ptc2	0.97	0.85	0.77	0.24
<i>ric1</i>	Ypt/Rab-specific guanyl-nucleotide exchange factor (GEF) subunit Ric1	0.53	0.23	1.26	1.28
<b>snoRNA metabolic process</b>		<b>Fold change in <math>\Delta cid1</math></b>	<b>Effect change in <math>\Delta cid1</math></b>	<b>Fold change in <math>\Delta dis3L2</math></b>	<b>Effect change in <math>\Delta dis3L2</math></b>
<i>cid14</i>	TRAMP complex poly(A) polymerase subunit Cid14	0.89	0.58	0.74	0.26
<b>Telomere organization</b>		<b>Fold change in <math>\Delta cid1</math></b>	<b>Effect change in <math>\Delta cid1</math></b>	<b>Fold change in <math>\Delta dis3L2</math></b>	<b>Effect change in <math>\Delta dis3L2</math></b>
<i>ssb1</i>	DNA replication factor A subunit Ssb1	1.09	1.64	1.48	6.48
<i>tcg1</i>	single-stranded telomeric binding protein Tgc1	1.30	2.34	1.76	3.60
<b>Transmembrane transport</b>		<b>Fold change in <math>\Delta cid1</math></b>	<b>Effect change in <math>\Delta cid1</math></b>	<b>Fold change in <math>\Delta dis3L2</math></b>	<b>Effect change in <math>\Delta dis3L2</math></b>
<i>amt2</i>	ammonium transmembrane transporter Amt2	0.98	0.94	0.78	0.25
<i>atp2</i>	F1-ATPase beta subunit Atp2	0.97	0.55	1.29	4.29
<i>cch1</i>	calcium ion channel Cch1	0.86	0.53	0.83	0.30
<i>mrs2</i>	magnesium ion transmembrane transporter Mrs2 (predicted)	0.82	0.53	0.66	0.29
<i>per1</i>	plasma membrane amino acid permease Per1	0.73	0.50	0.63	0.31
<i>SPAC12G12.07c</i>	conserved fungal protein	1.20	3.03	1.42	4.71
<i>SPAC23H3.12c</i>	mitochondrial hydrogen/potassium transport system protein (predicted)	0.75	0.58	0.58	0.25
<i>SPAC24H6.11c</i>	sulfate transmembrane transporter	0.86	0.64	0.68	0.30

	(predicted)				
<i>SPAC30D11.06c</i>	Lazarus1 family transmembrane transporter	0.76	0.56	0.64	0.34
<i>SPAC323.07c</i>	MatE family transmembrane transporter (predicted)	0.74	0.58	0.58	0.24
<i>SPAC328.09</i>	mitochondrial 2-oxoadipate and 2-oxoglutarate transmembrane transporter (predicted)	0.79	0.44	0.68	0.26
<i>SPBPB2B2.01</i>	amino acid transmembrane transporter (predicted)	0.78	0.73	2.44	2.92
<i>SPCC1235.11</i>	mitochondrial pyruvate transmembrane transporter subunit Mpc1 (predicted)	0.73	0.50	0.54	0.21
<i>SPCC320.08</i>	transmembrane transporter (predicted)	1.03	1.13	0.73	0.31
<i>SPCC4B3.13</i>	MatE family transmembrane transporter (predicted)	1.12	1.44	1.51	3.52
<i>SPCC965.11c</i>	amino acid transmembrane transporter (predicted)	0.81	0.70	0.69	0.29
<i>str1</i>	siderophore-iron transmembrane transporter Str1	1.09	1.16	2.41	4.77
<i>vma6</i>	V-type ATPase V0 subunit d (predicted)	0.84	0.30	0.94	0.66
<b>tRNA metabolic process</b>		<b>Fold change in <math>\Delta</math>cid1</b>	<b>Effect change in <math>\Delta</math>cid1</b>	<b>Fold change in <math>\Delta</math>dis3L2</b>	<b>Effect change in <math>\Delta</math>dis3L2</b>
<i>pgp1</i>	mitochondrial metalloproteinase, tRNA N6-threonyl-carbamoyl-adenosine (t6A), a modification protein Pgp1	0.65	0.34	0.66	0.38
<i>SPBC16A3.06</i>	tRNA specific adenosine-37 deaminase Tad1 (predicted)	1.57	2.35	1.74	3.30
<i>SPCC63.07</i>	tRNA guanylyltransferase Thg1 (predicted)	1.53	2.94	1.28	2.11
<i>vrs2</i>	mitochondrial valine-tRNA ligase Vrs2/Vas2	0.77	0.33	0.84	0.47
<b>Vesicle-mediated transport</b>		<b>Fold change in <math>\Delta</math>cid1</b>	<b>Effect change in <math>\Delta</math>cid1</b>	<b>Fold change in <math>\Delta</math>dis3L2</b>	<b>Effect change in <math>\Delta</math>dis3L2</b>
<i>chc1</i>	clathrin heavy chain Chc1 (predicted)	0.91	0.61	1.24	5.08
<i>cog1</i>	Golgi transport complex subunit Cog1 (predicted)	0.77	0.29	0.91	0.48
<i>end4</i>	Huntingtin-interacting protein homolog	0.92	0.57	1.15	3.07
<i>sec26</i>	coatamer beta subunit (predicted)	1.08	1.43	1.29	3.39
<i>SPBC18H10.20c</i>	arrestin-related endocytic adaptor Any1	1.00	1.02	0.74	0.15
<i>SPBC4.03c</i>	COPII-coated vesicle component Sfb3 (predicted)	1.19	1.73	1.32	2.83
<i>SPBPJ4664.04</i>	coatamer alpha subunit Cop1 (predicted)	1.03	1.34	1.22	7.41
<i>SPCC18.15</i>	WD repeat protein, involved in diphthamide biosynthesis Dph7 (predicted)	1.87	1.74	2.60	6.65

<i>SPCC970.06</i>	COP II adaptor Erv29 (predicted)	1.18	2.36	1.23	3.33
<b>Vitamin metabolic process</b>		<b>Fold change in <math>\Delta cid1</math></b>	<b>Effect change in <math>\Delta cid1</math></b>	<b>Fold change in <math>\Delta dis3L2</math></b>	<b>Effect change in <math>\Delta dis3L2</math></b>
<i>SPCC4G3.16</i>	CMP deaminase family/ methyltransferase bifunctional enzyme involved in riboflavin biosynthesis and tRNA pseudouridine biosynthesis Rib2 (predicted)	0.63	0.30	0.66	0.30

**Table B3: Genes up-regulated at least 1.8-fold in the *S. pombe*  $\Delta cid1$  deletion strain compared to wildtype *S. pombe*.**

Gene	Protein name	Fold up in $\Delta cid1$
<i>SPAC19D5.02c</i>	peroxisomal membrane protein Pex22	5.00
<i>rps1101</i>	40S ribosomal protein S11	2.57
<i>SPBC23E6.02</i>	ATP-dependent DNA helicase	2.16
<i>SPNCRNA.636</i>	antisense RNA (predicted)	2.10
<i>SPAC806.06c</i>	nicotinamide mononucleotide (NMN) adenylyltransferase	2.05
<i>erg10</i>	acetylCoA C acetyltransferase Erg10	1.97
<i>suc22</i>	ribonucleotide reductase small subunit Suc22	1.90
<i>tif212</i>	translation initiation factor eIF2 beta subunit	1.82
<i>mug74</i>	sequence orphan; with a role in meiosis	1.82
<i>rik1</i>	Silencing protein Rik1, component of the rik1-associated E3 ubiquitin ligase complex	1.82
<i>SPBC1711.08</i>	chaperone activator Aha1	1.81
<i>SPAC27E2.11c</i>	sequence orphan	1.80

**Table B4: Genes down-regulated at least 1.8-fold in the *S. pombe*  $\Delta cid1$  deletion strain compared to wildtype *S. pombe*.**

Gene	Protein name	Fold Down in $\Delta cid1$
<i>SPAC869.02c</i>	nitric oxide dioxygenase Yhb1	3.15
<i>nsc1</i>	neuronal calcium sensor related protein Ncs1	2.30
<i>ric1</i>	Ypt/Rabspecific guanylnucleotide exchange factor (GEF) subunit Ric1	2.13

<i>SPCC70.03c</i>	proline dehydrogenase	2.04
<i>SPNCRNA.1548</i>	SPBC336.13c-antisense-1	2.03
<i>nta1</i>	protein Nterminal amidase Nta1	2.02
<i>qcr1</i>	mitochondrial processing peptidase (MPP) complex beta subunit Mas1	2.02
<i>SPBC839.16</i>	C15,6,7,8tetrahydrofolate (THF) synthase, trifunctional enzyme Thf1	2.01
<i>edc3</i>	enhancer of mRNA decapping Edc3	2.00
<i>SPNCRNA.1212</i>	antisense RNA (predicted)	1.97
<i>SPBC1683.04</i>	glycosyl hydrolase family 3	1.93
<i>SPCC548.05c</i>	ubiquitinprotein ligase E3 Dbp5	1.88
<i>SPBC29A10.17</i>	<i>Schizosaccharomyces</i> specific protein	1.86

**Table B5: Genes down-regulated at least 1.8-fold in the *S. pombe*  $\Delta$ *dis3L2* deletion strain compared to wildtype *S. pombe*.**

Gene	Protein name	Fold Down in $\Delta$ <i>dis3L2</i>
<i>SPBC13G1.09</i>	bystin family U3 and U14 snoRNA associated protein Enp1	4.29
<i>frg1</i>	FRG1 family protein, involved in mRNA splicing	3.63
<i>SPAPB8E5.07c</i>	rRNA processing protein Rrp12	3.63
<i>leu2</i>	3-isopropylmalate dehydratase Leu2	3.31
<i>SPBC19F5.04</i>	aspartate kinase	3.30
<i>utp20</i>	U3 snoRNP protein Utp20	3.16
<i>gpd2</i>	glycerol-3-phosphate dehydrogenase Gpd2	3.10
<i>gap1</i>	GTPase activating protein Gap1	3.08
<i>SPCC550.15c</i>	ribosome biogenesis protein Rei1	3.02
<i>SPNCRNA.1451</i>	cdc28-antisense-1 RNA	2.93
<i>SPBC56F2.08c</i>	pumilio family RNA-binding protein Puf1	2.91
<i>SPNCRNA.1467</i>	mdl1-antisense-1 RNA	2.86
<i>SPBC18H10.20c</i>	arrestin-related endocytic adaptor Any1	2.78
<i>adh4</i>	alcohol dehydrogenase Adh4	2.68
<i>SPNCRNA.1138</i>	antisense RNA (predicted)	2.64
<i>SPNCRNA.1297</i>	intergenic RNA (predicted), possible alternative UTR	2.64
<i>SPNCRNA.1548</i>	SPBC336.13c-antisense-1 RNA	2.62
<i>SPAC22A12.06c</i>	serine hydrolase-like	2.61



<i>rrn3</i>	ribosomal DNA (rDNA) transcription factor Rrn3	2.60
<i>SPNCRNA.935</i>	intergenic RNA (predicted), possible alternative UTR	2.41
<i>tea3</i>	cell end marker Tea3	2.40
<i>mcp7</i>	meiosis specific coiled-coil protein Mcp7	2.40
<i>arg1</i>	acetylornithine aminotransferase	2.34
<i>SPBP8B7.32</i>	Unassigned	2.34
<i>SPCC1235.11</i>	mitochondrial pyruvate transmembrane transporter subunit Mpc1	2.24
<i>SPNCRNA.1204</i>	rsm1-antisense-1 RNA	2.21
<i>SPBC2G5.02c</i>	CK2 family regulatory subunit Ckb2	2.17
<i>SPNCRNA.579</i>	ggc1-antisense-1 RNA	2.13
<i>ulp2</i>	SUMO deconjugating cysteine peptidase Ulp2	2.12
<i>SPCC306.06c</i>	ER membrane protein, BIG1 family mitochondrial iron-sulfur cluster assembly	2.11
<i>SPBC1709.19c</i>	protein Nfu1	2.10
<i>fub2</i>	PI31 proteasome regulator Fub2	2.09
<i>SPAC323.07c</i>	MatE family transmembrane transporter	2.07
<i>ptc2</i>	protein phosphatase 2C Ptc2	2.05
<i>SPCC18.17c</i>	proteasome assembly chaperone	2.04
<i>SPNCRNA.606</i>	isp3-antisense-1 RNA	2.01
<i>amt2</i>	ammonium transmembrane transporter Amt2	2.00
<i>cob</i>	cytochrome b	2.00
<i>cnt6</i>	centaurin ADOP ribosylation factor GTPase activating protein family	2.00
<i>dna2</i>	DNA replication endonuclease-helicase Dna2	1.98
<i>SPAC23H3.12c</i>	mitochondrial hydrogen/potassium transport system protein	1.98
<i>cid14</i>	TRAMP complex poly(A) polymerase subunit Cid14	1.97
<i>SPAC328.09</i>	mitochondrial 2-oxoadipate and 2-oxoglutarate transmembrane transporter	1.96
<i>SPBC2G2.14</i>	mitotic centromere-SPB clustering protein Csi1	1.95
<i>SPCC548.05c</i>	ubiquitin-protein ligase E3 Dbl5	1.95
<i>swd3</i>	WD repeat protein Swd3	1.93
<i>pso2</i>	DNA 5' exonuclease	1.92

<i>cdc23</i>	MCM-associated protein Mcm10	1.92
<i>SPCC830.09c</i>	RNase P and RNase MRP subunit	1.90
<i>SPAC7D4.05</i>	hydrolase	1.88
<i>vps901</i>	guanyl-nucleotide exchange factor Vps902	1.86
<i>SPBC2G2.17c</i>	beta-glucosidase Psu2	1.85
<i>cnp3</i>	kinetochore protein, CENP-C ortholog Cnp3	1.84
<i>SPBC582.04c</i>	RNAi protein, Dsh1	1.83
<i>dph2</i>	diphthamide biosynthesis protein	1.82

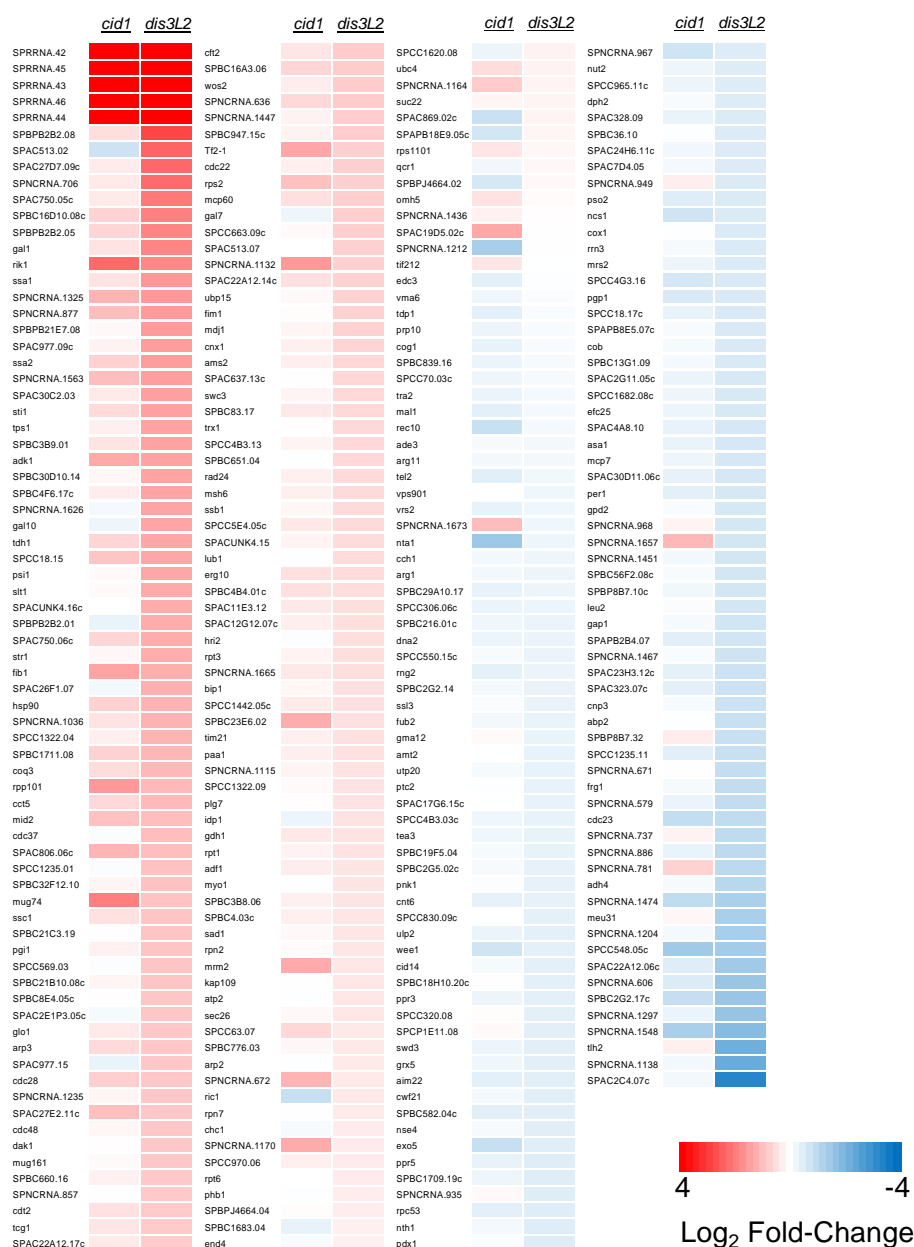
**Table B6: Genes up-regulated at least 1.8-fold in the *S. pombe*  $\Delta$ *dis3L2* deletion strain compared to wildtype *S. pombe*.**

<b>Heat shock proteins/chaperones</b>		<b>Fold up in <math>\Delta</math><i>dis3L2</i></b>
<i>SPBC4F6.17c</i>	mitochondrial heatshock protein Hsp78	3.33
<i>SPBC1711.08</i>	chaperone activator Aha1	2.90
<i>psi1</i>	DNAJ domain protein Psi1	2.87
<i>tim21</i>	TIM23 translocase complex subunit Tim21	2.73
<i>hsp90</i>	Hsp90 chaperone	2.70
<i>bip1</i>	ER heat shock protein BiP	2.51
<i>SPCC569.03</i>	mug2/mug135/meu2 family	2.47
<i>sti1</i>	chaperone activator Sti1	2.42
<i>SPBC16D10.08c</i>	heat shock protein Hsp104	2.37
<i>ssa1</i>	heat shock protein Ssa1	2.37
<i>ssc1</i>	mitochondrial heat shock protein Hsp70	2.36
<i>ssa2</i>	heat shock protein Ssa2	2.31
<i>cdc37</i>	Hsp90 co-chaperone Cdc37	1.90
<i>SPBC3B9.01</i>	Hsp70 nucleotide exchange factor Fes1	2.12
<b>Metabolism</b>		<b>Fold change</b>
<i>SPAC22A12.17c</i>	short chain dehydrogenase	3.67
<i>SPAC977.09c</i>	phospholipase	3.42
<i>SPBC947.15c</i>	mitochondrial NADH dehydrogenase (ubiquinone) Nde1	3.40
<i>gal1</i>	galactokinase Gal1	3.09
<i>SPAC2E1P3.05c</i>	fungal cellulose binding domain protein	3.04

<i>gdh1</i>	NADP-specific glutamate dehydrogenase Gdh1	2.82
<i>SPCC18.15</i>	WD repeat protein, involved in diphthamide biosynthesis Dph7	2.73
<i>SPAC806.06c</i>	nicotinamide mononucleotide (NMN) adenylyltransferase	2.71
<i>gal10</i>	UDP-glucose 4-epimerase/aldose 1-epimerase Gal10	2.60
<i>SPACUNK4.16c</i>	alpha,alpha-trehalose-phosphate synthase	2.40
<i>SPBC660.16</i>	phosphogluconate dehydrogenase, decarboxylating	2.29
<i>glo1</i>	glyoxalase I	2.28
<i>SPBC30D10.14</i>	dienelactone hydrolase family	2.25
<i>pgi1</i>	glucose-6-phosphate isomerase	2.25
<i>SPCC1322.04</i>	UTP-glucose-1-phosphate uridylyltransferase	2.22
<i>SPAC26F1.07</i>	glucose 1-dehydrogenase (NADP+)	2.21
<i>SPBC21B10.08c</i>	antibiotic biosynthesis monooxygenase-like domain	2.13
<i>erg10</i>	acetyl-CoA C-acetyltransferase Erg10	2.12
<i>idp1</i>	isocitrate dehydrogenase Idp1	2.11
<i>atp2</i>	F1-ATPase beta subunit Atp2	2.10
<i>tps1</i>	alpha,alpha-trehalose-phosphate synthase [UDP-forming]	2.03
<i>SPCC663.09c</i>	short chain dehydrogenase (predicted)	2.02
<i>SPBC8E4.05c</i>	fumarate lyase superfamily, bacterial 3-carboxy-cis,cis-muconate cycloisomerase related	1.99
<i>plg7</i>	phospholipase A2, PAF family homolog	1.96
<i>SPBC32F12.10</i>	phosphoglucomutase	1.91
<i>SPCC5E4.05c</i>	mitochondrial acylglycerol lipase Mgl1	1.88
<i>SPACUNK4.15</i>	2',3'-cyclic-nucleotide 3'-phosphodiesterase	1.87
<i>gal7</i>	galactose-1-phosphate uridylyltransferase Gal7	1.81
<b>Oxidative Stress</b>		<b>Fold change</b>
<i>trx1</i>	cytosolic thioredoxin Trx1	2.81
<i>SPAC11E3.12</i>	mitochondrial thioredoxin family protein	2.02
<b>Protein degradation</b>		<b>Fold change</b>
<i>cdc48</i>	AAA family ATPase involved in ubiquitin-mediated protein degradation Cdc48	2.33
<i>rpn2</i>	19S proteasome regulatory subunit Rpn2	2.33

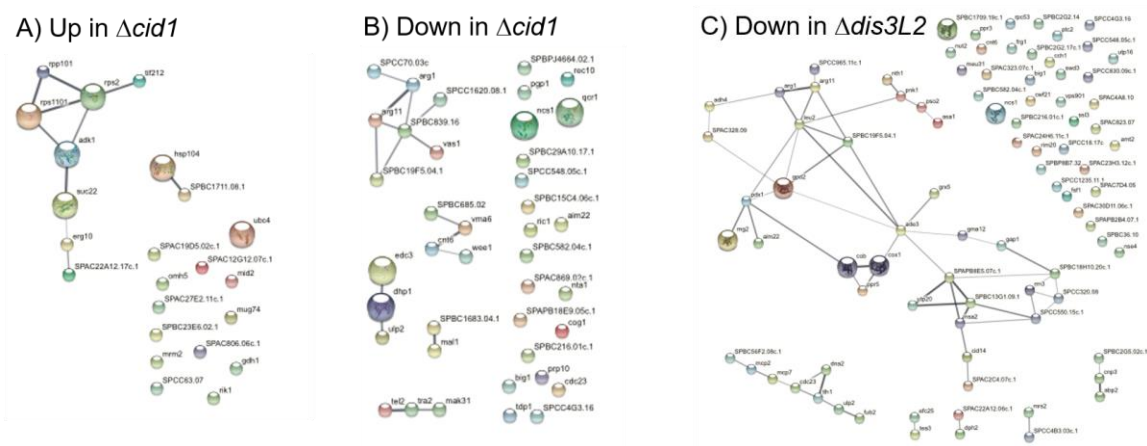
<i>rpt3</i>	19S proteasome regulatory subunit Rpt3	1.96
<i>rpt1</i>	19S proteasome regulatory subunit Rpt1	1.92
<i>rpn7</i>	19S proteasome regulatory subunit Rpn7	1.92
<b>Vesicles</b>		<b>Fold change</b>
<i>SPBPJ4664.04</i>	coatomer alpha subunit Cop1 (predicted)	2.89
<i>chc1</i>	clathrin heavy chain Chc1	2.35
<b>Antisense RNA</b>		<b>Fold change</b>
<i>SPNCRNA.1036</i>	wis2-antisense-1 RNA	3.89
<i>SPNCRNA.1447</i>	uds1- antisense RNA	3.10
<i>SPNCRNA.1626</i>	antisense RNA	2.84
<b>Cell cycle</b>		<b>Fold change</b>
<i>ssb1</i>	DNA replication factor A subunit Ssb1	2.69
<i>cft2</i>	cleavage factor two Cft2/polyadenylation factor CPSF-73	2.48
<i>msh6</i>	MutS protein homolog	2.37
<i>SPBC651.04</i>	<i>Schizosaccharomyces</i> specific protein	2.26
<i>paa1</i>	protein phosphatase regulatory subunit Paa1	2.15
<i>mid2</i>	medial ring protein Mid2	2.11
<i>ams2</i>	cell cycle regulated GATA-type transcription factor Ams2	2.03
<i>cdt2</i>	WD repeat protein Cdt2	2.00
<i>tcg1</i>	single-stranded telomeric binding protein Tgc1	1.85
<i>rad24</i>	14-3-3 protein Rad24	1.83
<b>Cytoskeleton</b>		<b>Fold change</b>
<i>myo1</i>	myosin type I	2.61
<i>arp2</i>	ARP2/3 actin-organizing complex subunit Arp2	2.50
<i>fim1</i>	fimbrin	2.20
<i>adf1</i>	actin depolymerizing factor, cofilin	1.92
<i>arp3</i>	Arp2/3 protein complex, actin-like protein subunit Arp3	1.91
<b>Other</b>		<b>Fold change</b>
<i>SPBPB2B2.05</i>	peptidase family C26 protein	2.52
<i>slt1</i>	<i>Schizosaccharomyces</i> specific protein Sl1	2.28

<i>str1</i>	siderophore-iron transmembrane transporter Str1	2.25
<i>SPAC12G12.07c</i>	conserved fungal protein	2.24
<i>SPBPB21E7.08</i>	Unassigned	2.23
<i>SPAC30C2.03</i>	Schizosaccharomyces specific protein	2.15
<i>mug161</i>	CwfJ family protein, splicing factor	2.11
<i>hri2</i>	eIF2 alpha kinase Hri2	2.02
<i>SPAC750.05c</i>	<i>S. pombe</i> specific 5Tm protein family	1.97
<i>SPBC83.17</i>	transcriptional coactivator, multiprotein bridging factor Mbf1	1.92
<i>SPAC750.06c</i>	<i>S. pombe</i> specific DUF999 protein family 4	1.92
<i>SPCC1322.09</i>	conserved fungal protein	1.91
<i>SPCC1442.05c</i>	MICOS complex subunit Mic26/27	1.90
<i>SPAC22A12.14c</i>	BSD domain protein, unknown biological role	1.89
<i>SPCC1235.01</i>	Unassigned	1.89
<i>SPAC27E2.11c</i>	<i>Schizosaccharomyces</i> specific protein	1.84
<i>SPCC4B3.13</i>	MatE family transmembrane transporter	1.81



**Figure B1: Colour map of gene expression changes between wildtype *S. pombe* and  $\Delta cid1$  or  $\Delta dis3L2$  strains.**

Rank-ordered false colour map showing the fold-change in gene expression (relative to the WT strain) for all genes identified as differentially expressed in either  $\Delta cid1$  or  $\Delta dis3L2$  strains. An effect size cut off of 1.5 was applied to identify differentially expressed genes in each mutant strain.



**Figure B2: Search tool for the retrieval of interacting genes/proteins (STRING) diagram of RNAs with altered expression levels in *cid1* and *dis3L2* deletion strains compared to wildtype.**

Expression of genes encoding respective proteins displayed are down-regulated more than 1.5-fold. The diagram was generated using the STRING database. Functional associations between proteins are shown, with confidence of the proposed association denoted by line thickness. **A)** RNAs up-regulated in a  $\Delta cid1$  strain compared to wildtype *S. pombe* cells. **B)** RNAs down-regulated in a  $\Delta cid1$  strain compared to wildtype *S. pombe* cells. **C)** RNAs down-regulated in a  $\Delta dis3L2$  strain compared to wildtype *S. pombe* cells. No significant functional enrichments were detected.

## Curriculum Vitae

**Name:** Christina Z. Chung

**Post-secondary Education and Degrees:** The University of Western Ontario, London, ON  
2014-2019 Ph.D.  
Department of Biochemistry  
Supervisor: Ilka U. Heinemann

The University of Western Ontario, London, ON  
2010-2014 B.M.Sc.  
Honors Specialization in Clinical Biochemistry with Distinction

**Honours and Awards:** A representative for the University of Western Ontario (Schulich School of Medicine & Dentistry – Biochemistry) at the 32<sup>nd</sup> Canadian Student Health Research Forum and recipient of a CIHR Silver Poster Award  
Winnipeg, Canada  
June 10-14, 2019

Natural Sciences and Engineering Council (NSERC)  
Alexander Graham Bell Canada Graduate Scholarship – Doctoral  
2017-2020

Queen Elizabeth II Graduate Scholarship in Science and Technology (QEII-GSST)  
Doctoral Scholarship  
2017-2018 (Declined to accept NSERC)

Ontario Graduate Scholarship (OGS)  
Doctoral Scholarship  
2016-2017

Western Graduate Research Scholarship  
2014-2019

Dr. D. S. M. Haines Gold Medal (Honors Specialization in Clinical Biochemistry at The University of Western Ontario)  
2014

Dr. D. S. M. Haines Book Prize (Honors Specialization in Clinical Biochemistry at The University of Western Ontario)  
2014



Biochemistry 4485E Award (Honors Specialization in Clinical Biochemistry at The University of Western Ontario)  
2014

Undergraduate Dean's Honor List  
The University of Western Ontario  
2010-2014

The Western Scholarship of Distinction (Entrance Scholarship)  
The University of Western Ontario  
2010

**Teaching Assistant Experience:** Biochemistry 3385 – Human Biochemistry  
The University of Western Ontario  
2015-2016

Biochemistry 3390 – Advanced Methods for Biochemistry  
The University of Western Ontario  
2018

#### **Publications:**

**Chung CZ**, Balasuriya N, Manni E, Liu X, Li SSC, O'Donoghue P, & Heinemann IU. 2019. Gld2 activity is regulated by phosphorylation in the N-terminal domain. *RNA Biol.* doi:10.1080/15476286.2019.1608754.

Chen AW, Jayasinghe MI, **Chung CZ**, Rao BS, Kenana R, Heineman IU, & Jackman JE. 2019. The Role of 3' to 5' Reverse RNA Polymerization in tRNA Fidelity and Repair. *Genes (Basel)*. **10(3)**. doi: 10.3390/genes10030250.

**Chung CZ\***, Jaramillo JE\*, Ellis MJ, Bour DYN, Seidl LE, Jo DHS, Turk MA, Mann MR, Bi Y, Haniford DB, Duennwald ML, & Heinemann IU. 2018. RNA surveillance by uridylation-dependent RNA decay in *Schizosaccharomyces pombe*. *Nucleic Acids Res.* **47(6)**: 3045-3057 (\*These authors contributed equally).

Turk MA, **Chung CZ**, Manni E, Zukowski SA, Engineer A, Badakhshi Y, & Heinemann IU. 2018. MiRAR-miRNA Activity Reporter for Living Cells. *Genes (Basel)*. **9(6)**. doi: 10.3390/genes9060305.

Desai R\*, Kim, K\*, Buechsenschuetz HC, Chen AW, Bi Y, Mann MR, Turk MA, **Chung CZ**, & Heinemann IU. 2017. Minimal requirements for reverse polymerization and tRNA repair by tRNA<sup>His</sup> guanylyltransferase. *RNA Biol.* **15(4-5)**: 614-622 (\*These authors contributed equally).

**Chung CZ\***, Seidl LE\*, Mann MR\*, & Heinemann IU. 2017. Tipping the balance of RNA stability by 3' editing of the transcriptome. *Biochimica et Biophysica Acta*. **1861**: 2971-2979 (\*These authors contributed equally).

**Chung CZ**, Jo DHS, & Heinemann IU. 2016. Nucleotide specificity of the human terminal nucleotidyltransferase Gld2 (TUT2). *RNA N. Y. N.* **22(8)**: 1239-1249.