

**Human CNOT3 in Cell Proliferation
and Disease: Insights into
the Function of the Ccr4-Not Complex**

Ashkan Bakhtiari

**Thesis submitted to the University of Nottingham for the
degree of Doctor of Philosophy**

January 2025

*With them the Seed of Wisdom did I sow,
And with my own hand laboured it to grow:
And this was all the Harvest that I reap'd—
“I came like Water, and like Wind I go.”*

Omar Khayyam (1048–1131)

These lines are taken from Omar Khayyam's Rubaiyat, translated into English by Edward FitzGerald in the 19th century.

Declaration

Except where acknowledged in the text, I declare that this thesis is my own work and it is based on the research work that was undertaken by myself, Ashkan Bakhtiari, in the Gene Regulation & RNA Biology group, School of Pharmacy, Faculty of Science, University of Nottingham, UK.

Abstract

The Ccr4-Not complex, a conserved multi-subunit regulator of mRNA metabolism, transcription, and translation, plays a pivotal role in maintaining cellular homeostasis. The work described in this thesis investigated the role of human CNOT3, a core subunit of the Ccr4-Not complex. The protein contains two conserved domains linked by an intrinsically disordered region. The N-terminal domain can directly interact with the ribosome, while the C-terminal Not-Box region is required for incorporation into the Ccr4-Not complex, which involves interactions with the C-terminal domain of the CNOT1 subunit and the Not-Box region of CNOT2. To examine the structural and functional significance of CNOT3, stable HEK293 cell lines with inducible expression of siRNA-resistant CNOT3 variants were generated. These cell lines allowed dissection of CNOT3's role in cell proliferation by siRNA-mediated knockdown of endogenous CNOT3 combined with overexpression of siRNA-resistant CNOT3 variants. Western blot analyses confirmed that the CNOT3 variants studied were expressed at comparable levels, validating the functionality of the experimental system.

The use of AlamarBlue and crystal violet cell viability assays revealed that deletion of the N-terminal and C-terminal domains of CNOT3 failed to rescue cell proliferation following siRNA knockdown of endogenous CNOT3, highlighting their essential roles in efficient cell proliferation in the HEK293 cell model. Furthermore, co-immunoprecipitation experiments revealed that the C-terminal domain is essential for the incorporation of CNOT3 into the Ccr4-Not complex, whereas the N-terminal domain is not required for this interaction. These findings demonstrate the importance of the structural integrity of CNOT3 in promoting cellular growth and maintaining its incorporation into the complex. Mutations in CNOT3, particularly within its N-terminal domain, have been associated with T-cell acute lymphoblastic leukaemia (T-ALL) and neurodevelopmental disorders (NDDs). This research integrates computational tools, such as SIFT, PolyPhen-2, PROVEAN, and AlphaMissense, with experimental validation to evaluate the impact of missense mutations. While these mutations preserved the fundamental interactions within the Ccr4-Not complex, they appear to modulate its function through subtle mechanisms. Specifically, given the established role of the N-terminal domain in ribosome binding, these mutations may influence both regulatory pathways and translational efficiency.

In addition, mutations in CNOT2 were investigated, demonstrating that truncating mutations result in haploinsufficiency by disrupting its incorporation into the Ccr4-Not

complex, as shown through co-immunoprecipitation experiments using plasmid-based expression systems. In contrast, the missense mutation does not affect core interactions but may have subtler functional consequences that require further study. These findings contribute to understanding how mutations in *CNOT3* and *CNOT2* influence the structural and functional integrity of the Ccr4-Not complex. By providing a foundation for future investigations into their broader biological and pathological significance, this work emphasises the need for detailed functional studies to elucidate the precise mechanisms underlying these mutations' contributions to disease.

Acknowledgements

I would like to take this opportunity to express my deepest gratitude to my primary supervisor, Dr. Sebastiaan Winkler. His unwavering guidance, invaluable advice, and consistent encouragement throughout my PhD journey have been vital to my development as a researcher. Dr. Winkler has not only provided me with the expertise I needed to navigate the complexities of my research topic, but he has also instilled in me the confidence to explore innovative ideas and push the boundaries of my understanding. His mentorship has been instrumental in shaping this research and has significantly contributed to my growth as a scientist.

I am also sincerely thankful to my second supervisor, Dr. Claire Seedhouse, for her insightful feedback and constructive discussions. Her ability to challenge my thinking and encourage me to refine my arguments has greatly enriched my work. Dr. Seedhouse's support and willingness to share her extensive knowledge have been essential in helping me overcome various hurdles throughout my studies.

I extend my heartfelt appreciation to Dr. Hilary Collins, the Scientific Research Officer of the Gene Regulation and Biology group, whose assistance significantly contributed to the progress of this work.

I am also deeply grateful to Dr. David Heery, the head of Biomolecular Sciences, for his leadership.

I would like to take a moment to express my heartfelt appreciation for my friend, Alfredo Smart. Your wisdom and support have been invaluable throughout my time in the lab and during the process of writing my thesis. I am truly grateful for the countless hours you dedicated to proofreading my work, helping me enhance the clarity of my writing. Your offer to have me come and live with you in Cambridge was a testament to your generosity and genuine friendship, and I will never forget it.

I am thankful to Alexander Kondrashov for his patience, guidance, and support during challenging moments, especially when I struggled to understand complex concepts. His willingness to teach and clarify difficult questions has been invaluable.

I would also like to acknowledge Chris Roberts for his assistance in teaching cloning techniques and his support during this journey.

My warmest thanks go to every member of the Gene Regulation and Biology lab for their camaraderie, collaborative spirit, and insightful discussions.

I also extend my gratitude to the GSW group for their invaluable support, advice, and contributions, which have enriched this research in countless ways. Working alongside such talented individuals in a group has been a rewarding experience.

I wish to express my heartfelt thanks to my dear friends, Daniel Mohammadian and Farzad Barati, for being true friends during the difficult times of being away from Iran. Their kindness has been a source of comfort and strength.

Finally, my deepest and most heartfelt gratitude goes to my wife, Nazanin, and my little daughter, Peransa, who was born just as I started my course in February 2021. Her arrival marked the beginning of both a challenging and profoundly meaningful chapter of my life. Their love, patience, and support have been my greatest source of strength throughout this journey.

TABLE OF CONTENTS

DECLARATION	I
ABSTRACT	II
ACKNOWLEDGEMENTS.....	IV
LIST OF FIGURES.....	XIII
LIST OF TABLES	XVI
LIST OF ABBREVIATIONS	XVIII
CHAPTER 1. INTRODUCTION	1
1.1 EUKARYOTIC GENE REGULATION.....	2
1.2 EPIGENETIC REGULATION	4
1.3 TRANSCRIPTION	4
1.3.1 TRANSCRIPTION INITIATION	4
1.3.2 TRANSCRIPTION ELONGATION.....	5
1.3.3 TRANSCRIPTION TERMINATION.....	6
1.4 POST-TRANSCRIPTIONAL REGULATION.....	6
1.4.1 MRNA CAPPING.....	7
1.4.2 MRNA SPLICING	7
1.4.3 MRNA POLYADENYLATION	7
1.4.4 MRNA NUCLEAR EXPORT	8
1.5 TRANSLATION	8
1.5.1 EUKARYOTIC RIBOSOME	8
1.5.2 TRANSLATION INITIATION	10
1.5.3 TRANSLATION ELONGATION	11
1.5.4 <i>TRANSLATION TERMINATION</i>	11
1.6 TRANSLATION REGULATION	12
1.7 CYTOPLASMIC MESSENGER RNA DECAY	17
1.8 DEADENYLATION: SHORTENING OF POLY (A) TAIL	20
1.8.1 THE PAN2/PAN3 COMPLEX	20
1.8.2 POLY (A)-SPECIFIC RIBONUCLEASE (PARN).....	23
1.8.3 THE CCR4-NOT COMPLEX.....	24

1.9 DECAPPING.....	26
1.9.1 THE DCP1-DCP2 COMPLEX.....	27
1.9.2 DECAPPING SCAVENGER (DCPS) ENZYME	27
1.9.3 THE LSM1-7-PAT1 COMPLEX	28
1.10 DEGRADATION	28
1.10.1 XRN1 ENZYME.....	28
1.10.2 EXOSOME ENZYME	30
1.11 OTHER MRNA DECAY PATHWAYS.....	30
1.11.1 NONSENSE-MEDIATED DECAY (NMD).....	30
1.11.2 NON-STOP DECAY	33
1.11.3 NO-GO DECAY	34
1.12 THE CCR4-NOT COMPLEX.....	35
1.13 THE N-TERMINAL MODULE.....	38
1.14 THE NUCLEASE MODULE	39
1.15 THE NOT MODULE	42
1.15.1 CNOT1 SUBUNIT	42
1.15.2 CNOT2-CNOT3 SUBUNITS	42
1.16 PHYSIOLOGICAL ROLE OF THE NOT MODULE	46
1.16.1 THE PHYSIOLOGICAL ROLE OF CNOT1	46
1.16.2 PHYSIOLOGICAL ROLE OF CNOT2-CNOT3	46
1.16.3 FUNCTION OF CNOT3 IN MRNA DECAY AND TRANSLATION.....	47
1.17 INVOLVEMENT OF CNOT3 IN DISEASES	51
1.17.1 INVOLVEMENT OF CNOT3 IN CANCER.....	51
1.17.2 INVOLVEMENT OF CNOT3 IN OTHER DISEASE.....	52
1.18 INVOLVEMENT OF CNOT2 IN DISEASES	54
1.19 AIM AND OBJECTIVES.....	56
CHAPTER 2. MATERIALS AND METHODS	58
2.1 BACTERIAL GROWTH AND TRANSFORMATION	59
2.1.1 REAGENTS, STOCK SOLUTIONS AND BUFFERS FOR BACTERIAL METHODS.....	59
2.1.2 CULTURE OF E. COLI DH5 α	59

2.1.3 TRANSFORMATION OF COMPETENT CELLS	59
2.2 MOLECULAR BIOLOGY	60
2.2.1 REAGENTS, STOCK SOLUTIONS AND BUFFERS FOR USE IN MOLECULAR BIOLOGY METHODS.	60
2.2.2 SMALL-SCALE PLASMID DNA PREPARATION.....	61
2.2.3 LARGE-SCALE PLASMID DNA PREPARATION	61
2.2.4 DETERMINATION OF DNA/RNA CONCENTRATION	61
2.2.5 AGAROSE GEL ELECTROPHORESIS.....	61
2.2.6 DNA EXTRACTION AND PURIFICATION FROM AGAROSE GEL.....	62
2.2.7 RESTRICTION ENZYME DIGESTION OF DNA	62
2.2.8 LIGATION OF DNA FRAGMENTS	62
2.2.9 POLYMERASE CHAIN REACTION (PCR).....	63
2.2.10 SITE-DIRECTED MUTAGENESIS (INVERSE PCR)	63
2.2.11 CONSTRUCTION OF INDUCED FLAG-CNOT3 EXPRESSION VECTORS	66
2.2.12 CONSTRUCTION OF EUKARYOTIC EXPRESSION PLASMIDS CONTAINING DISEASE- ASSOCIATED MUTATIONS IN CNOT3 AND CNOT2.....	68
2.2.13 COLONY PCR.....	68
2.2.14 DNA SEQUENCING	69
2.3 CELL CULTURE	69
2.3.1 REAGENTS, STOCK SOLUTIONS AND BUFFERS FOR USE IN CELL CULTURE	70
2.3.2 FREEZING CELLS	71
2.3.3 THAWING CELLS	71
2.3.4 DNA PLASMIDS TRANSIENT TRANSFECTION.....	71
2.3.5. GENERATION OF STABLE CELL LINES	72
2.3.6 DETERMINATION OF HYGROMYCIN KILL CURVE	73
2.3.7 DOXYCYCLINE-INDUCED EXPRESSION OF CNOT3 VARIANTS.....	73
2.3.8 siRNA TRANSFECTION OF HEK293 CELLS	74
2.3.9 CRYSTAL VIOLET ASSAY.....	76
2.3.10 ALAMARBLUE ASSAY	77
2.4 PROTEIN ANALYSIS	78
2.4.1 REAGENTS, STOCK SOLUTIONS AND BUFFERS FOR USE IN PROTEIN ANALYSIS	78
2.4.2 PREPARATION OF PROTEIN LYSATES FOR WESTERN BLOTTING	79
2.4.3 BRADFORD ASSAY TO DETERMINE PROTEIN CONCENTRATION.....	80
2.4.4 SODIUM DODECYL SULPHATE-POLYACRYLAMIDE GEL ELECTROPHORESIS (SDS-PAGE) ...	80
2.4.5 WESTERN BLOTTING AND IMMUNODETECTION.....	82
2.5 CO-IMMUNOPRECIPITATION.....	82

2.5.1 REGENTS, STOCK SOLUTIONS AND BUFFERS FOR USE IN CO-IMMUNOPRECIPITATION	82
2.5.2 PREPARATION OF PROTEIN LYSATE FOR CO-IMMUNOPRECIPITATION.....	83
2.5.3 CO-IMMUNOPRECIPITATION	84
2.6 COMPUTATIONAL WORK.....	84
2.6.1 PREDICTION OF AMINO ACID SUBSTITUTION EFFECTS USING POLYPHEN-2	84
2.6.2 PREDICTION OF MISSENSE VARIANT PATHOGENICITY USING ALPHAMISSENSE	87
2.6.3 PREDICTION OF AMINO ACID SUBSTITUTION EFFECTS USING SIFT	88
2.6.4 CHIMERA X SOFTWARE.....	90
2.7 STATISTICAL ANALYSIS.....	90
CHAPTER 3. GENERATION OF STABLE CELL LINES FOR THE CONDITIONAL EXPRESSION OF SIRNA-RESISTANT FLAG-CNOT3 VARIANTS.....	91
3.1 INTRODUCTION	92
3.2 CONSTRUCTION OF PLASMIDS EXPRESSING FLAG-CNOT3 VARIANTS.....	95
3.2.1 DESIGN OF SIRNA-RESISTANT CDNA EXPRESSION PLASMIDS	95
3.2.2 DESIGN OF CDNA VARIANTS ENCODING FLAG-CNOT3 LACKING N- AND C-TERMINAL DOMAINS	99
3.3 TRANSIENT EXPRESSION OF FLAG-CNOT3 MUTANTS IN HEK293 CELL LINE	102
3.4 GENERATING FLAG-CNOT3 CELL LINES USING G-418.....	104
3.5 CONSTRUCTION OF VECTORS FOR INDUCIBLE EXPRESSION OF FLAG-CNOT3 VARIANTS.....	104
3.6 GENERATION OF FLP-IN TREX HEK293 STABLE CELL LINES CONTAINING INDUCIBLE FLAG-CNOT3 VARIANTS.....	107
3.7 FLAG-CNOT3R1 AND FLAG-CNOT3R2 MRNAS ARE RESISTANT AGAINST SIRNA#1 AND #2, RESPECTIVELY.....	110
3.8 DISCUSSION	112
CHAPTER 4. ESSENTIAL DOMAINS OF CNOT3 FOR CELL PROLIFERATION	116
4.1 INTRODUCTION	117
4.2 DOXYCYCLINE-INDUCED EXPRESSION OF FLAG-CNOT3 VARIANTS IN STABLY TRANSFECTED CELL LINES DOES NOT AFFECT CELL PROLIFERATION	118

4.3 CNOT3 KNOCKDOWN LEADS TO A DECREASE IN CELL PROLIFERATION IN HEK293 CELLS	120
4.4 CNOT3 KNOCKDOWN IN HEK293 AFFECTS CCR4-NOT COMPLEX FORMATION	123
4.5 THE C-TERMINAL DOMAIN OF CNOT3 IS REQUIRED FOR CELL PROLIFERATION	125
4.6 THE N-TERMINAL DOMAIN OF CNOT3 IS REQUIRED FOR CELL PROLIFERATION	129
4.7 DISCUSSION	132
CHAPTER 5. CHARACTERISING DISEASE-ASSOCIATED MUTATIONS IN THE GENES ENCODING THE CNOT3 AND CNOT2 SUBUNITS OF THE CCR4-NOT COMPLEX.....	136
5.1 INTRODUCTION	137
5.2 IMPORTANCE OF C-TERMINAL DOMAIN OF CNOT3 FOR INCORPORATION OF CNOT3 INTO THE CCR4-NOT COMPLEX.....	138
5.3 LITERATURE-BASED IDENTIFICATION OF CNOT3 MUTATIONS ASSOCIATED WITH T-ALL AND THE RARE NEURODEVELOPMENTAL DISORDER IDDSADF	141
5.4 THE ANALYSIS OF CNOT3 VARIANTS FOUND IN T-ALL AND NEURODEVELOPMENTAL DISORDERS USING PREDICTIVE ALGORITHMS	142
5.4.1 CNOT3 VARIANTS FOUND IN T-ALL AND NEURODEVELOPMENTAL DISORDERS ANALYSED USING SORTING INTOLERANT FROM TOLERANT (SIFT)	144
5.4.2 CNOT3 VARIANTS FOUND IN T-ALL AND NEURODEVELOPMENTAL DISORDERS ANALYSED USING POLYPHEN-2	146
5.4.3 CNOT3 VARIANTS FOUND IN T-ALL AND NEURODEVELOPMENTAL DISORDERS ANALYSED USING ALPHAMISSENSE	148
5.5 CONSTRUCTION OF PLASMIDS CONTAINING CNOT3 WITH PATIENT-DERIVED MUTATIONS LINKED TO T-ALL AND IDDSADF	151
5.6 EXPRESSION OF CNOT3 MUTATIONS ASSOCIATED WITH T-ALL AND IDDSADF IN HEK293 CELL LINES	153
5.7 PATIENT-DERIVED MUTATIONS LINKED WITH T-ALL AND IDDSADF IN CNOT3 DO NOT AFFECT ITS INCORPORATION INTO THE CCR4-NOT COMPLEX	155

5.8 LITERATURE-BASED IDENTIFICATION OF CNOT2 MUTATIONS ASSOCIATED WITH IDNADFS.....	158
5.9 CNOT2 VARIANTS FOUND IN NEURODEVELOPMENTAL DISORDERS ANALYSED USING PREDICTIVE ALGORITHMS	160
5.9.1 CNOT2 VARIANTS FOUND IN NEURODEVELOPMENTAL DISORDERS ANALYSED USING SORTING INTOLERANT FROM TOLERANT (SIFT)	160
5.9.2 CNOT2 VARIANTS FOUND IN NEURODEVELOPMENTAL DISORDERS ANALYSED USING POLYPHEN-2	161
5.9.3 CNOT2 VARIANTS FOUND IN NEURODEVELOPMENTAL DISORDERS ANALYSED USING ALPHAMISSENSE	163
5.9.4 CNOT2 VARIANTS FOUND IN NEURODEVELOPMENTAL DISORDERS ANALYSED USING PROVEAN	163
5.10 CONSTRUCTING PLASMIDS CONTAINING PATIENT-DERIVED MUTATIONS LINKED TO RARE NEURODEVELOPMENTAL DISORDERS IN CNOT2	165
5.11 TRANSIENT EXPRESSION OF FLAG-CNOT2 MUTANTS IN HEK293 CELL LINE	165
5.12 PATIENT-DERIVED MUTATIONS LINKED TO IDNADFS DISRUPT THE INTERACTION OF CNOT2 WITH THE CCR4-NOT COMPLEX	168
5.13 DISCUSSION	171
5.13.1 IMPORTANCE OF CNOT3 TERMINI FOR CCR4-NOT COMPLEX FORMATION.....	171
5.13.2 IN SILICO ANALYSIS OF CNOT3 AND CNOT2 MUTATIONS	172
CHAPTER 6. CONCLUDING REMARKS AND FUTURE OUTLOOK.....	174
6.1 CONCLUDING REMARKS AND FUTURE OUTLOOK.....	175
6.2 GENERATION OF STABLE CELL LINES FOR THE CONDITIONAL EXPRESSION OF SIRNA-RESISTANT FLAG-CNOT3 VARIANTS	177
6.3 ESSENTIAL DOMAINS FOR CELL PROLIFERATION.....	178
6.4 CHARACTERISING THE INTERACTIONS AND DISEASE-ASSOCIATED MUTATIONS OF CNOT3 AND CNOT2 VARIANTS IN THE CCR4-NOT COMPLEX.....	179
6.5 FUTURE OUTLOOK.....	181
REFERENCES	184
APPENDIX 1. COVID STATEMENT	199

APPENDIX 2. SEQUENCING RESULTS	201
SEQUENCING RESULTS FOR CNOT3 MUTATIONS LINKED WITH IDDSADF	202
APPENDIX 3. PLASMID MAPS	206

List of figures

FIGURE 1.1 OVERVIEW OF GENE REGULATION IN EUKARYOTES.....	3
FIGURE 1.2 STRUCTURE OF THE RIBOSOME.	10
FIGURE 1.3 OVERVIEW OF EUKARYOTIC TRANSLATION PROCESS.....	12
FIGURE 1.4 TRANSLATION AND CODON USAGE.....	16
FIGURE 1.5 THE MAIN mRNA DEGRADATION PATHWAYS.	19
FIGURE 1.6 THE TRIMERIC <i>S. CEREVISIAE</i> PAN2 – (PAN3) ₂ COMPLEX STRUCTURE.	22
FIGURE 1.7 STRUCTURE OF PARN.	25
FIGURE 1.8 STRUCTURE OF THE <i>D. MELANOGASTER</i> XRN1 ENZYME.....	29
FIGURE 1.9 NONSENSE-MEDIATED mRNA DECAY (NMD).....	32
FIGURE 1.10 MECHANISM OF NON-STOP DECAY (NSD).....	34
FIGURE 1.11 THE CCR4-NOT COMPLEX SUBUNITS.	37
FIGURE 1.12 STRUCTURE OF THE N-TERMINAL MODULE OF THE CCR4-NOT COMPLEX.	39
FIGURE 1.13 STRUCTURE OF THE HUMAN CCR4-NOT NUCLEASE MODULE.	41
FIGURE 1.14 STRUCTURAL ORGANISATION OF THE HUMAN NOT MODULE.	45
FIGURE 1.15 MECHANISM OF CNOT3-MEDIATED mRNA DEGRADATION AT STALLED RIBOSOMES..	50
FIGURE 2.1: POLYPHEN-2 WEB INTERFACE FOR PREDICTING THE FUNCTIONAL IMPACT OF AMINO ACID SUBSTITUTIONS.	86
FIGURE 2.2 ALPHAMISSENSE WEB INTERFACE FOR MISSENSE VARIANT PATHOGENICITY SEARCH.	88
FIGURE 2.3 SIFT SEQUENCE WEB INTERFACE FOR PREDICTING AMINO ACID SUBSTITUTION EFFECTS.....	89
FIGURE 3.1. OVERVIEW OF THE GENERATION OF FLP-IN TREX HEK293 STABLE CELL LINE.	94
FIGURE 3.2 DESIGN OF siRNA-RESISTANT CNOT3 cDNAs.	97
FIGURE 3.3 CONSTRUCTION OF VARIANTS OF FLAG-CNOT3.	98
FIGURE 3.4 DESIGN OF VARIANTS OF CNOT3 WITH TRUNCATIONS AT THE N-TERMINAL AND C- TERMINAL DOMAINS, AND NOT-BOX REGION.	100
FIGURE 3.5 CONSTRUCTION OF VARIANTS OF CNOT3 LACKING N-TERMINAL AND C-TERMINAL DOMAINS.....	101
FIGURE 3.6 TRANSIENT EXPRESSION OF FLAG-CNOT3 VARIANTS IN HEK293 CELL LINES.....	103
FIGURE 3.8 CONSTRUCTION OF VECTORS FOR INDUCIBLE EXPRESSION OF FLAG-CNOT3 VARIANTS.	106

FIGURE 3.9 EXPRESSION OF FLAG-CNOT3 VARIANTS IN STABLY TRANSFECTED FLP-IN TREX HEK293 CELL LINES.	110
FIGURE 3.10 MRNA ENCODING FLAG-CNOT3R1 AND FLAG-CNOT3R2 ARE RESISTANT TO KNOCKDOWN BY siRNA#1 AND siRNA#2, RESPECTIVELY.	111
FIGURE 4.1 DOXYCYCLINE-INDUCED EXPRESSION OF FLAG-CNOT3 VARIANTS IN STABLY TRANSFECTED CELL LINES DOES NOT AFFECT CELL PROLIFERATION.	119
FIGURE 4.2 CNOT3 KNOCKDOWN DECREASES CELL PROLIFERATION IN THE HEK293 CELL LINE.	122
FIGURE 4.3 CNOT3 KNOCKDOWN IN HEK293 AFFECTS Ccr4-NOT COMPLEX FORMATION.	124
FIGURE 4.4 EXPRESSION OF FLAG-CNOT3R1, FLAG-CNOT3R1 Δ CTD, AND FLAG- CNOT3R1 Δ Not-Box ARE RESISTANT TO KNOCKDOWN BY siRNA#1.	126
FIGURE 4.5 THE C-TERMINAL DOMAIN OF CNOT3 IS IMPORTANT FOR CELL PROLIFERATION.	128
FIGURE 4.6 STABLE HEK293 CELLS EXPRESSING FLAG-CNOT3R2 AND FLAG-CNOT3R2 Δ NTD ARE RESISTANT AGAINST KNOCKDOWN BY siRNA#2.	131
FIGURE 4.7 THE N-TERMINAL DOMAIN OF CNOT3 IS IMPORTANT FOR EFFICIENT CELL PROLIFERATION.	132
FIGURE 5.1 THE C-TERMINAL DOMAIN OF CNOT3 IS REQUIRED FOR INCORPORATION INTO THE Ccr4-NOT COMPLEX.	140
FIGURE 5.2 DISTRIBUTION OF MISSENSE MUTATIONS IN CNOT3 OBSERVED IN T-CELL ACUTE LYMPHOBLASTIC LEUKAEMIA (T-ALL) AND IDDSADF WITHIN THE CNOT3-N.	143
FIGURE 5.3 POLYPHEN-2 ANALYSIS OF CNOT3 MISSENSE MUTATIONS.	147
FIGURE 5.4 ALPHAFOLD PREDICTED 3D STRUCTURES OF WILD-TYPE AND MUTANT CNOT3 PROTEIN.	150
FIGURE 5.5 CONSTRUCTION OF VARIANTS OF FLAG-CNOT3. A SCHEMATIC REPRESENTATION OF SITE-DIRECTED MUTATION BY INVERSE PCR.	152
FIGURE 5.6 EXPRESSION OF CNOT3 MUTATIONS ASSOCIATED WITH T-ALL AND IDDSADF IN HEK293 CELL LINES.	154
FIGURE 5.7 PATIENT-DERIVED MUTATIONS LINKED WITH T-ALL IN CNOT3 DO NOT AFFECT ITS INCORPORATION INTO THE Ccr4-NOT COMPLEX.	156
FIGURE 5.8 PATIENT-DERIVED MUTATIONS LINKED WITH (IDDSADF) IN CNOT3 DO NOT AFFECT ITS INCORPORATION INTO THE Ccr4-NOT COMPLEX.	157
FIGURE 5.9 SCHEMATIC REPRESENTATION OF THE CNOT2 MUTATIONS.	159
FIGURE 5.10 THREE-DIMENSIONAL SCHEMATIC OF ASN466H MUTATION IN CNOT2.	162
FIGURE 5.11 TRANSIENT EXPRESSION OF FLAG-CNOT2 VARIANTS IN HEK293 CELL LINE.	167

FIGURE 5.12 PATIENT-DERIVED MUTATIONS LINKED TO IDNADFS DISRUPT THE INTERACTION OF CNOT2 WITH THE CCR4-NOT COMPLEX.....	170
APPENDIX 2 FIGURE 1. SEQUENCING RESULT OF L48V.	202
APPENDIX 2 FIGURE 2. SEQUENCING RESULT OF R57W.....	202
APPENDIX 2 FIGURE 3. SEQUENCING RESULT OF R57Q.	203
APPENDIX 2 FIGURE 4. SEQUENCING RESULT OF K119E.....	203
APPENDIX 2 FIGURE 5. SEQUENCING RESULT OF E147K.....	204
APPENDIX 2 FIGURE 6. SEQUENCING RESULT OF A188C.....	204
APPENDIX 2 FIGURE 7. SEQUENCING RESULT OF A188H.	205
APPENDIX 3 FIGURE 1. THE PLASMID MAP OF PCMV-FLAG-CNOT3.	207
APPENDIX 3 FIGURE 2. THE PLASMID MAP OF pcDNA5-FRT-TETO-FSNAPC.....	208
APPENDIX 3 FIGURE 3. THE PLASMID MAP OF RECOMBINASE.	209
APPENDIX 3 FIGURE 4. THE PLASMID MAP OF PCMV-FLAG-CNOT2.	210

List of tables

TABLE 1.1 ENZYMES INVOLVED IN mRNA DEADENYLATION, DECAPPING AND DEGRADATION	26
TABLE 1.2 NAME OF THE CCR4-NOT COMPLEX IN <i>DROSOPHILA</i> , <i>YEAST</i> AND HUMANS.....	36
TABLE 2.1 REAGENTS AND BUFFERS ARE USED IN BACTERIAL GROWTH AND TRANSFORMATION.....	59
TABLE 2.2 REAGENTS AND BUFFERS USED IN MOLECULAR BIOLOGY METHODS	60
TABLE 2.3 STANDARD PHUSION POLYMERASE REACTION.....	63
TABLE 2.4 THE LIST OF PRIMERS USED IN SITE-DIRECTED MUTAGENESIS.....	64
TABLE 2.5 LIST OF PLASMIDS CONSTRUCTED	67
TABLE 2.6 STANDARD TAQ DNA POLYMERASE PCR REACTION.....	69
TABLE 2.7 LIST OF SEQUENCING PRIMERS	69
TABLE 2.8 LIST OF REAGENTS AND BUFFERS USED IN CELL CULTURE	70
TABLE 2.9 TRANSFECTION STEPS FOR TRANSIENT TRANSFECTION.....	72
TABLE 2.10 siRNAs USED IN THIS RESEARCH	75
TABLE 2.11 TRANSFECTION CONDITIONS IN DIFFERENT CULTURE VESSELS	76
TABLE 2.12 ANTIBODIES USED FOR WESTERN BLOTTING	78
TABLE 2.13 BUFFERS USED IN WESTERN BLOTTING.....	79
TABLE 2.14 LIST OF THE SDS-PAGE GELS COMPONENT.....	81
TABLE 2.15 REAGENTS AND BUFFERS USED IN CO-IMMUNOPRECIPITATION.....	82
TABLE 2.16 TRANSFECTION STEPS USED FOR TRANSIENT TRANSFECT	84
TABLE 5.1 CNOT3 MUTATIONS ASSOCIATED WITH BOTH T-ALL AND NEURODEVELOPMENTAL DISORDERS	142
TABLE 5.2 LIST OF CNOT3 VARIANTS ASSOCIATED WITH T-ALL AND NEURODEVELOPMENTAL DISORDERS (IDDSADF) EVALUATED WITH SIFT	144
TABLE 5.3 ANALYSIS OF CNOT3 VARIANTS ASSOCIATED WITH T-ALL AND IDDSADF EVALUATED WITH POLYPHEN-2	146
TABLE 5.4 ANALYSIS OF CNOT3 VARIANTS ASSOCIATED WITH T-ALL AND IDDSADF EVALUATED WITH ALPHAMISSENSE	149
TABLE 5.5 CNOT2 MUTATIONS ASSOCIATED WITH IDNADFS	158
TABLE 5.6 ANALYSIS OF CNOT2 VARIANT ASSOCIATED WITH IDNADFS EVALUATED WITH SIFT	160
TABLE 5.7 ANALYSIS OF CNOT2 VARIANT ASSOCIATED WITH IDNADFS EVALUATED WITH POLYPHEN-2	161
TABLE 5.8 ANALYSIS OF CNOT2 VARIANT ASSOCIATED WITH IDNADFS EVALUATED WITH ALPHAMISSENSE	163

TABLE 5.9 ANALYSIS OF CNOT2 VARIANT ASSOCIATED WITH IDNADFS EVALUATED WITH PROVEAN	165
--	-----

List of abbreviations

AF	AlphaFold
ASD	Autism Spectrum Disorder
CAM	Chorioallantoic Membrane
Ccr4-Not	Carbon catabolite repression 4 (Ccr4)–negative on TATAless
CDKI	Cyclin-dependent Kinase Inhibitor
cDNA	Complementary DNA
CMV	Cytomegalovirus promoter
Co-IP	Co-immunoprecipitation
CRC	Colorectal Cancer
CS	Conserved Sequence
CTD	Carboxy-Terminal Domain
DDD	Deciphering Developmental Disorders
DEDD	Aspartate (D), Glutamate (E), Aspartate (D), Aspartate (D)
DNA	Deoxyribonucleic acid
DPE	Downstream Promoter Elements
DSIF	DRB Sensitivity-Inducing Factor
E. coli	Escherichia coli
eIFs	Eukaryotic Initiation Factors
EJC	Exon Junction Complex
ES	Mouse Embryonic Stem
FRT site	Flippase recognition target site
GTFs	General Transcription Factors
HEK293	Human Embryonic Kidney 293 Cells
IDDSADF	Intellectual Disability with Dysmorphic Syndrome and Autistic Features
IDNADFS	Intellectual Developmental Disorder with Nasal Speech, Dysmorphic Facies, and Skeletal anomalies
IRPs	Iron Regulatory Proteins
KLF2	Krüppel-Like Factor 2
L7A	Ribosomal Protein L7A
LSqCC	Lung Squamous Cell Carcinoma
m7G	7-methylguanosine

Met-tRNAi	Methionine transfer RNA
miRNAs	Micro Ribonucleic Acid
mRNA	Messenger RNA
NAR	CNOT1 Anchor Region
NDDs	Neurodevelopmental Disorders
NGD	No-go Decay
NMD	Nonsense-mediated Decay
NSCLC	Non-Small Cell Lung Cancer
NSD	Non-Stop Decay
NTD	N terminal domain
ORF	Open Reading Frame
PABP	Poly(A)-binding Protein
PAP	Poly(A) Polymerase
PCR	Polymerase Chain Reaction
PIC	Pre-initiation Complex
PolyPhen-2	Polymorphism Phenotyping v2
P-TEFb	Transcription Elongation Factor b
PMR1	Polysomal Ribonuclease 1
pre-mRNA	Precursor Messenger RNA
PROVEAN	Protein Variation Effect Analyser
RB	Retinoblastoma
RNA	Ribonucleic Acid
RNAi	RNA Interference
RNAPII	RNA Polymerase II
RNPs	Ribonucleoprotein Particles
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SIFT	Sorting Intolerant From Tolerant
SNP	Single Nucleotide Polymorphism
SNVs	Single Nucleotide Variants
snRNAs	Small Nuclear RNAs
siRNA	short interference RNA
SURF	SMG1, UPF1, and Release Factors
T-ALL	T-cell Acute Lymphoblastic Leukaemia

TBP	TATA-Binding Protein
TCGA	The Cancer Genome Atlas
TNBC	Triple-negative Breast Cancer
TREx	Tetracycline-Regulated Expression
TSS	Transcription Start Site
tRNA	Transfer Ribonucleic Acid
UCH	Ubiquitin C-terminal Hydrolase
UTR	Untranslated Region
VEGF	Vascular Endothelial Growth Factor
WT	Wild Type

Chapter 1. Introduction

1.1 Eukaryotic gene regulation

Just as civil society requires checks and balances to regulate freedom of expression, cells have developed various mechanisms to control the expression of their constituent genes. Regulation of gene regulation, includes a wide range of mechanisms that are used by cells to increase or decrease the production of specific gene products (protein or RNA) (Lee and Young 2013). Sophisticated programs of gene expression are widely observed in biology, for example to trigger developmental pathways (Pollex et al. 2024), respond to environmental stimuli (Penner-Goeke and Binder 2024; Rahman and McGowan 2022), or adapt to new food sources (Dwaraka et al. 2024).

In eukaryotic cells, the regulation of gene expression is a highly complex process that can occur at multiple levels. Unlike prokaryotic genes, which are often organised into operons and regulated collectively, eukaryotic genes are regulated individually. Additionally, eukaryotic cells possess a significantly larger number of genes compared to prokaryotic cells, necessitating more intricate regulatory mechanisms (Remenyi, Scholer, and Wilmanns 2004). Gene expression can be modulated at any stage during the processes of transcription of DNA into mRNA and the subsequent translation of mRNA into protein (Dahan, Gingold, and Pilpel 2011). For clarity and simplicity, the regulation of gene expression is typically categorised into three distinct levels: epigenetic and transcriptional, post-transcriptional, translational, and post-translational (**Figure 1.1**).

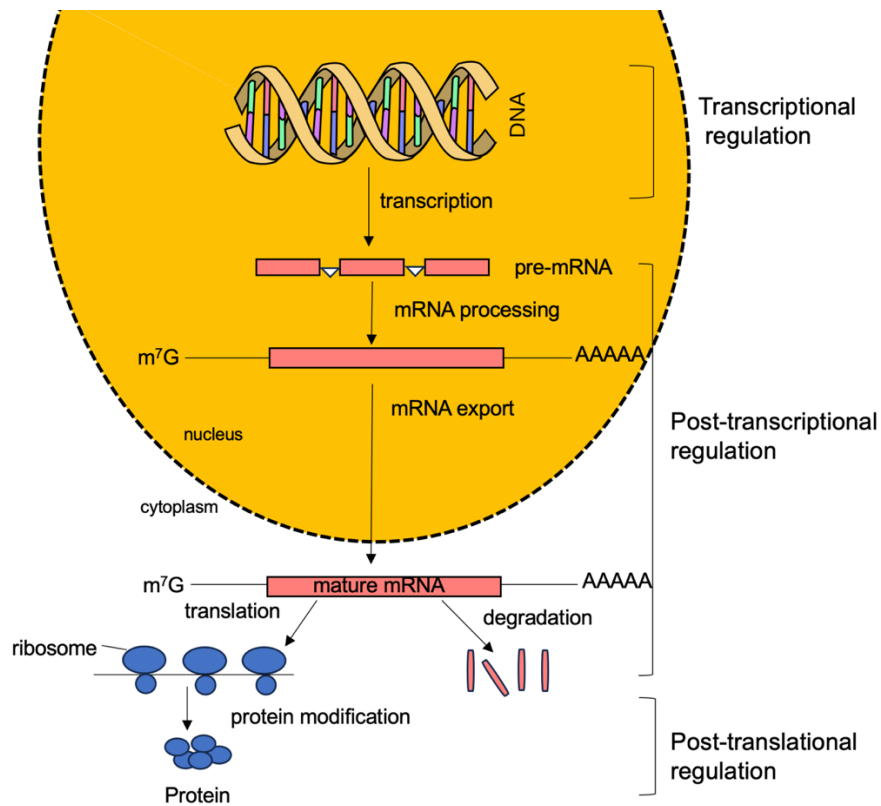


Figure 1.1 Overview of gene regulation in eukaryotes.

In eukaryotic cells, gene expression is regulated on three different levels: Epigenetic and transcriptional, post-transcriptional, translational and post-translational.

1.2 Epigenetic regulation

Epigenetics, derived from the Greek prefix "epi," meaning "above" or "on top of," refers to changes in gene expression or phenotype that occur through mechanisms other than alterations in the underlying DNA sequence (Stephens et al. 2013; Berger et al. 2009). Epigenetics adds an additional layer of regulatory control over gene expression, enabling organisms to adapt to environmental changes. Epigenetic mechanisms have the ability to modify chromatin structure and nuclear organisation. These changes, either individually or in combination, significantly impact gene expression, leading to long-lasting molecular and functional effects (Stephens et al. 2013).

1.3 Transcription

1.3.1 Transcription initiation

Transcription, the process by which DNA is transcribed into RNA, and nuclear RNA processing, which transforms the precursor mRNA (pre-mRNA) into a mature RNA molecule, are intertwined and mutually dependent processes (Persson and Mueller 2015).

Transcription initiation begins with the formation of the pre-initiation complex (PIC) at the core promoter, a minimal region close to the transcription start site (TSS) (Thomas and Chiang 2006). The process of transcription initiation begins with the binding of TBP to the TATA box, which induces a sharp 90° bend in the DNA, facilitating the recruitment of additional general transcription factors (GTFs) (Kim et al. 1993). However, it is now well established that many eukaryotic promoters, particularly in *metazoans*, do not contain consensus or even non-consensus TATA boxes. These TATA-less promoters can still initiate transcription either from a single defined nucleotide or from multiple closely spaced start sites. In such cases, transcription initiation is typically directed by alternative core promoter elements, most notably the initiator (Inr) element, which overlaps the transcription start site and can function independently of a TATA box. The Inr is sufficient to support accurate transcription initiation by RNA polymerase II and is functionally analogous to the TATA box in many respects. Additionally, other core promoter elements such as the downstream promoter element (DPE) can cooperate with the Inr to enhance transcription, particularly in TATA-less contexts. Therefore, while the TATA box plays a well-characterised role in recruiting the transcriptional machinery, its absence does not preclude precise and efficient transcription initiation (Smale 1997).

Following the establishment of the PIC, DNA unwinding is mediated by the translocase activity of Transcription Factor IIH (TFIIH) (Fishburn et al. 2015). After DNA unwinding, the process transitions to promoter clearance or promoter escape. In this step, RNA polymerase II (RNAPII), a multi-subunit enzyme central to eukaryotic transcription and responsible for synthesising precursor messenger RNA (pre-mRNA) from protein-coding genes, synthesises 10–16 nucleotides before disengaging from the initiation complex into elongation. The GTFs remain at the promoter, acting as a scaffold for subsequent rounds of transcription (Sharma et al. 2024).

1.3.2 Transcription elongation

Once RNAPII is engaged, it moves along the template strand of DNA towards the transcription start site, reading the template DNA strand in a 3' to 5' direction and synthesising a complementary pre-mRNA transcript in the 5' to 3' direction. However, elongation is a highly regulated and dynamic process rather than a continuous one. Early in elongation, RNAPII often pauses near the promoter-proximal region, a step that allows additional regulatory inputs to fine-tune transcription and facilitates the recruitment of essential RNA processing factors such as those involved in capping, splicing, and polyadenylation (Jonkers and Lis 2015). This regulatory checkpoint is mediated by the Negative Elongation Factor (NELF) and DRB Sensitivity-Inducing Factor (DSIF), which bind to RNAPII and stabilise its paused state (Kwak and Lis 2013).

The release of RNAPII from this paused state is orchestrated by the Positive Transcription Elongation Factor b (P-TEFb), a critical regulator of elongation. P-TEFb phosphorylates the serine 2 (Ser2) residues on the C-terminal domain (CTD) of RNAPII, as well as components of NELF and DSIF. This phosphorylation converts DSIF into a positive elongation factor and results in the eviction of NELF, allowing RNAPII to transition into productive elongation (Taube et al. 2002). Importantly, Ser2 phosphorylation not only releases RNAPII from the pause but also facilitates the recruitment of co-transcriptional RNA processing factors, including splicing machinery, capping enzymes, and polyadenylation complexes (Jonkers and Lis 2015; Taube et al. 2002). This coupling ensures that RNA synthesis and RNA processing occur simultaneously, streamlining the maturation of nascent RNA (Jonkers and Lis 2015; Taube et al. 2002).

1.3.3 Transcription termination

Transcription termination is a critical phase in the gene expression process, ensuring that RNA synthesis concludes precisely and effectively. This stage involves the disassembly of the transcription machinery, the release of RNAPII from the DNA template, and the detachment of the newly synthesised RNA transcript (Eaton and West 2020; Proudfoot 2011). The mechanism of transcription termination is still unclear, typically explained by two models: the allosteric model and the torpedo model (Connelly and Manley 1988; Sharma et al. 2024).

The torpedo model provides one explanation for transcription termination. After RNAPII synthesises the nascent RNA, cleavage occurs at the polyadenylation site, facilitated by the cleavage and polyadenylation complex. Following cleavage, the 5'-3' exonuclease *XRN2* (Rat1 in yeast) rapidly degrades the downstream RNA fragment, chasing RNAPII and destabilising the elongation complex. This process ultimately results in the release of RNAPII from the DNA, terminating transcription. The torpedo mechanism emphasises the role of RNA degradation in physically dislodging RNAPII from the template (Connelly and Manley 1988).

The allosteric model offers an alternative perspective. Here, the binding of termination factors to the polyadenylation signal induces a conformational change in RNAPII. This structural shift reduces its processivity, causing RNAPII to pause and eventually dissociate from the DNA (Logan et al. 1987). Concurrently, the loss of elongation factors or the recruitment of termination-specific proteins further destabilises the transcription complex, facilitating termination. This model highlights the importance of molecular signalling and protein dynamics in orchestrating the final steps of transcription (Ahn, Kim, and Buratowski 2004).

Termination is intricately coupled with RNA processing, ensuring that the nascent transcript undergoes modifications such as 5' capping, splicing, and 3' polyadenylation. The carboxy-terminal domain (CTD) of RNAPII serves as a scaffold for the sequential recruitment of RNA processing factors (Ahn, Kim, and Buratowski 2004).

1.4 Post-transcriptional regulation

The process of nuclear RNA processing is fundamental to the conversion of the primary RNA transcript, also known as pre-mRNA, into a fully functional, mature mRNA that is capable of guiding protein synthesis (Han et al. 2011). Pre-mRNA undergoes a series of modifications that transform it from an unstable, nascent molecule into a mature mRNA

capable of navigating the complex intracellular environment and facilitating accurate translation (Darnell 1978). These modifications—5' capping, splicing, and 3' polyadenylation—are tightly coordinated and integrated with transcription itself, illustrating the remarkable efficiency and complexity of eukaryotic gene regulation.

1.4.1 mRNA Capping

One of the earliest modifications to the pre-mRNA is the addition of a 7-methylguanosine cap to its 5' end. This process begins almost immediately after transcription initiation when the nascent RNA transcript is just 20–30 nucleotides in length (Ramanathan, Robb, and Chan 2016). The capping reaction is catalysed by a series of enzymatic steps involving RNA triphosphatase, guanylyltransferase, N7-methyltransferase and Cap1 methyltransferase enzymes. These enzymes sequentially remove the terminal phosphate group from the 5' triphosphates of the RNA, attach a guanosine residue via a 5'-5' triphosphate linkage, and methylate the guanosine at the N7 position to produce the mature cap structure (Shatkin 1976). The 5' cap on mRNA acting like a shield, protecting the mRNA from being chopped up by enzymes that normally destroy unprotected RNA ends (Ghosh and Lima 2010; Liu and Kiledjian 2006).

1.4.2 mRNA Splicing

Splicing is another of the most complex and dynamic aspects of RNA processing, responsible for removing non-coding introns and ligating coding exons to generate a continuous open reading frame (ORF). This process is carried out by the spliceosome, a massive and highly dynamic ribonucleoprotein complex composed of small nuclear RNAs (snRNAs) and associated proteins, collectively referred to as a small nuclear ribonucleoprotein (snRNPs) (Abelson 2008).

1.4.3 mRNA Polyadenylation

Polyadenylation is a fundamental post-transcriptional modification in eukaryotic gene expression, involving the addition of a poly(A) tail to the 3' end of precursor messenger RNA (pre-mRNA). This modification is crucial for mRNA stability, nuclear export, and translational efficiency. The process of polyadenylation can be broadly categorised into nuclear and cytoplasmic phases, each with distinct mechanisms and regulatory roles (Colgan and Manley 1997).

In the nucleus, polyadenylation occurs co-transcriptionally and is tightly coupled with transcription termination. The process is initiated by the recognition of specific sequence elements within the pre-mRNA, notably the highly conserved hexamer AAUAAA, located upstream of the cleavage site, and a downstream GU-rich sequence. Following cleavage, poly(A) polymerase (PAP) adds a poly(A) tail, typically consisting of 200-250 adenine residues, to the newly formed 3' end (Colgan and Manley 1997). However, mRNAs encoding histone proteins lack a poly(A) tail and instead contain a unique stem-loop structure at their 3' end (Marzluff, Wagner, and Duronio 2008).

1.4.4 mRNA nuclear export

After polyadenylation, which marks the final step in pre-mRNA processing by adding a poly(A) tail to the transcript's 3' end, the mRNA becomes primed for nuclear export. This transition involves the recruitment of nuclear export machinery, ensuring that only properly processed and mature mRNAs are transported out of the nucleus. The poly(A) tail, along with other processing marks such as the 5' cap and splicing-dependent exon junction complexes (EJCs), signals that the mRNA is export-ready (Palazzo, Qiu, and Kang 2024). These processing marks attract export adaptors like the TREX complex and Yra1, which facilitate the assembly of export-competent messenger ribonucleoprotein (mRNP) complexes (Stewart 2010). Once assembled, these complexes interact with nuclear pore complexes (NPCs) to ensure selective and efficient export to the cytoplasm, while defective or misprocessed transcripts are retained and degraded within the nucleus to maintain gene expression fidelity (Stewart 2010).

1.5 Translation

1.5.1 Eukaryotic Ribosome

The human 80S ribosome, a molecular machine essential for protein synthesis, consists of two subunits: the 40S small subunit and the 60S large subunit. The 40S subunit is composed of 18S ribosomal RNA (rRNA) and 33 ribosomal proteins (RPs), and its primary role is to bind, unwind, and scan messenger RNAs (mRNAs) during translation. Meanwhile, the 60S subunit, which contains 5S, 5.8S, and 28S rRNAs along with 47 RPs, is responsible for catalysing peptide bond formation and ensuring the quality control of nascent peptide chains. Together, these subunits form a highly efficient system for

decoding genetic information into functional proteins (Pelletier, Thomas, and Volarevic 2018).

Ribosome biogenesis primarily occurs in the nucleolus, a specialised subnuclear structure that forms around regions of ribosomal DNA (rDNA). These regions, known as nucleolar organizer regions (NORs), harbour hundreds of tandemly arranged rDNA gene repeats on the short arms of five acrocentric chromosomes. These genes are transcribed to produce a 47S precursor rRNA (pre-rRNA), which contains the sequences for 18S, 5.8S, and 28S rRNAs. The transcription of the 47S pre-rRNA by RNA Polymerase I is accompanied by its co-transcriptional assembly into a 90S processome, a precursor to the mature ribosome. This complex involves the association of rRNA with ribosomal proteins, 5S rRNA (transcribed separately by RNA Polymerase III), and numerous small nucleolar RNAs (snoRNAs), enzymes, and assembly factors. These components work together to process and chemically modify the rRNA, ensuring proper folding and functional maturation (Pelletier, Thomas, and Volarevic 2018).

The assembly process of ribosomal subunits is a multistep mechanism. During the early steps, dynamic interactions between the rRNA and ribosomal proteins occur, forming key functional sites such as the mRNA entry and exit channels. In the 40S subunit, proteins like rpS26e and rpS15e play crucial roles in stabilising mRNA binding and maintaining the decoding site's fidelity. These ribosomal proteins interact with specific mRNA regions, ensuring that the codon-anticodon pairing occurs accurately and efficiently during translation (Graifer and Karpova 2012).

The ribosome's architecture is highly adapted to its translational role. The small subunit (40S) contains the decoding centre (**Figure 1.2**), which interacts with mRNA and tRNA and includes three key binding sites for tRNA: the aminoacyl (A), peptidyl (P), and exit sites (E). The A site accommodates incoming aminoacyl-tRNA, the P site holds the tRNA carrying the growing peptide chain, and the E site is where the uncharged tRNA exits the ribosome. The large subunit (60S) contains the peptidyl transferase centre (PTC), which catalyses peptide bond formation. These sites are arranged to facilitate the sequential binding and movement of tRNA molecules, guided by the ribosome's thermal ratchet mechanism. This mechanism rectifies random thermal motions into unidirectional translocation of tRNA and mRNA during elongation (Graifer and Karpova 2012).

Once assembled into pre-40S and pre-60S subunits, these immature ribosomal particles are exported to the cytoplasm, where additional modifications and assembly steps occur,

rendering the ribosome competent for translation (Pelletier, Thomas, and Volarevic 2018). The involvement of specific export factors, like Ccrm1 and Nmd3, facilitates this process. Furthermore, cytoplasmic maturation steps involve the proofreading and final incorporation of late-binding ribosomal proteins, ensuring the production of functional ribosomes (Spirin 2002)..

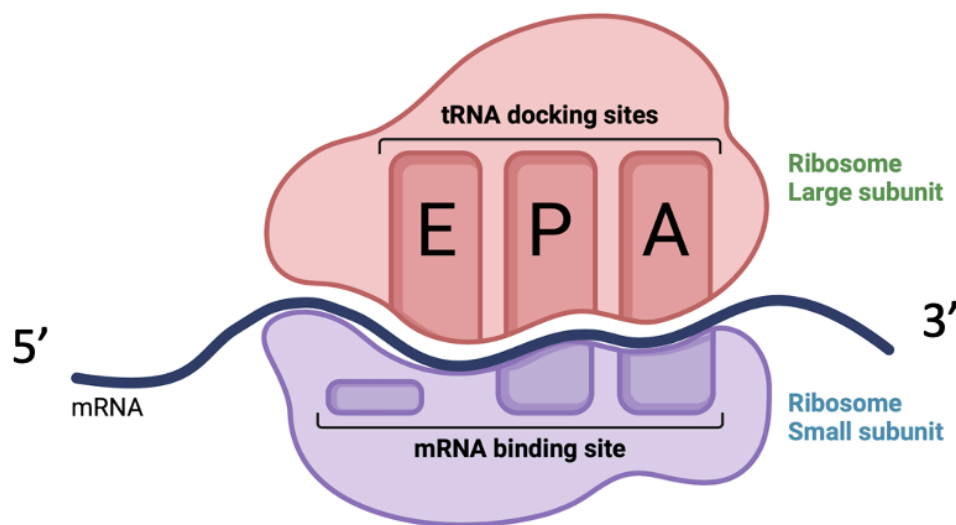


Figure 1.2 Structure of the Ribosome.

Schematic representation of the ribosome and its functional sites during translation. The ribosome consists of two subunits: the large subunit in pink and the small subunit in purple. The large subunit contains three tRNA docking sites: the A (aminoacyl) site, which accommodates incoming aminoacyl-tRNA; the P (peptidyl) site, which holds the tRNA attached to the growing peptide chain; and the E (exit) site, where uncharged tRNA exits the ribosome. The small subunit contains the mRNA binding site, which interacts with the mRNA strand to guide translation. Together, these subunits coordinate the sequential binding and movement of tRNA and mRNA to facilitate protein synthesis. The figure was created using Biorender.com.

1.5.2 Translation initiation

Translation is the cellular process by which messenger RNA (mRNA) is decoded to synthesise proteins, following the genetic instructions encoded in DNA. This process occurs in the cytoplasm and involves three primary stages: initiation, elongation, and termination (Sonnenberg and Hinnebusch 2009).

During initiation, the small ribosomal subunit binds to initiation factors and the initiator tRNA is charged with methionine (Met-tRNAⁱ). This complex recognises the 5' cap of the mRNA and scans in the 5' to 3' direction to locate the start codon (Sonenberg and Hinnebusch 2009). Upon finding AUG start codon, typically embedded within the Kozak consensus sequence (Kozak 1989), the initiator tRNA pairs with it, establishing the reading frame for translation. Subsequently, the large ribosomal subunit joins the complex, forming a complete ribosome ready for elongation. This assembly positions the initiator tRNA in the P site of the ribosome (Sonenberg and Hinnebusch 2009; Jackson, Hellen, and Pestova 2010) **(Figure 1.3)**.

1.5.3 Translation elongation

Elongation involves the sequential addition of amino acids to the growing polypeptide chain. Aminoacyl-tRNA, carrying the appropriate amino acid, enters the A site of the ribosome and pairs with the corresponding mRNA codon through Watson-Crick base pairing between the codon on the mRNA and the anticodon on the tRNA. The ribosome catalyses the formation of a peptide bond between the amino acid in the A site and the growing polypeptide chain attached to the tRNA in the P site through the enzymatic activity of the peptidyl transferase centre, which is part of the ribosome's large subunit (Polacek and Mankin 2005). The ribosome then moves one codon forward along the mRNA, translocating the tRNA with the growing polypeptide to the P site, vacating the A site for the next aminoacyl-tRNA, and moving the uncharged tRNA to the E site, from which it exits the ribosome. This cycle repeats, elongating the polypeptide chain until a stop codon is encountered (Sonenberg and Hinnebusch 2009; Jackson, Hellen, and Pestova 2010) **(Figure 1.3)**.

1.5.4 Translation termination

Termination concludes translation when the ribosome encounters a stop codon (UAA, UAG, or UGA) on the mRNA. Release factors recognise the stop codon and bind to the A site, prompting the release of the polypeptide chain from the tRNA in the P site. The ribosomal subunits, mRNA, and remaining tRNA then dissociate, making the components available for future rounds of translation (Sonenberg and Hinnebusch 2009; Jackson, Hellen, and Pestova 2010) **(Figure 1.3)**.

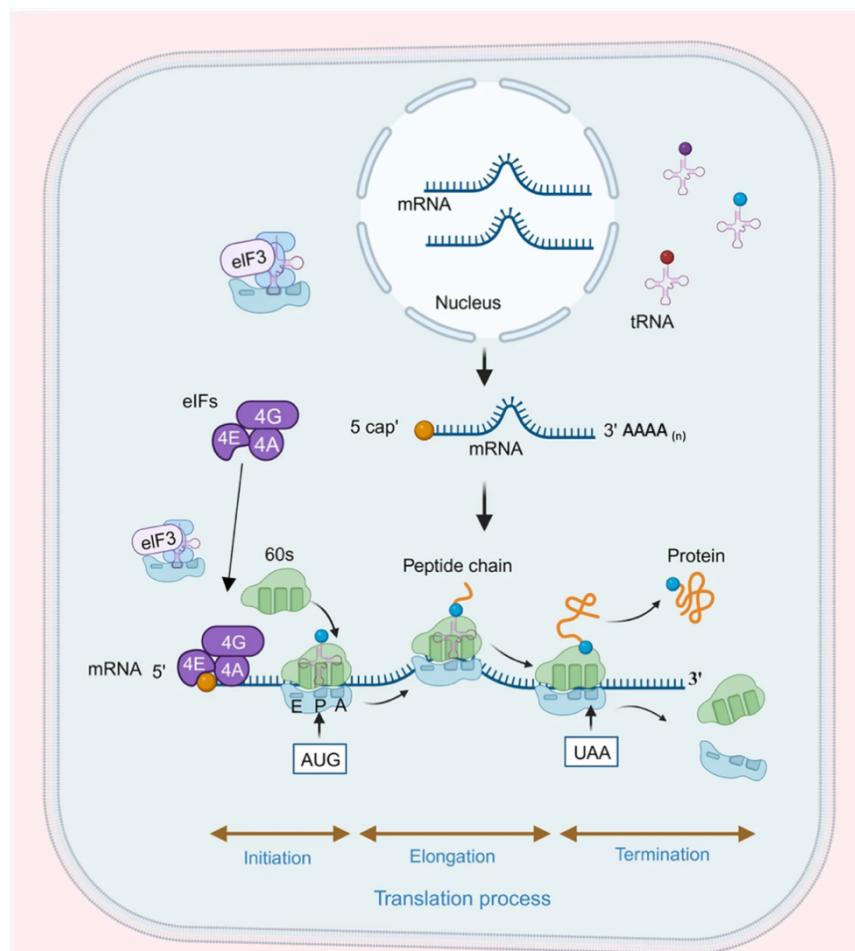


Figure 1.3 Overview of Eukaryotic Translation Process.

The figure outlines eukaryotic translation, starting with mRNA export from the nucleus. Initiation involves eIFs binding the 5' cap and assembling the ribosome. During elongation, tRNAs add amino acids as the ribosome moves along the mRNA. Termination occurs at stop codons, releasing the completed protein (Jia et al. 2024).

1.6 Translation regulation

Translational regulation is a fundamental aspect of gene expression control, ensuring that proteins are synthesised at the right time, place, and quantity within a cell. During

translation initiation, the assembly of the ribosome on mRNA is a highly regulated process. Eukaryotic initiation factors (eIFs) play crucial roles in this assembly. For instance, eIF2, in its GTP-bound form, escorts the initiator methionyl-tRNA to the small ribosomal subunit. Phosphorylation of the eIF2 alpha subunit can inhibit this process, leading to a decrease in global protein synthesis, which is a common cellular response to stress condition, such as innate immune response to viral infections (Kashiwagi et al. 2016).

The structure of the mRNA itself also influences translation. Secondary structures in the 5' untranslated region (UTR) can impede the scanning process of the ribosome, thereby affecting initiation efficiency. For example, the iron-responsive element is a small stem-loop structure found in the 5' UTR of mRNAs involved in iron metabolism. In low iron conditions, iron regulatory proteins (IRPs) bind to the iron-responsive elements (IREs), physically blocking the ribosome from accessing the mRNA and initiating translation. This prevents the production of proteins involved in iron storage and uptake when iron levels are already low (Leppek, Das, and Barna 2018).

Moreover, microRNAs (miRNAs) are small non-coding RNAs, approximately 22 nucleotides in length, that play a crucial role in translational repression. They achieve this by binding to complementary sequences within the 3' UTRs of target mRNAs, leading to translational repression or mRNA degradation, thereby fine-tuning protein synthesis (Sonenberg and Hinnebusch 2009).

The length of the poly(A) tail at the 3' end of mRNA has a complex relationship with stability and translational efficiency. While Cytoplasmic polyadenylation, regulated by elements in the 3' UTR and associated binding proteins, can activate the translation of specific mRNAs during developmental processes such as oocyte maturation (Villalba, Coll, and Gebauer 2011), recent findings challenge the notion that longer poly(A) tail universally correlate with increased mRNA stability or translational efficiency (Passmore and Collier 2022).

Passmore et al highlighted that highly translated, stable mRNAs often have relatively short poly (A) tails (30 adenosines), suggesting that the relationship between poly(A) tail length and mRNA metabolism is not straightforward but influenced by transcript-specific factors and cellular context. Moreover, Poly(A)-binding protein (PABP) is a vital protein that plays a multifaceted role in the regulation of eukaryotic translation. Its primary function arises from its interaction with the poly(A) tail. This interaction, in conjunction

with PABPA's associations with other translation factors, significantly influences the efficiency and control of protein synthesis (Gorgoni and Gray 2004; Vicens, Kieft, and Rissland 2018).

For more than 20 years, it was believed that various codons were translated at varying speeds. However, a radio-labelled amino acid incorporation assay demonstrated that the identity of the codon could influence the rate of translation elongation (Akashi 1994; Gingold and Pilpel 2011; Sorensen and Pedersen 1991). The relationship between codon usage and translation is a finely tuned system that reflects the evolutionary optimisation of gene expression processes. Codon usage bias, or the preferential use of specific synonymous codons, is a feature observed in all genomes and significantly affects translation efficiency, accuracy, and protein folding (Yu et al. 2015). This bias arises primarily from selection pressures for efficient and accurate protein synthesis and from co-evolution with tRNA availability (Yu et al. 2015).

During translation, each codon in an mRNA sequence corresponds to a specific amino acid, which is added to the growing polypeptide chain by the ribosome. However, not all synonymous codons are translated with the same efficiency. Codons that are more frequently used, termed frequent codons, are usually paired with abundant tRNAs, enabling rapid and accurate decoding. Conversely, rare codons correspond to less abundant tRNAs, resulting in slower elongation rates and increased likelihood of ribosome pausing (Pedersen 1984; Yu et al. 2015). Numerous studies have investigated codon pair biases across different species and, in some cases, between cell types (Gingold et al. 2014). Notably, the codon pair GGGCUU was the most consistently favoured, being over-represented in 76% of the species analysed, while the least favoured pair, UUCGCA, was found to be under-represented in 86% of the genomes studied (Komar 2016).

One study used a *Neurospora crassa* cell-free translation system to directly measure the effects of codon usage on translation elongation. By analysing the translation of firefly luciferase (De Keersmaecker et al. 2013) mRNAs engineered with different codon usage patterns, it was demonstrated that optimal codons significantly increased the rate of translation elongation compared to rare codons (**Figure 1.4**). For instance, codon optimisation of the Luc mRNA led to faster translation and earlier detection of the bioluminescent signal, an indirect proxy for translation speed. In contrast, the use of rare codons delayed translation by several minutes, highlighting the profound effect of codon usage on translational dynamics (Yu et al. 2015).

Barrington et al emphasise that codon usage can regulate translation elongation and initiation, expanding its role in translational control beyond previously understood mechanisms. Codon usage within the open reading frame (ORF) influences the binding of critical initiation factors, such as eIF4E and eIF4G1, which mediate cap-dependent translation. Rare codons reduce the association of these factors with the mRNA, leading to repression of translation initiation (Barrington et al. 2023).

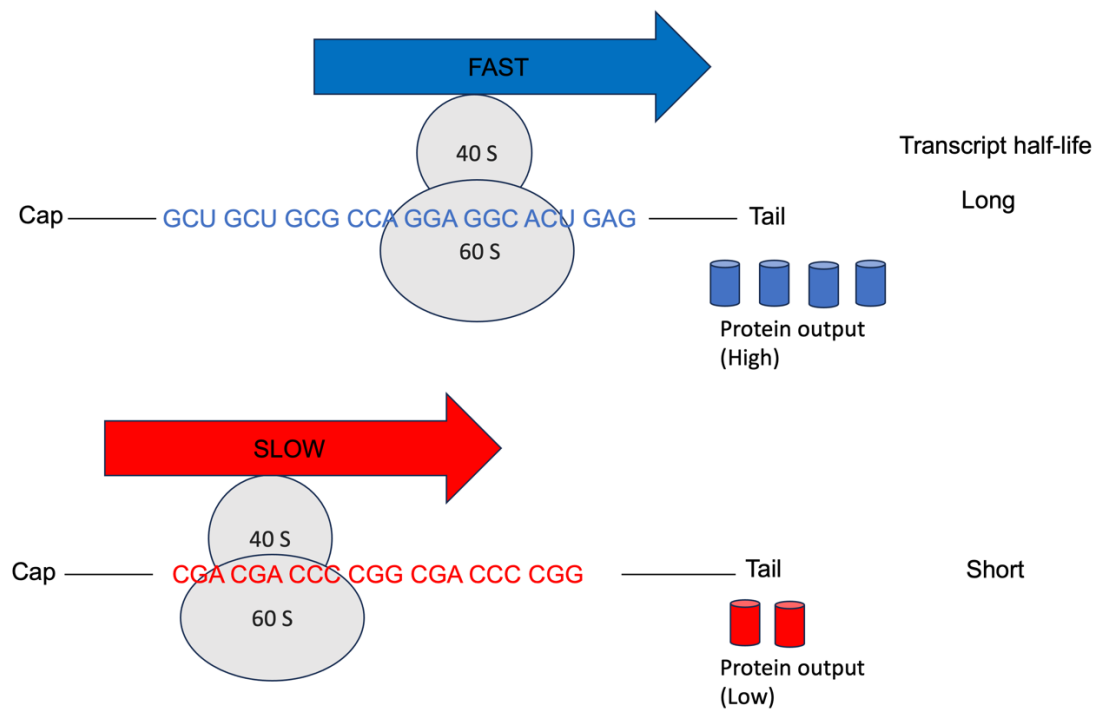


Figure 1.4 Translation and codon usage.

The diagram illustrates the relationship between codon usage, translation speed, and protein output. The top portion represents a transcript with optimal codons that match abundant tRNAs, leading to fast ribosomal movement, increased protein synthesis, and a longer transcript half-life. The bottom portion shows a transcript with rare codons, which correspond to rare tRNAs, resulting in slower ribosomal movement, reduced protein output, and a shorter transcript half-life. The 40S and 60S ribosomal subunits are depicted interacting with the mRNA during translation. The arrows indicate the overall translation speed and protein output.

1.7 Cytoplasmic messenger RNA decay

One of the processes in the post-transcriptional regulation of gene expression is mRNA turnover, which significantly influences the rates at which genes are expressed in cells. mRNA level is determined by the balance between transcription and degradation (Korner and Wahle 1997). Each mRNA molecule has a unique half-life, which can be modified in response to various internal or external stimuli (Korner and Wahle 1997). Rapid mRNA degradation is particularly critical for transcripts encoding proteins that are only needed temporarily, such as those involved in early responses to serum stimulation (Korner and Wahle 1997). While the instability of mRNA is clearly essential for the swift termination of protein synthesis following transient transcription, it also facilitates rapid induction, as a short mRNA half-life allows for quicker turnover and response to changing cellular conditions (Korner and Wahle 1997).

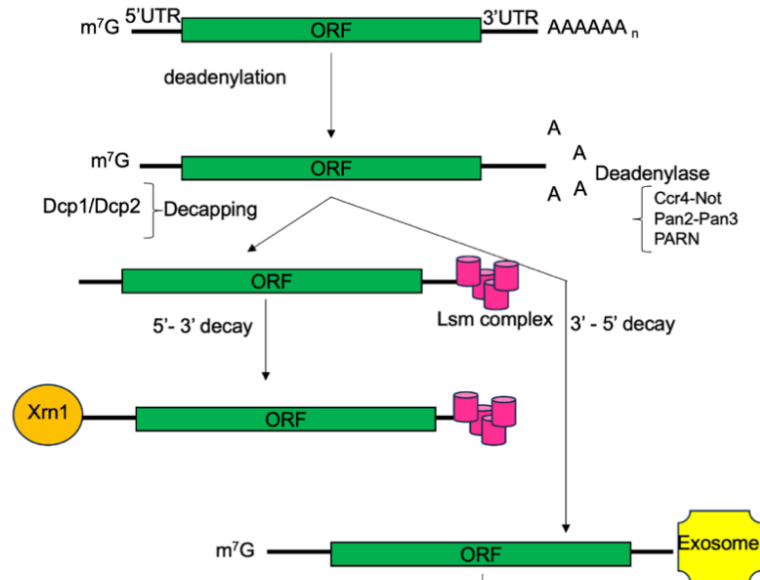
Research conducted primarily in yeast and mammals has identified two main pathways of mRNA decay in eukaryotic cells (Tucker and Parker 2000) (**Figure 1.5 A**). The degradation process typically begins with the removal of the poly(A) tail at the 3' end of the mRNA, a process known as deadenylation, which is carried out by various mRNA deadenylases (discussed below). Following deadenylation, the 5' cap structure of the transcript is removed by a decapping complex composed of two subunits, *Dcp1p* and *Dcp2p*, making the mRNA susceptible to degradation by the 5'→3' exonuclease *Xrn1p*.

Alternatively, after deadenylation, mRNAs may undergo degradation in the 3'→5' direction by the cytoplasmic exosome, a multiprotein complex with diverse 3'→5' exonucleases. In this pathway, the resulting capped oligonucleotide fragments are hydrolysed by the scavenger decapping enzyme *DcpS* (Parker and Song 2004).

In certain eukaryotes, mRNA degradation is initiated by endonucleolytic cleavage within the mRNA molecule. This process generates two fragments that are subsequently degraded by the exosome and *Xrn1* (**Figure 1.5 B**). Several endonucleases are known to target mRNAs, including polysomal ribonuclease 1 (PMR1) (Yang and Schoenberg 2004), inositol-requiring enzyme 1 (IRE1) (Schoenberg 2011), and some nucleolar ribosomal RNA (rRNA)-processing enzymes such as RNase MRP. Additionally, short interfering RNAs (siRNAs) can trigger mRNA decay through cleavage, along with other unidentified endonucleases involved in the degradation of faulty mRNAs (Gatfield and Izaurralde

2004). PMR1 and IRE1 primarily target actively translating mRNAs and are only activated under specific conditions (Hollien and Weissman 2006). While PMR1 is associated with polysomes, IRE1 specifically targets transcripts localised to the endoplasmic reticulum.

A



B

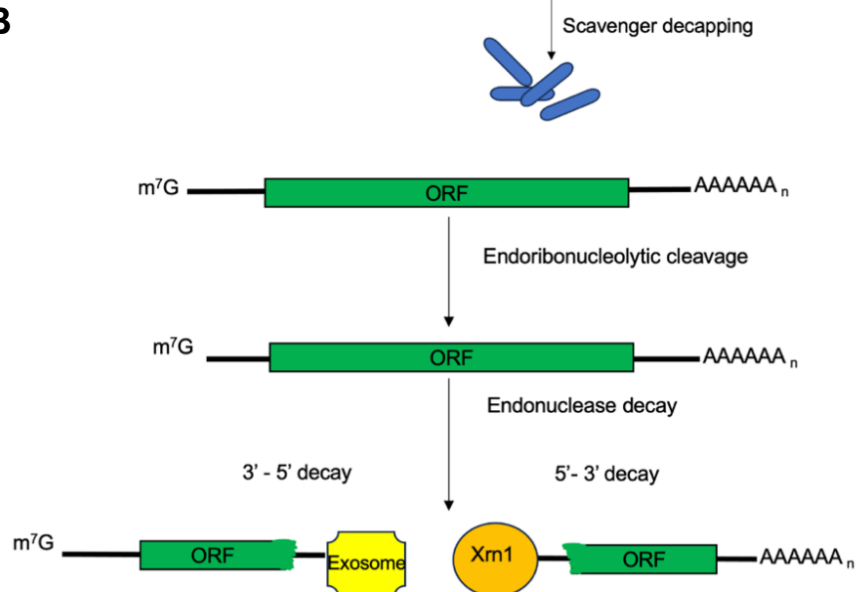


Figure 1.5 The main mRNA degradation pathways.

(A) The majority of mRNAs are degraded exonucleolytically through either a 5'→3' or 3'→5' decay pathway, both of which begin with deadenylation. The removal of the poly(A) tail is carried out by the Ccr4-Not complex, the PAN2-PAN3 complex, or PARN. In the 5'→3' pathway, the LSM complex binds to the 3' end of the mRNA following deadenylation and facilitates decapping via the Dcp1/Dcp2 complex, allowing the exonuclease Xrn1 to degrade the mRNA from the 5' end. Alternatively, in the 3'→5' pathway, the exosome complex degrades the mRNA after deadenylation.

(B) For certain mRNAs, degradation begins with endonucleolytic cleavage within the mRNA body. This process generates cleavage fragments with exposed ends, which are subsequently degraded by the exosome complex at the 3' end and Xrn1 at the 5' end.

1.8 Deadenylation: shortening of Poly (A) tail

The dynamic regulation of the poly(A) tail is a crucial factor in controlling gene expression, which directly influences the cellular proteome and affects various life processes in eukaryotic organisms (Liu et al. 2022). In the cytoplasm, poly(A) tails are bound by poly(A)-binding proteins (PABPC, specifically PABPC1 in mammals and Pab1 in yeast) and are progressively shortened in a 3'-to-5' direction by two evolutionarily conserved multiprotein complexes: *Pan2-Pan3* and *Ccr4-Not*. This shortening occurs in a biphasic manner (Yi et al. 2018). During the first phase, *Pan2-Pan3* mediates the gradual removal of the distal region of the poly(A) tail (Yamashita et al. 2005), reducing its length from approximately 200–110 nucleotides in mammals and 90 nucleotides in yeast. In the subsequent, faster phase, the *Ccr4-Not* complex primarily shortens the remaining poly(A) tail from around 110 nucleotides in mammals to a minimal length, thereby initiating mRNA decay (Schafer et al. 2019). *Ccr4-Not* is likely the primary deadenylase recruited for target mRNAs, as it has established numerous interactions with RNA-binding proteins, and the mechanisms of recruitment have been elucidated in detail in various studies (Zhao et al. 2023).

The key exonucleases (deadenylases) involved in these processes include *Pan2* from the *Pan2-Pan3* complex and *CNOT6 /CNOT6L* (Wang et al. 2010) and *CNOT7/CNOT8* (Maryati, Airhihen, and Winkler 2015) from the *Ccr4-Not* complex (Lau et al. 2009) (Wahle and Winkler 2013).

1.8.1 The *Pan2/Pan3* complex

The *Pan2-Pan3* complex is an essential player in mRNA deadenylation, a critical step in mRNA turnover. *Pan2* (**Figure 1.6**), the catalytic subunit of the complex, consists of three main domains: an N-terminal WD40 domain, a ubiquitin C-terminal hydrolase (UCH)-like domain, and a C-terminal exonuclease domain that belongs to the DEDD family. The WD40 domain, with its seven-blade β -propeller structure, facilitates interactions with *Pan3*, the regulatory subunit of the complex. *Pan2*'s enzymatic activity depends on

magnesium ions (Mg^{2+}) and breaks down the poly(A) tail of mRNA, releasing 5-AMP as a byproduct (Jonas et al. 2014; Zhao et al. 2023).

Pan3 interacts with *Pan2* in a unique ratio of one *Pan2* molecule to two *Pan3* molecules. *Pan3* features a pseudokinase domain that (**Figure 1.6**), while not catalytically active, has a functional ATP-binding site that is crucial for enhancing the RNA decay activity of *Pan2*. This structural arrangement allows *Pan3* to position the poly(A) substrate and direct it toward *Pan2*'s active site, increasing the efficiency of deadenylation. Together, the Pan2–Pan3 complex targets the distal portion of the poly(A) tail for initial trimming, which is a slower and more controlled phase of deadenylation compared to the action of other deadenylases like Ccr4–Not (Schafer et al. 2014).

The complex is recruited to the poly(A) tail via its interaction with cytoplasmic poly(A)-binding protein (PABPC). *Pan2*–*Pan3* shows a preference for longer poly(A) tails associated with poly(A)-binding proteins, and its recruitment typically requires at least two PABPC molecules. Additionally, the GW182 protein, a component of the miRNA repression machinery, has been shown to facilitate the recruitment of Pan2–Pan3 to specific mRNAs (Zhao et al. 2023).

Interestingly, Pan2–Pan3 and Ccr4–Not work in a sequential manner during mRNA deadenylation. Pan2–Pan3 carries out the initial shortening of the poly(A) tail, which is followed by the processive action of Ccr4–Not. This process is tightly regulated, as *Pan2*–*Pan3* is activated by PABPC, while Ccr4–Not is inhibited by it. This interplay ensures a controlled and efficient removal of the poly(A) tail, balancing mRNA stability and degradation (Schafer et al. 2014).

The structural organisation of *Pan2* reveals that its UCH and RNase domains form a tightly associated module, creating a stable platform for enzymatic activity. While the WD40 domain primarily interacts with Pan3 to stabilise the complex, the pseudokinase domain of Pan3 acts as a regulator rather than an enzyme. This domain helps orient the poly(A) substrate, ensuring effective interaction with Pan2's active site. Although Pan2–Pan3 can function in vitro without PABPC under specific conditions, the interaction with PABPC is critical for its recruitment and activity in a cellular context (Schafer et al. 2014). Mutations in Pan3's ATP-binding site have been shown to reduce the activity of Pan2, highlighting the importance of this site in enhancing the deadenylation process. Overall, the Pan2–Pan3 complex represents a highly coordinated system that ensures precise and

regulated removal of the poly(A) tail, a key step in the life cycle of mRNA (Schafer et al. 2014).

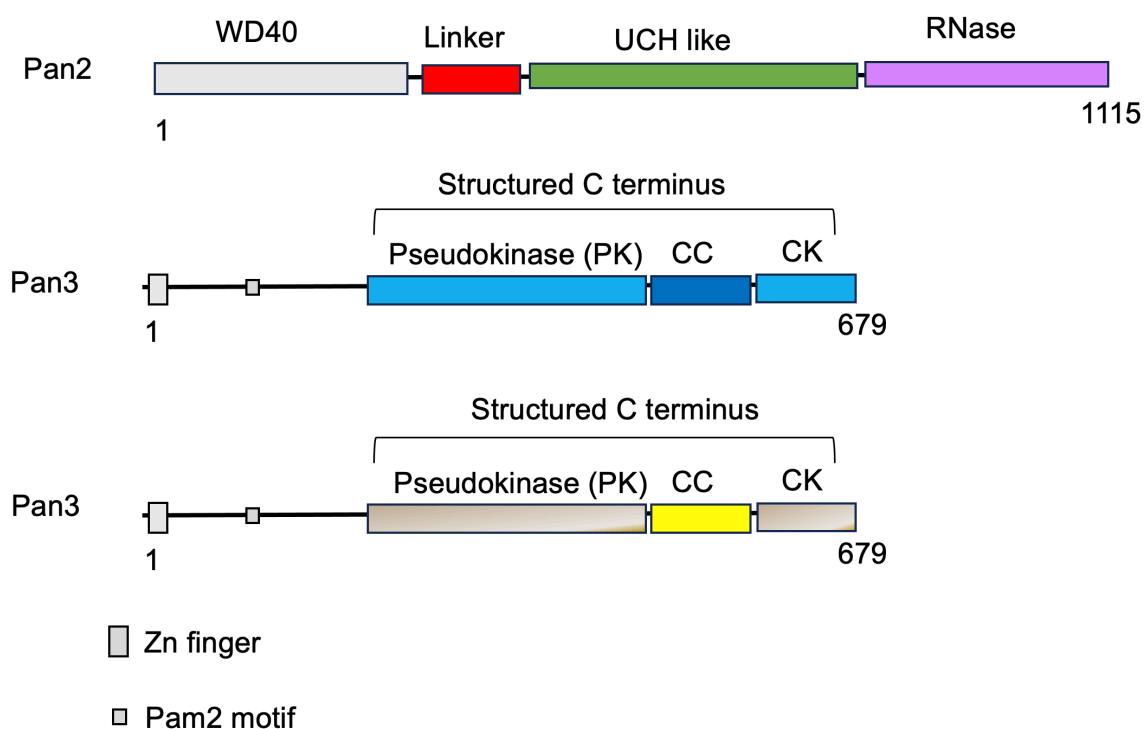


Figure 1.6 The trimeric *S. cerevisiae* PAN2 - (PAN3)₂ complex structure.

The *Pan2* subunit contains an N-terminal WD40 domain (grey), a linker region (Nagy and Watzele) wraps around the asymmetric PAN3 homodimer, a UCH-like domain (green), and a C-terminal RNase domain (purple). These domains contribute to *Pan2*'s role as a catalytic subunit in the deadenylation process.

The *Pan3* subunit features a Zn finger (grey) at the N-terminal, a *Pan2* interaction motif (grey square), and a structured C-terminal region. The structured C-terminal includes a pseudokinase domain (PK, blue/beige), a coiled-coil domain (CC, blue/yellow), and a knob domain (CK, blue/yellow).

1.8.2 Poly (A)-specific ribonuclease (*PARN*)

In mammalian cells, an additional enzyme involved in deadenylation is the poly(A)-specific exonuclease, commonly referred to as poly(A)-specific ribonuclease (*PARN*) (Korner and Wahle 1997). *PARN* is characterised by its high specificity for single-stranded poly(A) sequences and its dependence on the presence of a 3' hydroxyl group for enzymatic activity (Wu et al. 2005). Sequence analysis indicates that *PARN* belongs to the DnaQ-like superfamily of 3'-5' exonucleases, which are characterised by a conserved catalytic core made up of four acidic amino acids (aspartic acid, glutamic acid, aspartic acid, and aspartic acid) (**Figure 1.7 A and B**) known as the DEDD motif (Wu et al. 2005; Virtanen et al. 2013). Additionally, it contains a conserved R3H domain, which is thought to function as a single-stranded nucleic acid-binding domain, enabling *PARN* to bind effectively to polyadenylated mRNAs.

While *PARN* is conserved across many eukaryotic species, it is notably absent in *Saccharomyces cerevisiae* and *Drosophila melanogaster*, indicating that this enzyme is not universally required by all eukaryotes (Parker and Song 2004). In addition, *PARN*, a major deadenylase in mammalian cell extracts, is inhibited by the poly(A)-binding protein (Pab1p) under physiological salt conditions (Korner and Wahle 1997).

PARN's exonuclease activity is allosterically regulated, which means that the enzyme undergoes conformational changes upon binding of cofactors or interacting molecules. Recent studies have demonstrated that *PARN* interacts with the 5' cap structure of mRNA through its cap-binding domain, enhancing its deadenylation efficiency (Nilsson et al. 2007). This interaction is believed to promote a more processive mode of action, ensuring efficient degradation of poly(A) tails. Additionally, *PARN*'s processivity is influenced by the presence of divalent metal ions such as Mg²⁺, which are required for its catalytic activity (Wu et al. 2005).

Structural studies have further highlighted that *PARN* consists of distinct domains (**Figure 1.7 A**), including the nuclease domain, the R3H domain, and the RPM (Regulatory of Poly(A) Metabolism) domain. The nuclease domain contains the DEDD motif crucial for catalytic function, while the R3H domain is involved in RNA recognition. The RPM domain is believed to contribute to the regulation of *PARN*'s activity by mediating interactions with regulatory proteins (Virtanen et al. 2013). **Figure 1.7 B** illustrates the nuclease domain of *PARN* in complex with a substrate RNA molecule, emphasising its role in deadenylation.

1.8.3 The Ccr4-Not complex

While *PARN* acts as an important deadenylase, it operates in parallel with other deadenylase enzymes, such as the Ccr4-Not complex (Chalabi Hagkarim and Grand 2020), to mediate efficient mRNA turnover and degradation. The Ccr4-Not complex is a multiprotein assembly that plays a central role in deadenylation and is functionally distinct from *PARN*, as discussed in detail in section 1.12 of this thesis. Together, these deadenylases highlight the complexity and coordination required for post-transcriptional mRNA regulation.

The deadenylase enzymes involved in mRNA degradation are summarised in **Table 1.1**.

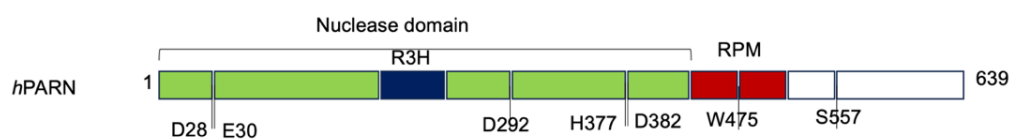
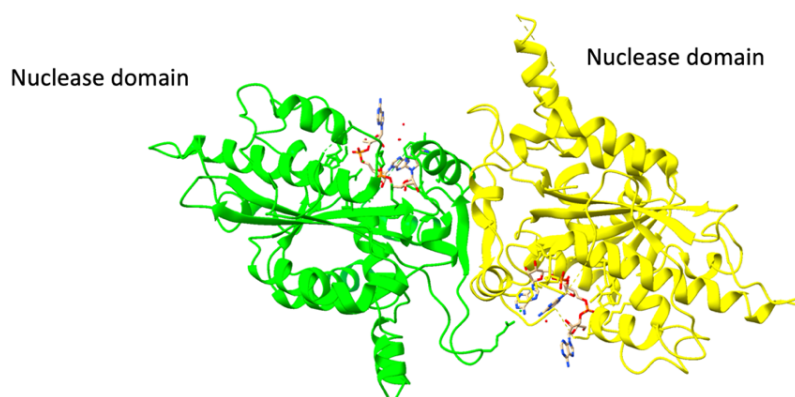
A**B**

Figure 1.7 Structure of PARN.

(A) Schematic representation of the human *PARN* (*hPARN*) protein, highlighting its functional domains and key residues. The nuclease domain (green) spans most of the protein and is responsible for its deadenylase activity. The R3H motif (dark blue) is a conserved RNA-binding domain involved in substrate recognition. Critical catalytic residues (D28, E30, D292, H377, D382) and an RPM motif (Nagy and Watzele), which plays a role in protein interactions, are indicated. The total length of *hPARN* is 639 amino acids (Virtanen et al. 2013).

(B) A ribbon diagram of the PARNn–RNA complex. The two molecules are shown in yellow and green. Nucleotides are shown in stick model (Wu et al. 2005) (PDB ID: 2A1R) .

Table 1.1 Enzymes involved in mRNA deadenylation, decapping and degradation

Enzyme	Human name	Function
Deadenylase enzyme	<i>CNOT7</i>	DEDD-type deadenylases, member of Ccr4-Not complex
	<i>CNOT8</i>	DEDD-type deadenylases, member of Ccr4-Not complex
	<i>PARN</i>	DEDD-type deadenylases, member of Ccr4-Not complex
	<i>Pan2/Pan3</i>	DEDD-type deadenylase that exists as dimer with PAN3.
	<i>CNOT6</i>	EEP-type deadenylase, member of Ccr4-Not complex
	<i>CNOT6L</i>	EEP-type deadenylase, Member of Ccr4-Not complex
Decapping enzyme	<i>Dcp1a, Dcp1b</i>	Member of decapping complex with Dcp2
	<i>Dcp2</i>	Catalytic pyrophosphatase subunit of decapping complex
	<i>Lsm 1-7</i>	Member of Lsm complex with subunits 1-7
Cytoplasmic exonuclease enzyme	<i>Xrn1</i>	Cytoplasmic 5'-3' exonuclease
	<i>Xrn2</i>	Multisubunit exonuclease

1.9 Decapping

Decapping is the removal of the 7-methylguanosine (m7G) cap from mRNA, a process critical for the degradation of transcripts marked for turnover. This cap, added post-transcriptionally, stabilises mRNA and facilitates translation, with decapping serving as a key regulatory step in mRNA decay (Parker 2012).

Decapping enzymes serve **(Table 1.1)** as pivotal regulators of mRNA stability and turnover, playing a critical role in gene expression control and cellular homeostasis. These enzymes act by removing the protective 5' cap structure of mRNAs, a step that commits transcripts to degradation and prevents further translation.

1.9.1 The *Dcp1-Dcp2* complex

The process of mRNA decapping is conducted by a complex formed by the *Dcp1* and *Dcp2* proteins, and it is affected by various additional factors (Parker 2012). The primary decapping enzyme, *Dcp2*, is central to the 5'-to-3' mRNA decay pathways. *Dcp2* belongs to the NUDIX hydrolase family and is responsible for hydrolysing the 5' cap structure, resulting in the production of m⁷GDP (7-methylguanosine diphosphate) and a 5' monophosphorylated RNA (She et al. 2008). This decapped RNA is then rapidly degraded by the exonuclease *Xrn1*, completing the 5'-to-3' degradation process. *Dcp2* activity is highly regulated to ensure precise control over mRNA decay. It interacts with a range of protein cofactors, including *Dcp1*, which functions as an activator to enhance the enzymatic activity of *Dcp2* and stabilise its binding to target mRNAs (She et al. 2008). Additional regulatory proteins, such as *Edc3* and *Edc4*, further influence the activity and recruitment of *Dcp2* to mRNAs, integrating decapping into broader RNA processing and degradation pathways (Coller and Parker 2004).

The decapping process by *Dcp2* is not random but highly selective, targeting specific transcripts often marked for decay due to errors in translation, the presence of premature stop codons, or signals from other decay pathways like nonsense-mediated decay (NMD) (Coller and Parker 2004). *Dcp2* is also involved in the regulation of mRNA during stress responses, where stalled or untranslated mRNAs are directed to processing bodies (P-bodies), dynamic cytoplasmic granules that serve as sites for mRNA storage, decay, and regulation. Within P-bodies, *Dcp2* operates alongside other decapping activators and repressors, fine-tuning mRNA decay under various cellular conditions (Coller and Parker 2004; Li and Kiledjian 2010).

1.9.2 Decapping Scavenger (*DcpS*) enzyme

In the 3'-to-5' decay pathway, the scavenger decapping enzyme (*DcpS*) plays a complementary role. *DcpS* proteins are members of the HIT family of pyrophosphatases and use a histidine triad to perform catalysis (Coller and Parker 2004). Unlike *Dcp2*, which acts on full-length mRNAs, *DcpS* specifically targets short-capped oligonucleotides

generated as byproducts of exosome-mediated mRNA degradation. *DcpS* cleaves these cap remnants to release m⁷GDP, recycling this metabolite for cellular use and ensuring the efficient removal of residual decay products. This activity prevents the accumulation of inactive RNA fragments that could interfere with normal cellular processes. *DcpS* is particularly important for maintaining the efficiency of the exosome complex, a major player in 3'-to-5' mRNA decay (Li and Kiledjian 2010).

1.9.3 The *Lsm1-7-Pat1* complex

The *Lsm1-7* complex, along with *Pat1*, plays a pivotal role in the 5'-to-3' mRNA degradation pathways by linking deadenylation to decapping (Tharun and Parker 2001). The *Lsm1-7* complex is composed of seven subunits organised in a heptameric ring that binds short poly(A) oligonucleotides *in vivo* and is localised to P-bodies (Sharif and Conti 2013). *Pat1p*, although its precise biochemical function remains unclear, is also part of this complex and contributes to its activity (Coller and Parker 2004).

While several proteins are critical for efficient decapping of most normal mRNAs *in vivo*, they are not strictly required for the decapping process itself (Coller and Parker 2004). This group includes the *Lsm1-7* complex, *Pat1p*, and *Dhh1p*, a member of the DEAD box ATPase family (Coller and Parker 2004).

1.10 Degradation

1.10.1 *Xrn1* enzyme

Following decapping, mRNAs are degraded in the 5' to 3' direction by the *Xrn1* nuclease (**Figure 1.8**), which exhibits a preference for mRNA substrates containing a 5' monophosphate (Hsu and Stevens 1993). *Xrn1* contains two highly conserved domains that form the enzyme's active site, stabilised through interactions with additional structural domains. The active site is capable of coupling duplex unwinding with the enzyme's processivity, enabling *Xrn1* to degrade structured RNAs without requiring a helicase (Parker 2012; Jinek, Coyle, and Doudna 2011). The active site of *Xrn1* interacts with two Mg²⁺ ions and is notably unable to accommodate 5'-groups larger than a 5'-phosphate group (Jinek, Coyle, and Doudna 2011). This limitation prevents *XRN1* from binding to and degrading capped mRNAs or RNA duplexes. However, *Xrn1* is capable of unwinding short RNA duplexes if a sufficiently long single-stranded RNA 5'-end is available (Jinek, Coyle, and Doudna 2011).

A related enzyme, *Rat1*, is primarily localised to the nucleus, where it participates in nuclear RNA processing and degradation. However, when *Rat1* is mislocalised to the cytoplasm, it can substitute for *Xrn1* in mRNA degradation, demonstrating that *Xrn1*-specific protein–protein interactions are not essential for its function. *Xrn1* activity is regulated by adenosine 3', 5' biphosphate (pAp), a product of sulfate assimilation, which acts as an inhibitor. This regulatory mechanism allows cells to modulate *Xrn1* activity during specific cellular responses (Parker 2012).

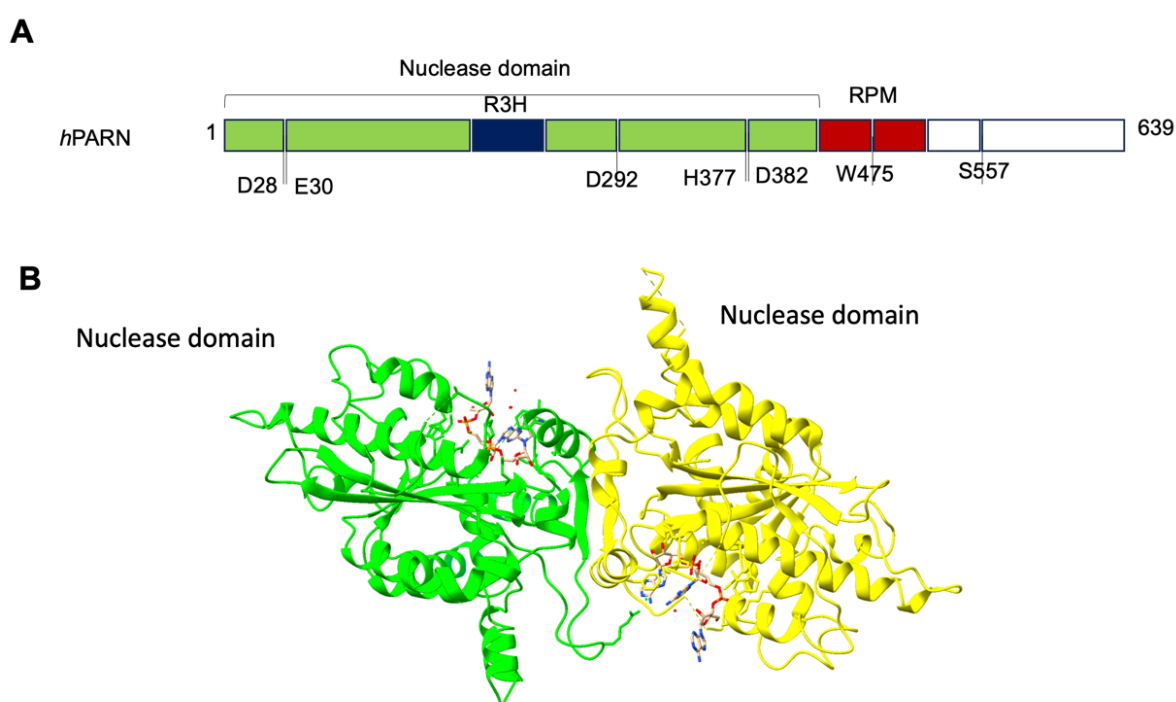


Figure 1.8 Structure of the *D. melanogaster* *Xrn1* Enzyme.

The exoribonuclease domain, depicted in tan, houses the active site that binds a Mg^{2+} ion and the substrate, both represented in blue. Additional regions of *XRN1* are highlighted in distinct colours: PAZ in red, KOW in green, helix in purple, and SH3L in yellow (Jinek, Coyle, and Doudna 2011) (PDB ID: 2Y35).

1.10.2 Exosome enzyme

The second pathway of mRNA decay following deadenylation involves degradation in the 3' to 5' direction, a process mediated by the exosome. The exosome is a multiprotein complex comprising 10 core subunits, including six RNase PH family members and three small RNA-binding proteins, which together form a ring-shaped structure similar to bacterial PNPs (Parker 2012). Additionally, the Rrp44/Dis3 protein, a key component of the exosome, possesses both exonuclease and endonuclease activities (Schaeffer et al. 2009). Beyond its cytoplasmic role in mRNA decay, the exosome is also actively involved in a variety of nuclear RNA processing and degradation pathways. In the nucleus, the exosome interacts with additional factors, including Rrp6, a 3' to 5' exonuclease, as well as Rrp47 and Mpp6, which contribute to its nuclear-specific functions. These interactions enable the exosome to play a critical role in maintaining RNA homeostasis in both the cytoplasm and nucleus (Parker 2012).

1.11 Other mRNA decay pathways

Defective mRNAs are degraded more rapidly than normal mRNAs through a process called mRNA surveillance. This mechanism allows eukaryotic cells to detect abnormalities in mRNAs and utilise specific enzymes to selectively degrade them (van Hoof and Wagner 2011). There are three translation-linked mRNA surveillance pathways that facilitate the degradation of defective mRNAs, thereby preventing the synthesis of potentially harmful proteins. Nonsense-mediated decay (NMD) identifies mRNAs containing premature stop codons and inhibits the production of truncated proteins. No-go decay detects mRNAs where ribosomes have stalled during the process of translation. Nonstop decay (NSD) serves as a quality control mechanism for nonstop mRNAs, recognising stalled ribosomes at the 3' end of transcripts and directing them for rapid degradation. While these specialised pathways utilise the same decay enzymes involved in the degradation of normal mRNAs, they differ in their mechanisms for identifying and targeting faulty transcripts for degradation (Jamar, Kritsiligkou, and Grant 2018).

1.11.1 Nonsense-mediated decay (NMD)

Nonsense-mediated decay (NMD) is one of the most extensively studied mRNA surveillance mechanisms (**Figure 1.9**). NMD is responsible for identifying and degrading

transcripts that contain premature termination codons (PTCs). PTCs can result from a variety of causes, including mutations, frame-shifts, errors in RNA processing, inefficient translation initiation, or extended 3' untranslated regions (UTRs). If left untranslated, these transcripts could produce truncated proteins with potentially harmful or dysfunctional properties. The NMD pathway is conserved across all eukaryotes, and its core components, UPF1, UPF2, and UPF3, exhibit a high degree of conservation (Conti and Izaurralde 2005).

The NMD pathway operates through the sequential assembly and transition of protein complexes. When a ribosome encounters a PTC, the SURF complex forms, comprising UPF1, SMG1, and the translation release factors eRF1 and eRF3. Upon interaction with UPF2 and UPF3, which are often bound to the exon junction complex (EJC) downstream of the PTC, the SURF complex transitions to the decay-inducing (DECID) complex. This transition releases the ribosome and eRFs while UPF1 is phosphorylated by SMG1 (Hug, Longman, and Caceres 2016; Kashima et al. 2006; Melero et al. 2014). Phosphorylated UPF1 then recruits additional factors such as SMG5, SMG6, and SMG7, which subsequently initiate mRNA degradation through exonucleolytic and endonucleolytic pathways (Schoenberg and Maquat 2012). This mechanism ensures that aberrant mRNAs are efficiently targeted and degraded, maintaining the fidelity of gene expression (Hug, Longman, and Caceres 2016).

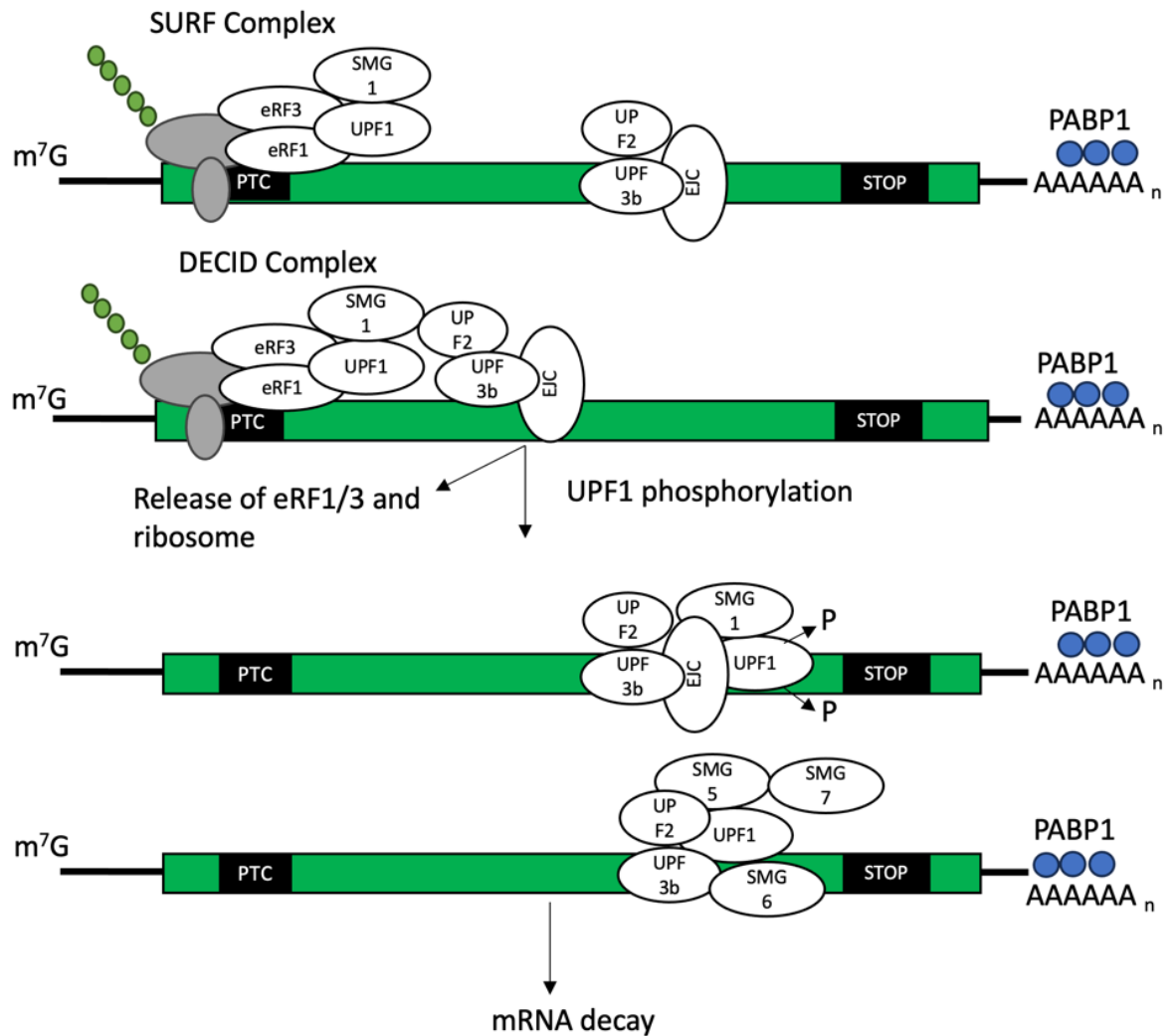


Figure 1.9 Nonsense mediated mRNA decay (NMD).

When a ribosome encounters a premature termination codon (PTC), it facilitates the recruitment of UPF1 and SMG1 through interactions with the release factors eRF1 and eRF3, forming the SURF complex. In the presence of UPF2 and UPF3, which may be associated with the exon junction complex downstream of the PTC on the mRNA, the ribosome releases eRF1 and eRF3. SMG1 then phosphorylates UPF1, which subsequently initiates the recruitment of SMG5, SMG6, and SMG7. The SMG5-SMG7 heterodimer, in turn, attracts mRNA decay factors, ultimately leading to the degradation of the faulty mRNA transcript.

1.11.2 Non-stop decay

Non-stop decay (NSD) targets mRNAs that fail to contain a stop codon (**Figure. 1.10**). These defective transcripts can arise from events such as breakage or the absence of an in-frame stop codon, leading to translation extending into the poly(A) tail. Premature polyadenylation is believed to be a significant contributor to the generation of mRNAs targeted by this pathway. In addition to preventing the production of faulty proteins, NSD also plays a role in ribosome recycling (Hug, Longman, and Caceres 2016).

Current evidence supports the existence of two distinct NSD pathways, which are likely to function together. The original model of NSD, conserved across yeast and mammalian cells, involves the cytoplasmic exosome, the SKI complex (comprising Ski2, Ski3, and Ski8), and the adaptor protein Ski7 (Hug, Longman, and Caceres 2016). In this pathway, the stalled ribosome at the 3' end of the transcript initiates the process. Ski7, through its C-terminal region, which shares structural similarity with the GTPase domains of elongation factor-1A (EF1A) and eRF3, binds to the empty A site of the ribosome, facilitating its release. Subsequently, Ski7 recruits the exosome and its associated SKI complex, which degrade the transcript by deadenylation followed by 3'→5' exonucleolytic decay (Hug, Longman, and Caceres 2016; Frischmeyer et al. 2002).

In cases where Ski7 is absent, a secondary pathway can mediate NSD through a 5'→3' mRNA-decay mechanism, as observed in *S. cerevisiae* (Inada and Aiba 2005). This alternative pathway likely results from the removal of PABP by the translating ribosome, rendering the transcript vulnerable to decapping and reducing translational efficiency. These two mechanisms collectively ensure the efficient degradation of defective mRNAs lacking stop codons, maintaining cellular quality control.

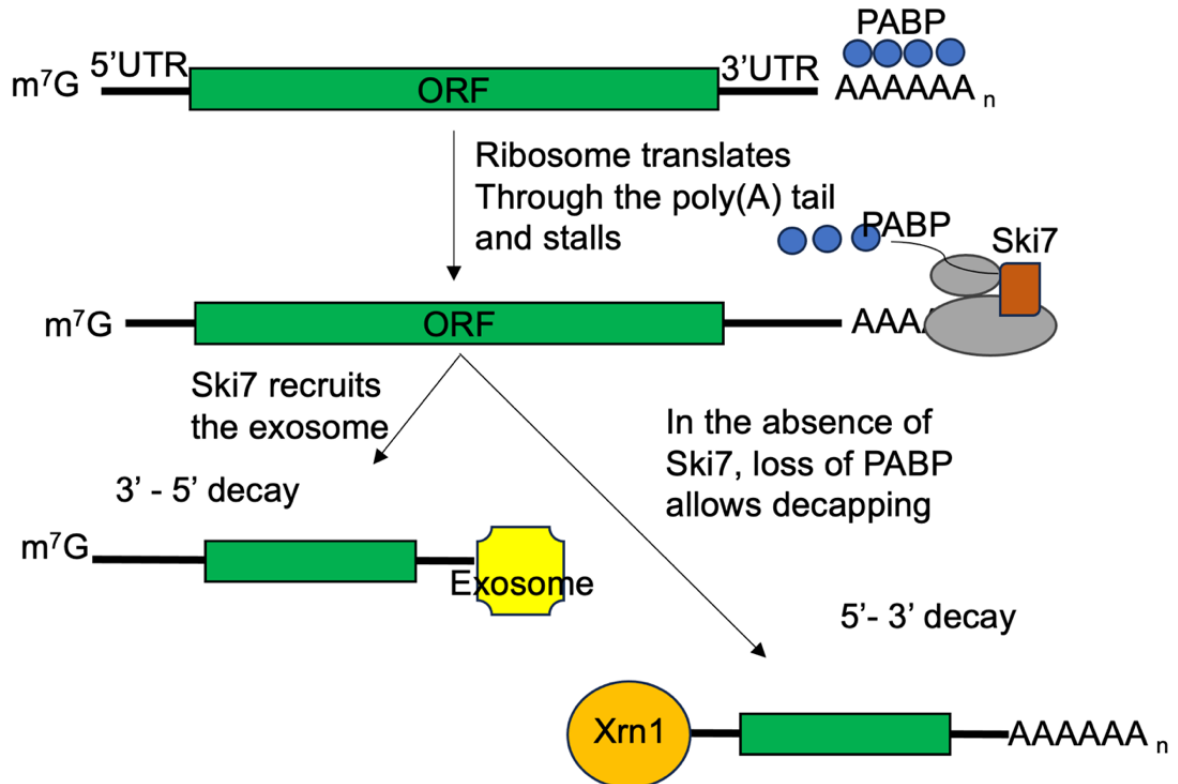


Figure 1.10 Mechanism of Non-Stop Decay (NSD).

The process of non-stop decay (NSD) is a surveillance mechanism that targets mRNAs lacking stop codons. When a ribosome stalls through the poly(A) tail and stalls, Ski7 is recruited to the empty A site of the ribosome. Ski7 then facilitates the recruitment of the exosome complex, which degrades the transcript in the 3'→5' direction. In the absence of Ski7, the removal of PABP (poly(A)-binding protein) from the transcript makes the mRNA susceptible to decapping. Once decapped, the mRNA undergoes 5'→3' decay mediated by the exonuclease *Xrn1*.

1.11.3 No-go decay

No-go decay was first discovered in *yeast* and remains the least understood among the pathways. NGD prevents the accumulation of faulty transcripts by detecting stalled ribosomes on mRNA and triggering endonucleolytic cleavage near the stall site (Garneau, Wilusz, and Wilusz 2007). This releases the ribosome and generates mRNA fragments, which are subsequently degraded by *Xrn1* and the cytoplasmic 3'-5' exosome. While the precise mechanism and identity of the endonuclease remain unclear, the process requires Dom34 and Hbs1, which are homologous to eRF1 and eRF3, respectively, and

may interact directly with stalled ribosomes. Hbs1 is also related to the NSD factor Ski7, suggesting a similar role in ribosome recycling (Doma and Parker 2006). Although the significance of NGD for cellular viability is not fully understood, evidence from *Arabidopsis thaliana* implies that ribosome stalling and NGD may be evolutionarily conserved processes (Doma and Parker 2006; Onouchi et al. 2005).

1.12 The Ccr4-Not complex

The Ccr4-Not (carbon catabolite repression – negative on TATA-less) complex is a multi-subunit protein assembly found in all eukaryotic organisms, playing a critical role in managing RNA metabolism throughout its various stages, from synthesis to degradation (Pavanello, Hall, and Winkler 2023). Over time, it has become recognised as a vital regulator of gene expression balance in eukaryotes (Collart 2016). It promotes the deadenylation process, during which the removal of poly(A) tails from mRNA leads to inhibited translation and signals the mRNA for degradation (Collart 2003).

In early studies conducted on yeast, the Ccr4-Not complex exists in two distinct sizes, with approximate masses of 1.0 MDa and 1.9 MDa (Xu et al. 2014; Nasertorabi et al. 2011). The *yeast* Ccr4-Not complex consists of a scaffold protein, Not1, and a number of evolutionarily conserved core proteins that dock onto Not1 (**Table 1.2**) (Collart 2016). The primary proteins that make up the core include the Not2-Not5 heterodimer, the Caf40 subunit, and the catalytic subunits Ccr4 and Caf1 (Pop2). In mammals, including humans, their homologs assemble into a comparable multi-subunit complex that plays a crucial role in regulating various cellular processes (Albert et al. 2000; Viswanathan et al. 2003). While the *yeast* and human Ccr4-Not complexes share similarities, there are notable differences between them (**Figure 1.11 A**).

Additional subunits in the human Ccr4-Not complex include *RQCD1* (also known as *CNOT9/Not9/Rcd1/Caf40*), *Caf130*, *Not10 (CNOT10)*, *C2orf29 (CNOT11)*, and *TAB182*, are species specific. *Caf130* is unique to *yeast* (Collart and Panasenko 2012), whereas *CNOT10*, *C2orf29*, and *TAB182* are components of the human complex (Lau et al. 2009). Among these, *Not10* and *C2orf29* are conserved across *metazoans* (Xu et al. 2014).

Another notable difference is that the Not4 RING E3 ubiquitin ligase is a conserved protein but serves as a stable component of the Ccr4-Not complex only in *yeast*. *CNOT4* is not as stably associated as the other subunits in mammalian cells (Mostafa et al. 2020). The human Ccr4-Not complex comprises five core subunits: *CNOT1*, *CNOT2* and *CNOT3*, which together form the NOT module, and the catalytic module *CNOT7/CNOT8* (equivalent to *Caf1* in *yeast*) and *CNOT6/6L* (equivalent to Ccr4). *CNOT1* has three primary regions: the N-terminal domain, which associated with *CNOT11*, which interacts with *CNOT10*. Together, these proteins create the N-terminal module. The deadenylase module associates with the central region through a MIF4G domain found in *CNOT1*, while *CNOT7/CNOT8* acts as a bridge between *CNOT1* and *CNOT6/CNOT6 L* (Collart and Panasencko 2017). Additionally, a neighbouring domain, DUF3819, of unknown function in the central region serves as the binding site for *CNOT9*. The NOT module interacts with the C-terminal part of *CNOT1*. According to their interaction map, *CNOT2-CNOT3* heterodimer tethers to *CNOT1* via their N-terminal NOT1 anchor regions (NARs), while *CNOT3* directly interacts with *CNOT2* through their respective Not-Box domains; the Not-Box domains do not directly contact *CNOT1* (Xu et al. 2014) (Figure 1.11 B).

Table 1.2 Name of the Ccr4-Not complex in *Drosophila*, *Yeast* and Humans (Collart 2016)

Core name	<i>Drosophila</i> Gene name	<i>Yeast</i> (<i>Saccharomyces cerevisiae</i>)	<i>Homo sapiens</i> Gene name
Not1	<i>Not1</i>	<i>Not1/CDC39</i>	<i>CNOT1</i>
Not2	<i>Regena</i>	<i>Not2/CDC36</i>	<i>CNOT2</i>
Not5	<i>Not3</i>	<i>Not5</i>	<i>CNOT3</i>
Caf1	<i>Pop2</i>	<i>Caf1/Pop2</i>	<i>CNOT7 and CNOT8</i>
Ccr4	<i>twin</i>	<i>Ccr4</i>	<i>CNOT6 and CNOT6L</i>
Caf40	<i>Rcd1</i>	<i>Caf40</i>	<i>CNOT9/Rcd1/Caf40/RQCD1</i>
Not4	<i>Not4</i>	<i>Not4</i>	<i>CNOT4</i>
Not10	<i>Not10</i>		<i>CNOT10</i>
Not11	<i>Not11</i>		<i>CNOT10</i>
Caf130		<i>Caf130</i>	
Not3		<i>Not3</i>	
TAB182			<i>TAB182</i>

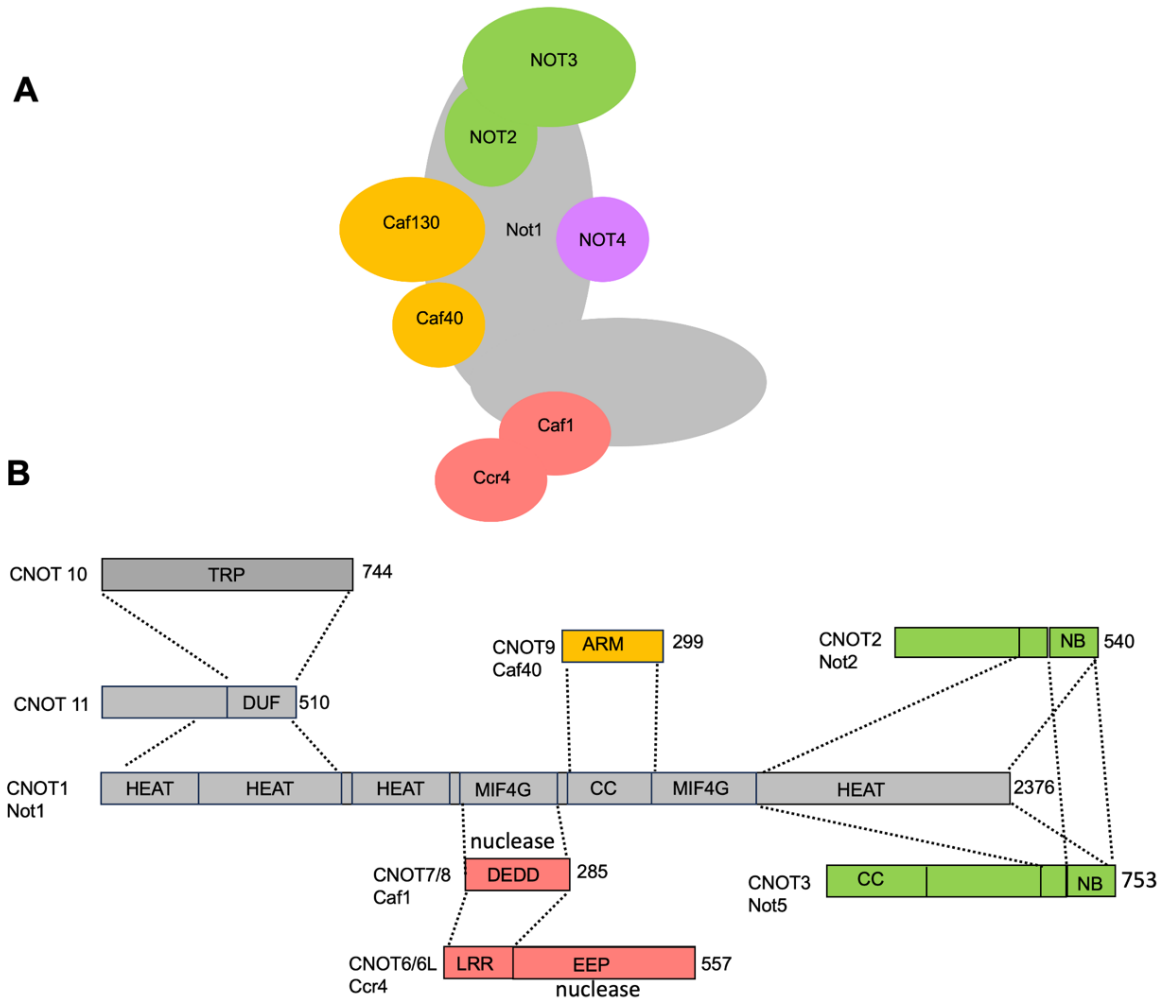


Figure 1.11 The Ccr4-Not complex subunits.

(A) The structure of the yeast Ccr4–Not complex shows Not1 as the central scaffold interacting with Not2 and Not3 (green), Not4 (purple), Caf40 and Caf130 (yellow), and the catalytic subunits Caf1 and Ccr4. (B) The human Ccr4–Not complex highlights CNOT1 as the scaffold protein, containing HEAT and MIF4G domains. It anchors CNOT2 and CNOT3 via their Not-Box regions (green and orange respectively). The catalytic subunits, CNOT7/8(Caf1) and CNOT6/6L (Ccr4), contain DEDD and EEP nuclease domains. Additional subunits include CNOT4 (with RING and RRM domains) and CNOT9 (Caf40), which interact through its ARM domain. CNOT11 (blue) and CNOT10 (magenta) associate with the N-terminal domain of CNOT1 directly and indirectly respectively (Chalabi Hagkarim and Grand 2020; Collart 2016).

1.13 The N-terminal module

The N-terminal module of the Ccr4-Not complex is built around the N-terminal domain of CNOT1 (CNOT1N) and, in *metazoans*, incorporates CNOT10 and CNOT11 (also referred to as C2ORF29 in humans) (**Figure 1.12**). The exact role of the CNOT10/CNOT11 module remains uncertain; however, since it is absent in yeast, it has been suggested that it may either be an optional component or serve as a functional equivalent to the *yeast* protein Caf130 (Bawankar et al. 2013; Xu et al. 2014). The purification of endogenous Ccr4-Not complexes from human and *Drosophila* cells confirmed that CNOT10 and CNOT11 are genuine core subunits (Mauxion et al. 2023). In vitro experiments using recombinant proteins demonstrated that both CNOT10 and CNOT11 can bind RNA and enhance the deadenylation activity of the Ccr4-Not complex. At the cellular level, these subunits were identified as regulators of interferon-responsive genes during HIV infection in primary T cells (Gordon et al. 2020). Additionally, in mouse embryonic stem cells, the depletion of CNOT10 resulted in the accumulation of specific transcripts. In *trypanosomes*, the deletion of the CNOT10 orthologue disrupted the structural integrity of the Ccr4-Not complex and hindered the degradation of numerous mRNAs (Farber et al. 2013; Mauxion et al. 2023). Despite these findings, the exact organisation and roles of the CNOT1N-CNOT10-CNOT11 module remain unclear (Mauxion et al. 2023).

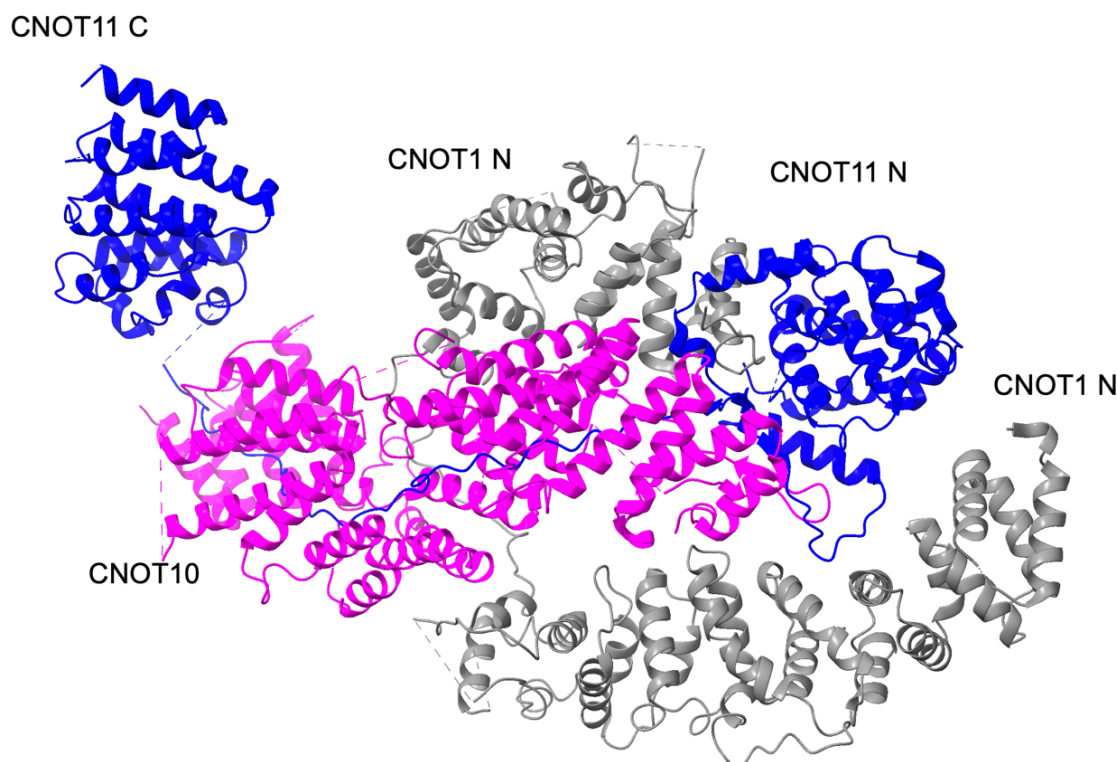


Figure 1.12 Structure of the N-terminal Module of the Ccr4-Not Complex.

The image depicts the structural organization of the N-terminal module of the Ccr4-Not complex, highlighting the interactions between CNOT1N, CNOT10, and CNOT11. The N-terminal domain of CNOT1 (CNOT1N) is shown in grey, serving as the structural scaffold for the assembly of the module. CNOT10 is represented in magenta, while CNOT11 is shown in blue, with its N-terminal domain (CNOT11 N) and C-terminal domain (CNOT11 C) labelled (PDB ID: 8BFI).

1.14 The nuclease module

The Ccr4-Not nuclease complex (**Figure 1.13**) is defined by two key subunits responsible for ribonuclease activity: Ccr4 and Caf1 in yeast. Ccr4 belongs to the EEP superfamily as was originally identified as a regulator of ADH2 in the yeast *Saccharomyces cerevisiae* (Shirai et al. 2014). The Ccr4 is distinguished by two highly conserved domains: a C-terminal deadenylase domain and an N-terminal leucine-rich repeat (LRR) domain. The LRR domain plays a crucial role in its interaction with CAF1 and its integration into the Ccr4-Not complex (Winkler and Balacco 2013). The structure and composition of Ccr4 exhibit notable differences across species. Firstly,

the activation domain located in the N-terminal domain of *Saccharomyces cerevisiae* Ccr4p is absent in orthologues from other eukaryotes, including mammals (Shirai et al. 2014). Secondly, while yeasts and metazoans such as *Drosophila melanogaster* and *Caenorhabditis elegans* possess a single Ccr4, vertebrates—including *Xenopus laevis*, *Danio rerio*, *Mus musculus*, and *Homo sapiens*—have two orthologues: CNOT6 (*Ccr4a*) and CNOT6L (*Ccr4b*) (Winkler and Balacco 2013; Shirai et al. 2014).

Caf1, a core deadenylase subunit of the Ccr4–Not complex, contains a ribonuclease (RNase) D domain that belongs to the DEDD superfamily (Collart and Panasenko, 2012; Winkler and Balacco, 2013). Initially, yeast *Caf1* (Caf1p) was discovered as a potential transcriptional regulator necessary for glucose derepression in *Saccharomyces cerevisiae* (Sakai et al. 1992). Later, research by Tucker et al. demonstrated that *Caf1p*, in conjunction with *Ccr4p*, functions as a key cytoplasmic mRNA deadenylase (Tucker et al. 2002). In mammals, including humans, *CAF1* is represented by two orthologues: CNOT7 (hCAF1a) and CNOT8 (hCAF1b). Analysis of the amino acid sequences of human CNOT7 and CNOT8 reveals a high level of similarity between the two proteins (Shirai et al. 2014; Winkler and Balacco 2013; Mittal et al. 2011). Human CNOT7 and CNOT8 interact with the middle region of CNOT1 via its MIF4G domain, which consists of pairs of antiparallel α -helices (Shirai et al. 2014). Furthermore, yeast two-hybrid analysis demonstrated that CNOT6 directly interacts with CNOT7 and CNOT8, but not with CNOT1. These results indicate that the association of CNOT7 and CNOT8 with CNOT1 is crucial for the incorporation of CNOT6 and CNOT6L into the Ccr4–Not complex, enabling them to perform their deadenylation functions (Lau et al. 2009).

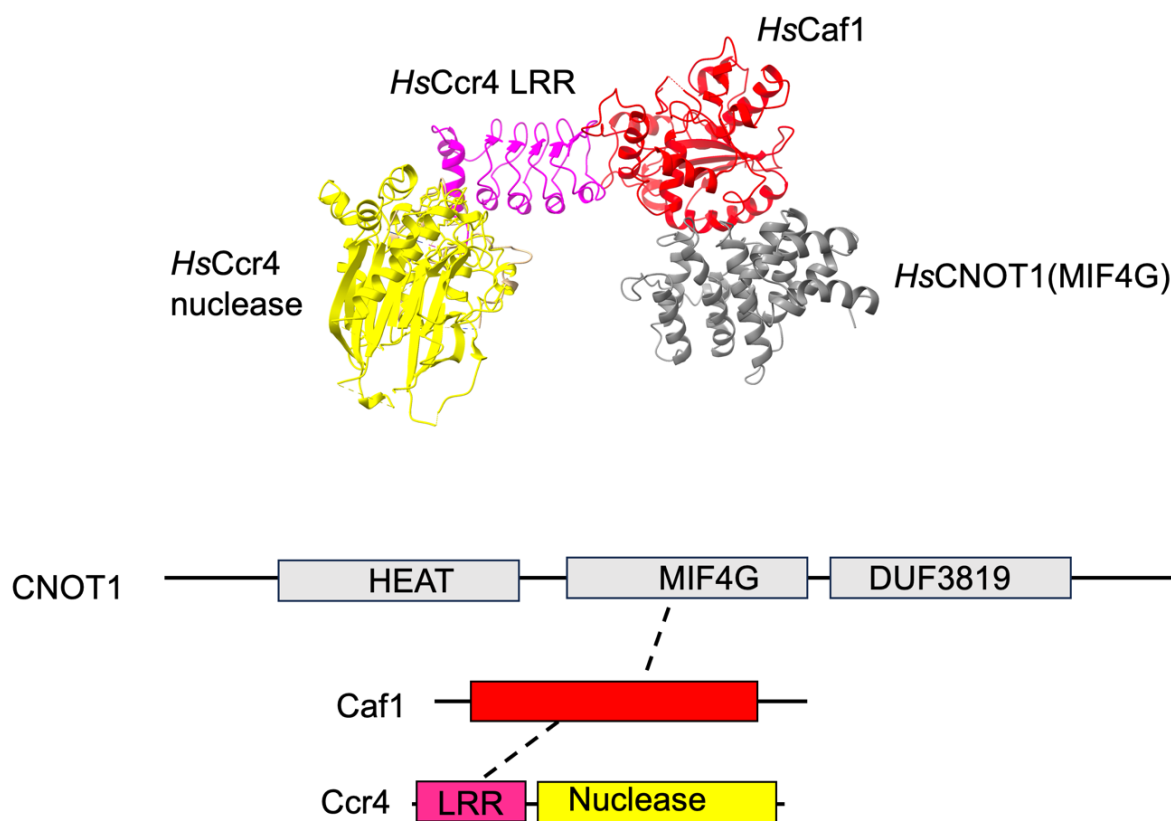


Figure 1.13 Structure of the Human Ccr4-Not nuclease module.

The structure shows HsCNOT1 (grey) anchoring HsCaf1 (red) and HsCcr4 (yellow and magenta) via its MIF4G domain. The LRR region (magenta) supports complex assembly, while the nuclease domain of HsCcr4 (yellow) drives mRNA deadenylation. The schematic highlights CNOT1 domains and binding sites for Caf1 and Ccr4 (Zhang et al. 2022) (PBD ID: 7VOI and 3NGQ).

1.15 The Not module

Another key module of the Ccr4–Not complex consists of CNOT1, CNOT2, and CNOT3 in humans and their counterparts Not1, Not2, and either Not3 or Not5 in yeast (Miller and Reese 2012; Xu et al. 2014).

1.15.1 CNOT1 subunit

CNOT1 acts as a scaffold and the largest subunits of the complex (>200 kD) (Chalabi Hagkarim and Grand 2020). Structurally, CNOT1 can be divided into three primary regions: the N-terminal, middle, and C-terminal regions, based on structural data and predictions. The N-terminal domain of CNOT1 interacts with CNOT11, which in turn associates with CNOT10. The middle region, known as the MIF4G (middle domain of eukaryotic initiation factor 4G) domain, directly binds to CAF1a (CNOT7) or CAF1b (CNOT8). Additionally, CAF40 (CNOT9) interacts with the remaining portion of the middle region of CNOT1, which includes the DUF3819 domain. The C-terminal domain of CNOT1 associates with the CNOT2/CNOT3 complex, facilitating the assembly of the NOT module (**Figure 1.14 A**). These interactions collectively enable CNOT1 to function as a scaffold, supporting the formation and structural integrity of the Ccr4–Not complex (Boland et al. 2013; Xu et al. 2014).

1.15.2 CNOT2-CNOT3 subunits

CNOT2 and CNOT3 are characterised by distinct structural domains. CNOT3 is comprised of three primary regions: an N-terminal domain featuring coiled-coil domains, a central region, and a C-terminal domain. In contrast, CNOT2 is composed of an N-terminal domain and a C-terminal domain but lacks the distinctly defined middle region present in CNOT3 (Boland et al. 2013) (**Figure 1.14 A**). In the human complex, both CNOT2 and CNOT3 contain a Not-Box motif, which facilitates their interaction Field (Miller and Reese 2012; Xu et al. 2014). Understanding the regulation and function of the NOT module necessitates an investigation into its three-dimensional structure, particularly the C-terminal domains of CNOT1, CNOT2, and CNOT3, which are responsible for the assembly of the module Field (Boland et al. 2013).

Recent three-dimensional structures of the human CNOT1–CNOT2–CNOT3 complexes have provided insights into the structure and assembly of the NOT module (Boland et al. 2013; Xu et al. 2014). These structures reveal a similar overall architecture, where CNOT2 and CNOT3 form a symmetric heterodimer through their Not-Box motifs

(**Figure 1.14 B**), creating what is known as the symmetric lobe. The Not-Boxes of CNOT2/Not2 and CNOT3/Not5 are located within their globular C-terminal domains. The structure of the Not-Box comprises three N-terminal α -helices and a C-terminal β -sheet made up of four β -strands in yeast and five in humans, with notable bending between β 3/ β 4 in yeast or β 4/ β 5 in humans (Xu et al. 2014). These domains share similarities with Sm domains but lack the characteristic Sm sequence motifs and an additional β -strand required for Sm–Sm dimerisation. The Not-Box regions of CNOT2/Not2 and CNOT3/Not5 enable dimerisation through their N-terminal helices, forming a highly symmetrical heterodimer with enhanced complementarity compared to a CNOT3 homodimer. A CNOT2 homodimer adopts a distinct dimerisation mechanism, involving domain swapping via the N-terminal helices, which differs from the arrangement seen in both CNOT3 homodimers and CNOT2/CNOT3 heterodimers. Additionally, the Not-Box structures of CNOT2/Not2 and CNOT3/Not5 are highly similar, with an r.m.s.d. of less than 1.3 Å when superimposed (Xu et al. 2014; Boland et al. 2013). The heterodimer interface is further stabilised by N-terminal sequences, referred to as connector sequences (CSs), which encircle the Not-Box regions of their respective partners, functioning like clamps (Boland et al. 2013).

Moreover, the CNOT2–CNOT3 heterodimer is tethered to the conserved helical surface of the N-terminal subdomain (N-SD) of CNOT1 exclusively via the N-terminal NOT1 anchor regions (NARs) of CNOT2 and CNOT3. Importantly, the Not-Box domains themselves do not contact CNOT1. (**Figure 1.14 C**) (Xu et al. 2014; Boland et al. 2013). Each NAR comprises an N-terminal α -helix (α N) and a C-terminal domain (NAR-C), the latter of which lacks any defined secondary structure. The α N helices insert into grooves along the lateral surface of CNOT1, forming stable contacts. The extreme N-terminal residues of CNOT2 penetrate a cleft formed by loops L15 and L19 and helix α 21 within the C-terminal subdomain (C-SD) of CNOT1 (Xu et al. 2014; Boland et al. 2013). Meanwhile, the NAR-Cs conform to the conserved helical surface of the CNOT1 N-SD, creating additional points of interaction. The CNOT3 NAR-C, being shorter, crosses the surface of helices α 4, α 6, and α 8 of CNOT1, engaging these regions through a combination of hydrophobic and polar interactions before entering the junction. In contrast, the CNOT2 NAR-C takes a longer, zigzagging path across the surface of CNOT1. This region contains several aromatic residues (Y396, F399, and W403) that insert into a hydrophobic cleft formed by helices α 8, α 9, and α 11 of CNOT1. Additionally,

residues 382–395 of CNOT2 form extensive contacts with helix $\alpha 11$, resulting in well-defined interactions that contribute to the stability of the assembly (Xu et al. 2014; Boland et al. 2013).

The NAR-Cs connect to the (CSs) at the junction between the two lobes of the structure. The core of this junction is stabilised by hydrophilic interactions involving the N-terminal domains of the CSs, with contributions from specific residues in CNOT2 (R409, Q411, D412, and D414) and CNOT3 (E632, R633, R635, and R640). Additionally, the relative orientation of the two lobes is maintained by the C-terminal tail of CNOT2, which clamps the CNOT3 CS within a tunnel and loops back to interact with CNOT1 (specifically helix $\alpha 11$), crossing over the CNOT3 NAR-C in the process. These interactions, particularly those mediated by the C-terminal tail of CNOT2, are likely to occur at a late stage in the assembly of the NOT module and impose additional topological constraints. This structural arrangement prevents the dissociation of CNOT3 in the presence of CNOT2, thereby reinforcing the stability of the module. Furthermore, this explains why the CNOT3 Not-box domain is essential for incorporation of CNOT3 into the Ccr4–Not complex: deletion of the Not-Box prevents formation of the CNOT2–CNOT3 heterodimer, resulting in a loss of CNOT3 interaction with CNOT1, as the isolated NAR of CNOT3 is insufficient to maintain stable binding to CNOT1. Moreover, surface conservation analysis suggests that the interfaces within the trimeric complex are evolutionarily conserved, indicating that this mode of assembly is preserved across multiple species (Boland et al. 2013; Xu et al. 2014).

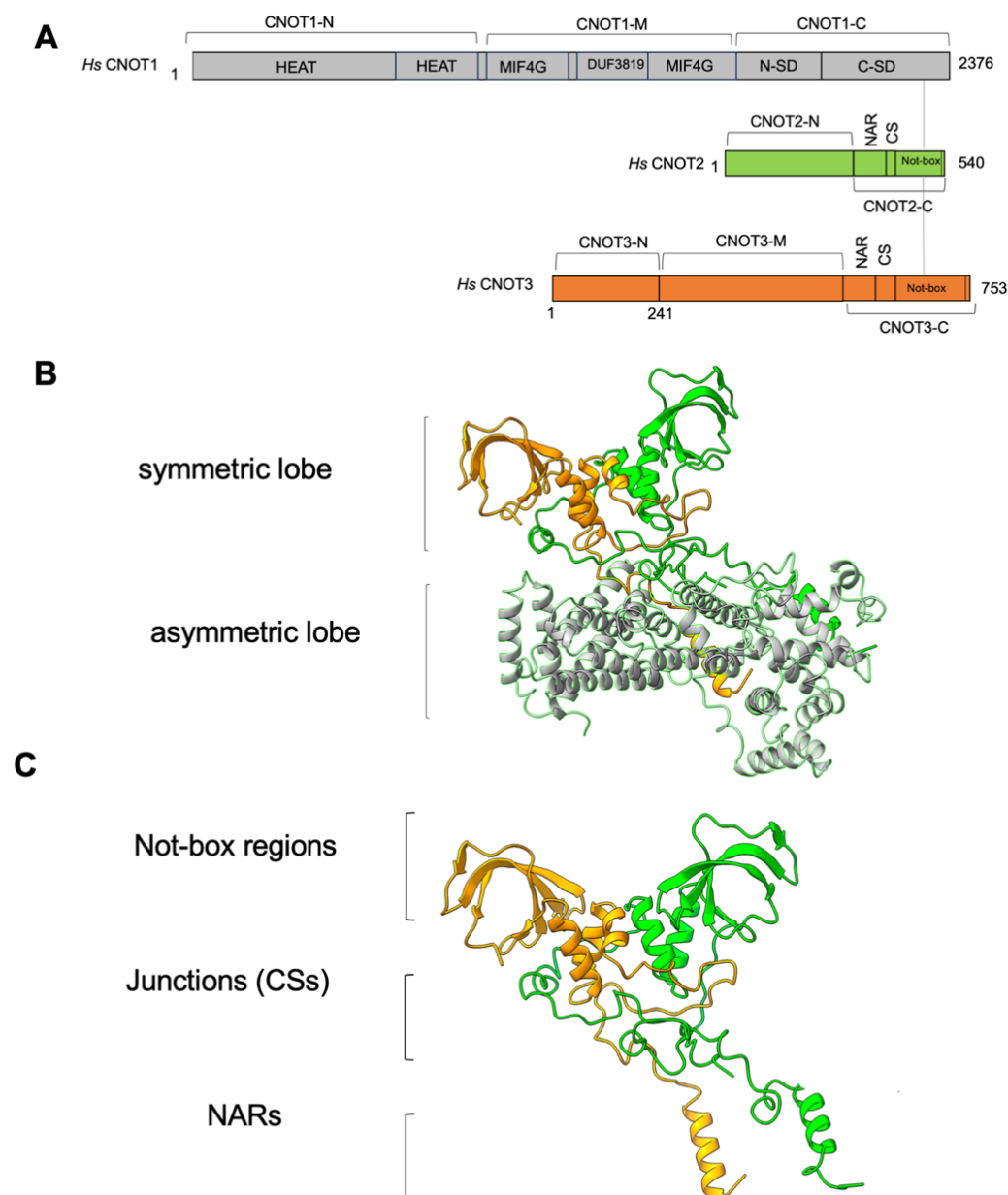


Figure 1.14 Structural organisation of the Human NOT Module.

(A) Domain architectures of CNOT1, CNOT2, and CNOT3. CNOT2 and CNOT3 contain N-terminal anchor regions (NAR), connector sequences (CS), and C-terminal Not-Box regions that interact with CNOT1. (B) The structure of the human NOT module shows CNOT1 Field as the scaffold protein interacting with CNOT2 (green) and CNOT3 (orange) via their Not-Box regions. (C) The CNOT2–CNOT3 heterodimer shows their interlocking interaction through the Not-Box regions and junctions (CSs). The NAR domains in CNOT2 and CNOT3 anchor these subunits to CNOT1, stabilising the assembly of the entire complex.

1.16 Physiological role of the NOT Module

1.16.1 *The physiological role of CNOT1*

Even though CNOT1 is not known to have any enzymatic activity, its role as a scaffold is crucial for the deadenylase functions of the CNOT complex. Research indicates that when CNOT1 is depleted in HeLa cells, the levels of CHOP mRNA rise, leading to the activation of caspase-4 and resulting in apoptosis due to ER stress. This demonstrates that CNOT1 is vital for cell viability and proliferation (Chalabi Hagkarim and Grand 2020; Shirai et al. 2014). Additionally, the Depletion of CNOT1 results in the destabilisation of the whole complex and degradation of some other subunits, such as CNOT2, CNOT6L, CNOT7, and CNOT9, but not CNOT3, in HeLa cells (Chalabi Hagkarim and Grand 2020).

1.16.2 *Physiological role of CNOT2-CNOT3*

The precise cellular function of the CNOT2 and CNOT3 modules also remains unclear, as neither protein possesses a catalytic domain. It is believed that this module primarily acts as a protein interaction platform, modulating the enzymatic activities of the Ccr4–Not complex. Supporting this, multiple studies have demonstrated that the downregulation of either Not2 or Not5 affects the levels of specific mRNAs and, in some cases, the length of poly(A) tails. These effects are likely due to the module's role in facilitating the recruitment of the Ccr4–Not complex to particular mRNAs Field (Collart and Panasenko 2017; Morita et al. 2011). It has also been noted to play a role in decapping fields (Collart and Panasenko 2017). It is important to consider that any function attributed to the phenotypes of cells deficient in Not2 or Not5 may be a consequence of compromised integrity of the Ccr4-Not complex. However, there is at least one instance in *Drosophila* in which Not3 interacts with the RNA-binding protein Bicaudal-C, facilitating the assembly of the complete Ccr4-Not complex onto mRNAs for the purpose of their deadenylation (Chicoine et al. 2007; Collart and Panasenko 2017). This module is also associated with transcription, where it interacts with a wide range of general and specific transcription factors. It is found at transcription sites and influences the presence of transcription factors at promoters, as well as global histone acetylation levels (Collart and Panasenko 2017; Collart 2016) . The Not2-Not5 heterodimer also plays a crucial role in translation, as it is required to prevent the aggregation of the largest subunit of RNA polymerase II during its synthesis (Collart and Panasenko 2017; Villanyi et al. 2014).

Several studies have demonstrated the role of CNOT3 in regulating gene expression. Research in MCF-7 cells revealed that CNOT3 suppresses transcription driven by ER alpha, as knockdown of CNOT3 resulted in increased expression of ER alpha target genes (Winkler et al. 2006). Another study conducted in HEK293T cells found that while the full-length CNOT3 does not influence transcription, its Not-Box domain functions as a repressor of reporter gene activity (Zwartjes et al. 2004). Besides, in mouse embryonic stem cells (ESCs), CNOT3 is essential for self-renewal and survival. Its depletion leads to rapid differentiation and a significant loss of the epiblast cell population, with downregulation of pluripotency markers such as *Oct4*, *Nanog*, and *Sox2*. Notably, in vivo deletion of *Cnot3* causes peri-implantation lethality, and ESCs lacking CNOT3 fail to maintain colony formation even under ground-state conditions. Mechanistically, CNOT3 co-occupies promoter regions of target genes with factors like *TRIM28*, *c-MYC*, and *ZFX*, forming a transcriptional module that regulates genes associated with the cell cycle and survival (Hu et al. 2009; Zheng et al. 2016).

CNOT3 has been identified as a critical factor for cell cycle progression and is essential for the proliferation of MCF-7 cells. Its knockdown results in cell accumulation in the G1 phase, suggesting its role in facilitating the transition into the S phase. However, CNOT3 is not required for cell survival (Mittal et al. 2011). Another study (Takahashi et al. 2012) reported the importance of CNOT3 during mitotic progression. Knockdown of *CNOT3* in Hela cells resulted in mitotic arrest. Evidence was presented that CNOT3 controls mRNA levels of MAD1, which is a protein involved in mitotic spindle assembly. CNOT3 might therefore represent a player in controlling the spindle assembly checkpoint.

1.16.3 Function of CNOT3 in mRNA decay and translation

CNOT3, as a component of the versatile Ccr4-Not complex, has been attributed with various molecular roles. These roles encompass the regulation of transcription, the removal of poly(A) tails from RNA, the degradation of mRNA, and the assessment of codon usage efficiency (Ghashghaei et al. 2024). The Ccr4-Not complex plays a critical role in linking translation and mRNA decay, with CNOT3/Not5 (Human/yeast) serving as a key subunit in this process (Absmeier et al. 2023). CNOT3/Not5 integrates signals from the translational machinery to modulate mRNA stability, thereby ensuring precise regulation of gene expression. Its function is particularly evident during conditions of translational stalling, where ribosomes encounter challenges in elongation due to

various factors, such as rare codons or limited availability of cognate tRNAs (described in section 1.6.1) (Pavanello, Hall, and Winkler 2023; Buschauer et al. 2020).

The relationship between these processes has been difficult to identify, but recent experiments conducted on the budding yeast *S. cerevisiae* suggest a potential mechanism (Buschauer et al. 2020). Work by Buschauer et al. (Buschauer et al. 2020) showed that when a ribosome stalls during translation, the A-site on the ribosome remains unoccupied due to the low abundance of the incoming t-RNA-aminoacyl complex, signalling a slowdown in decoding. In this context, Not5, the orthologue of human CNOT3 is recruited to the ribosome through its ability to bind the E-site (**Figure 1.15 A and B**). Cryo-electron microscopy studies have revealed the molecular details of this interaction, showing that the N-terminal domain of Not5 inserts into the E-site and stabilises the stalled ribosome by locking the L1 stalk in an open conformation. This conformational change effectively halts further elongation by preventing translocation of the ribosome. The engagement of Not5 with the ribosome not only indicates a stalled condition but also prepares the linked mRNA for deadenylation and eventual degradation through the 5'-3' decay pathway (Buschauer et al. 2020; Absmeier et al. 2023). The interaction of the N-terminal domain of CNOT3 with the ribosomal E-site is preserved in mammalian cells (Absmeier et al. 2023).

As it was thought in yeast the presence of rare codons in the A-site is a strong determinant for CNOT3 recruitment, recent research in HEK293T cells has provided new insights into the codon-specific recruitment of CNOT3 during translation (Zhu et al. 2024). In mammalian cells, the P-site codon plays a more significant role (Zhu et al. 2024). When the E-site becomes vacant due to the dissociation of the E-site tRNA, it creates an opportunity for CNOT3 to enter the E-site and probe the P-site tRNA. This probing involves assessing the structural compatibility of the P-site tRNA, particularly its D-arm and the conserved U13:A22:A46 triplet. If the P-site tRNA meets these structural requirements, CNOT3 stabilises on the ribosome and engages in downstream mRNA decay processes (Zhu et al. 2024).

Specific arginine codons, such as CGG, CGA, and AGG, are highly associated with CNOT3 recruitment to translating ribosomes. High-resolution structural analyses revealed that these codons promote recruitment due to their interaction with the D-arm of the P-site tRNA, facilitating stabilising hydrogen bonds with CNOT3 and enhancing its affinity for the ribosomal E-site (Zhu et al. 2024; Collart, Audebert, and Bushell 2023). Conversely,

tRNAs that lack this structural compatibility, such as those with an additional nucleotide in the D-loop, generate steric clashes that prevent CNOT3 recruitment (Zhu et al. 2024). This distinction ensures that only specific ribosomal conformations and codon-tRNA combinations are targeted by CNOT3 for mRNA degradation. Such selective interactions underline a sophisticated regulatory mechanism, where codon usage and tRNA structure jointly dictate the stability of the associated mRNA (Collart, Audebert, and Bushell 2023; Zhu et al. 2024).

Once recruited to a stalled ribosome, CNOT3/Not5 serves as a platform for the Ccr4-Not complex's enzymatic machinery. The deadenylase module of the complex, comprising CNOT6 and CNOT7, is activated to remove the poly(A) tail of the mRNA. Deadenylation destabilises the mRNA and primes it for decapping, a process mediated by the decapping machinery, which exposes the mRNA to 5'-to-3' exonucleolytic degradation by enzymes such as *Xrn1*. This sequential process ensures efficient degradation of the defective mRNA. Additionally, the recruitment of the E3 ubiquitin ligase CNOT4 by the Ccr4-Not complex facilitates the ubiquitination of ribosomal proteins, such as eS7, marking the stalled ribosome for quality control and potentially triggering its recycling or degradation (Absmeier et al. 2023; Collart, Audebert, and Bushell 2023).

Likewise, CNOT3 enhances translation efficiency by recruiting highly translated mRNAs, such as *c-MYC*, to active polysomes, ensuring ribosome occupancy and elongation on critical transcripts essential for rapid cell division. While global protein synthesis remains unaffected, CNOT3 selectively boosts AU3-biased codons linked to proliferation and destabilises less efficient or GC3-biased codons associated with differentiation (Ghashghaei et al. 2024).

Additionally, CNOT3 interacts with ribosomal proteins and elongation factors like EEF1G and EEF1E, which are crucial for translation elongation. Its depletion disrupts these interactions, impairing elongation and reducing synthesis of key growth-promoting proteins. These findings highlight CNOT3's role in codon-specific translation regulation and cell survival (Ghashghaei et al. 2024).

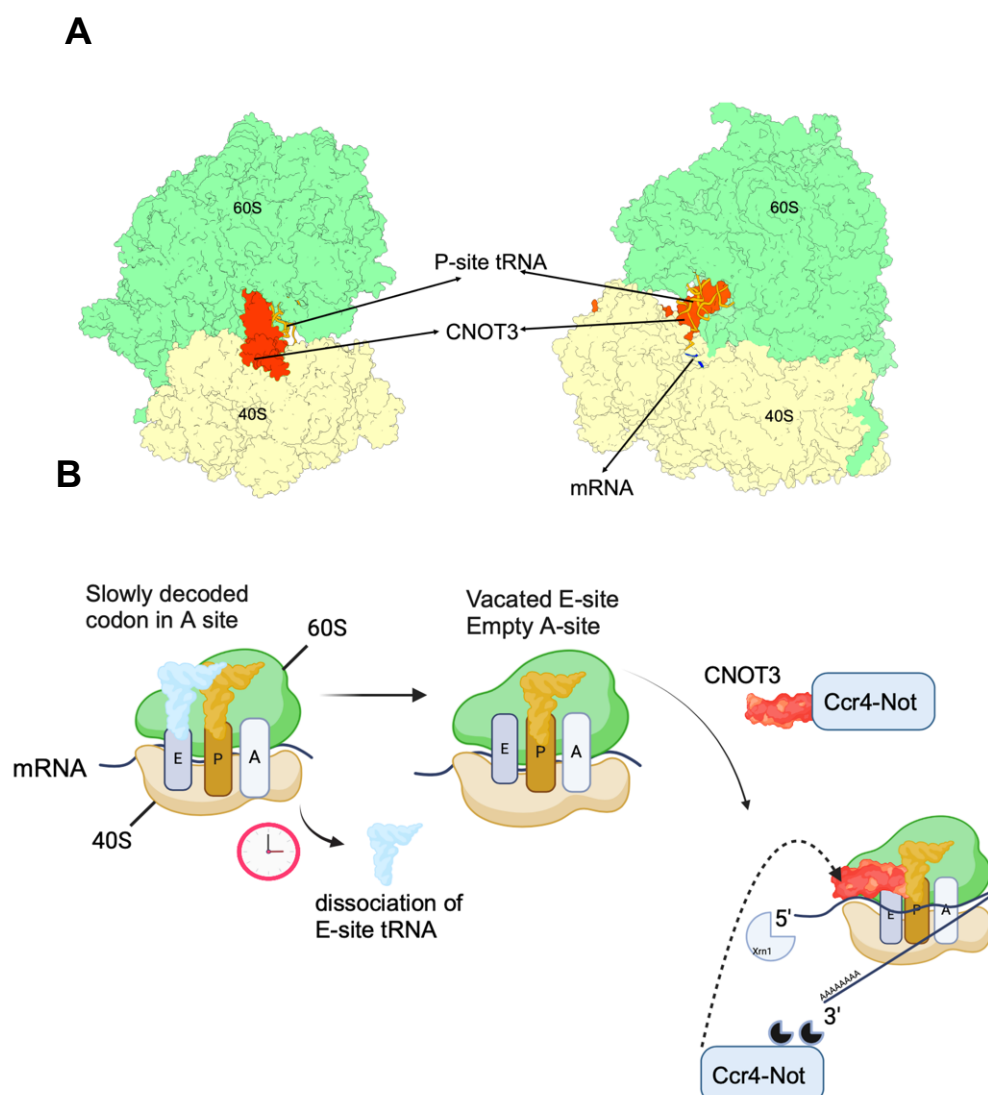


Figure 1.15 Mechanism of CNOT3-mediated mRNA degradation at stalled ribosomes.

(A) Structural representation of the ribosome highlighting the interaction between CNOT3 and the ribosomal P-site during translational stalling. The 60S large ribosomal subunit is shown in green, the 40S small ribosomal subunit in yellow, the P-site tRNA in orange, and CNOT3 in red (Zhu et al. 2024) (PDB ID 9C3H).

(B) Schematic diagram illustrating the mechanism of translational stalling and subsequent recruitment of CNOT3. A slowly decoded codon in the A-site results in dissociation of the E-site tRNA, leaving the E-site vacated and the A-site empty. This allows CNOT3 (red) to bind to the vacant E-site and probe the P-site tRNA (orange), assessing its structural features. This interaction enables CNOT3 to determine whether the ribosome state is compatible for its recruitment. When compatible P-site tRNAs are present, CNOT3 stabilises on the ribosome and recruits the Ccr4-Not complex. This complex initiates mRNA degradation through deadenylation, decapping, and 5'-to-3' exonucleolytic decay, selectively targeting stalled transcripts for degradation (Buschauer et al. 2020). The schematic was created using Biorender.com.

1.17 Involvement of CNOT3 in diseases

1.17.1 Involvement of CNOT3 in cancer

Potential roles for *CNOT3* have been identified in a number of diseases, with particularly significant implications in cancer. Exome sequencing studies have identified *CNOT3* as a tumour suppressor gene, with mutations found in approximately 7.9% of adult T-cell acute lymphoblastic leukaemia (T-ALL) cases (De Keersmaecker et al. 2013). These mutations frequently localise to the Arg57 residue within the N-terminal domain of the protein, which has been highlighted as a mutation hotspot. The Arg57 (corresponding to Lys-58 in *yeast*) mutations lead to splicing defects and significantly reduced *CNOT3* mRNA expression, undermining its tumour-suppressive functions. This effect was confirmed by mRNA expression analysis, which demonstrated that mutations in this region result in transcripts that are either absent or expressed at very low levels (De Keersmaecker et al. 2013).

To further investigate the role of *CNOT3* in cancer, a *Drosophila* eye cancer model was employed. In this model, reducing *Not3* (the *Drosophila* homolog of *CNOT3*) expression in a sensitised background led to a significant increase in tumour formation. For instance, tumour incidence increased from 8% in control RNA interference (RNAi) flies to as high as 90% with *Not3* knockdown. These findings underscore that *CNOT3* loss is sufficient to drive tumour development in a genetically sensitised environment, reinforcing its role as a tumour suppressor (De Keersmaecker et al. 2013).

Similarly, in colorectal cancer (CRC), *CNOT3* has been identified as a critical regulator of tumour aggressiveness and progression, further underscoring its versatile role in human diseases. Unlike its limited expression in normal colonic epithelial cells, *CNOT3* is significantly upregulated in CRC, particularly in advanced stages. Intriguingly, its expression is confined to a subpopulation of tumour cells with stem cell-like properties, capable of self-renewal and sustaining the malignancy. Mechanistically, *CNOT3* operates through dual regulatory roles. On the one hand, it activates transcriptional programs associated with embryonic stem cell-like traits, fostering tumour cell proliferation and survival. On the other hand, it represses genes silenced in poorly differentiated tumours, thereby preventing cellular differentiation and maintaining the malignant phenotype. Unlike its cytoplasmic role in other cancers, *CNOT3* in CRC predominantly functions in

the nucleus, interacting with chromatin to orchestrate the transcriptional dynamics that drive tumour aggressiveness (Cejas et al. 2017).

Moreover, recent analyses of RNA-sequencing datasets from The Cancer Genome Atlas (TCGA) revealed a striking increase in *CNOT3* expression in non-small cell lung cancer (NSCLC) tumours, particularly in subtypes like lung adenocarcinoma (LADC) and lung squamous cell carcinoma (LSqCC) (Shirai et al. 2019). Unlike some oncogenes, this overexpression does not stem from gene amplification but appears to be driven by alterations in upstream pathways or the tumour microenvironment (Shirai et al. 2019). One of the most profound effects of *CNOT3* in NSCLC is its regulation of the cell cycle through *p21*, a key cyclin-dependent kinase inhibitor (CDKI). In NSCLC cells, *CNOT3* suppresses *p21* expression by promoting the degradation of its mRNA (*CDKN1A*). This suppression allows tumour cells to evade growth restraints, facilitating rapid proliferation. When *CNOT3* is depleted, this balance shifts dramatically (Shirai et al. 2019).

What's more, in leukaemia, regulatory functions of *CNOT3* are essential for maintaining the malignant phenotype. Elevated levels of *CNOT3* in leukaemia cells correlate with poor patient prognosis, and its depletion results in reduced proliferation, induction of differentiation, and apoptosis. A key finding from the study is the specific impact of *CNOT3* on the translation of *c-MYC*, a master regulator of cell proliferation. Despite having no effect on *c-MYC* mRNA levels, *CNOT3* depletion significantly reduces *c-MYC* protein levels, highlighting its role in promoting the translation of this and other oncogenic transcripts. This selective regulation further emphasises *CNOT3*'s importance in leukemogenesis (Ghashghaei et al. 2024).

1.17.2 Involvement of *CNOT3* in other disease

Research on *Cnot3*^{+/-} mice highlighted its role in metabolic control. These mice exhibited enhanced glucose tolerance, resistance to high-fat diets, and reduced body and tissue weight. This was linked to increased expression of lipid catabolism-related genes and decreased expression of lipogenesis-related genes. The stability of specific mRNAs involved in lipid and glucose metabolism was directly regulated by *CNOT3* (Morita et al. 2011).

One study showed that *Cnot3*^{+/-} mice exhibited cardiac contractile abnormalities and were prone to heart failure under stress. Reduced *CNOT3* levels were associated with changes in histone modifications (H3K4me3 and H3K9ac), impacting gene regulation.

Treatment with HDAC inhibitors restored cardiac function even under stress conditions. Additionally, a SNP in the *CNOT3* promoter has been linked to cardiac malfunctions in humans (Neely et al. 2010).

Furthermore, *CNOT3* has been identified as a critical modifier of *PRPF31* mutations, which cause autosomal dominant retinitis pigmentosa, a degenerative retinal disorder (Venturini et al. 2012). Not all carriers of *PRPF31* mutations develop symptoms due to incomplete penetrance, and *CNOT3* plays a key role in determining this variability. *CNOT3* negatively regulates *PRPF31* by binding directly to its promoter and repressing transcription. In asymptomatic carriers, lower levels of *CNOT3* allow higher expression of the functional *PRPF31* allele, compensating for the mutated allele. Conversely, higher *CNOT3* expression reduces *PRPF31* levels, increasing the risk of retinal degeneration. Experimental silencing of *CNOT3* in retinal cells confirmed its repressive effect, as *PRPF31* expression significantly increased when *CNOT3* was inhibited (Venturini et al. 2012).

Besides, *CNOT3* has protective roles in maintaining bone mass. In mouse models, the absence of *CNOT3* reduced bone volume and increased osteoclast numbers, promoting bone resorption while not affecting bone formation. This effect was linked to increased stability of the 3'UTR of *RANK* mRNA, a key factor in osteoclast differentiation. Importantly, the absence of *CNOT3* exacerbated osteoporosis in aged mice, further demonstrating its role in bone health (Watanabe et al. 2014).

Additionally, recent research has linked *CNOT3* mutations to neurodevelopmental disorders, reflecting its diverse roles in human biology (Martin et al. 2019). This discovery, spearheaded by the Deciphering Developmental Disorders (DDD) study, highlights the significant impact of de novo variants in *CNOT3* on developmental processes. These findings reveal a spectrum of clinical features ranging from intellectual disabilities to subtle physical traits, marking *CNOT3* as a key contributor to these complex conditions.

The clinical manifestations associated with *CNOT3* variants are diverse yet share some recurring features. All individuals studied exhibited developmental delays and learning difficulties, although the severity varied significantly. Speech development was particularly affected, with some children unable to speak, while others demonstrated limited vocabulary or the ability to form sentences. Behavioural challenges were prominent, including traits consistent with autism spectrum disorder (ASD), anxiety, and

aggression, affecting about half of the cases studied (Martin et al. 2019). Genetic analyses identified two main types of *CNOT3* variants: missense mutations and protein-truncating variants. Missense mutations generally resulted in milder phenotypes, as evidenced by some children attending mainstream schools with additional learning support. In contrast, truncating mutations, which lead to the loss of functional protein regions, were associated with more severe developmental impairments, including profound intellectual disabilities. The study hypothesised that these effects arise from haploinsufficiency, where the presence of only one functional copy of the gene fails to produce sufficient protein for normal development. Supporting this hypothesis, data from large genetic databases show that *CNOT3* is highly intolerant of loss-of-function mutations (Martin et al. 2019).

1.18 Involvement of *CNOT2* in diseases

CNOT2, an essential subunit of the Ccr4-Not complex, plays a critical role in regulating mRNA stability and transcription, both of which are vital for normal neurodevelopment. Haploinsufficiency of *CNOT2* has been identified as the cause of Intellectual Developmental Disorder with Nasal Speech, Dysmorphic Facies, and Variable Skeletal Anomalies (IDNADFS), a rare condition characterised by a consistent and recognisable spectrum of developmental and physical anomalies. This disorder also shares a significant overlap with chromosome 12q15 deletion syndrome, further emphasising the importance of *CNOT2* in brain development and cellular function (Uehara et al. 2019; Niceta et al. 2023).

Individuals with *CNOT2* haploinsufficiency present with developmental delays affecting both motor and speech milestones. Walking is often delayed, and many individuals exhibit nasal speech, which is a hallmark feature of the condition. Intellectual disability is universally observed, with severity ranging from mild to severe, significantly impacting daily functioning. In addition to these developmental challenges, affected individuals display distinctive craniofacial features, including a high, flat forehead, upslanted palpebral fissures, anteverted nares, a short triangular nose, thin upper lips, and low-set ears. Skeletal anomalies, such as scoliosis, vertebral defects, and brachydactyly, further contribute to the phenotypic variability of the disorder (Uehara et al. 2019).

At the molecular level, *CNOT2* mutations disrupt the function of the Ccr4-Not complex, impairing its role in mRNA regulation. Structural modelling has shown that missense

mutations, such as p. Asn466His, destabilise the Not-Box region of the CNOT2 protein, leading to loss of function. This loss of function impairs gene regulatory processes that are critical for brain development, explaining the wide-ranging neurodevelopmental deficits observed in patients (Uehara et al. 2019).

In addition, CNOT2 has emerged as a significant player in cancer progression due to its role in regulating transcription and mRNA stability. Recent research (Sohn et al. 2018) has revealed its overexpression in various malignancies, including breast, prostate, and pancreatic cancers, with particularly high levels observed in aggressive subtypes such as triple-negative breast cancer (TNBC). Elevated CNOT2 expression correlates with advanced tumour stages and poor prognosis, suggesting its critical role in driving cancer cell proliferation and survival.

In cancer, *CNOT2* exerts its oncogenic effects through multiple mechanisms. It directly activates the expression of vascular endothelial growth factor (VEGF), a key driver of angiogenesis, by binding to its promoter. Additionally, CNOT2 regulates the expression of proliferation-related genes such as *c-Myc*, *β-catenin*, and *cyclin D1*, which are pivotal for cell cycle progression. Experimental studies have shown that depletion of CNOT2 significantly impairs cancer cell proliferation, reduces colony-forming ability, and inhibits migratory capacity. These findings highlight its dual role in promoting both tumour growth and metastasis (Sohn et al. 2018). The angiogenic role of CNOT2 has been further validated through in vivo models. Knockdown of CNOT2 in mouse xenografts and chick embryo chorioallantoic membrane (CAM) assays led to a marked reduction in tumour vascularisation and size. Tumours with suppressed CNOT2 expression exhibit decreased levels of VEGF and PCNA, a marker of cellular proliferation. These results underscore the importance of CNOT2 in sustaining tumour vascular networks and its potential as a target for anti-angiogenic therapies (Sohn et al. 2018).

Additionally, Recent studies have identified *CNOT2* oncogenic potential, particularly in colorectal cancer (CRC), where it contributes to tumour growth, resistance, and apoptosis regulation (Jung et al. 2021). CNOT2 is frequently overexpressed in cancer cells, including CRC, and its high levels are associated with poorer survival rates. The oncogenic role of *CNOT2* is largely mediated through its interaction with key tumour suppressor genes such as p53. It has been shown that *CNOT2* suppresses *p53* expression, reducing its tumour-suppressive activities such as apoptosis induction and cell cycle

arrest. This mechanism highlights how *CNOT2* can drive tumour proliferation and evade programmed cell death mechanisms (Jung et al. 2021)

1.19 Aim and Objectives

CNOT3 and CNOT2 are integral components of the Ccr4-Not complex, a key regulator of gene expression and mRNA metabolism. Specifically, CNOT3 has been shown to be essential for cell proliferation in human cells, making it a potential factor in oncogenesis and therapeutic targets. Understanding which domains of CNOT3 are important for its role in cell proliferation in human cell lines is particularly interesting, especially since Suzuki et al. identified domains of CNOT3 for proliferation in mouse embryonic stem cells.

Notably, the N-terminal domain of CNOT3 has been identified as a hotspot for mutations implicated in T-cell acute lymphoblastic leukaemia (T-ALL) and neurodevelopmental disorders. This domain is also crucial due to its interaction with ribosomes, which is essential for its role in translation regulation.

Mutations in *CNOT2* have also been associated with neurodevelopmental disorders. These mutations can disrupt the normal functioning of *CNOT3* and *CNOT2*, leading to altered protein-protein interactions within the Ccr4-Not complex. Such disruptions can significantly affect gene regulation and translation, contributing to the pathology of these diseases. By delineating the functional domains of CNOT3 and assessing the impact of disease-associated mutations, this research aims to provide a deeper understanding of how CNOT3 and CNOT2 contribute to gene regulation and translation, ultimately informing potential therapeutic strategies.

To achieve the aim, the specific objectives are:

1. To generate stable cell lines for the conditional expression of siRNA-resistant Flag-CNOT3 variants. I will develop and validate a toolkit to investigate the function of CNOT3 in human cells. This will involve constructing plasmids to express wild-type and mutant variants of CNOT3, such as CNOT3 lacking N-terminal and C-terminal domains, and Not-Box regions and generating stable cell lines with these constructs under the control of an inducible promoter.
2. To characterise the functional roles of the N-terminal and C-terminal domains, and Not-Box regions of CNOT3 in driving the proliferation and interaction with other subunits of Ccr4-Not in human cell lines. This will be done using the established

stable cell lines containing CNOT3 expression cassettes. This will allow to pinpoint the structural regions of the CNOT3 that are required for cell proliferation.

3. To investigate the impact of missense mutations in CNOT3 and CNOT2 in neurodevelopmental disorders and T-ALL. This will be done by introducing missense mutations identified from literature on T-ALL and neurodevelopmental disorders into *CNOT3* and *CNOT2* using site-directed mutagenesis. The impact of these mutations on the interaction between CNOT3 and CNOT2 and their interactions with other subunits of the Ccr4-Not complex will be examined.

Together, these objectives will provide a more detailed insight into the molecular mechanism by which CNOT3 and CNOT2 impact on cell proliferation and gene regulation.

Chapter 2.

Materials and Methods

2.1 Bacterial growth and transformation

2.1.1 Reagents, stock solutions and buffers for bacterial methods

Table 2.1 Reagents and buffers are used in bacterial growth and transformation

Reagent/ Buffer	Composition
Luria-Bertoni (LB) media	10 g/l tryptone, 5 g/l yeast extract, 5 g/l NaCl, pH 7.2 (adjusted using NaOH)
LB agar	10 g/l tryptone, 5 g/l yeast extract, 5 g/l NaCl, pH 7.2 (adjusted using NaOH and 15 g/l bacterial agar
Ampicillin 1000X stock solution	100 mg/ml in 50% ethanol/water

2.1.2 Culture of *E. coli* DH5 α

The bacterial strain *E.coli* DH5 α (NEB, C2987H) was utilised for all DNA manipulations. To obtain isolated colonies, the bacteria were streaked onto agar plates containing the appropriate antibiotics and grown at a temperature of 37°C. Once colonies were formed, they were stored at a temperature of 4°C for up to several weeks. Liquid cultures were inoculated with a single colony and grown in universal screw-cap tubes containing LB medium with antibiotics, incubated at 37°C and 200rpm for 12-24 hours. For 200-300 ml cultures, glass conical flasks were used.

2.1.3 Transformation of Competent Cells

100 μ l of competent cells *E.coli* DH5 α cells were used for the transformation process. To achieve a good transformation yield when transforming plasmids, <500 ng of DNA was utilised. When transforming ligations or mutagenesis reactions, 10 μ l of ligation reaction mix was used.

Pre-chilled round-bottomed universal tubes (14 ml) were used to mix the competent *E. coli* cells and DNA, which were then incubated on ice for 10 minutes. The transformation

mix was then heat-shocked by incubation at 42°C for 30 seconds, followed by incubation on ice for 2 minutes. Afterwards, 1 ml of LB medium (pre-warmed at 37°C) was added to the cells, and the mixture was incubated at 37°C, 200rpm for 1 hour. For plasmid transformation, 100 µl of the cell suspension was spread onto LB agar plates containing the appropriate antibiotic.

2.2 Molecular biology

2.2.1 Regents, stock solutions and buffers for use in molecular biology methods

Table 2.2 Reagents and buffers used in molecular biology methods

Reagent/ Buffer	Catalogue number	Supplier
GenElute Plasmid Miniprep Kit	PLN350	Sigma-Aldrich
GenElute™ HP Plasmid Maxiprep Kit	NA0310-1KT	Sigma-Aldrich
The Monarch DNA Gel Extraction Kit	T1020S	NEB
<i>Afl</i> III-HF restriction enzyme	R0520S (20000 units/ml)	NEB
<i>Psp</i> OMI-HF restriction enzyme	R0653S (20000 units/ml)	NEB
<i>Not</i> I-HF restriction enzyme	R3189S (20000 units/ml)	NEB
Phusion High-Fidelity DNA polymerase	M0530S	NEB
Deoxynucleotide (dNTP) Solution Mix	N0447S	NEB
<i>Taq</i> DNA Polymerase with Standard <i>Taq</i> Buffer	M0273S	NEB
Nuclease-free Water	B1500S	NEB
<i>T4</i> DNA Ligase	M0202S	NEB

2.2.2 Small-scale plasmid DNA preparation

The GenElute Plasmid Miniprep Kit was used for the small-scale preparation of plasmid DNA. A single bacterial colony was picked and used to inoculate either 2 ml (for small scale) or 5 ml (for use in mammalian cell transient transfections) of LB containing the appropriate antibiotic ampicillin and grown overnight in an orbital shaker at 37°C, 200 rpm. All the overnight cultures were pelleted by centrifugation in a microfuge tube at 4000 g for 2 minutes. The plasmid DNA was then extracted according to the manufacturer's instructions. Small-scale plasmid extractions were eluted in 50 µl of elution buffer, or on 100 µl buffer when DNA was isolated from 5 ml. All plasmid DNA was assessed using NANODrop ONE C (Thermo Fisher Scientific) To check the concentration and purity before storing at -20°C.

2.2.3 Large-scale plasmid DNA preparation

The GenElute HP maxiprep Kit (Sigma-Aldrich,) was utilised to manufacture a large supply of high-quality plasmid DNA for transient transfections. Initially, a singular *E.coli* colony was used to inoculate 5 ml of LB (which had antibiotic). This mixture was cultured for 4-6 hours in an orbital shaker at 37°C, 200 rpm. This culture was then used to inoculate 200 ml of LB containing ampicillin. The culture was allowed to mature overnight in an orbital shaker at 37°C, 200rpm. The overnight culture was then re-suspended and processed by the manufacturer's standard protocol. The plasmid DNA was eluted in 500 µl of elution buffer and stored at -20 °C.

2.2.4 Determination of DNA/RNA concentration

To determine DNA concentration, the NanoDrop ND-1000 UV-Vis Spectrophotometer (NanoDrop Technologies) was utilised. To assess sample purity, A_{260}/A_{280} ratio values were calculated. A ratio between 1.8 to 2.0 was deemed acceptable for DNA samples, whereas RNA samples were expected to have a ratio between 2.0 to 2.2.

2.2.5 Agarose gel electrophoresis

A 1X TBE buffer was used to prepare the agarose gels. Next, the DNA samples were mixed with 1X loading dye and loaded onto the 1% gel. Depending on the sample, the gels were then run at 80V for 45 minutes to 1 hour. After the electrophoresis process, DNA was

visualized by transillumination with ultraviolet light. The results were then documented using Bio-Rad Gel Doc 2000 (Catalogue number 170-8100) and BioRad Quantity One software.

2.2.6 DNA extraction and purification from agarose gel

To extract DNA fragments from agarose gel, a surgical scalpel was used to cut out the DNA bands from the agarose gel. Monarch DNA Gel Extraction Kit instructions were followed to carry out the purification process. Initially, the gel slice was dissolved in a binding buffer at 65°C. To do this, 4 volumes of Monarch Gel Dissolving Buffer were added to the tube containing the gel slice (e.g. 400 µl buffer per 100 mg agarose). Samples were then processed through spin columns and briefly centrifuged with a benchtop microcentrifuge, with flow-through being discarded. The columns were washed twice with Wash Solution, and eventually, DNA was eluted in 50 µl of the pre-heated 65°C elution buffer provided.

2.2.7 Restriction enzyme digestion of DNA

Plasmid DNA underwent restriction endonuclease treatment (Danna and Nathans 1971) to produce compatible ends for cloning and to verify the generation of plasmids (Di Felice, Micheli, and Camilloni 2019). All required restriction endonucleases and buffers were obtained from New England Biolabs and employed following the manufacturer's guidelines. Specifically, the restriction enzymes *Afl*II-HF (Catalogue number R0520S), *Psp*OMI-HF (Catalogue number R0653S), and *Not*I-HF (Catalogue number R3189S) were employed. For analytical digestion, 100 ng of plasmid DNA was digested in a total volume of 10 µl using 1 µl of the appropriate restriction enzyme and 1 µl of 10X NEBuffer, with nuclease-free water added to the final volume. For preparative digests used in cloning, 1 µg of plasmid DNA was digested in a 50 µl reaction mixture, composed of 5 µl of 10X NEBuffer, 1 µl of the appropriate restriction enzyme (*Afl*II, *Psp*OMI, or *Not*I-HF), and nuclease-free water to a final volume of 50 µl. Reactions were incubated at 37°C for 1 hour. For double digests involving two restriction enzymes, NEBuffer CutSmart (Catalogue number B6004S) was used when both enzymes were compatible.

2.2.8 Ligation of DNA fragments

To generate recombinant plasmids, *T4* DNA ligase was utilised at a stock concentration 400,000 units/ml, to ligate DNA fragments. Ligation reactions involved combining

approximately 100 ng of a digested vector with a 3-fold molar excess of insert, along with 1 µl of *T4* DNA ligase in 1X *T4* DNA ligase buffer (10X) to achieve a final reaction volume of 10 µl. The reactions were incubated overnight at room temperature.

2.2.9 Polymerase Chain Reaction (PCR)

Polymerase Chain Reaction (PCR) was carried out using a thermal cycler to generate DNA fragments suitable for cloning and screening recombinant plasmids (Zhu et al. 2020). To synthesize DNA fragments for cloning, the Phusion DNA polymerase with 5'-3' proofreading activity was used. A standard reaction was set up according to **Table 2.3**. Samples were initially denatured at 98°C for 5 minutes followed by 30 cycles consisting of a denaturing step of 95°C for 30 seconds, annealing at 60- 68°C for 30 seconds, then elongation at 72°C for 1 minute per kilobase. A final elongation step for 10 minutes at 72°C was included. To ensure that the newly generated PCR fragments were of the correct size, 10 µl of the reaction was run on a 1% agarose gel.

Table 2.3 Standard Phusion Polymerase Reaction

Volume	Component/Reagent
4 µl	5X Phusion HF buffer
0.5 µl	10 mM dNTPs
1 µl	10 µM Forward Primer
1 µl	10 µM Reverse Primer
100 ng (1 µl)	Template DNA
0.2 µl	Phusion High Fidelity DNA Polymerase
Up to 20 µl	Nuclease-free water

2.2.10 Site-directed mutagenesis (Inverse PCR)

Site-directed mutagenesis involves the design of non-overlapping primers that contain desired mutations (Green and Sambrook 2019; Silva et al. 2017). These primers were oriented against each other as opposed to regular primers and the 5' ends were phosphorylated in advance by the manufacturer (Silva et al. 2017). SnapGene software (version 7.0.3, available at <https://www.snapgene.com>) and the NEB Primer Design Tool (available at <https://nebiocalculator.neb.com>) were used to design primers (**Table 2.4**).

To initiate the reactions, they were first incubated at a temperature of 98°C for 5 minutes. This was followed by 30 cycles, wherein each cycle included a denaturing step of 95°C for 30 seconds, annealing for an additional 30 seconds, and elongation at 72°C for a minute per kilobase. At the end of these cycles, a final elongation step lasting 10 minutes at 72°C was incorporated. The annealing temperature was adjusted within a range of 60-68°C.

To eliminate template DNA from site-directed mutagenesis reactions, the Dpn I (20,000 U/ml) enzyme was employed. 1 µl of *DpnI* was added to every 50 µl of mutagenesis reaction and left to incubate at 37°C for 1 hour. Following this, 10 µl of the reaction mixture was utilised to transform *E. coli* DH5α, after which DNA isolation (mini prep/Maxi prep) and sequencing were carried out to verify the occurrence of the desired mutations and unwanted mutations.

Table 2.4 The list of primers used in site-directed mutagenesis

Plasmid	Primers sequence 5'-3'
Flag.CNOT3R1	Forward: GGGATGAGTCAACATAGTATTCAGCGTCATCCTTGATCTTGCGGATGG Reverse: CCATCCGCAAGATCAAGGATGACGCTGAATACTATGTTGACTCATCCC
Flag.CNOT3R2	Forward: GCTGCCAGATACTGAGCCTTTGTGCCCTCCAGATA Reverse: TATCTGGAGGGCACAAGGCTCAGTATCTGGCAGC
Flag.CNOT3R2 ΔNTD	Forward: TCCCCCCCAGCCACAGC Reverse: TGGTACCTTGTCATCGTCATCTTTGTAATCCATGG
Flag- CNOT3R1ΔCTD	Forward: TAGCGGCCGCTCGAGCATG Reverse: GGGCACAGGGCCCAGTGG
Flag- CNOT3R1ΔNot- Box	Forward: TAGCGGCCGCTCGAGCATG Reverse: TGGGGGTGGCATCTGGTG

Flag- CNOT3.Arg57G ln	Forward: CAACA A CTGAGGGACCAAATCAAGACATGG Reverse: TAGCTTCTTAATCTCCTTCTTTAGGTCAGCC
Flag- CNOT3.ARG57 Trp	Forward: CAATGGCTGAGGGACCAAATCAAGACATGGG Reverse: TAGCTTCTTAATCTTAGGTCAGCCTAGCCTCATACTTTTC
Flag- CNOT3.Leu48V al	Forward: GTAAAGAAGGAGATTAAGAAGCTAVAACGGCTGA Reverse: GTCAGCCTCATACTTTTCTTTCTGGTTCGCGTTG
Flag- CNOT3.Glu147 Lys	Forward: TGA A GTCACTGTCAGTGCAGACACGCAAGA Reverse: CTTCACTCTCAA A CTGGTCCACCTGCATGTT
Flag- CNOT3.Arg188 His	Forward: CTGCACATGCTGGACAATGACTCCATCCTCG Reverse: GATGGTCTCTAGCATGCGCACGTGGTAG
Flag- CNOT3.Arg188 Cys	Forward: TATGCTGGACAATGACTCCATCCTCGTTGAC Reverse: CACAGGATGGTCTCTAGCATGCGCACG
Flag- CNOT3.Lys119 Glu T7 BGHR	Forward: AGAGGTTGGCCAGTGGCTCACGAATACCAT Reverse: TCCTTCTCTT C CTGGGCAGGATCTACCTT TAATACGACTCACTATTAGGG TAGAAGGCACAGTCGAGG
Flag- CNOT2.Asn466 His	Forward: GTGGAGCTTTTT C ACCGTGATTGGAGA Reverse: TGCAGCTAAAAGTTGTAATACGTC
Flag- CNOT2.Tyr494 Ter	Forward: AAAACCAATACCTAGGAGAGGGGAACA Reverse: CATTGTTGGCTCCATGCCTG

Flag-	Forward: ACAGATGGACCCTAATTCCTGGAGAT
CNOT2.Lys316	Reverse: ACTTGAAGTTGTCTTGCCAGATGTATTC
Ter	
<i>Afl</i> III	Forward: CACTATCTTAAGACCCAAGCTTCCACCATGGAT
Sp6	Forward: ATTTAGGTGACACTATAG

2.2.11 Construction of induced FLAG-CNOT3 expression vectors

To generate cell lines with inducible expression of CNOT3 variants, cDNA fragments encoding Flag-CNOT3, Flag-CNOT3R1, Flag-CNOT3R2, Flag-CNOT3 Δ NTD, Flag-CNOT3R1 Δ CTD, and Flag-CNOT3R1 Δ Not-Box were cloned into the pcDNA5-FRT-TetO expression vector (Braun et al. 2018). First, the pcDNA5-FRT-TetO vector (which originally contained an fSNAP tag) was digested with *Afl*III (20,000 units/ml) and *Psp*OMI (20,000 units/ml) to remove the fSNAP tag. Next, an *Afl*III restriction site was introduced at the 5' end of each Flag-CNOT3 cDNA by PCR amplification using high-fidelity Phusion polymerase. Following gel extraction and PCR amplification, the pcDNA5-FRT-TetO vector and all inserts were digested with *Afl*III and *Not*I restriction enzymes. The inserts were then ligated into the linearized pcDNA5-FRT-TetO vector. These constructs were chemically transformed into competent *E. coli* DH5 α cells following the manufacturer's protocol. The primers used for PCR amplification were designed using SnapGene software (version 7.0.3, available at <https://www.snapgene.com>) and the NEB Primer Design Tool (available at <https://nebiocalculator.neb.com>) (Table 2.4). Table 2.5 lists all plasmids constructed in this research using inverse PCR and restriction enzyme cloning procedures.

Table 2.5 List of plasmids constructed

Plasmid	Backbone	Insert/modification
pCMV-Flag-CNOT3R1	pCMV-Flag	Three-point mutations in CNOT3 cDNA
pCMV-Flag-CNOT3R2	pCMV-Flag	Two-point mutations in CNOT3 cDNA
pCMV-Flag-CNOT3 Δ NTD	pCMV-Flag	Removal of N terminal of CNOT3
pCMV-Flag-CNOT3R1 Δ CTD	pCMV-Flag	Removal of C terminal of CNOT3
pCMV-Flag-CNOT3R1 Δ Not-Box	pCMV-Flag	Removal of Not-Box
pcDNA-5FRT-TetO-fsNAP (empty plasmid)	CDM Lab	Removal of sNAP tag
Flp-In- recombinase	pCAG-FLpo	N/A
pCDNA5-FRT-TetO-Flag-CNOT3	pCDNA5-FRT-TetO	Insert Flag-CNOT3 cDNA
pCDNA5-FRT-TetO-Flag-CNOT3R1	pCDNA5-FRT-TetO	Insert Flag-CNOT3R1 cDNA
pCDNA5-FRT-TetO-Flag-CNOT3R2	pCDNA5-FRT-TetO	Insert Flag-CNOT3R2 cDNA
pCDNA5-FRT-TetO-Flag-CNOT3 Δ NTD	pCDNA5-FRT-TetO	Insert Flag-CNOT3R2 Δ NTD cDNA

pCDNA5-FRT-TetO-Flag- CNOT3.ΔCTD	pCDNA5-FRT-TetO	Insert CNOT3R1ΔCTD cDNA	Flag-
pCDNA5-FRT-TetO-Flag- CNOT3.ΔNot-Box	pCDNA5-FRT-TetO	Insert CNOT3R1ΔNot-Box cDNA	Flag-

2.2.12 Construction of eukaryotic expression plasmids containing disease-associated mutations in *CNOT3* and *CNOT2*

Plasmids containing mutations in *CNOT3/CNOT2* (Martin et al. 2019; De Keersmaecker et al. 2013; Niceta et al. 2023; Uehara et al. 2019) that cause T-ALL and neurodevelopmental disorders were constructed using inverse PCR. The primers, listed in **Table 2.4**, were designed using SnapGene software (version 7.0.3, available at <https://www.snapgene.com>). The method described in section 2.2.10 was followed to construct these plasmids, and the plasmids constructed are listed in **Table 2.5**.

2.2.13 Colony PCR

To screen the newly generated plasmids for the presence of the correct insert, single clones from the plate were selected and inoculated into 1 ml of LB with the appropriate antibiotic (ampicillin). This screening step is critical to confirm that the bacterial colonies contain plasmids with the desired insert before further analysis. A standard Taq Polymerase PCR (NEB, Catalogue number M0273S) reaction was set up according to **Table 2.6**. The samples were initially denatured at 95°C for 5 minutes, followed by 30 cycles of denaturation at 95°C for 30 seconds, annealing at 60°C to 68°C for 30 seconds, and elongation at 72°C for 1 minute per kilobase. A final elongation step for 10 minutes at 72°C was included. To verify the size of the insert, 10 µl of the reaction was loaded onto a 1% agarose gel (Bergkessel and Guthrie 2013).

Table 2.6 Standard Taq DNA polymerase PCR reaction

Volume	Component/Reagent
2.5 µl	10X Standard Taq Reaction Buffer
0.5 µl	10 mM dNTPs
0.5 µl	10 µM Forward Primer
0.5 µl	10 µM Reverse Primer
2 µl	Template DNA
0.125 µl	Taq DNA Polymerase (5000 units/m)
Up to 25 µl	Nuclease free water

2.2.14 DNA sequencing

Plasmids isolated via miniprep or maxiprep were sent to Source Bioscience for Sanger sequencing alongside their appropriate sequencing primer (**Table 2.7**). 5 µl of plasmid at a concentration of 100 µg/ml and 5 µl of sequencing primer at 3.2 ng/ml.

Table 2.7 List of sequencing primers

Primer	Sequence
BGH Reverse	TAG AAG GCA CAG TCG AGG
CMVF_pCDNA3	CAA CGG GAC TTT CCA AAA TG
T7 promoter (F)	TAA TAC GAC TCA CTA TAG GG
pCMV forward	GAG CTC GTT TAG TGA ACC GTC

2.3 Cell Culture

HEK293 and Flp-In TREx-HEK293 cells (Ward, Alvarez-Curto, and Milligan 2011) were grown in Dulbecco's modified Eagle media (DMEM) with 10% fetal bovine serum (FBS) in a humidified incubator at 37 °C with 5% CO₂. To maintain aseptic conditions a Class II biological safety cabinet was used. Cells were kept in T75 flasks for routine culture. Once the cells reached 80% confluence, they were washed two times with PBS and transferred to a new flask by trypsinisation for two minutes. After spinning down the cells at 200 g for 5 minutes at room temperature in 15 ml falcon tubes, the medium was removed, and the cells were resuspended in fresh media before being added to a new flask. The cells were typically split 1:10 three times a week to maintain their growth.

Cells were thawed from liquid nitrogen storage when needed by quickly warming the cryovials in a 37°C water bath, followed by resuspension in warm DMEM medium containing 10% FBS. For long-term storage, cells were frozen in freezing medium (10% DMSO in FBS) and stored in cryovials at -80°C before being transferred to liquid nitrogen. Experiments were conducted on cells between passages 6 and 25.

2.3.1 Reagents, stock solutions and buffers for use in cell culture

Table 2.8 List of reagents and buffers used in cell culture

Product name	Catalogue number	Supplier
Fetal Bovine Serum	A5256701	ThermoFisher Scientific
DMEM complete	D5796	Sigma-Aldrich
Trypan Blue solution	T6146-5G	Sigma-Aldrich
AlamarBlue cell viability	DAL1025 (25 ml)	ThermoFisher Scientific
Crystal Violet cell Solution	V5265-250ML	Sigma-Aldrich
Opti-MEM I Reduced Serum Medium	319850063	ThermoFisher Scientific
Fetal Bovine Serum, Tet system approved, USDA-approved regions	A4736401	ThermoFisher Scientific
Hygromycin B Gold (solution)	Ant-hg-1 (100 mg/ml)	Invitrogen
INTERFERin, siRNA transfection reagent	101000036	Polyplus-transfection
FuGENE6 Transfection Reagent	E2691	Promega
DMSO	34869-100ML	Sigma-Aldrich
Mr. Frosty cryo-container	1562-1EA	NALGENE Labware

2.3.2 Freezing cells

To facilitate storage, the cells were cryopreserved in liquid nitrogen. Before cryopreservation, the cells were grown in 10 cm dishes to approximately 80% confluency for 24 hours. The cells from each dish were then harvested, with each dish providing enough cells for one vial. After trypsinisation, the cells were pelleted by centrifugation at 200 g for 5 minutes at 4°C in 15 ml falcon tubes for HEK293 and Flp-In TREx HEK293. The supernatant was then removed, and the cells were resuspended in a freezing mixture consisting of 1 ml FBS with 10% DMSO. Finally, the cells were stored at -80°C in a Nalgene Cryo 1°C Freezing Container (Mr frosty) to ensure cooling at a rate of -1°C per minute before transfer to liquid nitrogen.

2.3.3 Thawing cells

Upon removal from the liquid nitrogen, the ampoule was promptly thawed in a 37°C water bath. Next, the cell aliquot was introduced into a DMEM medium containing 10% FBS. After mixing, the cells were centrifuged at 200 g for 5 minutes at room temperature and the supernatant was discarded. The cells were then resuspended in a new 10% FBS DMEM solution for routine upkeep in a T75 flask, as previously outlined. The cells were given a week to recuperate before any experimentation occurred.

2.3.4 DNA plasmids transient transfection

The expression plasmids were constructed and transiently transfected to observe the expression of CNOT3 variants in the HEK293 cell line. The transfection was performed using FuGENE6 (Nagy and Watzele 2006). In a 6-well plate, cells were seeded and grown to 80% confluency to prepare for transfection. The plasmids were transfected in a 1:3 ratio using 2 µg of DNA plasmid and 6 µl of the transfection reagent (**Table 2.9**). The transfection mixture was prepared by combining the appropriate amount of Opti-MEM with 6 µl of the transfection reagent and leaving it at room temperature for 5 minutes. Then, 2 µg of plasmid DNA was added to the mixture and left at room temperature for 15 minutes. Finally, the mixture was added dropwise onto the cells. The next day, the media was removed and replaced with warm media. Cells were then incubated in an incubator for 12 hours before harvesting.

Table 2.9 Transfection steps for transient transfection

Component/Reagent	Volume
2 µg of each DNA plasmids	X µl (depends on plasmids concentration)
Transfection reagent (FuGENE 6)	6 µl
Opti-MEM	Up to 100 µl

2.3.5. Generation of stable cell lines

To transfect the Flp-In TREx HEK293 cell line, two plasmids were used. The pcDNA5-FRT-TetO expression vector carried the *CNOT3* gene under the control of a tetracycline-inducible promoter, enabling controlled expression. The second plasmid encoded pCAG-Flpo, a Flp recombinase enzyme, (Xue, Atallah, and Scanziani 2014), which mediates site-specific recombination, ensuring stable integration of the expression vector into the cell's genome. This integration allows for consistent and reliable expression of the gene of interest.

Flp-in TREx HEK293 cells were seeded onto 6 cm plates and grown to approximately 80% confluency before transfection. The plasmids were transfected in a 1:9 ratio with one µg of the pcDNA5-FRT-TetO-CNOT3 expression vector and nine µg of the Flp-in recombinase vector, resulting in a final concentration of 10 µg/ml plasmids. To prepare the transfection mixture, 470 µl of Opti-MEM was combined with 20 µl of FuGENE and 10 µl of plasmids per well. The mixture was then mixed and incubated for 5 minutes at room temperature. Prior to transfection, the cells were washed twice with PBS, and 2.5 ml of 10% FBS DMEM media was added to each plate. The transfection mixture was then added dropwise onto the cells. After 24 hours, the media containing the transfection mixture was removed and replaced with fresh 10% FBS DMEM media.

Hygromycin (250 µg/ml) was added to the cells 48 hours after initial transfection to select successfully transfected cells.

2.3.6 Determination of Hygromycin Kill Curve

To determine the most effective concentration of Hygromycin B for selecting successfully transfected Flp-In TREx HEK293 cells, a kill curve experiment was conducted. This involved seeding the cells in a 6-well plate and allowing them to grow to around 50% confluency. Varying concentrations of Hygromycin B (1000 µg/ml, 750 µg/ml, 500 µg/ml, 250 µg/ml, 100 µg/ml, and a control with 0 µg/ml) were then added to the wells. The cells were incubated for ten days, with the Hygromycin-containing media replaced every 2-3 days.

Throughout the incubation period, cell viability was closely monitored daily using microscopic observation for cell death. The goal was to identify the lowest concentration of Hygromycin B that resulted in complete cell death within seven days. This concentration was deemed optimal for selecting only those cells that had successfully integrated the transfected DNA, as they would be the only ones able to survive in the presence of the antibiotic.

The results of the kill curve showed that 250 µg/ml of Hygromycin B effectively eliminated all non-transfected cells within the seven-day timeframe. This concentration was then used in subsequent experiments to select for the successfully transfected cells, ensuring a pure population for further analysis.

Successfully transfected cells were then cultured in a medium containing 250 µg/ml Hygromycin, and six independent mass populations were generated and maintained in parallel.

2.3.7 Doxycycline-Induced Expression of CNOT3 Variants

To induce the expression of CNOT3 variants, inducible TetO cassette cell lines were seeded in 6-well plates. After 24 hours of incubation at 37°C with 5% CO₂ to reach approximately 80% confluency, expression was induced by adding 2 µg/ml doxycycline hyclate (Sigma-Aldrich, Catalogue number D9891-1G) directly to the culture media. This concentration was selected based on the manufacturer's recommendations for optimal induction of Tet-On systems. The plates were gently swirled to ensure even distribution of doxycycline and prevent localised high concentrations. Cells were treated for 48 hours, with individual wells harvested. All remaining wells were harvested simultaneously at the end of the experiment for further analysis (see Section 2.3.6).

2.3.8 siRNA Transfection of HEK293 cells

For a 6-well plate, HEK293 cells were transfected with a final concentration of 30 nM siRNA, INTERFERin, and Opti-MEM medium. The siRNAs used were siGENOME duplex #1 and #2 targeting the N terminal and C terminal of CNOT3 respectively, along with a non-targeting negative control (**Table 2.10**). One knockdown was performed for both siRNAs.

Cells were seeded 24 hours before transfection and grown to a confluency of approximately 30%. The transfection was done using INTERFERin with one hit of 30 nM siRNA. The transfection mixture was assembled based on the culture vessel, as shown in **Table 2.11**.

Table 2.10 siRNAs used in this research

Reagent	Sequence	Specificity	Supplier	Concentration
siGENOME duplex, CNOT3 #1	Sense: GGACGACGUUGAGUACUAAUUU Antisense: AUAGUACUCAACGUCGUCCUU	Human	Horizon Discovery	30 nM
siGENOME duplex, CNOT3 #2	Sense: GCACUAAGGCACAGUAUCUUU Antisense: AGAUACUGUGCCUUAGUGCUU	Human	Horizon Discovery	30 nM
sicontrol non- targeting siRNA #1	UAGCGACUAAACACAUCAA	Human	Horizon Discovery	30 nM

Table 2.11 Transfection conditions in different culture vessels

Culture vessel	Volume of INTERFERin	Volume of medium w/o serum for complex formation	Volume of complete medium on cells	Final volume
24-well	4 μ l	100 μ l	500 μ l	600 μ l
12-well	8 μ l	200 μ l	1 ml	1.2 ml
6-well	12 μ l	200 μ l	2 ml	2.2 ml

To prepare for transfection on a 6-well plate, 200 μ l of Opti-MEM was used to dilute 30 nM of siRNA duplexes. The solution was mixed by pipetting it up and down. Next, 12 μ l of INTERFERin was added, and it was mixed by vortexing for 10 seconds. The mixture was incubated for 10 minutes at room temperature to allow the formation of transfection complexes between siRNA duplexes and the transfection reagent.

During the formation of the complex, the growth medium was removed, and a fresh DMEM medium containing 10% FBS was added. After that, 200 μ l of the mixture was added to the cells. The plate was gently swirled and incubated at 37°C.

After 24 hours of transfection, the medium was replaced with a fresh one. Gene silencing was measured 72 hours after knockdown.

2.3.9 Crystal violet assay

The crystal violet assay involves staining cells adhered to the cell culture plates (Feoktistova, Geserick, and Leverkus 2016). This assay assumes that all the cells attached to the plate's bottom are viable, and the intensity of staining is proportional to the number of cells present.

After 72 hours following the knockdown, the cells underwent treatment with crystal violet to evaluate the impact of the gene of interest knockdown on the proliferation rate of the HEK293 cell line. The first step was to remove all the medium from the cells, which were washed twice with PBS. Next, the cells were fixed with 10% solution HCl for at least 10 minutes at room temperature. Then, 500 μ l of 0.05% crystal violet staining solution

was pipetted onto the cells and left to incubate at room temperature for 15 minutes. Once the incubation was complete, the crystal violet solution was discarded, and the cells were washed with deionised water until no colour was released from the plate. Once this step was done, the plate was left in the chemical fumes hood in a tilted position to dry. After the plate was completely dry, photos were photographed after 7 days.

2.3.10 AlamarBlue Assay

Cells in a six-well plate (from section 2.3.6) were exposed to AlamarBlue solution (O'Brien et al. 2003). To carry out the assay, 10% of the culture volume in each well was replaced with AlamarBlue. For example, in a well containing 1 ml of culture, 100 μ l of AlamarBlue was introduced, and the cells were incubated for 2 hours at 37°C in the incubator. Subsequently, the fluorescence was measured using a plate reader (BioTek Synergy H1, 208985) at an excitation wavelength of 530-560 nm and an emission wavelength of 590 nm.

The AlamarBlue assay was also used to measure the effect of doxycycline induction on the proliferation rate of stable cell lines. The stably induced cell lines were initially cultured in a 24-well and allowed to reach 70% confluency. The following day, one set of cells was induced with doxycycline (2 μ g/ml) while the other was not induced (no treatment of Dox) and was incubated under standard conditions at 37°C. After 48 hours, the AlamarBlue assay was conducted, as previously described.

2.4 Protein analysis

2.4.1 Reagents, stock solutions and buffers for use in Protein Analysis

Table 2.12 Antibodies used for western blotting

Antibody	Dilution		Supplier
Anti-CNOT3	1:1000		Abnova (4B8)
Anti-CNOT1	1:1000		Cell signalling (#44613)
Anti-CNOT2	1:1000		Cell signalling (#34214)
Anti-CNOT7	1:1000		Cell signalling (#86665)
Anti-S3	1:1000		Cell signalling (#9538)
Anti- β tubulin	1:1000		Abcam (ab15568)
Anti-Flag M2	1:1000		Sigma (F1805)
Anti-mouse (secondary)	HRP	1:2000	Proteintech (SA00001-1)
Anti-Rabbit (secondary)	HRP	1:2000	Invitrogen

Table 2.13 Buffers used in western blotting

Buffer/ Reagent	Composition	Application
RIPA buffer	150 mM NaCl, 1.0% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS, 50 mM Tris, 1 mM EDTA, pH 8.0	WB
4X Upper buffer	0.5 M Tris base, 0.4% SDS, pH 6.8; stored at room temp	WB
4X Lower buffer	1.5 M tris base, 0.4% SDS, pH 8.8; stored at room temp	WB
10X Running buffer	0.25 M Tris base, 1.0% SDS, 1.92 M glycine; stored at room temp	WB
10X Transfer buffer	0.25 M Tris base, 1.92 M glycine; stored at room temp	WB
1X TBS-T	50 mM Tris-HCL pH 7.8, 150 mM NaCl, 0.1% Tween-20. Stored at room temp	WB
Blocking buffer	N/A	WB
4X SDS loading dye	2.4 ml 1 M Tris buffer pH 6.8, 4 ml 100% glycerol, 0.5ml β -mercaptoethanol, 0.8 g SDS, 4 mg bromophenol blue and 3.1 ml distilled water. Stored at -20°C	WB
10% APS	2.0 g ammonium persulphate in 20 ml water (WB ,Thermo Fisher Scientific, Catalogue number 17875)

2.4.2 Preparation of protein lysates for western blotting

The cells bearing HEK293 were transfected with the appropriate expression vectors using FuGENE6 (Nagy and Watzele 2006). After 48 hours post-transfection, the cells

were harvested for further analysis. Before harvesting, the cells were washed twice with PBS and replaced with a fresh warm medium. After 48 hours post-transfection, the cells were harvested by pipetting the culture medium up and down to dislodge them. The cell suspensions were then transferred into 15 ml falcon tubes and centrifuged at 4000 g, 4°C for 5 minutes, and the media was discarded. The cell pellets were washed twice with cold PBS and the centrifugation step was repeated. Following this, the cell pellets were re-suspended in 500 µl of cold cell lysis buffer (RIPA buffer) and transferred into 1.5 ml tubes. Cells were lysed twice by freeze/thawing at -80°C, followed by centrifugation at 16,000 g, 4°C for 10 minutes. The supernatant was transferred into new 1.5 ml tubes and either stored at -80°C or used directly for the assay.

2.4.3 Bradford assay to determine protein concentration

The concentration of protein lysates was measured using the Bradford assay (Kielkopf, Bauer, and Urbatsch 2020), which utilised Bradford reagent (Sigma Aldrich, Catalogue number B6916-500ML) and Bovine serum albumin (BSA) with a concentration of 20 mg/ml was diluted to 10 µg/µl with lysis buffer for protein isolation. Standard curves were established with varying concentrations of BSA, ranging from 0, 2, 4, 6, 8 to 10 µg/µl. To conduct the assay, 10 µl of protein sample was introduced to 200 µl of Bradford reagent, and the total volume was adjusted to 1 ml with water. The mixture was then vortexed and left to incubate for 5 minutes at room temperature. Once incubation was completed, the samples were moved to 1 ml cuvettes, and the absorbance was measured at 600 nm. The protein concentration of the samples was calculated by evaluating the standard curve obtained via linear regression analysis using Microsoft Excel.

2.4.4 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

The Invitrogen X-Cell SureLock Mini Cell system (ThermoFisher Scientific) was used to conduct SDS-PAGE analysis (Nowakowski, Wobig, and Petering 2014). The separation gels were prepared at different concentrations, namely, 10% and 14%, based on the molecular mass of the protein of interest. The separation gel was initially cast in a gel cassette while leaving a room of 1.5 cm at the top for the stacking gel to be added. After pouring the separation gel onto the cassette, 500 µl of isopropanol was promptly added to the top of the gel, following which the gel was allowed to set for 15-20 minutes. Once the gel was formed, the isopropanol was removed by inverting the gel and draining it onto tissue paper. Subsequently, the stacking gel was cast on top, and the appropriate

Chapter 2. Materials and Methods

comb was inserted immediately after pouring. The gel was allowed to stand at room temperature for 30 minutes before being used or stored at 4°C. The comb was removed before use, and the gel was washed with 1X running buffer. Before loading, protein extracts were denatured by boiling for 5 minutes in 1X SDS sample buffer. The protein ladder employed was the NEB colour protein standard, broad range 11-2456 kDa (Catalogue number P7719S). Gels were run in 800 ml of 1X running buffer at 180V for approximately 1.5 hour.

Table 2.14 List of the SDS-PAGE gels component

Reagent	Resolving Gel		Stacking gel
	10%	14%	4%
Protein MW range, kDa	30-200	10-80	-
4X Lower buffer	2 ml	2 ml	-
4X Upper buffer	-	-	750 µl
40% acrylamide:	2 ml	2.8 ml	300 µl
Bisacrylamide 29:1			
10% APS	80 µl	80 µl	60 µl
TEMED	8 µl	8 µl	6 µl
H ₂ O	4 ml	3.2 ml	1950 µl

2.4.5 Western blotting and immunodetection

After separation by SDS-PAGE, proteins were transferred to nitrocellulose membranes (Whatman Protran 0.45 μ m) via the Invitrogen X-Cell SureLock Mini Cell system for 1 hour at 25V in a transfer buffer. The membranes were then blocked with a solution of 5% dried milk dissolved in 1X TBST for 1 hour at room temperature. Following this, the membranes were transferred to 50 ml Falcon tubes containing 1X TBST, the specific primary antibody, and 5% dried milk, and incubated on a rotator overnight at 4°C. Before proceeding to incubate the membranes with the secondary antibody, they were washed three times with 1X TBST buffer for 5 minutes each time. The membranes were further incubated with the specific secondary antibody in 1X TBST and 5% dried milk for an hour on a rotator at room temperature. Finally, the membranes were incubated with 600 μ l of ECL western blotting substrate (ThermoFisher Scientific, Catalogue number 32106) for 1 minute before imaging. The Fujifilm LAS-4000 system and image reader computer software (Fujifilm, Japan) were used for imaging and image analysis was carried out using Image J. **Table 2.12** lists all the antibodies used.

2.5 Co-Immunoprecipitation

2.5.1 Regents, stock solutions and buffers for use in Co-Immunoprecipitation

Table 2.15 Regents and buffers used in Co-Immunoprecipitation

Reagent/Buffers	Catalogue number	Supplier
Bradford Reagent	B6916-500ML	Sigma-Aldrich
DynaMag-2	12321D	Thermo Fisher Scientific
Anti-Flag M2 magnetic beads	M8823-1ML	Sigma-Aldrich
Protease inhibitor cocktail	11836170001	Sigma-Aldrich

0.2 NP-40 buffer (50 mM Tris-HCL pH 8.0, 150 mM NaCl, 5 mM MgCl ₂ , EDTA 0.5 mM, 0.2% NP-40, 5% Glycerol, 1 mM DTT)	N/A	N/A
2X Laemmli buffer (0.125 M Tris base, 0.14 M SDS, glycerol 20%, 2-mercaptoethanol 10%, Bromophenol blue)	N/A	N/A

2.5.2 Preparation of protein lysate for co-immunoprecipitation

HEK293 cells were seeded and grown at 80% confluency in 6 cm dishes. After 24 hours, cells were transfected with plasmids encoding (Flag-CNOT3Arg57Gln, Flag-CNOT3ARG57Trp, Flag-CNOT3Leu48Val, Flag-CNOT3Glu147Lys, Flag-CNOT3Arg188His, Flag-CNOT3Arg188Cys, Flag-CNOT3.Lys119Glu, Flag-CNOT2Asn466His, Flag-CNOT2Tyr494Ter and Flag-CNOT2Lys316Ter) using FuGENE6. In each culture dish, a total of 5 µg of plasmid DNA was transfected as shown in **Table 2.16**. After 48 hours post-transfection, the cells were put on an ice-cold tray and harvested by pipetting the culture medium up and down to dislodge them. The cell suspensions were then transferred into 15 ml Flacon tubes and centrifuged at 4000 g, 4°C for 5 minutes, and the media was discarded. The cell pellets were washed twice with cold PBS and the centrifugation step was repeated. Following this, the cell pellets were re-suspended in 500 µl - 1 ml of cold cell lysis buffer (0.2 NP-40) and transferred into 1.5 ml tubes. The cells were lysed twice by freeze/thawing at -80°C, followed by centrifugation at 16,000 g, 4°C for 10 minutes. The supernatant was transferred into new 1.5 ml tubes and either stored at -80°C or used directly for the assay. Following that, the cell lysate protein concentration was measured using Bradford assay.

Table 2.16 Transfection steps used for transient transfect

Component/ Reagent	Volume
5 µg of DNA plasmids	X µl
Transfection reagent (FuGENE 6)	30 µl
Opti-MEM	Up to 600 µl

2.5.3 Co-Immunoprecipitation

For co-immunoprecipitation (Lin and Lai 2017) using anti-Flag M2 magnetic beads, the protocol was scaled according to the number of samples. Firstly, 10 µl of anti-Flag magnetic beads were mixed in a 1.5 ml tube, then the beads were washed three times with 500 µl of 1X TBS (50 mM Tris HCl, 150 mM NaCl, pH 7.4) to prepare the beads. The tube containing beads and TBS was thoroughly mixed and then placed in the magnetic rack to collect the beads. The storage buffer and TBS mixture were then discarded. Afterwards, the equilibrated beads and protein extract were incubated for 1 hour at room temperature, with gentle mixing using the rotating device to capture the Flag proteins. Once the incubation was complete, the tube was placed in the magnetic stand to collect the beads and remove the supernatant. Next, the beads were washed three times with 500 µl of 1X TBS to remove non-specifically bound proteins. Finally, the beads were eluted with 2X Laemmli buffer for 10 minutes at 95°C. The tube was then placed in the magnetic beads to collect the supernatant, which was used for further analysis using the SDS-PAGE gel.

2.6 Computational work

2.6.1 Prediction of Amino Acid Substitution Effects Using PolyPhen-2

The PolyPhen-2 tool (Adzhubei, Jordan, and Sunyaev 2013) was utilised to predict the impact of missense mutations. The tool was accessed online via <http://genetics.bwh.harvard.edu/pph2/>. The analysis was performed using the HumVar model, which is designed to predict the potential effects of both rare Mendelian variants and common polymorphisms.

Step-by-Step Procedure

1. **Protein Sequence Retrieval:** Protein sequences for the relevant proteins, including CNOT3 and CNOT2, were obtained from the UniProt database. The UniProt accession numbers of these proteins were used to ensure accurate and consistent input data.
2. **Input of Amino Acid Substitutions:** The specific amino acid substitutions of interest were identified based on experimental data. For each mutation, the following details were entered into the PolyPhen-2 interface:
 - Position of the substitution within the protein sequence.
 - Wild-type amino acid (the original amino acid in the protein).
 - Mutant amino acid (the amino acid replacing the original one).
3. **Submission of Query:** After inputting the required data, the "Submit Query" button was clicked to start the analysis. In cases where multiple substitutions were analysed, the batch processing feature was used to input multiple variants in a single query, by preparing a file in the required format.
4. **Data Processing and Prediction:** PolyPhen-2 then processed the submitted data using various features, including sequence-based analysis, evolutionary conservation, and protein structure. For each substitution, PolyPhen-2 generated a prediction score ranging from 0 to 1, indicating the likelihood of the mutation being damaging to the protein's function. The output was provided as a qualitative prediction:
 - Benign (no predicted impact on protein function),
 - Possibly damaging (some potential for functional impact),
 - Probably damaging (high confidence of a deleterious effect).
5. **Interpretation of Results:** The scores and classification for each substitution were reviewed. Scores closer to 1 were interpreted as indicating a higher probability of a detrimental effect on the protein.
6. **Download and Storage of Results:** After analysis, the results were downloaded from the website in a tabular format for further examination and incorporation into the study.



PolyPhen-2 prediction of functional effects of human nsSNPs

Home About Help Downloads Batch query WRESS.de

PolyPhen-2 (Polymorphism Phenotyping v2) is a tool which predicts possible impact of an amino acid substitution on the structure and function of a human protein using straightforward physical and comparative considerations. Please, use the form below to submit your query.

21-Jun-2021: Server has been migrated to new hardware. Note, all queries were terminated and user sessions data discarded in the process, hence you will need to resubmit your query if affected. We apologize for the inconvenience caused.

Query Data

Protein or SNP identifier

Protein sequence in FASTA format

Position

Substitution

AA₁ A R N D C E Q G H I L K M F P S T W Y V

AA₂ A R N D C E Q G H I L K M F P S T W Y V

Query description

[Display advanced query options](#)

Figure 2.1: PolyPhen-2 Web Interface for Predicting the Functional Impact of Amino Acid Substitutions.

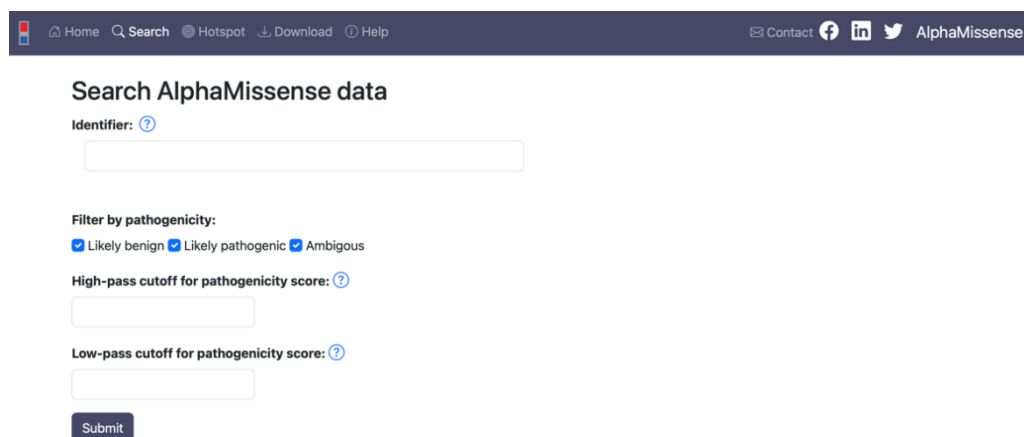
This figure shows the web interface of PolyPhen-2 (Polymorphism Phenotyping v2). Input Protein name and Protein sequence in FASTA format were entered in protein identifier and protein sequence respectively. The specific substitution is entered by indicating the position of the mutation and the respective wild-type and mutant amino acids.

2.6.2 Prediction of missense variant pathogenicity using AlphaMissense

To predict the functional consequences of missense mutations, AlphaMissense was employed (Ljungdahl et al. 2023; Cheng et al. 2023). AlphaMissense utilises deep learning algorithms to assess the likelihood of a missense variant being pathogenic or benign. The analysis was conducted via the AlphaMissense online platform, which can be accessed at <https://alphamissense.hegelab.org/results>.

Step-by-Step Procedure.

1. Gene Input: To initiate the analysis, the gene name was directly entered into the AlphaMissense query field. In this case, the genes *CNOT3* and *CNOT2* were analysed. Upon entering the gene name, the "Submit Query" button was pressed, which redirected the tool to the result page.
2. Analysis of Missense Mutations: On the result page, AlphaMissense automatically generated a comprehensive list of missense mutations for each amino acid position in the protein encoded by the entered gene. For each mutation, AlphaMissense provided a pathogenicity score (ranging from 0 to 1), where a higher score indicated a greater likelihood of the mutation being pathogenic. Additionally, each variant was classified into one of two categories: "Pathogenic" or "Benign", based on the calculated score.
3. Pathogenicity Data Collection: The list of missense mutations, along with their respective pathogenicity scores and classifications, was downloaded from the results page for further analysis. These predictions were then used to evaluate the potential functional impact of specific mutations within the *CNOT3* and *CNOT2* proteins.



The screenshot shows the AlphaMissense web interface. At the top is a dark navigation bar with links for Home, Search, Hotspot, Download, and Help on the left, and Contact, social media icons (Facebook, LinkedIn, Twitter), and the AlphaMissense logo on the right. Below the navigation bar, the main heading is "Search AlphaMissense data". Under this heading is an "Identifier:" label with a help icon and a text input field. Below the input field is a "Filter by pathogenicity:" section with three radio buttons: "Likely benign" (selected), "Likely pathogenic", and "Ambiguous". Below the filters are two more input fields: "High-pass cutoff for pathogenicity score:" and "Low-pass cutoff for pathogenicity score:", each with a help icon. At the bottom of the form is a "Submit" button.

Figure 2.2 AlphaMissense Web Interface for Missense Variant Pathogenicity Search.

This figure displays the AlphaMissense web interface used for predicting the pathogenicity of missense variants. The Identifier field allows for the input of the gene name to retrieve missense variant data. Filters are available to categorise the predicted variants as likely benign, likely pathogenic, or ambiguous. Additional options include setting high-pass and low-pass cutoff values for the pathogenicity score. After inputting the required data and adjusting the filters, results are generated by pressing the Submit button, providing insights into the pathogenic potential of the variants based on the prediction model.

2.6.3 Prediction of Amino Acid Substitution Effects Using SIFT

The analysis was conducted via the SIFT (Kumar, Henikoff, and Ng 2009) online platform, which can be accessed at https://sift.bii.a-star.edu.sg/www/SIFT_seq_submit2.html.

Step-by-Step Procedure

1. **Protein Sequence Input:** To begin the analysis, the protein sequence of the gene of interest (such as *CNOT3*) was prepared in FASTA format. The sequence was either pasted directly into the designated text box on the SIFT interface or uploaded as a FASTA file.
2. **Mutation Data Input:** After entering the protein sequence, the specific substitutions of interest (missense mutations) were entered in the provided text box in the required format (i.e., indicating the position of the mutation and the amino acids involved, such as 188 A > H). Alternatively, a file containing the list of substitutions was uploaded, depending on the scale of the analysis.
3. **Submission of Query:** Once the protein sequence and the desired substitutions were input, the "Submit" button was clicked to begin the analysis. SIFT then performed a homology search against a database of related protein sequences to assess how likely the substitution is to be tolerated or damaging to the protein's function.

4. Data Processing and Result Generation: SIFT generated a score for each substitution, ranging from 0 to 1. Scores closer to 0 indicated a higher likelihood that the substitution would be damaging, while scores closer to 1 suggested that the substitution was likely tolerated. Each substitution was also categorised as “damaging” or “tolerated” based on the score. Results were typically available within 10-15 minutes after submission.

SIFT Sequence

SIFT Sequence provides SIFT predictions for a given protein FASTA sequence. This will take 10-15 min because we must search your protein sequence against a database to pick the related sequences. You can also [submit your protein sequence and related sequences](#) or [aligned sequences](#) if you already have them.

Results are deleted after 24 hours, so please save them!
[\[Preventing connection failures\]](#)

Protein sequence

Name of file containing protein query sequence ([fasta format](#)).
Choose File no file selected

-OR-

Paste in your protein query sequence ([Upload example](#)) ([fasta format](#)).

Enter the substitutions of interest ([format](#)):

-OR-

Upload a file containing substitutions of interest ([format](#)):
Choose File no file selected

Figure 2.3 SIFT Sequence Web Interface for Predicting Amino Acid Substitution Effects.

This figure shows the SIFT (Sorting Intolerant From Tolerant) Sequence web interface. The interface allows for the input of a protein query sequence either by uploading a FASTA format file or by pasting the sequence into the provided text box. Below the protein sequence input, specific substitutions of interest can be entered directly or uploaded in a separate file. SIFT processes the query by comparing the sequence against a database of related sequences to predict whether the substitutions are likely to be damaging or tolerated. Results are typically generated within 10-15 minutes and are automatically deleted after 24 hours unless saved.

2.6.4 ChimeraX software

For the structural analysis of CNOT3 mutations, ChimeraX (version 1.8) (Meng et al. 2023) was employed to visualise and analyse the effects of predicted missense mutations on the protein's three-dimensional structure. Protein structure files in PDB format and Uniprot ID were imported into ChimeraX, and the regions of interest corresponding to the mutated residues were highlighted. The software's visualization tools were used to assess any potential structural alterations.

2.7 Statistical analysis

Prior to performing statistical tests, data sets were assessed for normality using the Shapiro–Wilk test in GraphPad Prism 10. If data were normally distributed, parametric tests were used; otherwise, non-parametric tests were applied. In this study, all data sets were confirmed to be normally distributed; therefore, parametric tests were used for all statistical comparisons.

For the Alamar Blue assay, data were normalised to the mean of the untreated siControl group (set to 100%) and are presented as metabolic activity (% of control). Experiments were performed in biological triplicates. Statistical significance was determined using a two-way ANOVA to assess the effects of CNOT3 variant expression and siRNA treatment, followed by Tukey's multiple comparisons test. Data are expressed as mean \pm standard error of the mean (SEM), and a p-value < 0.05 was considered statistically significant. All statistical analyses were conducted using GraphPad Prism 10.

Chapter 3.
Generation of stable cell lines for
the conditional expression of siRNA-
resistant Flag-CNOT3 variants

3.1 Introduction

This chapter focuses on the development of a specialised molecular toolkit designed to investigate the role of CNOT3 in a human cell line. Specifically, the objective was to generate stable human cancer cell lines capable of conditionally expressing various siRNA-resistant Flag-CNOT3 variants. To achieve this, several key methodologies were employed. First, siRNA-resistant variants of CNOT3 cDNAs were engineered using site-directed mutagenesis. This process involved introducing specific point mutations into the *CNOT3* gene (Medema 2004; Aslam et al. 2009), rendering it resistant to degradation by siRNA. This approach was essential in allowing the exogenous (experimentally introduced) Flag-CNOT3 variants to be studied without interference from the endogenous CNOT3, which can still be silenced by siRNA-mediated degradation.

In addition to creating these resistant variants, truncated versions of CNOT3 were constructed (Sarkar et al. 2021; Suzuki et al. 2015), informed by prior structural and functional studies on the protein domains of CNOT3 (Lau et al. 2009; Boland et al. 2013; Buschauer et al. 2020). These truncated variants, lacking the N-terminal and the C-terminal domains, and the Not-Box region, were designed and constructed using the inverse PCR (Silva et al. 2017; Green and Sambrook 2019) to identify the regions of CNOT3 that are essential for its role in regulating cancer cell proliferation. By selectively removing these domains (Albert et al. 2000; Raisch et al. 2019; Collart 2003), the toolkit allows a comparison of the functional properties of these mutant CNOT3 proteins to the wild-type protein aimed to determine the importance of each region for CNOT3 activity in human cells.

Furthermore, to provide precise control over the expression of these CNOT3 variants, the Flp-In TREx system (O'Gorman, Fox, and Wahl 1991; Sauer 1994) was employed to establish stable cell lines for inducible expression of CNOT3. This system was selected for its ability to regulate gene expression in a temporal manner, meaning that the expression of the Flag-CNOT3 variants could be turned on or off as needed by the addition or removal of doxycycline. This capability is particularly valuable for minimising confounding factors that often arise from the variability associated with transient transfections, thereby ensuring more reliable experimental outcomes.

The Flp-In system, derived from *Saccharomyces cerevisiae*, was employed to streamline the creation of stable mammalian expression cell lines using a DNA recombination system. This system harnesses the recombinase enzyme Flp and site-specific recombination to facilitate the integration of the gene of interest into a specific location in the genome of mammalian cells. The Flp-In system relies on three vectors to generate isogenic stable mammalian cell lines expressing the gene of interest.

The first essential component of the Flp-In system is the pFRT/lacZeo target site vector (**Figure 3.1 A**), which is utilised to create a Flp-In host cell line. This vector contains a lacZ-Zeocin fusion gene regulated by the SV40 early promoter and an FRT site that serves as the binding and cleavage site for the Flp recombinase. The second major component is the pcDNA5-FRT-TetO expression vector (**Figure 3.1 B**), which facilitates the integration of the gene of interest into the genome of mammalian cells. The third crucial component is the pCAG-Flpo plasmid (**Figure 3.1 B**), which induces site-specific recombination between the FRT sites of the pFRT/lacZeo and pcDNA5-FRT vectors (O'Gorman, Fox, and Wahl 1991; Sauer 1994). In this study, the Flp-In T-REx host cell line, pre-engineered and generously provided by the Cornelia de Moor (CDM) group, eliminated the requirement for transfection with the pFRT/lacZeo plasmid. This streamlined approach facilitated the stable and inducible expression of the gene of interest (**Figure 3.1 C**), thereby providing a robust platform for assessing the impact of CNOT3 mutants on cellular proliferation.

The methodologies detailed in this chapter—including the construction of siRNA-resistant CNOT3 cDNAs, the generation of truncated CNOT3 variants, and the establishment of stable cell lines for inducible expression of CNOT3—collectively form a robust platform for conducting functional analyses of CNOT3 in human cells.

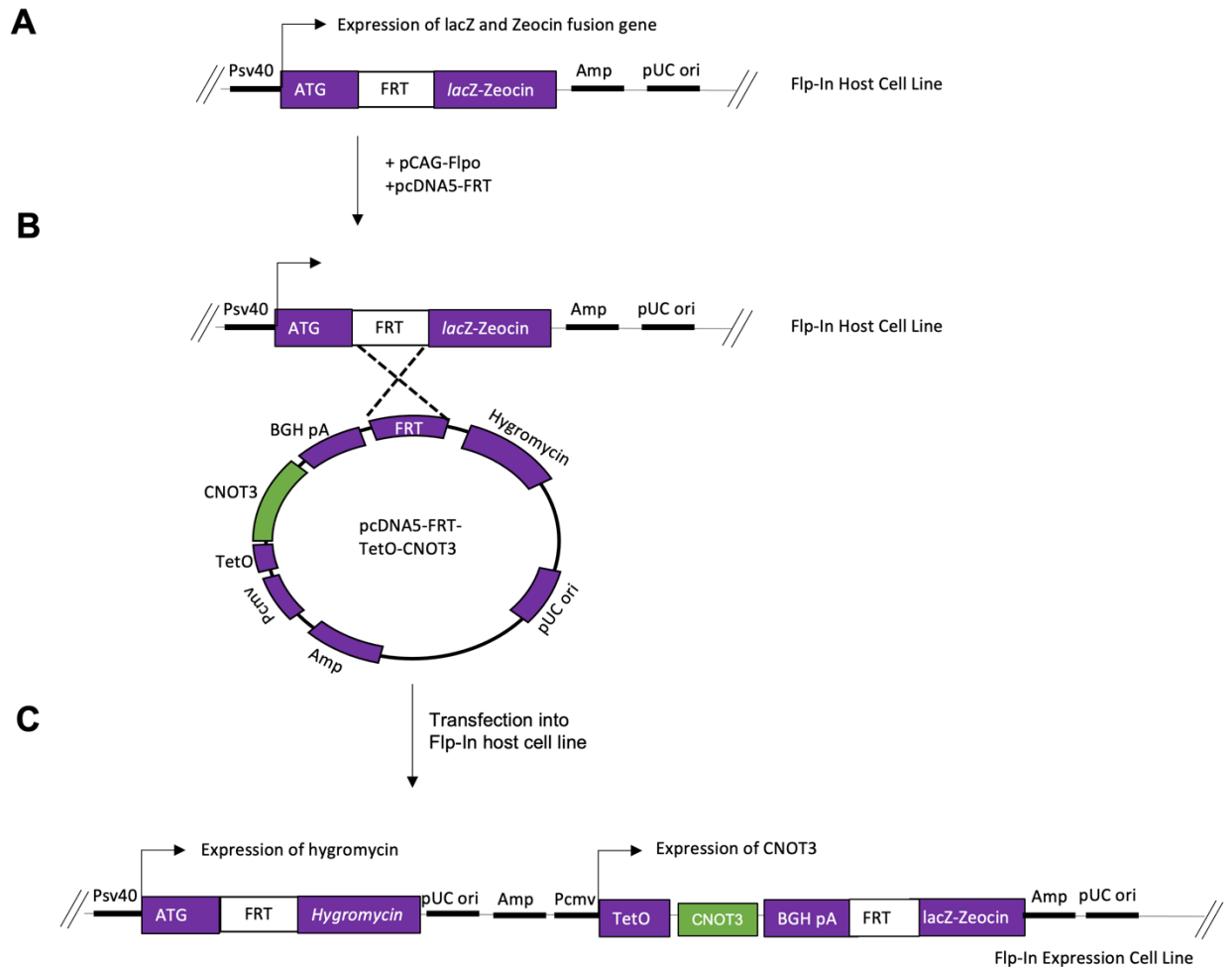


Figure 3.1. Overview of the generation of Flp-In TREX HEK293 stable cell line.

The schematic depicts the procedures of making the Flp-In TREX HEK293 and integration of the gene of interest. (A) The first step was to transfect the pFRT/lacZeo into HEK293 cells to generate a Zeocin-resistant Flp-In host cell line. (B) Next, the pcDNA5-FRT-TetTO containing the gene of interest (GIO) was cotransfected with the recombinase vector, pCAG-Flpo, into the Flp-In TREx HEK293. (C) The Flp plasmid conducts recombination between the FRT sites between the cell line and the expression vector.

3.2 Construction of plasmids expressing Flag-CNOT3 variants

3.2.1 Design of siRNA-resistant cDNA expression plasmids

To design siRNA-resistant *CNOT3* cDNA sequences, silent mutations were introduced at specific sites to avoid degradation by siRNAs while maintaining the amino acid sequence of the protein. This approach was chosen to ensure that the exogenously expressed CNOT3 remains unaffected by siRNA treatment, which specifically targets endogenous CNOT3 mRNA (**Figure 3.2 A**).

The introduction of these silent mutations was done by altering nucleotide sequences at the siRNA recognition sites without changing the encoded amino acids. This strategy ensures that the functionality of the CNOT3 protein remains intact while rendering the mRNA resistant to siRNA-mediated degradation (**Figure 3.2 B and D**). The sites for these mutations were selected based on several criteria outlined by (Medema 2004): first, the mutations were introduced in regions of the CNOT3 mRNA that are specifically recognised by the siRNA. By altering this region, the exogenous CNOT3 mRNA avoids being targeted by the siRNAs, ensuring stable expression in the presence of siRNA-mediated knockdown. The regions targeted by siRNA#1 and siRNA#2 were independently altered. Secondly, the mutations were designed to be silent, meaning they do not alter the amino acid sequence of the CNOT3 protein. At least two silent mutations were introduced—one around position 5-10 and another around position 10-15 of the siRNA target sequence, ideally separated by two codons. These mutations were selected to avoid possible G wobble base pairs, which could otherwise impact the specificity of the siRNA knockdown. Importantly, this strategy was applied to the *CNOT3* cDNA R2 but not to the *CNOT3* cDNA R1, where this precaution was not necessary.

This is crucial for maintaining the biological function of the protein while making the mRNA resistant to siRNA degradation. By introducing these mutations, the endogenous CNOT3 mRNA remains susceptible to siRNA knockdown, allowing for the study of CNOT3 function by comparing the knockdown effects on endogenous versus exogenously.

Mutations were introduced using site-directed mutagenesis (Zhang et al. 2021), employing inverse PCR (Silva et al. 2017) as one possible approach (**Figure 3.3 A**). The PCR reaction was treated with *Dpn* I to degrade the methylated parental template DNA, leaving the amplified PCR product intact (**Figure 3.3 B**). Following the digestion, the linear product was circularised by ligation to construct the expression plasmids with

siRNA-resistant Flag-CNOT3 cDNAs (**Figure 3.3 C**). The inverse PCR was used to introduce mutations at the positions corresponding to Asp216, Val218, and Glu219 to generate plasmid pCMV-Flag-CNOT3R1 containing a Flag-CNOT3 cDNA resistant against siRNA#1, and at positions corresponding to Thr692 and Ala694 to generate pCMV-Flag-CNOT3R2 containing a Flag-CNOT3 cDNA resistant against siRNA#2. All cDNA sequences were confirmed using the Sanger sequencing (**Figure 3.2 C and E**).

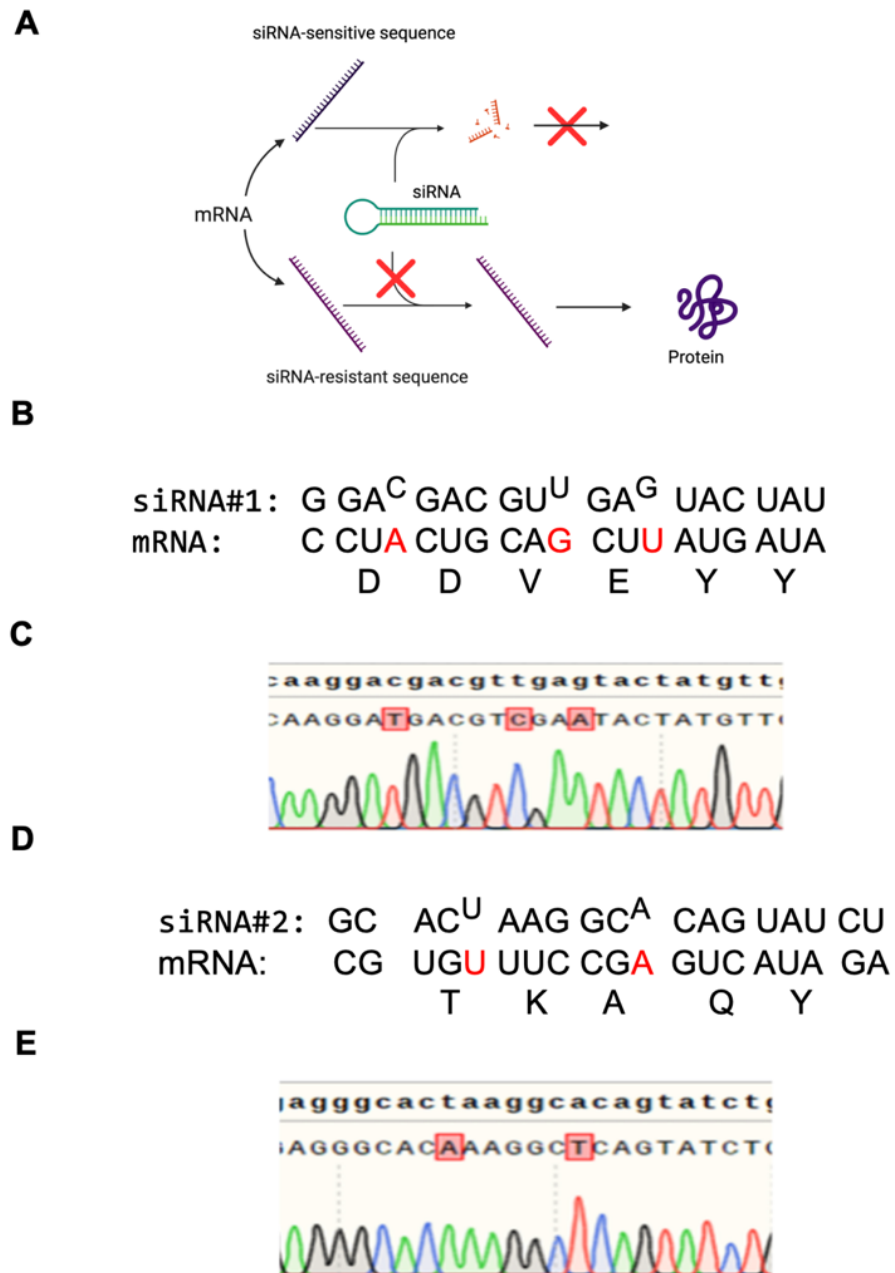


Figure 3.2 Design of siRNA-resistant CNOT3 cDNAs.

(A) The siRNA-sensitive sequence binds to the siRNA, leading to mRNA degradation and the inhibition of protein production. In contrast, the siRNA-resistant sequence, designed with point mutations that disrupt siRNA binding without altering the protein sequence, avoids degradation, allowing translation to proceed and protein expression to occur. (B) Three-point mutations were designed in the region recognised by siRNA#1, targeting the 5' end of CNOT3 mRNA. (C) Verification of the desired mutations in the sequences of Flag-CNOT3R1 using Sanger Sequencing. (D) Two-point mutations were designed in the region recognised by siRNA#2, targeting the 3' end of CNOT3 mRNA. (E) Verification of the desired mutations in the sequences of Flag-CNOT3R2 using Sanger sequencing.

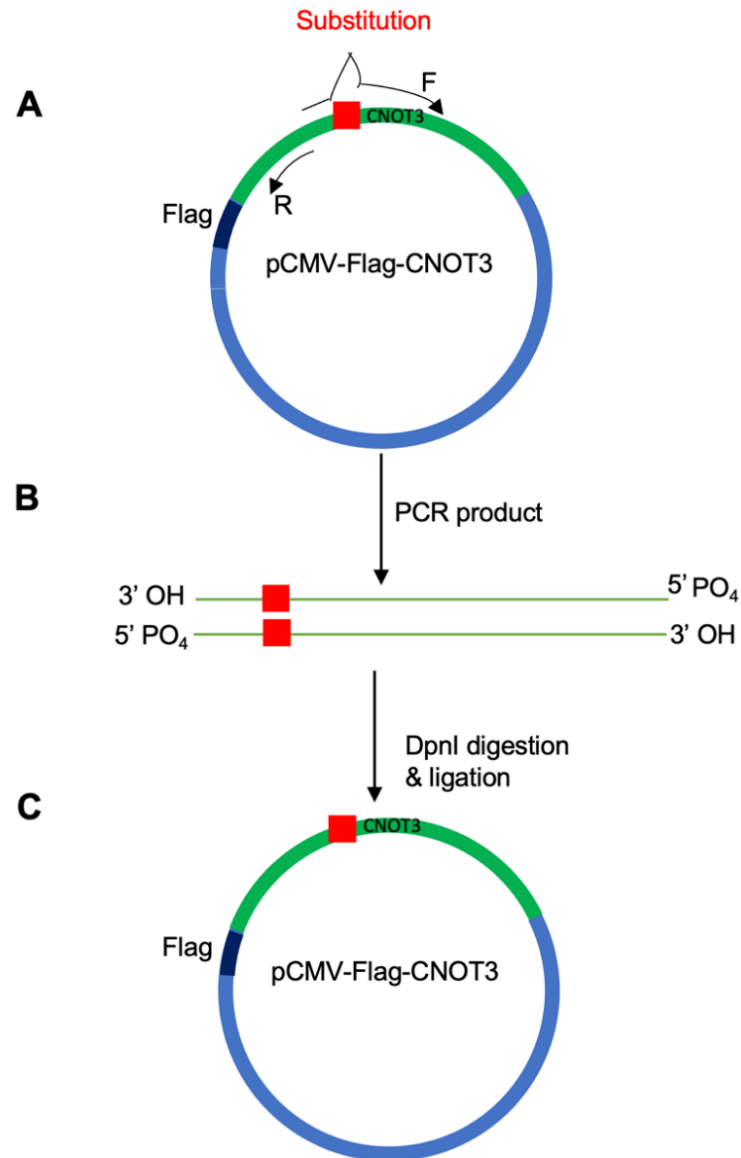


Figure 3.3 Construction of variants of Flag-CNOT3.

A schematic representation of site-directed mutation by inverse PCR. (A) Schematic representation of the pCMV-Flag-CNOT3 plasmid. The plasmid contains the CNOT3 gene, with forward (F) and reverse (R) primers indicated for site-directed mutagenesis. The substitution site for the mutation is also marked in red square. (B) Illustration of the PCR product obtained after amplification using the forward and reverse primers. The PCR product contains the desired mutation at the substitution site, with 5' (PO₄) phosphate and 3' (OH) hydroxy groups necessary for subsequent ligation. (C) Diagram showing the final mutant pCMV-Flag-CNOT3 plasmid after DpnI digestion and ligation. The red square represents the site of mutation within the CNOT3 gene.

3.2.2 Design of cDNA variants encoding Flag-CNOT3 lacking N- and C-terminal domains

To identify the role for specific regions of CNOT3, a series of plasmids were constructed using inverse PCR. The constructed plasmids encoded variants of the CNOT3 protein with specific deletions. These deletions were in the N-terminal and C-terminal domains, and the Not-Box region of CNOT3. The resulting truncated proteins were named CNOT3 Δ NTD 1-241, CNOT3 Δ CTD 651-751, and CNOT3 Δ Not-Box 701-751, indicating deletions at the N-terminal, the C-terminal, and the Not-Box region, respectively (**Figure 3.4 A and E**). The last construct which was made was Flag-CNOT3NTD-only, lacking the C-terminal and the middle region of CNOT3 protein.

To confer resistance to siRNA knockdown in the truncated CNOT3 variants, truncated plasmids were constructed using the parental plasmids Flag-CNOT3R1 and Flag-CNOT3R2. Specifically, Flag-CNOT3- Δ NTD was derived from Flag-CNOT3R2, while Flag-CNOT3- Δ CTD and Flag-CNOT3- Δ Not-Box were derived from Flag-CNOT3R1. This approach was chosen to ensure that the truncated CNOT3 variants remained expressed despite siRNA knockdown. To construct Flag-CNOT3only, the wild type of Flag-CNOT3, without any point mutation, was used as a parental fragment.

To create truncated variants of the Flag-CNOT3 protein (Sarkar et al. 2021), a technique called inverse PCR was employed. Unlike traditional PCR, where primers bind to the ends of the DNA sequence, inverse PCR uses primers that bind within the sequence, facing outwards (**Figure 3.5 A**). This allows for the amplification of a specific internal segment of the DNA.

Following the PCR reaction, the enzyme *Dpn* I was used to remove the original, unmodified DNA (**Figure 3.5 B**). The remaining, amplified DNA fragments were then joined together to create the new, shorter variants of Flag-CNOT3 (**Figure 3.5 C**).

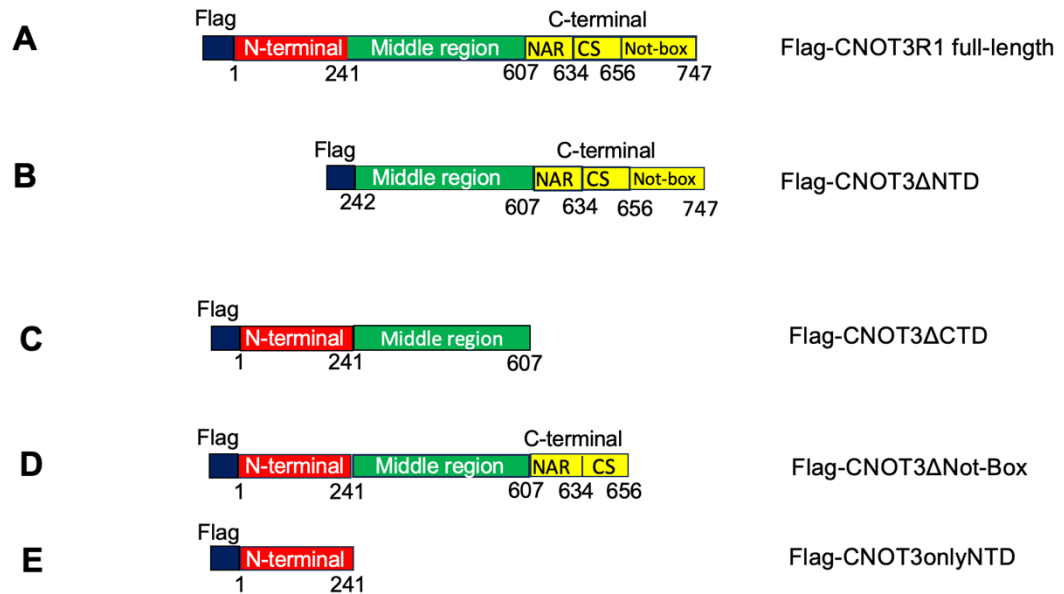


Figure 3.4 Design of variants of CNOT3 with truncations at the N-terminal and C-terminal domains, and Not-Box region.

(A) Overview of the domains of human CNOT3, which is characterised by a coiled-coil N-terminal domain (in red), a central linker region (in green) and a C-terminal (in red) that includes the Not-Box region, a CNOT1 anchor region (NAR) and a connector sequence (CS) (Boland et al. 2013). (B) Deletion of N-terminal in CNOT3. (C) CNOT3 lacking the C-terminal. (D) CNOT3 lacking the Not-Box region. (E) CNOT3 lacking C-terminal domain and middle region.

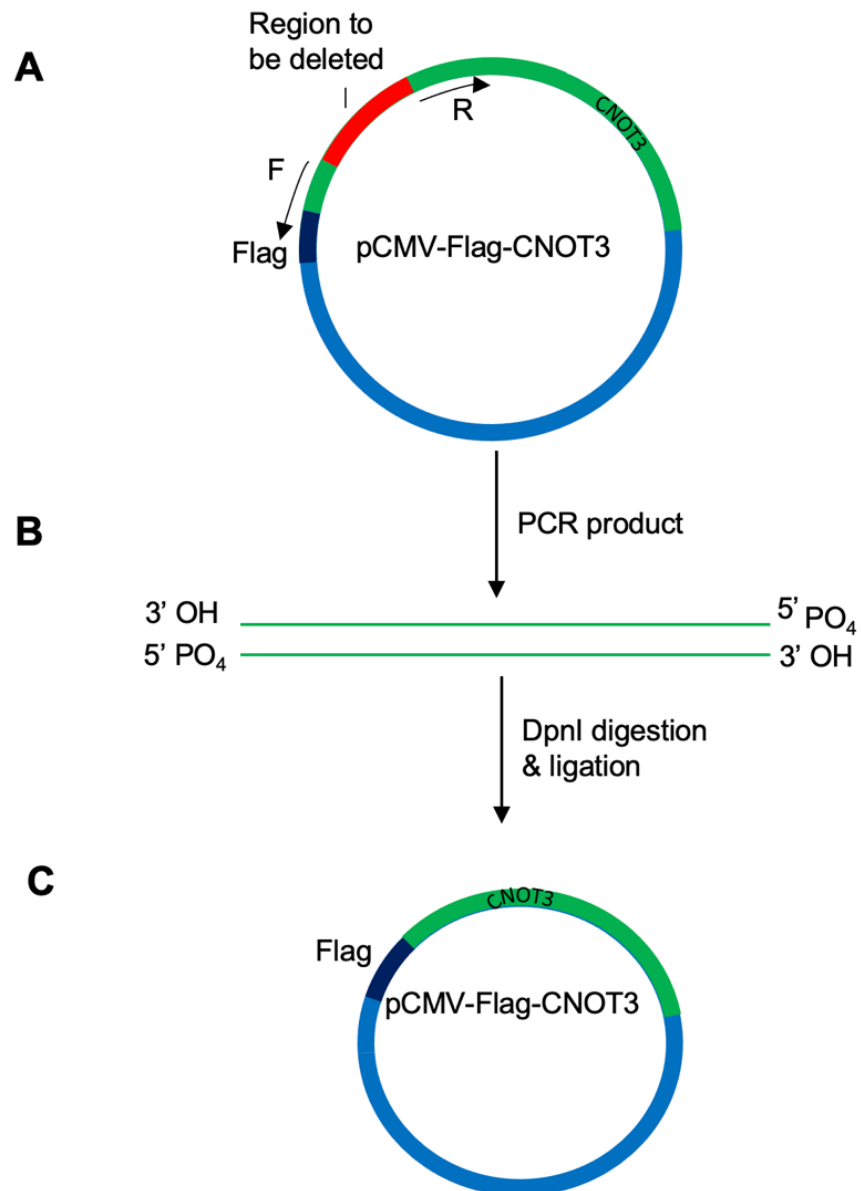


Figure 3.5 Construction of variants of CNOT3 lacking N-terminal and C-terminal domains.

Constructing of plasmids lacking N-terminal and C-terminal domains using the inverse PCR procedure. (A) The pCMV-Flag-CNOT3 plasmid, containing the *CNOT3* gene, was used as a template for PCR amplification. The forward (F) and reverse (R) primers were designed to amplify the plasmid, excluding the deletion region. (B) The resulting PCR product contained phosphorylated 5' and 3' hydroxy ends (indicated as 5' PO₄ and 3' OH). This product contains the plasmid sequence without the region targeted for deletion (C) The PCR reaction was treated with *DpnI* to digest the methylated, non-mutated parental DNA template. The linear PCR product was then ligated to form a circular plasmid, resulting in the pCMV-Flag-CNOT3 plasmid with the desired deletion.

3.3 Transient expression of Flag-CNOT3 mutants in HEK293 cell line

To evaluate the protein expression levels of Flag-CNOT3 variants, transient transfection of the constructed plasmids was carried out. The plasmids containing pCMV-FlagCNOT3, pCMV-Flag-CNOT3R1, pCMV-Flag-CNOT3R2, pCMV-Flag-CNOT3 Δ NTD, pCMV-Flag-CNOT3 Δ CTD, pCMV-Flag-CNOT3 Δ Not-Box and Flag-CNOT3onlyNTD were transiently transfected into HEK293 cell lines. Subsequently, 48 hours post-transfection, cell lysates were prepared, and protein expression was analysed through western blotting (**Figure 3.6**).

The utilisation of the Flag antibody (monoclonal rabbit antibody, Sigma-Aldrich) facilitated the assessment of the protein expression of the transfected Flag-CNOT3 variants without the contribution of endogenous CNOT3. Notably, the results demonstrated that all the CNOT3 variants were expressed at comparable levels in the HEK293 cell lines after transient expression, with the exception of the Flag-CNOT3onlyNTD variant, which exhibited significantly lower expression (**Figure 3.6**). The reduced expression level was consistently observed across multiple experiments, suggesting that the N-terminal domain alone may not be as efficiently expressed or stable as full-length CNOT3 or other truncated versions of CNOT3. Due to this marked decrease in expression, it was concluded that the Flag-CNOT3onlyNTD variant would not be suitable for further functional studies. This decision was made to ensure that subsequent experiments could be interpreted based on the function of the CNOT3 variant rather than its expression level.

The presence of endogenous CNOT3 was identified through the use of a CNOT3 antibody (mouse monoclonal antibody 4B8, AbNova). Endogenous CNOT3 expression was markedly lower than that of all Flag-CNOT3 variants. Truncated and full-length Flag-CNOT3 variants showed consistent expression levels, with full-length variants exhibiting slightly higher expression (**Figure 3.6**).

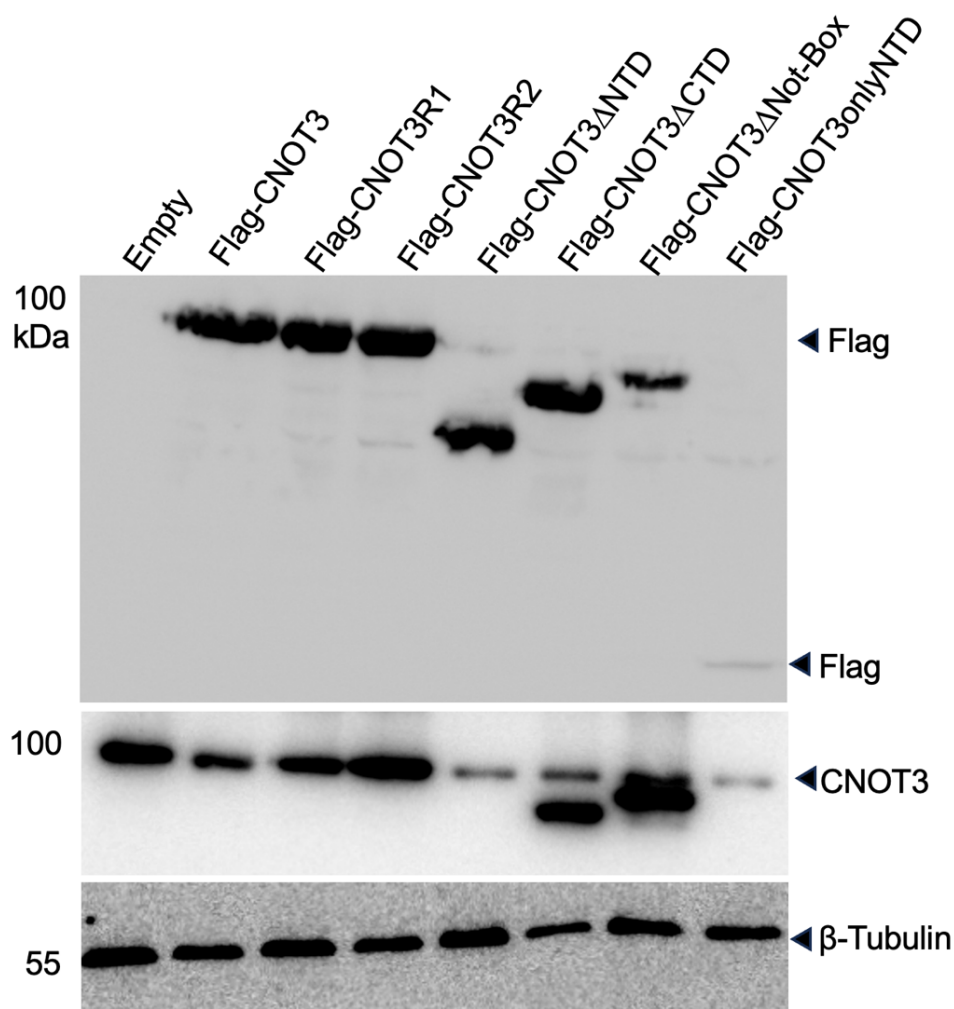


Figure 3.6 Transient expression of Flag-CNOT3 variants in HEK293 cell lines.

The Flag antibody was used to detect Flag-CNOT3 variants. Flag-CNOT3 variants were expressed equally for all the variants of CNOT3 except the Flag-CNOT3onlyNTD using Flag antibody (rabbit monoclonal antibody, Sigma-Aldrich). The Flag-CNOT3onlyNTD was expressed at the lowest level. Endogenous CNOT3 was detected using the CNOT3 antibody (mouse monoclonal antibody 4B8, AbNova). In Flag-CNOT3ΔNTD, the epitope recognised by the anti-CNOT3 antibody is not present. β-tubulin was used as a loading control.

In addition to using anti-CNOT3 antibodies, the effect of Flag-CNOT3 expression on the CNOT1 and CNOT2 subunits of the Ccr4-Not complex was also evaluated, as well as the expression of ribosomal proteins S3 and L7a. The analysis of the levels of CNOT1 and CNOT2 indicated that their expression levels were relatively similar regardless of the expression of Flag-CNOT3 variants, except for Flag-CNOT3 Δ NTD, which led to increased CNOT1 levels. It is unclear why Flag-CNOT3 Δ NTD led to increased CNOT1 levels, and further studies are needed to understand this increase. However, the expression of S3 and L7a remained consistent across all Flag-CNOT3 variants. To control for protein loading, β -tubulin was utilised as a loading control.

Figure 3.7 shows the transient expression of Flag-CNOT3 variants without Flag-CNOT3onlyNTD in HEK293 cell lines

3.4 Generating Flag-CNOT3 Cell Lines using G-418

Having confirmed Flag-CNOT3 expression via transient transfection, the focus shifted towards establishing a stable cell line model. This was crucial for downstream applications. To achieve this, stable HEK293 cell lines expressing Flag-tagged CNOT3 variants were initially pursued. A pCMV backbone vector carrying the Flag-CNOT3 constructs was transfected into HEK293 cells, and G-418 selection (1.0 mg/ml) was applied to isolate a stable integrant. However, the cells unexpectedly exhibited resistance to G-418, precluding the establishment of stable cell lines. The underlying cause of this G-418 resistance, despite troubleshooting efforts, remained elusive. Potential factors include anomalies within the plasmid construct, suboptimal transfection efficiency, or an unforeseen mechanism of G-418 resistance in this cellular context. Therefore, an alternative strategy for generating stable cell lines was developed, involving the use of the Flp-In T-REx HEK293 system.

3.5 Construction of vectors for inducible expression of Flag-CNOT3 variants

Because the initial attempt to generate stable cell lines expressing Flag-CNOT3 variants in HEK293 cells using G418 selection did not succeed, an alternative approach was adopted. This involved using the Flp-In T-REx HEK293 (Ward, Alvarez-Curto, and Milligan

2011; O'Gorman, Fox, and Wahl 1991) cell lines to generate stable cell lines. The Flp-In TREx system is a more controlled and reliable method for creating stable, inducible expression of the gene of interest.

To generate the expression plasmids suitable for this approach, the CNOT3 cDNAs from the pCMV-Flag-CNOT3 were first amplified by polymerase chain reaction (PCR) to incorporate an *Afl*III restriction site at the 5' terminus of the Flag-CNOT3 cDNAs. The PCR product was then digested with *Afl*III and *Not*I, producing compatible 5' overhangs with *Psp*OMI (5' GGCC). The digested Flag-CNOT3 cDNAs were subsequently purified and used for further cloning (**Figure 3.8 A**). This process involved the use of specially designed primers to ensure the incorporation of the *Afl*III restriction site into the amplified product. Next, plasmid pcDNA5-FRT-TetO-fSNAP (**Figure 3.8 B**) (Braun et al. 2018) was subjected to digestion with the restriction endonucleases *Afl*III and *Psp*OMI to excise the fSNAP cDNA insert (**Figure 3.8 B**). Following the digestion, the linearised plasmid backbone was purified by agarose gel electrophoresis.

The purified Flag-CNOT3 cDNAs were then ligated into the digested pcDNA5-FRT-TetO vector, resulting in the generation of the following expression plasmids: pcDNA5-FRT-TetO-Flag-CNOT3, pcDNA5-FRT-TetO-Flag-CNOT3R1, pcDNA5-FRT-TetO-Flag-CNOT3R2, pcDNA5-FRT-TetO-Flag-CNOT3 Δ NTD, pcDNA5-FRT-TetO-Flag-CNOT3R1 Δ CTD, and pcDNA5-FRT-TetO-Flag-CNOT3R1 Δ Not-Box (**Figure 3.8 C**).

All cDNA sequences were confirmed by DNA sequencing to ensure the accuracy of the constructs.

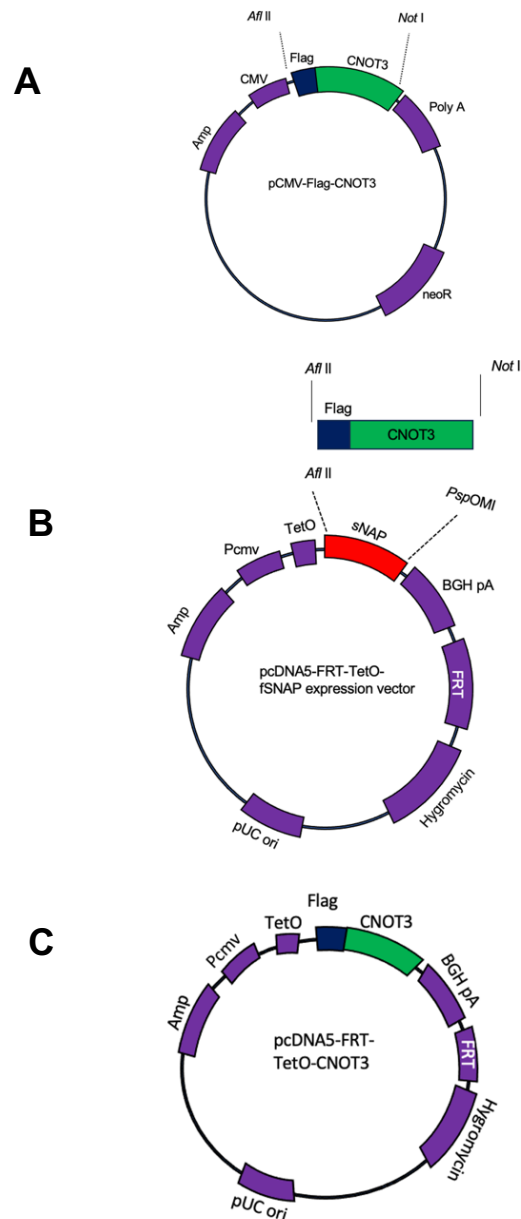


Figure 3.8 Construction of vectors for inducible expression of Flag-CNOT3 variants.

(A) Subcloning of Flag-CNOT3 cDNA into the expression vector pcDNA5-FRT-TetO. The Flag-CNOT3 cDNAs were amplified by PCR using primers that incorporated an AflII site at the 5' end, as well as the BGHR primers (see Materials and Methods for further details). After amplification, the template DNA was removed by DpnI digestion.

(B) The expression vector was digested with AflII and PspOMI restriction enzymes to generate cohesive ends compatible with the CNOT3 fragments. The recombinant plasmids were obtained after ligation and transformation of *E. coli* DH5 α .

(C) pcDNA5-FRT-TetO plasmid derived from (B), where the Flag-CNOT3 insert has been cloned, resulting in a Tet-inducible expression constructs for Flag-tagged CNOT3 variants.

3.6 Generation of Flp-In TREx HEK293 stable cell lines containing inducible Flag-CNOT3 variants

To generate stable cell lines with inducible Flag-CNOT3 expression cassettes, Flp-In TREx 293 cells were co-transfected with pCAG-Flpo (Xue, Atallah, and Scanziani 2014) encoding the Flp recombinase and the following expression plasmids: pcDNA5-FRT-TetO-FlagCNOT3, pcDNA5-FRT-TetO-FlagCNOT3R1, pcDNA5-FRT-TetO-FlagCNOT3R2, pcDNA5-FRT-TetO-Flag-CNOT3 Δ NTD, pcDNA5-FRT-TetO-Flag-CNOT3 Δ CTD, and pcDNA5-FRT-TetO-Flag-CNOT3 Δ Not-Box.

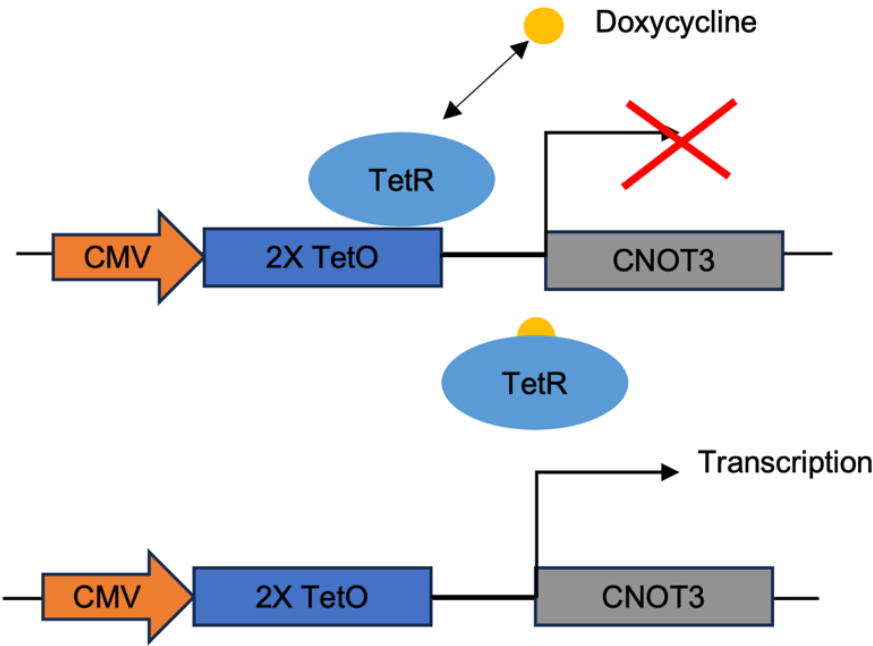
After transfection, hygromycin (250 μ g/ml) was added to the culture medium 48 hours post-transfection to select transfected cells. These cells were then cultured in a medium containing 250 μ g/ml hygromycin, resulting in the generation and parallel maintenance of six mixed cell populations.

After cell populations were established, the expression of Flag-CNOT3 variants was evaluated by western blotting. To this end, cells were seeded in six-well culture plates in a medium with doxycycline (1.0 μ g/ml) to induce the expression of Flag-CNOT3 cDNA, while cells were also cultured without doxycycline. Subsequently, total cell lysates were prepared. Following the separation of proteins by SDS-PAGE, immunoblotting was performed to detect the expression of Flag-CNOT3 variants. The results confirmed the expression of Flag-CNOT3 variants, using the Flag antibody, in the presence of doxycycline and negligible expression in the absence of doxycycline (**Figure 3.9 A and B**). It also shows that the expression cassette was specifically integrated into the Flp-In site in the host genome (O'Gorman, Fox, and Wahl 1991). By utilising the Flp-In system and targeting the FRT site, the expression cassette is inserted into a known genomic location with a defined regulatory environment. This not only minimises the positional effect but also ensures a more consistent and predictable gene expression profile across different cell clones. The FRT site is typically located in a "safe harbour" region within the genome, which is devoid of strong regulatory elements that could interfere with gene expression. This reduces the likelihood of variations in gene expression levels due to the influence of neighbouring sequences, resulting in a more controlled and reproducible experimental system.

Based on the Western blot analysis using the anti-CNOT3 antibody (mouse monoclonal antibody 4B8, AbNova), it was observed that endogenous CNOT3 was detected in the Flag-CNOT3 and Flag-CNOT3R1 variants at similar levels in the presence of doxycycline but not in the absence of doxycycline. However, in the case of Flag-CNOT3R2, the expression level was found to be higher in the presence of doxycycline. Additionally, the expression of endogenous CNOT3 in the other variants remained comparable, both in the presence and absence of doxycycline. This suggests that the overexpression of Flag-CNOT3 variants has no effect on the expression level of endogenous CNOT3.

Furthermore, the expression levels of various components of Ccr4-Not, including CNOT1, CNOT2, CNOT7, and the ribosomal protein S3 was analysed. This investigation revealed that the expression levels of these components remained similar and were unaffected by the overexpression of Flag-CNOT3 mutants, as illustrated in **Figure 3.9 B**.

A



B

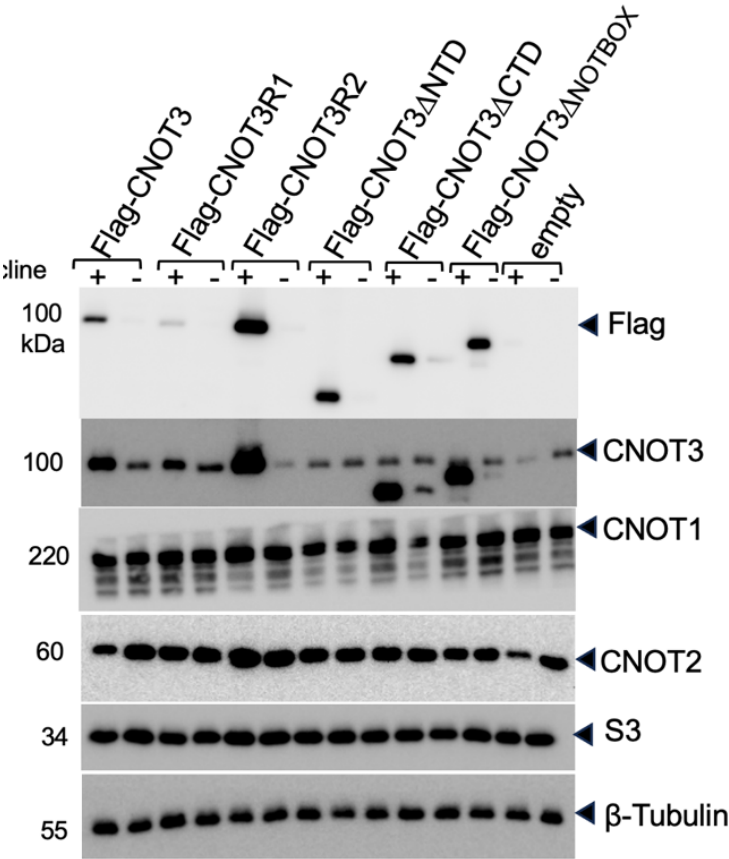


Figure 3.9 Expression of Flag-CNOT3 variants in stably transfected Flp-In TREx HEK293 cell lines.

(A) Schematic representation of the inducible Flag-CNOT3 expression cassette. In the absence of doxycycline, the Tet repressor (TetR) binds to the TetO operator sequence thereby inhibiting transcription. Doxycycline binds TetR, preventing binding to TetO sequences, allowing CMV promoter-driven CNOT3 transcription. (B) Western blot analysis of stable cell lines containing inducible Flag-CNOT3 expression cassettes in the presence and absence of doxycycline (+/-). The results confirmed the expression of Flag-CNOT3 variants, using the Flag antibody, in the presence of doxycycline and no expression in the absence of doxycycline. The expression of CNOT3 endogenous was detected using anti-CNOT3 antibody (mouse monoclonal antibody 4B8). The expression level of other components of the Ccr4-Not complex was detected using CNOT1, CNOT2, and S3 as ribosomal antibodies. β -Tubulin was used as the loading control.

3.7 Flag-CNOT3R1 and Flag-CNOT3R2 mRNAs are resistant against siRNA#1 and #2, respectively

Next, we confirmed the resistance of the mRNAs encoding Flag-CNOT3R1 and Flag-CNOT3R2 to siRNA#1 and siRNA#2, respectively. To determine the effectiveness of the silent mutations introduced to confer resistance of the CNOT3 variants against siRNA knockdown, both stable cell lines expressing the Flag-CNOT3R1 and Flag-CNOT3R2 variants were transfected with siRNA duplexes #1 and #2. After 48 hours, total cell lysates were prepared and subjected to western blotting using anti-CNOT3 and anti-Flag antibodies (**Figure 3.10 A**). The results revealed that the point mutations in the N-terminal domain of Flag-CNOT3R1 resulted in resistance against siRNA#1 but not siRNA#2. By contrast, endogenous CNOT3 proteins levels were reduced upon transfection with siRNA#1 and siRNA#2.

Further analysis of the siRNA-resistant mRNA encoding Flag-CNOT3R2 demonstrated that point mutations in the C-terminal domain of Flag-CNOT3R2 conferred resistance against siRNA duplex 2 but not siRNA duplex 1 (**Figure 3.10 B**). This was evidenced by the decrease in the expression level of endogenous CNOT3 upon transfection with siRNA#1 and siRNA#2. Notably, transfection of non-targeting siRNAs in both cell lines did not result in a significant reduction at the protein level.

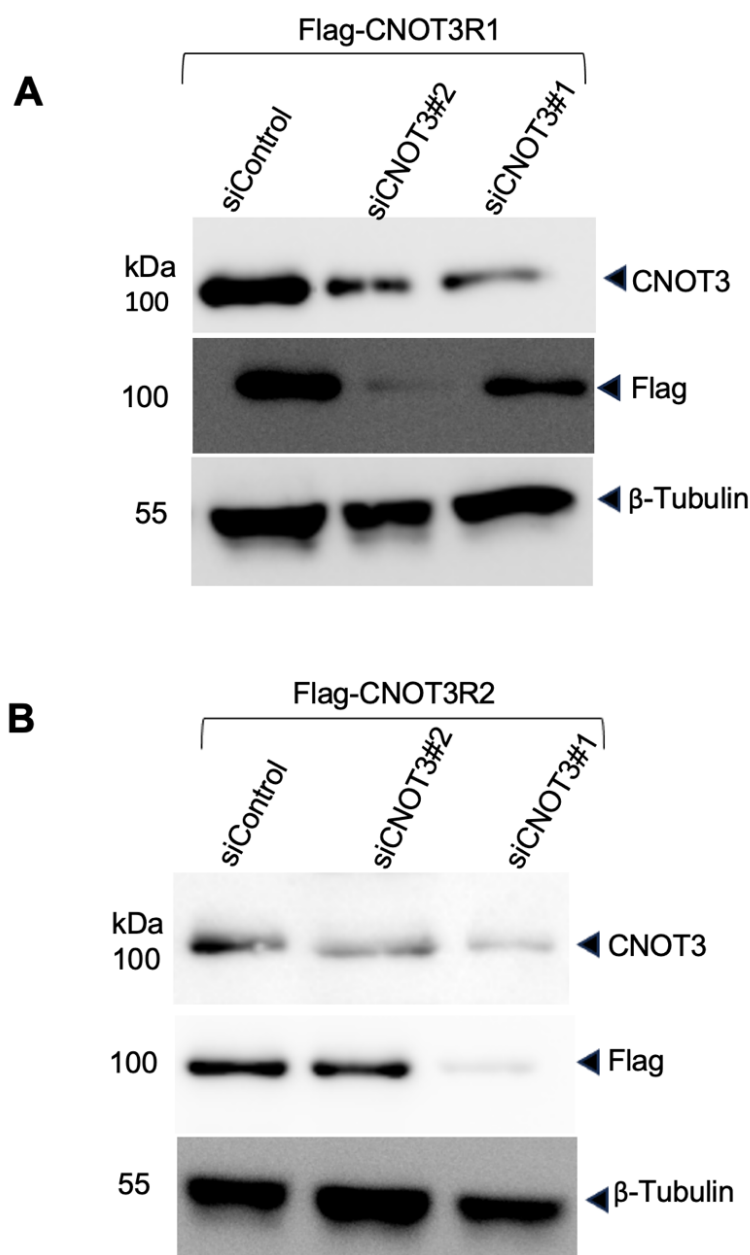


Figure 3.10 mRNA encoding Flag-CNOT3R1 and Flag-CNOT3R2 are resistant to knockdown by siRNA#1 and siRNA#2, respectively.

(A) The Flag-CNOT3R1 mRNA was resistant to siRNA#1, whereas knockdown was observed in cells transfected with siRNA#2. The endogenous CNOT3 was sensitive to siRNA#1 and #2. No knockdown was observed in cells transfected with non-targeting siRNA. (B) The Flag-CNOT3R2 mRNA was resistant to siRNA#2 but sensitive to siRNA#1. No knockdown was observed in cells transfected with non-targeting siRNA. The endogenous CNOT3 was knocked down in stable cell lines expressing Flag-CNOT3R2. β -tubulin was used as the loading control.

3.8 Discussion

In this chapter, the design of siRNA-resistant Flag-CNOT3 cDNAs (with mutations that prevent siRNA from targeting the corresponding mRNA) was described. In addition, the design of Flag-CNOT3 deletion variants and the generation of stable cell lines containing inducible Flag-CNOT3 expression cassettes were reported.

The design of siRNA-resistant cDNA expression constructs (**Figure 3.2**) and the subsequent validation of their protein expression (**Figure 3.10**) were essential to ensure the successful implementation of downstream applications. Specifically, three silent point mutations were introduced into the region of the CNOT3 cDNA that corresponds to the N-terminal domain of the encoded protein, and two additional mutations were introduced in the region corresponding to the C-terminal domain. These silent mutations ensured that the exogenous cDNA was resistant to siRNA-mediated degradation while maintaining the integrity of the protein sequence. This approach allowed for the selective knockdown of endogenous CNOT3 while inducing the expression of siRNA-resistant variants via doxycycline. The overexpression of these variants enables the assessment of their ability to complement the cell growth defect observed in HEK293 cells, shedding light on the functional roles of CNOT3 in regulating cell growth. The successful design of variants of Flag-CNOT3 resistant to siRNA aligns with the widespread use of this method, as demonstrated in numerous studies (Aslam et al. 2009; Medema 2004; Zheng, Tang, and Tao 2005; Morita et al. 2012; Jiang and Price 2014). By introducing targeted point mutations into the CNOT3 cDNA, the constructs were engineered to resist siRNA targeting, allowing the selective silencing of endogenous CNOT3. This strategy effectively decoupled the effects of exogenous Flag-CNOT3 variants from those of the endogenous protein, facilitating a precise interpretation of the function of CNOT3 through complementation analysis—a classic genetic technique commonly used in mammalian cells (Medema 2004).

The generation of truncated variants, lacking the N-terminal and C-terminal domains, and the Not-Box region, will enable the identification of critical regions within the CNOT3 protein that influence cell proliferation and gene expression in human cancer cells. The similar levels of expression of these truncated variants, confirmed by western blotting, suggest that the truncated proteins were stable and functional within the cellular environment, except for the low expression of the Flag-CNOT3onlyNTD. As a result, this variant was not further investigated.

The Flp-In TREx system (O'Gorman, Fox, and Wahl 1991; Sauer 1994; Schlake and Bode 1994) provided an efficient and reliable method for generating inducible stable cell lines, essential for controlled expression studies. Although the system is designed to integrate a single copy of the cDNA into the same genomic locus in all cells, the resulting populations may still contain heterogeneity due to incomplete integration or selection. This could result in some cells lacking the expression cassette while others carry it, potentially leading to variable expression levels. Nevertheless, the system offers an overall advantage by minimising positional effects and promoting uniform expression in successfully integrated clones. The absence of background expression in the absence of doxycycline further validated the inducibility of the system, minimising potential confounding factors in downstream analyses.

Upon overexpressing the Flag-CNOT3 variants in TREx Flp-In cell lines, a notable difference was observed in the expression levels of Flag-CNOT3R1 and Flag-CNOT3R2, despite their identical transient expression profiles in HEK293 cells. Specifically, the expression of endogenous CNOT3 in Flag-CNOT3R2 was higher compared to the other variants, suggesting potential differences in the regulatory mechanisms governing these variants. One possible explanation for the lower expression of Flag-CNOT3R1 could be the synonymous mutations introduced in the cDNA sequences. These mutations, while not altering the amino acid sequence, might have influenced codon usage, potentially changing rare codons to more abundant ones or vice versa (Komar 2007). Although this is a plausible explanation, it is less likely because no difference was observed in the transient transfection experiments in HEK293 cells. A more likely explanation is that the differences in expression levels may stem from a heterogeneous population of cells within the TREx Flp-In cell lines, where some cells do not contain the expression cassette while others do. This mixture could lead to variable expression levels, with only a subset of cells fully expressing the Flag-CNOT3 variants. These results align with studies by (Quax et al. 2015) and (Plotkin and Kudla 2011), which showed that rare codons can reduce translation efficiency, but this effect may not have been significant in the transient expression system.

Additionally, the higher expression of Flag-CNOT3R2 might be due to a more optimal codon usage that enhances translation efficiency, as supported by (Zhou et al. 2016; Presnyak et al. 2015). However, cell population variability remains a more likely contributor to the observed differences in the stable expression system.

Interestingly, the expression levels of endogenous CNOT3 in Flag-CNOT3 Δ NTD, Flag-CNOT3R1 Δ CTD, and Flag-CNOT3R1 Δ Not-Box were comparable, suggesting that the removal of these specific domains did not significantly affect stability of the endogenous protein.

Investigation into the expression levels of other components of the Ccr4-Not complex revealed that expression of Flag-CNOT3 variants did not lead to alterations in the expression of other Ccr4-Not components. This suggests that CNOT3 is not a rate-limiting component in the assembly of the complex. In addition, it suggests that the overexpression of Flag-CNOT3 does not interfere with the assembly of the complex. These findings are in line with previous studies, which demonstrated that individual subunits such as CNOT3 contribute to the complex's function but do not necessarily alter the expression or assembly of other components (Collart and Panasenko 2012; Chalabi Hagkarim and Grand 2020).

This study is in line with the findings of Suzuki et al (Suzuki et al. 2015), who investigated the role of CNOT3 in mouse embryonic fibroblasts (MEFs). While both studies demonstrate that CNOT3 is crucial for cell viability and Ccr4-Not complex integrity, they employ distinct but complementary experimental approaches.

The thesis work utilises human HEK293 cells with an inducible Flp-In T-REx system, enabling precise control over CNOT3 expression and systematic analysis of domain function through various CNOT3 mutants. This approach can reveal structural requirements for CNOT3 function, particularly the importance of both N-terminal and C-terminal domains for cell viability. While using human cells provides direct relevance to human biology, the transformed nature of HEK293 cells may affect certain cellular responses.

The distinct cellular contexts - transformed human cells versus primary mouse cells - may explain any divergent phenotypes between the studies. However, the core findings regarding CNOT3's importance for complex stability and cell survival are conserved across both systems, suggesting fundamental roles for CNOT3 that transcend species and cell type differences. Together, these studies provide complementary insights into CNOT3 biology, combining detailed structural analysis with mechanistic understanding of its role in cell survival.

The key advantages of the thesis approach include the ability to precisely control CNOT3 expression through the inducible system, systematic analysis of domain functions through mutants, and direct relevance to human biology. However, this system has limitations compared to Suzuki et al.'s approach - notably the use of transformed cells that may have altered cellular responses, particularly regarding cell death pathways. Additionally, the thesis approach is more complex, time-consuming, and technically demanding to establish than their direct knockout strategy in primary cells.

Despite these methodological differences, the constructs and cell lines generated in this chapter form the basis for assessing the role of CNOT3 in cell proliferation in later chapters.

Chapter 4.
Essential Domains of CNOT3 for
Cell Proliferation

4.1 introduction

Cell proliferation is regarded as a fundamental process in cellular biology, essential for growth, development, and tissue repair (Hanahan 2022). Dysregulation of cell proliferation is known to be a hallmark of many diseases, including cancer (Almalki 2023). Previous studies have highlighted the pivotal role of the Ccr4-Not complex, a multi-subunit protein assembly involved in mRNA metabolism and gene expression regulation, in controlling cell proliferation (Aslam et al. 2009). Within this complex, the CNOT3 protein has emerged as a key regulator for integrity and cell viability (Cejas et al. 2017; Suzuki et al. 2015). However, it remains unclear which domain of CNOT3 is important for cell proliferation in human cancer cell lines.

Building on prior research, this chapter focuses on addressing this gap by investigating the functional roles of different CNOT3 domains in regulating human cell proliferation. Specifically, it examines the N-terminal and C-terminal domains, and Not-Box regions of CNOT3 in relation to their ability to drive proliferation and their interactions with other subunits of the Ccr4-Not complex. The established stable cell lines (described in Chapter 3) were employed as a toolkit to explore these structural regions in detail.

To ensure reliable experimental conditions, the influence of doxycycline on cell proliferation in the induced stable cell lines expressing various CNOT3 constructs was first examined. It was confirmed that doxycycline induction did not interfere with cell growth, validating the experimental conditions for subsequent investigations.

This chapter investigated whether CNOT3 knockdown affects cell proliferation in HEK293 cells. siRNA-mediated silencing was employed, and the results aligned with previous studies (Ghashghaei et al. 2024; Aslam et al. 2009; Cejas et al. 2017) showing a significant reduction in proliferation. These findings underscore the essential role of CNOT3 in promoting cellular growth in HEK293 cells, consistent with observations in other cell lines.

Finally, to further delineate the roles of the N-terminal and C-terminal domains of CNOT3, siRNA-resistant stable cell lines expressing various CNOT3 mutants were developed. In line with findings from Suzuki et al. on mouse embryonic fibroblast stem cells, this study demonstrated the vital roles of both the N-terminal and C-terminal domains in

Chapter 4. Essential domains of CNOT3 for cell proliferation

promoting cell proliferation. Specific mutations within these regions conferred resistance to siRNA knockdown, allowing cells to maintain normal proliferation rates. In summary, this chapter provides a detailed analysis of the functional roles of the N and C termini of CNOT3 in regulating human cell line. These findings contribute to an understanding of cellular growth mechanisms, in particular in diseases characterised by abnormal cell proliferation. Furthermore, this work adds to the broader context of mRNA metabolism and gene expression regulation in cellular biology.

4.2 Doxycycline-induced expression of Flag-CNOT3 variants in stably transfected cell lines does not affect cell proliferation

To investigate the effects of overexpressing CNOT3 variants on cell proliferation, the rate of cell proliferation was monitored in stable cell lines expressing different CNOT3 variants, including Flag-CNOT3 (wild type), Flag-CNOT3R1, Flag-CNOT3R2, Flag-CNOT3 Δ NTD, Flag-CNOT3R1 Δ CTD, Flag-CNOT3R1 Δ Not-Box, along with an empty vector control. These stably induced cell lines were cultured with and without doxycycline, and their metabolic activity was measured 48 hours post-induction using the AlamarBlue assay. This assessment provides an estimate of cell viability at the time of measurement, where higher metabolic activity indicates a greater number of viable cells.

The obtained data revealed that overexpression of CNOT3 variants in cell lines using doxycycline did not have a discernible impact on cell proliferation, as depicted in **Figure 4.1**. This finding suggests that in the context of this study, overexpression of CNOT3 variants using a doxycycline-inducible system did not significantly influence the proliferation activity of the investigated cell lines.

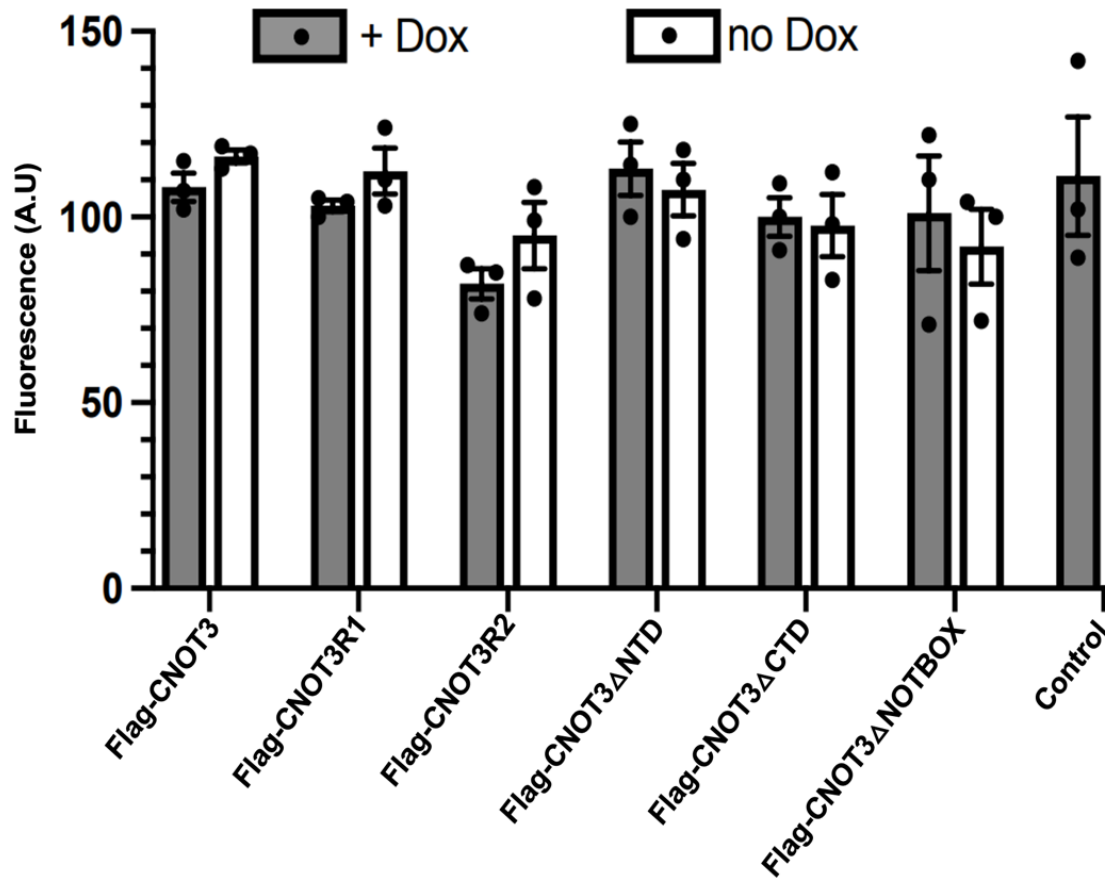


Figure 4.1 Doxycycline-induced expression of Flag-CNOT3 variants in stably transfected cell lines does not affect cell proliferation.

Stable cell lines expressing Flag-CNOT3, Flag-CNOT3R1, Flag-CNOT3R2, Flag-CNOT3 Δ NTD, Flag-CNOT3R1 Δ CTD, and Flag-CNOT3R1 Δ Not-Box were seeded, and expression was induced by adding doxycycline. The metabolic activity of the cell lines was measured after 48 hours using the AlamarBlue assay. A stable cell line containing an empty expression cassette was used as control. Error bars indicate the standard deviation of three independent experiments (n=3).

4.3 CNOT3 knockdown leads to a decrease in cell proliferation in HEK293 cells

To assess the effect of knocking down the expression of CNOT3 on cellular proliferation, HEK293 cells were transfected with siRNAs specifically designed to target CNOT3 mRNA. Three distinct siRNA duplexes were employed in this study: siRNA#1, which targets the CNOT3 mRNA encoding the N-terminal domain; siRNA#2, which targets the CNOT3 mRNA encoding the C-terminal domain; and a non-targeting control siRNA. These siRNAs were selected to ensure that observed effects upon knockdown of CNOT3 expression was not due to off-target effects of the siRNAs.

The effectiveness of CNOT3 knockdown was validated through western blot (**Figure 4.2 A**). HEK293 cells were seeded and transfected with siRNA#1, siRNA#2, and two distinct non-targeting control siRNAs. Total cell lysate was collected at 72 hours after transfection for western blot analysis. Western blot results showed that transfection with siRNA#1 and siRNA#2 resulted in a significant reduction in CNOT3 protein levels in the HEK293 cells. In contrast, cells treated with non-targeting control siRNAs showed no discernible impact on CNOT3. Moreover, siRNA transfection did not result in changes in the level of β -Tubulin. These experiments validated the specificity and efficacy of the siRNAs, as the absence of knockdown in the non-targeting control groups confirmed that the observed effects were due to specific targeting of CNOT3 mRNA.

Next, to determine the effect of CNOT3 knockdown on cell proliferation in HEK293 cells, two different phenotypic assays AlamarBlue and Crystal Violet assays, were conducted. The same siRNAs used for Western blotting were utilized in these assays. The AlamarBlue assay was performed concurrently with the western blotting (72 h after transfection) to assess immediate effects, while the Crystal Violet assay was performed five days post-transfection to reveal the sustained effects of the siRNAs at a later time point.

The AlamarBlue assay results indicated a significant decrease in cellular metabolic activity in the CNOT3 knockdown groups compared to the control (**Figure 4.2 B**). This suggests that CNOT3 is important for cell proliferation in HEK293.

Chapter 4. Essential domains of CNOT3 for cell proliferation

Similarly, the Crystal Violet phenotypic assay further corroborated these findings by showing a reduction in cell viability and proliferation five days post-transfection in cells treated with siRNA#1 and siRNA#2, but not in cells treated with the non-targeting control siRNA (**Figure 4.2 C**).

These findings contribute to the growing body of knowledge surrounding the functional significance of CNOT3 in cellular processes. This study demonstrates the role of CNOT3 in cellular proliferation and highlights the effectiveness of using targeted siRNA duplexes to dissect the functions of specific genes.

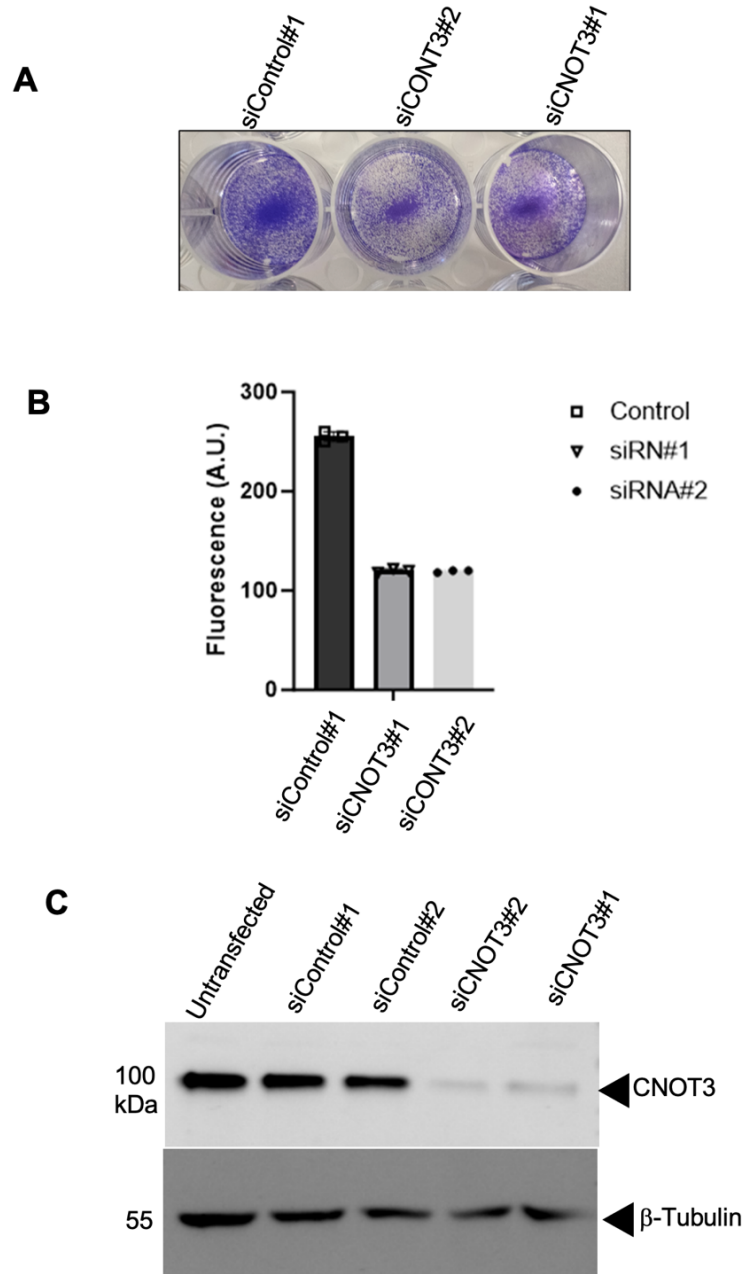


Figure 4.2 CNOT3 knockdown decreases cell proliferation in the HEK293 cell line.

(A) Knockdown of CNOT3 results in decreased in cell viability. Cells were transfected with siRNA#1 and #2 and control siRNA. Five days after transfection, cells were stained with Crystal violet. (B) Decreased metabolic activity as a result of knock down of CNOT3 with siRNA#1 and #2 but not in control siRNA. (C) Western blotting following siRNA-mediated knockdown of CNOT3 using siRNA#1, siRNA#2, sicontrols, and untransfected cells. Error bars represent the standard error of the mean of three independent experiments (n=3).

4.4 CNOT3 knockdown in HEK293 affects Ccr4-Not complex formation

In the context of previous findings, it has been well-documented that the depletion of CNOT1 and CNOT2 disrupts the integrity of the Ccr4-Not complex, impacting various cellular processes (Ito, Inoue, et al. 2011; Ito, Takahashi, et al. 2011). It has also been demonstrated that depleting CNOT3 affects the integrity of the Ccr4-Not complex by reducing the expression level of other components, such as CNOT1, CNOT2, and CNOT7, in mouse embryonic fibroblasts (Suzuki et al. 2015).

However, the current study was designed to explore the effects of CNOT3 knockdown on the expression of subunits within the Ccr4-Not complex in HEK293 cells. The experimental procedure involved transfecting a specified number of cells with siRNA#1 and #2 to silence the expression of CNOT3 and determine the effect of CNOT3 knockdown on other subunits of the Ccr4-Not complex.

The aim of this experiment was to determine the effect of CNOT3 knockdown on the protein levels of other components of the Ccr4-Not complex, specifically CNOT1, CNOT2, and CNOT7. To achieve this, cells were transfected with two different siRNAs targeting CNOT3 (siCNOT3#1 and siCNOT3#2), and a non-targeting siRNA (siControl) was used as a control. Protein expression was then assessed by Western blotting.

As shown in **Figure 4.3**, CNOT3 knockdown was confirmed by the reduction in CNOT3 protein levels in both siCNOT3#1 and siCNOT3#2 lanes compared to the control lane. Knockdown of CNOT3 also resulted in a decrease in CNOT1, CNOT2, and CNOT7 protein levels. In contrast, the control (siControl) maintained normal expression levels of all components. β -Tubulin was used as a loading control, showing consistent protein loading across all samples.

These findings suggest that CNOT3 depletion leads to a downregulation of other subunits within the Ccr4-Not complex, highlighting the critical role of CNOT3 in maintaining the stability of this complex.

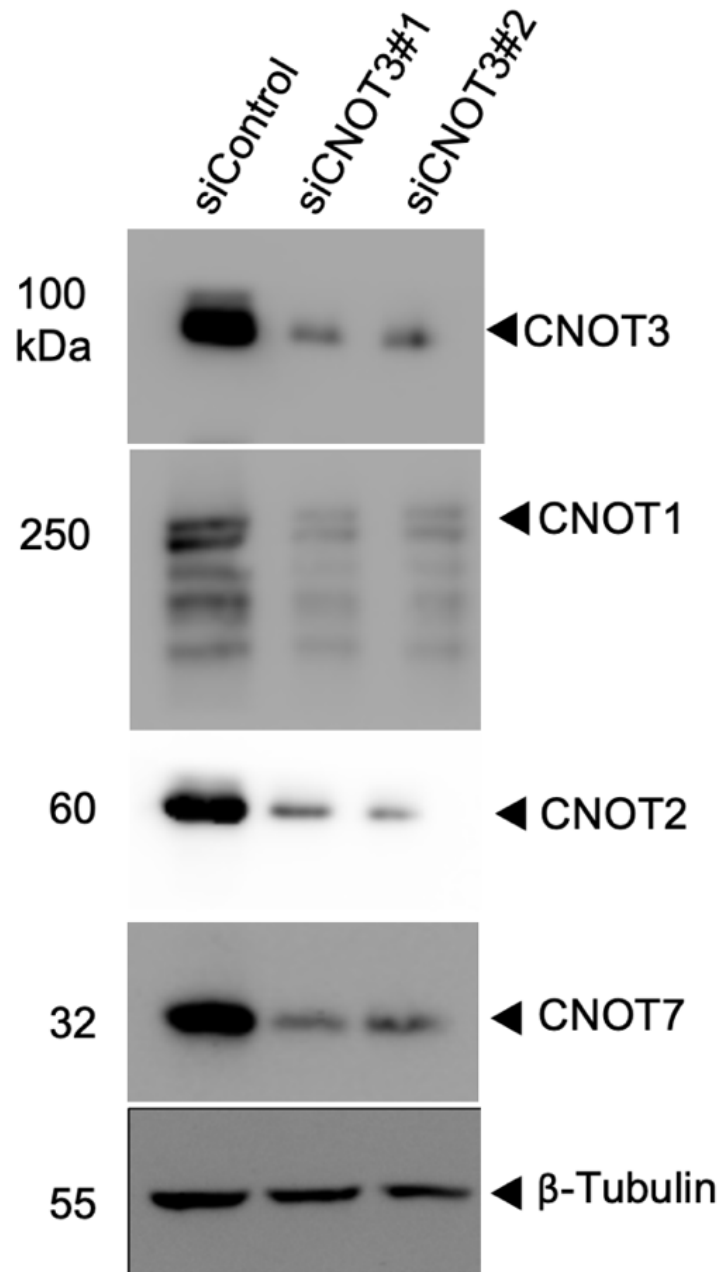


Figure 4.3 CNOT3 knockdown in HEK293 affects Ccr4-Not complex formation.

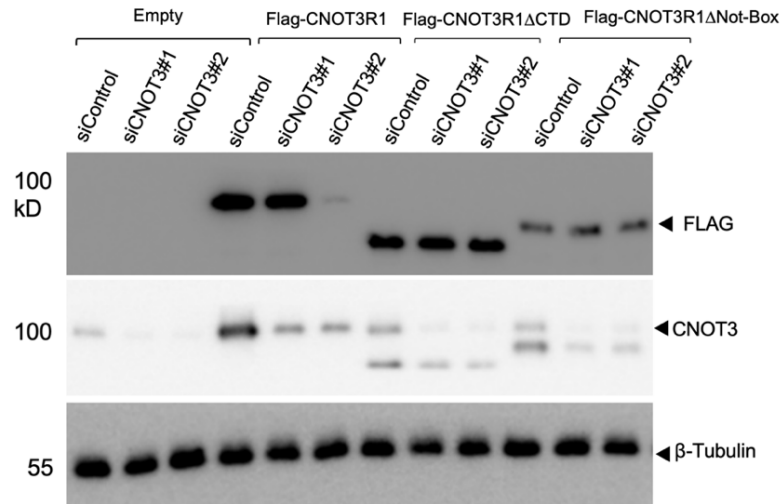
Cells were transfected with siRNAs#1 siRNA#2, and control siRNA. After transfection (72 h), cell lysates were prepared and subjected to Western blotting. Antibodies used were: CNOT1 monoclonal antibody (Cell Signalling), CNOT2 monoclonal antibody, (Cell Signalling) and CNOT7 monoclonal antibody, (Cell Signalling). β -Tubulin was used as a loading control.

4.5 The C-terminal domain of CNOT3 is required for cell proliferation

The results from previous experiments demonstrated the significance of CNOT3 in maintaining the integrity of the Ccr4-Not complex. Notably, a substantial decrease in the proliferation of HEK293 cell lines was observed following the knockdown of CNOT3. Building on these observations, the specific regions of CNOT3 crucial for its role in cell proliferation were further delineated. This investigation focused on the use of stable cell lines expressing different variants of Flag-CNOT3, including Flag-CNOT3R1, Flag-CNOT3R1 Δ CTD, and Flag-CNOT3R1 Δ Not-Box.

As previously discussed, the expression of Flag-CNOT3R1 Δ CTD and Flag-CNOT3R1 Δ Not-Box was achieved by using mutant mRNAs that conferred resistance to siRNA knockdown. Consequently, stable cell lines expressing Flag-CNOT3R1 Δ CTD and Flag-CNOT3R1 Δ Not-Box harboured point mutations in the cDNA encoding the N-terminal domain, and a deletion of the cDNA encoding the C-terminal domain. The resistance of Flag-CNOT3R1 and Flag-CNOT3R2 against siRNAs targeting different regions of CNOT3 is shown in Chapter 3. To confirm that the expression of Flag-CNOT3R1, Flag-CNOT3R1 Δ CTD, and Flag-CNOT3R1 Δ Not-Box was not affected by transfection of siRNA#1, western blot analysis was carried out (**Figure 4.4**). As expected, only Flag-CNOT3R1 expression was sensitive to siRNA#2. By contrast, expression of Flag-CNOT3R1 Δ CTD, and Flag-CNOT3R1 Δ Not-Box was not affected by transfection of siRNA#2, because the region targeted by siRNA#2 was not present in the mRNAs encoding these variants. Moreover, endogenous CNOT3 was effectively knocked down in all cell lines using siRNA#1 and siRNA#2, whereas no knockdown was observed in cells transfected with non-targeting siControl. Cells carrying an empty expression cassette served as a control, and the presence of the Flag peptide was undetected due to its size.

A



B

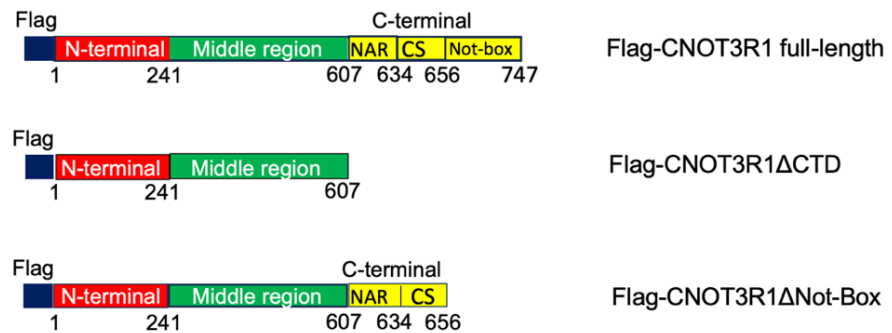


Figure 4.4 Expression of Flag-CNOT3R1, Flag-CNOT3R1ΔCTD, and Flag-CNOT3R1ΔNot-Box are resistant to knockdown by siRNA#1.

(A) Stable cell lines expressing Flag-CNOT3R1 (full-length, siRNA#1-resistant), Flag-CNOT3R1ΔCTD (lacking the C-terminal domain) and Flag-CNOT3R1ΔNot-Box (lacking the Not-Box domain) were cultured in medium containing doxycycline for 48 hours before transfection with siRNAs targeting CNOT3 (siRNA#1, siRNA#2) or a non-targeting control siRNA. Lysates were prepared 72 hours after transfection and subjected to western blotting. Flag-CNOT3 was detected using an anti-Flag rabbit antibody (Sigma-Aldrich). CNOT3 was detected using mouse monoclonal antibody 4B8 (Abnova). β -tubulin served as a loading control.

(B) Domain organisation of human CNOT3 constructs used in Figure 4.4 A. The full-length Flag-CNOT3R1 construct contains the N-terminal conserved region (red), middle region (green), and the C-terminal Not-Box domain (all in yellow). Flag-CNOT3R1ΔCTD lacks the C-terminal region including the NAR, CS, and Not-Box domains. Flag-CNOT3R1ΔNot-Box retains the N-terminal and middle regions but lacks the Not-Box domain.

Chapter 4. Essential domains of CNOT3 for cell proliferation

Then, the question was to determine which domain is essential for the proliferation rate of HEK293 cells in these cell lines. To address this, the set of stable cell lines was cultured in the presence of doxycycline to induce expression of Flag-CNOT3 variants. Subsequently, 48 hours post-induction, the cell lines were transfected with siRNA#1 and siRNA#2. After a further incubation for 72 h, cells were analysed using the AlamarBlue assay. The data revealed that the overexpression of wild-type Flag-CNOT3 and the empty cell lines failed to complement the decrease in the proliferation rate caused by transfection of siRNA#1 and siRNA#2, indicating successful knockdown of endogenous CNOT3 and Flag-CNOT3 (**Figure 4.5**). In contrast, the siRNA#1-resistant variant Flag-CNOT3R1 rescued reduced proliferation rate comparable to that observed upon transfection of non-targeting control siRNA, while the siRNA#2-sensitive Flag-CNOT3R1 could not. This shows the proliferation defect is specific to the knockdown of CNOT3 and not due to off-target effects.

Furthermore, the overexpression of stable cell lines expressing Flag-CNOT3R1 Δ CTD and Flag-CNOT3R1 Δ Not-Box did not rescue the proliferation defect after knockdown of endogenous CNOT3, showing the requirement of the C-terminal and the Not-Box regions for proliferation of HEK293 cells (**Figure 4.5**).

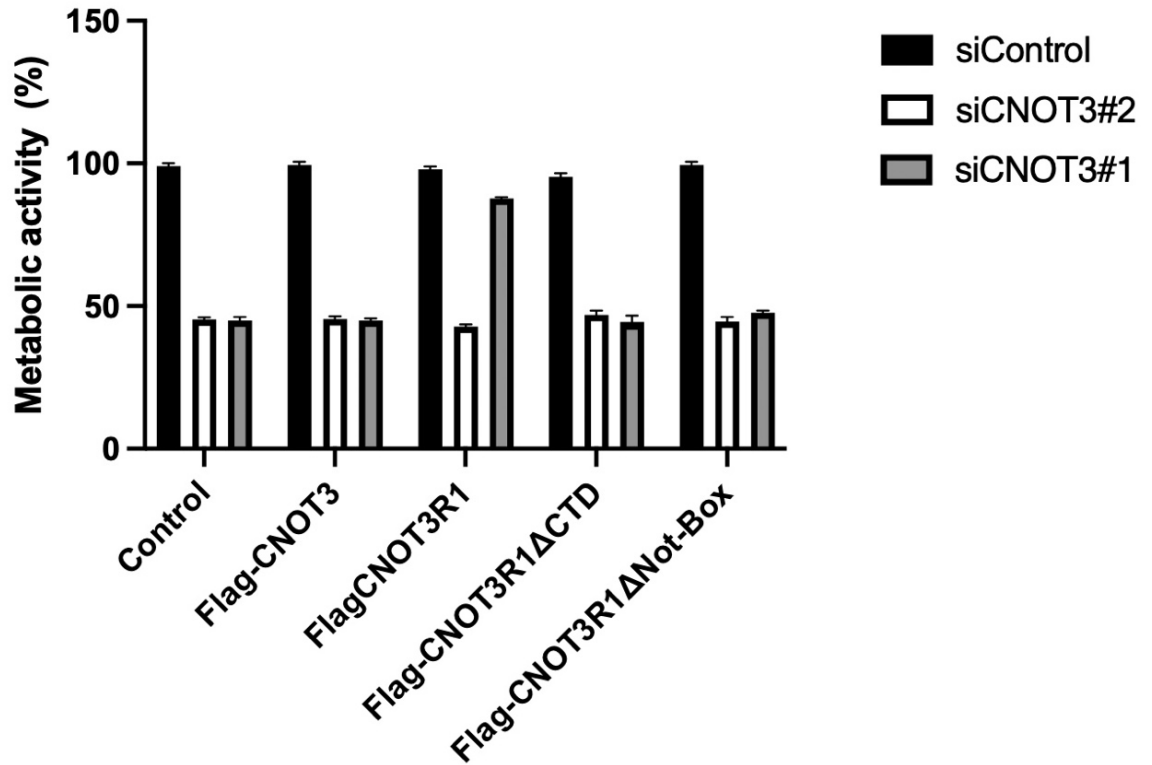


Figure 4.5 The C-terminal domain of CNOT3 is important for cell proliferation.

Stable cell lines containing inducible Flag-CNOT3, Flag-CNOT3R1, Flag-CNOT3R1ΔCTD, and Flag-CNOT3R1ΔNot-Box expression cassettes were cultured in the presence of doxycycline for 48 hours. Cells were then transfected with siRNAs targeting CNOT3 (siRNA#1 and #2) or a non-targeting control siRNA. After 72 hours, cell proliferation was assessed using the AlamarBlue assay. Error bars represent the standard error of the mean of three independent experiments.

4.6 The N-terminal domain of CNOT3 is required for cell proliferation

After defining the role of the C-terminal domain in proliferation, the significance of the N-terminal domain of CNOT3 was investigated. To explore this, stable cell lines expressing the Flag-CNOT3R2 and Flag-CNOT3R2 Δ NTD variants were studied for their resistance to siRNA-mediated knockdown.

First, the resistance of cell lines against siRNAs #1 and #2 was assessed. To confer resistance to siRNA knockdown, Flag-CNOT3R2 Δ NTD was engineered using the Flag-CNOT3R1 construct. Consequently, the stable cell lines expressing Flag-CNOT3R2 Δ NTD mRNA harboured specific point mutations in the sequence encoding the C-terminal domain in addition to a deletion of the sequences encoding the N-terminal domain. The resistance of these stable cell lines, along with those expressing Flag-CNOT3R2 shown in Chapter 3, and the control was verified through transfection with siRNAs targeting two different regions of CNOT3. After siRNA transfection of cell lines containing inducible expression cassettes containing Flag-CNOT3R2, Flag-CNOT3 Δ NTD, or an empty expression cassette, cell lysates were prepared and subjected to western blotting. The western blot analysis (**Figure 4.6**) demonstrated that the expression of Flag-CNOT3R2 was resistant against siRNA#2 but sensitivity to siRNA#1. Expression of Flag-CNOT3R2 Δ NTD was also resistant to transfection with siRNA#2. In addition, however, expression of Flag-CNOT3R2 Δ NTD was also unaffected by siRNA#1, because the mRNA region targeted by siRNA#1 was deleted. Moreover, endogenous CNOT3 was effectively knocked down in all cell lines using siRNA#1 and siRNA#2, whereas no knockdown was observed in cells transfected with non-targeting siControl. Next, after confirming that expression of Flag-CNOT3R2 and Flag-CNOT3R2 Δ NTD was resistant to siRNA#2, the role of the N-terminal domain of CNOT3 in cell proliferation was investigated. To determine the role of the N-terminal domain, stable cell lines expressing the empty pcDNA5-FRT-TetO, Flag-CNOT3, Flag-CNOT3R2, and Flag-CNOT3R2 Δ NTD were seeded and cultured in a medium containing doxycycline for induction of expression. Subsequently, 48 hours post-induction, the cell lines were transfected with siRNA#1 and siRNA#2. After a further incubation of 72 h, cells were subjected to analysis using the AlamarBlue assay.

Chapter 4. Essential domains of CNOT3 for cell proliferation

The AlamarBlue assay data revealed that the overexpression of the siRNA#2-resistant Flag-CNOT3R2 was able to complement the reduced proliferation rate as a result of endogenous CNOT3 knockdown. This was shown to be specific because it was not observed upon expression of wild-type Flag-CNOT3 or after transfection with siRNA#1 (**Figure 4.7**). By contrast, overexpression of Flag-CNOT3R2 Δ NTD did not rescue the defect in proliferation after knockdown of endogenous CNOT3. This implies the importance of the N-terminal domain for the proliferation rate of HEK293 cells.

Together, the data presented in this chapter demonstrates that efficient proliferation of human HEK293 cells depends on both the N-terminal domain of CNOT3 and the C-terminal domain containing the Not-Box region.

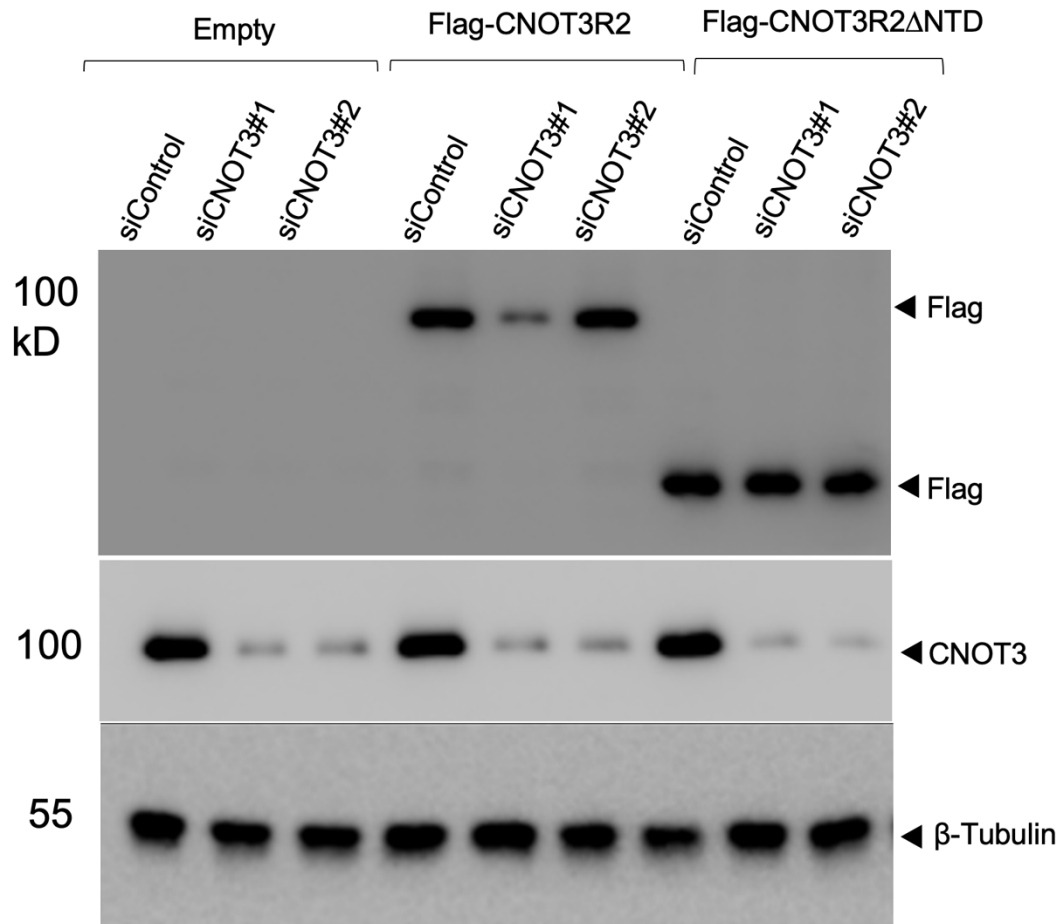


Figure 4.6 Stable HEK293 cells expressing Flag-CNOT3R2 and Flag-CNOT3R2ΔNTD are resistant against knockdown by siRNA#2.

HEK293 cell lines containing the indicated inducible Flag-CNOT3 expression cassettes were cultured in medium containing doxycycline for 48 hours before transfection with siRNAs targeting endogenous CNOT3 (siRNA#1 and siRNA#2) or a non-targeting control siRNA. Cell lysates were prepared 72 hours after transfection and subjected to western blotting. Expression of Flag-CNOT3R2 and Flag-CNOT3R2ΔNTD was detected using an anti-Flag rabbit antibody (Sigma-Aldrich). Endogenous CNOT3 was detected using mouse monoclonal antibody 4B8 (Abnova). β-tubulin served as a loading control.

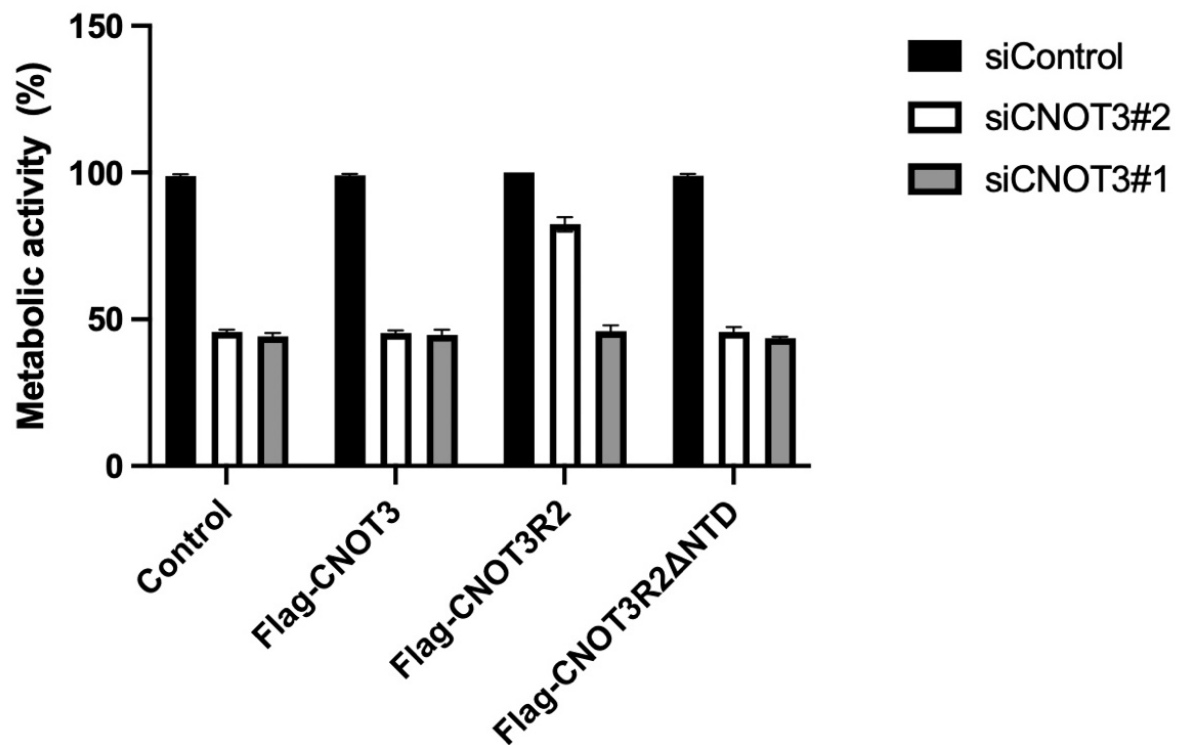


Figure 4.7 *The N-terminal domain of CNOT3 is important for efficient cell proliferation.*

Stable HEK293 cell lines carrying inducible expression cassettes containing Flag-CNOT3, Flag-CNOT3R2, and Flag-CNOT3R2ΔNTD were cultured in doxycycline for 48 hours. Subsequently, the cells were transfected with siRNA#1 and #2 or non-targeting control siRNA. After 72 hours, the AlamarBlue assay was performed.

4.7 Discussion

The findings presented in this chapter advance our understanding of the role of CNOT3 in cell proliferation, highlighting the intricate mechanisms through which this protein influences cellular growth. The experimental strategies employed, using doxycycline-induced cell lines and siRNA-mediated knockdown, facilitated a detailed investigation into the specific contributions of the N- and C- terminal regions of CNOT3. Also, the

Chapter 4. Essential domains of CNOT3 for cell proliferation

toolkit of stable cell lines described in the previous chapter were used to discover which regions of CNOT3 are essential for efficient proliferation in HEK293 cells.

In the first step, the lack of significant impact on cell proliferation following the induction of expression of the Flag-CNOT3 variants validates the experimental approach, indicating that overexpression of Flag-CNOT3 did not affect cell proliferation. This validation was crucial, as it ensured that subsequent observations and conclusions regarding the role of CNOT3 were not due to overexpression.

Following CNOT3 knockdown in HEK293 cells, a significant reduction in cell proliferation was observed, highlighting the important role of CNOT3 in cellular growth. This finding is consistent with previous studies that have implicated CNOT3 in cell viability (Jing et al. 2019; Cejas et al. 2017; Aslam et al. 2009; Ghashghaei et al. 2024) . However, the function contrasts with the previously characterised role of CNOT3 as a tumour suppressor in T-cell acute lymphoblastic leukaemia (De Keersmaecker et al. 2013). The observed reduction in proliferation suggests that CNOT3 is not just a structural component of the Ccr4-Not complex but is actively involved in critical cellular processes governing proliferation and might be regarded as a driver for oncogenesis.

CNOT3 is a conserved component of the Ccr4-Not complex, establishing stable connections with the scaffolding subunit CNOT1 and subunit CNOT2 (Boland et al. 2013). The disruption of the Ccr4-Not complex integrity is evidenced by altered expression levels of CNOT1, CNOT2, and CNOT7 following CNOT3 knockdown, highlighting the interdependence of these subunits for the stability of the complex. The data revealed from this experiment was in line with the previous study, which showed the knockdown of CNOT3 effects on the expression level of other subunits of the Ccr4-Not complex in mouse embryonic fibroblast (Suzuki et al. 2015; Ghashghaei et al. 2024). One possible explanation is that, in the absence of CNOT3, the unassembled subunits become destabilised and are subject to degradation by proteolytic pathways, or that CNOT3 knockdown leads to transcriptional downregulation of other subunits. Either mechanism would result in lower steady-state levels of the affected components. The disruption of the complex also suggests that the effects on cell proliferation observed

Chapter 4. Essential domains of CNOT3 for cell proliferation

after CNOT3 knockdown may be mediated by broader impairments in the complex's regulatory functions, such as mRNA decay and translational control.

The development of stable cell lines expressing various siRNA-resistant CNOT3 mutants provided further insights into the functional roles of the amino and carboxylic acid terminal domains. The resistance to siRNA knockdown conferred by specific synonymous mutations allowed for the delineation of the contributions of these regions to role of CNOT3 in cell proliferation (Aslam et al. 2009). The function of each domain of CNOT3 was investigated by knocking down the endogenous CNOT3 and simultaneously overexpressing the exogenous Flag-tagged variants of CNOT3 to determine the function of each domain. It was previously shown (Suzuki et al. 2015) that the N and C termini are important for cell proliferation in mouse embryonic fibroblast cells. The results described in this chapter focused on the role of each region of CNOT3 in the human HEK293 cell line. The AlamarBlue assay, a widely used method to assess cell viability and proliferation, was employed to evaluate the impact of CNOT3 knockdown on cell proliferation. Following CNOT3 knockdown in cell lines expressing CNOT3 variants, it was shown that the Flag-CNOT3R1 variant could complement the depletion of CNOT3, maintaining normal proliferation rates. In contrast, cell lines expressing CNOT3 lacking the C-terminal (CTD) and Not-Box regions could not complement the depletion of endogenous CNOT3, resulting in reduced proliferation rates. This suggests that the C-terminal and Not-Box regions are crucial for the role of CNOT3 in promoting efficient cell proliferation. Mutations in the C-terminal (C) and Not-Box regions of CNOT3 significantly disrupt its function by impairing critical protein-protein interactions, particularly with CNOT1 and CNOT2, which are essential for the structural and functional integrity of the Ccr4-Not complex (Boland et al. 2013). These domains play a pivotal role in anchoring CNOT3 within the complex and facilitating its interactions with other subunits that are crucial for mRNA degradation. The deletion of these domains hinders CNOT3's ability to effectively regulate the stability and degradation of mRNA, a process critical for maintaining cellular homeostasis. As a result, the loss of function in these regions compromises the complex's ability to support efficient cell proliferation. The impaired ability to regulate key cellular transcripts can lead to dysregulated growth signals and failure to maintain normal cellular functions, highlighting the importance of these

Chapter 4. Essential domains of CNOT3 for cell proliferation

domains in both the structural integrity of CNOT3 (Boland et al. 2013) and its role in cellular proliferation (Suzuki et al. 2015). As a result, they are unable to compensate for the knockdown of endogenous CNOT3, which is necessary for promoting cell proliferation (Suzuki et al. 2015). These findings are consistent with the research by (Suzuki et al. 2015) in mouse embryonic fibroblast models, which demonstrated the importance of these regions for CNOT3's function in cellular growth.

Similarly, the AlamarBlue assay was used to investigate the effects of CNOT3 knockdown in cell lines expressing the N-terminal deletion (NTD) alongside the Flag-CNOT3R2 variant and a control. The results indicated that the Flag-CNOT3R2 variant could rescue the proliferation defect caused by CNOT3 knockdown, whereas the Flag-CNOT3R2ΔNTD variant could not. The inability of Flag-CNOT3R1ΔNTD to rescue proliferation defects, suggests multiple crucial regulatory roles for the N-terminal domain. Our findings, supported by previous study (Suzuki et al. 2015), suggests that the N-terminal coiled-coil domain may serve as a critical platform for target recognition, potentially through direct RNA interactions or by recruiting additional regulatory factors that guide the complex to specific mRNAs. Moreover, the N-terminal domain likely plays a vital role in ribosome association, suggesting that its absence could disrupt the complex's ability to coordinate translation regulation (Absmeier et al. 2023; Zhu et al. 2024) with mRNA decay, thereby affecting both the stability and translation efficiency of its target transcripts.

This multifaceted hypothesis is particularly compelling given the discovery of N-terminal domain mutations in T-cell acute lymphoblastic leukaemia, where disrupted proliferation control is a hallmark feature (De Keersmaecker et al. 2013). This comprehensive regulatory failure could explain the inability to rescue proliferation defects in CNOT3-depleted cells.

Overall, this chapter analyses the functional significance of the N-terminal and C-terminal domains of CNOT3 in cell proliferation. The data presented support the hypothesis that the structural integrity of CNOT3 is crucial for its role in the Ccr4-Not complex and for the proliferation of HEK293. These insights contribute to a deeper understanding of the molecular mechanisms underlying cellular growth.

Chapter 5.
Characterising disease-associated mutations
in the genes encoding the CNOT3 and CNOT2
subunits of the Ccr4-Not Complex

5.1 Introduction

Neurodevelopmental disorders (NDDs) pose a significant challenge, affecting a portion of the population and often leading to lifelong difficulties (Niceta et al. 2023; Martin et al. 2019). These disorders encompass a wide range of conditions, including intellectual disability and autism spectrum disorder, which often lead to lifelong difficulties in learning, communication, and social interaction (Martin et al. 2019). Recent research has highlighted the role of genetic factors in these disorders. For example, *de novo* variants in *CNOT3* identified by the Deciphering Developmental Disorders (DDD) study have provided strong evidence for the involvement of this gene in developmental disorders (Martin et al. 2019).

Interestingly, the Ccr4-Not complex, a key regulator of gene expression in which *CNOT3* is a core component, has been implicated in both neurodevelopmental disorders and T-cell acute lymphoblastic leukaemia (T-ALL) (Vicente et al. 2018; De Keersmaecker et al. 2013). In NDDs, mutations in *CNOT3*—such as c.142C>G (p. Leu48Val), c.355A>G (p. Lys119Glu), c.439G>A (p. Glu147Lys), c.562C>T (p. Arg188Cys), and c.563G>A (p. Arg188His)—have been associated with intellectual disability, developmental delay, and behavioural abnormalities. Patients with these mutations often exhibit distinctive facial features, including a thin upper lip, low-set ears, and micrognathia, aiding in clinical diagnosis. Additional features may include hypotonia, seizures, and other neurological symptoms, underscoring the broad impact of *CNOT3* dysfunction in human development (Martin et al. 2019).

In T-ALL, specific mutations in *CNOT3*—namely, c.169C>T (p. Arg57Trp) and c.170G>A (p. Arg57Gln)—have been identified as significant contributors to disease pathogenesis. These mutations are believed to impair the regulatory functions of the Ccr4-Not complex, contributing to leukemogenesis (De Keersmaecker et al. 2013).

The interaction between *CNOT3* and *CNOT2* is crucial for the stability and function of the complex. Mutations in *CNOT2* have also been identified in patients with neurodevelopmental disorders. The clinical manifestations of these *CNOT2* mutations include nasal speech, dysmorphic facial features, and variable skeletal anomalies. Shared phenotypic features among affected individuals include upslanted palpebral fissures, anteverted nares, a thin upper lip, low-set ears, and micrognathia. Additionally, variable

skeletal abnormalities such as clinodactyly and short fifth fingers have been observed. The study provided further evidence supporting *CNOT2* as a gene associated with NDDs. The findings emphasise the role of *CNOT2* in contributing to the clinical features observed in patients with the 12q15 microdeletion syndrome and highlight its relevance as a novel gene involved in human NDDs (Uehara et al. 2019; Niceta et al. 2023).

Coupled with truncating mutations (described in chapter 3), these findings suggest that mutations in *CNOT3* and *CNOT2* may disrupt their interaction, thereby impairing the functionality of the Ccr4-Not complex. Previous work in mouse embryonic fibroblasts has demonstrated the importance of the C-terminal domain of *CNOT3* for its interaction with *CNOT2*, highlighting the need to further investigate this interaction in other systems (Suzuki et al. 2015).

This chapter is aimed at investigating the molecular consequences of mutations in *CNOT3* and *CNOT2*, with a focus on their impact on the interaction between human *CNOT3* and *CNOT2*. It is hypothesised that the binding interface between *CNOT3* and *CNOT2* will be disrupted by these mutations, leading to the destabilisation of the Ccr4-Not complex. To test this hypothesis, computational predictions and experimental validation were combined. Missense mutations identified in patients with T-ALL and neurodevelopmental disorders were introduced into human *CNOT3* and *CNOT2* using site-directed mutagenesis. In silico prediction tools (SIFT, PolyPhen-2, and AlphaMissense) were utilised to assess the potential effects of these mutations on protein structure and function. Finally, co-immunoprecipitation (co-IP) assays were employed to directly measure the impact of these mutations on the *CNOT3-CNOT2* interaction in mammalian cells.

5.2 Importance of C-terminal domain of *CNOT3* for incorporation of *CNOT3* into the Ccr4-Not complex

To confirm the requirement for the N-terminal domain of *CNOT3* for incorporation into the Ccr4-Not complex in the context of mammalian cells, co-immunoprecipitation (co-IP) analysis was conducted. Plasmids encoding Flag-*CNOT3*, Flag-*CNOT3*ΔNTD, Flag-*CNOT3*R1ΔCTD, and Flag-*CNOT3*R1ΔNot-Box were transiently transfected into HEK293

cells along with an empty vector control. After 48 hours, cell lysates were collected, and co-IP was carried out.

Western blot analysis of the total cell lysate was carried out as a preliminary step before co-immunoprecipitation to ascertain the equal expression of various CNOT3 variants. The results from the input analysis revealed similar expression levels of all variants in HEK293 cells. As anticipated, no expression was detected in the cell line transfected with the empty vector, indicating the specificity of the assay **(Figure 5.1 A)**.

The co-immunoprecipitation of both Wildtype Flag-CNOT3 and Flag-CNOT3R2ΔNTD successfully demonstrated precipitation of CNOT1, CNOT2, and CNOT7 using anti-flag beads. This confirms that the N-terminal domain is not required for the interaction of CNOT3 with CNOT1, CNOT2, and CNOT7. In contrast, the co-immunoprecipitation of Flag-CNOT3R1ΔCTD and Flag-CNOT3R1ΔNot-Box with anti-flag beads did not yield any precipitation of CNOT1, CNOT2 and CNOT7, indicating that the C-terminal and Not-Box region play a crucial role in incorporating CNOT3 into the Ccr4-Not complex. This is consistent with the notion that the Not-Box of CNOT3 is required for the integrity of the Not-module (Bawankar et al. 2013; Buschauer et al. 2020). Additionally, as expected, no precipitation was observed in the empty plasmid control, thus affirming the specificity of the interactions observed **(Figure 5.1 B)**.

These findings confirm the indispensable role of the C-terminal and Not-Box region of CNOT3 in mediating these interactions in HEK293 cells.

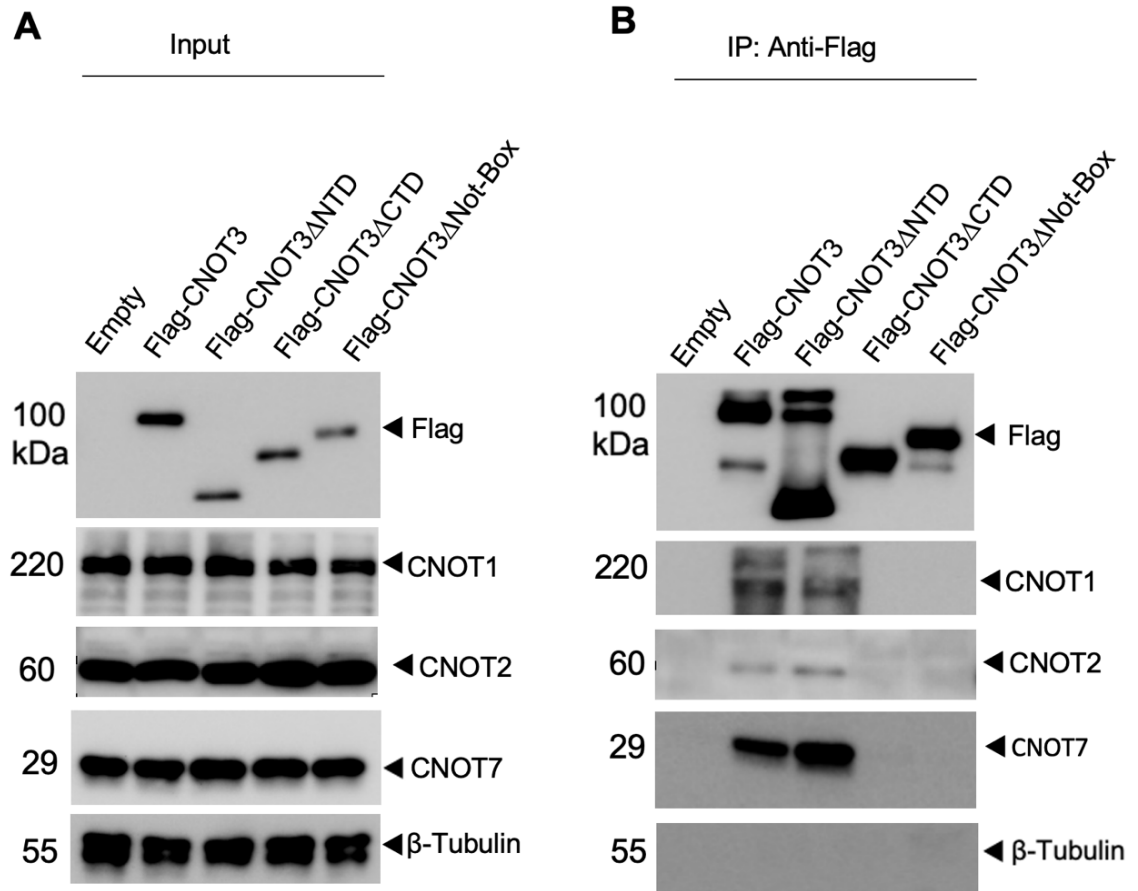


Figure 5.1 The C-terminal domain of CNOT3 is required for incorporation into the Ccr4-Not complex.

(A) Similar expression of Flag-CNOT3 variants. Cell lysates were separated by 10% SDS-PAGE and subjected to western blotting. (B) Immunoprecipitation of Flag-CNOT3 variants. Proteins were identified using antibodies recognising the Flag epitope, CNOT1 (monoclonal antibody, Cell Signaling), CNOT2 (monoclonal antibody, Cell Signaling), and CNOT7 (monoclonal antibody, Cell Signaling). β -tubulin was used as a loading control.

5.3 Literature-based identification of CNOT3 mutations associated with T-ALL and the rare neurodevelopmental disorder IDDSADF

Mutations in CNOT3 (**Table 5.1**), those located within the N-terminal domain, have been implicated in both T-cell acute lymphoblastic leukaemia (T-ALL) and a rare neurodevelopmental disorder known as Intellectual Disability with Dysmorphic Syndrome and Autistic Features (IDDSADF) (Martin et al. 2019; De Keersmaecker et al. 2013) (**Figure 5.2**).

A study by Martin et al. analysed 16 probands with de novo variants in CNOT3, revealing that these mutations in the N-terminal domain (**Figure 5.2 A and B**) are strongly associated with neurodevelopmental disorders, specifically IDDSADF. This rare disorder is characterised by intellectual disability, distinctive facial dysmorphisms, and autistic behaviours. Notable de novo mutations identified in this study include c.142C>G (p. Leu48Val), c.355A>G (p. Lys119Glu), c.439G>A (p. Glu147Lys), c.562C>T (p. Arg188Cys), and c.563G>A (p. Arg188His). These mutations have been associated with clinical manifestations such as hypotonia, intellectual disability, and behavioural problems, though no specific phenotype was consistently observed across all cases.

In T-cell acute lymphoblastic leukaemia (T-ALL), CNOT3 mutations are significant contributors to the disease's pathogenesis. (De Keersmaecker et al. 2013) identified two mutations within the N-terminal domain of CNOT3, including c.169C>T (p. Arg57Trp) and c.170G>A (p. Arg57Gln).

Table 5.1 CNOT3 mutations associated with both T-ALL and neurodevelopmental disorders

Variant	Effect	Type of variant	Disease	Reference
c.169C>T	p. (Arg57Trp)	Missense	T-ALL	A
c.170G>A	p. (Arg57Gln)	Missense	T-ALL	A
c.142C>G	p. (Leu48Val)	Missense	IDDSADF	B
c.355A>G	p. (Lys119Glu)	Missense	IDDSADF	B
c.439G>A	p. (Glu147Lys)	Missense	IDDSADF	B
c.562C>T	p. (Arg188Cys)	Missense	IDDSADF	B
c.563G>A	p. (Arg188His)	Missense	IDDSADF	B

A (De Keersmaecker, 2013 #247), B (Martin et al. 2019).

5.4 The analysis of CNOT3 variants found in T-ALL and neurodevelopmental disorders using predictive algorithms

To investigate the potential impact of CNOT3 variants identified in T-ALL and neurodevelopmental disorders, ChimeraX was initially used to visualize the protein structure and map the positions of the mutations (**Figure 5.2**). This provided a structural context for understanding their potential effects. Following this, predictive algorithms including SIFT, AlphaMissense, and PolyPhen-2 were employed to evaluate whether these coding missense variants could lead to functional changes in the protein, offering insights into their possible roles in disease pathogenesis.

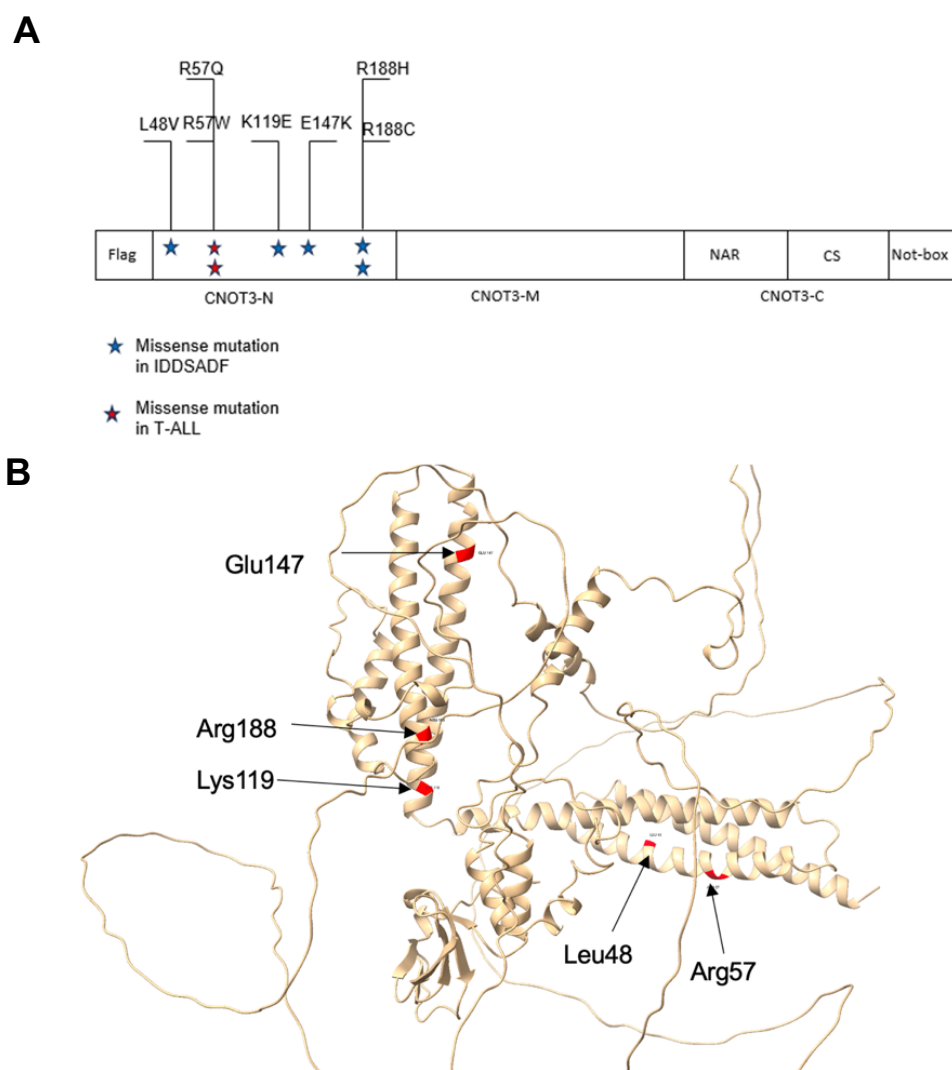


Figure 5.2 Distribution of missense mutations in CNOT3 observed in T-cell acute lymphoblastic leukaemia (T-ALL) and IDDSADF within the CNOT3-N.

(A) Schematic Representation of CNOT3 protein domains and missense mutations. The diagram illustrates the domain structure of the CNOT3 protein, highlighting the positions of identified missense mutations. The protein is divided into its major domains: CNOT3-N (N-terminal), CNOT3-M (Middle), CNOT3-C (C-terminal), NAR (Not1 Associated Region), CS (Conserved Sequence), and Not-Box. The mutations R57W and R57Q in T-cell acute lymphoblastic leukaemia (T-ALL), along with L48V K119E, E147K, R188C, and R188H, are highlighted with red and blue stars, respectively, indicating their specific locations within the CNOT3-N region. (B) AlphaFold-predicted model of CNOT3 (AF-O75175-F1, UniProt ID: O75175). This 3D model of the CNOT3 protein was visualised using ChimeraX software (version 1.8) (Meng et al. 2023). The protein backbone is displayed in tan, with alpha-helices represented as coiled ribbons and beta-strands as arrows. The positions of key residues where missense mutations have occurred, specifically Leu48, Arg57, Lys119, Glu147, and Arg188, are shown in red.

Chapter 5. Characterising disease-associated mutations in CNOT3 and CNOT2

5.4.1 CNOT3 variants found in T-ALL and neurodevelopmental disorders analysed using Sorting intolerant from tolerant (SIFT)

To assess the potential impact of these mutations, Sorting Intolerant From Tolerant (SIFT), a bioinformatics tool was utilised that predicts the effects of amino acid substitutions on protein function (Kumar, Henikoff, and Ng 2009). SIFT is an algorithm that predicts whether an amino acid substitution in a protein sequence is likely to affect protein function or not. It categorises each variant as "tolerated", or "damaging" based on the degree of conservation of the amino acid position and the physicochemical properties of the amino acids involved. SIFT calculates the probability score of changing the amino acid at the position that appears in the alignment from 0 to 1, with a cut-off value=0.05. The amino acid substitution is predicted damaging if the score ≤ 0.05 and tolerated if the score is >0.05 .

Specific missense mutations in *CNOT3*, associated with T-ALL and neurodevelopmental disorders, are focused on in this analysis, including those reported in the literature: L48V, R57W, R57Q, K119E, E147K, R188C, and R188H. Each of these mutations represents a single amino acid change in the CNOT3 protein sequence (**Table 5.2**).

Table 5.2 List of CNOT3 variants associated with T-ALL and neurodevelopmental disorders (IDDSADF) evaluated with SIFT

Variant	Score	Prediction	Disease
L48V	0.00	Damaging	IDDSADF
R57W	0.00	Damaging	T-ALL
R57Q	0.03	Damaging	T-ALL
K119E	0.08	Tolerated	IDDSADF
E147K	0.05	Tolerated	IDDSADF
R188C	0.00	Damaging	IDDSADF
R188H	0.01	Damaging	IDDSADF

The L48V variant has a SIFT score of 0.00, which strongly suggests that it is highly damaging. This variant is associated with neurodevelopmental disorders (IDDSADF), implying that it may significantly disrupt the normal function of CNOT3, contributing to developmental issues.

Similarly, the R57W variant, with a SIFT score of 0.00, is predicted to be damaging, and it is associated with T-ALL. This strongly suggests that this mutation might play a critical role in leukemogenesis. The R57Q variant, with a SIFT score of 0.03, is also considered damaging and is similarly linked to T-ALL, indicating that alterations at the R57 residue are likely critical for CNOT3's activity and might contribute to cancer development.

The K119E variant has a SIFT score of 0.08, categorized as tolerated, suggesting this mutation may not severely impact CNOT3's functionality. However, SIFT's tolerance prediction may not fully capture the variant's effect, as it could still contribute to neurodevelopmental disorders through subtle mechanisms or in combination with other genetic or environmental factors.

Another variant, E147K, with a SIFT score of 0.05, is also predicted to be tolerated. This mutation is associated with neurodevelopmental disorders. While not predicted to severely impair CNOT3's function, its association with developmental disorders suggests that it could play a role in disease development in combination with other mutations or factors.

The R188C and R188H variants, with SIFT scores of 0.00 and 0.01, respectively, are both predicted to be damaging. These variants are associated with neurodevelopmental disorders. The damaging predictions for both mutations at this residue suggest that alterations in this region of CNOT3 may lead to significant functional impairments, contributing to both cancer and developmental abnormalities.

In summary, the variants with low SIFT scores, particularly L48V, R57W, R57Q, R188C, and R188H, are predicted to be damaging and are associated with serious diseases like T-ALL and neurodevelopmental disorders. These damaging mutations likely represent alterations that affect CNOT3's regulatory functions. Conversely, K119E and E147K are predicted to be tolerated, suggesting a more subtle impact on protein function, though they may still be involved in disease under certain circumstances.

5.4.2 CNOT3 variants found in T-ALL and neurodevelopmental disorders analysed using PolyPhen-2

Following the initial evaluation of CNOT3 variants using SIFT and AlphaMissense, PolyPhen-2 (Polymorphism Phenotyping v2) was employed to further elucidate the potential functional impact of these (Adzhubei, Jordan, and Sunyaev 2013). PolyPhen-2 complements these analyses by incorporating both sequence-based and structure-based features, such as the physical properties of amino acids, their spatial locations within the protein's 3D structure, and the evolutionary conservation of the substitution site. It employs a Naive Bayes classifier trained on large datasets, which helps to distinguish between mutations likely to be "benign," "possibly damaging," or "probably damaging" based on a combination (Bawankar et al. 2013) of biochemical, evolutionary, and structural attributes (Adzhubei, Jordan, and Sunyaev 2013). The Poly-Phen-2 assigns each variant a qualitative descriptor—either "benign," "possibly damaging," or "probably damaging"—along with a quantitative score ranging from 0 (benign) to 1 (damaging), where higher scores indicate a greater probability of a damaging effect. **Table 5.3 and Figure 5.3** below present the PolyPhen-2 results for the CNOT3 variants examined in this study.

Table 5.3 Analysis of CNOT3 variants associated with T-ALL and IDDSADF evaluated with PolyPhen-2

Variant	Score	Prediction	Disease
L48V	0.997	Probably pathogenic	IDDSADF
R57W	0.994	Probably pathogenic	T-ALL
R57Q	0.998	Probably pathogenic	T-ALL
K119E	0.457	possibly pathogenic	IDDSADF
E147K	0.995	Probably pathogenic	IDDSADF
R188C	0.998	Probably pathogenic	IDDSADF
R188H	0.997	Probably pathogenic	IDDSADF

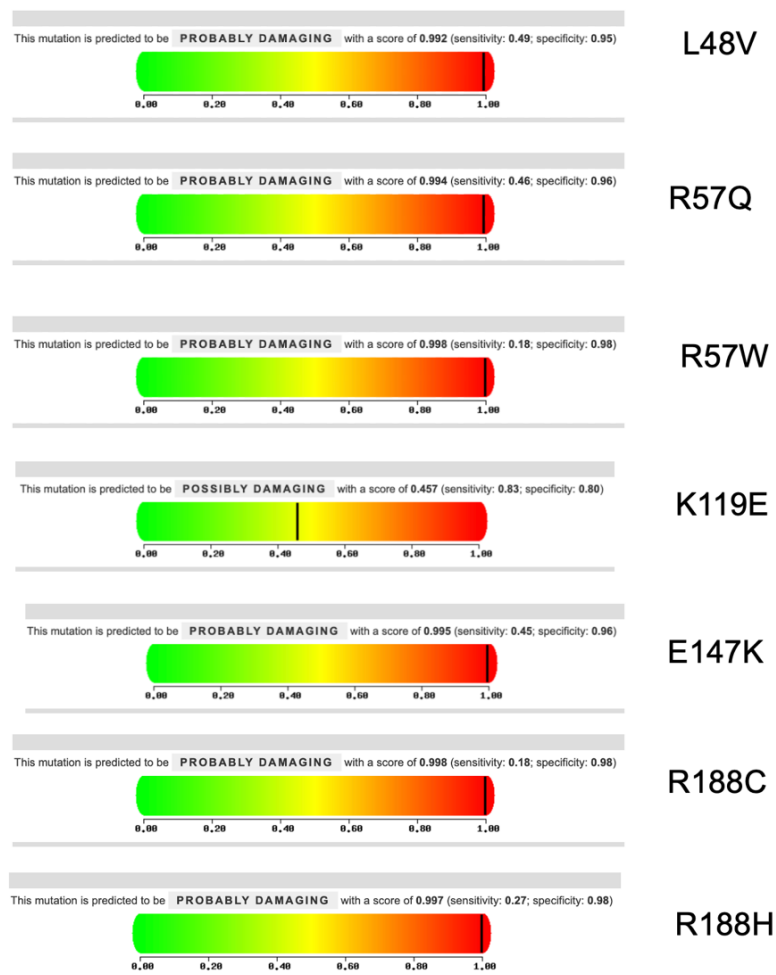


Figure 5.3 PolyPhen-2 Analysis of CNOT3 Missense Mutations.

This figure presents the predicted functional impact of seven missense mutations in the CNOT3 protein as analysed by PolyPhen-2. Each panel represents a specific mutation, showing a colour-coded bar indicating the likelihood of the mutation being damaging. The green to red gradient represents the score range from 0 (benign) to 1 (probably damaging), with the specific score for each mutation marked by a vertical black line. L48V: Predicted to be probably damaging with a score of 0.992 (sensitivity: 0.49, specificity: 0.95). R57Q: Predicted to be probably damaging with a score of 0.994 (sensitivity: 0.46, specificity: 0.96). R57W: Predicted to be probably damaging with a score of 0.998 (sensitivity: 0.18, specificity: 0.98). K119E: Predicted to be possibly damaging with a score of 0.457 (sensitivity: 0.83, specificity: 0.80). E147K: Predicted to be probably damaging with a score of 0.995 (sensitivity: 0.45, specificity: 0.96). R188C: Predicted to be probably damaging with a score of 0.998 (sensitivity: 0.18, specificity: 0.98). R188H: Predicted to be probably damaging with a score of 0.997 (sensitivity: 0.27, specificity: 0.98).

The analysis reveals that six out of the seven CNOT3 variants (L48V, R57Q, R57, E147K, R188C, and R188H) are predicted to be "probably damaging," with scores approaching 1.0. These predictions suggest a high probability that these variants significantly impair the structure or function of the CNOT3 protein. The variants R57 and R188C have the highest scores (0.998), strongly indicating their potential to disrupt normal protein function.

The variant K119E, on the other hand, is classified as "possibly damaging" with a score of 0.457, suggesting a moderate likelihood of a damaging effect. While this score is lower than those of the other variants, it still indicates a possible impact on protein functionality, although with less confidence.

5.4.3 CNOT3 variants found in T-ALL and neurodevelopmental disorders analysed using AlphaMissense

Missense variants are a significant source of genetic variation contributing to human diseases, including T-cell acute lymphoblastic leukaemia (T-ALL) and various neurodevelopmental disorders (IDDSADF). AlphaMissense, a machine learning-based tool that combines structural context, evolutionary conservation, and population frequency data, was utilised to predict structure of CNOT3 wild-type, mutant (**Figure 5.4 A and B**) and the pathogenicity of missense variants. This tool provides a valuable approach to understanding which variants may have a significant clinical impact without relying on potentially biased human-curated data (Cheng et al. 2023).

In this study, AlphaMissense was applied to assess missense variants related to T-ALL and neurodevelopmental disorders in CNOT3, and the results are presented in **Table 5.3**.

Table 5.4 Analysis of CNOT3 variants associated with T-ALL and IDDSADF evaluated with AlphaMissense

Variant	Score	Prediction	Disease
L48V	0.997	likely pathogenic	IDDSADF
R57W	1.0	likely pathogenic	T-ALL
R57Q	1.0	likely pathogenic	T-ALL
K119E	0.97	likely pathogenic	IDDSADF
E147K	1.0	likely pathogenic	IDDSADF
R188C	0.999	likely pathogenic	IDDSADF
R188H	0.997	likely pathogenic	IDDSADF

The AlphaMissense model classified all analysed variants as "Likely Pathogenic," with pathogenicity scores nearing or at 1.0. This classification indicates that these variants have a high likelihood of being involved in disease processes. AlphaMissense has predicted variant pathogenicity by integrating structural information and leveraging patterns in biological data (Cheng et al. 2023; Ljungdahl et al. 2023).

The high pathogenicity scores obtained for the CNOT3 variants suggest that these variants may significantly contribute to the pathogenesis of T-cell acute lymphoblastic leukaemia (T-ALL) and neurodevelopmental disorders (IDDSADF) by disrupting the normal function of the Ccr4-Not complex. Given the Ccr4-Not complex's critical role in regulating mRNA decay and transcriptional control, the disruption caused by these variants could lead to the dysregulation of key genes involved in T-cell development and neuronal function.

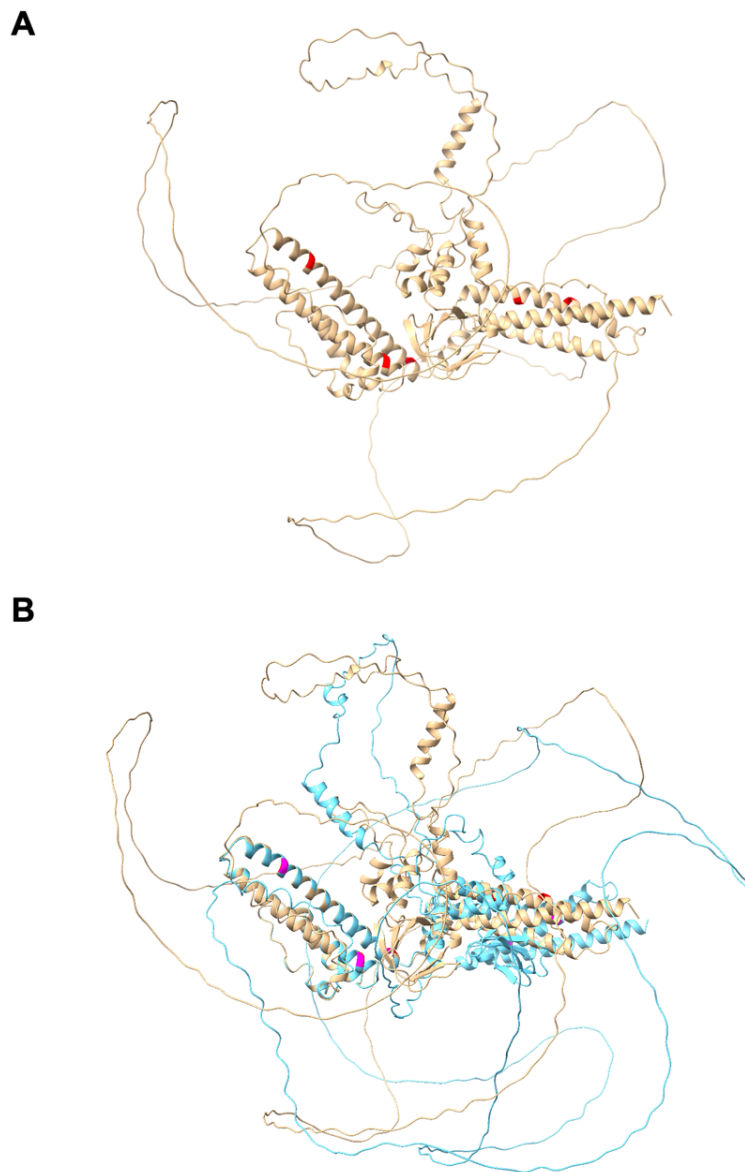


Figure 5.4 AlphaFold predicted 3D Structures of wild-type and mutant CNOT3 Protein.

(A) Predicted structure of wild-type human CNOT3 generated using AlphaFold 3. The model reveals an extended α -helical architecture consistent with its scaffold function within the Ccr4–Not complex. Residues corresponding to the disease-associated point mutations analysed in this study are highlighted in red. These mutations are situated within structured α -helical regions, potentially impacting local conformation or interaction surfaces.

(B) Structural overlay of the AlphaFold 3-predicted wild-type and mutant CNOT3 proteins. The wild-type structure is shown in tan, and the mutant variant is displayed in sky blue. Mutation sites are indicated in red on the wild-type model and in magenta on the mutant model. While the global secondary and tertiary structure is retained in the mutant, subtle local deviations are observed near the mutation sites. These may reflect changes inside

chain positioning, local packing, or flexibility that could impact protein–protein interactions or functional stability.

5.5 Construction of plasmids containing CNOT3 with patient-derived mutations linked to T-ALL and IDDSADF

To investigate the functional consequences of these mutations, site-directed mutagenesis was employed to introduce the specific variants into a plasmid expressing the *CNOT3* gene. The plasmid used in this study is pCMV-Flag-CNOT3, which includes a Flag tag to facilitate downstream protein detection and purification.

The process, illustrated in Figure 5.4, began with the design of specific forward (F) and reverse (R) primers to introduce the desired mutations into the *CNOT3* gene (**Figure 5.5 A**). These primers were constructed to incorporate the nucleotide changes corresponding to each patient-derived mutation. The plasmid was then amplified using PCR with the designed primers, resulting in a linear product (**Figure 5.5 B**) that contained the mutation(s) of interest. The PCR product was designed to maintain 5' phosphate (5' PO₄) and 3' hydroxyl (3' OH) groups to facilitate subsequent ligation.

Following PCR amplification, the product was treated with *Dpn* I, an enzyme that specifically digests methylated DNA. Since the original plasmid DNA template is methylated, *Dpn* I digestion selectively removed the parental DNA, leaving only the newly synthesized, mutated plasmid DNA. The mutated linear DNA was then ligated to re-circularize the plasmid (**Figure 5.5 C**). After ligation, the plasmid was transformed into competent *E. coli* cells for propagation and subsequent extraction. Post-transformation, colonies were screened, and plasmid DNA was extracted. The presence of the desired mutations was confirmed by Sanger sequencing, ensuring that the plasmids with the correct sequence were used for further experiments.

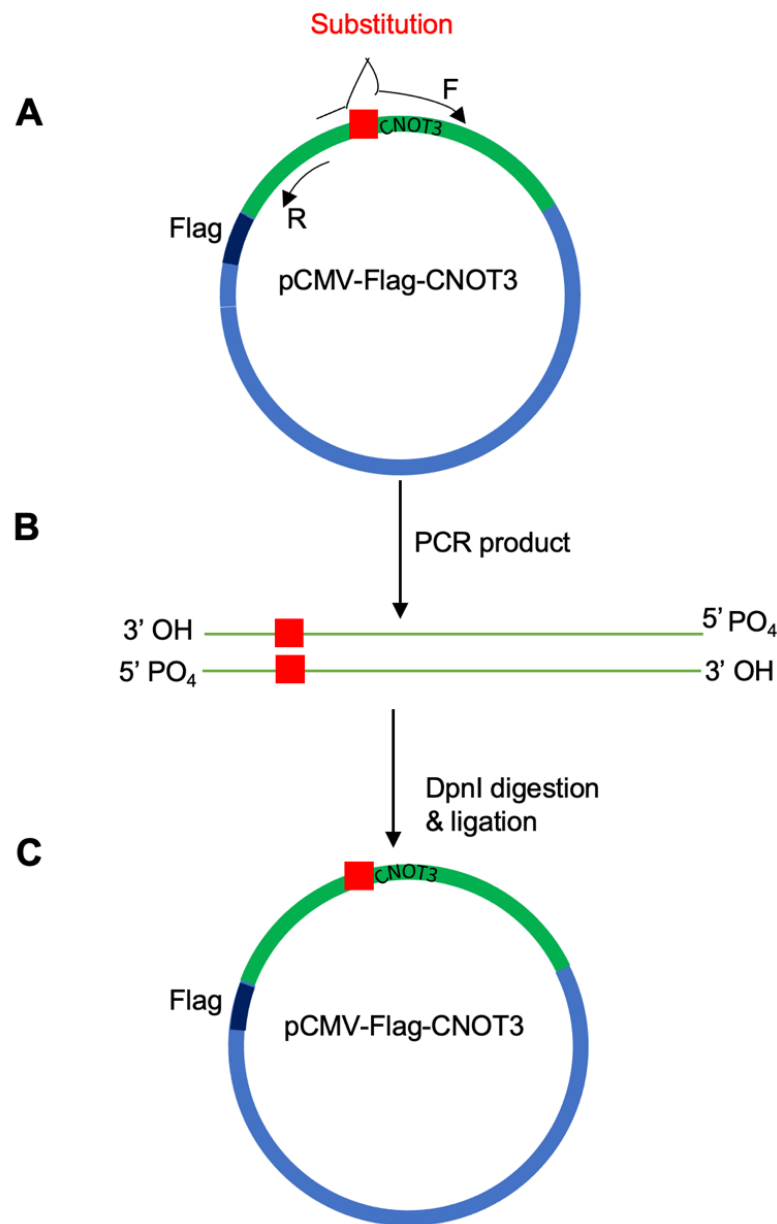


Figure 5.5 Construction of variants of Flag-CNOT3. A schematic representation of site-directed mutation by inverse PCR.

(A) Schematic representation of the pCMV-Flag-CNOT3 plasmid. The plasmid contains the *CNOT3* gene, with forward (F) and reverse (R) primers indicated for site-directed mutagenesis. (B) Illustration of the PCR product obtained after amplification using the forward and reverse primers. The PCR product contains the desired mutation at the substitution site, with 5' (PO₄) phosphate and 3' (OH) hydroxy groups necessary for subsequent ligation. (C) Diagram showing the final mutant pCMV-Flag-CNOT3 plasmid after *Dpn* I digestion and ligation. The red square represents the site of mutation within the *CNOT3* gene.

5.6 Expression of CNOT3 mutations associated with T-ALL and IDDSADF in HEK293 cell lines

To analyse the expression of Flag-CNOT3 variants harbouring mutations associated with CNOT3 in T-All (T-cell Acute Lymphoblastic Leukaemia) and neurodevelopmental disorders (IDDSADF), all plasmids, wild type Flag-CNOT3, empty vector, Flag-CNOT3L48V, Flag-CNOT3R57W, Flag-CNOT3R57Q, Flag-CNOT3K119E, Flag-CNOT3E147K, Flag-CNOT3R188C and Flag-CNOT3R188H were transiently transfected into the HEK293 cell line. Subsequent immunoblot analysis revealed that all variants were expressed at similar levels indicating that the mutations do not lead to unstable or misfolded proteins. The expression levels of variants of CNOT3 were evaluated using Flag antibody, enabling a comprehensive assessment of their respective levels (**Figure 5.6**).

Expression of CNOT3 did not affect expression of other subunits of the Ccr4-Not complex, specifically CNOT1 and CNOT2, or the L7A ribosomal protein. This suggests that the mutations in CNOT3 do not disrupt the overall stability or expression of the Ccr4-Not complex.

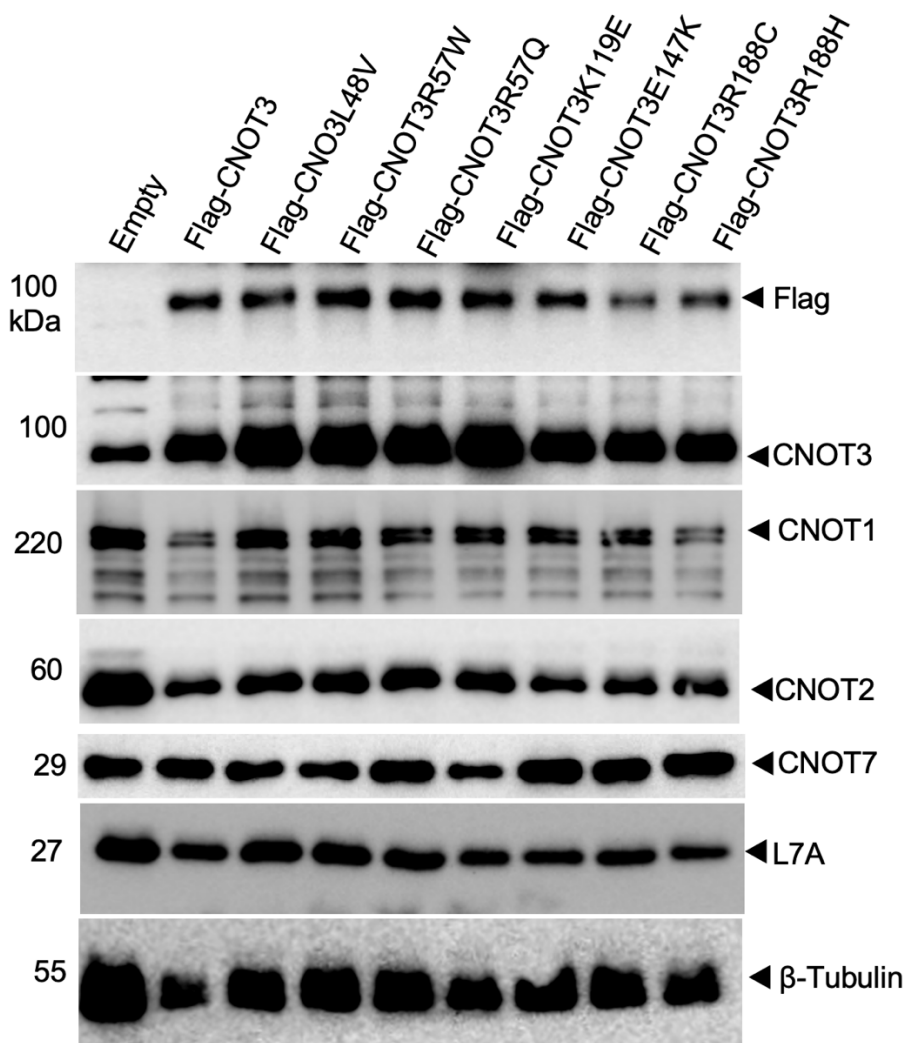


Figure 5.6 Expression of CNOT3 mutations associated with T-ALL and IDDSADF in HEK293 Cell lines.

(R57Q, R57W are associated with T-ALL, while mutations L48V, K119E, E147K, A188H and A188C are associated with IDDSADF) Flag-CNOT3 variants were expressed equally in the HEK293 cell line using anti-Flag. The expression levels of other components of the Ccr4-Not complex were assessed using CNOT1 (monoclonal antibody Cell Signalling), CNOT2 (monoclonal antibody, Cell Signalling), CNOT7 (monoclonal antibody Cell Signalling), and L7A (monoclonal antibody Cell Signalling), a ribosomal antibody. β-tubulin was used as a loading control.

5.7 Patient-derived mutations linked with T-ALL and IDDSADF in CNOT3 do not affect its incorporation into the Ccr4-Not complex

To investigate the impact of patient-derived CNOT3 mutations on its interaction with the Ccr4-Not complex, co-immunoprecipitation experiments were performed in HEK293 cells. Cells were transiently transfected with plasmids encoding wild-type Flag-CNOT3, an empty vector, and CNOT3 variants harbouring mutations associated with either T-cell acute lymphoblastic leukaemia (T-ALL) and intellectual disability-dysmorphic facies-autism spectrum disorder (IDDSADF).

After 48 hours, cell lysates were harvested, and co-immunoprecipitation was performed using anti-flag beads. Consistent expression of all CNOT3 variants was confirmed by Western blot analysis of the total cell lysates (**Figures 5.7 A and 5.8 A**). As expected, no CNOT3 expression was detected in the empty vector control.

In the co-immunoprecipitation experiments, CNOT1, CNOT2, and CNOT7 were effectively co-precipitated by both wild-type Flag-CNOT3 and the mutant variants (**Figures 5.7 B and 5.8 B**). This indicates that the ability of CNOT3 to interact with these core components of the Ccr4-Not complex is not impaired by either the T-ALL-associated mutations (Arg57Gln and Arg57Trp) or the IDDSADF-associated mutations (L48V, K119E, E147K, R188H, and R188C). Furthermore, as anticipated, no precipitation was observed in the empty plasmid control. This outcome reaffirms the specificity of the interactions observed and underscores the validity of the experimental findings.

These findings suggest that disrupted complex formation is not responsible for the pathogenic effects of these CNOT3 mutations. Instead, it is possible that they may result from subtle functional alterations within the assembled complex or involve CNOT3 interactions with other proteins or pathways. Further investigation is needed to elucidate the precise mechanisms by which these mutations contribute to the disease.

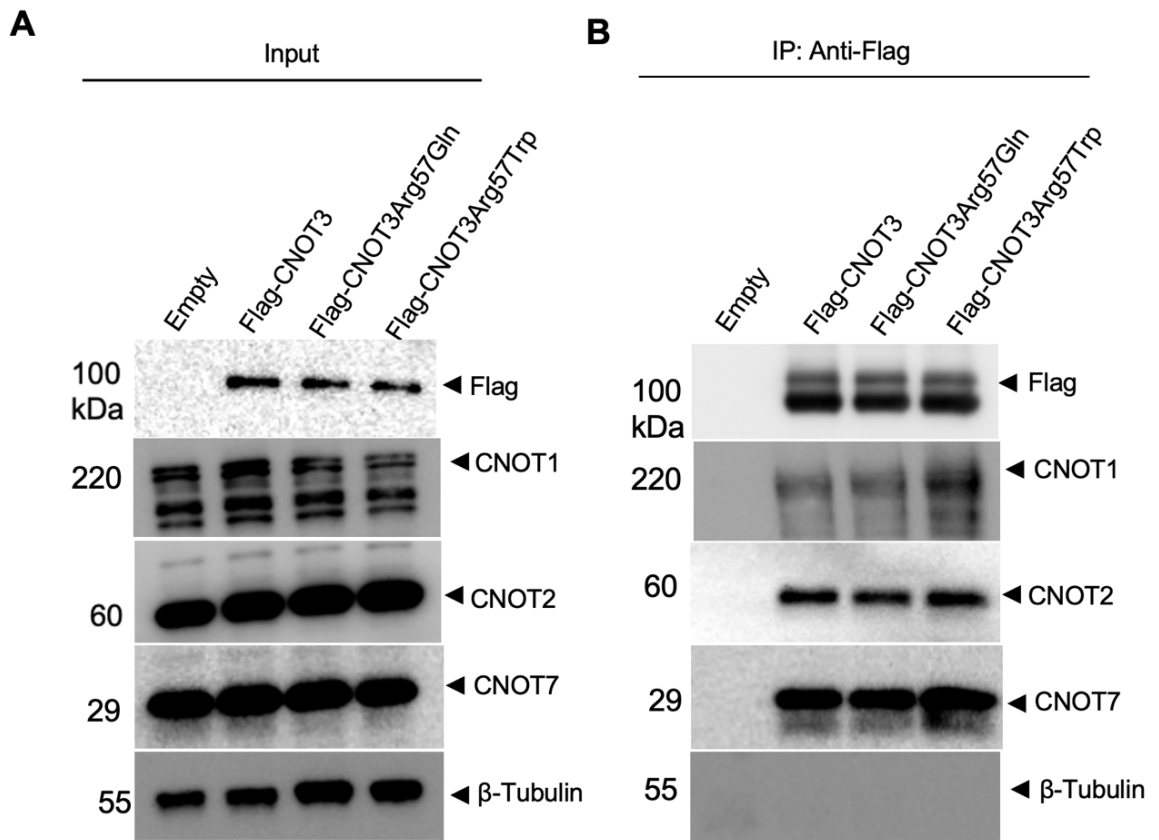


Figure 5.7 Patient-derived mutations linked with T-ALL in CNOT3 do not affect its incorporation into the Ccr4-Not complex.

(A) Immunoblot analysis of cell lysates. (B) Immunoprecipitation of Flag-CNOT3 co-precipitated, Flag (rabbit monoclonal antibody, Sigma-Aldrich), CNOT1 (rabbit monoclonal antibody, Cell Signalling), CNOT2 (rabbit monoclonal antibody, Cell Signaling), and CNOT7 (rabbit monoclonal antibody, Cell Signalling). β -tubulin was used as a loading control.

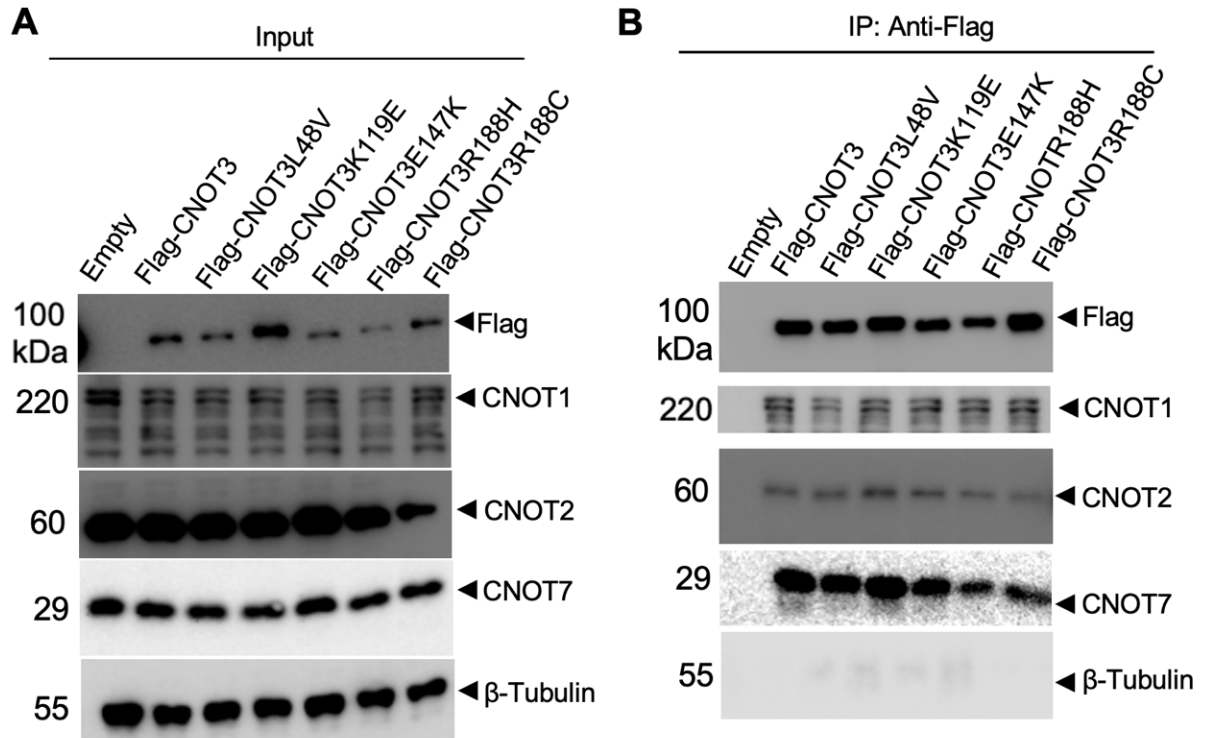


Figure 5.8 Patient-derived mutations linked with (IDDSADF) in CNOT3 do not affect its incorporation into the Ccr4-Not complex.

(A) Western blot analysis of total cell lysates before co-immunoprecipitation. (B) Immunoprecipitation of Flag-CNOT3 co-precipitated Flag (rabbit monoclonal antibody, Sigma-Aldrich), CNOT1 (rabbit monoclonal antibody, Cell Signaling), CNOT2 (rabbit monoclonal antibody, Cell Signaling), and CNOT7 (rabbit monoclonal antibody, Cell Signaling). β-tubulin was used as a loading control.

5.8 Literature-based identification of CNOT2 mutations associated with IDNADFS

Mutations in CNOT2 have been associated with a rare neurodevelopmental disorder known as Intellectual Developmental Disorder with Nasal Speech, Dysmorphic Facies, and variable Skeletal anomalies (IDNADFS) (Niceta et al., 2022; Uehara et al., 2019).

A study by Uehara et al. (2019) described a patient with a de novo heterozygous nonsense mutation in the *CNOT2* gene (c.946A>T, p. Lys316Ter), which was absent in large population databases, indicating its rarity and potential pathogenicity. This mutation, along with other structural variants such as microdeletions involving the *CNOT2* gene, contributes to a clinical spectrum characterised by developmental delays, intellectual disability, distinctive facial dysmorphisms, and hypotonia.

Niceta et al. further expanded the understanding of this disorder by reporting five unrelated individuals with de novo CNOT2 variants, including one with a truncating mutation (c.1482T>G, p. Tyr494Ter) and another with a missense change (c.1396A>C, p. Asn46His). These mutations (**Table 5.5 and Figure 5.9**) along with microdeletions at the 12q15 locus encompassing CNOT2, suggest that *CNOT2* mutation is a significant factor in the pathogenesis of IDNADFS.

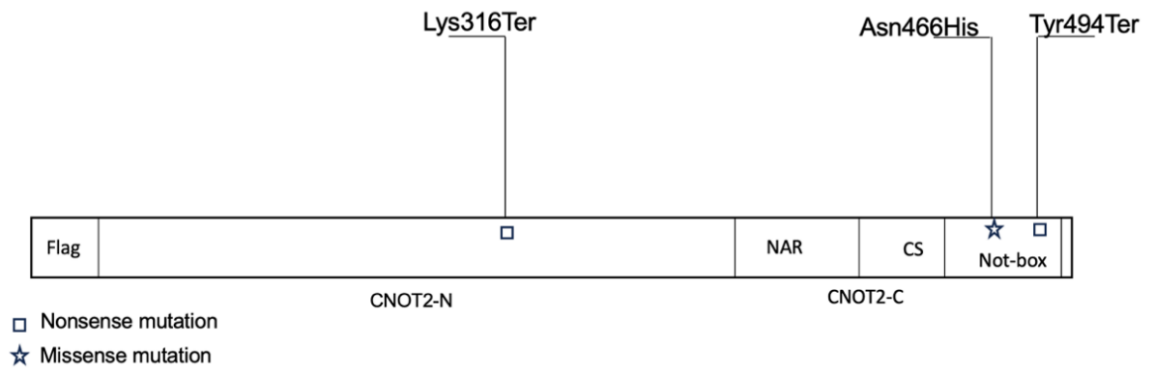
CNOT2 mutations are linked to nasal speech, facial dysmorphism, and skeletal anomalies like up-slanted palpebral fissures and micrognathia. The Deciphering Developmental Disorders (DDD) study reinforces *CNOT2*'s role in neurodevelopmental disorders (Uehara et al. 2019; Niceta et al. 2023).

Table 5.5 CNOT2 mutations associated with IDNADFS

Variant	Effect	Type of Variant	Disease	Reference
c.946A>T	p. (Lys316Ter)	Nonsense	IDNADFS	A
c.1482T>G	p. (Tyr494Ter)	Nonsense	IDNADFS	B
c.1396A>C	p. (Asn466His)	Missense	IDNADFS	B

A (Uehara et al. 2019), B (Niceta et al. 2023).

A



B

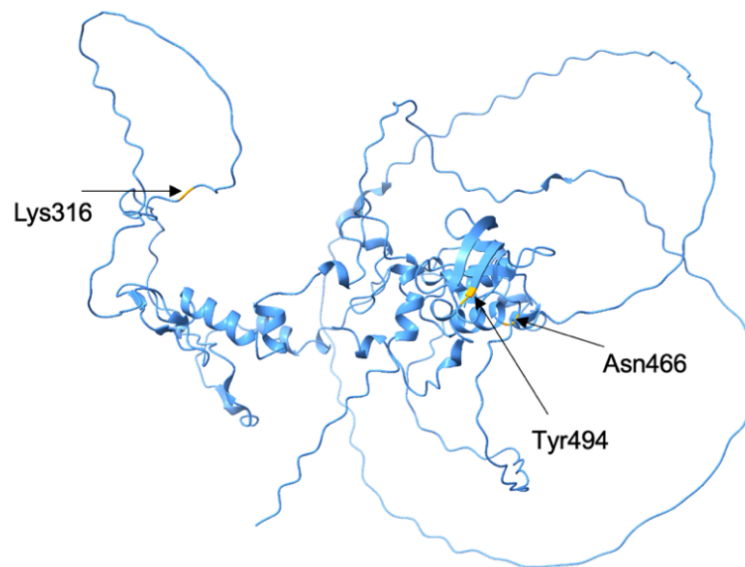


Figure 5.9 Schematic representation of the CNOT2 mutations.

(A) Domains are indicated (NAR: Not1-interacting domain, CS: CNOT2-specific domain, Not-Box). The positions of identified mutations are marked with symbols: squares for nonsense mutations, and a star for a missense mutation. (B) Three-dimensional structure of the CNOT2 protein predicted by AlphaFold (AF-Q9NZN8-F1). The polypeptide chain is shown in blue. The locations of the amino acids affected by panel (A) mutations are highlighted in yellow.

5.9 CNOT2 variants found in neurodevelopmental disorders analysed using predictive algorithms

To distinguish whether these SNVs may lead to functional variation and potentially contribute to neurodevelopmental disorders (IDNADFS), three algorithms (SIFT, PolyPhen-2 and AlphaMissense) were used to analyse and predict the effects of coding non-synonymous variants on protein function.

5.9.1 CNOT2 variants found in neurodevelopmental disorders analysed using Sorting intolerant from tolerant (SIFT)

The potential impact of these mutations was assessed using Sorting Intolerant From Tolerant (SIFT), as described in section 5.4.1. This analysis focuses on specific missense mutations in CNOT2 associated with neurodevelopmental disorders, including one previously reported in the literature: a missense change (c.1396A>C, p. AsN466His) that results in a single amino acid change in the CNOT2 protein sequence (Uehara et al. 2019).

Table 5.6 Analysis of CNOT2 variant associated with IDNADFS evaluated with SIFT

Variant	Score	Pathogenicity	Disease
N466His	0.04	damaging	IDNADFS

The analysis of the CNOT2 N466His variant, associated with intellectual disability, autism spectrum disorder, dysmorphic facial features, and seizures (IDNADFS), using SIFT yielded a score of 0.04 (**Table 5.6**), falling well below the threshold for tolerance. This strongly suggests that the substitution of asparagine with histidine at position 466 is deleterious and likely disrupts the normal function of the CNOT2 protein. The only damaging nature of this mutation underscores its potential role in the development of neurodevelopmental disorders and highlights the importance of CNOT2 in maintaining proper neurological function.

5.9.2 CNOT2 variants found in neurodevelopmental disorders analysed using PolyPhen-2

Following the initial evaluation of CNOT3 variants using SIFT and AlphaMissense, PolyPhen-2 (Polymorphism Phenotyping v2), described in section 5.4.3), was employed to further elucidate the potential functional impact of these mutations (Adzhubei, Jordan, and Sunyaev 2013)

The analysis (**Table 5.7**) reveals that CNOT variant (N466H) is predicted to be benign," with score of 0.297. This prediction suggests a high probability that this variant does not impair the structure or function of the CNOT3 protein.

Figure 5.10 provides a visual representation of the CNOT2 protein structure, highlighting the significance of the asparagine residue at position 466 (N466). The visualisation suggests that N466 plays a role in maintaining protein integrity by forming interactions (potentially hydrogen bonds) with other amino acids within the protein, such as ARG 534 and GLU 463. These interactions likely contribute to the stability and proper folding of CNOT2.

However, in silico predictions, which utilise computational algorithms to forecast the impact of amino acid substitutions, suggest that a change from asparagine (N) to histidine (H) at position 466 (N466H mutation) might not significantly affect the protein's structure or function. This prediction contradicts the structural insights from the image, which implies the importance of N466 for protein stability.

Table 5.7 Analysis of CNOT2 Variant associated with IDNADFS evaluated with PolyPhen-2

Variant	Pathogenicity	Score	Disease
N466H	benign	0.266	IDNADFS

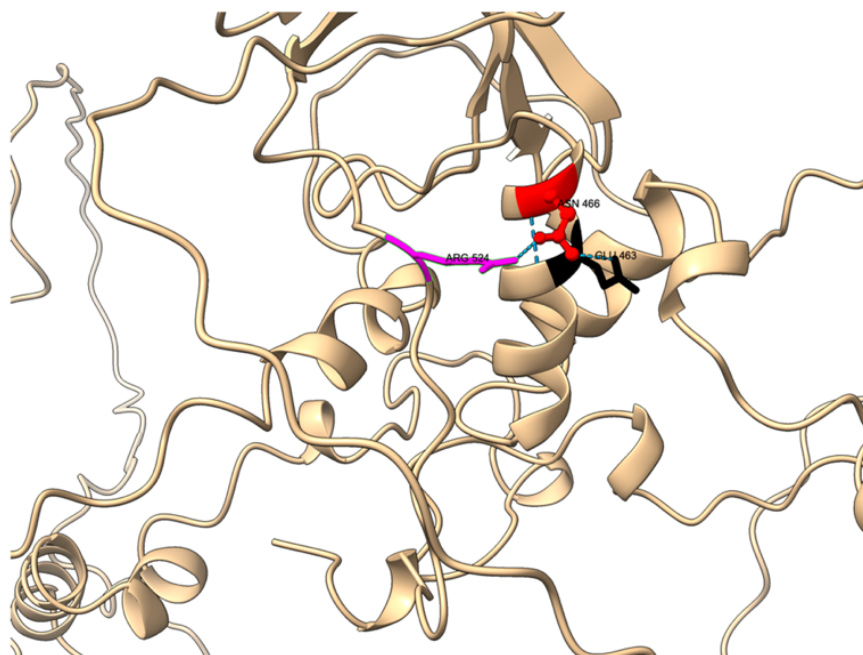


Figure 5.10 Three-dimensional schematic of Asn466H mutation in CNOT2.

This image represents the three-dimensional model of the CNOT2 protein (AF-Q9NZN8-F1-v4, UniProt ID: Q9NZN8), visualised using ChimeraX software version 1.8. The model highlights key structural features and interactions within the CNOT2 protein. The backbone of the protein is represented in a tan colour, with alpha-helices depicted as coiled ribbons and beta-strands as broad arrows, illustrating the secondary structure elements of the protein. Key residues involved in critical interactions are shown in different colours: Asn466 is shown in red, indicating its importance in the protein's structure. Arg524 is marked in magenta, highlighting its involvement in hydrogen bonding interactions. Glu463 is represented in cyan, demonstrating its role in stabilising the protein through hydrogen bonds. Hydrogen bonds between Asn466 and Arg524 are depicted as dashed lines.

5.9.3 *CNOT2* variants found in neurodevelopmental disorders analysed using AlphaMissense

Next, to assess the potential pathogenicity of this missense variant, the AlphaMissense tool (described in section 5.4.2) was used (Cheng et al. 2023).

The AlphaMissense analysis assigned a pathogenicity score of 0.297 (**Table 5.8**) to the Asn466His (N466H) variant in CNOT2, classifying it as likely benign. This score indicates that the substitution of Asparagine (N) with Histidine (H) at position 466 is unlikely to significantly disrupt the normal function of the Ccr4-Not complex or contribute to the pathogenesis of IDNADFS. However, this prediction seems to contradict the findings from the 3D modelling assessment, which indicated that the N466His mutation could destabilise the Not-Box region of the CNOT2 protein due to the loss of a hydrogen bond with Arg524. This discrepancy highlights the limitations of relying solely on computational predictions for pathogenicity assessment.

Despite the AlphaMissense classification, the structural analysis suggests that the N466H variant may indeed affect protein stability and function. Therefore, further investigation is necessary to determine the true impact of this variant on the CNOT2 protein and its role in IDNADFS pathogenesis.

Table 5.8 Analysis of CNOT2 Variant associated with IDNADFS evaluated with AlphaMissense

Variant	Score	Pathogenicity	Disease
N466H	0.297	benign	IDNADFS

5.9.4 *CNOT2* variants found in neurodevelopmental disorders analysed using Provean

In this study, the functional impact of two truncating mutations in *CNOT2*, K316Ter and Tyr494Ter, was evaluated using PROVEAN (Protein Variation Effect Analyzer). PROVEAN (Choi et al. 2012) is a highly suitable tool for analysing truncating mutations because it assesses protein sequence changes based on evolutionary conservation and potential biological impact across homologous sequences. Unlike in silico tools such as

SIFT and PolyPhen-2, which are optimised for missense mutations, PROVEAN is particularly effective for predicting the effects of truncating mutations that introduce premature stop codons, as it evaluates the impact of such mutations in terms of overall protein stability and loss of function.

For PROVEAN, a score below -2.5 is considered deleterious, indicating that the mutation is likely to impair protein function significantly (Choi et al. 2012). The K316Ter mutation yielded a PROVEAN score of -716.934 (**Table 5.9**), suggesting a profound loss of function due to premature termination at amino acid position 316. This early truncation likely results in a severely shortened protein that lacks critical domains necessary for CNOT2's stability, its interactions within the Ccr4-Not complex, and its overall functionality in mRNA regulation. The extremity of this score emphasises the likelihood that K316Ter severely compromises CNOT2's cellular roles, aligning with the pathogenesis of disorders associated with Ccr4-Not dysfunction.

Similarly, the Tyr494Ter mutation produced a PROVEAN score of -137.890 (**Table 5.9**), indicating a highly deleterious effect, though somewhat less severe than K316Ter. This truncation halts protein translation at residue 494, removing essential domains near the C-terminal domain. Although this mutation retains more of the protein structure than K316Ter, the substantial PROVEAN score suggests that it disrupts function significantly, possibly by impairing critical intermolecular interactions within the Ccr4-Not complex and affecting downstream regulatory activities.

In conclusion, these PROVEAN scores demonstrate that both K316Ter and Tyr494Ter are likely pathogenic mutations that cause functional loss in CNOT2. By destabilising protein structure and removing domains essential for Ccr4-Not complex integrity, these mutations contribute to the mechanistic basis of neurodevelopmental disorders and other pathologies linked to CNOT2 haploinsufficiency.

Table 5.9 Analysis of CNOT2 Variant associated with IDNADFS evaluated with Provean

Variant	Score	Pathogenicity	Disease
K316Ter	-716.934	deleterious	IDNADFS
Y493Ter	-137.890	deleterious	IDNADFS

5.10 Constructing plasmids containing patient-derived mutations linked to rare neurodevelopmental disorders in CNOT2

To investigate the functional consequences of these mutations, site-directed mutagenesis was employed to introduce them into a plasmid expressing the CNOT2 gene. The pCMV-Flag-CNOT2 plasmid used in this study includes a Flag tag to facilitate downstream protein detection and purification. All plasmids were constructed as described in section 5.5 of this chapter.

5.11 Transient expression of Flag-CNOT2 mutants in HEK293 cell line

Following the construction of these plasmids, their expression levels were evaluated in a series of experiments. To assess the impact of *CNOT2* mutations associated with IDNADFS (intellectual developmental disorder with nasal speech, dysmorphic facies, and variable skeletal anomalies, OMIM 618608), the plasmids—including wild-type Flag-CNOT2, empty vector, Flag-CNOT2N466H, Flag-CNOT2Y494Ter, and Flag-CNOT2K316Ter—were transiently transfected into the HEK293 cell line. Plasmid expression levels were then assessed using a Flag antibody. Western blot analysis (**Figure 5.111**) of the total cell lysate was carried out to validate the expression of various Flag-CNOT2 variants. The results of this analysis revealed the expression of distinct Flag-CNOT2 variants. Similar expression levels of the wild-type Flag-CNOT2 and Flag-CNOT2N466H variants were observed. However, the Flag-CNOT2Y494Ter and Flag-CNOT2K316Ter variants were expressed at different molecular weights, suggesting

that nonsense-mediated decay of the mutant mRNA may lead to reduced expression compared to wild-type Flag-CNOT2. The Flag-CNOT2Y494Ter variant showed very low expression, while the Flag-CNOT2K316Ter variant exhibited robust expression. Due to the low expression of Flag-CNOT2-Y494Ter, it was not further analysed.

Moreover, expression levels of the Ccr4-Not subunits CNOT1, CNOT3, and CNOT7, as well as the L7A and S3 ribosomal proteins were unchanged.

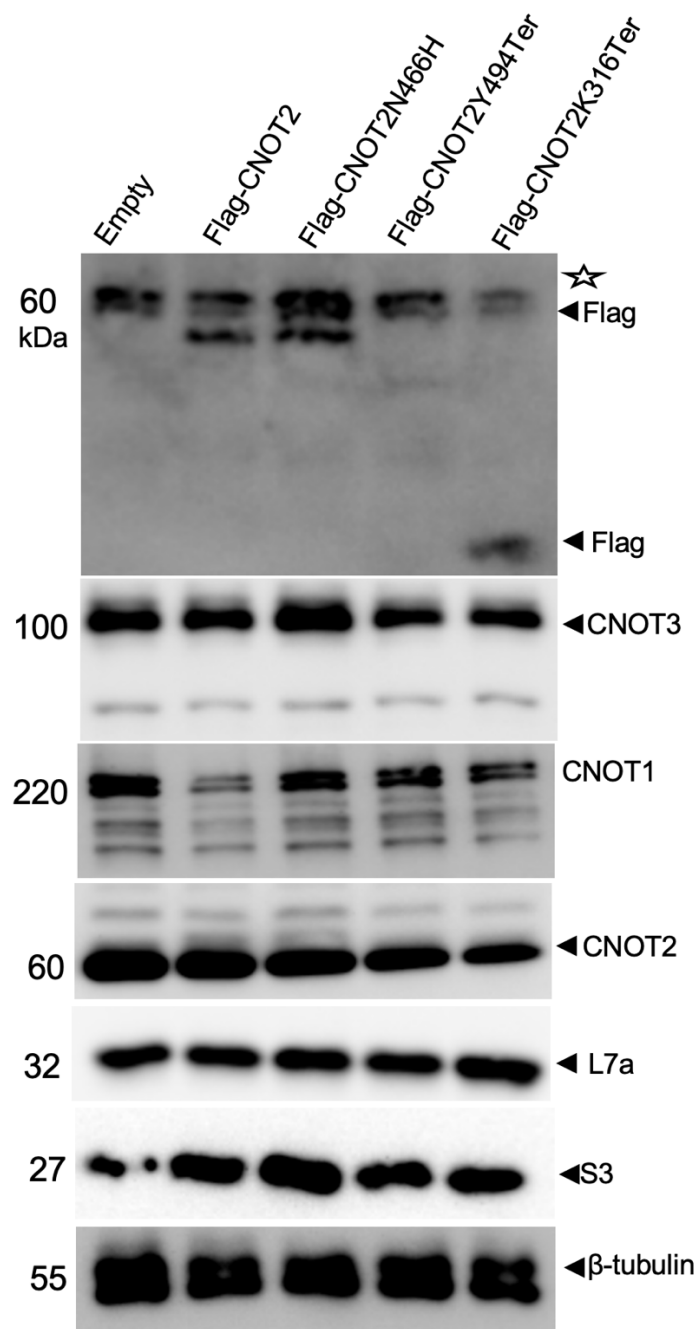


Figure 5.11 Transient expression of Flag-CNOT2 variants in HEK293 cell line.

The protein expression level of other components of the Ccr4-Not complex, CNOT1, CNOT2, CNOT3, and two ribosomal proteins, L7a and S3, was assessed using antibodies Flag (rabbit monoclonal antibody, Sigma-Aldrich), CNOT1 (rabbit monoclonal antibody, Cell Signaling), CNOT2, (rabbit monoclonal antibody, Cell Signaling), CNOT3 (mouse monoclonal antibody 4B8, AbNova), L7a (rabbit monoclonal antibody, Cell Signaling) and S3 (rabbit monoclonal antibody, Cell Signaling). β-tubulin was used as a loading control.

5.12 Patient-derived mutations linked to IDNADFS disrupt the interaction of CNOT2 with the Ccr4-Not complex

Similar to the CNOT3 investigation, the potential effects of IDNADFS-associated mutations on CNOT2's interaction with the Ccr4-Not complex were explored. For this analysis, a series of plasmids containing desired mutations (Flag-CNOT2-N466H, Flag-CNOT2K316Ter), along with wildtype Flag-CNOT2 and an empty vector, were transiently transfected into HEK293 cells, followed by co-immunoprecipitation. This approach aimed to identify any alterations in binding affinity or functional consequences of these mutations on the CNOT2-CNOT3-CNOT1 interaction and the overall Ccr4-Not complex. After 48 hours, cellular lysates were harvested and subjected to co-immunoprecipitation.

Western blot analysis of the total cell lysate conducted before co-immunoprecipitation confirmed the expression of various CNOT2 variants (**Figure 5.12 A**). The results indicated that all Flag-CNOT2 variants exhibited similar expression levels to wild-type Flag-CNOT2 in HEK293 cells. As expected, no expression was observed in the empty vector, confirming the assay's specificity. Furthermore, the expression levels of CNOT1, CNOT3, and CNOT7, components of the Ccr4-Not complex, were also investigated and found to be similar (**Figure 5.12 A**).

Wild-type Flag-CNOT2, empty vector, and the mutated variants Flag-CNOT2N466H and Flag-CNOT2K316Ter were utilised in co-immunoprecipitation experiments. The results revealed the precipitation of CNOT1, CNOT3, and CNOT7 in both wild-type Flag-CNOT2 and Flag-CNOT2N466H using anti-flag beads (**Figure 5.12 B**). This suggests that the missense mutation associated with IDNADFS in CNOT2 does not completely affect the incorporation of CNOT2 into the Ccr4-Not complex. However, a noticeable reduction in the amount of co-precipitated subunits was observed for the N466H mutant compared to the wild-type, which may reflect partial disruption of complex assembly or reduced expression or stability of the mutant protein.

Notably, no precipitation was detected with Flag-CNOT2K316Ter (**Figure 5.12 B**). This suggests that the stop codon produces an incomplete protein, preventing its integration into the Ccr4-Not complex. Similarly, the empty vector failed to precipitate any subunits

of the Ccr4-Not complex, confirming the specificity of the observed interactions and the validity of the experimental findings.

These findings demonstrate that the CNOT2 missense mutation (Flag-CNOT2N466H) partially compromise the incorporation of CNOT2 into the Ccr4-Not complex, as evidenced by its reduced pull down of CNOT1, CNOT3, and CNOT7. In addition to that, the truncating mutation (Flag-CNOT2K316Ter) results in a truncated protein incapable of integrating into the complex, thereby disrupting its function.

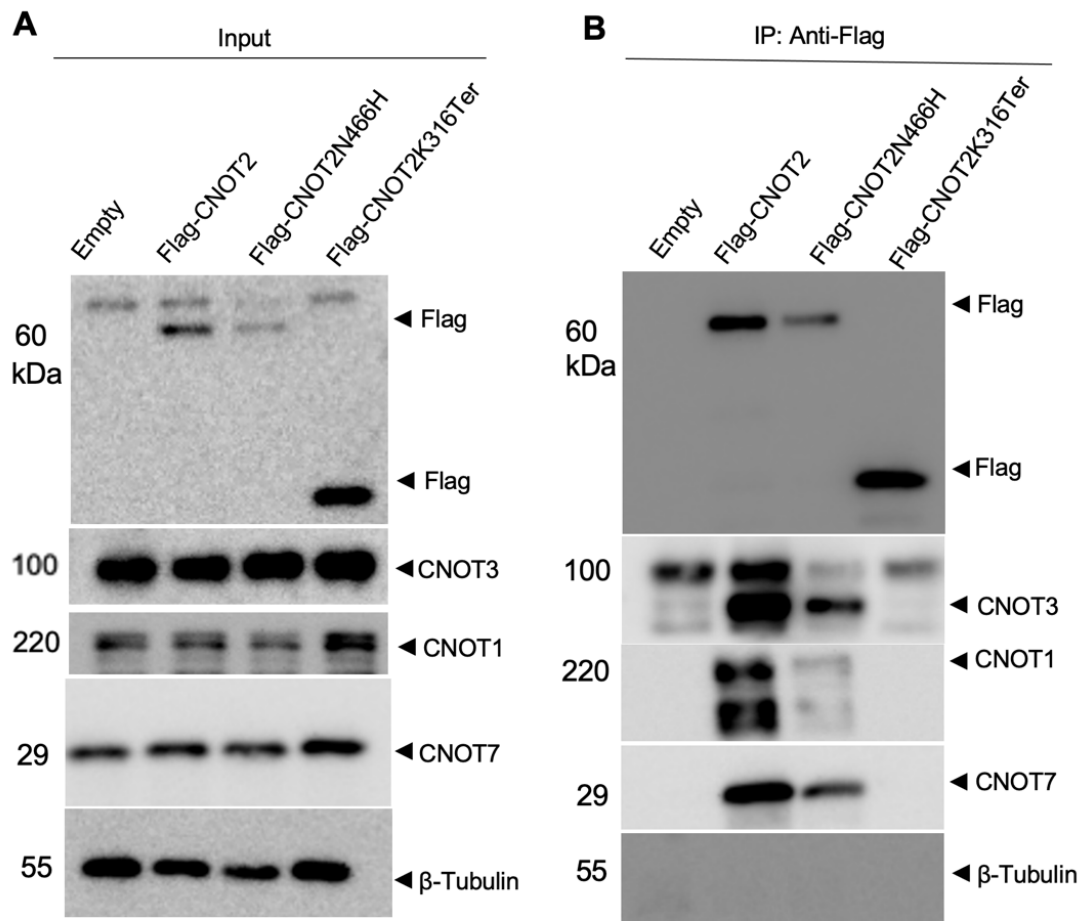


Figure 5.12 Patient-derived mutations linked to IDNADFS disrupt the interaction of CNOT2 with the Ccr4-Not complex.

(A) Western blot analysis of the total cell lysate conducted before co-immunoprecipitation. (B) Immunoprecipitation of Flag-CNOT2 variants. Proteins were detected using antibodies, Flag (rabbit monoclonal antibody, Sigma-Aldrich), CNOT1 (rabbit monoclonal antibody, Cell Signaling), CNOT3 (mouse monoclonal antibody 4B8, AbNova) and CNOT7 (rabbit monoclonal antibody, Cell Signaling). β-tubulin was used as a loading control.

5.13 Discussion

The intricate interplay between CNOT3 and CNOT2 within the Ccr4-Not complex (Ito, Inoue, et al. 2011; Collart and Panasenko 2012, 2017) has been investigated in this chapter through a combination of in silico predictions and experimental validation, focusing on patient-derived variants. These analyses have provided valuable insights into the structural and functional roles of these key subunits.

The findings corroborate previous structural studies and further illuminate the functional roles of specific domains within CNOT3, particularly C-terminal for its interaction with CNOT2, a finding that is consistent with recent investigations (Boland et al. 2013; Suzuki et al. 2015; Collart 2016). Moreover, the necessity of integrating both computational and experimental approaches to fully understand the impact of disease-associated mutations on protein function has been highlighted by the observed discrepancy between computational predictions and experimental results.

5.13.1 Importance of CNOT3 termini for Ccr4-Not complex formation

The critical importance of the C-terminal domain and, specifically, the Not-Box region of CNOT3 for its incorporation into the complex was elucidated by co-immunoprecipitation assays. This finding is consistent with prior structural studies that have underscored the pivotal role of these domains in mediating protein-protein interactions within the complex (Suzuki et al. 2015; Buschauer et al. 2020). The indispensability of these regions for complex assembly and stability was further solidified by the observation that CNOT3 variants lacking these regions failed to interact with other subunits.

Notably, findings also demonstrate that the N-terminal domain (NTD) of CNOT3 is dispensable for its interaction with CNOT2, aligning with previous studies (Suzuki et al. 2015; Buschauer et al. 2020). This consensus is further supported by our findings, which collectively suggest that the NTD may not be directly involved in mediating the core interaction between CNOT3 and CNOT2 within the Ccr4-Not complex. While the NTD plays a role in mediating CNOT3's interactions with ribosomes (Absmeier et al. 2023; Buschauer et al. 2020; Collart, Audebert, and Bushell 2023)

5.13.2 *In silico* analysis of CNOT3 and CNOT2 mutations

CNOT3 and *CNOT2* mutations found in the literature were analysed using *in silico*. The *in silico* predictions generated by SIFT, PolyPhen-2, AlphaMissense and PROVEAN provided valuable insights into the potential pathogenicity of *CNOT3* and *CNOT2* variants. In general, a high degree of concordance was observed between these tools, with most variants predicted to be damaging or likely pathogenic. However, some discrepancies were also noted, highlighting the importance of considering multiple prediction algorithms and incorporating experimental validation to accurately assess the functional impact of mutations. For instance, the K119E and E147K variants in *CNOT3* was classified as tolerated by SIFT but as likely pathogenic by AlphaMissense and possibly damaging by PolyPhen-2. variants.

Similarly, The *CNOT2* Asn466His (N466H) variant, located within the conserved Not-box region, demonstrates contrasting interpretations across predictive tools. While AlphaMissense and PolyPhen-2 classify this variant as benign, analysis with SIFT and structural evidence underscores its critical role in protein stability, suggesting pathogenicity. Structural predictions indicate that the N466H substitution disrupts essential hydrogen bonds within the Not-box region, potentially destabilising the protein structure. This variant, confirmed as *de novo*, is associated with phenotypic features that overlap with those observed in truncating mutations, such as K316Ter, highlighting its likely functional significance (Uehara et al., 2019; Niceta et al., 2023).

The truncating mutation K316Ter, located in the N-terminal domain of *CNOT2*, was analysed using PROVEAN and found to be deleterious. This truncation removes significant portions of the protein, which is essential for its function. Both N466H and K316Ter mutations lead to core phenotypic features, indicating severe impairment of *CNOT2* functionality. However, their mechanisms of disruption differ: N466H likely destabilises the protein through structural alterations in the Not-Box region, while K316Ter results in a loss of functionality (Uehara et al. 2019; Niceta et al. 2023).

Furthermore, although *CNOT3* is highly intolerant to missense variation, as reflected by its significantly reduced number of observed variants compared to the expected number

(expected: 292.1; observed: 156; $Z = 3.89$) (Martin et al. 2019), co-immunoprecipitation experiments demonstrated that the missense mutations did not disrupt the interaction with the core Ccr4-Not complex. This finding suggests that these mutations may still contribute to T-ALL and IDDSADF pathogenesis through more subtle impacts or other mechanisms such as interaction between the N-terminal domain of CNOT3 and the ribosome, which could influence translation efficiency (Zhu et al. 2024; Absmeier et al. 2023).

The truncating mutation in *CNOT2* (K316Ter) that causes IDNADFS likely acts through haploinsufficiency (Uehara et al. 2019). PROVEAN analysis and co-immunoprecipitation experiments demonstrated that Flag-CNOT2K316Ter fails to incorporate into the Ccr4-Not complex, severely compromising CNOT2 function. In contrast, for the missense mutation N466H, although structural analysis suggested that the amino acid substitution to histidine would be detrimental to the integrity of the Not-Box region and causing loss of function, *in vitro* experiments showed partial disruption of its incorporation into the Ccr4-Not complex. This suggests that N466H may exert more subtle effects that require further investigation to fully elucidate its role and impact.

The disparity between *in silico* predictions and experimental results serves as a cautionary reminder of the limitations inherent in computational models. While these tools are invaluable for generating hypotheses and guiding research, they may not always accurately predict the functional consequences of mutations within the complex cellular environment.

While this study provides insights into the impact of *CNOT3* and *CNOT2* mutations on Ccr4-Not complex formation, it's important to acknowledge its limitations. The use of *in vitro* assays, although valuable, may not fully recapitulate the complexities of *in vivo* cellular environments. Future studies should employ *in vivo* models and patient-derived samples to validate and extend our findings.

Chapter 6.

Concluding remarks and Future outlook

6.1 Concluding remarks and future outlook

This study focuses on understanding the role of human CNOT3, a key subunit of the Ccr4-Not complex, in cell proliferation and in disease. Insights into how CNOT3 contributes to these processes are essential for uncovering its broader significance in human health and pathology.

One of the initial steps in mRNA degradation is deadenylation, the shortening and removal of the poly(A) tail from mRNAs (Park et al. 2023; Pavanello, Hall, and Winkler 2023; Passmore and Collier 2022; Webster et al. 2017). The Ccr4-Not complex is a central player in this process, acting as a multi-subunit deadenylase that regulates the stability and turnover of mRNAs (Chen and Shyu 2011). The complex is critical for various cellular functions, including mRNA degradation, translation, transcription regulation, and protein ubiquitination, underscoring its importance in maintaining cellular homeostasis (Collart 2016).

The human Ccr4-Not complex consists of nine subunits, each with a specific role. CNOT1 serves as a scaffold protein, binding other subunits and organising them into functional modules (Collart and Panasenko 2017). The catalytic module includes CNOT6/6L (Mittal et al. 2011) and CNOT7/8 (Mostafa et al. 2020), responsible for removing the poly(A) tail (Yi et al. 2018). However, the non-catalytic subunits, such as CNOT2 (Ito, Inoue, et al. 2011) and CNOT3 (Zhu et al. 2024), play equally important roles in the regulatory mechanisms that guide the complex's interaction with mRNAs. In particular, CNOT3 has emerged as a crucial regulatory subunit, maintaining the structural integrity of the complex and ensuring its proper functioning.

This thesis focused on CNOT3, specifically its N-terminal and C-terminal domains, to better understand its role within the Ccr4-Not complex. The C-terminal domain of CNOT3, which contains the Not-Box, facilitates the interaction with CNOT2, securing incorporation of CNOT3 and CNOT2 within the Ccr4-Not complex (Suzuki et al. 2015; Ghashghaei et al. 2024; Boland et al. 2013; Xu et al. 2014). In contrast, the N-terminal domain of CNOT3, while less studied, has gained attention due to its involvement in critical protein-protein interactions

essential for translation regulation (Buschauer et al. 2020; Ghashghaei et al. 2024; Absmeier et al. 2023; Collart, Audebert, and Bushell 2023; Zhu et al. 2024).

Significantly, mutations in the N-terminal domain of CNOT3 have been linked to T-cell acute lymphoblastic leukaemia (T-ALL) and neurodevelopmental disorders (NDDs). Mutations have been identified in patients with these diseases (De Keersmaecker et al. 2013; Martin et al. 2019), indicating that alterations in this domain may disrupt normal cellular functions. These mutations are believed to affect the regulatory capacity of CNOT3, impairing its ability to maintain proper mRNA degradation, transcriptional control and translational regulation, leading to abnormal cellular growth and developmental abnormalities (Martin et al. 2019; De Keersmaecker et al. 2013; Uehara et al. 2019; Niceta et al. 2023; Ghashghaei et al. 2024). While this project focused on understanding the broader role of CNOT3 in HEK293 cells, these disease-associated mutations highlight the critical importance of the N-terminal domain in human health and disease.

Additionally, mutations in CNOT2 have also been implicated in neurodevelopmental disorders. Recent studies (Uehara et al. 2019; Niceta et al. 2023) have identified CNOT2 mutations as a cause of severe developmental delays, intellectual disability, and autism spectrum disorders. These mutations can alter the integrity and function of the Ccr4-Not complex, as CNOT2 plays a pivotal role in coordinating interactions between other subunits like CNOT1 and CNOT3. Disruption in the regulatory function of CNOT2 can lead to abnormal gene expression (Sohn et al. 2018; Jung et al. 2021; Ito, Inoue, et al. 2011), further highlighting the importance of non-catalytic subunits in maintaining cellular and developmental processes (Sohn et al. 2018; Jung et al. 2021; Ito, Inoue, et al. 2011).

Moreover, while previous research (Suzuki et al. 2015), has demonstrated the essential role of CNOT3 in mouse fibroblast embryonic stem cells (MEFs), this study extended the focus to a human cell model by investigating the function of CNOT3 in HEK293 cell line. This allowed for a more in-depth understanding of how CNOT3 operates in human cells, emphasising its pivotal role in maintaining

cellular proliferation and regulating the Ccr4-Not complex's mRNA degradation activities.

6.2 Generation of stable cell lines for the conditional expression of siRNA-resistant Flag-CNOT3 variants

The generation of stable cell lines for the conditional expression of siRNA-resistant Flag-CNOT3 variants directly supports the first objective of this thesis. These cell lines allow me to investigate the functional contributions of specific regions of CNOT3, including the N-terminal domain, C-terminal domain, and Not-box region. By providing a tightly controlled system to selectively express both wild-type and mutant variants under an inducible promoter, these cell lines facilitated detailed studies into the distinct roles of CNOT3 variants in cellular processes.

To achieve this, a series of siRNA-resistant Flag-CNOT3 expression plasmids, including Flag-CNOT3R1 (resistant variant 1), Flag-CNOT3R2 (resistant variant 2), and truncated variants, were successfully constructed. These plasmids were used to establish stable Flp-In TREx HEK293 cell lines, which offer tightly controlled and inducible expression of CNOT3 variants. The Flp-In TREx system (O'Gorman, Fox, and Wahl 1991) integrates the gene of interest into a defined genomic locus, ensuring consistent expression levels across different cell lines. This system was chosen due to its ability to mitigate clonal variability and its compatibility with siRNA-mediated knockdown of endogenous CNOT3. This approach allows for the conditional knockdown of endogenous CNOT3 while simultaneously expressing siRNA-resistant CNOT3 variants, enabling the functional analysis of these variants in a cellular context.

While siRNA-mediated knockdown was successfully employed in this study to deplete endogenous CNOT3, potential alternative strategies that could have strengthened the robustness and versatility of our approach. For instance, the dTAG system (Nabet et al. 2018) offers a powerful and highly specific alternative for conditional protein depletion. This system employs a tag-based chemical biology approach in which the protein of interest is fused to an engineered

FKBP12 variant (FKBP12^{F36V}), which is selectively targeted for degradation upon treatment with a small molecule degrader such as dTAG-13. The degrader functions by bridging FKBP12^{F36V} to the CRBN E3 ubiquitin ligase, thereby initiating ubiquitination and proteasomal degradation of the tagged protein. A key advantage of the dTAG system is its rapid and reversible mode of action, allowing for the study of immediate cellular consequences following protein loss—something not easily achievable with RNAi, which often has slower kinetics and incomplete knockdown. Moreover, because dTAG targets the protein directly rather than its transcript, it circumvents concerns related to mRNA stability, off-target transcript effects, and the generation of siRNA-resistant mutants (Nabet et al. 2018).

Another effective approach for gene silencing involves designing siRNAs that specifically target the 3' untranslated region (3' UTR) of the endogenous mRNA. This allows for the selective knockdown of the native transcript while enabling the expression of exogenous constructs that lack the targeted 3' UTR sequence. Since the coding sequence of the exogenous gene remains unchanged, this strategy avoids the need for introducing silent mutations to create siRNA resistance, which can sometimes affect translation efficiency, mRNA stability, or protein folding (Yoo et al. 2007).

The cell lines generated in this study provided a valuable resource for dissecting the contributions of specific CNOT3 domains to cell proliferation. Additionally, these cell lines serve as a robust platform to study the ability of different CNOT3 variants to rescue cellular phenotypes following siRNA knockdown of endogenous CNOT3, thereby offering insights into the molecular mechanisms underlying CNOT3's roles.

6.3 Essential domains for cell proliferation

Using the panel of siRNA-resistant stable HEK293 cell lines developed in this study, the essential domains of CNOT3 required for cell proliferation were identified. These findings directly align with the second objective, which aimed to characterise the functional roles of the N-terminal and C-terminal domains, and Not-Box regions of CNOT3 in driving proliferation.

As detailed in Chapter 4, the N-terminal and C-terminal domains of CNOT3 were shown to be indispensable for cell proliferation. Truncated variants lacking the N-terminal and C-terminal domains failed to rescue growth defects following the siRNA-mediated knockdown of endogenous CNOT3. These results highlight the essential roles of these domains in stabilising CNOT3 within the Ccr4-Not complex and supporting its regulatory functions in mRNA degradation and gene expression.

Additionally, the Not-Box region was also demonstrated to be essential, as variants lacking this domain were unable to complement the proliferation defect. These observations underscore the structural requirements for CNOT3's activity, revealing that the disruption of any of these key regions compromises its ability to maintain normal cellular growth. For instance, experiments using the AlamarBlue assay confirmed that only siRNA-resistant variants with intact N-terminal and C-terminal domains could restore proliferation, further validating these domains' importance.

These findings not only clarify the structural contributions of CNOT3 but also suggest a potential for diseases, such as cancer, where disruptions in cell proliferation are a hallmark.

6.4 Characterising the interactions and disease-associated mutations of CNOT3 and CNOT2 variants in the Ccr4-Not Complex

The third objective of this thesis was addressed to investigate how missense mutations in *CNOT3* and *CNOT2* contribute to neurodevelopmental disorders (Martin et al. 2019) and T-cell acute lymphoblastic leukaemia (T-ALL)(De Keersmaecker et al. 2013), alongside determining the functional significance of N-terminal and C-terminal domains, and Not-Box region in the Ccr4-Not complex interactions. Disease-associated mutations identified from the literature were introduced into Flag-tagged CNOT3 and CNOT2 constructs, and their effects on protein interactions within the Ccr4-Not complex were examined through expression in HEK293 cells.

Investigation of CNOT3 domain functionality was conducted through co-immunoprecipitation analysis in HEK293 cells using various CNOT3 constructs. The N-terminal domain was found to be dispensable for complex formation, as both wildtype Flag-CNOT3 and Flag-CNOT3R2ΔNTD successfully co-precipitated with CNOT1, CNOT2, and CNOT7. However, the deletion of the C-terminal domain and the Not-Box region abolished these interactions, demonstrating their crucial role in incorporating CNOT3 into the Ccr4-Not complex. These findings align with previous structural studies (Suzuki et al. 2015; Boland et al. 2013) indicating the importance of the C-terminal in maintaining Not-module integrity.

The assessment of these disease-associated mutations in *CNOT3* and *CNOT2* was conducted using both computational and experimental approaches. High concordance in predicting damaging effects for several CNOT3 missense mutations (L48V, R57W, R57Q, R188C, and R188H) was demonstrated through computational analyses utilising prediction algorithms including SIFT (Kumar, Henikoff, and Ng 2009), PolyPhen-2 (Adzhubei, Jordan, and Sunyaev 2013), and AlphaMissense (Ljungdahl et al. 2023). While varied predictions across different tools were observed for some mutations, such as K119E and E147K, most were consistently classified as pathogenic.

Normal expression levels of CNOT3 variants in HEK293 cells were detected by Western blot analysis, suggesting these mutations did not significantly impact protein abundance or stability. The preservation of interactions between missense mutations in CNOT3's N-terminal domain and core Ccr4-Not components CNOT1, CNOT2, and CNOT7 was demonstrated through co-immunoprecipitation experiments. While these mutations did not disrupt the interaction of CNOT3 and CNOT2 with the core Ccr4-Not complex, their involvement in disease pathogenesis might occur through alternative mechanisms. For instance, these mutations could interfere with the ability of the Ccr4-Not complex to bind ribosomes, a process critical for its role in regulating mRNA stability and translation (Absmeier et al. 2023; Zhu et al. 2024).

The missense mutation associated with *CNOT2*, identified as N466H, was evaluated in relation to IDNADFS (Niceta et al. 2023). The Co-

immunoprecipitation experiments revealed that interactions with core Ccr4-Not components, such as CNOT1, CNOT3, and CNOT7, remained intact in both the wild-type and N466H-mutant. Although Niceta et al. (Niceta et al. 2023) demonstrated through structural analyses that missense mutation in the Not-Box region, N466H, would perturb interactions within the Ccr4-Not complex and compromise its stability, the experimental results presented in this study indicated that these interactions remain intact. However, subtle functional effects or impacts on non-core interactions cannot be ruled out and require further investigation to fully understand.

In contrast, truncating mutation in CNOT2, K316Ter, cause significant functional impairment. Analysis using PROVEAN (Choi et al. 2012), a tool validated for predicting the impact of premature stop codons, classified this mutation as deleterious. Co-immunoprecipitation experiments confirmed that this truncation prevents CNOT2 incorporation into the Ccr4-Not complex. The truncating mutation has been strongly associated with severe neurodevelopmental phenotypes, including profound developmental delays, hypotonia, significant speech delays, and behavioural abnormalities, consistent with haploinsufficiency caused by complete loss of function (Uehara et al. 2019; Niceta et al. 2023). The comparison of clinical phenotypes did not reveal clear differences between missense and truncating mutation in CNOT2. The patient with the missense mutation N466H exhibited similar core features to those with truncating mutations, including intellectual disabilities, developmental delays, and distinctive facial features (Uehara et al. 2019; Niceta et al. 2023).

6.5 Future Outlook

The findings of this study open several avenues for future research to expand our understanding of CNOT3 and CNOT2 in cellular regulation and disease mechanisms. These directions include the exploration of molecular pathways, therapeutic targeting, and the mechanistic impacts of disease-associated mutations:

Gene Profiling and Pathway Analysis

Stable cell lines developed in this study provide a platform for gene profiling to uncover pathways directly regulated by CNOT3. RNA sequencing could offer a comprehensive analysis of transcriptome-wide changes, helping to identify specific target genes and pathways modulated by CNOT3 in the context of diseases like T-ALL and neurodevelopmental disorders. By examining gene expression changes following CNOT3 overexpression or knockdown, future studies could delineate the regulatory networks influenced by CNOT3 and evaluate how different variants affect these pathways. This approach may provide mechanistic insights into how CNOT3 dysfunction contributes to disease pathogenesis and identify potential molecular targets for therapeutic intervention.

Investigating Mutational Impacts Beyond Complex Formation

While this study demonstrated that missense mutations in *CNOT3* and *CNOT2* retain interactions with the core Ccr4-Not complex, future work should explore other potential pathogenic mechanisms. These include examining the effects of mutations on protein stability, catalytic activity, and interactions with other proteins or cofactors. Ribosome profiling and co-immunoprecipitation assays could be used to investigate the impact of mutations on ribosome association, particularly given recent evidence of interactions between CNOT3 and ribosomes through its N-terminal domain (Absmeier et al. 2023; Zhu et al. 2024). Understanding how these mutations influence translational control and mRNA metabolism could reveal novel aspects of their involvement in disease

Therapeutic Targeting of CNOT3 Domains

Future research should focus on the therapeutic potential of targeting critical CNOT3 domains, particularly the N-terminal and C-terminal domains, and Not-Box regions. Identifying small molecules that disrupt their function could be utilised to inhibit CNOT3-mediated cell proliferation, offering promising applications in cancer treatment. Aberrant CNOT3 expression has been implicated in tumour formation (Cejas et al. 2017), progression, and drug

resistance in various cancers, including non-small cell lung cancer (NSCLC) (Jing et al. 2019) and renal cancer (Mao et al. 2021). For example, CNOT3 has been shown to contribute to cisplatin resistance and dampen sensitivity to EGFR-TKIs like gefitinib. Targeting CNOT3 could, therefore, not only inhibit tumour growth but also overcome therapeutic resistance, potentially enhancing the efficacy of existing anti-cancer treatments (Jing et al. 2019).

References

- Abelson, J. 2008. 'Is the spliceosome a ribonucleoprotein enzyme?', *Nat Struct Mol Biol*, 15: 1235-7.
- Absmeier, E., V. Chandrasekaran, F. J. O'Reilly, J. A. W. Stowell, J. Rappsilber, and L. A. Passmore. 2023. 'Specific recognition and ubiquitination of translating ribosomes by mammalian CCR4-NOT', *Nat Struct Mol Biol*, 30: 1314-22.
- Adzhubei, I., D. M. Jordan, and S. R. Sunyaev. 2013. 'Predicting functional effect of human missense mutations using PolyPhen-2', *Curr Protoc Hum Genet*, Chapter 7: Unit7 20.
- Ahn, S. H., M. Kim, and S. Buratowski. 2004. 'Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing', *Mol Cell*, 13: 67-76.
- Akashi, H. 1994. 'Synonymous codon usage in *Drosophila melanogaster*: natural selection and translational accuracy', *Genetics*, 136: 927-35.
- Albert, T. K., M. Lemaire, N. L. van Berkum, R. Gentz, M. A. Collart, and H. T. Timmers. 2000. 'Isolation and characterization of human orthologs of yeast CCR4-NOT complex subunits', *Nucleic Acids Res*, 28: 809-17.
- Almalki, S. G. 2023. 'The pathophysiology of the cell cycle in cancer and treatment strategies using various cell cycle checkpoint inhibitors', *Pathol Res Pract*, 251: 154854.
- Aslam, A., S. Mittal, F. Koch, J. C. Andrau, and G. S. Winkler. 2009. 'The Ccr4-NOT deadenylase subunits CNOT7 and CNOT8 have overlapping roles and modulate cell proliferation', *Mol Biol Cell*, 20: 3840-50.
- Barrington, Chloe L., Gabriel Galindo, Amanda L. Koch, Emma R. Horton, Evan J. Morrison, Samantha Tisa, Timothy J. Stasevich, and Olivia S. Rissland. 2023. 'Synonymous codon usage regulates translation initiation', *Cell Reports*, 42.
- Bawankar, P., B. Loh, L. Wohlbold, S. Schmidt, and E. Izaurralde. 2013. 'NOT10 and C2orf29/NOT11 form a conserved module of the CCR4-NOT complex that docks onto the NOT1 N-terminal domain', *RNA Biol*, 10: 228-44.
- Berger, S. L., T. Kouzarides, R. Shiekhattar, and A. Shilatifard. 2009. 'An operational definition of epigenetics', *Genes & development*, 23: 781-3.
- Bergkessel, M., and C. Guthrie. 2013. 'Colony PCR', *Methods Enzymol*, 529: 299-309.
- Boland, A., Y. Chen, T. Raisch, S. Jonas, D. Kuzuoglu-Ozturk, L. Wohlbold, O. Weichenrieder, and E. Izaurralde. 2013. 'Structure and assembly of the

- NOT module of the human CCR4-NOT complex', *Nat Struct Mol Biol*, 20: 1289-97.
- Braun, Joerg E., Larry J. Friedman, Jeff Gelles, and Melissa J. Moore. 2018. 'Synergistic assembly of human pre-spliceosomes across introns and exons', *eLife*, 7.
- Buschauer, R., Y. Matsuo, T. Sugiyama, Y. H. Chen, N. Alhusaini, T. Sweet, K. Ikeuchi, J. Cheng, Y. Matsuki, R. Nobuta, A. Gilmozzi, O. Berninghausen, P. Tesina, T. Becker, J. Collier, T. Inada, and R. Beckmann. 2020. 'The Ccr4-Not complex monitors the translating ribosome for codon optimality', *Science*, 368.
- Cejas, P., A. Cavazza, C. N. Yandava, V. Moreno, D. Horst, J. Moreno-Rubio, E. Burgos, M. Mendiola, L. Taing, A. Goel, J. Feliu, and R. A. Shivdasani. 2017. 'Transcriptional Regulator CNOT3 Defines an Aggressive Colorectal Cancer Subtype', *Cancer Res*, 77: 766-79.
- Chalabi Hagkarim, N., and R. J. Grand. 2020. 'The Regulatory Properties of the Ccr4-Not Complex', *Cells*, 9.
- Chen, C. Y., and A. B. Shyu. 2011. 'Mechanisms of deadenylation-dependent decay', *Wiley Interdiscip Rev RNA*, 2: 167-83.
- Cheng, J., G. Novati, J. Pan, C. Bycroft, A. Zemgulyte, T. Applebaum, A. Pritzel, L. H. Wong, M. Zielinski, T. Sargeant, R. G. Schneider, A. W. Senior, J. Jumper, D. Hassabis, P. Kohli, and Z. Avsec. 2023. 'Accurate proteome-wide missense variant effect prediction with AlphaMissense', *Science*, 381: eadg7492.
- Chicoine, J., P. Benoit, C. Gamberi, M. Paliouras, M. Simonelig, and P. Lasko. 2007. 'Bicaudal-C recruits CCR4-NOT deadenylase to target mRNAs and regulates oogenesis, cytoskeletal organization, and its own expression', *Dev Cell*, 13: 691-704.
- Choi, Y., G. E. Sims, S. Murphy, J. R. Miller, and A. P. Chan. 2012. 'Predicting the functional effect of amino acid substitutions and indels', *PLoS One*, 7: e46688.
- Colgan, D. F., and J. L. Manley. 1997. 'Mechanism and regulation of mRNA polyadenylation', *Genes & development*, 11: 2755-66.
- Collart, M. A. 2003. 'Global control of gene expression in yeast by the Ccr4-Not complex', *Gene*, 313: 1-16.
- . 2016. 'The Ccr4 Not complex is a key regulator of eukaryotic gene expression', *Wiley Interdiscip Rev RNA*, 7: 438-54.
- Collart, M. A., L. Audebert, and M. Bushell. 2023. 'Roles of the CCR4-Not complex in translation and dynamics of co-translation events', *Wiley Interdiscip Rev RNA*, 15: e1827.
- Collart, M. A., and O. O. Panasenko. 2012. 'The Ccr4-not complex', *Gene*, 492: 42-53.
- . 2017. 'The Ccr4-Not Complex: Architecture and Structural Insights', *Subcell Biochem*, 83: 349-79.

- Coller, J., and R. Parker. 2004. 'Eukaryotic mRNA decapping', *Annu Rev Biochem*, 73: 861-90.
- Connelly, S., and J. L. Manley. 1988. 'A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II', *Genes & development*, 2: 440-52.
- Conti, E., and E. Izaurralde. 2005. 'Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species', *Curr Opin Cell Biol*, 17: 316-25.
- Dahan, O., H. Gingold, and Y. Pilpel. 2011. 'Regulatory mechanisms and networks couple the different phases of gene expression', *Trends Genet*, 27: 316-22.
- Danna, K., and D. Nathans. 1971. 'Specific cleavage of simian virus 40 DNA by restriction endonuclease of *Hemophilus influenzae*', *Proc Natl Acad Sci U S A*, 68: 2913-7.
- Darnell, J. E., Jr. 1978. 'Implications of RNA-RNA splicing in evolution of eukaryotic cells', *Science*, 202: 1257-60.
- De Keersmaecker, K., Z. K. Atak, N. Li, C. Vicente, S. Patchett, T. Girardi, V. Gianfelici, E. Geerdens, E. Clappier, M. Porcu, I. Lahortiga, R. Luca, J. Yan, G. Hulselmans, H. Vranckx, R. Vandepoel, B. Sweron, K. Jacobs, N. Mentens, I. Wlodarska, B. Cauwelier, J. Cloos, J. Soulier, A. Uyttebroeck, C. Bagni, B. A. Hassan, P. Vandenberghe, A. W. Johnson, S. Aerts, and J. Cools. 2013. 'Exome sequencing identifies mutation in CNOT3 and ribosomal genes RPL5 and RPL10 in T-cell acute lymphoblastic leukemia', *Nat Genet*, 45: 186-90.
- Di Felice, Francesca, Gioacchino Micheli, and Giorgio Camilloni. 2019. 'Restriction enzymes and their use in molecular biology: An overview', *Journal of Biosciences*, 44.
- Doma, M. K., and R. Parker. 2006. 'Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation', *Nature*, 440: 561-4.
- Dwaraka, V. B., L. Aronica, N. Carreras-Gallo, J. L. Robinson, T. Hennings, M. M. Carter, M. J. Corley, A. Lin, L. Turner, R. Smith, T. L. Mendez, H. Went, E. R. Ebel, E. D. Sonnenburg, J. L. Sonnenburg, and C. D. Gardner. 2024. 'Unveiling the epigenetic impact of vegan vs. omnivorous diets on aging: insights from the Twins Nutrition Study (TwINS)', *BMC Med*, 22: 301.
- Eaton, J. D., and S. West. 2020. 'Termination of Transcription by RNA Polymerase II: BOOM!', *Trends Genet*, 36: 664-75.
- Farber, V., E. Erben, S. Sharma, G. Stoecklin, and C. Clayton. 2013. 'Trypanosome CNOT10 is essential for the integrity of the NOT deadenylase complex and for degradation of many mRNAs', *Nucleic Acids Res*, 41: 1211-22.

- Feoktistova, M., P. Geserick, and M. Leverkus. 2016. 'Crystal Violet Assay for Determining Viability of Cultured Cells', *Cold Spring Harb Protoc*, 2016: pdb prot087379.
- Fishburn, J., E. Tomko, E. Galburt, and S. Hahn. 2015. 'Double-stranded DNA translocase activity of transcription factor TFIIH and the mechanism of RNA polymerase II open complex formation', *Proc Natl Acad Sci U S A*, 112: 3961-6.
- Frischmeyer, P. A., A. van Hoof, K. O'Donnell, A. L. Guerrierio, R. Parker, and H. C. Dietz. 2002. 'An mRNA surveillance mechanism that eliminates transcripts lacking termination codons', *Science*, 295: 2258-61.
- Garneau, N. L., J. Wilusz, and C. J. Wilusz. 2007. 'The highways and byways of mRNA decay', *Nat Rev Mol Cell Biol*, 8: 113-26.
- Gatfield, D., and E. Izaurralde. 2004. 'Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in *Drosophila*', *Nature*, 429: 575-8.
- Ghashghaei, M., Y. Liu, J. Ettles, G. Bombaci, N. Ramkumar, Z. Liu, L. Escano, S. S. Miko, Y. Kim, J. A. Waldron, K. Do, K. MacPherson, K. A. Yuen, T. Taibi, M. Yue, A. Arsalan, Z. Jin, G. Edin, A. Karsan, G. B. Morin, F. Kuchenbauer, F. Perna, M. Bushell, and L. P. Vu. 2024. 'Translation efficiency driven by CNOT3 subunit of the CCR4-NOT complex promotes leukemogenesis', *Nat Commun*, 15: 2340.
- Ghosh, A., and C. D. Lima. 2010. 'Enzymology of RNA cap synthesis', *Wiley Interdiscip Rev RNA*, 1: 152-72.
- Gingold, H., and Y. Pilpel. 2011. 'Determinants of translation efficiency and accuracy', *Mol Syst Biol*, 7: 481.
- Gingold, H., D. Tehler, N. R. Christoffersen, M. M. Nielsen, F. Asmar, S. M. Kooistra, N. S. Christophersen, L. L. Christensen, M. Borre, K. D. Sorensen, L. D. Andersen, C. L. Andersen, E. Hulleman, T. Wurdinger, E. Ralfkiaer, K. Helin, K. Gronbaek, T. Orntoft, S. M. Waszak, O. Dahan, J. S. Pedersen, A. H. Lund, and Y. Pilpel. 2014. 'A dual program for translation regulation in cellular proliferation and differentiation', *Cell*, 158: 1281-92.
- Gordon, D. E., A. Watson, A. Roguev, S. Zheng, G. M. Jang, J. Kane, J. Xu, J. Z. Guo, E. Stevenson, D. L. Swaney, K. Franks-Skiba, E. Verschueren, M. Shales, D. C. Crosby, A. D. Frankel, A. Marson, I. Marazzi, G. Cagney, and N. J. Krogan. 2020. 'A Quantitative Genetic Interaction Map of HIV Infection', *Mol Cell*, 78: 197-209 e7.
- Gorgoni, B., and N. K. Gray. 2004. 'The roles of cytoplasmic poly(A)-binding proteins in regulating gene expression: a developmental perspective', *Brief Funct Genomic Proteomic*, 3: 125-41.
- Graifer, D., and G. Karpova. 2012. 'Structural and functional topography of the human ribosome', *Acta Biochim Biophys Sin (Shanghai)*, 44: 281-99.

- Green, M. R., and J. Sambrook. 2019. 'Inverse Polymerase Chain Reaction (PCR)', *Cold Spring Harb Protoc*, 2019.
- Han, J., J. Xiong, D. Wang, and X. D. Fu. 2011. 'Pre-mRNA splicing: where and when in the nucleus', *Trends Cell Biol*, 21: 336-43.
- Hanahan, D. 2022. 'Hallmarks of Cancer: New Dimensions', *Cancer Discov*, 12: 31-46.
- Hollien, J., and J. S. Weissman. 2006. 'Decay of endoplasmic reticulum-localized mRNAs during the unfolded protein response', *Science*, 313: 104-7.
- Hsu, C. L., and A. Stevens. 1993. 'Yeast cells lacking 5'→3' exoribonuclease 1 contain mRNA species that are poly(A) deficient and partially lack the 5' cap structure', *Mol Cell Biol*, 13: 4826-35.
- Hu, G., J. Kim, Q. Xu, Y. Leng, S. H. Orkin, and S. J. Elledge. 2009. 'A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal', *Genes & development*, 23: 837-48.
- Hug, N., D. Longman, and J. F. Cáceres. 2016. 'Mechanism and regulation of the nonsense-mediated decay pathway', *Nucleic Acids Res*, 44: 1483-95.
- Inada, T., and H. Aiba. 2005. 'Translation of aberrant mRNAs lacking a termination codon or with a shortened 3'-UTR is repressed after initiation in yeast', *Embo j*, 24: 1584-95.
- Ito, K., T. Inoue, K. Yokoyama, M. Morita, T. Suzuki, and T. Yamamoto. 2011. 'CNOT2 depletion disrupts and inhibits the CCR4-NOT deadenylase complex and induces apoptotic cell death', *Genes Cells*, 16: 368-79.
- Ito, K., A. Takahashi, M. Morita, T. Suzuki, and T. Yamamoto. 2011. 'The role of the CNOT1 subunit of the CCR4-NOT complex in mRNA deadenylation and cell viability', *Protein Cell*, 2: 755-63.
- Jackson, R. J., C. U. Hellen, and T. V. Pestova. 2010. 'The mechanism of eukaryotic translation initiation and principles of its regulation', *Nat Rev Mol Cell Biol*, 11: 113-27.
- Jamar, N. H., P. Kritsiligkou, and C. M. Grant. 2018. 'Loss of mRNA surveillance pathways results in widespread protein aggregation', *Sci Rep*, 8: 3894.
- Jia, X., X. He, C. Huang, J. Li, Z. Dong, and K. Liu. 2024. 'Protein translation: biological processes and therapeutic strategies for human diseases', *Signal Transduct Target Ther*, 9: 44.
- Jiang, Yan, and David H. Price. 2014. 'Rescue of the TTF2 Knockdown Phenotype with an siRNA-Resistant Replacement Vector', *Cell Cycle*, 3: 1149-51.
- Jinek, M., S. M. Coyle, and J. A. Doudna. 2011. 'Coupled 5' nucleotide recognition and processivity in Xrn1-mediated mRNA decay', *Mol Cell*, 41: 600-8.
- Jing, L., M. E. Zhai, J. Cui, X. Y. Fan, Y. Y. Cheng, J. L. Jiang, and Z. N. Chen. 2019. 'CNOT3 contributes to cisplatin resistance in lung cancer through inhibiting RIPK3 expression', *Apoptosis*, 24: 673-85.

- Jonas, S., M. Christie, D. Peter, D. Bhandari, B. Loh, E. Huntzinger, O. Weichenrieder, and E. Izaurralde. 2014. 'An asymmetric PAN3 dimer recruits a single PAN2 exonuclease to mediate mRNA deadenylation and decay', *Nat Struct Mol Biol*, 21: 599-608.
- Jonkers, I., and J. T. Lis. 2015. 'Getting up to speed with transcription elongation by RNA polymerase II', *Nat Rev Mol Cell Biol*, 16: 167-77.
- Jung, J. H., D. Lee, H. M. Ko, and H. J. Jang. 2021. 'Inhibition of CNOT2 Induces Apoptosis via MID1IP1 in Colorectal Cancer Cells by Activating p53', *Biomolecules*, 11.
- Kashima, I., A. Yamashita, N. Izumi, N. Kataoka, R. Morishita, S. Hoshino, M. Ohno, G. Dreyfuss, and S. Ohno. 2006. 'Binding of a novel SMG-1-Upfl-eRF1-eRF3 complex (SURF) to the exon junction complex triggers Upfl phosphorylation and nonsense-mediated mRNA decay', *Genes & development*, 20: 355-67.
- Kashiwagi, K., M. Takahashi, M. Nishimoto, T. B. Hiyama, T. Higo, T. Umehara, K. Sakamoto, T. Ito, and S. Yokoyama. 2016. 'Crystal structure of eukaryotic translation initiation factor 2B', *Nature*, 531: 122-5.
- Kielkopf, C. L., W. Bauer, and I. L. Urbatsch. 2020. 'Bradford Assay for Determining Protein Concentration', *Cold Spring Harb Protoc*, 2020: 102269.
- Kim, Y., J. H. Geiger, S. Hahn, and P. B. Sigler. 1993. 'Crystal structure of a yeast TBP/TATA-box complex', *Nature*, 365: 512-20.
- Komar, A. A. 2007. 'Silent SNPs: impact on gene function and phenotype', *Pharmacogenomics*, 8: 1075-80.
- . 2016. 'The Yin and Yang of codon usage', *Hum Mol Genet*, 25: R77-R85.
- Korner, C. G., and E. Wahle. 1997. 'Poly(A) tail shortening by a mammalian poly(A)-specific 3'-exoribonuclease', *J Biol Chem*, 272: 10448-56.
- Kozak, M. 1989. 'The scanning model for translation: an update', *J Cell Biol*, 108: 229-41.
- Kumar, P., S. Henikoff, and P. C. Ng. 2009. 'Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm', *Nat Protoc*, 4: 1073-81.
- Kwak, H., and J. T. Lis. 2013. 'Control of transcriptional elongation', *Annu Rev Genet*, 47: 483-508.
- Lau, N. C., A. Kolkman, F. M. van Schaik, K. W. Mulder, W. W. Pijnappel, A. J. Heck, and H. T. Timmers. 2009. 'Human Ccr4-Not complexes contain variable deadenylase subunits', *Biochem J*, 422: 443-53.
- Lee, T. I., and R. A. Young. 2013. 'Transcriptional regulation and its misregulation in disease', *Cell*, 152: 1237-51.
- Leppek, K., R. Das, and M. Barna. 2018. 'Functional 5' UTR mRNA structures in eukaryotic translation regulation and how to find them', *Nat Rev Mol Cell Biol*, 19: 158-74.

- Li, Y., and M. Kiledjian. 2010. 'Regulation of mRNA decapping', *Wiley Interdiscip Rev RNA*, 1: 253-65.
- Lin, J. S., and E. M. Lai. 2017. 'Protein-Protein Interactions: Co-Immunoprecipitation', *Methods Mol Biol*, 1615: 211-19.
- Liu, H., and M. Kiledjian. 2006. 'Decapping the message: a beginning or an end', *Biochem Soc Trans*, 34: 35-8.
- Liu, J., X. Lu, S. Zhang, L. Yuan, and Y. Sun. 2022. 'Molecular Insights into mRNA Polyadenylation and Deadenylation', *Int J Mol Sci*, 23.
- Ljungdahl, A., S. Kohani, N. F. Page, E. S. Wells, E. M. Wigdor, S. Dong, and S. J. Sanders. 2023. 'AlphaMissense is better correlated with functional assays of missense impact than earlier prediction algorithms', *bioRxiv*.
- Logan, J., E. Falck-Pedersen, J. E. Darnell, Jr., and T. Shenk. 1987. 'A poly(A) addition site and a downstream termination region are required for efficient cessation of transcription by RNA polymerase II in the mouse beta maj-globin gene', *Proc Natl Acad Sci U S A*, 84: 8306-10.
- Mao, Q., Q. Zhuang, J. Shen, Z. Chen, D. Xue, T. Ding, and X. He. 2021. 'MiRNA-124 regulates the sensitivity of renal cancer cells to cisplatin-induced necroptosis by targeting the CAPN4-CNOT3 axis', *Transl Androl Urol*, 10: 3669-83.
- Martin, R., M. Splitt, D. Genevieve, E. Aten, A. Collins, C. I. de Bie, L. Faivre, N. Foulds, J. Giltay, R. Ibitoye, S. Joss, J. Kennedy, B. Kerr, E. Kivuva, M. Koopmans, R. Newbury-Ecob, N. Jean-Marçais, E. A. J. Peeters, S. Smithson, S. Tomkins, F. Tranmauthem, A. Piton, and A. van Haeringen. 2019. 'De novo variants in CNOT3 cause a variable neurodevelopmental disorder', *Eur J Hum Genet*, 27: 1677-82.
- Maryati, M., B. Airhihen, and G. S. Winkler. 2015. 'The enzyme activities of Caf1 and Ccr4 are both required for deadenylation by the human Ccr4-Not nuclease module', *Biochem J*, 469: 169-76.
- Marzluff, W. F., E. J. Wagner, and R. J. Duronio. 2008. 'Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail', *Nat Rev Genet*, 9: 843-54.
- Mauxion, F., J. Basquin, S. Ozgur, M. Rame, J. Albrecht, I. Schafer, B. Seraphin, and E. Conti. 2023. 'The human CNOT1-CNOT10-CNOT11 complex forms a structural platform for protein-protein interactions', *Cell Rep*, 42: 111902.
- Medema, R. H. 2004. 'Optimizing RNA interference for application in mammalian cells', *Biochem J*, 380: 593-603.
- Melero, R., A. Uchiyama, R. Castano, N. Kataoka, H. Kurosawa, S. Ohno, A. Yamashita, and O. Llorca. 2014. 'Structures of SMG1-UPFs complexes: SMG1 contributes to regulate UPF2-dependent activation of UPF1 in NMD', *Structure*, 22: 1105-19.

- Meng, E. C., T. D. Goddard, E. F. Pettersen, G. S. Couch, Z. J. Pearson, J. H. Morris, and T. E. Ferrin. 2023. 'UCSF ChimeraX: Tools for structure building and analysis', *Protein Sci*, 32: e4792.
- Miller, J. E., and J. C. Reese. 2012. 'Ccr4-Not complex: the control freak of eukaryotic cells', *Crit Rev Biochem Mol Biol*, 47: 315-33.
- Mittal, S., A. Aslam, R. Doidge, R. Medica, and G. S. Winkler. 2011. 'The Ccr4a (CNOT6) and Ccr4b (CNOT6L) deadenylase subunits of the human Ccr4-Not complex contribute to the prevention of cell death and senescence', *Mol Biol Cell*, 22: 748-58.
- Morita, E., J. Arai, D. Christensen, J. Votteler, and W. I. Sundquist. 2012. 'Attenuated protein expression vectors for use in siRNA rescue experiments', *Biotechniques*, 0: 1-5.
- Morita, M., Y. Oike, T. Nagashima, T. Kadomatsu, M. Tabata, T. Suzuki, T. Nakamura, N. Yoshida, M. Okada, and T. Yamamoto. 2011. 'Obesity resistance and increased hepatic expression of catabolism-related mRNAs in Cnot3^{+/-} mice', *Embo j*, 30: 4678-91.
- Mostafa, D., A. Takahashi, A. Yanagiya, T. Yamaguchi, T. Abe, T. Kureha, K. Kuba, Y. Kanegae, Y. Furuta, T. Yamamoto, and T. Suzuki. 2020. 'Essential functions of the CNOT7/8 catalytic subunits of the CCR4-NOT complex in mRNA regulation and cell viability', *RNA Biol*, 17: 403-16.
- Nabet, B., J. M. Roberts, D. L. Buckley, J. Paulk, S. Dastjerdi, A. Yang, A. L. Leggett, M. A. Erb, M. A. Lawlor, A. Souza, T. G. Scott, S. Vittori, J. A. Perry, J. Qi, G. E. Winter, K. K. Wong, N. S. Gray, and J. E. Bradner. 2018. 'The dTAG system for immediate and target-specific protein degradation', *Nat Chem Biol*, 14: 431-41.
- Nagy, Vivien, and Manfred Watzel. 2006. 'FuGENE® 6 Transfection Reagent: minimizing reagent-dependent side effects as analyzed by gene-expression profiling and cytotoxicity assays', *Nature Methods*, 3: iii-v.
- Nasertorabi, F., C. Batisse, M. Diepholz, D. Suck, and B. Bottcher. 2011. 'Insights into the structure of the CCR4-NOT complex by electron microscopy', *FEBS Lett*, 585: 2182-6.
- Neely, G. G., K. Kuba, A. Cammarato, K. Isobe, S. Amann, L. Zhang, M. Murata, L. Elmen, V. Gupta, S. Arora, R. Sarangi, D. Dan, S. Fujisawa, T. Usami, C. P. Xia, A. C. Keene, N. N. Alayari, H. Yamakawa, U. Elling, C. Berger, M. Novatchkova, R. Kogelgruber, K. Fukuda, H. Nishina, M. Isobe, J. A. Pospisilik, Y. Imai, A. Pfeufer, A. A. Hicks, P. P. Pramstaller, S. Subramaniam, A. Kimura, K. Ocorr, R. Bodmer, and J. M. Penninger. 2010. 'A global in vivo Drosophila RNAi screen identifies NOT3 as a conserved regulator of heart function', *Cell*, 141: 142-53.
- Niceta, M., S. Pizzi, F. Inzana, A. Peron, S. Bakhtiari, M. Nizon, J. Levy, C. Mancini, B. Cogne, F. C. Radio, E. Agolini, D. Cocciadiferro, A. Novelli, M. A. Salih, M. P. Recalcati, R. Arancio, M. Besnard, A. C. Tabet, M. C.

- Kruer, M. Priolo, B. Dallapiccola, and M. Tartaglia. 2023. 'Delineation of the clinical profile of CNOT2 haploinsufficiency and overview of the IDNADFS phenotype', *Clin Genet*, 103: 156-66.
- Nilsson, P., N. Henriksson, A. Niedzwiecka, N. A. Balatsos, K. Kokkoris, J. Eriksson, and A. Virtanen. 2007. 'A multifunctional RNA recognition motif in poly(A)-specific ribonuclease with cap and poly(A) binding properties', *J Biol Chem*, 282: 32902-11.
- Nowakowski, A. B., W. J. Wobig, and D. H. Petering. 2014. 'Native SDS-PAGE: high resolution electrophoretic separation of proteins with retention of native properties including bound metal ions', *Metallomics*, 6: 1068-78.
- O'Brien, J., I. Wilson, T. Orton, and F. Pognan. 2003. 'Investigation of the Alamar Blue (resazurin) fluorescent dye for the assessment of mammalian cell cytotoxicity', *Eur J Biochem*, 267: 5421-6.
- O'Gorman, S., D. T. Fox, and G. M. Wahl. 1991. 'Recombinase-mediated gene activation and site-specific integration in mammalian cells', *Science*, 251: 1351-5.
- Onouchi, H., Y. Nagami, Y. Haraguchi, M. Nakamoto, Y. Nishimura, R. Sakurai, N. Nagao, D. Kawasaki, Y. Kadokura, and S. Naito. 2005. 'Nascent peptide-mediated translation elongation arrest coupled with mRNA degradation in the CGS1 gene of Arabidopsis', *Genes & development*, 19: 1799-810.
- Palazzo, A. F., Y. Qiu, and Y. M. Kang. 2024. 'mRNA nuclear export: how mRNA identity features distinguish functional RNAs from junk transcripts', *RNA Biol*, 21: 1-12.
- Park, J., M. Kim, H. Yi, K. Baeg, Y. Choi, Y. S. Lee, J. Lim, and V. N. Kim. 2023. 'Short poly(A) tails are protected from deadenylation by the LARP1-PABP complex', *Nat Struct Mol Biol*, 30: 330-38.
- Parker, R. 2012. 'RNA degradation in *Saccharomyces cerevisiae*', *Genetics*, 191: 671-702.
- Parker, R., and H. Song. 2004. 'The enzymes and control of eukaryotic mRNA turnover', *Nat Struct Mol Biol*, 11: 121-7.
- Passmore, L. A., and J. Coller. 2022. 'Roles of mRNA poly(A) tails in regulation of eukaryotic gene expression', *Nat Rev Mol Cell Biol*, 23: 93-106.
- Pavanello, L., M. Hall, and G. S. Winkler. 2023. 'Regulation of eukaryotic mRNA deadenylation and degradation by the Ccr4-Not complex', *Front Cell Dev Biol*, 11: 1153624.
- Pedersen, S. 1984. 'Escherichia coli ribosomes translate in vivo with variable rate', *Embo j*, 3: 2895-8.
- Pelletier, J., G. Thomas, and S. Volarevic. 2018. 'Ribosome biogenesis in cancer: new players and therapeutic avenues', *Nat Rev Cancer*, 18: 51-63.
- Penner-Goeke, S., and E. B. Binder. 2024. 'Linking environmental factors and gene regulation', *Elife*, 13.

- Persson, P. B., and M. Mueller. 2015. 'Transcription', *Acta Physiol (Oxf)*, 215: 159-60.
- Plotkin, J. B., and G. Kudla. 2011. 'Synonymous but not the same: the causes and consequences of codon bias', *Nat Rev Genet*, 12: 32-42.
- Polacek, N., and A. S. Mankin. 2005. 'The ribosomal peptidyl transferase center: structure, function, evolution, inhibition', *Crit Rev Biochem Mol Biol*, 40: 285-311.
- Pollex, T., A. Rabinowitz, M. C. Gambetta, R. Marco-Ferreres, R. R. Viales, A. Jankowski, C. Schaub, and E. E. M. Furlong. 2024. 'Enhancer-promoter interactions become more instructive in the transition from cell-fate specification to tissue differentiation', *Nat Genet*, 56: 686-96.
- Presnyak, V., N. Alhusaini, Y. H. Chen, S. Martin, N. Morris, N. Kline, S. Olson, D. Weinberg, K. E. Baker, B. R. Graveley, and J. Collier. 2015. 'Codon optimality is a major determinant of mRNA stability', *Cell*, 160: 1111-24.
- Proudfoot, N. J. 2011. 'Ending the message: poly(A) signals then and now', *Genes & development*, 25: 1770-82.
- Quax, T. E., N. J. Claassens, D. Soll, and J. van der Oost. 2015. 'Codon Bias as a Means to Fine-Tune Gene Expression', *Mol Cell*, 59: 149-61.
- Rahman, M. F., and P. O. McGowan. 2022. 'Cell-type-specific epigenetic effects of early life stress on the brain', *Transl Psychiatry*, 12: 326.
- Raisch, T., C. T. Chang, Y. Levdansky, S. Muthukumar, S. Raunser, and E. Valkov. 2019. 'Reconstitution of recombinant human CCR4-NOT reveals molecular insights into regulated deadenylation', *Nat Commun*, 10: 3173.
- Ramanathan, A., G. B. Robb, and S. H. Chan. 2016. 'mRNA capping: biological functions and applications', *Nucleic Acids Res*, 44: 7511-26.
- Remenyi, A., H. R. Scholer, and M. Wilmanns. 2004. 'Combinatorial control of gene expression', *Nat Struct Mol Biol*, 11: 812-5.
- Sakai, A., T. Chibazakura, Y. Shimizu, and F. Hishinuma. 1992. 'Molecular analysis of POP2 gene, a gene required for glucose-derepression of gene expression in *Saccharomyces cerevisiae*', *Nucleic Acids Res*, 20: 6227-33.
- Sarkar, M., M. Martufi, M. Roman-Trufero, Y. F. Wang, C. Whilding, D. Dormann, P. Sabbattini, and N. Dillon. 2021. 'CNOT3 interacts with the Aurora B and MAPK/ERK kinases to promote survival of differentiating mesendodermal progenitor cells', *Mol Biol Cell*, 32: ar40.
- Sauer, Brian. 1994. 'Site-specific recombination: developments and applications', *Current Opinion in Biotechnology*, 5: 521-27.
- Schaeffer, D., B. Tsanova, A. Barbas, F. P. Reis, E. G. Dastidar, M. Sanchez-Rotunno, C. M. Arraiano, and A. van Hoof. 2009. 'The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities', *Nat Struct Mol Biol*, 16: 56-62.

- Schafer, I. B., M. Rode, F. Bonneau, S. Schussler, and E. Conti. 2014. 'The structure of the Pan2-Pan3 core complex reveals cross-talk between deadenylase and pseudokinase', *Nat Struct Mol Biol*, 21: 591-8.
- Schafer, I. B., M. Yamashita, J. M. Schuller, S. Schussler, P. Reichelt, M. Strauss, and E. Conti. 2019. 'Molecular Basis for poly(A) RNP Architecture and Recognition by the Pan2-Pan3 Deadenylase', *Cell*, 177: 1619-31 e21.
- Schlake, T., and J. Bode. 1994. 'Use of mutated FLP recognition target (FRT) sites for the exchange of expression cassettes at defined chromosomal loci', *Biochemistry*, 33: 12746-51.
- Schoenberg, D. R. 2011. 'Mechanisms of endonuclease-mediated mRNA decay', *Wiley Interdiscip Rev RNA*, 2: 582-600.
- Schoenberg, D. R., and L. E. Maquat. 2012. 'Regulation of cytoplasmic mRNA decay', *Nat Rev Genet*, 13: 246-59.
- Sharif, H., and E. Conti. 2013. 'Architecture of the Lsm1-7-Pat1 complex: a conserved assembly in eukaryotic mRNA turnover', *Cell Rep*, 5: 283-91.
- Sharma, S., S. Kapoor, A. Ansari, and A. K. Tyagi. 2024. 'The general transcription factors (GTFs) of RNA polymerase II and their roles in plant development and stress responses', *Crit Rev Biochem Mol Biol*, 59: 267-309.
- Shatkin, A. J. 1976. 'Capping of eucaryotic mRNAs', *Cell*, 9: 645-53.
- She, M., C. J. Decker, D. I. Svergun, A. Round, N. Chen, D. Muhlrads, R. Parker, and H. Song. 2008. 'Structural basis of dcp2 recognition and activation by dcp1', *Mol Cell*, 29: 337-49.
- Shirai, Y. T., A. Mizutani, S. Nishijima, M. Horie, C. Kikuguchi, O. Elisseeva, and T. Yamamoto. 2019. 'CNOT3 targets negative cell cycle regulators in non-small cell lung cancer development', *Oncogene*, 38: 2580-94.
- Shirai, Y. T., T. Suzuki, M. Morita, A. Takahashi, and T. Yamamoto. 2014. 'Multifunctional roles of the mammalian CCR4-NOT complex in physiological phenomena', *Front Genet*, 5: 286.
- Silva, D., G. Santos, M. Barroca, and T. Collins. 2017. 'Inverse PCR for Point Mutation Introduction', *Methods Mol Biol*, 1620: 87-100.
- Smale, S. T. 1997. 'Transcription initiation from TATA-less promoters within eukaryotic protein-coding genes', *Biochim Biophys Acta*, 1351: 73-88.
- Sohn, E. J., D. B. Jung, H. Lee, I. Han, J. Lee, H. Lee, and S. H. Kim. 2018. 'CNOT2 promotes proliferation and angiogenesis via VEGF signaling in MDA-MB-231 breast cancer cells', *Cancer Lett*, 412: 88-98.
- Sonenberg, N., and A. G. Hinnebusch. 2009. 'Regulation of translation initiation in eukaryotes: mechanisms and biological targets', *Cell*, 136: 731-45.
- Sorensen, M. A., and S. Pedersen. 1991. 'Absolute in vivo translation rates of individual codons in Escherichia coli. The two glutamic acid codons GAA and GAG are translated with a threefold difference in rate', *J Mol Biol*, 222: 265-80.
- Spirin, A. S. 2002. 'Ribosome as a molecular machine', *FEBS Lett*, 514: 2-10.

- Stephens, K. E., C. A. Miaskowski, J. D. Levine, C. R. Pullinger, and B. E. Aouizerat. 2013. 'Epigenetic regulation and measurement of epigenetic changes', *Biol Res Nurs*, 15: 373-81.
- Stewart, M. 2010. 'Nuclear export of mRNA', *Trends Biochem Sci*, 35: 609-17.
- Suzuki, T., C. Kikuguchi, S. Sharma, T. Sasaki, M. Tokumasu, S. Adachi, T. Natsume, Y. Kanegae, and T. Yamamoto. 2015. 'CNOT3 suppression promotes necroptosis by stabilizing mRNAs for cell death-inducing proteins', *Sci Rep*, 5: 14779.
- Takahashi, A., C. Kikuguchi, M. Morita, T. Shimodaira, N. Tokai-Nishizumi, K. Yokoyama, M. Ohsugi, T. Suzuki, and T. Yamamoto. 2012. 'Involvement of CNOT3 in mitotic progression through inhibition of MAD1 expression', *Biochem Biophys Res Commun*, 419: 268-73.
- Taube, R., X. Lin, D. Irwin, K. Fujinaga, and B. M. Peterlin. 2002. 'Interaction between P-TEFb and the C-terminal domain of RNA polymerase II activates transcriptional elongation from sites upstream or downstream of target genes', *Mol Cell Biol*, 22: 321-31.
- Tharun, S., and R. Parker. 2001. 'Targeting an mRNA for decapping: displacement of translation factors and association of the Lsm1p-7p complex on deadenylated yeast mRNAs', *Mol Cell*, 8: 1075-83.
- Thomas, M. C., and C. M. Chiang. 2006. 'The general transcription machinery and general cofactors', *Crit Rev Biochem Mol Biol*, 41: 105-78.
- Tucker, M., and R. Parker. 2000. 'Mechanisms and control of mRNA decapping in *Saccharomyces cerevisiae*', *Annu Rev Biochem*, 69: 571-95.
- Tucker, M., R. R. Staples, M. A. Valencia-Sanchez, D. Muhlrads, and R. Parker. 2002. 'Ccr4p is the catalytic subunit of a Ccr4p/Pop2p/Notp mRNA deadenylase complex in *Saccharomyces cerevisiae*', *Embo j*, 21: 1427-36.
- Uehara, T., T. Tsuchihashi, M. Yamada, H. Suzuki, T. Takenouchi, and K. Kosaki. 2019. 'CNOT2 haploinsufficiency causes a neurodevelopmental disorder with characteristic facial features', *Am J Med Genet A*, 179: 2506-09.
- van Hoof, A., and E. J. Wagner. 2011. 'A brief survey of mRNA surveillance', *Trends Biochem Sci*, 36: 585-92.
- Venturini, G., A. M. Rose, A. Z. Shah, S. S. Bhattacharya, and C. Rivolta. 2012. 'CNOT3 is a modifier of PRPF31 mutations in retinitis pigmentosa with incomplete penetrance', *PLoS Genet*, 8: e1003040.
- Vicens, Q., J. S. Kieft, and O. S. Rissland. 2018. 'Revisiting the Closed-Loop Model and the Nature of mRNA 5'-3' Communication', *Mol Cell*, 72: 805-12.
- Vicente, C., R. Stirparo, S. Demeyer, C. E. de Bock, O. Gielen, M. Atkins, J. Yan, G. Halder, B. A. Hassan, and J. Cools. 2018. 'The CCR4-NOT complex is a tumor suppressor in *Drosophila melanogaster* eye cancer models', *J Hematol Oncol*, 11: 108.

- Villalba, A., O. Coll, and F. Gebauer. 2011. 'Cytoplasmic polyadenylation and translational control', *Curr Opin Genet Dev*, 21: 452-7.
- Villanyi, Z., V. Ribaud, S. Kassem, O. O. Panasenko, Z. Pahi, I. Gupta, L. Steinmetz, I. Boros, and M. A. Collart. 2014. 'The Not5 subunit of the ccr4-not complex connects transcription and translation', *PLoS Genet*, 10: e1004569.
- Virtanen, A., N. Henriksson, P. Nilsson, and M. Nissbeck. 2013. 'Poly(A)-specific ribonuclease (PARN): an allosterically regulated, processive and mRNA cap-interacting deadenylase', *Crit Rev Biochem Mol Biol*, 48: 192-209.
- Viswanathan, P., J. Chen, Y. C. Chiang, and C. L. Denis. 2003. 'Identification of multiple RNA features that influence CCR4 deadenylation activity', *J Biol Chem*, 278: 14949-55.
- Wahle, E., and G. S. Winkler. 2013. 'RNA decay machines: deadenylation by the Ccr4-not and Pan2-Pan3 complexes', *Biochim Biophys Acta*, 1829: 561-70.
- Wang, H., M. Morita, X. Yang, T. Suzuki, W. Yang, J. Wang, K. Ito, Q. Wang, C. Zhao, M. Bartlam, T. Yamamoto, and Z. Rao. 2010. 'Crystal structure of the human CNOT6L nuclease domain reveals strict poly(A) substrate specificity', *Embo j*, 29: 2566-76.
- Ward, R. J., E. Alvarez-Curto, and G. Milligan. 2011. 'Using the Flp-In T-Rex system to regulate GPCR expression', *Methods Mol Biol*, 746: 21-37.
- Watanabe, C., M. Morita, T. Hayata, T. Nakamoto, C. Kikuguchi, X. Li, Y. Kobayashi, N. Takahashi, T. Notomi, K. Moriyama, T. Yamamoto, Y. Ezura, and M. Noda. 2014. 'Stability of mRNA influences osteoporotic bone mass via CNOT3', *Proc Natl Acad Sci U S A*, 111: 2692-7.
- Webster, M. W., J. A. W. Stowell, T. T. L. Tang, and L. A. Passmore. 2017. 'Analysis of mRNA deadenylation by multi-protein complexes', *Methods*, 126: 95-104.
- Winkler, G. S., and D. L. Balacco. 2013. 'Heterogeneity and complexity within the nuclease module of the Ccr4-Not complex', *Front Genet*, 4: 296.
- Winkler, G. S., K. W. Mulder, V. J. Bardwell, E. Kalkhoven, and H. T. Timmers. 2006. 'Human Ccr4-Not complex is a ligand-dependent repressor of nuclear receptor-mediated transcription', *Embo j*, 25: 3089-99.
- Wu, M., M. Reuter, H. Lilie, Y. Liu, E. Wahle, and H. Song. 2005. 'Structural insight into poly(A) binding and catalytic mechanism of human PARN', *Embo j*, 24: 4082-93.
- Xu, K., Y. Bai, A. Zhang, Q. Zhang, and M. G. Bartlam. 2014. 'Insights into the structure and architecture of the CCR4-NOT complex', *Front Genet*, 5: 137.
- Xue, Mingshan, Bassam V. Atallah, and Massimo Scanziani. 2014. 'Equalizing excitation-inhibition ratios across visual cortical neurons', *Nature*, 511: 596-600.

- Yamashita, A., T. C. Chang, Y. Yamashita, W. Zhu, Z. Zhong, C. Y. Chen, and A. B. Shyu. 2005. 'Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover', *Nat Struct Mol Biol*, 12: 1054-63.
- Yang, F., and D. R. Schoenberg. 2004. 'Endonuclease-mediated mRNA decay involves the selective targeting of PMR1 to polyribosome-bound substrate mRNA', *Mol Cell*, 14: 435-45.
- Yi, H., J. Park, M. Ha, J. Lim, H. Chang, and V. N. Kim. 2018. 'PABP Cooperates with the CCR4-NOT Complex to Promote mRNA Deadenylation and Block Precocious Decay', *Mol Cell*, 70: 1081-88 e5.
- Yoo, M. H., X. M. Xu, A. A. Turanov, B. A. Carlson, V. N. Gladyshev, and D. L. Hatfield. 2007. 'A new strategy for assessing selenoprotein function: siRNA knockdown/knock-in targeting the 3'-UTR', *RNA*, 13: 921-9.
- Yu, C. H., Y. Dang, Z. Zhou, C. Wu, F. Zhao, M. S. Sachs, and Y. Liu. 2015. 'Codon Usage Influences the Local Rate of Translation Elongation to Regulate Co-translational Protein Folding', *Mol Cell*, 59: 744-54.
- Zhang, K., X. Yin, K. Shi, S. Zhang, J. Wang, S. Zhao, H. Deng, C. Zhang, Z. Wu, Y. Li, X. Zhou, and W. Deng. 2021. 'A high-efficiency method for site-directed mutagenesis of large plasmids based on large DNA fragment amplification and recombinational ligation', *Sci Rep*, 11: 10454.
- Zhang, Q., L. Pavanello, A. Potapov, M. Bartlam, and G. S. Winkler. 2022. 'Structure of the human Ccr4-Not nuclease module using X-ray crystallography and electron paramagnetic resonance spectroscopy distance measurements', *Protein Sci*, 31: 758-64.
- Zhao, Q., L. Pavanello, M. Bartlam, and G. S. Winkler. 2023. 'Structure and function of molecular machines involved in deadenylation-dependent 5'-3' mRNA degradation', *Front Genet*, 14: 1233842.
- Zheng, X., P. Yang, B. Lackford, B. D. Bennett, L. Wang, H. Li, Y. Wang, Y. Miao, J. F. Foley, D. C. Fargo, Y. Jin, C. J. Williams, R. Jothi, and G. Hu. 2016. 'CNOT3-Dependent mRNA Deadenylation Safeguards the Pluripotent State', *Stem Cell Reports*, 7: 897-910.
- Zheng, Z. M., S. Tang, and M. Tao. 2005. 'Development of resistance to RNAi in mammalian cells', *Ann N Y Acad Sci*, 1058: 105-18.
- Zhou, Z., Y. Dang, M. Zhou, L. Li, C. H. Yu, J. Fu, S. Chen, and Y. Liu. 2016. 'Codon usage is an important determinant of gene expression levels largely through its effects on transcription', *Proc Natl Acad Sci U S A*, 113: E6117-E25.
- Zhu, H., H. Zhang, Y. Xu, S. Lassakova, M. Korabecna, and P. Neuzil. 2020. 'PCR past, present and future', *Biotechniques*, 69: 317-25.
- Zhu, X., V. E. Cruz, H. Zhang, J. P. Erzberger, and J. T. Mendell. 2024. 'Specific tRNAs promote mRNA decay by recruiting the CCR4-NOT complex to translating ribosomes', *Science*, 386: eadq8587.

- Zwartjes, C. G., S. Jayne, D. L. van den Berg, and H. T. Timmers. 2004.
'Repression of promoter activity by CNOT2, a subunit of the transcription
regulatory Ccr4-not complex', *J Biol Chem*, 279: 10848-54.

Appendix 1.

Covid statement

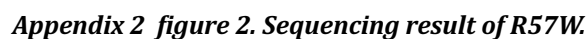
The COVID-19 pandemic significantly impacted the initial stages of my PhD research. During the first year of my program, strict restrictions were imposed on laboratory access and capacity to ensure safety and comply with government and institutional guidelines. As a result, I was limited to working in the lab for only a few hours each day, and these times had to be pre-booked due to reduced bench availability and restrictions on the number of researchers allowed in the lab simultaneously.

Additionally, social distancing measures requiring a minimum distance of 1.5 meters between individuals, along with the mandatory use of face shields and masks, made it challenging for others to effectively demonstrate experimental procedures to me. These restrictions significantly impacted my ability to gain hands-on training and refine my techniques. Consequently, my progress was slower than anticipated, and I was only able to construct two plasmids during my first year. This limited progress was noted with surprise by my internal examiner, further highlighting the unusual challenges faced during this time.

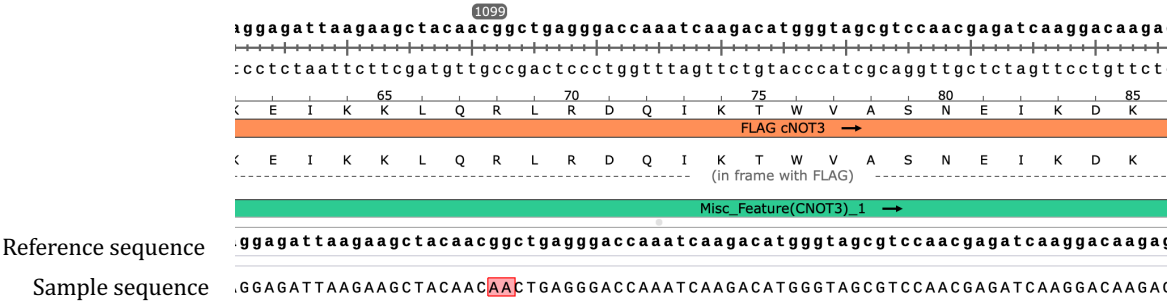
By the second year, as restrictions eased, I was able to accelerate my practical training and research progress.

Appendix 2.

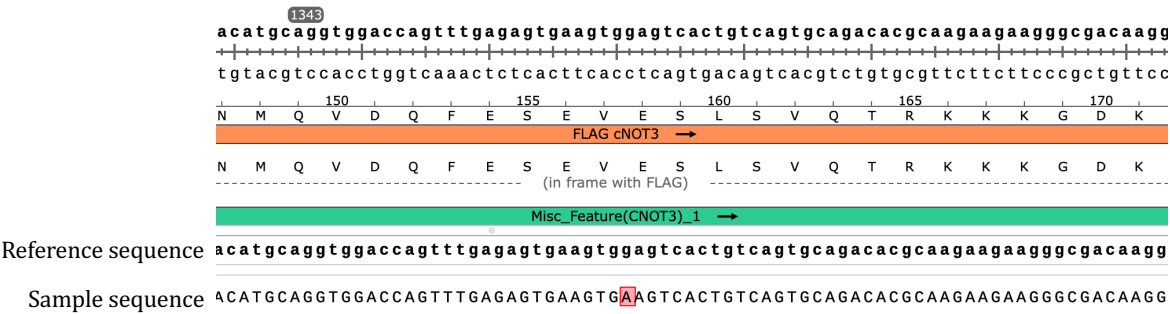
Sequencing results



Appendix 2. Sequencing results

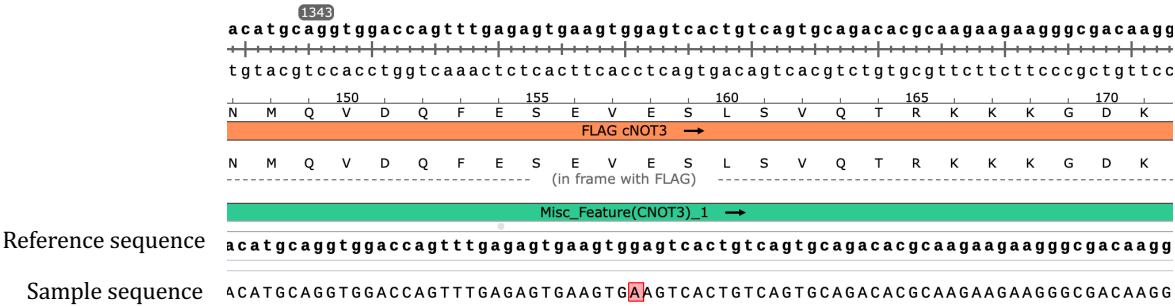


Appendix 2 figure 3. Sequencing result of R57Q.

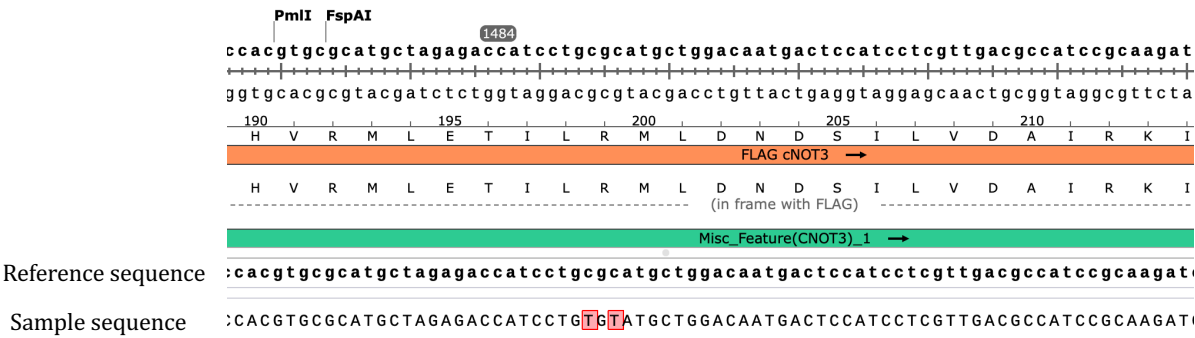


Appendix 2 figure 4. Sequencing result of K119E.

Appendix 2. Sequencing results

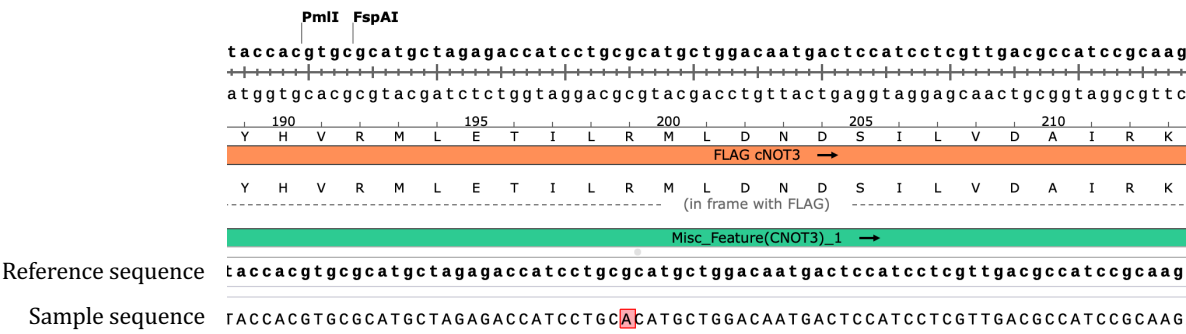


Appendix 2 figure 5. Sequencing result of E147K.



Appendix 2 figure 6. Sequencing result of A188C.

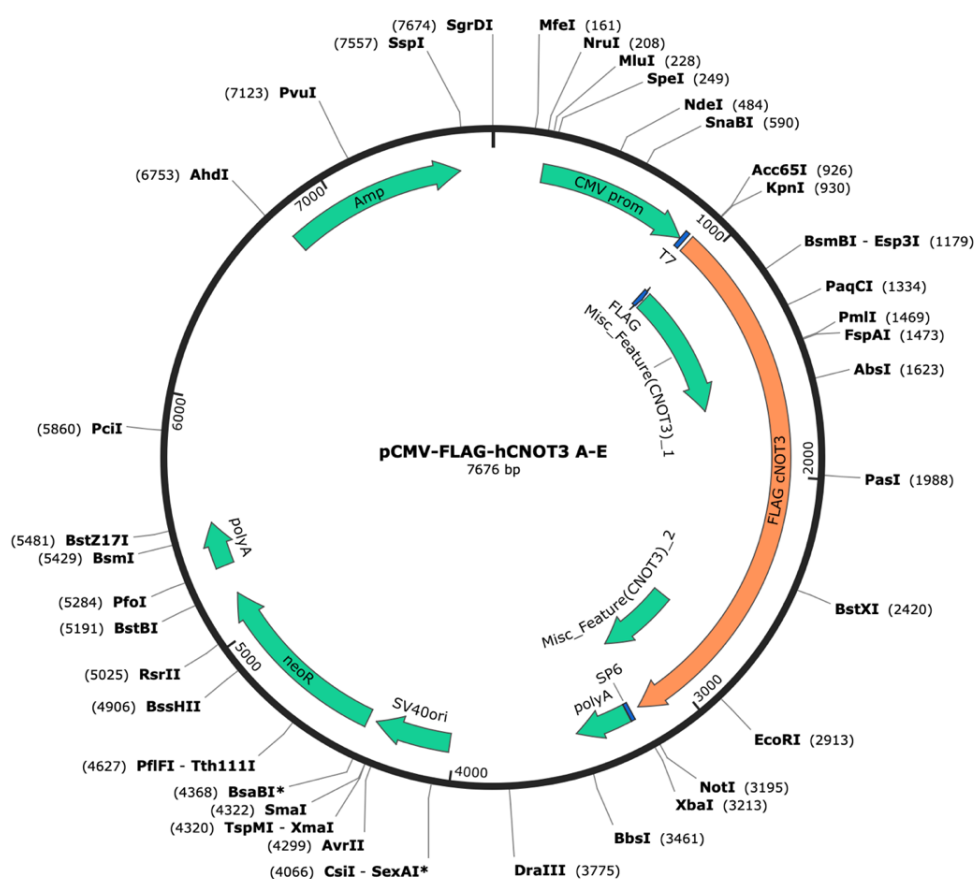
Appendix 2. Sequencing results



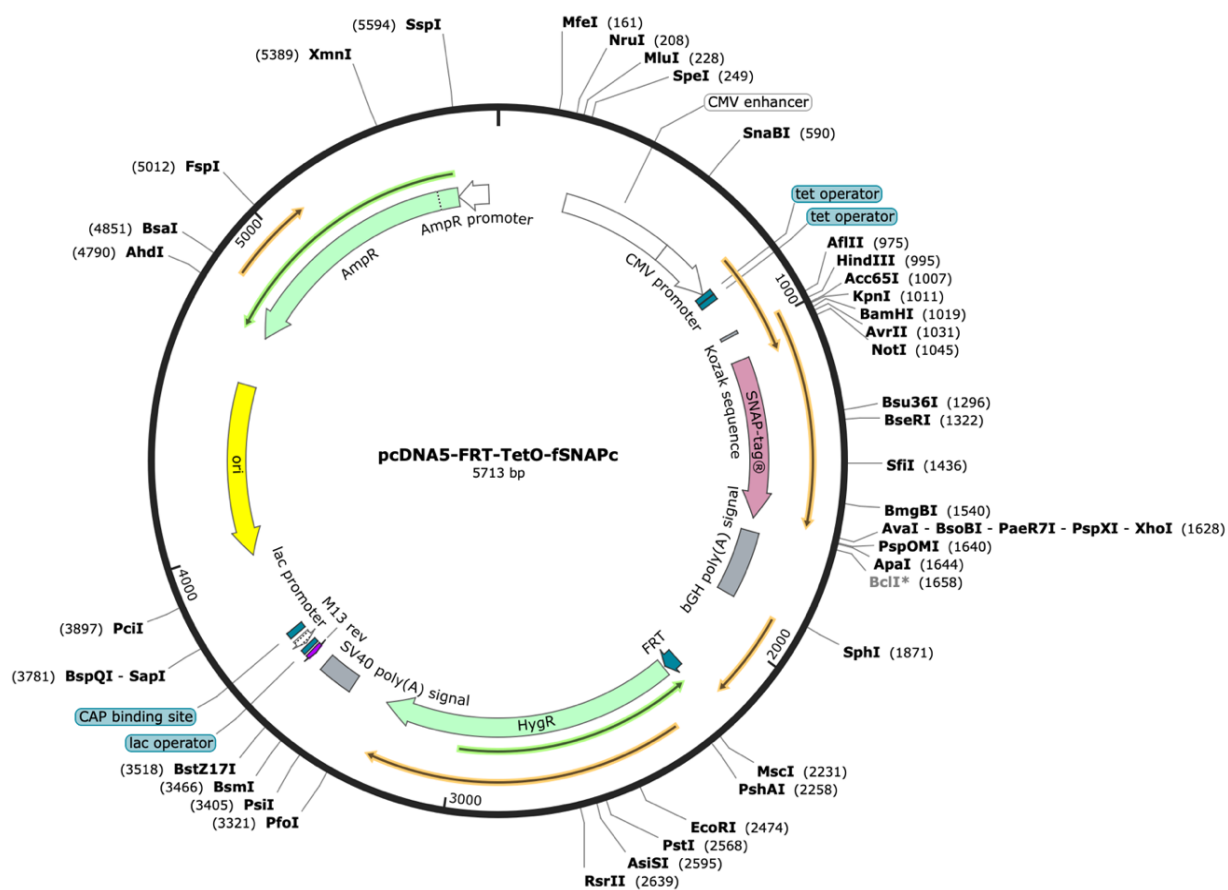
Appendix 2 figure 7. Sequencing result of A188H.

Appendix 3.

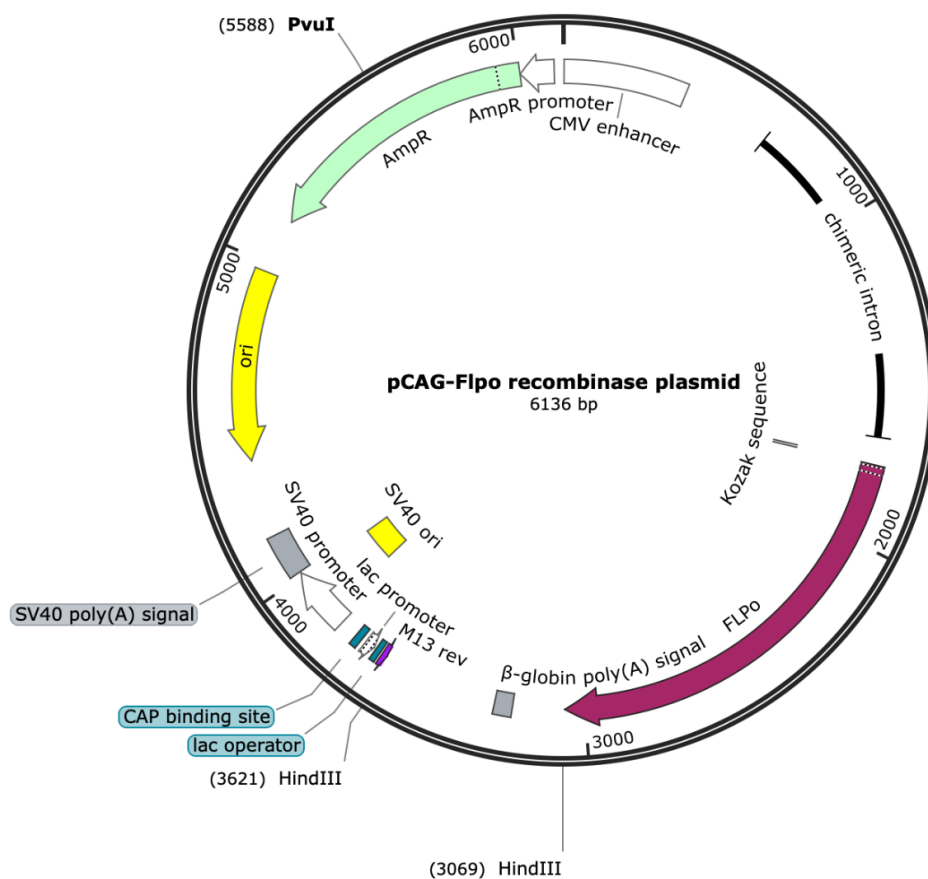
Plasmid maps



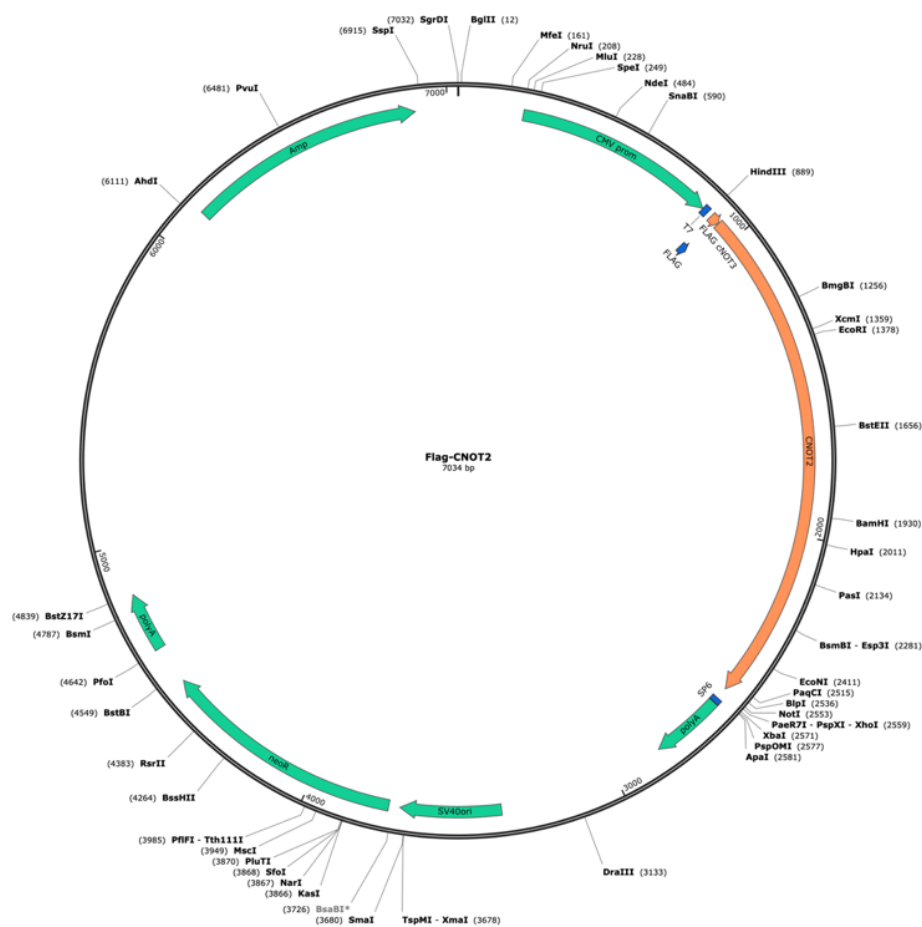
Appendix 3 figure 1. The plasmid map of pCMV-Flag-CNOT3.



Appendix 3 figure 2. The plasmid map of pcDNA5-FRT-TetO-fSNAPc.



Appendix 3 figure 3. The plasmid map of recombinase.



Appendix 3 Figure 4. The plasmid map of pCMV-Flag-CNOT2.