SUSANA MARGARIDA MARQUES COSTA

Licenciada em Biologia

ANATOMY OF THE DISEASE-RELATED DIS3L2 RIBONUCLEASE:

DISSECTING THE AMINO ACIDS RESPONSIBLE FOR SUBSTRATE SPECIFICITY

SUSANA MARGARIDA MARQUES COSTA

Licenciada em Biologia

ANATOMY OF THE DISEASE-RELATED DIS3L2 RIBONUCLEASE:

DISSECTING THE AMINO ACIDS RESPONSIBLE FOR SUBSTRATE SPECIFICITY

ANATOMY OF THE DISEASE-RELATED DIS3L2 RIBONUCLEASE: DISSECTING THE AMINO ACIDS RESPONSIBLE FOR SUBSTRATE SPECIFICITY

SUSANA MARGARIDA MARQUES COSTA

Licenciada em Biologia

Orientador:	Doutora Sandra Cristina Viegas, Investigadora Auxiliar, Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa
Coorientador:	Doutora Rute Gonçalves Matos, Investigadora Auxiliar, Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa

Júri:

Presidente:	Prof.ª Doutora Rita Sobral, Professora Auxiliar, NOVA School of Science and Technology | FCT NOVA
Arguente:	Doctor Benjamin P. Towler, Postdoctoral Researcher, Brighton and Sussex Medical School
Orientador:	Doutora Sandra C. Viegas, Investigadora Auxiliar, Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa

MESTRADO EM MICROBIOLOGIA MÉDICA

Universidade NOVA de Lisboa
Novembro, 2021

Anatomy of the disease-related Dis3L2 ribonuclease: dissecting the amino acids responsible for substrate specificity

“Giving up is for rookies.” – Phil

(Hercules)

Acknowledgments

This would only make sense to me in my native language, so here it goes…

Começo por agradecer à Professora Cecília Arraiano pela oportunidade de poder integrar o seu laboratório, por se ter sempre mostrado atenta e à disposição, permitindo-me ter excelentes condições de trabalho, e enchendo todos os dias em que tive o prazer de me poder cruzar com ela de carinho e entusiasmo.

Um obrigado gigante às minhas orientadoras, Sandra e Rute, que sempre foram impecáveis, me acompanharam de perto (quer na bancada como longe dela), aconselharam, e mantiveram focada. Obrigada pela paciência, mesmo quando falava sozinha (ou com as minhas maravilhosas – tinham dias! – proteínas), ou quando me esquecia que existe uma hora a que as pessoas costumam ir almoçar. Não tenho palavras que descrevam como vocês se tornaram importantes para mim, nem que demonstrem o quanto vocês me ensinaram e fizeram crescer como aspirante a cientista.

A TODOS os colegas do "CMA lab" que entraram na minha vida (quer tenha sido mais no início ou no fim desta jornada), que tive o privilégio de conhecer, poder partilhar os meus dias no ITQB, e com os quais senti que podia sempre contar para qualquer ajuda que precisasse. Um sentido obrigado à Vânia por me ter dado a formação que tanta curiosidade me despertava. Obrigada ao Zé pela indiscritível contribuição musical para os meus finais de tarde. Obrigada à Magá e à Pat pela animação que era estar naquele gabinete. Não sei se agradeça ao Fábio Caetano pelos sustos, mas definitivamente obrigada por me fazer chorar a rir com performances artísticas de alto calibre. E obrigada à *papaguena* do meu coração, à Vanessa, que me deixou assaltar cadernos e marcadores, alimentou com chocolates, e sem a qual muito não teria sido o que foi.

Aos Maristinhas, aos "Fecundados" e a todas as outras grandes amizades que nasceram da FCUL ou do voluntariado. Obrigada por estarem aí, todos estes anos, me terem acolhido e sempre me aturarem, até quando rio *um bocadinho* alto. Vocês são parte de quem eu sou! Um especial obrigada aos que, neste período, se tornaram diários. À minha "Carol Marie" por todas as chamadas aleatórias, companhia nas horas menos recomendadas, *gossip*, reclamações e fotografias motivacionais da gata. Ao Rodrigo pelos mimos, por me fazer sair de casa para passear, e ser capaz de me colar um sorriso na cara mesmo quando o cansaço era crescente.

Um último obrigada à minha família que me faz transbordar de amor, em especial: à minha Siss, sem a qual eu não saberia viver; aos meus pais, que foram sempre incansáveis, não só em boleias, mas sobretudo no apoio e suporte, tendo sempre as palavras certas para me motivarem; à minha avó Alice por todo o esparguete e companhia cá por casa; aos meus avós Helena e Romão por sempre se preocuparem, e se fazerem presentes apesar de distantes; à Sofia, aos meus tios, e às outras três ricas e incríveis primas.

A vocês e a todos aqueles que não mencionei acima pelo nome, um profundo e sincero OBRIGADA!

List of Communications

Oral Presentation

Costa, Susana M., Arraiano, Cecília M., Matos, Rute G., Viegas, Sandra C. "Chasing the crucial residues of Dis3L2 specificity". IMPSG2021 – II International Meeting of the Portuguese Society of Genetics, 1st – 2nd July 2021, Universidade de Trás-os-Montes e Alto Douro, Portugal (Online Meeting)

Poster Presentations

Costa, Susana M., Matos, Rute G., Viegas, Sandra C., Arraiano, Cecília M. "Uncovering the residues responsible for Dis3L2 specificity". MSACO2021 – Microbiology Society Annual Conference Online 2021, 26th – 30th April 2021 (Online Meeting)

Costa, Susana M., Arraiano, Cecília M., Matos, Rute G., Viegas, Sandra C. "Pursuing the key residues behind Dis3L2 specificity". SPB 2020 – XXI National Congress of Biochemistry "Tunning Biochemistry with Life and Society", 14th – 16th October 2021, Universidade de Évora, Portugal (Pitch Presentation on Presential Meeting)

Publications

Costa, Susana M., Saramago, Margarida, Matos, Rute G., Arraiano, Cecília M., Viegas, Sandra C. "How hydrolytic exoribonucleases impact human disease: Two sides of the same story". Invited Review to FEBS Open Bio, *accepted*

Abstract

Ribonucleases are enzymes which perform the processing and degradation of all types of RNA, being critical for the tight post-transcriptional regulation of gene expression. Three homologues from the RNase II/RNB family of exoribonucleases are found in eukaryotes: Dis3, Dis3L1, and Dis3L2. The first two associate with the RNA exosome, while Dis3L2 participates in an alternative degradation pathway, depending on 3'-uridylation by terminal uridylyl transferases and poly(U) polymerases. The first insights on protein–RNA interactions were obtained analysing the mouse Dis3L2–oligo(U) structure, supporting the distinctive preference of Dis3L2 for uridylated substrates. To date, Dis3L2 has already been associated with cellular processes as relevant as RNA surveillance, cell proliferation and differentiation. Human DIS3L2 defects have been related to several cancers, alongside with Perlman Syndrome, a rare foetal overgrowth disorder, consequently increasing the risk for Wilms Tumour, a malignant kidney tumour commonly affecting children. Nevertheless, much remains unknown about the mechanism of action of Dis3L2 and its peculiar substrate preference.

In this work, fission yeast was used as an eukaryotic microorganism model for studying a considerably important ribonuclease in human diseases. Our aim was the identification of particular amino acids potentially crucial for activity and substrate specificity of *Schizosaccharomyces pombe* Dis3L2 (SpDis3L2), namely those involved in its preference for uridylated substrates. Thus, wild-type and point mutants of SpDis3L2 were overexpressed and purified. A thermal shift assay allowed the improvement of protein stability and enhanced full-length protein purification. Subsequently, the exoribonucleolytic activity of SpDis3L2 versions was analysed over different RNA substrates, using an inactive mutant as control. We uncovered that all tested mutants were active, with several of them differing from the wild-type regarding their level of activity and their preference for uridylated substrates. Therefore, these results pave the way for the analysis of other protein versions with combinations of these point mutations.

Keywords: RNA decay, RNase II/RNB family, Uridylation, Dis3L2, *Schizosaccharomyces pombe*, Exoribonucleolytic activity

Resumo

As ribonucleases são enzimas responsáveis pelo processamento e degradação do RNA, sendo essenciais na regulação pós-transcricional da expressão génica. Em eucariotas, encontramos três homólogos das exorribonucleases da família RNase II/RNB: Dis3, Dis3L1 e Dis3L2. Os dois primeiros associam-se ao exossoma, enquanto a Dis3L2 participa numa via alternativa de degradação dependente da 3'-uridilação por uridilil transferases terminais e poli(U) polimerases. Atualmente, a Dis3L2 já foi associada a processos celulares pertinentes como o controlo de qualidade do RNA, e a proliferação e diferenciação celulares. Defeitos na DIS3L2 humana foram relacionados com vários cancros, bem como com a Síndrome de Perlman, uma doença rara de sobre-crescimento fetal que aumenta o risco para o Tumor de Wilms, um tumor renal geralmente encontrado em crianças. Contudo, ainda muito se desconhece do mecanismo de ação da Dis3L2 e da sua preferência distintiva por substratos uridilados.

Usando a levedura de fissão como modelo de microrganismo eucariótico para estudar esta ribonuclease extremamente relevante para doenças humanas, este trabalho pretendeu identificar aminoácidos da Dis3L2 de *Schizosaccharomyces pombe* (SpDis3L2) potencialmente cruciais para a sua atividade, especificidade de substrato, e, em particular, pela preferência por substratos uridilados. Assim, a SpDis3L2 selvagem e respetivos mutantes pontuais foram sobrexpressos e purificados, após a otimização das condições de purificação para que favorecessem a purificação da proteína no seu estado íntegro. A atividade exorribonucleolítica das versões da SpDis3L2 foi analisada sobre diversos substratos de RNA, tendo-se revelado que, à exceção do mutante inativo usado como controlo, todos os mutantes testados eram ativos. Comparando com a proteína SpDis3L2 selvagem, houve mutantes para os quais se observaram diferenças relativamente ao seu nível de atividade e à preferência por substratos uridilados. Tomando estes resultados em consideração, este trabalho constitui o primeiro passo para que se possa, futuramente, analisar versões desta ribonuclease com combinações destas mutações pontuais.

Palavras-Chave: Degradação do RNA, Família RNase II/RNB, Uridilação, Dis3L2, *Schizosaccharomyces pombe*, Atividade exorribonucleolítica

List of Contents

List of Figures

List of Tables

Abbreviations

3'-RACE	3'- Rapid Amplification of cDNA Ends
A	Adenosine (in the context of nucleotides)
A. thaliana	*Arabidopsis thaliana* (flowering plant)
aa	Amino acid(s)
Adh	Alcohol Dehydrogenase
Air2	Arginine methyltransferase-interacting RING finger protein 2
Amp	Ampicillin
APS	Ammonium persulfate
argU	tRNA gene that recognizes two rare codons in *E. coli* (AGG/AGA)
ATC	Anaplastic Thyroid Cancer
AtDis3L2	*Arabidopsis thaliana* Dis3L2 (Synonym: SOV)
ATP	Adenosine Triphosphate
bla	Beta-lactamase gene that confers resistance to ampicillin
BLASTp	Protein Basic Local Alignment Search Tool
BSA	Bovine Serum Albumin
C	Cytosine (in the context of nucleotides)
C. elegans	*Caenorhabditis elegans* (nematode)
C. reinhardtii	*Chlamydomonas reinhardtii* (green algae)
Cam	Chloramphenicol
CeDis3L2	*Caenorhabditis elegans* Dis3L2
Cid1	Caffeine-Induced Death suppressor protein 1
circRNA	circular RNA
CNP	C-type Natriuretic Peptide
COVID-19	COronaVIrus Disease 2019
CR3	Cysteine-Rich motif with three cysteines
CSD	Cold-Shock Domain
C-term	Carboxyl-terminus
CTEV	Congenital Talipes Equinovarus (or clubfoot)
D. melanogaster	*Drosophila melanogaster* (fly)
D. rerio	*Danio rerio* (zebrafish)
dbSNP	Database of Single Nucleotide Polymorphisms (SNPs) and multiple small-scale variations
Dcp-1/Dcp-2	Decapping complex proteins 1 and 2
DcpS	Decapping Scavenger
ddH$_2$O	Double-Distilled water

Dis3	Defective in sister chromatid joining
Dis3L, Dis3L1	Dis3-Like exonuclease 1
Dis3L2	Dis3-Like exonuclease 2
DmDis3L2	*Drosophila melanogaster* Dis3L2
dsRNA	double-stranded RNA
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EcRNase II	RNase II from *E. coli*
EDTA	Ethylenediamine tetraacetic acid
EMSA	Electrophoretic Mobility Shift Assay
Exo-10	Exo-9 + catalytic subunit Rrp6
Exo-11	Exo-9 with two catalytic subunits (Rrp6 and, most importantly, Dis3 or Dis3L1)
Exo-9	9 subunits that compose the structural core of the eukaryotic RNA exosome
EXOSC10	EXOSome Component 10 (Synonyms: Rrp6, PM/Scl-100)
g	Gram
G	Guanosine (in the context of nucleotides)
gRNA	guide RNA
h	Hour
H. sapiens	*Homo sapiens* (humans)
hDIS3, HsDis3	human Dis3
hDIS3L1, HsDis3L1	human Dis3L1
hDIS3L2, HsDis3L2	human Dis3L2
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethane sulfonic acid
His-SUMO	Tag composed of 6 histidine residues (6xHis) and a Small Ubiquitin-like MOdifier (SUMO)
HPA	Human Protein Atlas
Hs2	*Homo sapiens* protein 2, identified as PUP 6 in (61) (Synonyms: TUT7, TUTase 7, PAPD6, TENT3B, ZCCCHC6)
Hs3	*Homo sapiens* protein 3, identified as a PUP in (61) (Synonyms: TUT4, TUTase 4, PAPD3, TENT3A, ZCCCHC11)
HsGLD2, hGLD2	*Homo sapiens* Poly(A) RNA polymerase GLD2 (Defective in germ line development protein 2) (Synonyms: TENT2, TUTase 2, PAPD4)
ileY	tRNA gene that recognizes a rare codon in *E. coli* (AUA)
IMAC	Immobilized Metal Affinity Chromatography
INIBD	Intranuclear Inclusion Body Disease
IPTG	Isopropyl-β-D-Thiogalactopyranoside

kDa	Kilodalton
L	Litre
lacI	Gene that codifies for the repressor of *lac* operon
LB	Luria-Bertani Broth
let-7	lethal-7 microRNA
leuW	tRNA gene that recognizes a rare codon in *E. coli* (CUA)
Lin28a, Lin28b	Protein lin-28 homologues A and B
Lsm1-7	Sm-Like proteins 1 to 7
M	Molar (mol/L)
M. musculus	*Mus musculus* (mouse)
m^7GMP	7-methyl Guanosine Monophosphate
m^7GpppG	7-methyl Guanosine 5'cap
mDis3L2	mouse Dis3L2
mg	Milligram
Mg^{2+}	Magnesium ion
µg	Microgram
µl	Microlitre
µm	Micrometre
min	Minute
miR-562	microRNA 562
miRNA	microRNAs
ml	Millilitre
MLS	Marfan-Like Syndrome
mM	Millimolar (mmol/L)
mmol	Millimole
mol	Mole
MOPS	3-morpholinopropane-1-sulfonic acid
$MQ\ H_2O$	Milli-Q water
mRNA	messenger RNA
Mtr4	mRNA transport regulator 4 (or ATP-dependent RNA helicase DOB1)
MWCO	Molecular Weight Cut-Off
N	Normal
ncPAP	non-canonical Poly(A) Polymerase
ncRNA	non-coding RNA
NDP	Nucleoside Diphosphate
nm	Nanometre
NMD	Nonsense-Mediated Decay

NPPC	Gene that codifies for CNP
nt	nucleotide(s)
N-term	Amino-terminus
OB-fold	Oligosaccharide/Oligonucleotide Binding fold
ºC	Degree Celsius
OD_{600}	Optical Density at 600 nm
OFC	Orofacial Cleft
Oligo(dT)	deoxythymidine Oligonucleotide
PAA	Polyacrylamide
PABP	Poly(A) Binding Protein
PAGE	Polyacrylamide Gel Electrophoresis
PAP	Poly(A) Polymerase
PAPD3	PAP associated Domain containing 3 (Synonyms: TUT4, TUTase 4, Hs3, TENT3A, ZCCCHC11)
PAPD4	PAP-associated Domain-containing protein 4 (Synonyms: TENT2, TUTase 2, HsGLD2, hGLD2)
PAPD6	PAP associated Domain containing 6 (Synonyms: TUT7, TUTase 7, Hs2, TENT3B, ZCCCHC6)
Pat1	DNA topoisomerase 2-associated protein PAT1
P-bodies	Processing bodies
Pacman, Pcm	5'-3' exoribonuclease 1 in *D. melanogaster* (Synonym: Xrn1)
PCP	Poly(C) Polymerase
PGP	Poly(G) Polymerase
PIN	PilT N-terminus domain
piRNA	piwi-interacting RNA
PM/Scl-100	Polymyositis/Scleroderma autoantigen 100 kDa
PMSF	Phenylmethanesulfonyl fluoride
PNK	Polynucleotide kinase
PNPase	Polynucleotide Phosphorylase
pre-let-7	precursor of let-7 microRNA
PRLMNS	Perlman Syndrome
psi	Pound-force per square inch (or lbf/in^2)
PTC	Premature translation-Termination Codon
PUP	Poly(U) Polymerase
PVDF	Polyvinylidene Fluoride (membrane)
Rat1	5'-3' exoribonuclease 2 (Synonym: Xrn2)
RBP	RNA Binding Protein

RISC	RNA-Induced Silencing Complex
RNA	Ribonucleic Acid
RNAi	RNA interference
RNase	Ribonuclease
RNB	RNase II catalytic domain
rpm	revolutions per minute
rRNA	ribosomal RNA
Rrp44	Ribosomal RNA-processing protein 44 (or Exosome complex exonuclease RRP44; Synonym: Dis3)
Rrp6	Ribosomal RNA-processing protein 6 (or Exosome complex exonuclease RRP6; Synonyms: EXOSC10, PM/Scl-100)
RT	Room Temperature
S. cerevisiae	*Saccharomyces cerevisiae* (budding yeast)
S. pombe	*Schizosaccharomyces pombe* (fission yeast)
S1	RNA-binding domain, which was initially identified in ribosomal protein S1
ScDis3	*Saccharomyces cerevisiae* Dis3 (Synonym: Rrp44)
SDS	Sodium Dodecyl Sulphate
siRNA	small interfering RNA
snoRNA	small nucleolar RNA
snRNA	small nuclear RNA
SOV	Suppressor Of Varicose (Synonym: AtDis3L2)
SpDis3	*Schizosaccharomyces pombe* Dis3
SpDis3L2	*Schizosaccharomyces pombe* Dis3L2
ssRNA	single-stranded RNA
T	Thymine (in the context of nucleotides)
T. brucei	*Trypanosoma brucei* (trypanosome)
TAIL-Seq	Tail Sequencing
TB	Terrific Broth
TBE	Tris-Borate-EDTA
TBS	Tris-buffered saline
TBS-T	Tris-buffered saline with 0.1% Tween 20
TDMD	Target RNA-Directed miRNA Degradation
TDS	TUT-DIS3L2 Surveillance (pathway)
TEMED	Tetramethyl ethylenediamine
TENT	Terminal RNA Nucleotidyl Transferase
TENT2	Terminal RNA Nucleotidyl Transferase 2 (Synonyms: TUTase 2, HsGLD2, hGLD2, PAPD4)

TENT3A	Terminal RNA Nucleotidyl Transferase 3A (Synonyms: TUT4, TUTase 4, Hs3, PAPD3, ZCCCHC11)
TENT3B	Terminal RNA Nucleotidyl Transferase 3B (Synonyms: TUT7, TUTase 7, Hs2, PAPD6, ZCCCHC6)
T_m	Melting Temperature
TRAMP	Trf4/Air2/Mtr4 Polyadenylation (complex)
TrB	Transfer Buffer
Trf4	Topoisomerase one-Related Function protein 4
Trf5	Topoisomerase one-Related Function protein 5
tRNA	transfer RNA
TRUMP	Terminal RNA Uridylation-Mediated Processing (complex)
TSA	Thermal Shift Assay
TUT4, TUTase 4	Terminal Uridylyl Transferase 4 (Synonyms: Hs3, PAPD3, TENT3A, ZCCCHC11)
TUT7, TUTase 7	Terminal Uridylyl Transferase 7 (Synonyms: Hs2, PAPD6, TENT3B, ZCCCHC6)
TUTase	Terminal Uridylyl Transferase
TUTase 2	Terminal Uridylyl Transferase 2 (Synonyms: TENT2, HsGLD2, hGLD2, PAPD4)
U	Enzyme unit
U	Uridine (in the context of nucleotides)
UMP	Uridine Monophosphate
U-tail	Uridine tail
UTP	Uridine Triphosphate
UTR	Untranslated Region
UV	Ultraviolet
V	Volt
vol	Volume
wt	wild-type (in the context of proteins)
WT	Wilms' Tumour (in the context of human diseases)
X. laevis	*Xenopus laevis* (African clawed frog)
Xrn1	5'-3' exoribonuclease 1 (Synonym: Pcm)
Xrn2	5'-3' exoribonuclease 2 (Synonym: Rat1)
ZCCHC11	Zinc finger CCHC domain-containing protein 11 (Synonyms: TUT4, TUTase 4, Hs3, PAPD3, TENT3A)
ZCCHC6	Zinc finger CCHC-type containing 6 (Synonyms: TUT7, TUTase 7, Hs2, PAPD6, TENT3B)

1. Introduction

1.1.   Ribonucleases and their relevance in cellular metabolism

RNA is an extremely important molecule inside the cell, performing numerous functions. For example, messenger RNAs (mRNAs), transfer RNAs (tRNAs), and ribosomal RNAs (rRNAs) are involved in protein synthesis (1), whereas non-coding RNAs (ncRNAs) can play a panoply of regulatory roles both at transcriptional and post-transcriptional levels (2). At any given moment, the amount of RNA present in the cell results from the balance between its synthesis and degradation, according to the cellular needs or the urge to adapt to some environmental change. Therefore, ribonucleases (RNases), the enzymes responsible for all RNA species' processing and degradation, arise as remarkably relevant. They not only allow ribonucleotides recycling but also participate in several quality control mechanisms, namely by eliminating aberrant RNAs. There are two types of RNases in the cell: the endoribonucleases, which cleave the RNA molecules internally, and the exoribonucleases, that degrade the RNA starting on one of its extremities (3).

1.2. Eukaryotic mRNA Decay Pathways

In eukaryotes, mature mRNAs usually present a 5' 7-methyl guanosine cap (m^{7}GpppG) and a 3' terminal poly(A) tail, with the latter being firstly produced by poly(A) polymerases (PAPs) and secondly protected from degradation by the attachment of Poly(A) Binding Proteins (PABPs) (4, 5). In the cytoplasm, these tails can be further extended by non-canonical PAPs (ncPAPs) which will increase the messenger stability and favour translation (6). Moreover, the 5' and/or 3' untranslated regions (5'UTR and 3'UTR, respectively) of the eukaryotic mRNAs frequently include specific sequences able to either stabilize or promote its decay via binding of certain RNA Binding Proteins (RBPs) or regulatory microRNAs (miRNAs) (7–10). Nonetheless, if these molecules are abnormal (e.g., possessing nonsense codons), they can undergo an endonucleolytic cleavage, with the resulting products being eventually degraded by other mechanisms (11, 12). Afterwards, when these cytoplasmic mRNAs are no longer needed, they can be destroyed through several parallel though partially redundant pathways (13–15), involving either endo or exoribonucleases (Figure 1.1.).

The first step in these pathways is commonly the deadenylation or decapping of the transcripts. The deadenylation consists of the loss or shortening of the poly(A) tail carried out by deadenylases, which also causes the release of the PABPs and originates a mRNA with an unprotected 3'-end (16, 17). In turn, the decapping is triggered by the binding of the Lsm1-7/Pat1 complex (which works as a decapping activator, where the Pat1 protein recruits the Sm-like proteins Lsm1 to Lsm7) preferably to the 3' extremity of deadenylated mRNAs, blocking the access of exoribonucleases to this RNA extremity (18, 19). Next, the Dcp-1/Dcp-2 complex (composed of decapping enzymes 1 and 2) promotes the 5'cap removal, leading to the exposure of the 5'-end of the transcript (20). After this first phase has ensued, there are two canonical RNA degradation pathways following opposite directions.

On one hand, if the decapping occurs, that entails the decay to the 5'-3' direction, whereby the 5'-3' exoribonuclease 1 (Xrn1, also called Pacman, Pcm) is able to degrade completely the 5' monophosphate RNA starting on its 5'-end. This enzyme is the only known cytoplasmic exoribonuclease to act in this direction (13, 21–23). However, the 5PX superfamily to which Xrn1 belongs also includes a 5'-3' exoribonuclease 2 (Xrn2, also termed Rat1), which operates in the nucleus (13, 24).

On the other hand, if the mRNA is just deadenylated, then the RNA exosome, a multiprotein complex, is capable of attaching to the mRNA 3'-end and degrade the molecule in the 3'-5' direction (16, 25, 26). The eukaryotic RNA exosome is composed of a structural core, which is ring-shaped like the bacterial polynucleotide phosphorylase (PNPase) (27), encompassing nine subunits (Exo-9). PNPase has, amongst other functions, phosphorolytic activity and is able to processively degrade RNA molecules in the 3'-5' direction (27). Unlike PNPase, the subunits found in the eukaryotic Exo-9 core have lost that ability (except in the case of plants (28)), serving as a channel for the passage of RNA

molecules (27). This Exo-9 core becomes responsible for coupling a tenth and, sometimes, an eleventh variable subunits, which are the ones responsible for the catalytic activity of the RNA exosome (29). Depending on the cellular location of this multiprotein complex, it presents distinct isoforms. It has already been found that the tenth subunit of the eukaryotic RNA exosome is usually the Rrp6 protein (also known as EXOCS10 or PM/Scl100 in humans). Rrp6 is a distributive exoribonuclease of the DEDD superfamily, which also includes the *E. coli* RNase D (24). Rrp6 is present in all human isoforms of the RNA exosome (30, 31), whereas it exists only in the nuclear isoform of yeast (32). Noteworthy, it is generally the only catalytic subunit of the RNA exosome in the nucleolus, where a ten-subunits isoform (Exo-10) is present. In contrast, an eleventh catalytic subunit is attached to Exo-10, forming an Exo-11 complex in the nucleoplasm or cytoplasm. This eleventh subunit might be the Defective in sister chromatid joining (Dis3, also known as Rrp44) exoribonuclease or the Dis3-Like exonuclease 1 (Dis3L1 or Dis3L) (31, 33). Both of these proteins belong to the RNase II/RNB family and will later be discussed in further detail. In these scenarios where Dis3 belongs to the nucleoplasmic isoform, and Dis3L1 is included in the cytoplasmic isoform, Rrp6, which is still associated with the exosome, becomes a non-essential subunit, since Dis3 and Dis3L1 are highly processive (3, 34) and become the main responsible for RNA hydrolysis. In any case, after the digestion by the RNA exosome, the 5'cap in the remaining fragment (usually smaller than 10 nucleotides) is eventually decomposed by the Decapping Scavenger (DcpS) enzyme. DcpS is a m^7-specific pyrophosphatase, which ultimately generates a 7-methyl guanosine monophosphate (m^7GMP) and a nucleoside diphosphate (NDP) (35–37).

These were the most widely studied general processes of RNA decay. However, the discovery of Dis3-Like exonuclease 2 (Dis3L2), another member of the RNase II/RNB family, revealed a new alternative eukaryotic RNA decay pathway independent of deadenylation and decapping, but rather mediated by uridylation. This major finding took place almost simultaneously in fission yeast (38), mouse (39, 40), and humans (40, 41), and challenged all the previous assumptions on eukaryotic RNA degradation pathways. Thus, the mechanism of uridylation and the Dis3L2 ribonuclease involved in this RNA decay pathway uridylation-dependent will be next explored.
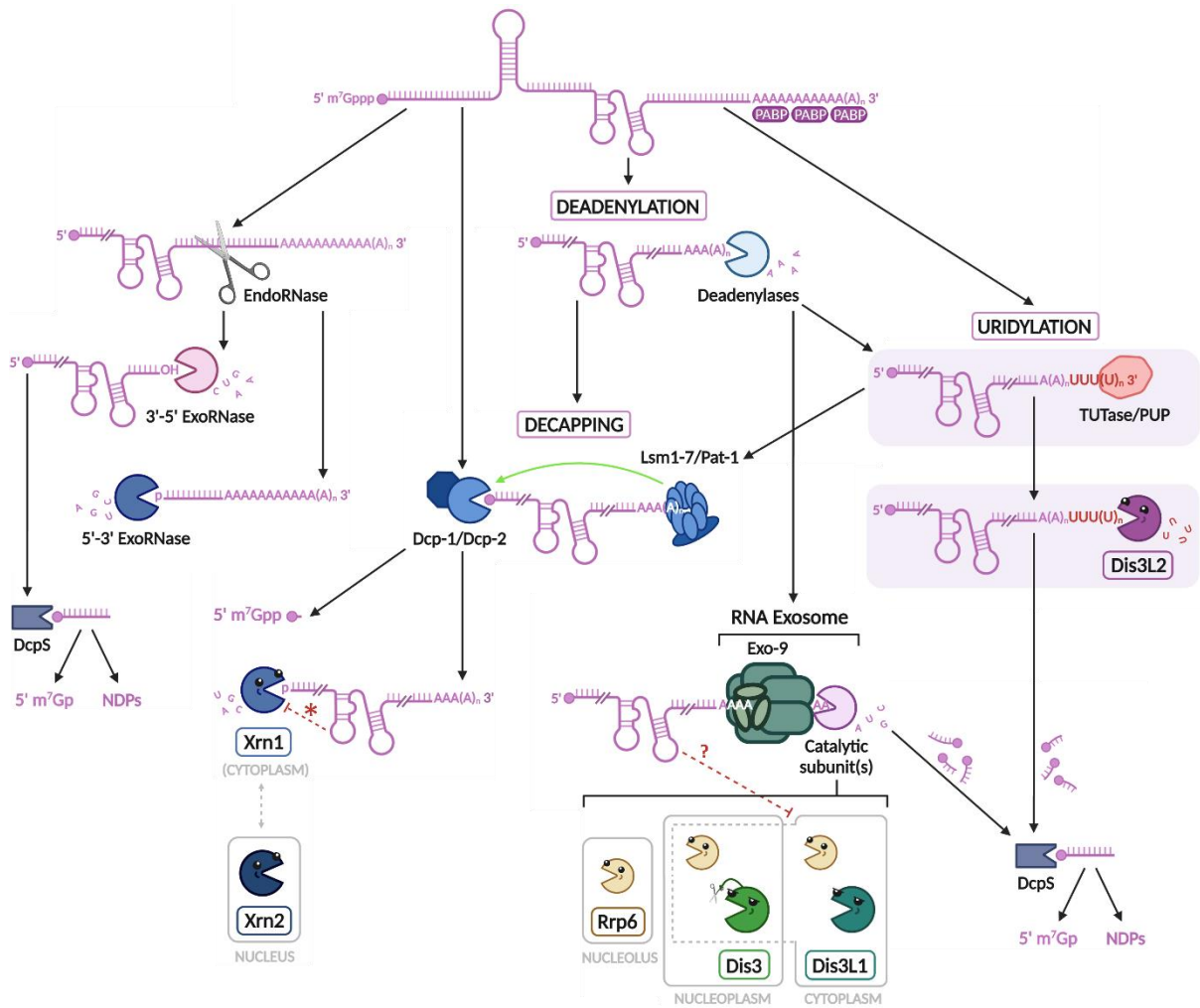
Figure 1.1. Overview of the main mRNA degradation pathways in eukaryotes. Mature cytoplasmic mRNAs are usually stabilised by a 5'cap and a long 3' poly(A) tail, to which poly(A) binding proteins (PABPs) may bind, thus, protecting that extremity from degradation. When transcripts' decay is triggered, distinct enzymes might intervene: endoribonucleases (EndoRNases), and/or exoribonucleases (ExoRNases). In the 5'-3' direction, decapped mRNAs might be digested by Xrn1 in the cytoplasm or by Xrn2 in the nucleus. ' ∗ ' represents that, despite Xrn1 being able to unwind highly structured RNA molecules, some of them block Xrn1 progression. In the 3'-5' direction, transcripts degradation might follow deadenylation (carried out by deadenylases) or uridylation (performed by terminal uridylyl transferases, TUTases, or poly(U) polymerases, PUPs). Deadenylated mRNAs can be readily degraded by the RNA exosome, which has distinct isoforms according to cellular localization, even though its nucleoplasmic isoform may also be present in the cytoplasm in smaller amounts (grey dashed line in the 'nucleoplasm' box). ' ? ' means that there is still no evidence of Dis3L1 degrading secondary structures. Uridylated mRNAs recruit specifically the Dis3L2 enzyme (pathway highlighted in purple). Small capped degradation products are fully hydrolysed by DcpS enzyme. (Created with BioRender.com.)

1.3. 3' Uridylation in eukaryotes

1.3.1. The mechanism of uridylation and how to escape it

Polyadenylation is the most well-documented 3'-end modification, but other changes may also take place in the RNA with some frequency, namely guanylation, cytidylation, and uridylation (42). These processes were largely underestimated with the use of oligo(dT)-based methodologies undoubtedly

skewing the terminal modifications that could be found in the nucleic acids. The rise of exciting innovative technologies, such as the 3'-RACE (43, 44) and the TAIL-Seq (42, 45), shed a new light on the whole 3'-terminome.

Some of the 3'-end modifications are executed by terminal RNA nucleotidyl transferases (TENTs), a subgroup of ncPAPs that belong to the DNA polymerase β-like nucleotidyl transferases superfamily, which is well-conserved in eukaryotes (46, 47). Although many of these enzymes perform polyadenylation in the most diverse scenarios (48–53), a subclass of them prefers uridine triphosphate (UTP) over adenosine triphosphate (ATP), thus uridylating their substrates (54). For example, in humans, only two of their eleven TENTs have already been reported to be involved in 3' uridylation of cytoplasmic mRNAs: TUT4 (also called ZCCHC11 (6), TUTase 4 (55), Hs3 (56), PAPD3 or TENT3A) and TUT7 (also named ZCCHC6 (6), TUTase 7 (55), Hs2 (56), PAPD6, or TENT3B). Hereafter, the used human TUTases nomenclature will be the one established by Yashiro *et al.*, 2018 (57) (which corresponds to the above-mentioned outside brackets).

Therefore, 3' uridylation is the process that comprises the addition of untemplated uridine monophosphate (UMP) residues to the 3'-end hydroxyl group of RNA molecules, thereupon being independent of the pre-existing sequence (6, 58–61). Hence, this post-transcriptional modification, which appears to be only present in eukaryotes, can be conducted by two types of TENTs: the terminal uridylyl transferases (TUTases) and the poly(U) polymerases (PUPs), which add, respectively, short [oligo(U)] or long [poly(U)] uridine tails (U-tails) to the RNA molecules (58, 60, 62). These enzymes intervene mainly in the cytoplasm (54, 55, 60), notwithstanding that they have also been described in other subcellular locations, specifically the nucleus (63) and mitochondria (53, 64–67). However, uridylation is yet to be described in chloroplasts (60).

TUTases and PUPs can act on plenty of RNA types, such as mRNAs, tRNAs, rRNAs, miRNAs, small interfering RNAs (siRNAs), piwi-interacting RNAs (piRNAs), guide RNAs (gRNAs), small nuclear RNAs (snRNAs), and small nucleolar RNAs (snoRNAs) (59). Nevertheless, the selectivity mechanisms of these targets are still largely unknown, and a further investigation of the structure of these proteins is required. In general, they exhibit a wide range of RNA substrates, albeit some of them display a higher specificity. The activity of TUTases may also vary according to the substrate in question and depending on if they are performing their function by themselves or as part of a larger complex.

Although in fungi like *Schizosaccharomyces pombe* uridylation is independent of deadenylation (68), in plants and mammals uridylation may be favoured by the small size of the poly(A) tail of the transcripts (69, 70). In fact, TUT4 and TUT7 select for uridylation the cytoplasmic mRNAs with poly(A) tails shorter than 25 nucleotides (nt), while the PAPBs predominantly attached to the longer poly(A) tails protect those transcripts from being targeted by these TUTases (42, 69). A widely studied example of one of the major roles of TUTases in human cellular metabolism is the control of pre-let-7 mRNA

maturation/degradation, depending on its mono-/oligo-uridylation, respectively, by TUT4/7 (71, 72), which will be detailed below. The involvement of orthologue TUTases in this same pathway has already been uncovered in other organisms, such as *Caenorhabditis elegans* (73), *Danio rerio* (74), and *Mus musculus* (75).

Unquestionably, template-independent uridylation arises as a ubiquitous process in the eukaryotic cell, being expected to be tightly regulated and to have a way of evading it. Therefore, the 2'-*O*-methylation, which resides in the addition of a methyl group to the 2' hydroxyl group of the last ribose in the 3'-end of RNAs by methyltransferases, has surfaced as a mechanism of protection from uridylation that is evolutionarily conserved amongst eukaryotes. This 3'-terminal modification prevents others from occurring simultaneously. In its absence, specific classes of ncRNAs (such as miRNAs, piRNAs and siRNAs) appear destabilized by extended 3' oligo-uridylation, resulting in their 3'-5' exonucleolysis (59, 62).

1.3.2. The implications of uridylation

Considering this panoply of modifying enzymes and substrates implicated, many of which are still unknown, if vulnerable unmethylated RNAs are uridylated by TUTases or PUPs, what consequences might this modification entail? This issue is extensively reviewed in several recent papers, explaining the main factors involved (namely, TUTases and exoribonucleases) and the cellular processes affected, covering a broad range of eukaryotic organisms (57, 62, 70, 76–78). Remarkably, 3' uridylation can either confer the target RNA molecules higher stability or, contrarily, mark them for the Dis3L2-dependent degradation pathway. It is pertinent to emphasize that the TUTases participating in each of these contexts and the size of the U-tails produced are usually different.

On the one side, uridylation is needed for the biogenesis of certain RNAs (79), can regulate translation efficiency (67), and might be possibly involved in the signalling for transport of micro- and other RNAs by extracellular vesicles (80). In addition, the most disparate, and sometimes unexpected, examples of uridylation-associated processes keep emerging from model organisms: the maturation of mitochondrial gRNAs and RNA editing, in *Trypanosoma brucei* (81); an antiviral response mechanism, in *C. elegans* (82); and the PABPs-dependent recovery of a specific size of oligo(A) tails in moderately deadenylated mRNAs, in *Arabidopsis thaliana* (83).

On the other side, uridylation may promote degradation in the two opposite directions. In what concerns the uridylated RNA decay in the 5'-3' direction, the U-tails can recruit the Lsm1-7 complex, consequently conducting to the targeted RNA decapping, and later digestion by Xrn1 (68, 69). Alternatively, these oligo(U) tails can directly induce the RNA decay in the 3'-5' direction by Dis3L2. This last path receives specific designations in case of contributing to target RNA-directed miRNA degradation (TDMD) (84), or if it targets substrates that are structured aberrant/defective ncRNAs

(TUT–Dis3L2 surveillance pathway, TDS) (85, 86). Of note, Dis3L2 and other decay factors seem to be more sensitive to oligo- rather than to mono-uridylated RNAs (69).

There are astonishing examples of uridylation-mediated decay that come from widely studied eukaryotes: the degradation of mature miRNAs and siRNAs, in *Chlamydomonas reinhardtii* (87); the programmed removal of maternal RNAs during embryogenesis, in *Xenopus laevis*, *D. rerio* and *M. musculus* (45, 74); and histone mRNAs decay, in *Homo sapiens* (55, 70). A striking example worth mentioning is the repression of mirtron (a type of non-canonical miRNAs, derived from intronic regions of pre-mRNAs) expression, in *Drosophila melanogaster* (88). This insect possesses a cytoplasmic TRUMP (terminal RNA uridylation-mediated processing) complex consisting of the Tailor TUTase and the Dis3L2 exoribonuclease. The TRUMP complex associates with mirtron precursors, uridylate them, and impair their maturation by the endoribonuclease Dicer (89). It should also be noted that this complex sets a parallelism with the nuclear Trf4/Air2/Mtr4 polyadenylation complex (TRAMP complex, first described in *Saccharomyces cerevisiae*), which is another RNA surveillance pathway that helps the RNA exosome bind to its substrates and might also modulate its catalytic activity (49, 90).

1.4. Dis3L2, the exoribonuclease with a crush on "U"

It has become increasingly evident how uridylation is crucial in RNA processing and turnover, which is also demonstrated by the already mentioned conservation in eukaryotes. This is extended to the exoribonuclease that is specifically recruited to degrade the uridylated RNAs, Dis3L2 (38–41, 89, 91, 92).

1.4.1. Involvement in cellular processes

Beyond all the examples that were given above, it is also worth highlighting other key cellular processes where the Dis3L2-dependent degradation pathway is essential. These include the Dis3L2 role in rRNA biogenesis (93), endoplasmic reticulum-associated mRNA translation (94), cell differentiation (39, 40, 95) and proliferation (96, 97), T cell activation (98) and spermatogenesis (99). It may also appear implicated in antagonistic phenomena such as mitosis (100, 101) and apoptosis (102–104). Furthermore, Dis3L2 even aids a quality control mechanism termed nonsense-mediated decay (NMD), which is triggered by the presence of premature translation-termination codons (PTCs) in mRNAs (44, 105).

The example of cell differentiation is one of the most studied processes, reinforcing the opposite effects that uridylation can exert particularly on the Lin28–let-7 pathway, and their consequences in development [reviewed in (57, 62)]. The let-7 family of miRNAs is crucial, being well-conserved in

animals with bilateral symmetry, and is responsible for differentiation, silencing numerous human oncogenes and thus operating as tumour suppressor (106). There are two types of double-stranded let-7 precursors (pre-let-7), which are classified according to the number of protruding nucleotides at their 3'-end. Group I precursors [hereby referred to as pre-let-7(I)] have a 3' overhang with 2 nt, while the ones from Group II [hereafter called pre-let-7(II)] have a 3' overhang with 1 nt (47, 62).

In general, pre-miRNAs have to be processed by Dicer to become part of functional RNA-induced silencing complexes (RISCs) (107), which assist the binding between mature miRNAs and their mRNA targets, resulting in either their degradation or translation inhibition (10). To bind the 3'-end extremity of pre-miRNAs, Dicer requires a 2 nt single-stranded region. Therefore, in somatic cells, pre-let-7(I) molecules are directly processed by Dicer. In opposition, pre-let-7(II) must be first mono-uridylated by TUT4/7 (or alternatively mono-adenylated by TENT2) for it to be processed. The resultant mature let-7 will impair the expression of several pluripotency-related genes, further promoting cell differentiation (62, 108, 109). In stem cells, in order to perpetuate a proliferative undifferentiated state, diverse pluripotency factors are expressed, such as Lin28a and Lin28b. Lin28a/b are cytoplasmic RBPs responsible for the negative regulation of let-7 expression (75, 110). In this sense, Lin28a binds pre-let-7(I)/(II) and promote their oligo-uridylation by TUT4/7 (72, 111). The addition of a longer U-tail prevents maturation by Dicer, and recruits Dis3L2 for pre-miRNA degradation (39, 40, 95, 109, 112).

In summary, if human TUTases TUT4/7 mono-uridylate pre-let-7(II), then they are contributing to miRNA maturation by Dicer, whereas if the same TUTases form a complex with Lin28, they promote oligo-uridylation of pre-let-7(I)/(II), marking them for Dis3L2-dependent degradation (62, 113). It is evident that this Lin28–let-7 pathway has been gaining more relevance since this loss of tumour suppressor function of let-7, induced by Lin28 and mediated by Dis3L2, has been associated with human tumorigenesis. This is because the absence of mature let-7 miRNA elicits the expression of its targets, contributing to pluripotency, cell reprogramming and, eventually, cancer (62, 114).

1.4.2. Intracellular localization of Dis3L2

Dis3L2 has been previously described as being homogeneously distributed in the cytosol of different cell types (38–41). Nevertheless, Malecki and associates (38) showed that this enzyme could be additionally located in cytoplasmic foci contiguous to processing bodies (P-bodies). They also noted that, under glucose deprivation, Dis3L2 has total co-localization with P-bodies. This could be a cellular stress response mechanism, allowing faster degradation of uridylated transcripts.

For a better understanding, P-bodies are eukaryotic cellular compartments deprived of membranes, whose presence is constitutive in some cell lines, and whose formation can also be induced by stress. They are aggregates of RNA molecules and proteins (115, 116), and their function is not yet totally clear, despite being established that they serve as a storage of mRNAs in a stable untranslational state.

They might also function as specific, though non-critical, sites of mRNA decay, since they have been shown to contain mRNA degradation intermediates, in combination with enzymes related to mRNA decapping and 5'-3' RNA decay (117, 118). Conversely, recent investigations state that P-bodies sequestrate inactive enzymes, conferring the cell higher plasticity when answering to stress, and allowing for a rapid translation re-initiation as soon as environmental conditions become favourable (115, 116, 119).

Importantly, the discovery that Dis3L2 can co-localize with P-bodies was the first evidence of a link between degradation in the 3'-5' direction and these compartments (38). Furthermore, TUT4 has already been identified in P-bodies (120). In response to human Dis3L2 (hDIS3L2) deletion, it was shown that the number of P-bodies increased, presumably by facing an accumulation of non-degraded untranslated mRNAs inside them (41).

1.4.3. A peculiar homologue within the RNase II/RNB family

1.4.3.1. A widespread family of 3'-5' exoribonucleases

As mentioned earlier, Dis3L2 belongs to the RNase II/RNB family, highly conserved both in prokaryotes and eukaryotes. All members of this family display a processive 3'-5' exonucleolytic activity (121).

In prokaryotes, the representative enzymes of the RNase II/RNB family are RNase II and RNase R, which are key players in RNA degradation and quality control (3). Importantly, RNase R has been associated with virulence in several Gram-negative bacteria, since it impacts growth, motility, competence, adaptation to cold, amongst other physiological processes (121, 122). In what concerns their catalytic properties, RNase II and RNase R differ in the size of the final degradation products, and in their ability to unwind and digest structured RNAs (123). While RNase II releases a 4 nt end-product (124, 125), stalling 5 to 7 nucleotides before RNA double-stranded regions (126), RNase R originates 2 nt end-products, being capable of cleaving double-stranded molecules if they have at least a 5 nt single-stranded overhang on their 3'-end (127, 128). Besides functioning as a nuclease, RNase R also possesses helicase activity, not requiring to be associated with another enzyme that has such function to be able to degrade double-stranded RNAs (dsRNAs) (128–130).

In turn, eukaryotes may express three different homologues of this family: Dis3 (or Rrp44), Dis3L1 (or Dis3L), and Dis3L2, showing meaningful differences between them (see below). Whereas the presence of Dis3L1 is constrained to metazoans, the other two paralogs also appear in plants and some fungi. Therefore, the three homologues can be simultaneously found almost uniquely in vertebrates.

Considering that *S. pombe* has two RNase II-like enzymes (Dis3 and Dis3L2) (38), as well as six ncPAPs (6, 54), some of them with TUTase activity (like the caffeine-induced death suppressor protein 1 (Cid1), an orthologue of human TUT4/7) (131), it is a suitable yeast model to study this uridylation/Dis3L2-mediated RNA decay pathway. Furthermore, fission yeast also shares other mechanisms of gene expression regulation with more complex eukaryotes, namely gene silencing by RNA interference (RNAi) (132). This strongly supports that the RNA metabolism of higher eukaryotes is closer to that of fission yeast (*S. pombe*) rather than that of budding yeast (*S. cerevisiae*). *S. cerevisiae* only possesses one homologue of the RNase II/RNB family (Rrp44/Dis3) (29), and seems to be the only eukaryotic model organism that lacks 3' uridylation (38, 68, 133). Despite having two ncPAPs, Trf4 and Trf5, none of them has TUTase activity (6).

1.4.3.2. The characteristic protein domain organization and the properties conferred to each homologue protein

Notably, all members of the RNase II/RNB family have a comparable protein domain organization (Figure 1.2.). They are usually composed by two cold-shock domains (CSD1 and CSD2), a RNase II catalytic domain (RNB), and a S1 domain. The CSD1, CSD2 and S1 are considered oligosaccharide/ oligonucleotide binding fold (OB-fold) domains, which means they assist RNA binding (125). Meanwhile, as its name implies, the RNB domain is the one responsible for the 3'-5' exonucleolytic decay of RNA molecules, being highly conserved (124).
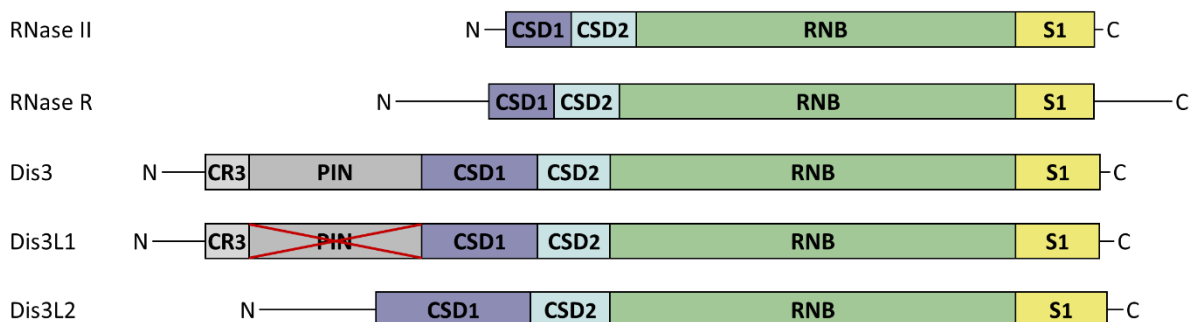
Figure 1.2. Protein domain organisation of enzymes from the RNase II/RNB family. Representation of protein domains present in representative members of the RNase II/RNB family: bacterial RNase II and RNase R, and eukaryotic Dis3 (or Rrp44), Dis3L1 (or Dis3L), and Dis3L2. While the PIN domain confers an additional endonucleolytic activity to Dis3, it is inactive in Dis3L1 (red cross on the domain). N: amino terminus. C: carboxyl terminus. Additional variable domains at N-term/C-term are not displayed. Dimensions are approximately in scale. (*From* Costa SM, Saramago M, Matos RG, Arraiano CM, Viegas SC. 2022. How hydrolytic exoribonucleases impact human disease: Two sides of the same story. FEBS Open Bio, *accepted*)

Whilst this arrangement is shared by prokaryotic RNase II/R and eukaryotic Dis3L2 (95), Dis3 and Dis3L1 evidence two additional N-term domains, both related to the association of these proteins to the RNA exosome: the Cysteine-Rich with three cysteines (CR3), and the PilT N-terminus (PIN) domains (34, 134). This last domain stipulates a dual role: a structural function since it is required for interacting with the exosome, as previously denoted (section 1.2.); and an endonucleolytic activity, enabling the enzyme to cut circular and linear single-stranded RNAs (circRNAs and ssRNAs, respectively) in a Mn^{2+}-dependent manner (135–137). This second feature relies on the presence of four essential residues that form the PIN active site (E97, D69, D177 and D146 in the human Dis3 homologue, hDIS3). So, while Dis3 (as part of the RNA exosome, mainly in the nucleoplasm) is able to cleave RNA molecules endonucleolytically, Dis3L1 (strictly belonging to the cytoplasmic isoform of the RNA exosome) does not, having an inactive PIN domain due to the absence of two of the four active site residues (E97 and D146 in hDIS3) (31).

Once CR3 and PIN domains are missing in Dis3L2, this enzyme acts independently of the RNA exosome, emphasizing its participation in an alternative RNA decay pathway. This autonomy was confirmed in *S. pombe* by the lack of co-localization of Dis3L2 (in the cytoplasm) with nuclear components of the exosome (38). It was further corroborated by not occurring co-precipitation of Dis3L2 with the subunits of the nuclear RNA exosome, both in *S. pombe* (38) and in human cell lines (39, 40).

It is also possible to distinguish the eukaryotic homologues according to their sequence conservation in their active site, within the RNB domain. This way, their amino acid signature might be DIDD (for Dis3 proteins), DVDD (if they are Dis3L1 enzymes), or DLDD (in the case of Dis3L2 ribonucleases) (38). Moreover, the aspartic acids (D) contained in these signatures are also engaged in the four aspartic acids that are critical for exoribonuclease activity, including an aspartic acid six amino acids upstream from these three (38).

In relation to their catalytic features, the homologues slightly differ in the size of the final degradation products they form: a minimum of 3 or 4 nt for Dis3 orthologues (budding yeast Rrp44 and hDIS3, respectively) (29, 31, 41); 4 nt for Dis3L1 orthologues (31); and 3 nt for Dis3L2 orthologues (38, 41).

Although there is no evidence of the ability of Dis3L1 to disrupt dsRNAs, it has already been reported that Dis3 can degrade structured RNAs, as long as they bear a single-stranded 3' overhang with at least 4 nt, which was corroborated both in the Dis3 orthologues from budding yeast and humans (29, 31, 34, 41). Surprisingly, studies based on Dis3L2 orthologues from *S. pombe* (SpDis3L2) and humans (hDIS3L2) demonstrated that these RNases just require 2 protruding nucleotides at the 3'-end of the structured RNAs in order to unwind them, and even manage to degrade them if they have blunt ends (38, 41).

1.4.4. The crystallographic structure of Dis3L2 and its preference for uridylated RNAs

Having as a starting point the knowledge about the protein domains of Dis3L2, inferred from multiple sequence alignments of diverse homologues, it remained to be understood how these domains were relatively arranged, and how the substrates would interact with them, be recognized, and subsequently degraded by the enzyme.

The resolution of the crystallographic structure of the mouse Dis3L2 (mDis3L2) in complex with a 14 nt oligo(U) (14U), by Faehnle and colleagues (95), revealed that the shape of the entrance channel for the RNA resembled a funnel (Figure 1.3.A and B), being overall akin to that found in *E. coli* RNase II, which is to be expected as both exoribonucleases have a similar protein domain organisation, albeit the CSD1 of Dis3L2 is slightly longer. Besides, it brought new insights into the structural interactions between the Dis3L2 enzyme and its RNA substrates (Figure 1.3.C). Using an endogenous substrate of the enzyme, pre-let-7 (whose critical importance was tackled above, in section 1.4.1.), they were able to test the preference *in vitro* of mDis3L2 for both the unmodified RNA and one with a 15U 3' stretch (similar to the one added by TUT4/7 *in vivo*), proving that the enzyme was significantly more efficient degrading the uridylated pre-let-7, which was supported by a stronger binding affinity to it (95).

This preference was supported by the fascinating finding of an extensive network of interactions between specific amino acid residues of mDis3L2 and the ribonucleotides of the 14U molecule trapped inside it, which allowed to establish a classification system with 3 zones of specificity for uridines: U-zones 1, 2 and 3 (Figure 1.3.C.). Accordingly, when the bound RNA was instead a ssRNA composed of just 15 adenosines (15A) or 15 cytosines (15C), many of those interactions were lost, rendering the binding affinity weaker (95). This was also in accordance with previous observations that poly(A) substrates had an inhibitory effect on the activity of the fission yeast orthologue, SpDis3L2 (38).

A.

B.

C.

U-zone 1

U-zone 2

U-zone 2

U-zone 3

Active site

13
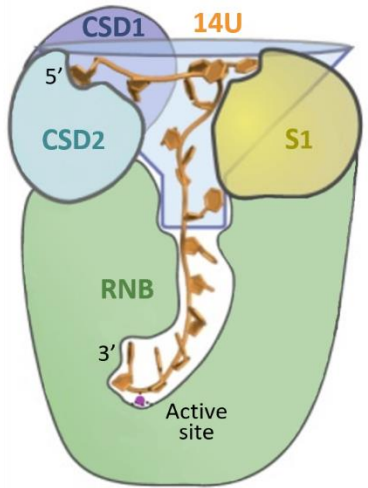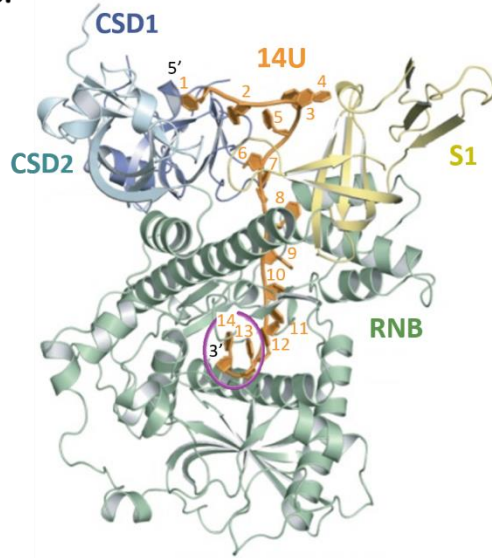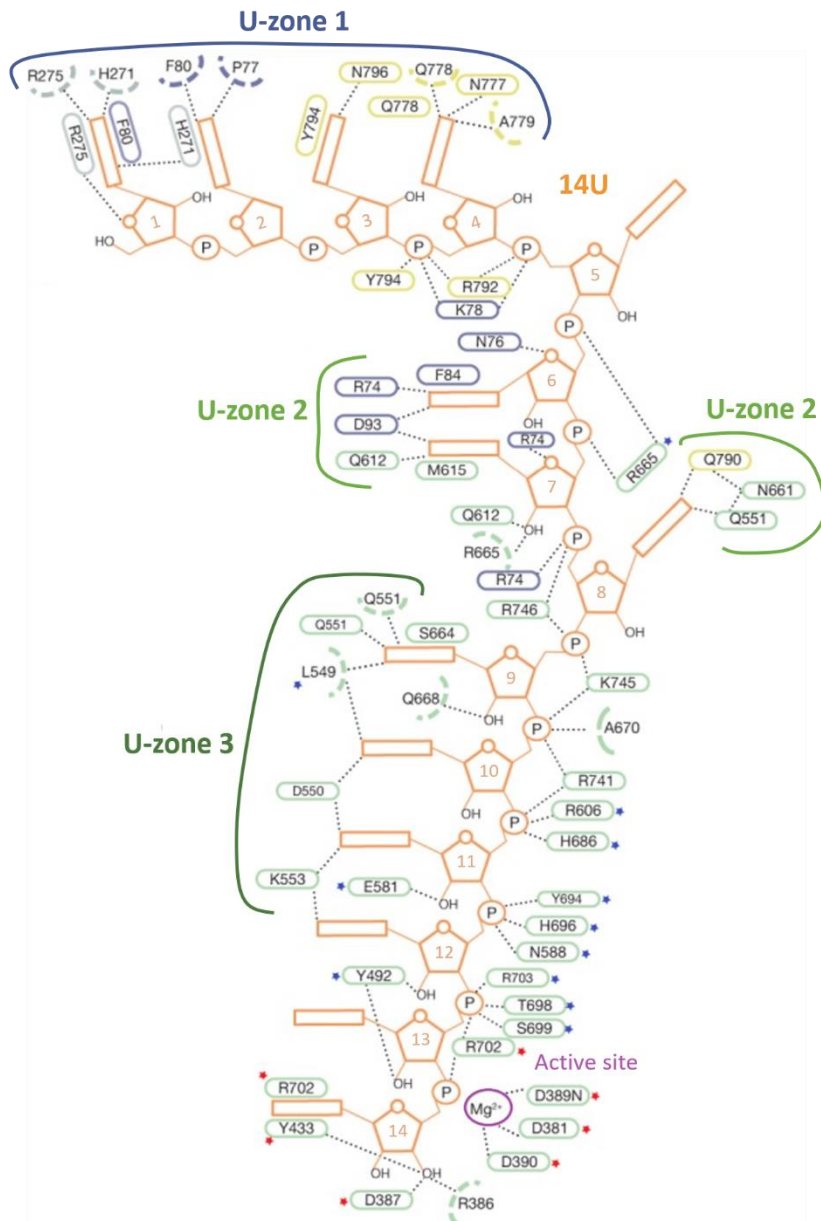
Figure 1.3. Schematic representation of mouse Dis3L2 structure in complex with an oligo(U) substrate, with protein–RNA interactions identified. RNA substrate is shown in orange (when nucleotides are numbered, they go from 1, at the 5'-end, to 14, at the 3'end). Protein domains are colour-coded according to Figure 1.2. **(A.)** Cartoon showing the funnel-like architecture of the enzyme when bound to a RNA substrate. The purple dot indicates the active centre of mDis3L2. **(B.)** Diagram of secondary structures formed by each domain and the positioning of the RNA substrate inside the protein. Purple circle marks the active centre of mDis3L2. **(C.)** Overview of interactions between specific amino acids of mDis3L2 and the RNA substrate, where the dashed lines illustrate hydrogen bonds and/or ion pairs. Interactions are established with the side chain (full circle) or with the main chain (dashed semi-circles) of the amino acid residues indicated. Tiny blue stars are pointing conserved residues of the protein which interact with the backbone of the 14U RNA, while tiny red stars are flagging conserved residues between mDis3L2 and budding yeast Rrp44 in the active site of the RNB domain. The three zones uridine-specifics are noted. U-zone 1 encompasses the "mouth" of the funnel, with the S1 domain along with the CSD lobe (CSD1 and CDS2 closely associated), covering ribonucleotides U1 – U4 of the RNA substrate, and works as a walkthrough for the RNA to the second zone. U-zone 2, keeping the metaphor, is the "stem" of the funnel, which is a narrower space including ribonucleotides U5 – U8, and provides an interface between the other two larger zones. U-zone 3 comprises the "core" of the RNB domain, harbouring ribonucleotides U9 – U12. The last 3'-end ribonucleotides, U13 and U14, are accommodated in the active centre of the enzyme interacting with the Mg^{2+} cation. [Adapted from Faehnle *et al.*, 2014 (95).]

In an attempt to obtain a structure from the fission yeast orthologue, Lv and associates (138) reported the resolution of the structure of a N-terminal truncated version of SpDis3L2 (since full-length SpDis3L2 failed to crystalize) without any RNA substrate bound to it. Considering this fact, it was not entirely unforeseen that the conformation of this SpDis3L2 RNA-free was distinct from its mouse counterpart. SpDis3L2 displays an open conformation in which the CSD2 appears side by side with the RNB domain instead of being placed on the top together with CSD1 and S1 (138). This way, the authors suggested that the binding of the RNA substrate would induce a radical conformational change leading to a funnel-like arrangement equivalent to the one studied for mDis3L2 (138). The precise mechanism by which this change of configuration would occur remains elusive.

1.4.5. Importance in human diseases

In the last decade, these eukaryotic enzymes of the RNase II/RNB family have entered the spotlight due to their implication in human diseases [reviewed in (139)], although the involvement of DIS3 and DIS3L1 have already been more widely characterised. For example, Multiple Myeloma has been associated with specific mutations in hDIS3 (140) and the intracellular ratio between its two possible isoforms (141), while Cerebellar Medulloblastoma has been related to hDIS3L1 expression levels (142).

Therewith, considering the diversity of processes in which Dis3L2 is engaged, it is conceivable that it may well be also a prominent contributing factor to the pathology of human diseases (Table 1.1.) [reviewed in (139, 143)].

The human *DIS3L2* gene (*hDIS3L2*) is annotated to chromosome 2 (Chr 2) (144), has fourteen possible splicing variants (145), and five confirmed isoforms (146). The canonical isoform 1 possesses 21 exons, resulting in 885 amino acids and a molecular weight of 99.3 kDa (147). Interestingly, when

consulting the platform of Human Protein Atlas (HPA, a project born in 2003 with the intention of mapping the whole human proteome relying on a wide range of omics approaches), it is possible to ascertain that the protein 'DIS3 mitotic control homolog-like 2' has detectable RNA expression in all analysed cell types, cell lines and tissues (including the brain). However, in what concerns protein expression, hDIS3L2 levels are more variable between tissues (148).

Germline mutations in *hDIS3L2* had already been associated with Perlman Syndrome (PRLMNS), a rare congenital overgrowth disease, with high neonatal mortality. Because it has an autosomal recessive inheritance, the PRLMNS only manifests itself if there is a biallelic abnormality in the gene that encodes for hDIS3L2 (100, 101, 143) (Table 1.1.). A few distinct heterozygous genomic variants were identified (100, 149, 150), although most cases appear to be related to homozygous deletion of exon 6 or 9 (100, 151). New-borns suffering from this syndrome usually have numerous foetal malformations, including altered facial features, neurological delays, organomegaly and renal abnormalities (100, 101, 143). The renal abnormalities may translate into nephroblastomatosis, a precursor of nephroblastoma, also known as Wilms' Tumour (WT) (152). This is a type of kidney cancer mainly found in children, namely in more than 60% of infants with PRLMNS who managed to survive beyond the neonatal period (100, 101, 153). There have already been some studies that examined sporadic Wilms' tumours. In one of them, 30% of the Wilms' tumours analysed had a partial or complete deletion of the *hDIS3L2* gene (100). In another study, it was proposed that, when there was a deletion in one allele of Chr 2 in the region where the *hDIS3L2* gene is located, the increased susceptibility to Wilms' Tumour in such cases could be due to the reduced expression of miR-562 miRNA, at least partially (154). The miR-562 is encoded in intron 9 of *hDIS3L2*, and one of its functions is to negatively regulate the expression of *EYA1* gene, which, in turn, is essential for proliferation in the early development of kidneys. As such, the downregulation of miR-562 would contribute to *EYA1* overexpression and, potentially, to tumorigenesis (154). A recent study has also shed light on some cellular pathways that might be affected by Dis3L2 to promote the overgrowth phenotype in both flies and humans (97).

Still regarding the Perlman Syndrome, it is curious to notice that, for example, the deletion in exon 6 resulted in a distinct subcellular expression pattern of hDIS3L2, with the enzyme being additionally found in the nucleus (100). A similar modification in subcellular localization was seen in patients with intranuclear inclusion body disease (INIBD). This is a neurodegenerative disease in which one of the hallmarks is the occurrence of nuclear inclusions in neurons and glial cells, which frequently contain proteins related to amyotrophic lateral sclerosis. In addition, these nuclear inclusions were found to retain the exoribonucleases XRN1 and hDIS3L2, which may therefore be responsible for their formation or degradation (114, 155).

Moreover, it was also uncovered that patients with Marfan-Like Syndrome (MLS) display chromosomal translocations involving Chr 2 that usually result in partial deletion of the *hDIS3L2* gene,

reducing its expression or abrogating the exonucleolytic activity of the enzyme (143, 156–158) (Table 1.1.). It was shown that this *hDIS3L2* truncation was possibly related to the overexpression of the C-type Natriuretic Peptide (CNP), encoded by the gene *NPPC* that regulates bone growth, subsequently leading to excessive bone growth and skeletal malformations, characteristics of this syndrome (158).

Impressively, mutations in *hDIS3L2* have so far been connected to various human cancers, namely colorectal cancer (159), anaplastic thyroid cancer (160) (Table 1.1.), hepatocellular carcinoma (161), and testicular germ cell tumour (162). However, it is intriguing that, when looking back at HPA, *hDIS3L2* expression was detected in all cancers tested, having presented weak to moderate cytoplasmic staining in the immunohistochemical data from most of them. In particular, in thyroid, liver and colorectal cancers, more than 90% of the studied patients had high/medium *hDIS3L2* expression. In fact, *hDIS3L2* higher expression has been proposed as a marker for adverse prognostic in liver cancer. By contrast, testis cancer was, in most cases, not stained in the presence of an anti-hDIS3L2 antibody (148).

Nonetheless, there is still a paucity of studies about the overall contribution of Dis3L2 to the progression of cancers. It is still controversial, as it seems to either encourage some or suppress the manifestation of others, varying considerably contingent on the type of cancer in question (114).

Table 1.1. Genomic variants in *hDIS3L2* associated with human diseases.

Disease in study	Alleles	Type	Region	Nomenclature Transcript, c.*	Polymorphism (dbSNP, NCBI)	Predicted effect (Protein, p.**)	Impact on hDIS3L2	Source
PRLMNS	Homozygous	Frameshift Deletion	e6	c.367-41553_602+40962del	n.a.	p.Val123Glufs*136	Loss of CDS2, RNB, S1 (truncated protein with 258 aa)	(100)
	Homozygous	In-frame Deletion	e9	c.951-?_1124+?del	n.a.	p.Gln318_Arg375del	Loss of RNB, and miR-562 binding region	
	Heterozygous	In-frame Deletion	e19	c.2394+5G>A (in i19 [a])	n.a.	p.Glu764_Asn798del	Partial loss of RNB, S1	
		Mutation	i	c.2289+37G>A	-	Not expressed	Not expressed	
	Heterozygous	In-frame Deletion	e9	c.951-?_1124+?del	-	p.Gln318_Arg375del	Loss of RNB, and miR-562 binding region	
		Missense Mutation	e13	c.1466G>A	rs387907116	p.Cys489Tyr [NV.PS (1)]	Single amino acid substitution in RNB (maintain 885 aa)	
	Homozygous	In-frame Deletion	e9	m.i.	n.a.	p.Gln318_Arg375del	Loss of RNB, and miR-562 binding region	(151)
	Heterozygous	Mutation	i5 [a]	c.367-2A>G	-	p.Val123Leufs*154	PTC after 32 aa, losing CSD1, CSD2, RNB, S1	(149)
		Mutation	e12	c.1328T>A	-	p.Met443Lys [NV.PS (2)]	Preserve CSD1, CSD2, S1, and RBN with single aa substitution (885 aa)	
OFCs and CTEV	Heterozygous	Missense Mutation	e8	c.878C>A	rs187563594	p.Pro293His	Likely deleterious and damaging	(150)
		Missense Mutation	e13	c.1570G>A	rs201308521	p.Glu524Lys	Likely deleterious and damaging	
n.a.	n.a.	Missense Mutation	e2	c.34C>A or c.34C>T	rs723044	p.Pro12Ser [NV-U]	Maintain 885 aa	UniProt (146)
		Missense Mutation	e13	c.1447C>G or c.1447C>T	rs186865544	p.Arg483Gly [NV-U.WT]	Maintain 885 aa	
		Missense Mutation	e14	c.1727G>A	rs200386096	p.Arg576His [NV-U.WT]	Maintain 885 aa	
MLS	Heterozygous	Deletion	Chr 2	2q37.1q37.3 deletion of 4,5 Mb	n.a.	m.i.	Predicted breakpoint at intron 6 of *hDIS3L2*, eliminating a 3'-end portion of *hDIS3L2* along with 45 genes	(158)
		m.i.	m.i.	m.i.	m.i.	m.i.	m.i.	
ATC	m.i.	Missense Mutation	e9	c.1082G>T	-	p.Trp361Leu	Likely tolerated, damaging or disease causing	(160)

Table 1.1. (continued)

_____

PRLMNS: Perlman Syndrome; **WT:** Wilms Tumour; **OFCs:** orofacial clefts; **CTEV:** congenital talipes equinovarus (or clubfoot) – constitutes a rare possible manifestation of PRLMNS, despite the remaining common symptoms of PRLMNS were not described in the patient of the referred study; **MLS:** Marfan-Like Syndrome; **ATC:** Anaplastic thyroid cancer; **Region:** Genomic region affected – might be a chromosome (Chr), or a particular exon (e) or intron (i) of the hDIS3L2 protein; **n.a.:** not applicable; **m.i.:** missing information (not mentioned on the original source); **aa:** amino acids

[a]   splice-site mutation;

*   in case of a known mutation, the denotation "c. n°A>B" means that, in the coding DNA reference sequence, the nucleotide A in position n° was replaced by nucleotide B;

**   in case of a known single amino acid substitution, the denotation "p. An°B" means that, on protein-level, the reference amino acid A in position n° was modified to an amino acid B;

[NV.PS] natural variant of the protein in Perlman Syndrome (manually annotated in UniProt);

[1]   p.Cys489Tyr (C489Y) assigned as pathogenic variant (ClinVar Accession Allele A: RCV000024121.3), being a substitution in a conserved residue, including in _S. pombe_, _D. melanogaster_, _A. thaliana_, and _M. musculus_;

[2]   p.Met443Lys (M443K) has unknown pathological significance, although possessing this mutation in at least one of the alleles enables long-term survival (due to retention of partial exonucleolytic activity);

[NV] natural variant of the protein (manually annotated in UniProt), apart from the two NV.PS;

[-U]   NV with unknown clinical implication in PRLMNS;

[-U.WT] NV with potential association with the pathogenicity of PRLMNS (still uncertain); was found in a patient with WT.

1.5. Objective of this Thesis

From all the aforementioned considerations, it is clear how Dis3L2 is an exoribonuclease extremely relevant for fundamental cellular functions, further contributing to the pathology of a variety of human diseases. It is equally evident that the fission yeast can represent an excellent microbial model in the study of RNA metabolism and decay in higher eukaryotes.

Although the enzyme has already been studied in a wide range of organisms (38–41, 89, 91, 92), it remains to be understood what impact it has on many cellular functions, as it has a vast range of targets, as well as what its precise mechanism of action is.

Therefore, this work aims to biochemically characterize the Dis3L2 of *Schizosaccharomyces pombe* (SpDis3L2), specifically by elucidating the role that some specific amino acids have upon the activity and specificity of the enzyme. In particular, we intend to identify residues potentially implicated in SpDis3L2 preference for uridylated substrates, which distinguishes this RNase from its counterparts.

To do so, we overexpressed and purified several mutant proteins of fission yeast Dis3L2 (with single amino acid substitutions), whose exonucleolytic activity was later examined over different RNA substrates.

2. Materials and Methods

All culture media, solutions and gels prepared are described in Table S.2.1. (Supplementary Materials section).

2.1. Protein sequence alignment

A multiple sequence alignment was constructed using the T-Coffee Server (available at http://tcoffee.crg.cat/apps/tcoffee/do:regular), and the amino acid sequences of ten eukaryotic homologues of the RNase II/RNB family: *Saccharomyces cerevisiae* Dis3, also named Rrp44 (ScDis3, NP_014621.1); *Schizosaccharomyces pombe* Dis3 (SpDis3, NP_596653.1); *Schizosaccharomyces pombe* Dis3L2 (SpDis3L2, NP_594510.1); *Caenorhabditis elegans* Dis3L2 (CeDis3L2, NP_498160.2); *Drosophila melanogaster* Dis3L2 (DmDis3L2, NP_728490.1); *Arabidopsis thaliana* Dis3L2, also designated SOV (AtDis3L2, NP_177891.1); *Mus musculus* Dis3L2 (MmDis3L2, NP_705758.1); *Homo sapiens* Dis3 (HsDis3, NP_055768.3); *Homo sapiens* Dis3L1 (HsDis3L1, NP_001137160.1); *Homo sapiens* Dis3L2 (HsDis3L2, NP_689596.4). The final figure was built using the BoxShade v3.2.1. Server (available at https://embnet.vital-it.ch/software/BOX_form.html).

2.2. Protein Overexpression

The plasmids containing the wild-type (wt) and mutant *SpDis3L2* genes (Table 2.1.) were previously transformed in BL21 (DE3) CodonPlus – RIL *E. coli* strain (*Stratagene*).

Table 2.1. List of plasmids used in the experimental work.

Plasmid	Relevant characteristics	Marker	Source
pDis3L2	Inducible expression vector, encoding wild-type *SpDis3L2* gene with a N-terminal His-SUMO tag	Amp^R	Rute Matos (CMA lab)
pDis3L2-A208E	pDis3L2 expressing a mutant SpDis3L2, where alanine (A) at position 208 was substituted by a glutamic acid (E)	Amp^R	Rute Matos (CMA lab)
pDis3L2-F209A	pDis3L2 expressing a mutant SpDis3L2, where phenylalanine (F) at position 209 was substituted by an alanine (A)	Amp^R	Rute Matos (CMA lab)
pDis3L2-D217P	pDis3L2 expressing a mutant SpDis3L2, where aspartic acid (D) at position 217 was substituted by a proline (P)	Amp^R	Rute Matos (CMA lab)
pDis3L2-L371R	pDis3L2 expressing a mutant SpDis3L2, where leucine (L) at position 371 was substituted by an arginine (R)	Amp^R	Rute Matos (CMA lab)
pDis3L2-L390F	pDis3L2 expressing a mutant SpDis3L2, where leucine (L) at position 390 was substituted by a phenylalanine (F)	Amp^R	Rute Matos (CMA lab)
pDis3L2-T456G	pDis3L2 expressing a mutant SpDis3L2, where threonine (T) at position 456 was substituted by a glycine (G)	Amp^R	Rute Matos (CMA lab)
pDis3L2-A457C	pDis3L2 expressing a mutant SpDis3L2, where alanine (A) at position 457 was substituted by a cysteine (C)	Amp^R	Rute Matos (CMA lab)
pDis3L2-R458T	pDis3L2 expressing a mutant SpDis3L2, where arginine (R) at position 458 was substituted by a threonine (T)	Amp^R	Rute Matos (CMA lab)
pDis3L2-L460I	pDis3L2 expressing a mutant SpDis3L2, where leucine (L) at position 460 was substituted by an isoleucine (I)	Amp^R	Rute Matos (CMA lab)
pDis3L2-L460V	pDis3L2 expressing a mutant SpDis3L2, where leucine (L) at position 460 was substituted by a valine (V)	Amp^R	Rute Matos (CMA lab)
pDis3L2-D461N	pDis3L2 expressing a mutant SpDis3L2, where aspartic acid (D) at position 461 was substituted by an asparagine (N)	Amp^R	Rute Matos (CMA lab)
pDis3L2-P512D	pDis3L2 expressing a mutant SpDis3L2, where proline (P) at position 512 was substituted by an aspartic acid (D)	Amp^R	Rute Matos (CMA lab)
pDis3L2-C560Y	pDis3L2 expressing a mutant SpDis3L2, where cysteine (C) at position 560 was substituted by a tyrosine (Y)	Amp^R	Rute Matos (CMA lab)

Table 2.1. (continued)

Plasmid	Relevant characteristics	Marker	Source
pDis3L2-A801C	pDis3L2 expressing a mutant SpDis3L2, where alanine (A) at position 801 was substituted by a cysteine (C)	Amp^R	Rute Matos (CMA lab)
pDis3L2-C804I	pDis3L2 expressing a mutant SpDis3L2, where cysteine (C) at position 804 was substituted by an isoleucine (I)	Amp^R	Rute Matos (CMA lab)
pDis3L2-K808H	pDis3L2 expressing a mutant SpDis3L2, where lysine (K) at position 808 was substituted by a histidine (H)	Amp^R	Rute Matos (CMA lab)
pDis3L2-R865L	pDis3L2 expressing a mutant SpDis3L2, where arginine (R) at position 865 was substituted by a leucine (L)	Amp^R	Rute Matos (CMA lab)
pDis3L2-R865T	pDis3L2 expressing a mutant SpDis3L2, where arginine (R) at position 865 was substituted by a threonine (T)	Amp^R	Rute Matos (CMA lab)

Amp^R: marker of resistance to ampicillin (the *bla* gene is encoded in the plasmid); **CMA lab**: Control of Gene Expression Laboratory, headed by professor Cecília Maria Arraiano (CMA) at ITQB NOVA.

For each *E. coli* strain of BL21 (DE3) CodonPlus – RIL harbouring each plasmid described in Table 2.1., a pre-inoculum (PI) was prepared in 50 ml Luria-Bertani Broth (LB) medium supplemented with 100 μg/ml ampicillin (Amp100) and 50 μg/ml chloramphenicol (Cam50), and incubated at 180 rpm, 37 ºC, overnight. In the next day, the culture was diluted in 1 L of Terrific Broth (TB) medium supplemented with Amp100 and Cam50, and grown at 150 rpm, 30 ºC to an OD_{600} of 0.5-0.8. Protein expression was then induced by the addition of 0.5 mM IPTG, and cells were incubated at 150 rpm, 16 ºC, overnight.

Cells were harvested by centrifugation at 5000 rpm, 15 ºC, for 10 minutes. Cellular pellets were stored at -80 ºC.

2.3. Cell lysis

The cellular pellets (acquired in the above section 2.2.) were resuspended in 20-25 ml Buffer C. Cell suspensions were lysed using a French® Press (*Thermo Electron*) at 1000 psi in the presence of 1 mM PMSF. Following lysis, 125 U of Benzonase® Nuclease 250 U/μl (*Sigma-Aldrich*) was added to the lysates to degrade all nucleic acids in solution. They were subsequently centrifuged at 15000 rpm, 4 ºC for 1 h to be clarified.

2.4. Protein Purification

The sample of pure RNase II protein from *E. coli* (EcRNase II) used as a control in some activity assays was kindly given by Vanessa Costa (CMA lab). The protein was stored in Buffer C (EcRNase II) with 50% (v/v) glycerol. Subsequent dilutions were made using the same Buffer C (EcRNase II) – see Table S.2.1. for detailed composition.

The wt and mutant variants of SpDis3L2 were purified using two chromatography techniques. Each protein was purified at least two times.

2.4.1. Affinity Chromatography

The cleared lysates (obtained in section 2.3.) were subjected to an immobilized metal affinity chromatography (IMAC), using HisTrap™ HP 1 ml columns (*Cytiva*) and the ÄKTA™ start (*Cytiva*) system. Columns were washed with 5 volumes of water and equilibrated with 5 volumes of Buffer A. After loading the sample, Buffer A continued to pass through the column until the UV absorbance stabilized, whereupon proteins were eluted after a linear gradient of Buffer B (from 25 mM to 500 mM of imidazole).

Small volume aliquots were taken and analysed by SDS-PAGE to monitor whether the full-length proteins of interest were present. Each sample, containing 5x Loading Buffer – Proteins, was boiled for 10 minutes, and subsequently loaded into a Novex™ WedgeWell™ 8-16% Tris-Glycine Protein Gel (*Invitrogen*). NZYColour Protein Marker II (*NZYTech*) was used as a molecular weight marker, and electrophoresis was conducted at 100-200 V in SDS-PAGE Running Buffer 1x. Thereafter, the gel was stained with BlueSafe (*NZYTech*), and the result was digitalized using the ImageLab v4.1. software implemented on the Gel Doc™ EZ Imager (*Bio-Rad*) system.

Afterwards, all fractions of each sample were gathered and concentrated in Vivaspin™ 6, 100 kDa MWCO columns (*Cytiva*) by centrifugation, at 4 ºC until the volume was lower than 500 µl. The sample was then centrifuged at 14000 rpm, 4 ºC for 20 minutes to remove any eventual precipitated proteins and salts before proceeding to the next purification step.

2.4.2. Size Exclusion Chromatography (Gel Filtration)

Concentrated samples were submitted to a size exclusion chromatography (also termed gel filtration), to improve protein purity. An ÄKTA Purifier 10 UPC (*Cytiva*) system was used with a Superdex™ 200 Increase 10/300 GL column (*Cytiva*), which was washed with 2 volumes of water and equilibrated with 2 volumes of Buffer C. The protein samples were injected, and UNICORN software

(*Cytiva*) was used to control the run and record the rise in UV absorbance that corresponded to the elution of the proteins. Protein samples were collected in 500 μl fractions.

An SDS-PAGE analysis was performed, in the same way as previously described (section 2.4.1.), to select the purer fractions to concentrate afterwards. Henceforth, the chosen fractions were pooled together and concentrated by centrifugation at 4 ℃, using Vivaspin™ 6, 100 kDa MWCO columns (*Cytiva*).

Lastly, 50% (v/v) glycerol was added to the concentrated protein samples. Protein concentration was determined using the Bradford Method (163), and they were stored at -20 ℃.

2.5. Thermal Shift Assay

Prior to the purification of all the SpDis3L2 variants, a Thermal Shift Assay (TSA) was performed with the wt protein purified with Buffers A1, B1 and C1 according to the procedures described above (section 2.4.), without the final step of 50% (v/v) glycerol addition. Protein sample was stored at 4 ℃.

To optimize the TSA, a prior [Protein] vs. [Dye] assay was performed. The purpose of this test was to identify the best concentrations of dye (SYPRO Orange, Protein Thermal Shift™ Dye kit, *Thermo Scientific*) and purified protein to use in the TSA to reach its full potential. Protein concentration was measured using NanoDrop™ 2000 Spectrophotometer (*Thermo Scientific*) and 20 μl reactions were prepared according to Table S.2.2.

The qTOWER³ G touch (230 V) (*Analytik Jena*) system was used to denature the samples, allowing a temperature gradient ranging from 20 ℃ to 95 ℃, 1 ℃/min, and fluorescence reading. The optimal concentration of dye and protein to use in the following TSAs were chosen through the analysis of the melting curves and the melting temperatures (T_ms) on qPCRsoft v4.0 software (*Analytik Jena*). Thereafter, 1x dye and 0.05 mg/ml of the wt SpDis3L2 protein (1 μg protein added to the 20 μl final reaction) were used in the TSAs.

TSA reactions (Tables S.2.3. and S.2.4.) were adapted from the composition of the reactions used in the commercial Durham pH and salt screenings (*Molecular Dimensions*), being added to each reaction: 10 μl screen reagent, 5 μl dye (SYPRO Orange), and 5 μl purified wt SpDis3L2 protein. Sample denaturation and melting curves analysis were conducted as described for the previous [Protein] vs. [Dye] assay.

2.6. Western Blot

Samples containing 1-2 µg of the purified proteins were mixed with 5x Loading Buffer – Proteins, boiled for 10 minutes, and resolved in a Novex™ WedgeWell™ 4-12% Tris-Glycine Protein Gel (*Invitrogen*).

A PVDF membrane (*Cytiva*) was activated by immersion in 100% (v/v) methanol, washed with di-distilled water (ddH$_2$O) and immersed during 15 minutes in transfer buffer 1x (TrB 1x, whose composition is in Table S.2.1.).

The western blot "sandwich" was assembled in a way all components should be embedded in TrB 1x, as follows: one sponge, three filter papers, the protein gel, the activated PVDF membrane, three filter papers and a sponge. This "sandwich" was placed inside the transfer tank, full of TrB 1x, and the transfer was carried out at 30 V, for 18 h, at 4 ºC, with continuous agitation of the buffer.

After transfer, the protein gel was stained with BlueSafe (*NZYTech*) to ensure that the transfer had occurred. All steps mentioned hereafter were performed at room temperature (RT). The PVDF membrane was blocked with 5% (w/v) non-fat milk in TBS-T 1x for 1 h, which was followed by a washing step of 15 minutes with TBS-T 1x. Thereafter, the membrane was incubated with primary antibody solution (Anti–His-tag diluted 1:5000 in TBS-T 1x) for 2 h, with gentle agitation, after which it was washed 3 times for 5 minutes with TBS-T 1x. This was followed by the incubation with the secondary antibody solution (Anti-IgG rat diluted 1:10000 in TBS-T 1x) for 1 h, and the washing procedure was then repeated. Afterwards, the chemiluminescent signal in the PVDF membrane was detected using the Western Lightning® Plus-ECL kit (*PerkinElmer*) and the ChemiDoc™ XRS+ (*Bio-Rad*) system.

Relative quantification of full-length SpDis3L2 protein was performed using the ImageLab™ v5.0 software (*Bio-Rad*).

2.7. Radioactive labelling of RNA substrates

RNA molecules used as substrates for the activity assays in the course of this work are presented in Table 2.2. Substrates were radioactively labelled at their 5'-end using [^{32}P]-γ-ATP and the T4 Polynucleotide Kinase (PNK, *Ambion*) enzyme.

Reactions were prepared with 10 µM RNA oligomer, 1x PNK Buffer A, 2 µl [^{32}P]-γ-ATP, 5 U PNK (*Ambion*) in a final volume of 15 µl, and incubated at 37 ºC, for 1 h. Enzyme inactivation was made at 80 ºC, for 5 minutes. Finally, the volume of the sample was adjusted to 25 µl with water, and illustra™ MicroSpin™ G-25 columns (*Cytiva*) were used, according to the manufacturers' instructions, to remove the unincorporated ribonucleotides.

Table 2.2. Nucleotide sequences of the RNA molecules used as substrates in the activity assays.

Substrate	Sequence
Adh[1]	5'-GUU UUG UAU AGA AAU CAA UG-3'
Adh4U	5'-GUU UUG UAU AGA AAU CAA UGU UUU-3'
Adh16U	5'-GUU UUG UAU AGA AAU CAA UGU UUU UUU UUU UUU UUU-3'
asAdh[2]	5'-CAU UGA UUU CUA UAC AAA AC-3'
Poly(U)	5'-UUU UUU UUU UUU UUU UUU UUU UUU UUU UUU-3'
Loop[3]	5'-GAG GCC UUU CGA GGC CUU GCU UUU UUU UUU UUU U-3'

[1] sequence at the 3'-end of *adh1* (alcohol dehydrogenase) mRNA (38); [2] antisense RNA complementary to Adh; [3] pre-let-7 miRNA mimic (Leemor Joshua-Tor laboratory) (164).

2.8. Activity Assays

2.8.1. Preparation of double-stranded RNA substrates

To obtain duplex structures, the RNA substrates were treated in two different ways.

In the case of dsAdh, the labelled Adh16U was annealed with the complementary unlabelled asAdh in a 1:10 molar ratio (8.3 nM Adh16U to 83.3 nM asAdh). The two RNAs were first diluted together in 20 mM Tris-HCl pH 7.5 and water was added to final volume. The mix was denatured at 80 ℃ for 5 minutes, and incubated for at least 45 minutes at RT to promote the annealing of the two molecules. This reaction was kept on ice until the moment of its addition to the activity assays' reactions.

In the case of Loop, the RNA was first diluted in 20 mM Tris-HCl pH 7.5 (or pH 8.0 for reactions with EcRNase II) and, to reach the final volume, it was added 5x Reaction Buffer (SpDis3L2 or EcRNase II, in accordance with the reactions concerned) diluted 1:100 in water. The sample was denatured at 80 ℃ for 5 minutes, and incubated for at least 2 h at RT to guarantee the folding of the RNA.

In an attempt to confirm the formation of the secondary structures, the migration of both substrates was analysed in a 20% polyacrylamide (PAA) non-denaturing gel at 200 V, 4 ℃, in TBE 1x buffer. The gel was exposed on a phosphor screen, and the radioactive signal was detected in a phosphor imager FLA-5100 (*FUJIFILM*) using the Image Reader FLA500 v.1 associated software (*FUJIFILM*).

2.8.2. Enzymatic reactions

Enzymatic reactions containing 8.3 or 83.3 nM of RNA substrate(s) and 1x Reaction Buffer were started by the addition of the enzyme at distinct concentrations (30, 40, 50 or 100 nM, which is indicated in the respective figure legend). The reactions were incubated at 30 ℃ (for assays with SpDis3L2) or at 37 ℃ (for assays with EcRNase II) for 45 or 60 minutes, based on whether the substrates were ssRNA or dsRNA molecules, respectively. Control reactions were run in parallel in the same conditions, but without adding the enzyme to the reaction.

Aliquots were withdrawn at different time points (specified in the respective figures) and a formamide-containing dye supplemented with 10 mM EDTA (STOP solution, Table S.2.1.) was added to the sample to stop the reaction.

Reaction products were denatured at 100 ℃, for 5 minutes, and resolved in a 20% PAA denaturing gel (of medium- or high-resolution, at 400 or 1800 V, respectively), in TBE 1x buffer. The gel was exposed on a phosphor screen at least 2 h or overnight, and results were obtained as above mentioned (section 2.8.1.).

When needed, the substrate consumption was quantified using ImageQuant TL software (*Cytiva*).

3. Results and Discussion

3.1. SpDis3L2 protein variants studied

3.1.1. Alignment of Dis3L2 homologues

In order to assess the influence of specific amino acid residues in the catalytic function performed by *S. pombe* Dis3L2 (SpDis3L2), they had to be individually modified. A multiple sequence alignment (Figure 3.1. and Figure S.3.1.) was constructed using the T-Coffee Server (available at http://tcoffee.crg.cat/apps/tcoffee/do:regular), submitting the sequences of diverse eukaryotic homologues of the RNase II/RNB family, being either Dis3, Dis3L1 or Dis3L2 proteins.

The modified residues were chosen across all protein domains, according to three criteria: (i) being conserved in Dis3L2 proteins but not in Dis3 nor Dis3L1; (ii) the corresponding residues in the mDis3L2 orthologue appear to interact with the oligo(U) substrate in the published RNA-bound structure (Figure 1.3., section 1.4.4.); (iii) the residues have already been detected as mutated in the hDIS3L2 orthologue, in a context of disease (Table 1.1., section 1.4.5.).

```
                        A208E           D217P              L371R            L390F
                        F209A
HsDis3L2    74 RINPKKF-HEAFIPSPDG--------DRDIFIDG.. 295 DYANTLFICRIVDWKEDCNFALGQLAKSL..
MmDis3L2    74 RINPKKF-HEAFIPSPDG--------DRDIFIDG.. 293 DFANTLFICRIIDWKEDCNFALGQLAKSL..
CeDis3L2    58 RINPKNY-QECFLDHPKG-----T-NHPDVLVLG.. 224 DFERFLYTAKITDWRAESVYADGRLVKLL..
AtDis3L2   172 RVNAHNR-NEAYCKIDGV--------PTDILING.. 401 NLEAELVAAQIVDWGEGSPFPVAQITHLF..
SpDis3L2   200 RILE-NH-RSAFACME-D--------IPDFYVDG.. 366 KHHDQLFSVEITRWSIYSRYPMGVLGEKL..
DmDis3L2   316 RVNRKNN-RQAFIIMSTD----REALERDGIVLL.. 485 DVSGLLYLAHILETD-CNGHCIAELIQPV..
SpDis3     272 NISTYNY-LEGSVVVPG--------YNKPVLVSG.. 416 -LVGRRIVVAIDLWDASSRYPEGHFVRDL..
ScDis3     278 QISEYNF-LEGSVSLPR--------FSKPVLIVG.. 444 -LLDKRIVISIDSWPTTHKYPLGHFVRDL..
HsDis3     252 RASRENY-LEATVWIHGD----NE-ENKEIILQG.. 393 -LEGRRIIVAIDGWPRNSRYPNGHFVRNL..
HsDis3L1   248 NVNKHRAQIEAFVRLQGASSKDSD-LVSDILIHG.. 387 -LQDFRVVVRIDSWESTSVYPNGHFVRVL..
```

CSD1   CSD2

```
                T456G A457C             P512D           C560Y         A801C C804I
                    R458T L460I/V                                         K808H
                         D461N
HsDis3L2   380 FTIDPSTARDLDDALS.. 438 QKVVPMLPRL.. 485 IIRSCTKLSY.. 732 LQKQADHCNDRRMAS
MmDis3L2   378 FTIDPSTARDLDDALA.. 436 QKVVPMLPRL.. 483 IIRSCTKLSY.. 730 LQKQADHCNDRRMAS
CeDis3L2   311 FTIDPKTARDLDDALH.. 375 QTVIPMLPRI.. 422 VIRSRVKLAY.. 667 IQEICTRCNDTKLAS
AtDis3L2   485 LTIDPSTATDLDDALS.. 543 QRKISMLPPL.. 590 IIRSCCKLSY.. 879 LSDVAAYCNERKLAA
SpDis3L2   450 ITIDPETARDLDDAVS.. 508 QKAIPMLPPL.. 556 VIKTCARLAY.. 797 CSLVAAHCNEKKEKS
DmDis3L2   569 FTIDPMTARDLDDAVS.. 627 NEVIHMLPQS.. 676 VINSCSQFAY.. 922 LHTLTKLANERKYNA
SpDis3     505 CSIDPPGCQDIDDALH.. 563 DKRIDMLPML.. 610 VIASKEAFSY.. 842 LIEICNGINYRHRMA
ScDis3     440 CSIDPPGCVDIDDALH.. 598 DKRIDMLPML.. 645 VIRSREAFSY.. 877 MDMICRNINRKHRNA
HsDis3     476 CSVDPPGCTDIDDALH.. 534 EKRIDMVPEL.. 581 VINSKASLTY.. 810 LADICKNLNFRHKMA
HsDis3L1   475 FSIDPKGCEDVDDTLS.. 533 DRRYDMLPSV.. 581 IIRSAYKLFY.. 833 LEELCRHINNRNQAA
```

RNB   / /   RNB   / /   RNB   / /   RNB

```
                R865L/T
HsDis3L2   789 YGVQKRIYCNAL.. 885
MmDis3L2   787 FGVQKRIYCNAL.. 870
CeDis3L2   724 YGVVKRVYVDKM.. 817
AtDis3L2   935 LGIERRIYYDQI.. 1055
SpDis3L2   860 YGISNRVDLSSD.. 927
DmDis3L2   979 SGHVISINYKMQ.. 1044
SpDis3     898 FGLEGIVYTKSL.. 970
ScDis3     933 FGVEGLIRLDNL.. 1001
HsDis3     867 YGLEGTVFFEEK.. 958
HsDis3L1   895 FGIKGAAYLKNK.. 1054
```

S1

Figure 3.1. Sequence alignment of several eukaryotic homologues from the RNase II/RNB family of exoribonucleases. Multiple sequence alignment was built using the T-Coffee and the amino acid sequences of ten homologues of the RNase II/RNB family: HsDis3L2 (*Homo sapiens* Dis3L2, NP_689596.4); MmDis3L2 (*Mus musculus* Dis3L2, NP_705758.1); CeDis3L2 (*Caenorhabditis elegans* Dis3L2, NP_498160.2); AtDis3L2 (*Arabidopsis thaliana* Dis3L2, also designated SOV, NP_177891.1); SpDis3L2 (*Schizosaccharomyces pombe* Dis3L2, NP_594510.1); DmDis3L2 (*Drosophila melanogaster* Dis3L2, NP_728490.1); SpDis3 (*Schizosaccharomyces pombe* Dis3, NP_596653.1); ScDis3 (*Saccharomyces cerevisiae* Dis3, also named Rrp44, NP_014621.1); HsDis3 (*Homo sapiens* Dis3, NP_055768.3); HsDis3L1 (*Homo sapiens* Dis3L1, NP_001137160.1). The alignment is partially presented, highlighting with different colours the residues mutated in the SpDis3L2 protein variants studied in this work. Complete alignment is in Figure S.3.1. The coloured boxes under the alignment indicate in which protein domain each region is located, and are colour-coded following Figure 1.2. (section 1.4.2.2.) and Figure 1.3. (section 1.4.4.); respective domain designations are inside each box. The two green boxes over the aligned sequences are pinpointing the four aspartic acid residues (D) of the active site of the enzymes, which are conserved in the RNB domain, and thus essential for its exonucleolytic function. The second green box includes the key residue that distinguishes the amino acid signature of the three paralogues: DLDD for Dis3L2 orthologues, DIDD in Dis3 proteins, and DVDD in Dis3L1 enzymes.

The residues in study are highly conserved and most of them exclusively found in Dis3L2 orthologues (namely L371, L390, T456, A457, R458, L460, P512, C560, A801, C804, K808 and R865), which may indicate they are relevant for Dis3L2 peculiarity, especially in what concerns its substrate preference. In the mDis3L2-14U structure, the corresponding residues to R458, K808 and R865 in *S. pombe* (thus R386, R741 and R792 in *M. musculus*) were demonstrated to interact with the backbone of the RNA substrate (95). Moreover, the substitution C489Y in the hDIS3L2, corresponding to C560Y in SpDis3L2, was found in a patient with Perlman Syndrome (100).

In particular, the substitution D461N was used as a control protein for the upcoming assays. The D461 residue is one of the four highly conserved aspartic acids of the catalytic centre of SpDis23L2 and was shown to be essential for the exoribonucleolytic activity (38). The role of this amino acid is conserved among other enzymes from the RNase II/RNB family (29, 33, 38, 39, 41, 128, 165, 166). It was already proven that it is necessary for catalysis, since its substitution, despite not undermining the RNA-binding capability of the RNase, completely supresses the exonuclease activity.

In this work one of the substitutions studied was D217P. A different substitution of this amino acid, D217A, had already been studied in SpDis3L2 (138). In a D461N background, the D217A variant revealed a decline in RNA-binding affinity, which also supported the role of the CSD1 domain, in which this residue is located, when engaging with the RNA substrates (138). Here, we studied the effect of the D217P substitution in a wild-type background and we expect that the aromatic ring pertaining to proline (P), and absent in alanine (A), could impact protein structure or the interaction with RNA in a different manner.

3.1.2. Optimization of protein purification conditions

Given that the purification of full-length SpDis3L2 is quite difficult, because the protein is very unstable and tends to degrade during the purification process, an optimization of the protein purification conditions was conducted. For that purpose, two Thermal Shift Assays (TSAs) were performed using the wt SpDis3L2 protein. Thus, we sought to find the conditions in which the stability of the protein was being maximised, hoping that the same would apply to the remaining variants, not losing awareness of the fact that the point mutations could also confer stability or instability to the different protein versions. This thermostability would be translated in a higher melting temperature (T_m) of the proteins, which was calculated through the analysis of the resultant melting curves.

Thus, wt SpDis3L2 protein was purified with a buffer containing 20 mM Tris-HCl pH 8.0 and 500 mM NaCl (A1, B1, and C1 buffers specified in Table S.2.1.). The first TSA tested the thermostability of the protein over different pH or salt conditions, using commercial solutions from Durham screenings (tested reagents are indicated in Table S.2.3.; T_ms obtained are in Table S.3.1.). The conditions in which the protein presented higher T_m were identified and, through further analysis of

the melting curves in each case (data not shown), we selected for subsequent use two buffers of different pH (100 mM MOPS pH 7.6, and 100 mM HEPES pH 8.0 – positions A12 and B3, respectively, in Tables S.2.3. and S.3.1.), and four salt-containing solutions (200 mM KCl, 5 mM MgCl2, 400 mM NaCl, and 400 mM NH4Cl – positions C10, C12, D5 and D11, respectively, in Tables S.2.3. and S.3.1.).

Each one of the buffers was combined with one or two salts, resulting in several different solutions (Table S.2.4.), which were used for the second TSA, aiming to find the mixture that would greatly promote the thermostability of the proteins. Considering the T_ms obtained for the proteins in this second TSA (Table S.3.2.) and the melting curves in each condition (data not shown), the solution which seemed most suitable, and which was henceforth used as a purification buffer for the wt and all SpDis3L2 variants, was 100 mM MOPS pH 7.6 + 200 mM KCl (position B2 in Tables S.2.4. and S.3.2.).

3.1.3. Purification and Quantification of the Proteins

All genes coding for the desired protein versions of SpDis3L2 were previously cloned with a N-terminal His-SUMO tag in a plasmid under the control of the T7 promoter. All proteins were overexpressed in BL21 (DE3) CodonPlus – RIL *E. coli* strain. This strain possesses a chromosomic copy of the T7 RNA Polymerase under the control of lacUV5 promoter, as well as the *lacI* gene. It also has an additional plasmid (pRIL) that confers resistance to chloramphenicol (Cam^R) and carries extra copies of three tRNA genes (*argU*, *ileY,* and *leuW*). Those tRNA genes allow the expression of rare codons in *E. coli* that are frequent in eukaryotic proteins, thus releasing the expression of the proteins of interest from possible constraints due to codon usage (167).

Protein purification comprised two steps relying on different purification methods to improve the yield of full-length SpDis3L2 obtained at the end. First, all recombinant His-tagged protein variants were purified by affinity chromatography, using HisTrap columns. The collected fractions were concentrated and then a size exclusion chromatography was performed, using a Superdex 200 Increase 10/300 GL column, suited for improving the purity of the purification of proteins with molecular weights between 10 and 600 kDa, therefore including full-length recombinant SpDis3L2 enzyme versions, with approximately 116 kDa. By analysing the peaks of the chromatograms obtained, as well as the SDS-PAGE gels of the collected fractions, we could conclude that full-length SpDis3L2 was eluted mainly between the 10 and 12 ml after sample injection in the system.

Representative results of the wt SpDis3L2 purification are displayed in Figure 3.2. The remaining results regarding the mutant variants purified, which were further used in the activity assays, are shown in the Supplementary Material section (Figure S.3.2. – Figure S.3.10.).

A.

B.
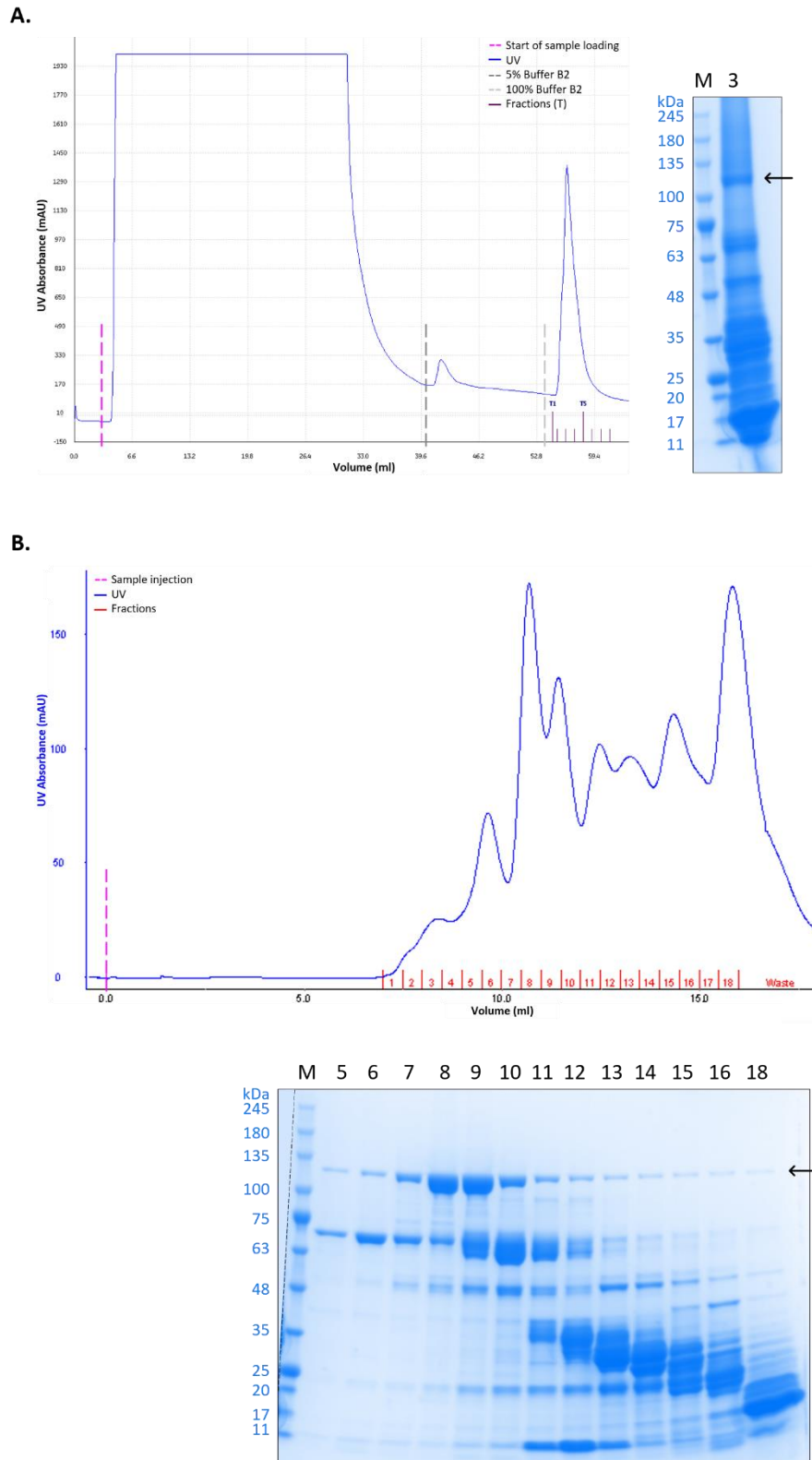
Figure 3.2. Representative results of the wt SpDis3L2 protein purification. **(A.) Left panel** – Protein elution profile in the affinity chromatography purification step. **Right panel** – SDS-PAGE with collected fractions during this 1st purification step. **(B.) Top panel** – Chromatogram obtained in the size exclusion chromatography step. **Bottom panel** – SDS-PAGE with collected fractions during this 2nd purification step. Samples loaded in the Novex WedgeWell 8-16% Tris-Glycine Protein Gels are numbered on the top of respective lanes according to the chromatograms. 'M' indicates the lane with the molecular marker (NZYColour Protein Marker II, *NZYTech*), and corresponding band sizes are denoted next to them (in kDa). Arrows are pointing the bands corresponding to full-length SpDis3L2.

From all the set of mutant versions of SpDis3L2 expressed, only four of them we were not able to obtain pure protein. These four mutants were D217P, L390F, R458T, and L460V. They were successfully purified by affinity chromatography but degraded after this first purification step. This may be due to the distinct biochemical properties of the amino acids that have been replaced. In fact, such point mutations cause changes in size, charge or contribute to the pH in a different manner. That might disturb the conformation of the full-length protein, destabilizing its structure, and subsequently leading to its degradation. Alternatively, it could also happen that proteins precipitated, in particular during the concentration step prior to the gel filtration.

To obviate these issues, we could try some strategies. For instance, we could perform the two purification steps on the same day, instead of performing them on successive days as it has been done so far. Also, the concentration of the samples before gel filtration could be reduced, for instance by keeping twice the sample volume and doing that second purification step twice. Another possibility would be to choose a fraction corresponding to the UV peak observed in the elution profile of the protein in the first purification step, and use it, without concentrating it, for the gel filtration. Finally, the addition of L-Glutamic acid and L-Arginine to the protein purification buffer could also be considered, as it has been reported that the combination of these two amino acids in solution enhances protein stability and solubility (168, 169).

Following protein purification, the concentration of each protein was determined by the Bradford Method (163). Given that each purified protein had a different degree of contamination by smaller degraded proteins, this type of measurement would not be an equally good approximation to the actual concentration of the full-length protein in the sample. For that reason, a subsequent Western Blot was carried out to correct the effective concentration of the full-length protein in solution, as exemplified in Figure 3.3.B. According to the concentrations determined by Bradford method, the same amount of each protein was applied in the gel for the western blot analysis. We then performed a relative quantification of the concentration of the band corresponding to the full-length protein, obtained in the previously mentioned western blot, and estimated the correction that needed to be done for the initial concentration of the protein in each case. A repetition of the western blot was performed with the corrected volumes to confirm the accuracy of this correction. It was observed that the concentration of the full-length proteins had become more similar, as illustrated in Figure 3.3.C.
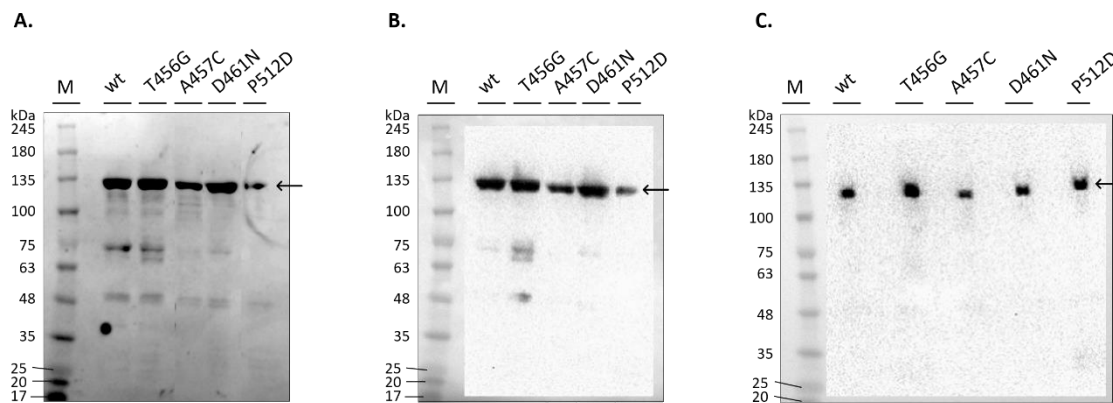
Figure 3.3. Representative SDS-PAGE and Western Blot analyses of recombinant SpDis3L2 protein variants. SpDis3L2 versions are identified on top of respective lanes. 'M' indicates the lanes with the molecular marker (NZYColour Protein Marker II, *NZYTech*), and corresponding band sizes are denoted next to them (in kDa). Arrows are pointing the bands corresponding to full-length SpDis3L2. (**A.**) 1.5 µg of proteins (according to the quantification by the Bradford Method) were loaded in a Novex WedgeWell 4-12% Tris-Glycine Protein Gel, which was stained with BlueSafe (*NZYTech*). (**B.**) 1.5 µg of proteins (according to the quantification by the Bradford Method) were loaded in a Novex WedgeWell 4-12% Tris-Glycine Protein Gel and a western blot was performed. The chemiluminescent signal detected in the PVDF membrane is presented. (**C.**) Protein concentrations were corrected based on the relative quantification of the bands corresponding to the full-length proteins (pointed with an arrow) in the PVDF membrane shown in **B.** Then, 1.0 µg of proteins (according to the corrected protein concentrations) was loaded in a Novex WedgeWell 4-12% Tris-Glycine Protein Gel and a western blot was performed. The chemiluminescent signal detected in the PVDF membrane is presented.

3.2. Exoribonucleolytic Activity Assays

To investigate whether the single amino acid modifications in the enzyme could affect the way it degrades its target RNA molecules, we analysed the exoribonucleolytic activity of each of the SpDis3L2 mutant variants on diverse synthetic substrates (either single- or double-stranded RNAs), and compared it with the regular activity of the wt protein.

Five types of activity assays were performed, whereby reaction times and enzyme concentrations were optimised, depending on the RNA substrates used in each case (whose sequences are presented in Table 2.2., section 2.7.): (i) a ssRNA – poly(U); (ii) two ssRNA substrates simultaneously – Adh and Adh4U; (iii) three ssRNA substrates simultaneously – Adh, Adh4U and Adh16U; (iv) a double-stranded RNA – dsAdh, formed by hybridization of two ssRNA substrates (Adh16U and asAdh), leaving a 16 nt 3' single-stranded overhang; (v) a structured RNA with a double-stranded region – dsLoop, originated by folding of a ssRNA substrate (Loop), leaving a 16 nt 3' single-stranded overhang. These RNA substrates were radioactively labelled at their 5'-ends to allow the monitoring of the reaction products, since the activity in study was 3'-5' exonucleolytic. Results are described and discussed in the next sub-sections.

3.2.1. Activity Assays with poly(U) substrate

3.2.1.1. Determination of the end-product of the Dis3L2 enzyme

First, a single poly(U) substrate was used to assess if the enzyme versions were active and able to degrade efficiently an uridine-based substrate.

It had already been demonstrated that Dis3L2 is capable of degrading the RNA molecules to a final size of 3 nucleotides (38, 41). So, to confirm the activity of the wild-type protein, similar activity assays were performed with the poly(U) substrate, and either the purified wt SpDis3L2 or the RNase II from *E. coli* (EcRNase II). This last enzyme was used as a size marker, since it has been established that the smallest product it produces has 4 nt (124, 125). A 20% PAA denaturing gel of high-resolution was performed to disclose the size of the reaction products at the nucleotide level (Figure 3.4.).
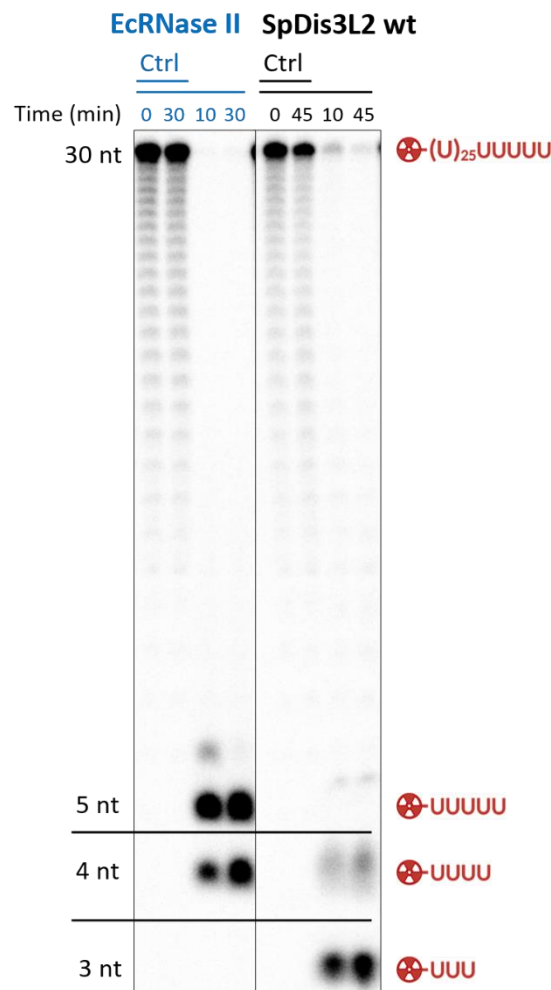
Figure 3.4. Exoribonucleolytic activity of *E. coli* RNase II and *S. pombe* Dis3L2 over poly(U) RNA substrate. 8.3 nM of poly(U) substrate was incubated either with 40 nM EcRNase II for 30 minutes at 37 ºC, or with 30 nM wt SpDis3L2 for 45 minutes at 30 ºC. Control reactions (Ctrl) were incubated for the same time without any enzyme. Samples were taken and reaction was stopped at the time-points indicated above each gel lane. RNA substrates and degradation products were separated through migration in a 7 M urea/20% PAA denaturing gel. The size of the molecules is depicted in line with the corresponding bands alongside the gel.

The results show that the substrate was being consumed in the presence of each of the enzymes. Thus, both RNases were active and presented the expected 3'-5' processive exoribonucleolytic activity, characteristic of the RNase II/RNB family to which they belong (124, 170). It was also possible to notice the accumulation of specific reaction products along time, with the smallest of the anticipated size in the two cases: 4 nt for EcRNase II, and 3 nt for wt SpDis3L2, in accordance with previous studies (38, 124).

Both EcRNase II and SpDis3L2 also presented products with up to 2 nt bigger than the final product, albeit in smaller quantities. When the substrate becomes smaller, it is known that these proteins start to present a distributive activity rather than processive (170). As RNA fragments are smaller, there are multiple cycles of release and re-association between the RNA and the enzyme (the so-called distributive activity), instead of the RNA degradation proceeding in a single event until the size of the final product is attained (processive activity).

3.2.1.2. Exoribonucleolytic activity of SpDis3L2 variants

Using again the poly(U) as a substrate, activity assays were carried out for all the protein versions (Figure 3.5.). The activity of all SpDis3L2 mutant variants was analysed by comparing it to the one exhibited by the wt protein.

As expected, D461N mutant did not show any catalytic activity (Figure 3.5.). This version was also used as control to the purification procedure, confirming that there were no contaminants in our samples. Because the same purification procedure was followed for all versions of the enzyme, this implies that the activity observed in the assays with all remaining SpDis3L2 variants is specific of each one of them.

All the other mutated versions of SpDis3L2 were active and able to degrade the poly(U) RNA. By looking at the gels, it is possible to see the full-size substrate disappearing, whereby the small specific reaction products accumulate over time. When focusing on the fraction of substrate left intact at the end of the total reaction time, the P512D and C560Y mutants appeared to be marginally less active.

As previously discussed, SpDis3L2 releases a mixture of 3, 4 and 5 nt products (Figure 3.4., section 3.2.1.1.). All the active mutant versions presented a similar degradation pattern when compared to the wt. However, while a second level of products larger than 3 nt was evident for most of the SpDis3L2 variants, it was much more subtle for P512D which appeared to be capable of being processive until the end, reaching directly the 3 nt. In contrast, the A457C mutant seemed to even accumulate a third level of small decay products, which could be a consequence of decreased processivity/higher distributivity of this enzyme version.
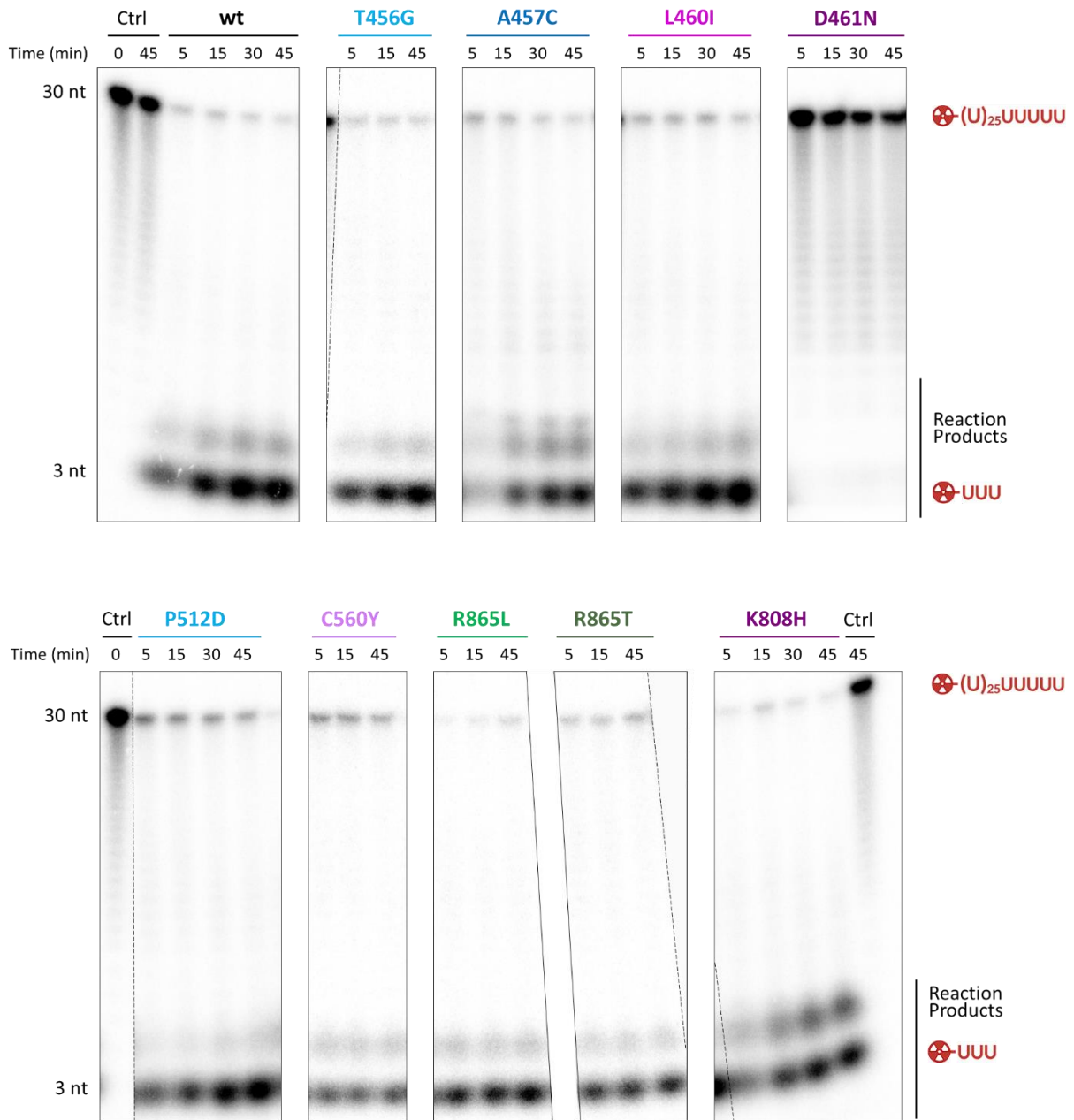
Figure 3.5. Exoribonucleolytic activity of SpDis3L2 wild-type and mutants on poly(U) RNA substrate. 8.3 nM of poly(U) substrate was incubated with 50 nM of each SpDis3L2 variant (named on the top) for 45 minutes at 30 ℃. Control reactions (Ctrl) were incubated for the same time without any enzyme. Samples were taken and reaction was stopped at the time-points indicated above each gel lane. RNA substrates and degradation products were separated through migration in a 7 M urea/20% PAA denaturing gel. The size of the molecules is depicted in line with the corresponding bands alongside the gel. The assay was performed one time (for D461N, C560Y, R865L, R865T and K808H) or two times (for wt, T456G, A457C, L460I and P512D) with proteins from independent purifications.

3.2.2. Activity assays with a mix of two ssRNA substrates (Adh and Adh4U)

To investigate whether the original residues in the modified positions have a prominent role in the preference of SpDis3L2 for uridylated substrates, which is one of the most remarkable singularities of Dis3L2 protein, we carried out activity assays with two different ssRNA present at the same time.

For this purpose, we used the mRNA sequence of the *adh1* gene from *S. pombe*, which encodes for an alcohol dehydrogenase (herein referred to as Adh) (38). This transcript has previously been reported to be a target of SpDis3L2. As a matter of fact, when the enzyme was first characterised in fission yeast, in strains with a *lsm1Δdis3l2Δ* background, the basal levels of *Adh* were significantly altered, and it was validated by Northern Blot that they were increased. By 3' RACE analysis it was further seen in these strains that most of these mRNAs were shortened and uridylated (38).

The assays presented in Figure 3.6. show the exoribonucleolytic activity of each SpDis3L2 version over an equimolar mix of Adh and Adh4U, with the second RNA corresponding to the Adh sequence with four extra uridines at its 3'-end.

C.
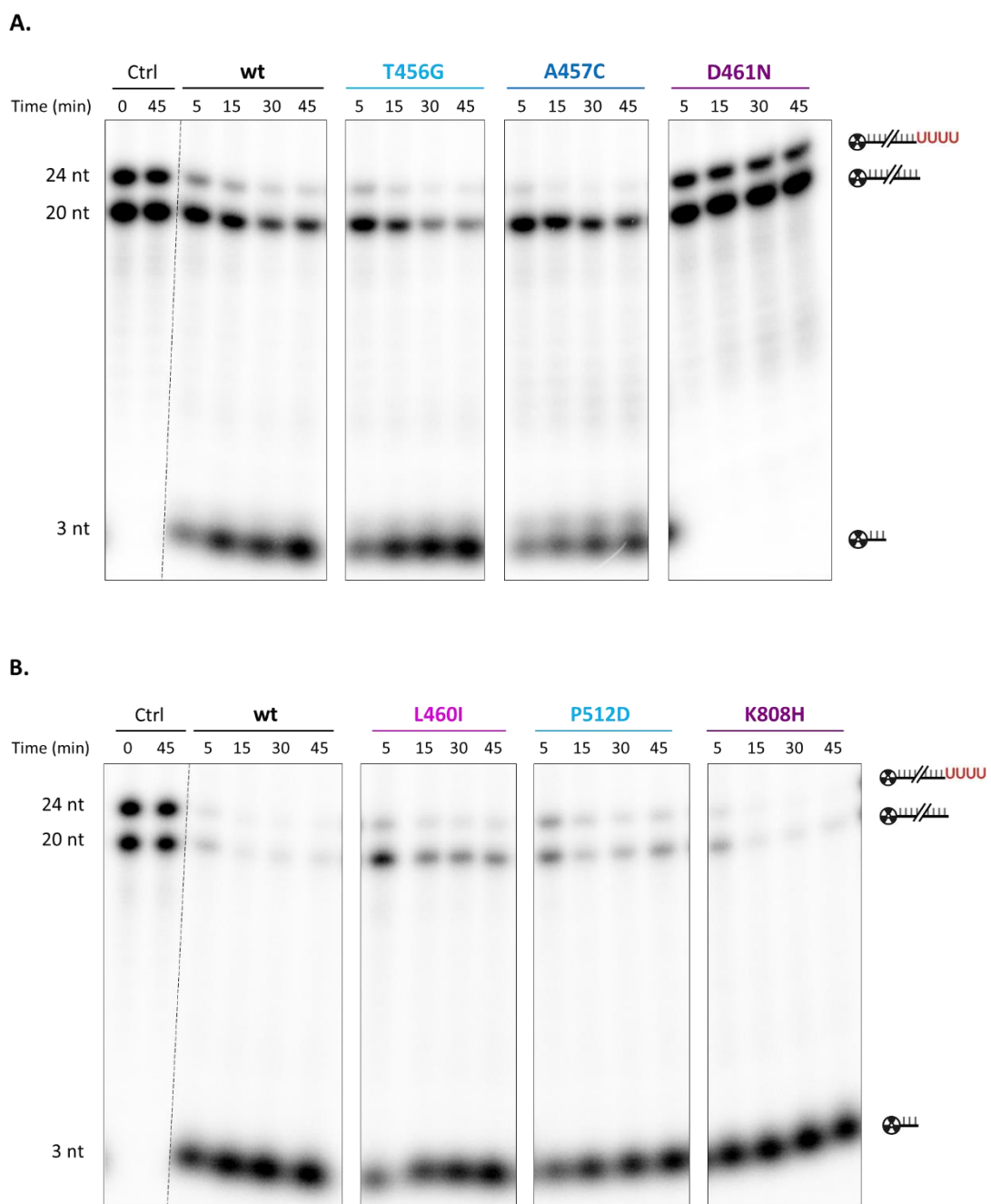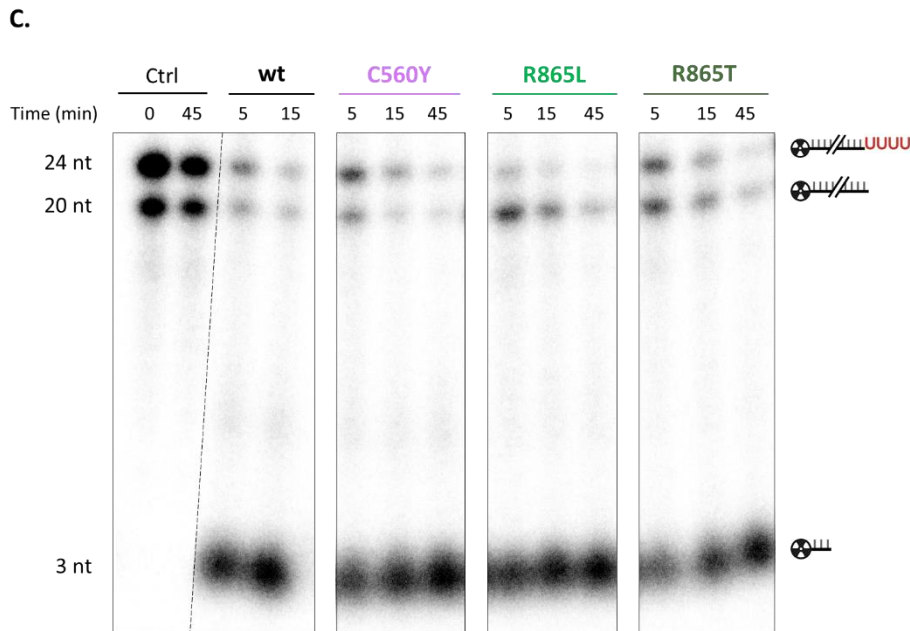
Figure 3.6. Exoribonucleolytic activity of SpDis3L2 wild-type and mutants on two different ssRNAs, illustrating their substrate preference. 8.3 nM of Adh and Adh4U were simultaneously incubated for 45 minutes at 30 ºC with different concentrations of each SpDis3L2 variant (named on the top): 40 nM (**A.**), 50 nM (**B.**), or 30 nM (**C.**). Control reactions (Ctrl) were incubated for the same time without any enzyme. Samples were taken and reaction was stopped at the time-points indicated above each gel lane. RNA substrates and degradation products were separated through migration in a 7 M urea/20% PAA denaturing gel. The size of the molecules is depicted in line with the corresponding bands alongside the gel. The assay was performed one time (for D461N, C560Y, R865L, R865T and K808H), two times (for T456G, A457C, L460I and P512D) or three times (for wt SpDis3L2) with proteins from independent purifications.

When these two ssRNAs were present at the same time in the reaction, it was possible to disclose the preference of the wt protein for the uridylated substrate, since Adh4U was degraded faster than Adh. In Figure 3.6.A. it is even visible that the wt enzyme started to consume the Adh RNA just after Adh4U had become scarce (after 30 minutes of reaction). The D461N mutant, as anticipated, had no catalytic activity. For the other mutants in study, the majority of them continued to exhibit the same aptitude, except for the C560Y mutant which appeared to degrade more rapidly the Adh instead of the Adh4U. This subtle inversion of preference needs to be further confirmed. For that we should repeat the protein purification and the activity assays in order to replicate these results. Also, the fact that the intensity of the radioactive signals of the two substrates were more disparate in the last gel (Figure 3.6.C.) may also difficult the interpretation, not allowing as much confidence in this visual qualitative comparison as in the gels above.

Curiously, some of the mutants seemed to have increased specificity towards the uridylated substrates, inasmuch they consumed less the Adh substrate than the wt protein which was serving as reference. Such observation was more noticeable for the A457C and L460I mutants. However, this specificity was not reproducible between assays with the A457C mutant, whereby it would still need to be confirmed. These modifications are found in relative proximity of the active centre, in the RNB

domain. The A457C mutant gains a polar amino acid and the possibility to establish additional disulfide bonds with other sulfur-containing amino acids of the protein. In turn, despite L460I being a substitution between amino acids with similar properties, the introduced amino acid corresponds to the one found in the equivalent position in the Dis3 paralogues. Considering this, we would expect that the catalytic activity could be impaired, instead of observing an increased preference for uridylated substrates. To help explaining this result, we could model this alteration in the protein.

3.2.3. Activity assays with a mix of three ssRNA substrates (Adh, Adh4U and Adh16U)

To further comprehend if the enzymes' preference for uridylated substrates would differ according to the U-tail size, an assay with three ssRNA substrates was performed (Figure 3.7.). In this reaction, besides Adh and Adh4U, we also added Adh16U, which corresponds to the Adh sequence with 16 additional uridines attached to its 3'-end.

If the enzymes kept their substrate predilection, it would be reasonable to suspect that they would predominantly hydrolyse both substrates with U-tails, and later the non-uridylated form of that RNA. We only used the wt protein and four other SpDis3L2 mutants (A457C, L460I, P512D and K808H) since this was meant to be an exploratory assay. Also, to guarantee the results reliability and proteins integrity, we only tested the proteins that had been most recently purified at the time of the assay.
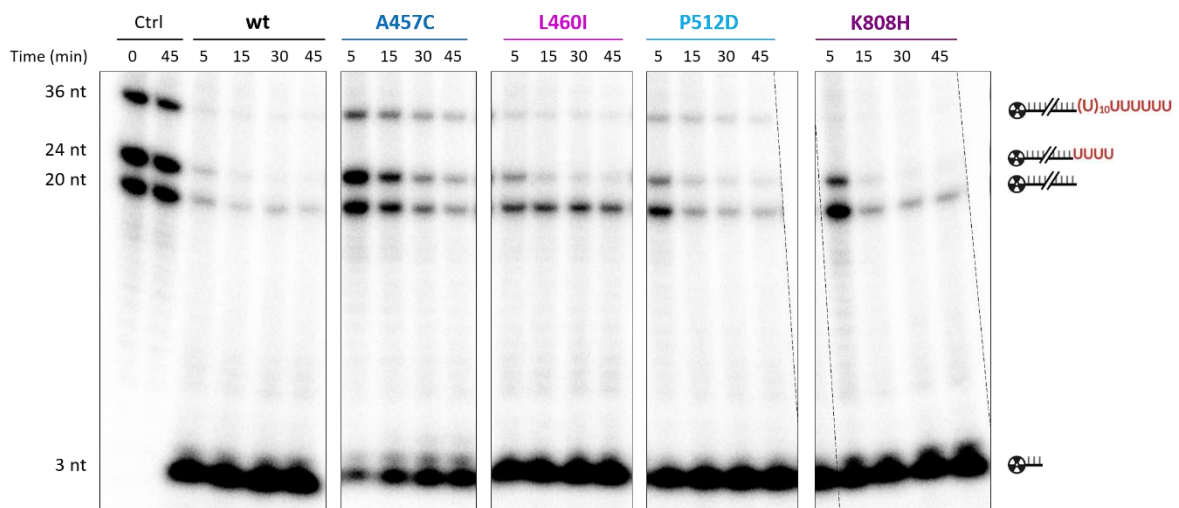
Figure 3.7. Exoribonucleolytic activity of SpDis3L2 wild-type and mutants on three ssRNA, demonstrating their preference for uridylated RNA substrates. 8.3 nM of Adh, Adh4U and Adh16U were simultaneously incubated with 100 nM of each SpDis3L2 variant (named on the top) for 45 minutes at 30 ºC. Control reactions (Ctrl) were incubated for the same time without any enzyme. Samples were taken and reaction was stopped at the time-points indicated above each gel lane. RNA substrates and degradation products were separated through migration in a 7 M urea/20% PAA denaturing gel. The size of the molecules is depicted in line with the corresponding bands alongside the gel. This assay was performed twice.

When comparing the degradation of the three ssRNA substrates by a certain protein variant, we should keep in mind that the radioactive signal of Adh16U in the images was slightly weaker when compared to the one of both Adh and Adh4U (see the lane corresponding to 0 min, Control in Figure 3.7. for comparison). It is perceivable that all analysed SpDis3L2 mutants preferred the uridylated substrates over the non-uridylated one, which resembled the behaviour seen in the previous assays with Adh and Adh4U (Figure 3.6., section 3.2.2.). Particularly for the L460I mutant, it is visible that it is more specific towards uridylated RNA molecules, which supports the above results. In this assay, it is more evident the A457C mutant was less active than the wt protein (Figure 3.7.).

To clarify our conclusions, we performed a quantitative analysis, by determining the disappearance of each substrate along time (Figure 3.8.).

These results further support the preference of all the tested enzymes for uridylated substrates. Moreover, Adh4U and Adh16U were degraded in a similar rate (Figure 3.8.), confirming that SpDis3L2 prefers uridylated substrates, and also showing that the size of the U-tail is not important. Comparing the activity of the different enzyme versions, we confirmed that the A457C mutant was indeed the least active. It retains the preference for uridylated substrates although the same is not strongly demarked.

It is also of note that, for the A457C mutant, the reaction velocity was practically held throughout the reaction, while for the wt and T456G, L460I, P512D and K808H mutants, the reaction velocity was higher in the first 5 to 15 minutes. This observation was made previously, whereby the degradation efficiency of SpDis3L2 was shown to decrease over time of the assay due to progressive loss of catalytic activity at 30 °C (38).

Regarding the L460I and K808H mutants, it is interesting to note a reduced appetence for non-uridylated substrates. That was demonstrated by an initial velocity of degradation that is considerably lower in the case of Adh (grey dashed line in the plots in Figure 3.8.), when comparing these mutants with the wt SpDis3L2. A bigger distance between the line corresponding to the degradation of Adh substrate and the ones of Adh4U and Adh16U was also seen in the plot (wt, L460I and K808H in Figure 3.8.). These observations corroborate our previous results.
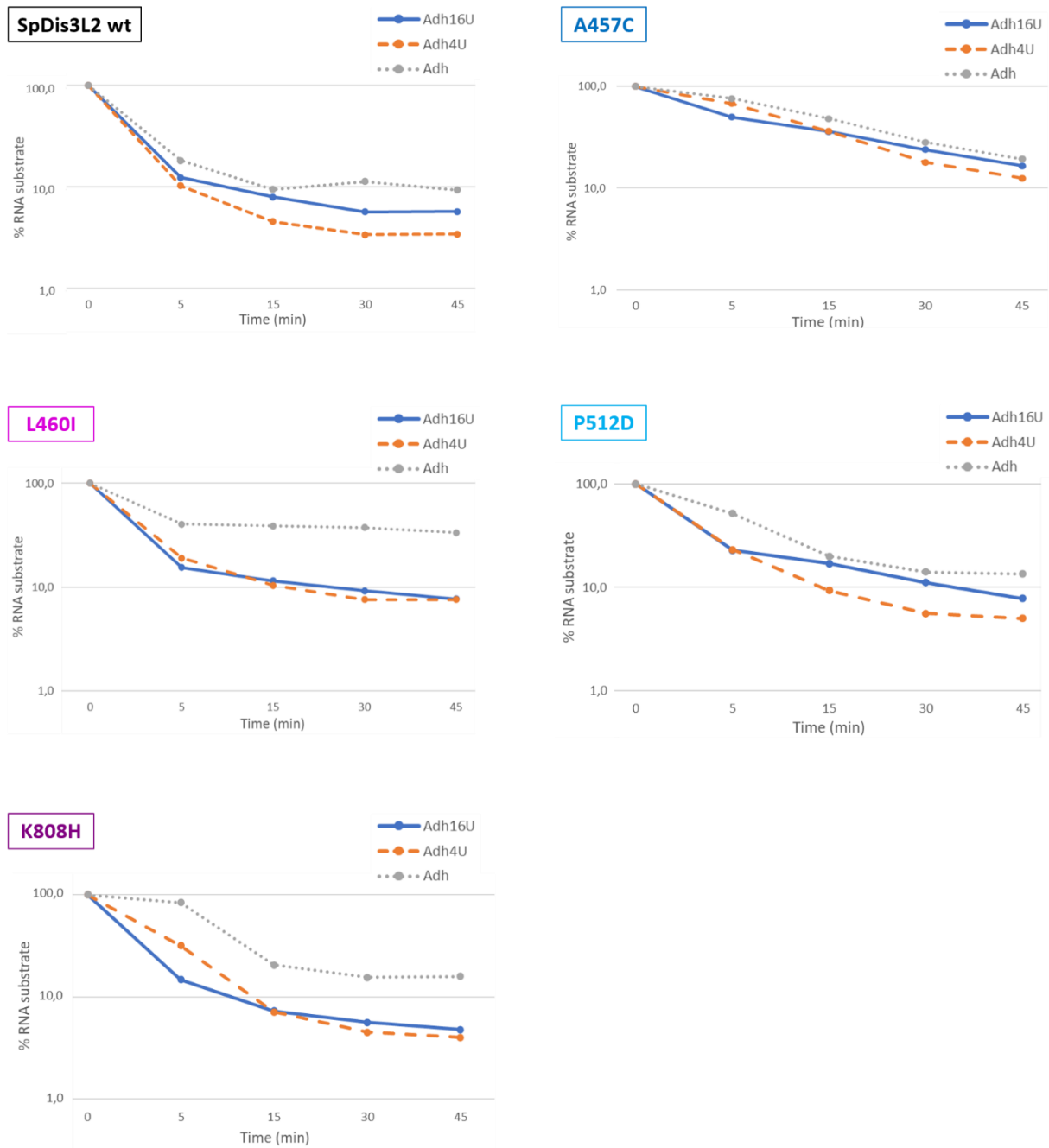
Figure 3.8. Quantification of the exoribonucleolytic activity of SpDis3L2 wild-type and mutants. Graphics represent the amount of remaining RNA at the time points indicated, for each of the three substrates (relative to the triple ssRNA substrate activity assays shown in Figure 3.7.). The amount of remaining RNA was quantified using ImageQuant and expressed as a percentage relative to the initial amount of substrate.

3.2.4. Activity assays with substrates forming double-stranded structures

In the cell, Dis3L2 targets a multiplicity of RNA classes, including highly structured ncRNAs and other aberrant transcripts that usually present a stable double-stranded structure near the oligo(U) tail (85, 86, 89, 171), as it is the case of let-7 precursors (pre-let-7) (172).

When comparing to its eukaryotic counterparts of the RNase II/RNB family, Dis3L2 is the exoribonuclease that has the best performance in hydrolysing RNAs with double-stranded regions, only requiring a protruding 3'-end with at least 2 nt (38). However, it has also been seen to succeed in degrading dsRNAs with blunt ends, although to a much smaller extent than if they had single-stranded tails (41).

To understand if the SpDis3L2 mutants conserved this degradation capacity, activity assays were carried out with two different structured substrates, and for all SpDis3L2 enzyme versions apart from D461N. The D461N mutant was not used in these cases since we had already confirmed in the previous assays that it was always inactive, behaving as expected, and there was no need to further test its absence of activity.

The first structured RNA substrate to be used was dsAdh, which has a perfect double-stranded structure. dsAdh was obtained through the hybridization of labelled Adh16U with unlabelled asAdh (that has a fully complementary sequence to Adh), having, therefore, a 3' overhang with 16 uridines (-16U) (Figure 3.9.A.).

To assure that, in the reaction mixtures, the dsAdh would be completely hybridized in the conditions used, we analysed this substrate in a 20% PAA non-denaturing gel, using the Adh16U molecule as a control (Figure 3.9.B.).
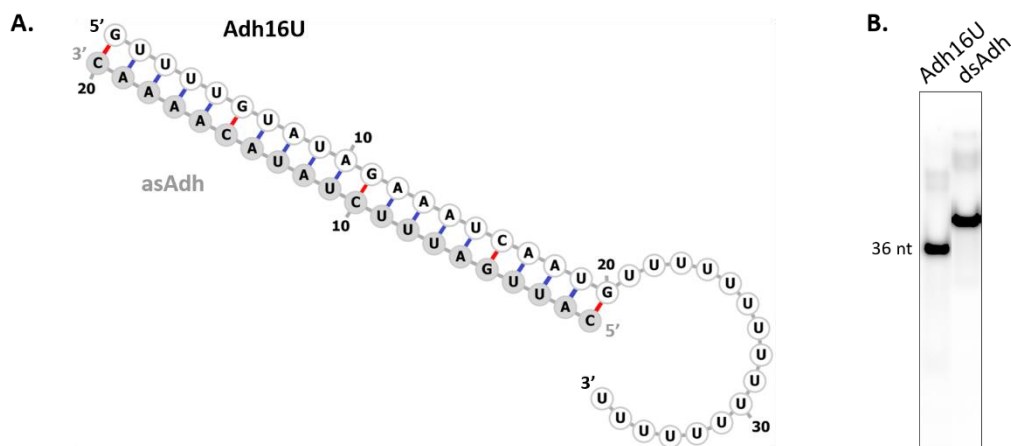
Figure 3.9. Formation of the dsAdh RNA substrate. (A.) Representation of the dsAdh substrate, originated by hybridization of Adh16U (labelled; represented in white) and asAdh (the complementary molecule to Adh, unlabelled; represented in grey). (**B.**) Non-denaturing 20% PAA gel of Adh16U (ssRNA) and dsAdh (dsRNA).

The results showed that in the dsAdh lane only one band was visible, and its migration was delayed compared to that of the control single-stranded Adh16U (Figure 3.9.B.). This confirmed that the RNA in that sample was in a double-stranded conformation and we could proceed with the assays (Figure 3.10.).
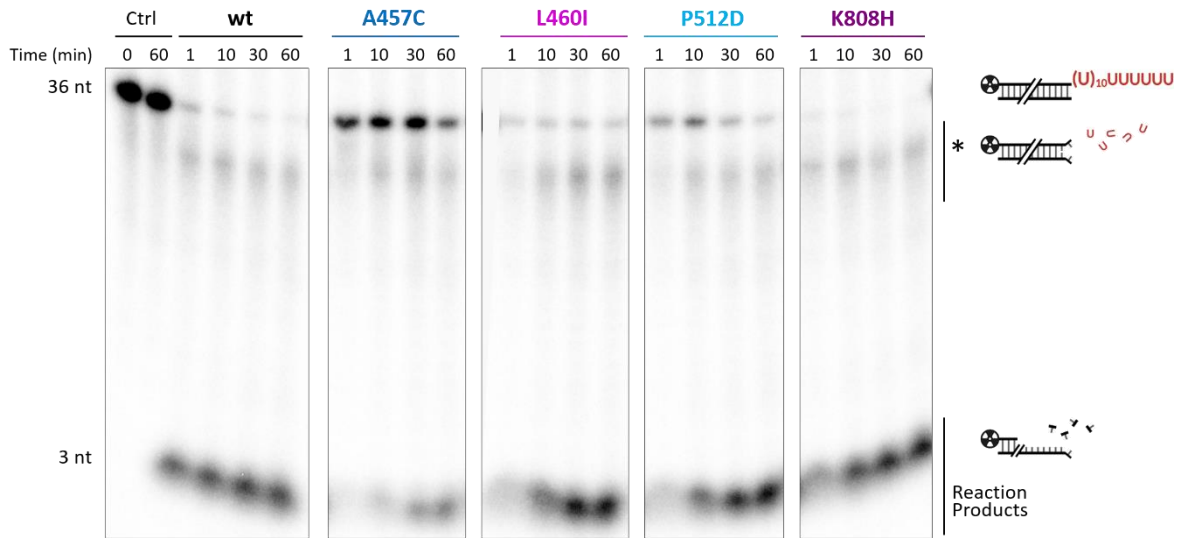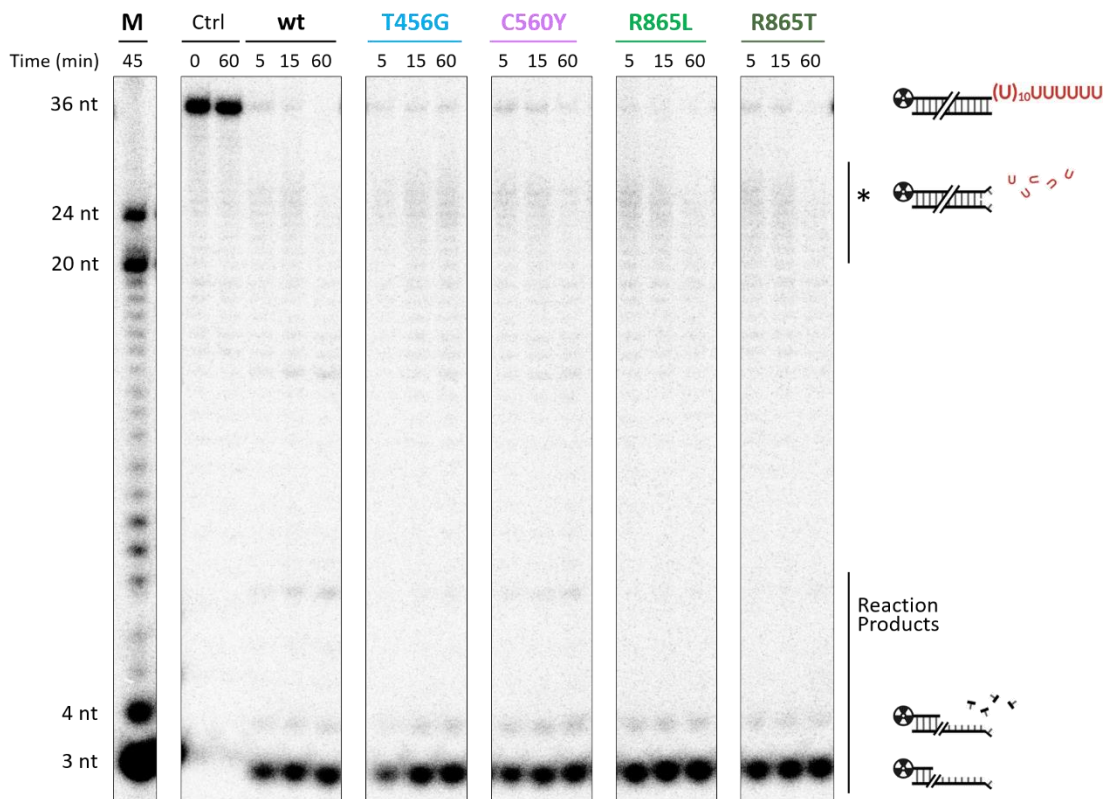
Figure 3.10. Exoribonucleolytic activity of SpDis3L2 wild-type and mutants on dsAdh RNA substrate. 8.3 nM of labelled Adh16U was hybridized with 83.3 nM of complementary unlabelled asAdh (to obtain the pretended 1:10 molar ratio between the two ssRNA substrates) to form dsAdh, which possesses a -16U 3' overhang. The dsAdh substrate was incubated with 100 nM of each SpDis3L2 variant (identified on the top) for 60 minutes at 30 °C. Control reactions (Ctrl) were incubated for the same time without any enzyme. Samples were taken and reaction was stopped at the time-points indicated above each gel lane. RNA substrates and degradation products were separated through migration in a 7 M urea/20% PAA denaturing gel of medium (**A.**) or high (**B.**) resolution. The size of the molecules is depicted in line with the corresponding bands alongside the gel. In the gel presented in **B.**, 'M' indicates the lane with a sample from an activity assay with the wt SpDis3L2 protein and a mix of two ssRNA (Adh and Adh4U), which was used as a size marker. ' * ' marks a range of intermediate products of dsRNA degradation. The assay was performed one time (for A457C, P512D, K808H, T456G, C560Y, R865L and R865T) or two times (for wt and L460I) with proteins from independent purifications.

Our results confirmed that SpDis3L2 was able to degrade structured substrates, releasing a final degradation product of 3 nt, as expected (Figure 3.10.). All the variants analysed continued to be able of degrading the dsAdh similarly to the wild-type. While the K808H mutant appeared to be slightly more active than the wt protein, the A457C mutant was considerably less active than the wt, which is consistent with the previous findings (Figure 3.7., section 3.2.3.) The P512D mutant also seemed less active than the wt, but in a minor extent.

It is also possible to observe the presence of intermediate products (denoted with ' * ' in Figure 3.10.). These intermediate products migrate close to the Adh4U (24 nt, as it can be seen in Figure 3.10.B.), which means that the protein slows down as the double-stranded zone approaches the entrance of the RNA pathway channel to the inside of the enzyme.

In this work we also used a second structured RNA substrate, hereafter designated as Loop, that mimics a secondary structure found in a particular let-7 precursor, pre-let-7g. The Loop ssRNA has a total of 34 nt, having the potential to form a double-stranded substrate with a stem-loop and a 3' overhang of 16 nt (-GC14U) (Figure 3.11.A.).

This RNA molecule was designed and is currently used in the laboratory headed by Leemor Joshua-Tor to study the mechanism of degradation of structured RNAs by the human DIS3L2. In a recent article of Meze *et al.* (164), using this substrate (referred there as hairpinA-GCU$_{14}$) and other shortened and adapted versions of this RNA, it was unveiled that the binding of dsRNA substrates to hDIS3L2 elicits a drastic change in the conformation of the enzyme, leading to the repositioning of CSD1 and CSD2.

Regarding the present work, to confirm the formation of the intended secondary structure, both the folded (after folding reaction) and unfolded Loop RNA substrates were run in a 20% PAA non-denaturing gel (Figure 3.11.).

Figure 3.11. Folding of Loop RNA substrate. (A.) Schematic representation of the Loop substrate. This model was adapted from the predictions obtained using the RNAfold Web Server (software RNAfold v2.4.18., available at http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi) and UNAFold Web Server (software mfold v3.6., available at http://www.unafold.org/mfold/applications/rna-folding-form-v2.php). (**B.**) Non-denaturing 20% PAA gel of both folded and unfolded forms of Loop substrate.

With this substrate, there was no visible delay in the gel migration of the RNA molecule after the folding. This was not surprising considering that it would be occurring an intramolecular folding, instead of a hybridization between two molecules as it happens for dsAdh, whose double-stranded form would have, consequently, a higher molecular weight.

Furthermore, to monitor if the folding treatment was being efficient, we carried out a control activity assay with EcRNase II. It is known that EcRNase II is sensible to double-stranded structures, stalling 5 to 7 nt before the double-stranded region (126).

As regards the control with EcRNase II (Figure S.3.11.), a fraction of the degradation products released were around 24 nt. This is likely to correspond to the 5-7 nt range before the double strand start, where RNase II would predictably be blocked. However, we can also see the formation of smaller degradation products, which were not predicted to appear if the RNA was completely folded. Therefore, we can say that the folding of this RNA molecule was partially efficient (Figure S.3.11. and Figure 3.12.).

Figure 3.12. Exoribonucleolytic activity of SpDis3L2 wild-type and mutants on dsLoop RNA substrate. 8.3 nM of Loop ssRNA was submitted to a folding treatment to form dsLoop, which possesses a stem-loop and a -GC14U 3' overhang. The dsLoop substrate was incubated with 100 nM of each SpDis3L2 variant (named on the top) for 60 minutes at 30 ℃. Control reactions (Ctrl) were incubated for the same time without any enzyme. Samples were taken and reaction was stopped at the time-points indicated above each gel lane. RNA substrates and degradation products were separated through migration in a 7 M urea/20% PAA denaturing gel of medium (**A.**) or high (**B.**) resolution. The size of the molecules is depicted in line with the corresponding bands alongside the gel. In the gel presented in **B.**, 'M' indicates the lane with a sample from an activity assay with the wt SpDis3L2 protein and a mix of two ssRNA (Adh and Adh4U), which was used as a size marker. ' * ' marks a range of intermediate products of dsRNA degradation. The assay was performed one time (for A457C, P512D, K808H, T456G, C560Y, R865L and R865T) or two times (for wt and L460I) with proteins from independent purifications.

In terms of the different versions of SpDis3L2, they all appeared to be active and degrade the substrate to the expected size of 3 nt (Figure 3.12.). The K808H mutant was shown to be the most active, as observed in assays with other substrates, while the A457C mutant remained remarkably the least active. In addition, the P512D mutant was slightly less active, as well as C560Y and R865T which curiously were found to digest dsAdh well, although they seemed somewhat less active on dsLoop than the wt protein. Consistent with the dsAdh assays, there was a modest accumulation of intermediate products (denoted with ' * ' in Figure 3.12.) around 20 nt of size. These products would either arise because of the enzyme slowing down as it moves closer to the double-stranded region, or they may reflect the increasing unwinding difficulty of the double strand itself, considering that this substrate has 4 consecutive GC pairs (Figure 3.11.). GC pairs are stronger interactions, since the two nucleotides establish between them an additional hydrogen bond in comparison to AU or other non-canonical pairs such as GU.

3.2.5. General conclusions

All the enzyme versions of SpDis3L2 that we were able to successfully purify (wt, T456G, A457C, L460I, D461N, P512D, C560Y, K808H, R865L and R865T), apart from the inactive D461N mutant that served as a control, were active, and generated 3 nt end-products. However, when poly(U) was used as a substrate, it became quite evident that an accumulation of larger size products occurred. Although the reason for this phenomenon is still concealed, the most logical explanation relies on the distributivity of the enzyme towards small fragments.

When evaluating the preference for the uridylated RNA molecules, most of them demonstrated the same preference when compared to the wild-type. It is interesting to note that the apparent absence of a changing preference according with the size of the U-tails actually makes sense when thinking about the diversity of possible biological RNA substrates of the enzyme within the eukaryotic cell. In physiological conditions, the *adh1* mRNAs, that were found in *S. pombe* to be uridylated, carried up to 7 uridines (38). If other substrates are taken into account, the tails are often longer. Other studies have seen that, for example *GADD45A* transcripts with premature STOP codons (or PTCs), targeted for degradation by Dis3L2, might possess untemplated tails with up to 15 uridines (44), while, for the more scrutinized case of pre-let-7, it was found that TUT4 is able to add 10 to 30 non-templated uridines to this miRNA precursor (72).

The ability of Dis3L2 to degrade structured RNAs is essential for the performance of its cellular functions (85, 86, 89, 171). Our results show that all the proteins were able to degrade structured RNA molecules. However, they slow down when approaching the double-stranded region, having some difficulties in overcoming its beginning, but then managing to prosecute with the decay of these substrates.

Therefore, it became clear that a single amino acid substitution may produce a significant difference. Even if they are in the same position, different amino acid modifications might impact differently the stability and the overall exoribonucleolytic activity of a protein. That was observed for L460I/L460V and R865L/R865T mutants.

On the one hand, for the L460I and L460V mutants, while we managed to purify the former to a good amount and degree of purity, the second could not be purified. Since valine (V) has a shorter main chain than isoleucine (I), it might be the case that the leucine (L) is responsible for establishing an interaction with another amino acid. In this situation, I would be able to complement this interaction, while V becomes insufficient and destabilises the protein. Regarding the L460I mutant, we observed that it had a higher specificity towards uridylated RNA substrates. We might hypothesise that the amino acid change is altering the orchestration of the remaining amino acids in the active centre, or even elsewhere in the protein, in a way that, against our expectations, the overall network of interactions along the RNA pathway within the enzyme would favour binding to uridines, thereby increasing the enzyme's preference for uridylated substrates. Further work is still necessary to completely understand this phenomenon.

On the other hand, for the R865L and R865T mutants, while the R865L substitution did not produce significant changes in activity, the R865T mutant was less active. This last result could be expected since the residue R865 in *S. pombe* and the corresponding R792 in *M. musculus* are located in S1 domain and are predicted to integrate a β-strand. Also, it has been described that, in mouse, this residue associates with the backbone of a the 14U RNA (95, 138). Therefore, we can postulate that SpDis3L2 R865 is likely to be important in binding the RNA substrates. Thus, modifying the amino acid in this position could affect the protein structure near the top of the RNA entry channel to the enzyme. To further investigate this possibility, we would have to perform binding assays (electrophoretic mobility shift assays, EMSAs), which was not possible due to lack of time derived from the constraints caused by the current COVID-19 pandemic.

Moreover, other mutants also revealed a certain degree of activity impairment, namely A457C, P512D and C560Y. In particular for the C560Y mutant, it appeared to prefer the non-uridylated RNA, albeit not in a significant way. By being a modification found in patients with PRLMNS, we can speculate that, in the context of disease, this preference not so marked, or even inverted, could translate in pathology. It could be related with a lower degradation efficiency of Dis3L2 over its targets, which would then accumulate, or a change of specificity towards other RNAs that would not be supposed to occur.

Finally, in what concerns all mutated versions of SpDis3L2 that we managed to purify, it is noteworthy that some of them were not in sufficient concentration to be further used in the activity assays or, otherwise, the results of the activity assays performed with those mutants were inconclusive,

revealing non-specific degradation. It was the case of A208E, F209A, L371R, A801C and C804I mutants. Possible explanations rely on protein precipitation, or high amount of small, degraded proteins after purification.

4. Final Remarks and Future Perspectives

It has become clear how RNases are extremely relevant in cellular metabolism, participating in the processing, degradation, and surveillance of all types of RNA. Likewise, it has also been shown how 3' uridylation is a modification widely encountered in eukaryotes, with enzymes dedicated to the addition of these U-tails to RNAs (the TUTases or PUPs), and an enzyme that preferentially binds and degrades them (the Dis3L2). Even though there are several TUTases/PUPs that can play redundant roles, it seems that Dis3L2 represents a bottleneck in this alternative RNA decay pathway (131). Moreover, hDIS3L2 has already been associated with several overgrowth-related human diseases, such as Perlman Syndrome, Wilms' Tumour and Marfan-like Syndrome (143), reinforcing the importance of the study of this enzyme.

The contribution of the RNase II/RNB family member RNase R to bacterial virulence has been well documented (122, 173–175), and given that ribonucleases of families other than RNase II/RNB have been associated with the virulence of some pathogenic fungi (176–178), one could wonder whether there is a possible implication of eukaryotic enzymes of the RNase II/RNB family, in particular Dis3L2, in the virulence of some parasites or fungi. The protein BLAST (BLASTp) of NCBI database was used

to search for homologues of Dis3L2, providing, as query, the amino acid sequence of the full-length Dis3L2 protein from *S. pombe* (SpDis3L2). From the resulting sequences producing significant alignments, the only proteins characterised unequivocally as Dis3L2 isoforms appeared only for *X. laevis*, *M. musculus* and numerous species of fish, whereas around half of the entries were 'hypothetical' or 'RNB-domain-containing' proteins assigned to fungi. Among those, some species are considered pathogenic for humans (179) and have already been detected as causing infection: *Basidiobolus meristosporus* (180), *Rhizopus microsporus* (including *R. microsporus* (181, 182), and *R. microsporus* var. *azygosporus* (183, 184)) and *Rhizopus arrhizus* (including *R. arrhizus* var. *delemar*, and *R. arrhizus* var. *arrhizus*, being the last one previously classified as *Rhizopus oryzae*) (185–187). Curiously, when searching for the amino acid signature of the catalytic centre of Dis3L2 proteins (DLDD, further explained in the next section, 1.4.2.2.) in those putative proteins from pathogenic fungi, it was found in all of them. These observations evidence that, presumably, these could be Dis3L2-like enzymes that exist in those organisms but have not yet been studied and, therefore, are not yet annotated in their genomes, rather than being absent as it might have been rashly thought at first. This is unsurprising given the enormous importance of uridylation/Dis3L2-mediated RNA decay and the high conservation of the enzyme in eukaryotes. Thus, there could be a link between the Dis3L2 enzyme and pathogenicity that has not yet started to be explored.

Furthermore, it was recently unravelled that *S. cerevisiae* Ssd1 homologues, a type of RNase II pseudonucleases that lost their catalytic activity during evolution, are descendants of active Dis3L2 proteins (188, 189). Besides retaining RNA-binding capacity, Ssd1 homologues were found to be relevant for the virulence of diverse fungal pathogens, particularly *Candida albicans* (190), *Aspergillus fumigatus* (191), and *Cryptococcus neoformans* (188).

In this work we used *S. pombe* as a microorganism model for the study of Dis3L2, a peculiar enzyme from the RNase II/RNB family that prefers uridylated RNA molecules. We managed to express SpDis3L2 mutant variants with single amino acid substitutions, we optimized their purification conditions, purified several of them to a good degree of purity, and evened out their concentrations by a western blot procedure. However, we were unable to purify some of the SpDis3L2 versions. Expression might be part of the problem since we are trying to overexpress an eukaryotic relatively long protein in *E. coli* bacteria. It should be investigated ways to overcome expression and purification issues, namely to mitigate protein degradation during the purification protocol. A possibility would be to use a distinct method in one of the proteins' purification steps, namely Ion Exchange or Q-Resource.

We tested the exoribonucleolytic activity of the SpDis3L2 variants over different single- and double-stranded RNA substrates (section 3.2.), observing that the D461N mutant was inactive as expected, and the remaining tested protein variants were active.

Different substitutions of single amino acids of SpDis3L2 had different effects on the protein purification yield, and then on its stability, activity, and preference for uridylated RNAs. That may come as a consequence of possible alterations in the structure of the protein, which might compromise some secondary structure (such as α-helixes or β-strands) or influence the overall folding of the enzyme, revealing the potential role of the original residues in those positions in the enzyme's activity and substrate specificity. It is noteworthy that the L460I and K808H mutants displayed higher specificity towards uridylated substrates, whereas this preference seemed to be abrogated to some extent in the C560Y mutant. This is even more curious since SpDis3L2 L460 is near the active site of the enzyme, the residue corresponding to SpDis3L2 K808 in mDis3L2 was previously reported to interact with an U14 substrate (95), and the substitution C560Y was found in the corresponding position in hDIS3L2 in patients with Perlman Syndrome (100).

Another striking example that reveals the preponderant role that point mutations can have on the activity of this enzyme, potentially making it more or less active, is that of residue R74 in mDis3L2, which corresponds to R200 in SpDis3L2. By changing this amino acid to an A in mDis3L2 (mutant R74A) it leads to a dramatic decrease in the *in vitro* specific activity of the enzyme on a 15U substrate (95). Conversely, this same modification in SpDis3L2 (mutant R200A) was previously studied by collaborators of the CMA lab and consistently resulted in increased activity (data not published). Additionally, it was seen that a significant part of the residues that were identified as being involved in interactions with the 14U substrate in the mDis3L2-U14 structure, are not conserved in all orthologues of Dis3L2, namely in fission yeast. This may highlight that the residues implicated in binding to RNA substrates in SpDis3L2 may not be the same as in mDis3L2, pointing to the fact that possibly the RNA recognition and binding mechanism of SpDis3L2 is different from its mouse counterpart. These findings underscore the importance of the current work.

Given the complex network of interactions that is established between the enzyme and its substrates, it is not unexpected that these point mutations have only a limited impact on the enzyme, and therefore they may not be sufficient to cause a significant change in activity or specificity. In mDis3L2 it has already been shown that two mutations that affected the enzyme's activity in the same sense when separately (both decreasing the activity of mDis3L2 on a 15U RNA), if they were present simultaneously there was a cumulative effect (95). Accordingly, the results we have now achieved for SpDis3L2 create the conditions for not only systematically evaluate a variety of new mutants with single amino acid substitutions, but also combinations of those substitutions in order to reach more interesting/outstanding conclusions. Thus, this work unveils a path of opportunities to go further into the world of SpDis3L2 amino acids contribution to the enzyme overall function.

Furthermore, it would be interesting to explore if differences in the mutated versions of SpDis3L2 activity over certain RNA substrates rely on eventual modifications in RNA-binding affinity. Hence,

EMSAs could be conducted with the wild-type protein and those enzyme variants (namely C560Y, as it was mentioned above, in section 3.2.5.) and the RNA substrates in question.

To conclude, it would be surely fascinating to "look closer into the cell" and analyse the cellular localization of mutant SpDis3L2 proteins by employing fluorescence microscopy.

5. Bibliography

1. Merrick WC, Pavitt GD. 2018. Protein synthesis initiation in eukaryotic cells. Cold Spring Harb Perspect Biol 10:a033092 doi:10.1101/cshperspect.a033092.
2. Kaikkonen MU, Lam MTY, Glass CK. 2011. Non-coding RNAs as regulators of gene expression and epigenetics. Cardiovasc Res 90:430–440 doi:10.1093/cvr/cvr097.
3. Arraiano CM, Andrade JM, Domingues S, Guinote IB, Malecki M, Matos RG, Moreira RN, Pobre V, Reis FP, Saramago M, Silva IJ, Viegas SC. 2010. The critical role of RNA processing and degradation in the control of gene expression. FEMS Microbiol Rev 34:883–923 doi:10.1111/j.1574-6976.2010.00242.x.
4. Zhao W, Manley JL. 1996. Complex alternative RNA processing generates an unexpected diversity of poly(A) polymerase isoforms. Mol Cell Biol 16:2378–2386 doi:10.1128/mcb.16.5.2378.
5. Kühn U, Wahle E. 2004. Structure and function of poly(A) binding proteins. Biochim Biophys Acta - Gene Struct Expr 1678:67–84 doi:10.1016/j.bbaexp.2004.03.008.
6. Schmidt MJ, Norbury CJ. 2010. Polyadenylation and beyond: Emerging roles for noncanonical poly(A) polymerases. Wiley Interdiscip Rev RNA 1:142–151 doi:10.1002/wrna.16.
7. Jia L, Mao Y, Ji Q, Dersh D, Yewdell JW, Qian SB. 2020. Decoding mRNA translatability and stability from the 5′ UTR. Nat Struct Mol Biol 27:814–821 doi:10.1038/s41594-020-0465-x.
8. Barreau C, Paillard L, Osborne HB. 2005. AU-rich elements and associated factors: Are there unifying principles? Nucleic Acids Res 33:7138–7150 doi:10.1093/nar/gki1012.
9. Murray EL, Schoenberg DR. 2007. A+U-Rich Instability Elements Differentially Activate 5′-3′ and 3′-5′ mRNA Decay. Mol Cell Biol 27:2791–2799 doi:10.1128/MCB.01445-06.
10. Huntzinger E, Izaurralde E. 2011. Gene silencing by microRNAs: Contributions of translational repression and mRNA decay. Nat Rev Genet 12:99–110 doi:10.1038/nrg2936.
11. Eberle AB, Lykke-Andersen S, Mühlemann O, Jensen TH. 2009. SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. Nat Struct Mol Biol 16:49–55 doi:10.1038/nsmb.1530.
12. Huntzinger E, Kashima I, Fauser M, Saulière J, Izaurralde E. 2008. SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. RNA 14:2609–2617 doi:10.1261/rna.1386208.
13. Łabno A, Tomecki R, Dziembowski A. 2016. Cytoplasmic RNA decay pathways - Enzymes and mechanisms. Biochim Biophys Acta - Mol Cell Res 1863:3125–3147 doi:10.1016/j.bbamcr.2016.09.023.
14. Hojka-Osinska A, Chlebowski A, Grochowska J, Owczarek EP, Affek K, Kłosowska-Kosicka K, Szczesny RJ, Dziembowski A. 2021. Landscape of functional interactions of human processive ribonucleases revealed by high-throughput siRNA screenings. iScience 24:103036 doi:10.1016/j.isci.2021.103036.
15. Towler BP, Newbury SF. 2018. Regulation of cytoplasmic RNA stability: Lessons from *Drosophila*. Wiley Interdiscip Rev RNA 9:1–20 doi:10.1002/wrna.1499.
16. Yamashita A, Chang TC, Yamashita Y, Zhu W, Zhong Z, Chen CYA, Shyu AB. 2005. Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. Nat Struct Mol Biol 12:1054–1063 doi:10.1038/nsmb1016.
17. Chen CYA, Shyu AB. 2011. Mechanisms of deadenylation-dependent decay. Wiley Interdiscip Rev RNA 2:167–183 doi:10.1002/wrna.40.
18. Tharun S. 2009. Lsm1-7-Pat1 complex: A link between 3′ and 5′-ends in mRNA decay? RNA Biol 6:228–232 doi:10.4161/rna.6.3.8282.
19. Sharif H, Conti E. 2013. Architecture of the Lsm1-7-Pat1 Complex: A Conserved Assembly in Eukaryotic mRNA Turnover. Cell Rep 5:283–291 doi:10.1016/j.celrep.2013.10.004.
20. Li Y, Kiledjian M. 2010. Regulation of mRNA decapping. Wiley Interdiscip Rev RNA 1:253–265 doi:10.1002/wrna.15.
21. Stevens A. 2001. 5′-exoribonuclease 1: Xrn 1. Methods in Enzymology doi:10.1016/S0076-6879(01)42549-3.

22. Braun JE, Truffault V, Boland A, Huntzinger E, Chang CT, Haas G, Weichenrieder O, Coles M, Izaurralde E. 2012. A direct interaction between DCP1 and XRN1 couples mRNA decapping to 5′ exonucleolytic degradation. Nat Struct Mol Biol 19:1324–1331 doi:10.1038/nsmb.2413.

23. Langeberg CJ, Welch WRW, McGuire JV, Ashby A, Jackson AD, Chapman EG. 2020. Biochemical Characterization of Yeast Xrn1. Biochemistry 59:1493–1507 doi:10.1021/acs.biochem.9b01035.

24. Zuo Y, Deutscher MP. 2001. Exoribonuclease superfamilies: Structural analysis and phylogenetic distribution. Nucleic Acids Res 29:1017–1026 doi:10.1093/nar/29.5.1017.

25. Liu Q, Greimann JC, Lima CD. 2006. Reconstitution, Activities, and Structure of the Eukaryotic RNA Exosome. Cell 127:1223–1237 doi:10.1016/j.cell.2006.10.037.

26. Bonneau F, Basquin J, Ebert J, Lorentzen E, Conti E. 2009. The Yeast Exosome Functions as a Macromolecular Cage to Channel RNA Substrates for Degradation. Cell 139:547–559 doi:10.1016/j.cell.2009.08.042.

27. Viegas SC, Matos RG, Arraiano CM. 2020. The Bacterial Counterparts of the Eukaryotic Exosome: An Evolutionary Perspective, p. 37–46. *In* LaCava, J, Vaňáčová, Š (eds.), The Eukaryotic RNA Exosome: Methods and Protocols. Humana, New York, USA doi:10.1007/978-1-4939-9822-7_2 ISBN:9781493998210.

28. Sikorska N, Zuber H, Gobert A, Lange H, Gagliardi D. 2017. RNA degradation by the plant RNA exosome involves both phosphorolytic and hydrolytic activities. Nat Commun 8 doi:10.1038/s41467-017-02066-2.

29. Dziembowski A, Lorentzen E, Conti E, Séraphin B. 2007. A single subunit, Dis3, is essentially responsible for yeast exosome core activity. Nat Struct Mol Biol 14:15–22 doi:10.1038/nsmb1184.

30. Lejeune F, Li X, Maquat LE. 2003. Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities. Mol Cell 12:675–687 doi:10.1016/S1097-2765(03)00349-6.

31. Tomecki R, Kristiansen MS, Lykke-Andersen S, Chlebowski A, Larsen KM, Szczesny RJ, Drazkowska K, Pastula A, Andersen JS, Stepien PP, Dziembowski A, Jensen TH. 2010. The human core exosome interacts with differentially localized processive RNases: hDIS3 and hDIS3L. EMBO J 29:2342–2357 doi:10.1038/emboj.2010.121.

32. Allmang C, Petfalski E, Podtelejnikov A, Mann M, Tollervey D, Mitchell P. 1999. The yeast exosome and human PM-Scl are related complexes of 3′→5′ exonucleases. Genes Dev 13:2148–2158 doi:10.1101/gad.13.16.2148.

33. Staals RHJ, Bronkhorst AW, Schilders G, Slomovic S, Schuster G, Heck AJR, Raijmakers R, Pruijn GJM. 2010. Dis3-like 1: A novel exoribonuclease associated with the human exosome. EMBO J 29:2358–2367 doi:10.1038/emboj.2010.122.

34. Makino DL, Baumgärtner M, Conti E. 2013. Crystal structure of an RNA-bound 11-subunit eukaryotic exosome complex. Nature 495:70–75 doi:10.1038/nature11870.

35. Liu H, Rodgers ND, Jiao X, Kiledjian M. 2002. The scavenger mRNA decapping enzyme DcpS is a member of the HIT family of pyrophosphatases. EMBO J 21:4699–4708 doi:10.1093/emboj/cdf448.

36. Chen N, Walsh MA, Liu Y, Parker R, Song H. 2005. Crystal structures of human DcpS in ligand-free and m^{7}GDP-bound forms suggest a dynamic mechanism for scavenger mRNA decapping. J Mol Biol 347:707–718 doi:10.1016/j.jmb.2005.01.062.

37. Milac AL, Bojarska E, del Nogal AW. 2014. Decapping Scavenger (DcpS) enzyme: Advances in its structure, activity and roles in the cap-dependent mRNA metabolism. Biochim Biophys Acta - Gene Regul Mech 1839:452–462 doi:10.1016/j.bbagrm.2014.04.007.

38. Malecki M, Viegas SC, Carneiro T, Golik P, Dressaire C, Ferreira MG, Arraiano CM. 2013. The exoribonuclease Dis3L2 defines a novel eukaryotic RNA degradation pathway. EMBO J 32:1842–1854 doi:10.1038/emboj.2013.63.

39. Chang HM, Triboulet R, Thornton JE, Gregory RI. 2013. A role for the Perlman syndrome exonuclease Dis3l2 in the Lin28-let-7 pathway. Nature 497:244–248 doi:10.1038/nature12119.

40. Ustianenko D, Hrossova D, Potesil D, Chalupnikova K, Hrazdilova K, Pachernik J, Cetkovska K, Uldrijan S, Zdrahal Z, Vanacova A. 2013. Mammalian DIS3L2 exoribonuclease targets the uridylated precursors of let-7 miRNAs. RNA 19:1632–1638 doi:10.1261/rna.040055.113.

41. Lubas M, Damgaard CK, Tomecki R, Cysewski D, Jensen TH, Dziembowski A. 2013. Exonuclease hDIS3L2 specifies an exosome-independent 3′-5′ degradation pathway of human cytoplasmic mRNA. EMBO J 32:1855–1868 doi:10.1038/emboj.2013.135.

42. Chang H, Lim J, Ha M, Kim VN. 2014. TAIL-seq: Genome-wide determination of poly(A) tail length and 3′ end modifications. Mol Cell 53:1044–1052 doi:10.1016/j.molcel.2014.02.007.

43. Green MR, Sambrook J. 2019. Rapid amplification of sequences from the 3′ ends of mRNAs: 3′-RACE. Cold Spring Harb Protoc 2019:426–430 doi:10.1101/pdb.prot095216.

44. da Costa PJ, Menezes J, Saramago M, García-Moreno JF, Santos HA, Gama-Carvalho M, Arraiano CM, Viegas SC, Romão L. 2019. A role for DIS3L2 over natural nonsense-mediated mRNA decay targets in human cells. Biochem Biophys Res Commun 518:664–671 doi:10.1016/j.bbrc.2019.08.105.

45. Chang H, Yeo J, Kim JG, Kim H, Lim J, Lee M, Kim HH, Ohk J, Jeon HY, Lee H, Jung H, Kim KW, Kim VN. 2018. Terminal Uridylyltransferases Execute Programmed Clearance of Maternal Transcriptome in Vertebrate Embryos. Mol Cell 70:72-82.e7 doi:10.1016/j.molcel.2018.03.004.

46. Aravind L, Koonin EV. 1999. DNA polymerase β-like nucleotidyltransferase superfamily: Identification of three new families, classification and evolutionary history. Nucleic Acids Res 27:1609–1618 doi:10.1093/nar/27.7.1609.

47. Martin G, Keller W. 2007. RNA-specific ribonucleotidyl transferases. RNA 13:1834–1849 doi:10.1261/rna.652807.

48. Wyers F, Rougemaille M, Badis G, Rousselle JC, Dufour ME, Boulay J, Régnault B, Devaux F, Namane A, Séraphin B, Libri D, Jacquier A. 2005. Cryptic Pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. Cell 121:725–737 doi:10.1016/j.cell.2005.04.030.

49. LaCava J, Houseley J, Saveanu C, Petfalski E, Thompson E, Jacquier A, Tollervey D. 2005. RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. Cell 121:713–724 doi:10.1016/j.cell.2005.04.029.

50. Wang L, Eckmann CR, Kadyk LC, Wickens M, Kimble J. 2002. A regulatory cytoplasmic poly(A) polymerase in *Caenorhabditis elegans*. Nature 419:312–316 doi:10.1038/nature01039.

51. Barnard DC, Ryan K, Manley JL, Richter JD. 2004. Symplekin and xGLD-2 are required for CPEB-mediated cytoplasmic polyadenylation. Cell 119:641–651 doi:10.1016/j.cell.2004.10.029.

52. Kwak JE, Drier E, Barbee SA, Ramaswami M, Yin JC, Wickens M. 2008. GLD2 poly(A) polymerase is required for long-term memory. Proc Natl Acad Sci USA 105:14644–14649 doi:10.1073/pnas.0803185105.

53. Tomecki R, Dmochowska A, Gewartowski K, Dziembowski A, Stepien PP. 2004. Identification of a novel human nuclear-encoded mitochondrial poly(A) polymerase. Nucleic Acids Res 32:6001–6014 doi:10.1093/nar/gkh923.

54. Rissland OS, Mikulasova A, Norbury CJ. 2007. Efficient RNA Polyuridylation by Noncanonical Poly(A) Polymerases. Mol Cell Biol 27:3612–3624 doi:10.1128/mcb.02209-06.

55. Mullen TE, Marzluff WF. 2008. Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5′ to 3′ and 3′ to 5′. Genes Dev 22:50–65 doi:10.1101/gad.1622708.

56. Kwak JE, Wickens M. 2007. A family of poly(U) polymerases. RNA 13:860–867 doi:10.1261/rna.514007.

57. Yashiro Y, Tomita K. 2018. Function and regulation of human terminal uridylyltransferases. Front Genet 9:1–14 doi:10.3389/fgene.2018.00538.

58. Aphasizhev R. 2005. RNA uridylyltransferases. Cell Mol Life Sci 62:2194–2203 doi:10.1007/s00018-005-5198-9.

59. Scott DD, Norbury CJ. 2013. RNA decay via 3′ uridylation. Biochim Biophys Acta - Gene Regul Mech 1829:654–665 doi:10.1016/j.bbagrm.2013.01.009.

60. Viegas SC, Silva IJ, Apura P, Matos RG, Arraiano CM. 2015. Surprises in the 3′-end: "U" can decide too! FEBS J 282:3489–3499 doi:10.1111/febs.13377.

61. Scheer H, Zuber H, De Almeida C, Gagliardi D. 2016. Uridylation Earmarks mRNAs for Degradation… and More. Trends Genet 32:607–619 doi:10.1016/j.tig.2016.08.003.

62. De Almeida C, Scheer H, Zuber H, Gagliardi D. 2018. RNA uridylation: a key posttranscriptional modification shaping the coding and noncoding transcriptome. Wiley Interdiscip Rev RNA 9 doi:10.1002/wrna.1440.

63. Trippe R, Richly H, Benecke BJ. 2003. Biochemical characterization of a U6 small nuclear RNA-specific terminal uridylyltransferase. Eur J Biochem 270:971–980 doi:10.1046/j.1432-1033.2003.03466.x.

64. Nagaike T, Suzuki T, Katoh T, Ueda T. 2005. Human mitochondrial mRNAs are stabilized with polyadenylation regulated by mitochondria-specific poly(A) polymerase and polynucleotide phosphorylase. J Biol Chem 280:19721–19727 doi:10.1074/jbc.M500804200.

65. Aphasizhev R, Sbicego S, Peris M, Jang SH, Aphasizheva I, Simpson AM, Rivlin A, Simpson L. 2002. Trypanosome mitochondrial 3′ terminal uridylyl transferase (TUTase): The key enzyme in U-insertion/deletion RNA editing. Cell 108:637–648 doi:10.1016/S0092-8674(02)00647-5.

66. Aphasizhev R, Aphasizheva I, Simpson L. 2003. A tale of two TUTases. Proc Natl Acad Sci USA 100:10617–10622 doi:10.1073/pnas.1833120100.

67. Aphasizheva I, Maslov D, Wang X, Huang L, Aphasizhev R. 2011. Pentatricopeptide Repeat Proteins Stimulate mRNA Adenylation/Uridylation to Activate Mitochondrial Translation in Trypanosomes. Mol Cell 42:106–117 doi:10.1016/j.molcel.2011.02.021.

68. Rissland OS, Norbury CJ. 2009. Decapping is preceded by 3′ uridylation in a novel pathway of bulk mRNA turnover. Nat Struct Mol Biol 16:616–623 doi:10.1038/nsmb.1601.

69. Lim J, Ha M, Chang H, Kwon SC, Simanshu DK, Patel DJ, Kim VN. 2014. Uridylation by TUT4 and TUT7 marks mRNA for degradation. Cell 159:1365–1376 doi:10.1016/j.cell.2014.10.055.

70. Zigáčková D, Vaňáčová Š. 2018. The role of 3′ end uridylation in RNA metabolism and cellular physiology. Philos Trans R Soc B Biol Sci 373 doi:10.1098/rstb.2018.0171.

71. Heo I, Joo C, Cho J, Ha M, Han J, Kim VN. 2008. Lin28 Mediates the Terminal Uridylation of let-7 Precursor MicroRNA. Mol Cell 32:276–284 doi:10.1016/j.molcel.2008.09.014.

72. Heo I, Joo C, Kim YK, Ha M, Yoon MJ, Cho J, Yeom KH, Han J, Kim VN. 2009. TUT4 in Concert with Lin28 Suppresses MicroRNA Biogenesis through Pre-MicroRNA Uridylation. Cell 138:696–708 doi:10.1016/j.cell.2009.08.002.

73. Lehrbach NJ, Armisen J, Lightfoot HL, Murfitt KJ, Bugaut A, Balasubramanian S, Miska EA. 2009. LIN-28 and the poly(U) polymerase PUP-2 regulate let-7 microRNA processing in *Caenorhabditis elegans*. Nat Struct Mol Biol 16:1016–1020 doi:10.1038/nsmb.1675.

74. Thornton JE, Du P, Jing L, Sjekloca L, Lin S, Grossi E, Sliz P, Zon LI, Gregory RI. 2014. Selective microRNA uridylation by Zcchc6 (TUT7) and Zcchc11 (TUT4). Nucleic Acids Res 42:11777–11791 doi:10.1093/nar/gku805.

75. Hagan JP, Piskounova E, Gregory RI. 2009. Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. Nat Struct Mol Biol 16:1021–1025 doi:10.1038/nsmb.1676.

76. De Almeida C, Scheer H, Gobert A, Fileccia V, Martinelli F, Zuber H, Gagliardi D. 2018. RNA uridylation and decay in plants. Philos Trans R Soc B Biol Sci 373 doi:10.1098/rstb.2018.0163.

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

78. Menezes MR, Balzeau J, Hagan JP. 2018. 3′ RNA uridylation in epitranscriptomics, gene regulation, and disease. Front Mol Biosci 5:1–20 doi:10.3389/fmolb.2018.00061.

79. Trippe R, Guschina E, Hossbach M, Urlaub H, Lührmann R, Benecke BJ. 2006. Identification, cloning, and functional analysis of the human U6 snRNA-specific terminal uridylyl transferase. RNA 12:1494–1504 doi:10.1261/rna.87706.

80. Koppers-Lalic D, Hackenberg M, Bijnsdorp IV, van Eijndhoven MAJ, Sadek P, Sie D, Zini N, Middeldorp JM, Ylstra B, de Menezes RX, Würdinger T, Meijer GA, Pegtel DM. 2014. Nontemplated nucleotide additions distinguish the small RNA composition in cells from exosomes. Cell Rep 8:1649–1658 doi:10.1016/j.celrep.2014.08.027.

81. Aphasizhev R, Suematsu T, Zhang L, Aphasizheva I. 2016. Constructive edge of uridylation-induced RNA degradation. RNA Biol 13:1078–1083 doi:10.1080/15476286.2016.1229736.

82. Le Pen J, Jiang H, Di Domenico T, Kneuss E, Kosałka J, Leung C, Morgan M, Much C, Rudolph KLM, Enright AJ, O'Carroll D, Wang D, Miska EA. 2018. Terminal uridylyltransferases target RNA viruses as part of the innate immune system. Nat Struct Mol Biol 25:778–786 doi:10.1038/s41594-018-0106-9.

83. Zuber H, Scheer H, Ferrier E, Sement FM, Mercier P, Stupfler B, Gagliardi D. 2016. Uridylation and PABP Cooperate to Repair mRNA Deadenylated Ends in *Arabidopsis*. Cell Rep 14:2707–2717 doi:10.1016/j.celrep.2016.02.060.

84. Haas G, Cetin S, Messmer M, Chane-Woon-Ming B, Terenzi O, Chicher J, Kuhn L, Hammann P, Pfeffer S. 2016. Identification of factors involved in target RNA-directed microRNA degradation. Nucleic Acids Res 44:2873–2887 doi:10.1093/nar/gkw040.

85. Łabno A, Warkocki Z, Kulínski T, Krawczyk PS, Bijata K, Tomecki R, Dziembowski A. 2016. Perlman syndrome nuclease DIS3L2 controls cytoplasmic non-coding RNAs and provides surveillance pathway for maturing snRNAs. Nucleic Acids Res 44:10437–10453 doi:10.1093/nar/gkw649.

86. Ustianenko D, Pasulka J, Feketova Z, Bednarik L, Zigackova D, Fortova A, Zavolan M, Vanacova S. 2016. TUT-DIS3L2 is a mammalian surveillance pathway for aberrant structured non-coding RNAs. EMBO J 35:2179–2191 doi:10.15252/embj.201694857.

87. Ibrahim F, Rymarquis LA, Kim EJ, Becker J, Balassa E, Green PJ, Cerutti H. 2010. Uridylation of mature miRNAs and siRNAs by the MUT68 nucleotidyltransferase promotes their degradation in *Chlamydomonas*. Proc Natl Acad Sci USA 107:3906–3911 doi:10.1073/pnas.0912632107.

88. Reimão-Pinto MM, Ignatova V, Burkard TR, Hung JH, Manzenreither RA, Sowemimo I, Herzog VA, Reichholf B, Fariña-Lopez S, Ameres SL. 2015. Uridylation of RNA Hairpins by Tailor Confines the Emergence of MicroRNAs in *Drosophila*. Mol Cell 59:203–216 doi:10.1016/j.molcel.2015.05.033.

89. Reimão-Pinto MM, Manzenreither RA, Burkard TR, Sledz P, Jinek M, Mechtler K, Ameres SL. 2016. Molecular basis for cytoplasmic RNA surveillance by uridylation-triggered decay in *Drosophila*. EMBO J 35:2417–2434 doi:10.15252/embj.201695164.

90. Schmidt K, Butler JS. 2013. Nuclear RNA surveillance: Role of TRAMP in controlling exosome specificity. Wiley Interdiscip Rev RNA 4:217–231 doi:10.1002/wrna.1155.

91. Weaver BP, Zabinsky R, Weaver YM, Lee ES, Xue D, Han M. 2014. CED-3 caspase acts with miRNAs to regulate non-apoptotic gene expression dynamics for robust development in *C. elegans*. eLife 3:1–22 doi:10.7554/eLife.04265.

92. Zhang W, Murphy C, Sieburth LE. 2010. Conserved RNaseII domain protein functions in cytoplasmic mRNA decay and suppresses Arabidopsis decapping mutant phenotypes. Proc Natl Acad Sci USA 107:15981–15985 doi:10.1073/pnas.1007060107.

93. Pirouz M, Munafò M, Ebrahimi AG, Choe J, Gregory RI. 2019. Exonuclease requirements for mammalian ribosomal RNA biogenesis and surveillance. Nat Struct Mol Biol 26:490–500 doi:10.1038/s41594-019-0234-x ISBN:4159401902.

94. Pirouz M, Wang CH, Liu Q, Ebrahimi AG, Shamsi F, Tseng YH, Gregory RI. 2020. The Perlman syndrome DIS3L2 exoribonuclease safeguards endoplasmic reticulum-targeted mRNA translation and calcium ion homeostasis. Nat Commun 11:1–13 doi:10.1038/s41467-020-16418-y.

95. Faehnle CR, Walleshauser J, Joshua-Tor L. 2014. Mechanism of Dis3l2 substrate recognition in the Lin28-let-7 pathway. Nature 514:252–256 doi:10.1038/nature13553.

96. Towler BP, Jones CI, Harper KL, Waldron JA, Newbury SF. 2016. A novel role for the 3′-5′ exoribonuclease Dis3L2 in controlling cell proliferation and tissue growth. RNA Biol 13:1286–1299 doi:10.1080/15476286.2016.1232238.

97. Towler BP, Pashler AL, Haime HJ, Przybyl KM, Viegas SC, Matos RG, Morley SJ, Arraiano CM, Newbury SF. 2020. Dis3L2 regulates cell proliferation and tissue growth though a conserved mechanism. PLoS Genet 16:1–29 doi:10.1371/journal.pgen.1009297.

98. Rodríguez-Galán A, Dosil SG, Gómez MJ, Fernández-Delgado I, Fernández-Messina L, Sánchez-Cabo F, Sánchez-Madrid F. 2021. MiRNA post-transcriptional modification dynamics in T cell activation. iScience 24 doi:10.1016/j.isci.2021.102530.

99.   Morgan M, Kabayama Y, Much C, Ivanova I, Di Giacomo M, Auchynnikava T, Monahan JM, Vitsios DM, Vasiliauskaitė L, Comazzetto S, Rappsilber J, Allshire RC, Porse BT, Enright AJ, O'Carroll D. 2019. A programmed wave of uridylation-primed mRNA degradation is essential for meiotic progression and mammalian spermatogenesis. Cell Res 29:221–232 doi:10.1038/s41422-018-0128-1.

100.   Astuti D, Morris MR, Cooper WN, Staals RHJ, Wake NC, Fews GA, Gill H, Gentle D, Shuib S, Ricketts CJ, Cole T, Van Essen AJ, Van Lingen RA, Neri G, Opitz JM, Rump P, Stolte-Dijkstra I, Müller F, Pruijn GJM, Latif F, Maher ER. 2012. Germline mutations in DIS3L2 cause the Perlman syndrome of overgrowth and Wilms tumor susceptibility. Nat Genet 44:277–284 doi:10.1038/ng.1071.

101.   Morris MR, Astuti D, Maher ER. 2013. Perlman Syndrome: Overgrowth, Wilms Tumor Predisposition and DIS3L2. Am J Med Genet Part C Semin Med Genet 163:106–113 doi:10.1002/ajmg.c.31358.

102.   Thomas MP, Liu X, Whangbo J, McCrossan G, Sanborn KB, Basar E, Walch M, Lieberman J. 2015. Apoptosis Triggers Specific, Rapid, and Global mRNA Decay with 3′ Uridylated Intermediates Degraded by DIS3L2. Cell Rep 11:1079–1089 doi:10.1016/j.celrep.2015.04.026.

103.   Liu X, Fu R, Pan Y, Meza-Sosa KF, Zhang Z, Lieberman J. 2018. PNPT1 Release from Mitochondria during Apoptosis Triggers Decay of Poly(A) RNAs. Cell 174:187-201.e12 doi:10.1016/j.cell.2018.04.017.

104.   Duncan-Lewis C, Hartenian E, King V, Glaunsinger BA. 2021. Cytoplasmic mRNA decay represses RNA Polymerase II transcription during early apoptosis. eLife 10:1–27 doi:10.7554/eLife.58342.

105.   Kurosaki T, Miyoshi K, Myers JR, Maquat LE. 2018. NMD-degradome sequencing reveals ribosome-bound intermediates with 3′-end non-templated nucleotides. Nat Struct Mol Biol 25:940–950 doi:10.1038/s41594-018-0132-7.

106.   Büssing I, Slack FJ, Großhans H. 2008. let-7 microRNAs in development, stem cells and cancer. Trends Mol Med 14:400–409 doi:10.1016/j.molmed.2008.07.001.

107.   Ha M, Kim VN. 2014. Regulation of microRNA biogenesis. Nat Rev Mol Cell Biol 15:509–524 doi:10.1038/nrm3838.

108.   Heo I, Ha M, Lim J, Yoon MJ, Park JE, Kwon SC, Chang H, Kim VN. 2012. Mono-uridylation of pre-microRNA as a key step in the biogenesis of group II let-7 microRNAs. Cell 151:521–532 doi:10.1016/j.cell.2012.09.022.

109.   Faehnle CR, Walleshauser J, Joshua-Tor L. 2017. Multi-domain utilization by TUT4 and TUT7 in control of let-7 biogenesis. Nat Struct Mol Biol 24:658–665 doi:10.1038/nsmb.3428.

110.   Piskounova E, Polytarchou C, Thornton JE, Lapierre RJ, Pothoulakis C, Hagan JP, Iliopoulos D, Gregory RI. 2011. Lin28A and Lin28B inhibit let-7 MicroRNA biogenesis by distinct mechanisms. Cell 147:1066–1079 doi:10.1016/j.cell.2011.10.039.

111.   Thornton JE, Chang HM, Piskounova E, Gregory RI. 2012. Lin28-mediated control of let-7 microRNA expression by alternative TUTases Zcchc11 (TUT4) and Zcchc6 (TUT7). RNA 18:1875–1885 doi:10.1261/rna.034538.112.

112.   Nowak JS, Hobor F, Velasco ADR, Choudhury NR, Heikel G, Kerr A, Ramos A, Michlewski G. 2017. Lin28a uses distinct mechanisms of binding to RNA and affects miRNA levels positively and negatively. RNA 23:317–332 doi:10.1261/rna.059196.116.

113.   Kim B, Ha M, Loeff L, Chang H, Simanshu DK, Li S, Fareh M, Patel DJ, Joo C, Kim VN. 2015. TUT 7 controls the fate of precursor micro RNA s by using three different uridylation mechanisms. EMBO J 34:1801–1815 doi:10.15252/embj.201590931.

114.   Luan S, Luo J, Liu H, Li Z. 2019. Regulation of RNA decay and cellular function by 3′-5′ exoribonuclease DIS3L2. RNA Biol 16:160–165 doi:10.1080/15476286.2018.1564466.

115.   Luo Y, Na Z, Slavoff SA. 2018. P-Bodies: Composition, Properties, and Functions. Biochemistry 57:2424–2431 doi:10.1021/acs.biochem.7b01162.

116.   Riggs CL, Kedersha N, Ivanov P, Anderson P. 2020. Mammalian stress granules and P bodies at a glance. J Cell Sci 133:1–9 doi:10.1242/jcs.242487.

117.   Sheth U, Parker R. 2003. Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. Science 300:805–808 doi:10.1126/science.1082320.

118. Decker CJ, Parker R. 2012. P-Bodies and Stress Granules: Possible Roles in the Control of Translation and mRNA Degradation. Cold Spring Harb Perspect Biol 4:1–16 doi:10.1101/cshperspect.a012286.

119. Eulalio A, Behm-Ansmant I, Schweizer D, Izaurralde E. 2007. P-Body Formation Is a Consequence, Not the Cause, of RNA-Mediated Gene Silencing. Mol Cell Biol 27:3970–3981 doi:10.1128/mcb.00128-07.

120. Hubstenberger A, Courel M, Bénard M, Souquere S, Ernoult-Lange M, Chouaib R, Yi Z, Morlot JB, Munier A, Fradet M, Daunesse M, Bertrand E, Pierron G, Mozziconacci J, Kress M, Weil D. 2017. P-Body Purification Reveals the Condensation of Repressed mRNA Regulons. Mol Cell 68:144-157.e5 doi:10.1016/j.molcel.2017.09.003.

121. Reis FP, Pobre V, Silva IJ, Malecki M, Arraiano CM. 2013. The RNase II/RNB family of exoribonucleases: Putting the "Dis" in disease. Wiley Interdiscip Rev RNA 4:607–615 doi:10.1002/wrna.1180.

122. Matos RG, Bárria C, Moreira RN, Barahona S, Domingues S, Arraiano CM. 2014. The importance of proteins of the RNase II/RNB-family in pathogenic bacteria. Front Cell Infect Microbiol 4 doi:10.3389/fcimb.2014.00068.

123. Matos RG, Barbas A, Gómez-Puertas P, Arraiano CM. 2011. Swapping the domains of exoribonucleases RNase II and RNase R: Conferring upon RNase II the ability to degrade ds RNA. Proteins Struct Funct Bioinforma 79:1853–1867 doi:10.1002/prot.23010.

124. Frazão C, McVey CE, Amblar M, Barbas A, Vonrhein C, Arraiano CM, Carrondo MA. 2006. Unravelling the dynamics of RNA degradation by ribonuclease II and its RNA-bound complex. Nature 443:110–114 doi:10.1038/nature05080.

125. Amblar M, Barbas A, Fialho AM, Arraiano CM. 2006. Characterization of the Functional Domains of *Escherichia coli* RNase II. J Mol Biol 360:921–933 doi:10.1016/j.jmb.2006.05.043.

126. Matos RG, Fialho AM, Giloh M, Schuster G, Arraiano CM. 2012. The *rnb* gene of synechocystis PCC6803 encodes a RNA hydrolase displaying RNase II and not RNase R enzymatic properties. PLoS One 7:1–12 doi:10.1371/journal.pone.0032690.

127. Lorentzen E, Basquin J, Tomecki R, Dziembowski A, Conti E. 2008. Structure of the Active Subunit of the Yeast Exosome Core, Rrp44: Diverse Modes of Substrate Recruitment in the RNase II Nuclease Family. Mol Cell 29:717–728 doi:10.1016/j.molcel.2008.02.018.

128. Matos RG, Barbas A, Arraiano CM. 2009. RNase R mutants elucidate the catalysis of structured RNA: RNA-binding domains select the RNAs targeted for degradation. Biochem J 423:291–301 doi:10.1042/BJ20090839.

129. Awano N, Rajagopal V, Arbing M, Patel S, Hunt J, Inouye M, Phadtare S. 2010. *Escherichia coli* RNase R has dual activities, Helicase and RNase. J Bacteriol 192:1344–1352 doi:10.1128/JB.01368-09.

130. Chu LY, Hsieh TJ, Golzarroshan B, Chen YP, Agrawal S, Yuan HS. 2017. Structural insights into RNA unwinding and degradation by RNase R. Nucleic Acids Res 45:12015–12024 doi:10.1093/nar/gkx880.

131. Chung CZ, Jaramillo JE, Ellis MJ, Bour DYN, Seidl LE, Jo DHS, Turk MA, Mann MR, Bi Y, Haniford DB, Duennwald ML, Heinemann IU. 2019. RNA surveillance by uridylation-dependent RNA decay in *Schizosaccharomyces pombe*. Nucleic Acids Res 47:3045–3057 doi:10.1093/nar/gkz043.

132. Moazed D. 2009. Small RNAs in transcriptional gene silencing and genome defence. Nature 457:413–420 doi:10.1038/nature07756.

133. Lee M, Kim B, Kim VN. 2014. Emerging roles of RNA modification: M6A and U-Tail. Cell 158:980–987 doi:10.1016/j.cell.2014.08.005.

134. Schaeffer D, Reis FP, Johnson SJ, Arraiano CM, Van Hoof A. 2012. The CR3 motif of Rrp44p is important for interaction with the core exosome and exosome function. Nucleic Acids Res 40:9298–9307 doi:10.1093/nar/gks693.

135. Lebreton A, Tomecki R, Dziembowski A, Séraphin B. 2008. Endonucleolytic RNA cleavage by a eukaryotic exosome. Nature 456:993–996 doi:10.1038/nature07480.

136. Schneider C, Leung E, Brown J, Tollervey D. 2009. The N-terminal PIN domain of the exosome subunit Rrp44 harbors endonuclease activity and tethers Rrp44 to the yeast core exosome. Nucleic Acids Res 37:1127–1140 doi:10.1093/nar/gkn1020.

137. Schaeffer D, Tsanova B, Barbas A, Reis FP, Dastidar EG, Sanchez-Rotunno M, Arraiano CM, Van Hoof A. 2009. The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities. Nat Struct Mol Biol 16:56–62 doi:10.1038/nsmb.1528.

138. Lv H, Zhu Y, Qiu Y, Niu L, Teng M, Li X. 2015. Structural analysis of Dis3l2, an exosome-independent exonuclease from *Schizosaccharomyces pombe*. Acta Crystallogr Sect D Biol Crystallogr 71:1284–1294 doi:10.1107/S1399004715005805.

139. Saramago M, da Costa PJ, Viegas SC, Arraiano CM. 2019. The Implication of mRNA Degradation Disorders on Human DISease: Focus on DIS3 and DIS3-Like Enzymes, p. 85–98. *In* Romão, L (ed.), The mRNA Metabolism in Human Disease. Springer, Cham, Lisboa, Portugal doi:10.1007/978-3-030-19966-1_4 ISBN:9783030199661.

140. Weißbach S, Langer C, Puppe B, Nedeva T, Bach E, Kull M, Bargou R, Einsele H, Rosenwald A, Knop S, Leich E. 2015. The molecular spectrum and clinical impact of DIS3 mutations in multiple myeloma. Br J Haematol 169:57–70 doi:10.1111/bjh.13256.

141. Robinson SR, Viegas SC, Matos RG, Domingues S, Bedir M, Stewart HJS, Chevassut TJ, Oliver AW, Arraiano CM, Newbury SF. 2018. Dis3 isoforms vary in their endoribonuclease activity and are differentially expressed within haematological cancers. Biochem J 475:2091–2105 doi:10.1042/BCJ20170962.

142. Palmer CJ, Galan-Caridad JM, Weisberg SP, Lei L, Esquilin JM, Croft GF, Wainwright B, Canoll P, Owens DM, Reizis B. 2014. Zfx facilitates tumorigenesis caused by activation of the Hedgehog pathway. Cancer Res 74:5914–5924 doi:10.1158/0008-5472.CAN-14-0834.

143. Pashler AL, Towler BP, Jones CI, Newbury SF. 2016. The roles of the exoribonucleases DIS3L2 and XRN1 in human disease. Biochem Soc Trans 44:1377–1384 doi:10.1042/BST20160107.

144. NCBI - Gene. 2021. DIS3L2 - DIS3 like 3′-5′ exoribonuclease 2 [*Homo sapiens* (human)]. https://www.ncbi.nlm.nih.gov/gene/129563/, accessed 2 November 2021.

145. EMBL-EBI - e!Ensembl. 2021. Human, Gene: DIS3L2. http://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000144535;r=2:231961713-232337192;t=ENST00000325385, accessed 2 November 2021.

146. Uniprot. 2021. UniProtKB - Q8IYB7 (DI3L2_HUMAN). https://www.uniprot.org/uniprot/Q8IYB7, accessed 2 November 2021.

147. neXtprot. 2021. DIS3L2 → DIS3-like exonuclease 2 [EC 3.1.13.-] (hDIS3L2). https://www.nextprot.org/entry/NX_Q8IYB7/exons, accessed 2 November 2021.

148. The Human Protein Atlas. DIS3L2. https://www.proteinatlas.org/ENSG00000144535-DIS3L2, accessed 2 November 2021.

149. Soma N, Higashimoto K, Imamura M, Saitoh A, Soejima H, Nagasaki K. 2017. Long term survival of a patient with Perlman syndrome due to novel compound heterozygous missense mutations in RNB domain of DIS3L2. Am J Med Genet Part A 173:1077–1081 doi:10.1002/ajmg.a.38111.

150. Gowans LJJ, Al Dhaheri N, Li M, Busch T, Obiri-Yeboah S, Oti AA, Sabbah DK, Arthur FKN, Awotoye WO, Alade AA, Twumasi P, Agbenorku P, Plange-Rhule G, Naicker T, Donkor P, Murray JC, Sobreira NLM, Butali A. 2021. Co-occurrence of orofacial clefts and clubfoot phenotypes in a sub-Saharan African cohort: Whole-exome sequencing implicates multiple syndromes and genes. Mol Genet Genomic Med 9:1–15 doi:10.1002/mgg3.1655.

151. Higashimoto K, Maeda T, Okada J, Ohtsuka Y, Sasaki K, Hirose A, Nomiyama M, Takayanagi T, Fukuzawa R, Yatsuki H, Koide K, Nishioka K, Joh K, Watanabe Y, Yoshiura KI, Soejima H. 2013. Homozygous deletion of dis3l2 exon 9 due to non-allelic homologous recombination between line-1s in a japanese patient with perlman syndrome. Eur J Hum Genet 21:1316–1319 doi:10.1038/ejhg.2013.45.

152. Dome JS, Huff V. 2003 [Updated 2016]. Wilms Tumor Predisposition. *In* Adam MP, Ardinger HH, Pagon RA, *et al.* (eds.). GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1294/.

153. Wegert J, Ishaque N, Vardapour R, Geörg C, Gu Z, Bieg M, Ziegler B, Bausenwein S, Nourkami N, Ludwig N, Keller A, Grimm C, Kneitz S, Williams RD, Chagtai T, Pritchard-Jones K, vanSluis P, Volckmann R, Koster J, Versteeg R, Acha T, O'Sullivan MJ, Bode PK, Niggli F, Tytgat GA, vanTinteren H, vandenHeuvel-Eibrink MM, Meese E, Vokuhl C, Leuschner I, Graf N, Eils R, Pfister SM, Kool M, Gessler M. 2015. Mutations in the SIX1/2 Pathway and the DROSHA/DGCR8 miRNA Microprocessor Complex Underlie High-Risk Blastemal Type Wilms Tumors. Cancer Cell 27:298–311 doi:10.1016/j.ccell.2015.01.002.

154. Drake KM, Ruteshouser EC, Natrajan R, Harbor P, Wegert J, Gessler M, Pritchard-Jones K, Grundy P, Dome J, Huff V, Jones C, Aldred MA. 2009. Loss of heterozygosity at 2q37 in sporadic Wilms' tumor: Putative role for miR-562. Clin Cancer Res 15:5985–5992 doi:10.1158/1078-0432.CCR-09-1065.

155. Mori F, Tanji K, Miki Y, Toyoshima Y, Sasaki H, Yoshida M, Kakita A, Takahashi H, Wakabayashi K. 2018. Immunohistochemical localization of exoribonucleases (DIS3L2 and XRN1) in intranuclear inclusion body disease. Neurosci Lett 662:389–394 doi:10.1016/j.neulet.2017.10.061.

156. Bocciardi R, Giorda R, Buttgereit J, Gimelli S, Divizia MT, Beri S, Garofalo S, Tavella S, Lerone M, Zuffardi O, Bader M, Ravazzolo R, Gimelli G. 2007. Overexpression of the C-type natriuretic peptide (CNP) is associated with overgrowth and bone anomalies in an individual with balanced t(2;7) translocation. Hum Mutat 28:724–731 doi:10.1002/humu.20511.

157. Moncla A, Missirian C, Cacciagli P, Balzamo E, Legeai-Mallet L, Jouve JL, Chabrol B, Le Merrer M, Plessis G, Villard L, Philip N. 2007. A cluster of translocation breakpoints in 2q37 is associated with overexpression of *NPPC* in patients with a similar overgrowth phenotype. Hum Mutat 28:1183–1188 doi:10.1002/humu.20611.

158. Tassano E, Buttgereit J, Bader M, Lerone M, Divizia MT, Bocciardi R, Napoli F, Pala G, Sloan-Béna F, Gimelli S, Gimelli G. 2013. Genotype-Phenotype Correlation of 2q37 Deletions Including *NPPC* Gene Associated with Skeletal Malformations. PLoS One 8 doi:10.1371/journal.pone.0066048.

159. Liu W, Yu Q, Ma J, Cheng Y, Zhang H, Luo W, Yao J, Zhang H. 2017. Knockdown of a DIS3l2 promoter upstream long noncoding RNA (AC105461.1) enhances colorectal cancer stem cell properties in vitro by down-regulating DIS3l2. Onco Targets Ther 10:2367–2376 doi:10.2147/OTT.S132708.

160. Pinto AT, Pojo M, Simões-Pereira J, Roque R, Saramago A, Roque L, Martins C, André S, Cabeçadas J, Leite V, Cavaco BM. 2019. Establishment and characterization of a new patient-derived anaplastic thyroid cancer cell line (C3948), obtained through fine-needle aspiration cytology. Endocrine 66:288–300 doi:10.1007/s12020-019-02009-5.

161. Xing S, Li Z, Ma W, He X, Shen S, Wei H, Li ST, Shu Y, Sun L, Zhong X, Huangfu Y, Su L, Feng J, Zhang X, Gao P, Jia WD, Zhang H. 2019. DIS3L2 promotes progression of hepatocellular carcinoma via hnRNP U-mediated alternative splicing. Cancer Res 79:4923–4936 doi:10.1158/0008-5472.CAN-19-0376.

162. Yao L, Cong R, Ji C, Zhou X, Luan J, Meng X, Song N. 2021. RNA-Binding Proteins Play an Important Role in the Prognosis of Patients With Testicular Germ Cell Tumor. Front Genet 12:1–21 doi:10.3389/fgene.2021.610291.

163. Bradford MM. 1976. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. Anal Biochem 72:248–254 doi:10.1006/abio.1976.9999.

164. Meze K, Axhemi A, Thomas DR, Doymaz A, Joshua-Tor L. 2021. A shape-shifting nuclease unravels structured RNA. bioRxiv doi:10.1101/2021.11.30.470623.

165. Amblar M, Arraiano CM. 2005. A single mutation in *Escherichia coli* ribonuclease II inactivates the enzyme without affecting RNA binding. FEBS J 272:363–374 doi:10.1111/j.1742-4658.2004.04477.x.

166. Charpentier X, Faucher SP, Kalachikov S, Shuman HA. 2008. Loss of RNase R induces competence development in *Legionella pneumophila*. J Bacteriol 190:8126–8136 doi:10.1128/JB.01035-08.

167. Agilent Technologies. 2015. BL21-CodonPlus Competent Cells - Instruction Manual. https://www.agilent.com/cs/library/usermanuals/public/230240.pdf, accessed 14 October 2020.

168. Golovanov AP, Hautbergue GM, Wilson SA, Lian LY. 2004. A simple method for improving protein solubility and long-term stability. J Am Chem Soc 126:8933–8939 doi:10.1021/ja049297h.

169. Shukla D, Trout BL. 2011. Understanding the synergistic effect of arginine and glutamic acid mixtures on protein solubility. J Phys Chem B 115:11831–11839 doi:10.1021/jp204462t.

170. Cannistraro VJ, Kennell D. 1994. The processive reaction mechanism of ribonuclease II. J Mol Biol doi:10.1006/jmbi.1994.1693.

171. Pirouz M, Du P, Munafò M, Gregory RI. 2016. Dis3l2-Mediated Decay Is a Quality Control Pathway for Noncoding RNAs. Cell Rep 16:1861–1873 doi:10.1016/j.celrep.2016.07.025.

172. Nam Y, Chen C, Gregory RI, Chou JJ, Sliz P. 2011. Molecular basis for interaction of let-7 MicroRNAs with Lin28. Cell 147:1080–1091 doi:10.1016/j.cell.2011.10.020.

173. Cheng ZF, Zuo Y, Li Z, Rudd KE, Deutscher MP. 1998. The vacB gene required for virulence in *Shigella flexneri* and *Escherichia coli* encodes the exoribonuclease RNase R. J Biol Chem 273:14077–14080 doi:10.1074/jbc.273.23.14077.

174. Erova TE, Kosykh VG, Fadl AA, Sha J, Horneman AJ, Chopra AK. 2008. Cold shock exoribonuclease R (VacB) is involved in *Aeromonas hydrophila* pathogenesis. J Bacteriol 190:3467–3474 doi:10.1128/JB.00075-08.

175. Haddad N, Matos RG, Pinto T, Rannou P, Cappelier JM, Prévost H, Arraiano CM. 2014. The RNase R from *Campylobacter jejuni* has unique features and is involved in the first steps of infection. J Biol Chem 289:27814–27824 doi:10.1074/jbc.M114.561795.

176. Mukherjee D, Gupta S, Ghosh A, Ghosh A. 2020. *Ustilago maydis* secreted T2 ribonucleases, Nuc1 and Nuc2 scavenge extracellular RNA. Cell Microbiol 22:1–15 doi:10.1111/cmi.13256.

177. Cánovas-Márquez JT, Falk S, Nicolás FE, Padmanabhan S, Zapata-Pérez R, Sánchez-Ferrer Á, Navarro E, Garre V. 2021. A ribonuclease III involved in virulence of Mucorales fungi has evolved to cut exclusively single-stranded RNA. Nucleic Acids Res 49:5294–5307 doi:10.1093/nar/gkab238.

178. Olombrada M, Lázaro-Gorines R, López-Rodríguez JC, Martínez-Del-Pozo Á, Oñaderra M, Maestro-López M, Lacadena J, Gavilanes JG, García-Ortega L. 2017. Fungal ribotoxins: A review of potential biotechnological applications. Toxins (Basel) 9 doi:10.3390/toxins9020071.

179. Köhler JR, Hube B, Puccia R, Casadevall A, Perfect JR. 2017. Fungi that infect humans. The Fungal Kingdom 813–843 doi:10.1128/9781555819583.ch39.

180. Sitterlé E, Rodriguez C, Mounier R, Calderaro J, Foulet F, Develoux M, Pawlotsky JM, Botterel F. 2017. Contribution of ultra deep sequencing in the clinical diagnosis of a new fungal pathogen species: *Basidiobolus meristosporus*. Front Microbiol 8:1–4 doi:10.3389/fmicb.2017.00334.

181. Bowers JR, Monroy-Nieto J, Gade L, Travis J, Refojo N, Abrantes R, Santander J, French C, Dignani MC, Hevia AI, Roe CC, Lemmer D, Lockhart SR, Chiller T, Litvintseva AP, Clara L, Engelthaler DM. 2020. *Rhizopus microsporus* Infections Associated with Surgical Procedures, Argentina, 2006–2014. Emerg Infect Dis 26:937–944 doi:10.3201/eid2605.191045.

182. Fernández-García O, Guerrero-Torres L, Roman-Montes CM, Rangel-Cordero A, Martínez-Gamboa A, Ponce-de-Leon A, Gonzalez-Lara MF. 2021. Isolation of *Rhizopus microsporus* and *Lichtheimia corymbifera* from tracheal aspirates of two immunocompetent critically ill patients with COVID-19. Med Mycol Case Rep 33:32–37 doi:10.1016/j.mmcr.2021.07.001.

183. Fujimoto A, Nagao K, Tanaka K, Yamagami J, Udagawa SI, Sugiura M. 2005. The first case of cutaneous mucormycosis caused by *Rhizopus azygosporus*. Br J Dermatol 153:428–430 doi:10.1111/j.1365-2133.2005.06593.x.

184. Kanwar A, Jordan A, Olewiler S, Wehberg K, Cortes M, Jackson BR. 2021. A fatal case of *Rhizopus azygosporus* pneumonia following COVID-19. J Fungi 7:1–6 doi:10.3390/jof7030174.

185. Hagel S, Ewald C, Doenst T, Sachse S, Roedel J, Pletz MW. 2015. Ventriculitis due to infection with *Rhizopus arrhizus*. Med Mycol Case Rep 10:18–20 doi:10.1016/j.mmcr.2015.12.004.

186. Walther G, Wagner L, Kurzai O. 2020. Outbreaks of Mucorales and the Species Involved. Mycopathologia 185:765–781 doi:10.1007/s11046-019-00403-1.

187. Maini A, Tomar G, Khanna D, Kini Y, Mehta H, Bhagyasree V. 2021. Sino-orbital mucormycosis in a COVID-19 patient: A case report. Int J Surg Case Rep 82:105957 doi:10.1016/j.ijscr.2021.105957.

188. Ballou ER, Cook AG, Wallace EWJ. 2021. Repeated Evolution of Inactive Pseudonucleases in a Fungal Branch of the Dis3/RNase II Family of Nucleases. Mol Biol Evol 38:1837–1846 doi:10.1093/molbev/msaa324.

189. Bayne RA, Jayachandran U, Kasprowicz A, Bresson S, Tollervey D, Wallace EWJ, Cook AG. 2021. Yeast Ssd1 is a non-enzymatic member of the RNase II family with an alternative RNA recognition site. Nucleic Acids Res 1–15 doi:10.1093/nar/gkab615.

190. Gank KD, Yeaman MR, Kojima S, Yount NY, Park H, Edwards JE, Filler SG, Fu Y. 2008. SSD1 is integral to host defense peptide resistance in *Candida albicans*. Eukaryot Cell 7:1318–1327 doi:10.1128/EC.00402-07.

191. Thammahong A, Dhingra S, Bultman KM, Kerkaert JD, Cramer RA. 2019. An Ssd1 Homolog Impacts Trehalose and Chitin Biosynthesis and Contributes to Virulence in *Aspergillus fumigatus*. mSphere 4 doi:10.1128/msphere.00244-19.

6. Supplementary Material

Table S.2.1. Culture Media, Solutions and Gels used in the experimental work.

CULTURE MEDIA	
Luria-Bertani Broth (LB)	**Terrific Broth (TB)**
10 g Bacto Tryptone 5 g Bacto Yeast Extract 10 g NaCl ddH2O to 1000 ml Adjust to pH 7.0 Autoclave	50.8 g Terrific Broth (*Grisp*) 0.4% (v/v) Glycerol ddH$_2$O to 1000 ml Autoclave

SOLUTIONS FOR GENERAL USE	
100 mg/ml Ampicillin (Amp)	**50 mg/ml Chloramphenicol (Cam)**
2.000 g Ampicillin MQ H$_2$O to 20 ml Sterilization by filtration (0,22 μm filter) Aliquot and store at -20 ℃	0.500 g Chloramphenicol 10 ml Ethanol 100% (v/v) Aliquot and store at -20 ℃
Tris-Borate-EDTA 10x (TBE 10x)	**5x Loading Buffer – Proteins**
121.1 g Tris base 61.8 g Boric acid 7.4 g Na$_2$EDTA MQ H$_2$O to 1000 ml	10% (w/v) SDS 10 mM β-mercaptoethanol 20% (v/v) Glycerol 0.2 M Tris-HCl pH 6.8 0.05% (w/v) Bromophenol Blue 0.3 M DTT in H$_2$O
SDS-PAGE Running Buffer 10x	
30 g Tris base 144 g Glycine 10 g SDS ddH2O to 1000 ml	

PROTEIN PURIFICATION BUFFERS (SpDis3L2)	
Initial Buffer A (A1)	**Initial Buffer B (B1)**
20 mM Tris-HCl pH 8.0 500 mM NaCl 20 mM Imidazole MQ H$_2$O to final volume	20 mM Tris-HCl pH 8.0 500 mM NaCl 500 mM Imidazole MQ H$_2$O to final volume
Initial Buffer C (C1)	**Optimised Buffer A (A2)**
20 mM Tris-HCl pH 8.0 500 mM NaCl MQ H$_2$O to final volume	100 mM MOPS pH 7.6 200 mM KCl 20 mM Imidazole MQ H$_2$O to final volume

Table S.2.1. (continued)

Optimised Buffer B (B2)		Optimised Buffer C (C2)	
100 mM	MOPS pH 7.6	100 mM	MOPS pH 7.6
200 mM	KCl	200 mM	KCl
500 mM	Imidazole		MQ H_2O to final volume
	MQ H_2O to final volume		

WESTERN BLOT SOLUTIONS

Transfer Buffer 10x (TrB 10x)		Transfer Buffer 1x (TrB 1x)	
15 g	Tris base	1x	TrB 10x
72 g	Boric acid	20% (v/v)	Methanol
7.4 g	Na_2EDTA		ddH_2O to 1000 ml
	MQ H_2O to 500 ml		

Tris-buffered saline 10x (TBS 10x)		TBS 1x with 0.1% Tween 20 (TBS-T 1x)	
24 g	Tris base (200 mM)	100 ml	TBS 10x (20 mM Tris, 150 mM NaCl)
88 g	NaCl (1500 mM)	1 ml	Tween 20 (0.1% (v/v))
Dissolve in 900 ml MQ H_2O			ddH_2O to 1000 ml
Adjust pH to 7.6 with 12 N HCl			
MQ H_2O to 1000 ml			

SOLUTIONS FOR ASSAYS WITH RNA SUBSTRATES AND RNASES

Buffer C (EcRNase II)	
10 mM	Tris-HCl pH 8.0
50 mM	KCl

Storage Buffer (EcRNase II)		Storage Buffer (SpDis3L2)	
1 vol	Buffer C (EcRNase II)	1 vol	Buffer C2 (SpDis3L2)
1 vol	Glycerol 100% (v/v)	1 vol	Glycerol 100% (v/v)

5x Reaction Buffer (EcRNase II)		5x Reaction Buffer (Dis3L2)	
500 mM	KCl	125 mM	NaCl
100 mM	$MgCl_2$	25 mM	$MgCl_2$
100 mM	Tris-HCl pH 8.0	50 mM	Tris-HCl pH 7.5
5 mM	DTT	1 mM	DTT
	Sigma H_2O to final volume		Sigma H_2O to final volume

5x Loading Buffer – PAA gels		STOP solution
98% (v/v)	Deionised Formamide	5x Loading Buffer – PAA gels
0.025% (w/v)	Xylene Cyanole	supplemented with 0.5 mg/ml total yeast RNA
0.025% (w/v)	Bromophenol Blue	
10 mM	EDTA pH 8.0	

Table S.2.1. (continued)

POLYACRYLAMIDE (PAA) GELS						
20% PAA denaturing gel				**20% PAA non-denaturing/native gel**		
For a medium- vs. high-resolution gel:				For a medium-resolution gel:		
21 g	29,4 g	Urea	(7 M)			
5 ml	7 ml	TBE 10x	(1x)	5 ml	TBE 10x	(1x)
25 ml	35 ml	PAA 40% 19:1 (20%)		25 ml	PAA 40% (w/v) 19:1 (20% (w/v))	
Add MQ H$_2$O to:				Add MQ H$_2$O to:		
50 ml	70 ml			50 ml		
To polymerize, add:				To polymerize, add:		
400 µl	300 µl	APS 10% (w/v)		400 µl	APS 10% (w/v)	
40 µl	50 µl	TEMED		40 µl	TEMED	

Table S.2.2. Reactions prepared for the [Protein] vs. [Dye] Assay. Volumes of each solution added to the reaction samples (final volume of 20 μl) are indicated in μl.

[Dye]	Solution to add:	[Protein]		
		1 μg	3 μg	5 μg
1x	Protein Buffer	18	16	14
	SYPRO Orange	1	1	1
	Protein	1	3	5
3x	Protein Buffer	16	14	12
	SYPRO Orange	3	3	3
	Protein	1	3	5
5x	Protein Buffer	14	12	10
	SYPRO Orange	5	5	5
	Protein	1	3	5

Table S.2.3. Matrix of tested conditions in the first TSA with the wt SpDis3L2 protein. Chosen conditions were adapted from Durham pH and salt screenings (*Molecular Dimensions*). Reactions were prepared in PCR tubes placed in a 96-well plate with 10 µl screen reagent (condition indicated in the matrix), 5 µl dye (SYPRO Orange), and 5 µl purified wt SpDis3L2 protein.

	1	2	3	4	5	6	7	8	9	10	11	12
A	water	water	4.0 M urea	100 mM bisTRIS pH 6.1	100 mM bisTRIS pH 6.6	100 mM bisTRIS pH 7.1	100 mM phosphate pH 6.3	100 mM phosphate pH 6.8	100 mM phosphate pH 7.3	100 mM MOPS pH 6.6	100 mM MOPS pH 7.1	100 mM MOPS pH 7.6
B	100 mM HEPES pH 7.0	100 mM HEPES pH 7.5	100 mM HEPES pH 8.0	100 mM TRIS pH 7.7	100 mM TRIS pH 8.2	100 mM TRIS pH 8.7	100 mM boric acid pH 8.6	100 mM boric acid pH 9.1	100 mM boric acid pH 9.6	100 mM glycine pH 9.2	100 mM glycine pH 9.7	100 mM glycine pH 10.2
C	500 mM Na_3 citrate	200 mM Na_3 citrate	1.5 M $(NH_4)_2SO_4$	1.0 M $(NH_4)_2SO_4$	800 mM $(NH_4)_2SO_4$	600 mM $(NH_4)_2SO_4$	400 mM $(NH_4)_2SO_4$	200 mM $(NH_4)_2SO_4$	500 mM KCl	200 mM KCl	400 mM $MgCl_2$	5 mM $MgCl_2$
D	1.5 M NaCl	1.0 M NaCl	800 mM NaCl	600 mM NaCl	400 mM NaCl	200 mM NaCl	1.5 M NH_4Cl	1.0 M NH_4Cl	800 mM NH_4Cl	600 mM NH_4Cl	400 mM NH_4Cl	200 mM NH_4Cl
E	1.0 M $MgSO_4$	800 mM $MgSO_4$	600 mM $MgSO_4$	400 mM $MgSO_4$	200 mM $MgSO_4$	5 mM $MnCl_2$	0.5 mM $MnCl_2$	5 mM DTT	5 mM β-mercapto-etanol			

Table S.2.4. Matrix of tested conditions in the second TSA with the wt SpDis3L2 protein. pH and salt conditions were chosen according the results obtained in the first TSA, whereas individual buffers or salt concentrations (1st TSA) were combined together in single reactions (2nd TSA) in multiple ways. Reactions were prepared in PCR tubes placed in a 96-well plate with 10 μl screen reagent (condition indicated in the matrix), 5 μl dye (SYPRO Orange), and 5 μl purified wt SpDis3L2 protein.

	1	2	3
A	water	100 mM MOPS pH 7.6	100 mM HEPES pH 8.0
B	200 mM KCl	100 mM MOPS pH 7.6 + 200 mM KCl	100 mM HEPES pH 8.0 + 200 mM KCl
C		100 mM MOPS pH 7.6 + 200 mM KCl + 5 mM MgCl$_2$	100 mM HEPES pH 8.0 + 200 mM KCl + 5 mM MgCl$_2$
D	400 mM NaCl	100 mM MOPS pH 7.6 + 400 mM NaCl	100 mM HEPES pH 8.0 + 400 mM NaCl
E		100 mM MOPS pH 7.6 + 400 mM NaCl + 5 mM MgCl$_2$	100 mM HEPES pH 8.0 + 400 mM NaCl + 5 mM MgCl$_2$
F	400 mM NH$_4$Cl	100 mM MOPS pH 7.6 + 400 mM NH$_4$Cl	100 mM HEPES pH 8.0 + 400 mM NH$_4$Cl
G	5 mM MgCl$_2$	100 mM MOPS pH 7.6 + 400 mM NH$_4$Cl + 5 mM MgCl$_2$	100 mM HEPES pH 8.0 + 400 mM NH$_4$Cl + 5 mM MgCl$_2$

Table S.3.1. Midpoint temperatures of the unfolding transition for the wt SpDis3L2 protein in the presence of different buffers (1st TSA). Reactions were prepared with solutions according Table S.2.3., and subjected to a temperature gradient from 20 ℃ to 95 ℃, 1 ℃/min. Numeric values represent the melting temperatures (T_ms) in degrees Celsius (℃) calculated automatically by the qPCRsoft touch software, implemented in the qTOWER³ G touch equipment (*Analytik Jena*).

	1	2	3	4	5	6	7	8	9	10	11	12
A	66.3	65.7	67.1	65.7	67.8	86.1	78.4	68.0	63.7	65.2	63.9	69.9
B	m.i.	90.6	67.2	67.5	79.4	79.7	64.4	62.9	67.9	77.3	66.9	72.6
C	82.6	91.0	91.3	83.9	81.9	73.6	71.7	80.7	67.6	68.6	82.1	80.9
D	76.7	65.1	88.0	67.3	78.3	74.0	88.5	79.4	82.9	70.6	86.1	69.0
E	87.5	86.8	66.4	64.6	64.9	64.9	71.4	68.8	80.8			

m.i.: missing information (data was not collected).

Table S.3.2. Midpoint temperatures of the unfolding transition for the wt SpDis3L2 protein in the presence of different buffers (2nd TSA). Reactions were prepared with solutions according Table S.2.4., and subjected to a temperature gradient from 20 ºC to 95 ºC, 1 ºC/min. Numeric values represent the melting temperatures (T_ms) in degrees Celsius (ºC) calculated automatically by the qPCRsoft touch software, implemented in the qTOWER³ G touch equipment (*Analytik Jena*).

	1	2	3
A	70,9	65,8	74,7
B	64,9	65,8	65,9
C		64,5	86
D	89,9	77,1	70
E		90,6	86,8
F	73	73,1	69,6
G	84	85	90,9

```
HsDis3L2      1 MSHPDYRM-----------------------------NLRPL---------------
MmDis3L2      1 MNHPDYKL-----------------------------NLRSP---------------
CeDis3L2      1 MFFRVKT------------------------------------ -------------
AtDis3L2      1 MKSASSEQ----------------------------SVERI-----E----------
SpDis3L2      1 MDLKPNIRRK---------------------EKRNLLKGEAALEK------------
DmDis3L2      1 MPYPLYPAHSQVMSETSSVNVDADRDCRRIEGLRQRSQRLAARL-----QQMQREMESNA
SpDis3        1 MSTVSGLKRPQSS--EKN-----HR----------DRVFV-RA-----T---------
ScDis3        1 MSVPAIAPRRKRLADGLS-----VT----------QKVFV-RS-----R---------
HsDis3        1 MLKSK------------------------TFLKKT-----R---------
HsDis3L1      1 MLQKR----------------------EKVLLLRT-----F---------


HsDis3L2     14 -------G-------------------------------------------------
MmDis3L2     14 -------G-------------------------------------------------
CeDis3L2      8 ---------------------------------------------------------
AtDis3L2     15 -----NGHKKK--R-------NRPQK-------------------------------
SpDis3L2     25 -----KGSIDRKTKNKAYPSTTHDPHQN-----------------------------
DmDis3L2     56 SSKDLSANNAVI---------PLQPLMRWAEPRPHPNQSQ------PSARSNSRQAIPYG
SpDis3       27 -----RGKVQKV---------VREQYLRNDIPCQSRACPLCRSKLPKDSRGNVLEPILSE
ScDis3       29 -----NGGATKI---------VREHYLRSDIPCLSRSCTKCPQIVVPDAQNELPKFILSD
HsDis3       13 -----AGGVMKI---------VREHYLRDDIGCGAPGCAACGGAH----EGPALEPQP-Q
HsDis3L1     15 -----QGRTLRI---------VREHYLRPCVPCHSPLCPQPA------ACS--------H


HsDis3L2     15 ---------------------------------------------------------
MmDis3L2     15 ---------------------------------------------------------
CeDis3L2      8 ---------------------------------------------------------
AtDis3L2     27 ---------------------------------------------------------
SpDis3L2     48 --------------DDSNIPGLGSGLLERIKDIVQ---------RPTDTQLKGQDSNHK
DmDis3L2    101 VPVSTPAPVAM--QLPYHV-------AFASYQATMQIPC-LQPLSH------------L
SpDis3       73 KPMFLE-KFGHHYLIPDSN-------TFYHCIDALEHPNNFFDVII------------L
ScDis3       75 SPLELSAPIGKHYVVLDTN-------VVLQAIDLLENPNCFFDVIV------------P
HsDis3       54 DPASSV-CPQPHYLLPDTN-------VLLHQIDVLEDPA-IRNVIV------------L
HsDis3L1     47 DGKLLS-SDVTHYVIPDWK-------VVQDYLEILEFPE-LKGIIF------------M


HsDis3L2     15 ------------------TPRGV-S--------------------------------
MmDis3L2     15 ------------------TPRGV-S--------------------------------
CeDis3L2      8 ------------------TPTKA-KP-------------------------------
AtDis3L2     27 -------Q---------N-RRSKQ-SSVP--------------------IEDAHVEESLD
SpDis3L2     84 KASLTETK---------T-EKAKV-KPKAKKKNSKE--KISKSSKQDEHKTDVHKES---
DmDis3L2    138 HVV----SSPSHRLIAKK-KSKRALRNEVEALQDMV--KHLSHLSPDVNAYACQLIR---
SpDis3      112 QTVFSEIS---------S-KSIPL-YNRMKRLCQEK--TKRFTPFSNEFFVDTFVER---
ScDis3      115 QIVLDEVR---------N-KSYPV-YTRLRTLCRDSDDHKRFIVFHNEFSEHTFVER---
HsDis3       92 QTVLQEVR---------N-RSAPV-YKRIRDVTNNQ--EKHFYTFTNEHHRETYVEQ---
HsDis3L1     85 QTACQAVQ---------HQRGRRQ-YNKLRNLLKDA--RHDCILFANEFQQCCYLPR---


HsDis3L2     21 ---------------------------------------------------------
MmDis3L2     21 ---------------------------------------------------------
CeDis3L2     15 ---------------------------------------------------------
AtDis3L2     49 GRDSSRSKAKDSTSSSKQQRPNTDELEAMRASNVAFNSMPPM----------------
SpDis3L2    128 ------------------------------------VSKL-----------------
DmDis3L2    188 -----------------------------------KNEILSQLEWSQTRSEAHP
SpDis3      156 -----------------------------------LD--DESANDRNDRAIRNA
ScDis3      161 -----------------------------------LP--NETINDRNDRAIRKT
HsDis3      136 -----------------------------------EQ--GENANDRNDRAIRVA
HsDis3L1    130 -----------------------------------ER--GESMEKWQTRSIYNA
```

```
                                                                                    A208E
                                                                                    F209A              D217P

HsDis3L2    21  ------------------------------------------------------------
MmDis3L2    21  ------------------------------------------------------------
Ceis3L2     15  ------------------------------------------------------------
AtDis3L2    91  -------------------------------------------RA---ESGY--------
SpDis3L2   132  ------------------------------------------------------------
DmDis3L2   207  QQVGYQGRHRH-----DSEVSATNNGSHVQQQPKLGRRQRNSKKN---DGSGSNGTQSQK
SpDis3     173  ASWFASHLASL----GIKIVLLTDDRENARLA------------A---EQGI--------
ScDis3     178  CQWYSEHLKPY----DINVVLVTNDRLNREAA-----------TKEVESNI--------
HsDis3     153  AKWYNEHLKKMSADNQLQVIFITNDRRNKEKA------------I---EEGI--------
HsDis3L1   147  AVWYYHHCQD-----RMPIVMVTEDEEAIQQY------------GS-ETEGV--------


HsDis3L2    21  -------------------------AVAGPHDIGAS-P-------GD-K-KSKNRSTR-
MmDis3L2    21  -------------------------SVVGPSAVGAS-P-------GD-K-KSKNKSMR-
CeDis3L2    15  -------------------------YVKPQN-------------------ANNLQNN-
AtDis3L2    97  -PRRSASPLLS----SPEVSKQLLSKSCPDPRACEQS-P-------GM-N-GELFQQIEG
SpDis3L2   132  -------------------SKNLESRNNRDENSAKRE-KNNSHQVEAD-T-NNATEMVSS
DmDis3L2   259  TELEAMRSYIN----KI-------------------VQ-------RH-VGMDHQLSE-D
SpDis3     206  -QVSTLKDYVQ----YLPDSEILL-----------DMVS-------AIADAIASKEQV-E
ScDis3     214  -ITKSLVQYTE----LLPNADDIR-----------DSIP-------QM-DSFDKDLER-D
HsDis3     190  -PAFTCEEYVK----SLTANPELI-----------DRLA-------CL---SEEGNEI-E
HsDis3L1   181  -FVITFKNYLDNFWPDLKAAHELC-----------DSIL-------QS-R-RERENES-Q


HsDis3L2    44  GKKKSIFETYMSKEDVSEGLKRGTLIQGVLRINPKKF-HEAFIPSPDG--------DRDI
MmDis3L2    44  GKKKSIFETYMSKEDVSEGLKRGTLIQGVLRINPKKF-HEAFIPSPDG--------DRDI
CeDis3L2    28  FNPRKIFTEYISKEETDAGIEDGSMFKGVLRINPKNY-QFCFLDHPKG-----T-NHPDV
AtDis3L2   142  SSQRKIFSSHWSLDAVTEALEKGEAFKALFRVNAHNR-NFAYCKIDGV--------PTDI
SpDis3L2   170  NAKKSVYPLYYDSATVKKGLKSGTIFKGTLRILE-NH-RSAFACME-D--------IPDF
DmDis3L2   286  LIFKGNFSDLEAHAQRLVAAGYGRIVEEEIRVNRKNN-RQAFIIMSTD----REALERDG
SpDis3     242  SGTKNVYELHWSMSRLLACIKNGEVHKGLINISTYNY-LEGSVVVPG---------YNKPV
ScDis3     249  TFSDFTFPEYYSTARVMGGLKNGVLYQGNIQISEYNF-LEGSVSLPR---------FSKPV
HsDis3     223  -SGKIIFSEHLPLSKLQQGIKSGTYLQGTFRASRENY-LEATVWIHGD----NE-ENKEI
HsDis3L1   219  ESHGKEYPEHLPLEVLEAGIKSGRYIQGILNVNKHRAQIEAFVRLQGASSKDSD-LVSDI
                                                                       CSD1


HsDis3L2    95  FIDGVVARNRALNGDLVVVKL-LPEEHWKVVKPESNDKETEAAYESDIPEELCGHHLPQQ
MmDis3L2    95  FIDGVVARNRALNGDLVVVKL-LPEDQWKAVKPESNDKEIEATYEADIPEEGCGHHPLQQ
CeDis3L2    81  LVLGQ-DRNRAMOGDVVAVKI-KPKEDWLVNYVEYVKWWAEH-KKG---D----------
AtDis3L2   193  LINGNVCQSRAVEGDTVVIKL-DPLSLWPKMKGFVTESAAKP--EG------TNSPPEKD
SpDis3L2   219  YVDGPIARNRAFHNDVVIVEP-VMNNDSPTEKSNFLQNGVEKVKIKDHDDELG-------
DmDis3L2   341  IVLLPVARRYAFDGDKVRAFVLNPGAQGSSKTAEPSSGE-----------ISGGKPS--
SpDis3     293  LVSGRENLNRAVQGDIVCIQI-LPQDQWKTEAEEIADD---------------------
ScDis3     300  LIVGQKNLNRAFNGDQVIVEL-LPQSEWKAPSSIVLDSE-----------------H--
HsDis3     276  ILQGLKHLNRAVHEDIVAVEL-LPKSQWVAPSSVVLHDE--------------------
HsDis3L1   278  LIHGMKARNRSIHGDVVVVEL-LPKNEWKGRTVALCEND--------------------
                                                                       CSD1
```

```
HsDis3L2   154  SLKSY-NDSPDVIVEAQFDGSDSEDGHGITQNVLVDGVKKLSVCVSEKGREDGDAPVTKD
MmDis3L2   154  SRKGW--SGPDVIIEAQFDDSDSEDRHGNT-SGLVDGVKKLSISTPDRGKEDSSTPVMKD
CeDis3L2   125  --------------------------------------------RNSGKTDNNSPN-KT
AtDis3L2   244  DKKARQKNGIDVV-----------------EGFEDGFSKNKSSVIGKGAKNGVTPSSPP
SpDis3L2   271  -------------------------------GAMEHLERL------------------
DmDis3L2   387  --LSL-ADGEELS-----DDTESQG--------------------------------
SpDis3     331  ---------EDVV--------VSTA---------------------------------
ScDis3     339  --FDV-NDNPDIE--AG-DDDDNNESSSNT-T----------------------VI--
HsDis3     314  ---------GQN--------EEDVE---------------------------------
HsDis3L1   316  ---------CDDK---------ASG---------------------------------
                                            CSD1


HsDis3L2   213  -ETTCISQ------------D---TR--AL--SEKSLQRSAKVVYILEKKHSR-AATGFL
MmDis3L2   211  -ENTPIPQ------------D---TR--GL--SEKSLQKSAKVVYILEKKHSR-AATGIL
CeDis3L2   139  -EKRCLRNEIQ--------DNGVTSD-EV--PDSCLITIGAIVHILEKKHFR-VAAGKL
AtDis3L2   286  SLDSCLGSFCEQKGNCSAVDK---LC--GILSSFPHKRPTGQVVAVVEKSLVRDSIVGLL
SpDis3L2   280  -----EI------------K---SV--AS--FKGDSRTRARVVAIEKRAEIS-KIVGIL
DmDis3L2   404  -----------SESDT---D----NV--VV--ISSDNCPKAFVIAITKRTELR-QIVGTI
SpDis3     339  -----------AEPDSARIND---LE--LI--TKRNAHPTAKVVGILRN-WR-PYVGHV
ScDis3     366  -----------SDKQRRLLAK---DAMIAQ--RSKKLQPTAKVVYIQRRS-WR-QYVGQL
HsDis3     322  -----------KEEETERML----KT--AV--SEKMLKPTGRVVGIKRN-WR-PYCGML
HsDis3L1   323  ------------------------E--SPSEPMPTGRVVGILQKN-WR-DYVVTF
                                     CSD1                      CSD2


HsDis3L2   252  KL-------------LAD-KNSELFRKYALFSPSDHRVPRIYVPLKDCPQD----FVARP
MmDis3L2   250  KL-------------LAD-KNSDLFKKYALFSPSDHRVPRIYVPLKDCPQD---FMTRP
CeDis3L2   185  QL-------------M---PN--SANPNVLFVATDSRVPRILIPKSDVDKE----FFSRP
AtDis3L2   341  DVKGWIHYKESDPKRCKS-PLSLSDDEYVQLMPADPRFPKLIVPFHVLPGS----IRARL
SpDis3L2   314  RAPGWS--------LKNV-EYVSKKSSYAIEIPKDKRLPFITIHKNDLSDLSGENWIENI
DmDis3L2   441  SF-------------TNP-TKLCDDQLEYKERPYDLRVPMVYVPKDACAAHIGNK---QQ
SpDis3     379  DN-------------ATIAQSKGGSQQTVLLTPMDRRVPKIRFRTRQAPR----------
ScDis3     408  AP-------------SSV-DPQSSSTQNVFVILMDKCLPKVRIRTRRAAE----------
HsDis3     361  SK-------------SD-----IKESRRHLFTPADKRIPRIRIETRQAST----------
HsDis3L1   350  PS-------------KEEVQSQGKNAQKILVTPWDYRIPKIRISTQQAET----------
                                        CSD2


                       L371R              L390F
HsDis3L2   294  ----KDYANTLFICRIVDWKEDCNEALGQIAKSLGQAGEIEPETEGILTEYGVD-FSDFS
MmDis3L2   292  ----KDFANTLFICRIIDWKEDCNEALGQIAKSLGQAGEIEPETEGILTEYGVD-FSDFS
CeDis3L2   223  ----KDFERFLYTAKITDWRAESVYADGRLVKLLGMSGEIDTETERIVYEHQID-HREFS
AtDis3L2   396  ENLDPNLEAELVAAQIVDWGEGSPFPVAQITELFGRGSELEPQINAILYQNSVC-DSDFS
SpDis3L2   365  ----LKHHDQLFSVEITRWSIYSRYPMGVLGEKLGNITDVEAYTNALLENGIS-SSPFS
DmDis3L2   484  ----IDVSGLLYLAHILETD-CNGHCIAELIQPVGRVGNLDDELKAILFHNGLRDIKPFE
SpDis3     416  ------LVGRRIVVAIDLWDASSRYPEGHFVRDLGEMETKEATEALLEYDVQ-HRPFP
ScDis3     444  ------LLDKRIVISIDSWPTTHKYPLGHFVRDLGTIESAQAETEALLEHDVE-YRPFS
HsDis3     393  ------LEGRRIIVAIDGWPRNSRYPNGHFVRNLGDVGEKETETEVLLEHDVP-HQPFS
HsDis3L1   387  ------LQDFRVVVRIDSWESTSVYPNGHFVRVLGRIGDLEGEIATILVENSIS-VIPFS
                                    CSD2                      RNB
```

79

T456G A457C R458T L460I/V D461N

```
HsDis3L2  349  SEVLECLPQG---LPWTIPPE----------EFSKRRDLRKD-CIFTIDPSTARDIDDA
MmDis3L2  347  SEVLECLPQS---LPWTIPPD----------EVGKRRDLRKD-CIFTIDPSTARDIDDA
CeDis3L2  278  DECLESLPIT-TAENWKVPDA----------EFEYRRDFRSD-IVFTIDPKTARDIDDA
AtDis3L2  455  PGSLTSLPR----VPWEVPEE----------EVQRRKDLRDL-CVLTIDPSTATDIDDA
SpDis3L2  420  DEVINCLPP----DDWIISHE----------EIKKRRDLRNE-LIITIDPETARDIDDA
DmDis3L2  539  QRFIDIYSQ----PPPPISQE----------DLRQRKDLRKM-CIFTIDPMTARDIDDA
SpDis3    469  KAVLDCLPEE--GHNWKVPAD------KT-HPLWKNRKDFRDK-LICSIDPPGCQDIDDA
ScDis3    497  KKVLECLPAE--GHDWKAPTKLDDPEAVSKDPLLTKRKDLRDK-LICSIDPPGCVDIDDA
HsDis3    446  QAVLSFLPK----MPWSITEK----------DMKNREDLRHL-CICSVDPPGCTDIDDA
HsDis3L1  440  EAQMCEMPVNTPESPWKVSPE----------EEQKRKDLRKSHLVFSIDPKGCEDVDDT
```

RNB

P512D

```
HsDis3L2  394  LSCKPLADG------NFKVGVHIADVSYFVPEGSDLDKVAAERATSVYLVQKVVPMLPRL
MmDis3L2  392  LACRRLTDG------TFEVGVHIADVSYFVPEGSSLDKVAAERATSVYLVQKVVPMLPRL
CeDis3L2  325  LHAKHIDDCDGKGTPGLEIGVHIADVTFFLKEGTELDKWASERGNSTYLSQTVIPMLPRI
AtDis3L2  499  LSVQSLPGG------FFRVGVHIADVSYFVLPETALDTEARFRSTSVYLMQRKISMLPPL
SpDis3L2  464  VSCRALDNG------TYEVGVHIADVTHFVKPDSALDKEAASRATTVYLVQKAIPMLPPL
DmDis3L2  583  VSIEKLGDN------EYEIGVHISDVSHFLIEDNELDNIVKERSTSIYLANEVIHMLPQS
SpDis3    519  LHACVLPNG------NYEVGVHIADVTHFVKPNISMDSEAASRGTTVYLVDKRIDMLPML
ScDis3    554  LHAKKLPNG------NWEVGVHIADVTHFVKPGTALDAEGAARGTSVYLVDKRIDMLPML
HsDis3    490  LHCRELENG------NLEVGVHIADVSHFIRPGNALDQESARRGTTVYLCEKRIDVVPEL
HsDis3L1  489  LSVRTLNNG------NLELGVHIADVTHFVAPNSYIDIEARTRATTYYLADRRYDMLPSV
```

RNB

C560Y

```
HsDis3L2  448  LCEELCSLNPMSDKLTFSVIWTLTPEG-K-ILD--EWFGRTIIRSCTKLSYEHAQSMIES
MmDis3L2  446  LCEELCSLNPMTDKLTFSVIWKLTPEG-K-ILE--EWFGRTIIRSCTKLSYDHAQSMIEN
CeDis3L2  385  LCEQLCSLNPGVDRLSFSTVFKMSYEA-E-LYD--VWFGRSVIRSRVKLAYEHAQDFIEN
AtDis3L2  553  LSENVGSLSPGADRLAFSILWDLNREG-D-VID--RWIGRTIIRSCCKLSYDHAQDIIDG
SpDis3L2  518  LCERLCSLNPNVERLAFSVFWKLDSNG-KEIGK--RWFGKTVIKICARLAYSEAQGVIEG
DmDis3L2  637  LCMR-CSLLPGQDKFAFSVFWRMNGKG-V-MLQKKPEFCRTVINSCSQFAYEHAQKIIDN
SpDis3    573  LGTDLCSLRPYVERFAFSCIWEMDENA-N-IIK--VHFTKSVIASKEAFSYADAQARIDD
ScDis3    608  LGTDLCSLKPYVDRFAFSVIWELDDSA-N-IVN--VNFMKSVIRSREAFSYEQAQLRIDD
HsDis3    544  LSSNLCSLKCDVDRLAFSCIWEMNHNA-E-ILK--TKFTKSVINSKASLTYAEAQLRIDS
HsDis3L1  543  LSADLCSLLGGVDRYAVSIMWELDKASYE-IKK--VWYGRTIIRSAYKLFYEAAQELLDG
```

RNB

```
HsDis3L2  504  PTEK---IPA-KELPPISPEHSSEEVHQAVINLHGIAKQLRQQRFVDGALRLDQLKLAFT
MmDis3L2  502  PTEK---IPE-EELPPISPEHSVEEVHQAVINLHSIAKQLRRQRFVDGALRLDQLKLAFT
CeDis3L2  441  PEKD---FTC-DELPDISDGNTPFEIKEKTLMLHRIAQVLRQKREDSGALRIELPRLKFA
AtDis3L2  609  KSD----VAE-NGWPALHGSFKWCDVTRSVKQLSEISTTLRQKRFRNGALQLENSKPVFL
SpDis3L2  575  KSWD---DA---VGKPIGGTHTPKDVETSILTLCEISRKLRKDRFAKGAVEINSTELKFQ
DmDis3L2  694  PNER---FTE-NDFPTILNGFNPDDIRNRVLWLHDIASSIRKTRLDNGALTINNAKLRFL
SpDis3    629  QK-------------------MQDPLTQGMRVLLKLSKILKQKRMDEGALNLASPEVRIQ
ScDis3    664  KT-------------------QNDELTMGMRALLKLSVKLKQKRLEAGALNLASPEVKVH
HsDis3    600  AN-------------------MNDDITTSLRGLNKLAKILKKRRIEKGALTLSSPEVRFH
HsDis3L1  600  NLSVVDDIPEFKDLDEKSRQAKLEEIVWAIGKLTIARHVRAKRDGCGALELEVECVQ
```

RNB

HsDis3L2 560 LDHETGLPQGCHIYEYRESNKLVEEFMLLANMAVAHKIHRAFPEQALLRRHPPPQTRMLS
MmDis3L2 558 LDHETGLPQGCHIYEYRDSNKLVEEFMLLANMAVAHKIFRTFPEQALLRRHPPPQTKMLS
CeDis3L2 497 LD-EDKKPQGVSIYEIKDSNKLVEEFMLLANMEVAKKIAENFPEHALLRNHPPPKEKMIK
AtDis3L2 664 FD-EHGVPYDFVTCSRKGSNFLVEEFMLLANMTAAEVISQAYRASSLLRRHPEPNTRKLK
SpDis3L2 629 LD-EYGMPNKCEVYEQTDANHLIEEFMLLANRSVAEHISKNESNNSLLRRHASPKEKQIN
DmDis3L2 750 LDPITGEPLSFEVEKQREANRLIEEFMLLANQAVARFIHDSFPDIAVLRNHPPPLIKSLK
SpDis3   670 TDNETSDPMDVEIKQLLETNSLVEEFMLLANISVAQKIYDAFPQTAVLRRHAAPPLTNFD
ScDis3   705 MDSETSDPNEVEIKKLLATNSLVEEFMLLANISVARKIYDAFPQTAMLRRHAAPPSTNFE
HsDis3   641 MDSETHDPIDLQTKELRETNSMVEEFMLLANISVAKKIHEEFSEHALLRKHPAPPPSNYE
HsDis3L1 660 LDDK-KNIHDLIPKQPLEVHETVAECMILANHWVAKKIWESFPHQALLRQHPPPHQEFFS

RNB

HsDis3L2 620 DLVEFCD-QMGLPVDFSSAGALNKSITQTFGDDKYSLARKEVLTNMCSRPMQMALYFCSG
MmDis3L2 618 DLVEFCD-QMGLPMDVSSAGALNKSITKTFGDDKYSLARKEVLTNMYSRPMQMALYFCSG
CeDis3L2 556 DVAEQCA-RIGFPLDGRTSGLLSTSIRKYQGKSRLDMCIRQVISSLTIKPMQQAKYFCTF
AtDis3L2 723 EFEGFCS-KHGMDLDISSSGQLQDSLEKITGNLKDDSVFVDILNNYAIKPMQLASYFCTG
SpDis3L2 688 EFCHFLK-SMNFDFDASSSAAFNASMVRLRS--TFNEELVELFENMAVRSLNRAEYFCTG
DmDis3L2 810 ALREKLL-ALGFELDYSSSKALQESMVRLCNEAPNPVAMNACLSQLLMKPMARATYFCSE
SpDis3   730 SLQDILRVCKGMHLKCDTSKSLAKSLDECV--DPKEPYFNTLLRILTTRCMLSAEYFCSG
ScDis3   765 ILNEMLNTRKNMSISLESSKALADSLDRCV--DPEDPYFNTLVRIMSTRCMMAAQYFYSG
HsDis3   701 ILVKAAR-SRNLEIKTDTAKSLAESLDQAE--SPTFPYLNTLLRILATRCMMQAVYFCSG
HsDis3L1 719 ELRECAK-AKGFFIDTRSNKTLADSLDNAN--DPHDPIVNRLLRSMATQAMSNALYFSTG

RNB

HsDis3L2 679 LLQD-PAQFRHYALNVPLYTHFTSPIRRFADVLVHRLLAAALGYRERLDM----------
MmDis3L2 677 MLQD-QEQFRHYALNVPLYTHFTSPIRRFADVLVHRLLAAALGYSEQPDV----------
CeDis3L2 615 EMP--LSFYHHFALNVDHYTHFTSPIRRYPDVLVHRQLAAALGYNERSER----------
AtDis3L2 782 NLKDSVAEWGHYALAVPLYTHFTSPLRRYPDIVVHRALAAALEAEELYSKQKQTAIDEGR
SpDis3L2 745 DFGE-KTDWHHYALSFNHYTHFTSPIRRYPDIIVHRLLERSLKNTSP-GI----------
DmDis3L2 869 GKSE-PADLWHYALSIPIYTHFTSPIRRYPDILVHRLLAAALKYCTPPK-----------
SpDis3   788 TFA--PPDFRHYGLASPIYTHFTSPIRRYADVLAHRQLAAAIDYETINP-----------
ScDis3   823 AYS--YPDFRHYGLAVDIYTHFTSPIRRYCDVVAHRQLAGAIGYEPLSL-----------
HsDis3   758 M----DNDFHHYGLASPIYTHFTSPIRRYADVIVHRLLAVAIGADCTYP-----------
HsDis3L1 776 SCA--EEEFHHYGLALDKYTHFTSPIRRYSDIVVHRLLMAAISKDKKME----------

RNB

A801C C804I K808H

HsDis3L2 728 -------------------------------APDTLQKQADHCNDRRMASKRVQELSTS
MmDis3L2 726 -------------------------------EPDTLQKQADHCNDRRMASKRVQELSIG
CeDis3L2 663 -----------------------------VPEEIQEICTRCNDTKLASKEASDESAM
AtDis3L2 842 SCFTGIHFNKDAAESIEGKEALSVAALKHGVPSTEILSDVAAYCNERKLAARKVRDACDK
SpDis3L2 793 -------------------------------DKKNCSLVAAHCNEKKEKSTTVQEDSQQ
DmDis3L2 917 ------------------------------RTPDDLHTLTKIANERKYNAKKAGEDSGN
SpDis3   835 ---------------------------SLSDKSRLIEICNGINYRHRMAQMAGRASIE
ScDis3   870 ---------------------------THRDKNKMDMICRNINRKHRNAQFAGRASIE
HsDis3   803 ---------------------------ELTDKHKLADICKNLNFRHKMAQYAQRASVA
HsDis3L1 823 --------------------------IKGNLFSNKDLEELCRHINNRQAAQHSQKQSTE

RNB

R865L/T

```
HsDis3L2  756  LFFAVLVKE------SGPLES----EAMVMGILKQAFDVLVLRYGVQKRIYCNALALR-
MmDis3L2  754  LFFAVLVKE------SGPLES----EAMVMGVLNQAFDVLVLRFGVQKRIYCNALALR-
CeDis3L2  691  LYFGVFIHQ------TGPMKC----QAVVIGVMDLSFDVLIVEYGVKRVYVDKMKRD-
AtDis3L2  902  LYTWFVLKQ-----KEIFPC----EARVMNLGSRFMTVYISKLGIERRIYYDQIEGL-
SpDis3L2  821  LLLSVYIAEYCKKHDKKSMPV-----QAFATRISGNSIDVYISEYGISNRVDLSSDDRI-
DmDis3L2  946  LYFKRYVHN------KQGIYM-----RAVVIEIFQHMMNVVTLESGHVISINYKMQKVLV
SpDis3    866  YYVGQALKG-------GVAEE-----DAYVLKVFKNGFVVFIARFGLEGIVYTKSLSSV-
ScDis3    901  YYVGQVMRN------NESTE-----TGYVLKVFNNGIVVLVPKFGVEGLIRLDNLTED-
HsDis3    834  FHTQLFFKS------KGIVSE-----EAYILFVRKNAIVVLIPKYGLEGTVFFEEKDK--
HsDis3L1  857  LFQCMYFKD------KDPATEERCISDGVIYSIRTNGVLLFIPRFGIKGAAYLKNKDGLV
```

RNB   S1


```
HsDis3L2  804  --------------SH-HFQKVGK----------------KPELTLVWEP---------
MmDis3L2  802  --------------SY-SFQKVGK----------------KPELTLVWEP---------
CeDis3L2  739  -----------------FNKST-------------------EKLTIYWPA---------
AtDis3L2  950  -------------CA-DWLEATSTL-IVDKLYSKRGGRGFFKPMKEAVYLVSPCEVCVAK
SpDis3L2  875  -------------KSFIVAPDDS----------------SVKITLF------------
DmDis3L2  995  DTHG-------VPNYI-----------LIAE----------------R---------
SpDis3    913  ----------LEPNVE------YVEDEYKLNIE----------------I---------
ScDis3    948  -----------PNSAA-----FDEVEYKLTFV----------------P---------
HsDis3    881  -----------PNPQ------LIYDD-EI----------------P---------
HsDis3L1  911  -ISCGPDSCSEWKPGSL------QR-FQNKITST----------------T---------
```

S1


```
HsDis3L2  823  ------------------EDM-EQ---EPAQQ--VITIFSLVEVVLQAEST---ALKYSA-
MmDis3L2  821  ------------------DDL-EE---EPTQQ--VITIFSLVDVVLQAEAT---ALKYSA-
CeDis3L2  753  ----------------DPNAESGNREEFSS--SIQMCNVVYILVPY-K---SIEVSA-
AtDis3L2  995  CSALSVHDTESPEAVSIDEV-AP--AVFPL--TIQLFSTIPVVLHAVGGDDGPLDIGAR
SpDis3L2  892  ------------------------DDSQK--TIALTDRFQVYLYSDYS---RTFFSI-
DmDis3L2  1009 --------------------NL---KQSPR--KLQLLSVVPICLIIWDQKLTG--FL-K
SpDis3    931  --------------------RD--QPKPQTVQLQMFQQVRVRVTTVRDEHSG--KQ-K
ScDis3    965  --------------------TN---SDKPR--DVYVFDKVEVQVRSVMDPITS--KR-K
HsDis3    893  --------------------SL---KIEDT--VFHVFDKVKVKIMLDSSNLQ---HQ-K
HsDis3L1  938  --------------------T--DGESV--TFHLFDHVTVRISIQASRCH---SD-T
```

S1


```
HsDis3L2  856  ---ILKRPG---TQ------------------------------------------
MmDis3L2  854  ---ILKRPG---LE------------------------------------------
CeDis3L2  789  ---TIVRPS---LE------------------------------------------
AtDis3L2  1049 LYMSSYY------------------------------------------
SpDis3L2  920  ------RCS---LV------------------------------------------
DmDis3L2  1040 MEEQ------------------------------------------
SpDis3    964  VQITLV------------------------------------------
ScDis3    996  AELLL------------------------------------------
HsDis3    923  IRMSLVEPQ---IP------------------------------------------
HsDis3L1  967  IRLEIISNKPYKIPNTELIHQSSPLLKSELVKEVTKSVEEAQLAQEVKVNIIQEEYQEYR
```

S1

```
HsDis3L2    864  ---GHLGPEKEEEES-----DGEPED------SSTS---
MmDis3L2    862  ------------KAS-----DEEPED------------
CeDis3L2    797  ------------QRNILKSTLKDMKE------TGSTILQ
AtDis3L2         ---------------------------------------
SpDis3L2    925  -----------SL----------------------N
DmDis3L2   1044  -----------------------------------I
SpDis3      970  -----------------------------------Y
ScDis3     1001  -----------------------------------K
HsDis3      934  ---GISI------PT-----DTSNMDLNGPKKKKMKLGK
HsDis3L1   1027  QTKGRSLYTLLEEIR-----DL------ALLDVSNNYGI
                                    S1
```

Figure S.3.1. Full-length sequence alignment of several eukaryotic homologues from the RNase II/RNB family of exoribonucleases, highlighting mutated residues. Multiple sequence alignment was built using the default T-Coffee and the amino acid sequences of ten homologues of the RNase II/RNB family: HsDis3L2 (*Homo sapiens* Dis3L2, NP_689596.4); MmDis3L2 (*Mus musculus* Dis3L2, NP_705758.1); CeDis3L2 (*Caenorhabditis elegans* Dis3L2, NP_498160.2); AtDis3L2 (*Arabidopsis thaliana* Dis3L2, also designated SOV, NP_177891.1); SpDis3L2 (*Schizosaccharomyces pombe* Dis3L2, NP_594510.1); DmDis3L2 (*Drosophila melanogaster* Dis3L2, NP_728490.1); SpDis3 (*Schizosaccharomyces pombe* Dis3, NP_596653.1); ScDis3 (*Saccharomyces cerevisiae* Dis3, also named Rrp44, NP_014621.1); HsDis3 (*Homo sapiens* Dis3, NP_055768.3); HsDis3L1 (*Homo sapiens* Dis3L1, NP_001137160.1). This final figure was constructed using BoxShade v3.2.1, where darker highlights on the residues indicate higher amino acid conservation in each position. The coloured boxes under the alignment indicate in which protein domain each region is located [colour-coded as in Figures 1.2. (section 1.4.2.2.) and 1.3. (section 1.4.4.)]; respective domain designations are inside each box. The positions outlined in red were modified in SpDis3L2, and the respective substitution is referred above each one of them. The two dark green boxes in the RNB domain are pinpointing the four aspartic acid residues (D) of the active site of the enzymes, which are conserved in the RNB domain, and thus essential for its exonucleolytic function. The second dark green box encompasses the key residue that distinguishes the amino acid signature of the three paralogues: DLDD for Dis3L2 orthologues, DIDD in Dis3 proteins, and DVDD in Dis3L1 enzymes. That position is outlined in light green and was also modified in SpDis3L2 to obtain the inactive mutant (D461N).
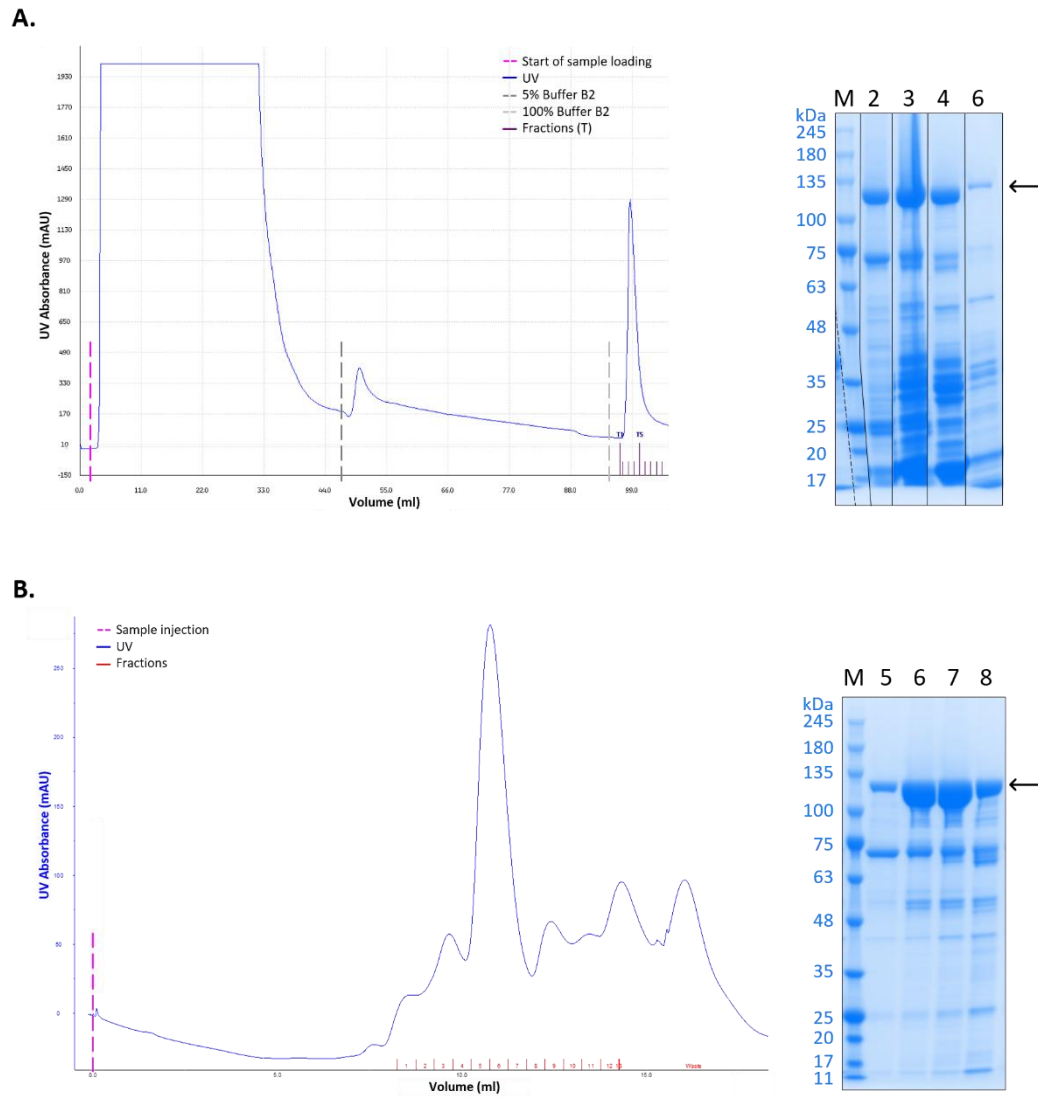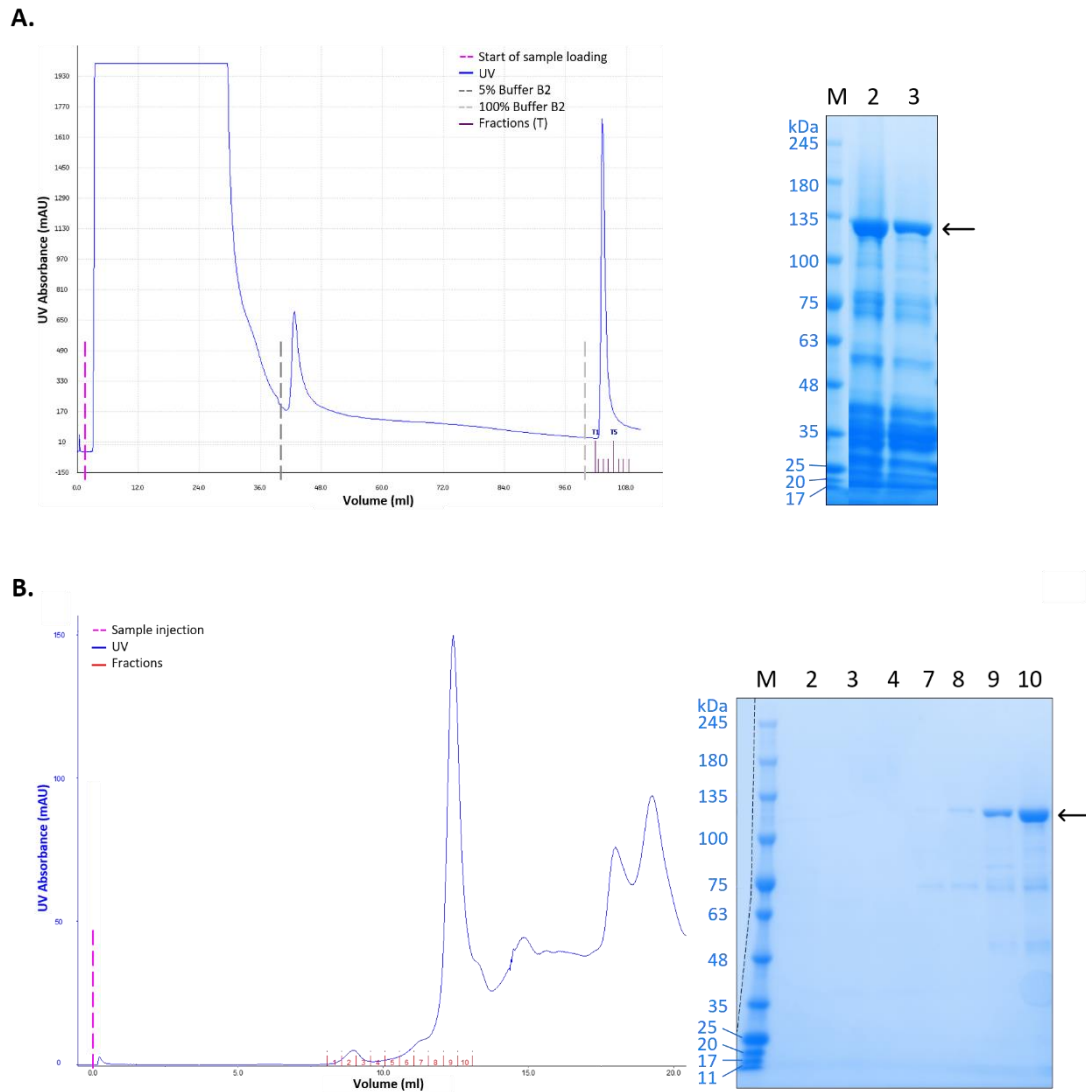
Figure S.3.2. Representative results of the T456G mutant purification. (**A.**) **Left panel** – Protein elution profile in the affinity chromatography purification step. **Right panel** – SDS-PAGE with collected fractions during this 1st purification step. (**B.**) **Left panel** – Chromatogram obtained in the size exclusion chromatography step. **Right panel** – SDS-PAGE with collected fractions during this 2nd purification step. Samples loaded in the Novex WedgeWell 8-16% Tris-Glycine Protein Gels are numbered on the top of respective lanes according to the chromatograms. 'M' indicates the lane with the molecular marker (NZYColour Protein Marker II, *NZYTech*), and corresponding band sizes are denoted next to them (in kDa). Arrows are pointing the bands corresponding to full-length SpDis3L2.
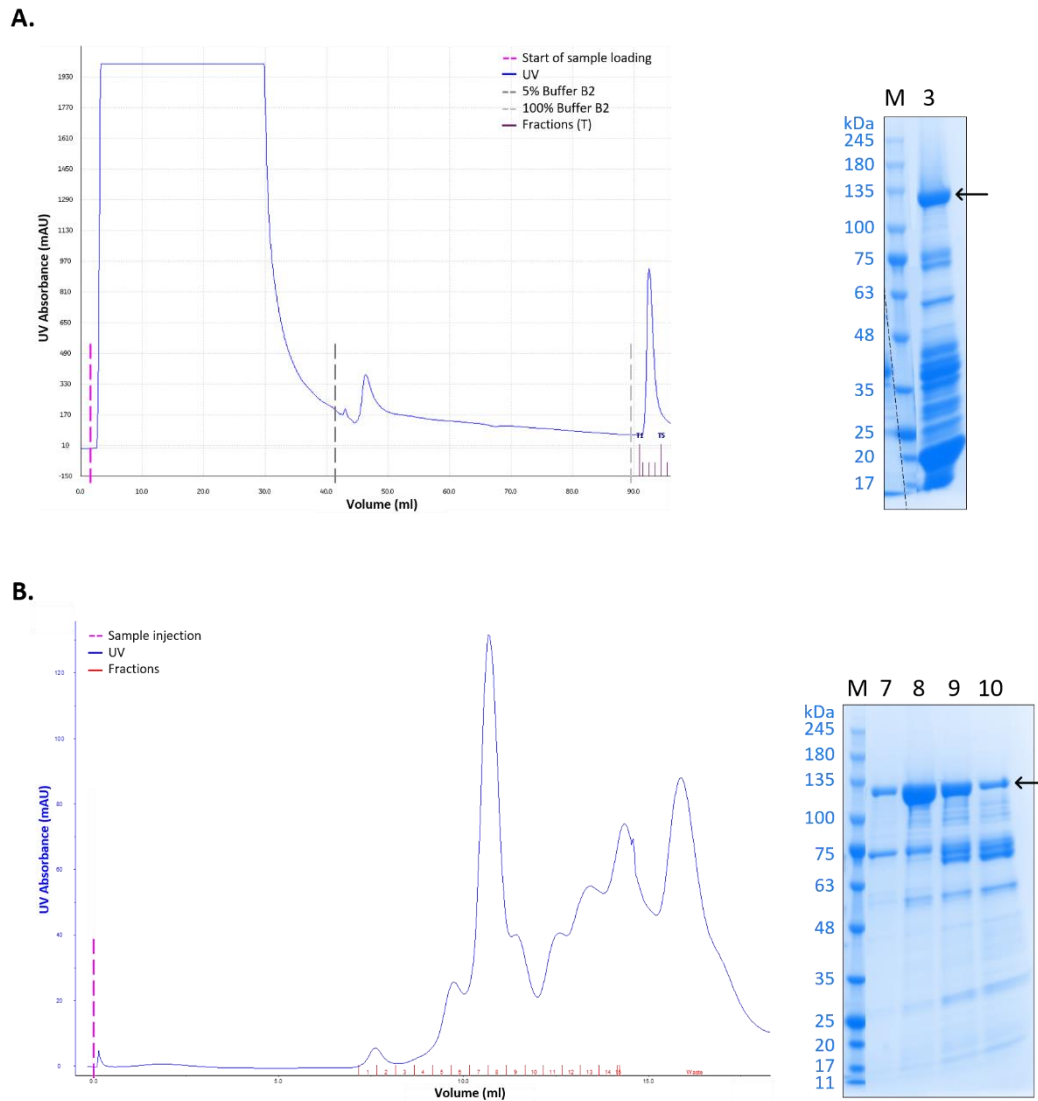
Figure S.3.3. Representative results of the A457C mutant purification. (A.) Left panel – Protein elution profile in the affinity chromatography purification step. **Right panel –** SDS-PAGE with collected fractions during this 1ˢᵗ purification step. **(B.) Left panel –** Chromatogram obtained in the size exclusion chromatography step. **Right panel –** SDS-PAGE with collected fractions during this 2ⁿᵈ purification step. Samples loaded in the Novex WedgeWell 8-16% Tris-Glycine Protein Gels are numbered on the top of respective lanes according to the chromatograms. 'M' indicates the lane with the molecular marker (NZYColour Protein Marker II, *NZYTech*), and corresponding band sizes are denoted next to them (in kDa). Arrows are pointing the bands corresponding to full-length SpDis3L2.
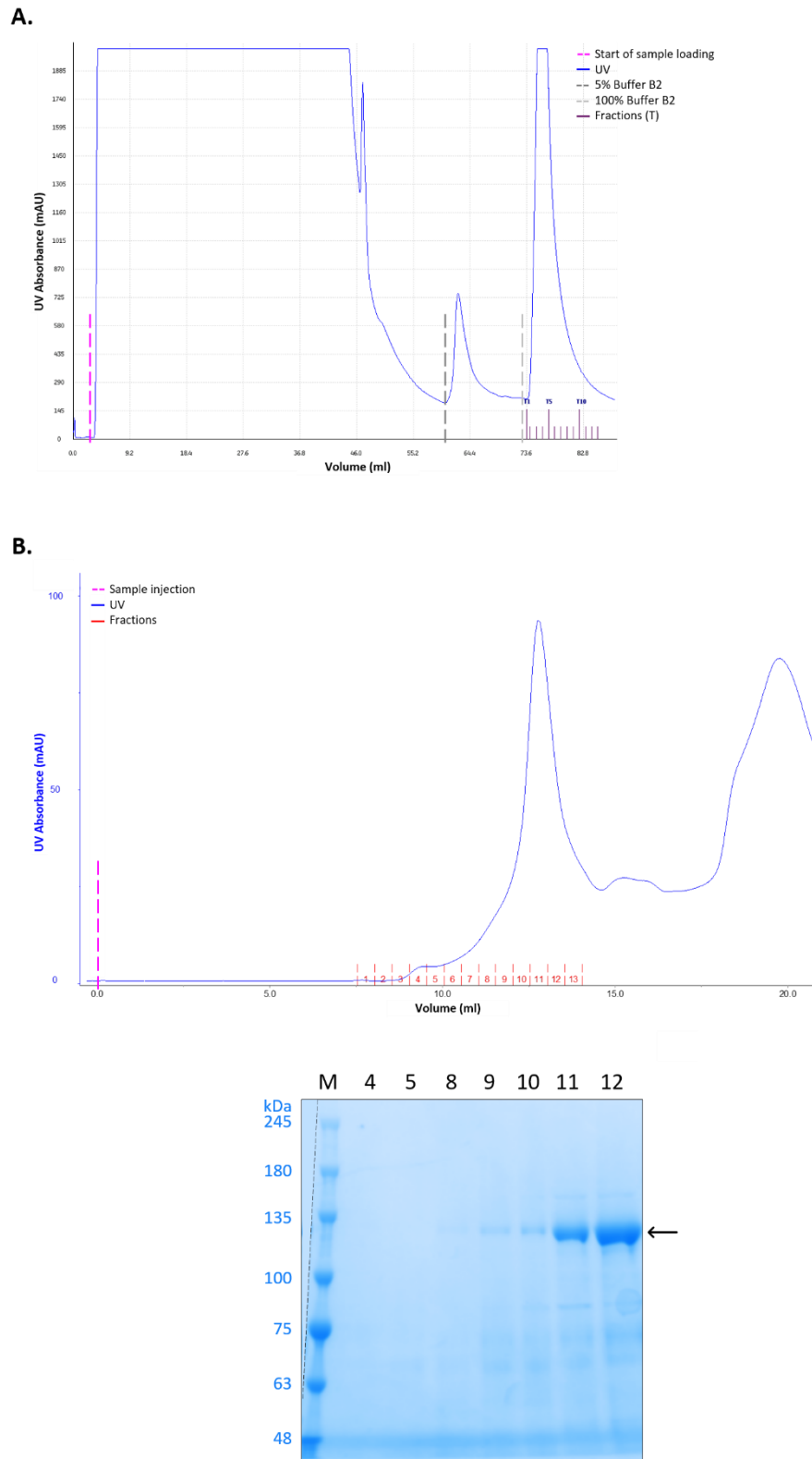
Figure S.3.4. Representative results of the L460I mutant purification. (A.) Left panel – Protein elution profile in the affinity chromatography purification step. **Right panel –** SDS-PAGE with collected fractions during this 1st purification step. **(B.) Left panel –** Chromatogram obtained in the size exclusion chromatography step. **Right panel –** SDS-PAGE with collected fractions during this 2nd purification step. Samples loaded in the Novex WedgeWell 8-16% Tris-Glycine Protein Gels are numbered on the top of respective lanes according to the chromatograms. 'M' indicates the lane with the molecular marker (NZYColour Protein Marker II, *NZYTech*), and corresponding band sizes are denoted next to them (in kDa). Arrows are pointing the bands corresponding to full-length SpDis3L2.
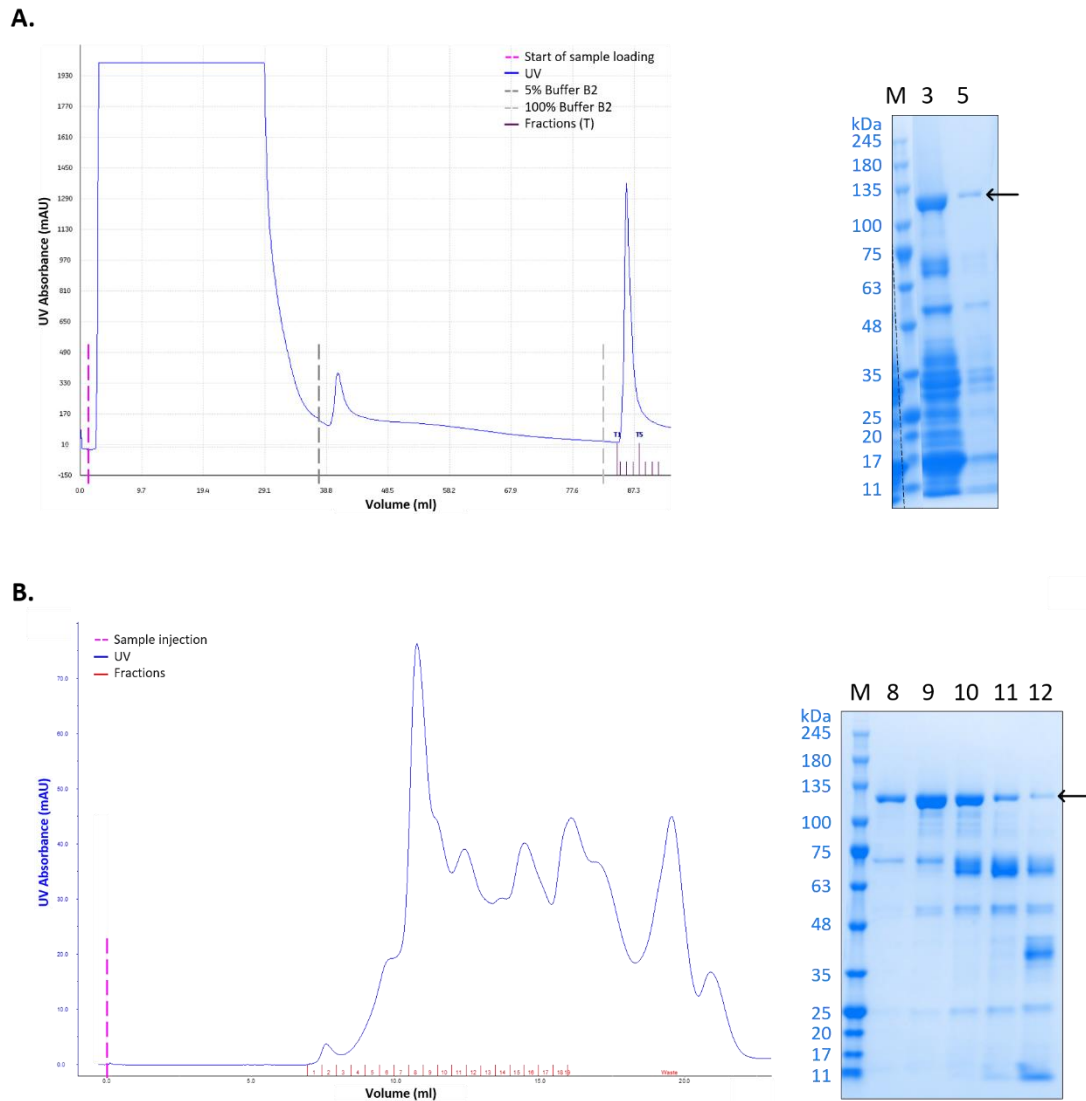
Figure S.3.5. Representative results of the D461N mutant purification. **(A.)** Protein elution profile in the affinity chromatography purification step. **(B.) Top panel –** Chromatogram obtained in the size exclusion chromatography step. **Bottom panel –** SDS-PAGE with collected fractions during this 2nd purification step. Samples loaded in the Novex WedgeWell 8-16% Tris-Glycine Protein Gels are numbered on the top of respective lanes according to the chromatograms. 'M' indicates the lane with the molecular marker (NZYColour Protein Marker II, *NZYTech*), and corresponding band sizes are denoted next to them (in kDa). Arrows are pointing the bands corresponding to full-length SpDis3L2.
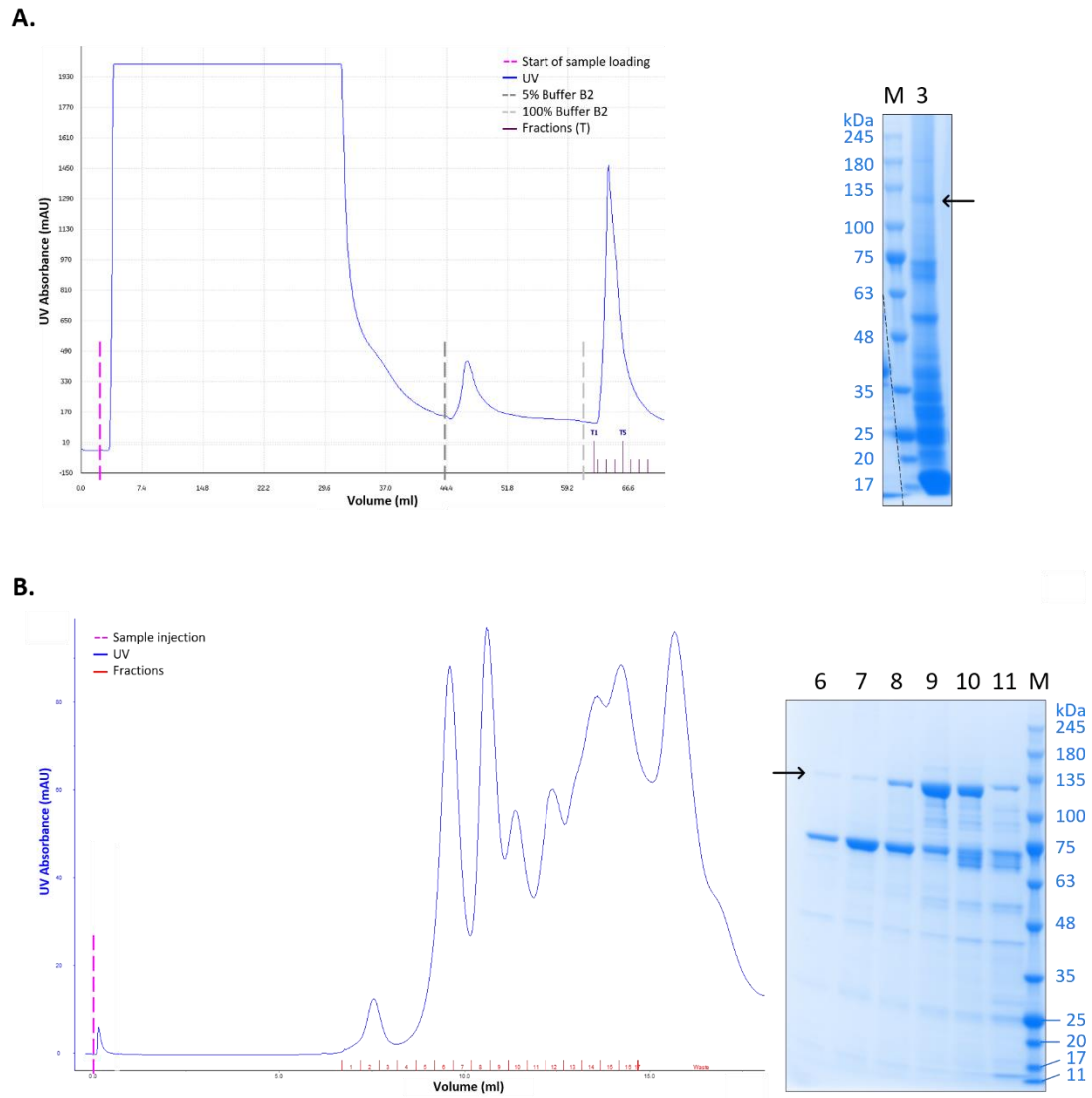
Figure S.3.6. Representative results of the P512D mutant purification. (A.) Left panel – Protein elution profile in the affinity chromatography purification step. **Right panel** – SDS-PAGE with collected fractions during this 1st purification step. **(B.) Left panel** – Chromatogram obtained in the size exclusion chromatography step. **Right panel** – SDS-PAGE with collected fractions during this 2nd purification step. Samples loaded in the Novex WedgeWell 8-16% Tris-Glycine Protein Gels are numbered on the top of respective lanes according to the chromatograms. 'M' indicates the lane with the molecular marker (NZYColour Protein Marker II, *NZYTech*), and corresponding band sizes are denoted next to them (in kDa). Arrows are pointing the bands corresponding to full-length SpDis3L2.

Figure S.3.7. Representative results of the C560Y mutant purification. **(A.) Left panel** – Protein elution profile in the affinity chromatography purification step. **Right panel** – SDS-PAGE with collected fractions during this 1st purification step. **(B.) Left panel** – Chromatogram obtained in the size exclusion chromatography step. **Right panel** – SDS-PAGE with collected fractions during this 2nd purification step. Samples loaded in the Novex WedgeWell 8-16% Tris-Glycine Protein Gels are numbered on the top of respective lanes according to the chromatograms. 'M' indicates the lane with the molecular marker (NZYColour Protein Marker II, *NZYTech*), and corresponding band sizes are denoted next to them (in kDa). Arrows are pointing the bands corresponding to full-length SpDis3L2.
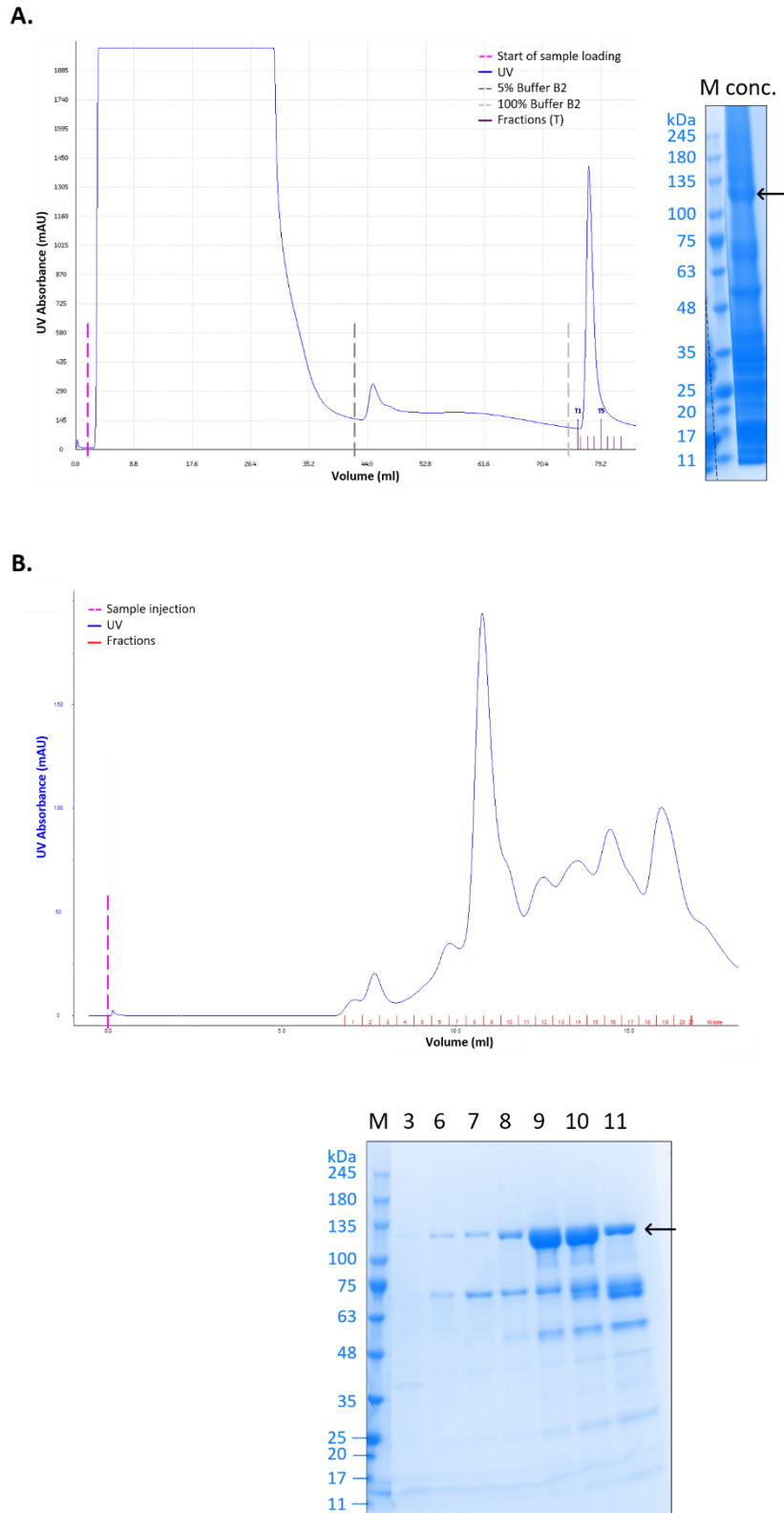
Figure S.3.8. Representative results of the K808H mutant purification. (A.) **Left panel** – Protein elution profile in the affinity chromatography purification step. **Right panel** – SDS-PAGE with a concentrated sample (cons.) obtained from the collected fractions in this 1st purification step. **(B.) Top panel** – Chromatogram obtained in the size exclusion chromatography step. **Bottom panel** – SDS-PAGE with collected fractions during this 2nd purification step. Samples loaded in the Novex WedgeWell 8-16% Tris-Glycine Protein Gels are numbered on the top of respective lanes according to the chromatograms. 'M' indicates the lane with the molecular marker (NZYColour Protein Marker II, *NZYTech*), and corresponding band sizes are denoted next to them (in kDa). Arrows are pointing the bands corresponding to full-length SpDis3L2.
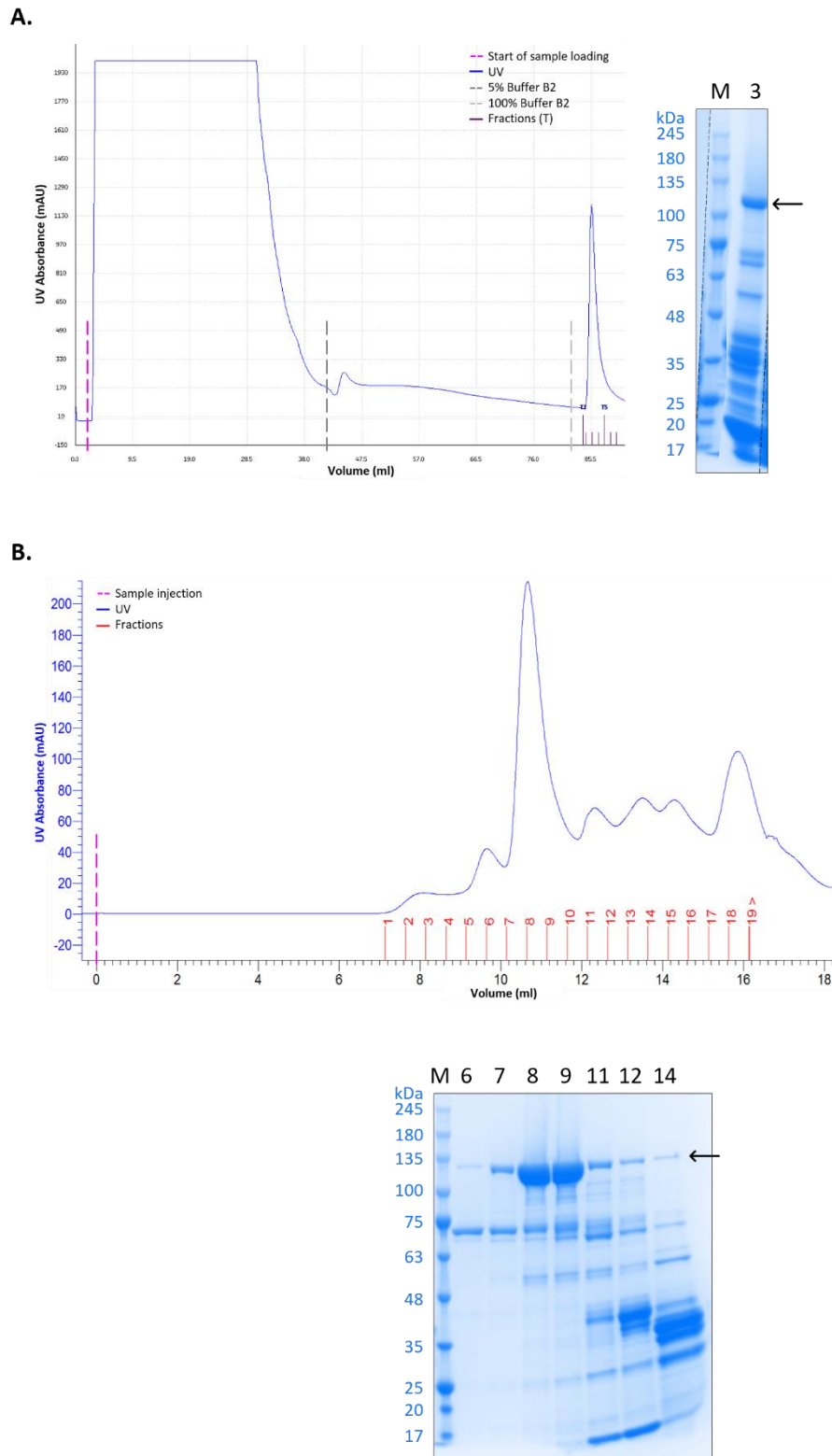
Figure S.3.9. Representative results of the R865L mutant purification. (A.) Left panel – Protein elution profile in the affinity chromatography purification step. **Right panel** – SDS-PAGE with collected fractions during this 1st purification step. **(B.) Top panel** – Chromatogram obtained in the size exclusion chromatography step. **Bottom panel** – SDS-PAGE with collected fractions during this 2nd purification step. Samples loaded in the Novex WedgeWell 8-16% Tris-Glycine Protein Gels are numbered on the top of respective lanes according to the chromatograms. 'M' indicates the lane with the molecular marker (NZYColour Protein Marker II, *NZYTech*), and corresponding band sizes are denoted next to them (in kDa). Arrows are pointing the bands corresponding to full-length SpDis3L2.
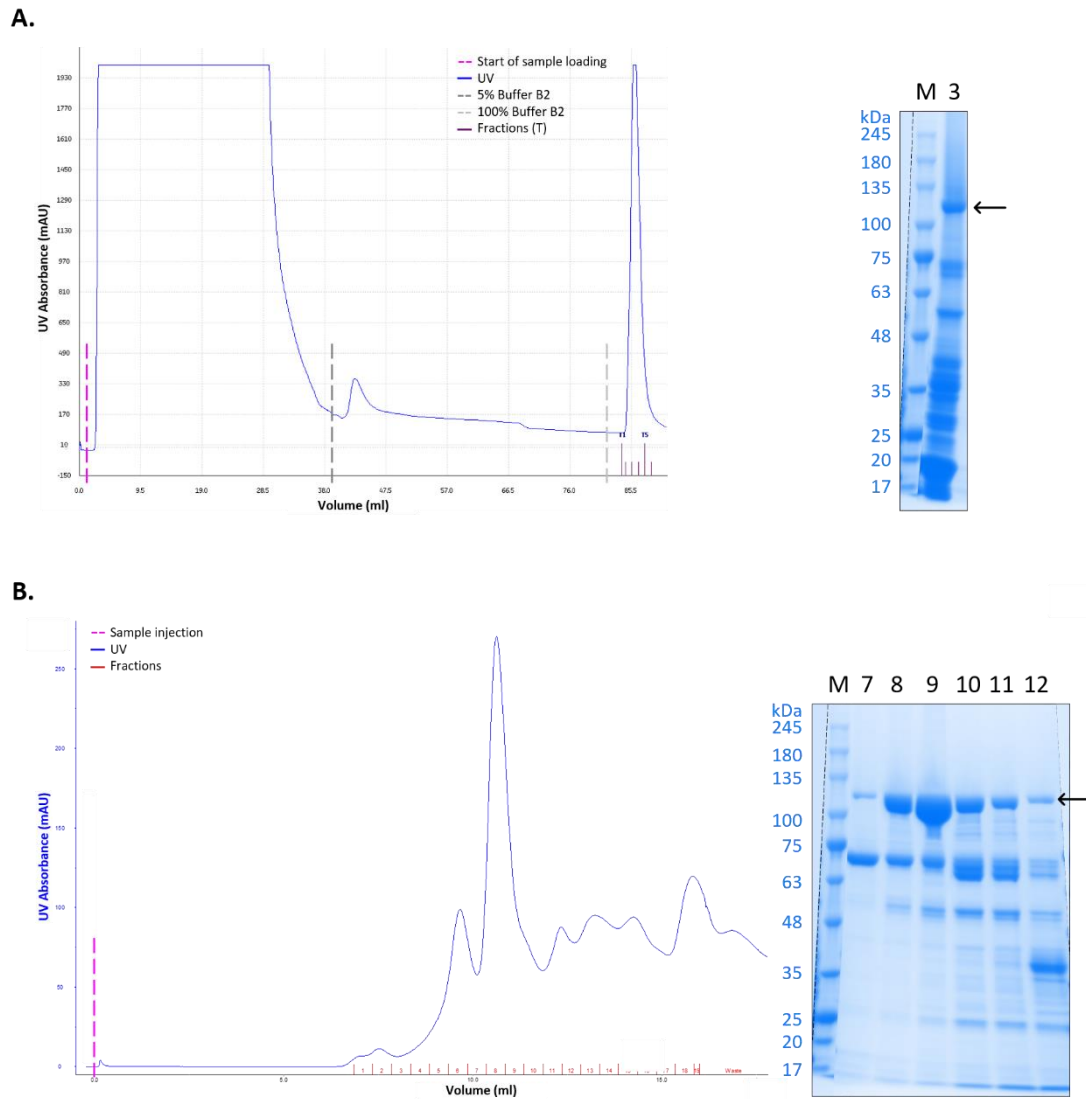
Figure S.3.10. Representative results of the R865T mutant purification. (A.) **Left panel** – Protein elution profile in the affinity chromatography purification step. **Right panel** – SDS-PAGE with collected fractions during this 1st purification step. **(B.) Left panel** – Chromatogram obtained in the size exclusion chromatography step. **Right panel** – SDS-PAGE with collected fractions during this 2nd purification step. Samples loaded in the Novex WedgeWell 8-16% Tris-Glycine Protein Gels are numbered on the top of respective lanes according to the chromatograms. 'M' indicates the lane with the molecular marker (NZYColour Protein Marker II, *NZYTech*), and corresponding band sizes are denoted next to them (in kDa). Arrows are pointing the bands corresponding to full-length SpDis3L2.
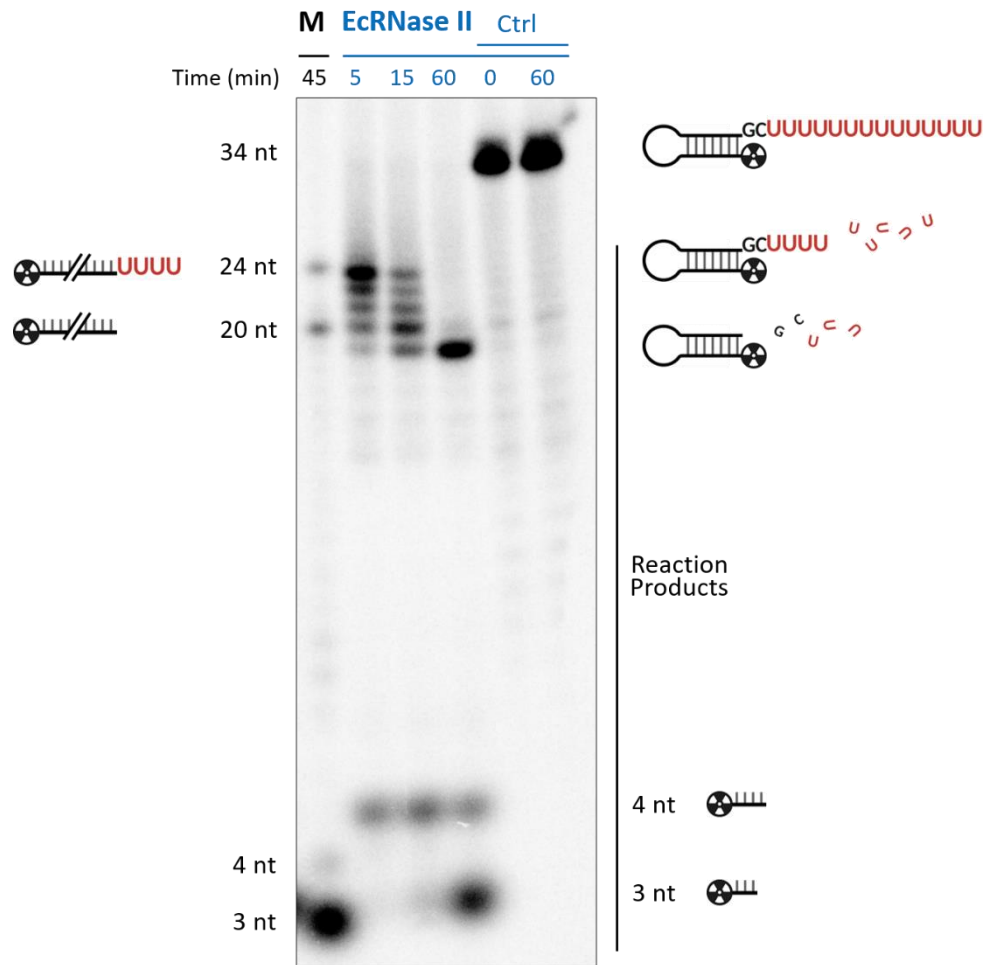
Figure S.3.11. Exoribonucleolytic activity of EcRNase II on dsLoop RNA substrate. This activity assay was a control on the effectiveness of the treatment given to the substrate to obtain its secondary structure. 8.3 nM of Loop ssRNA was submitted to a folding treatment to form dsLoop, which possesses a stem-loop and a -GC14U 3' overhang. The dsLoop substrate was incubated with 50 nM EcRNase II for 60 minutes at 37 ℃. Control reactions (Ctrl) were incubated for the same time without any enzyme. Samples were taken and reaction was stopped at the time-points indicated above each gel lane. RNA substrates and degradation products were separated through migration in a 7 M urea/20% PAA denaturing gel. The size of the molecules is depicted in line with the corresponding bands alongside the gel. 'M' indicates the lane with a sample from an activity assay with the wt SpDis3L2 protein and a mix of two ssRNA (Adh and Adh4U), which was used as a size marker.

ANATOMY OF THE DISEASE-RELATED DIS3L2 RIBONUCLEASE:
DISSECTING THE AMINO ACIDS RESPONSIBLE FOR SUBSTRATE SPECIFICITY

Susana M. Costa

2021