Roles of the putative RNA helicase HELZ in mRNA regulation

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Aoife Julia Hanet
aus Ottignies-Louvain-la-Neuve, Belgien

Tübingen
2019

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	24.04.2019
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Doron Rapaport
2. Berichterstatter:	Prof. Dr. Ralf-Peter Jansen

Declaration

This thesis describes work conducted in the laboratory of Prof. Dr. Elisa Izaurralde in the Department of Biochemistry at the Max Planck Institute for Developmental Biology, Tübingen, Germany, from September 2014 to January 2019. The work was supervised by Prof. Dr. Doron Rapaport at the Eberhard Karls University Tübingen, Germany, and was supported by a fellowship from the International Max Planck Research School from Molecules to Organisms, Tübingen, Germany. I declare that this thesis is the product of my own work. The parts that have been published or where other sources have been used were cited accordingly.

Table of Contents

Abbreviations

(E)GFP	(enhanced) green fluorescent protein
(m)RNP	(messenger) ribonucleoprotein
$(NH_4)_2SO_4$	ammonium sulfate
(q)PCR	(quantitative) polymerase chain reaction
1x/2x PSB	1x / 2x concentrated protein sample buffer
^{32}P	radioactive phosphorus with an extra neutron, 32 atomic mass units
40S	eukaryotic small ribosomal subunit, 40 Svedberg
5'/3'UTR	5'/3' untranslated region
5'AMP	5'adenosine monophosphate
60S	eukaryotic large ribosomal subunit, 60 Svedberg
80S	eukaryotic ribosome, 80 Svedberg
AQR	aquarius protein
As	adenosines
ATP	adenosine triphosphate
AU-rich	Adenylate-Uridylate-rich elements
BASP1	brain acid soluble protein 1
bp	base pairs
bs	binding sites
CAF40	CCR4-associated factor, 40kDa
CCCH-type	cysteine-cysteine-cysteine-histidine-type Zinc finger domain
CCR4	carbon catabolite repression 4 protein
CD	connector domain
cDNA	complementary DNA
CH	cysteine-histidine
CHX	cycloheximide
CN9BD	CNOT9-binding domain
CNOT	CCR4-NOT transcription complex subunit
CO_2	carbon dioxide
CRISPR	clustered regularly interspaced short palindromic repeats
crRNA	CRISPR RNA
Cys_2His_2	two cysteines, two histidines Zinc finger domain

DCP1/2	decapping proteins 1 and 2
DDX6	DEAD-box helicase 6
DEAA	aspartic acid-glutamic acid-alanine-alanine motif
DEAD-box	aspartic acid-glutamic acid-alanine-aspartic acid box family
DEAH/RHA	aspartic acid-glutamic acid-alanine-histidine/RNA Helicase A family
DEDD	aspartic acid-glutamic acid-aspartic acid-aspartic acid family
del. or Δ	deleted
Dm	*Drosophila melanogaster*
DMEM	Dulbecco's modified Eagle media
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
DNA2	DNA replication ATP-dependent helicase/nuclease 2
dNTPs	deoxyribonucleotide triphosphates
DPA	day of proliferation arrest
Dr	*Danio rerio*
dT	deoxythymine
DTT	dithiothreitol
E. coli	*Escherichia coli*
EDC4	enhancer of decapping 4
EDTA	ethylenediaminetetraacetic
EEP	exonuclease-endonuclease-phosphatase family
eIF4G	eukaryotic initiation factor 4G
EJC	exon junction complex
ERG	E26-related gene protein
F-Luc	Firefly luciferase
FastAP	fast alkaline phosphatase
FCS	foetal calf serum
FDR	false discovery rate
GAPDH	glyceraldehyde 3-phopshate dehydrogenase
GB1	protein G, B1 domain -tag
gDNA	genomic DNA
GFP	green fluorescent protein
Gg	*Gallus gallus*

GO	gene ontology
HA	human influenza hemagglutinin-tag
HCl	hydrogen chloride
HEAT	huntingtin, elongation factor 3, protein phosphatase 2A and TOR1 repeat
HEK293(T)	human embryonic kidney cell line 293 (SV40 large T antigen)
HeLa	human cervical cancer cell line
HELZ	helicase with a zinc finger
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HepKANO	human hepatoma cell line
HIF1 α/β	hypoxia-inducible factor 1 alpha and beta
His	Histidine
HLA-DMA	human leukocyte antigen class II histocompatibility antigen - DM α chain
HRP	horse-radish peroxidase
Hs	*Homo sapiens*
IFIH1	interferon-induced helicase C domain containing protein 1
IgG	immunoglobulin gamma
IGHMBP2	immunoglobulin mu-binding protein 2
IP	immunoprecipitation assay
IPTG	isopropyl ß-D-1-thiogalactopyranoside
KCl	potassium chloride
KD	knock-down
KH	hnRNPK homology domain
KH_2PO_4	potassium dihydrogen phosphate
KOH	potassium hydroxide
LB	lysogeny broth
lnRNA	long non-coding RNA
LxxLAP	leucine-x-x-leucine-alanine-proline motif
MALAT1	metastasis associated lung adenocarcinoma transcript 1 lnRNA
MAOA	monoamine oxidase A
MBP	maltose binding protein
$MgCl_2$	magnesium chloride
MIF4G	middle domain of eukaryotic initiation factor 4G
miRNA	micro-RNA

MLLE	methionine-leucine-leucine-glutamic acid domain
Mm	*Mus musculus*
MOPS	3-(*N*-morpholino) propanesulfonic acid
MOV10(L1)	moloney leukemia virus 10 (like 1) protein
mRNA	messenger RNA
Na_2HPO_4	di-sodium hydrogen phosphate
NaCl	sodium chloride
NaH_2PO_4	sodium dihydrogen phosphate
NaOAc	sodium acetate
NaOH	sodium hydroxide
ncRNA	non-coding RNA
NEB	New England Biolabs
NET	NaCl-EDTA-tris buffer
NGD	no-go decay pathway
Ni	nickel
NIM	NOT1-interacting motif
NMD	nonsense-mediated decay pathway
NOT	not on TATA-less
nPABP	nuclear polyA binding protein
NSD	non-stop decay pathway
OD	optical density
ORF	open reading frame
OSTN	osteocrin protein
P-bodies	processing-bodies
PABPC1	cytoplasmic polyA binding protein 1
PABPs	polyA binding proteins
PAGE	polyacrylamide gel electrophoresis
PAM	protospacer adjacent motif
PAM2	PABP-interacting motif 2
PAN2/3	polyA nuclease 2 and 3
PAP	polyA polymerase
PBS	phosphate-buffered saline
PCDH10	protocadherin 10 protein

R-Luc	*Renilla* luciferase
RecA-like	recombinase A-like
RIG1-like	retinoic acid-inducible gene 1 protein-like
RISC	RNA-induced silencing complex
RNA	ribonucleic acid
RNA pol II	RNA polymerase II
RNA-Seq	RNA-Sequencing
rpm	rotation per minute
RRM	RNA recognition motif
rRNA	ribosomal RNA
SD	standard deviation
SDS	sodium dodecylsulfate
SETX	senataxin protein
SF	super-family
Sf21	*Spodoptera frugiperda 21*
SG	stress granules
sgRNA	single guide RNA
sh	small hairpin
Ski2-like	superkiller2-like family
SLiM(s)	short linear motif(s)
SMD	staufen-mediated mRNA decay pathway
SMG7	suppressor of morphogenetic defects in genitalia 7 protein
SMYD2/3	SET and MYND domain-containing proteins 2/3
snRNA	small nucleolar RNA
SPARC	secreted protein acidic rich in cysteine
SQ	Serine-Glutamine
SSC	saline sodium citrate buffer
STREP	streptavidin-binding tag
TBE	tris-boric acid-EDTA buffer
TCA	trichloroacetic acid
TEMED	*N, N, N', N'*-tetramethylenediamine
TENM1	teneurin-1 protein
TMEM35	transmembrane protein 35A

TNRC6	trinucleotide repeats-containing gene 6 protein
TNRC6ASD	TNRC6A silencing domain
Tob	transducer of ErbB-2 protein
tracrRNA	trans-activating crRNA
tRNA	transfer RNA
TTP	tristetraprolin protein
TUT	terminal uridylyltransferase
UPF1	up-frameshift suppressor 1 protein
UV	ultraviolet
$V_{0/1}$	low ($_0$) or high ($_1$) titer virus
w/v	weight / volume
wt	wild type
XRN1	5'-to-3' exoribonuclease 1
Xt	*Xenopus tropicalis*
YFP	yellow fluorescent protein
ZGRF1	GRF-type zinc finger domain-containing protein 1
ZnF	zinc finger
ZNFX1	NFX1-type zinc finger-containing protein 1
β-GAP	reporter mRNA with beta-globin ORF in frame with GAPDH C-terminus
β-globin	beta-globin

Zusammenfassung

Der eukaryotische CCR4-NOT Komplex, ist ein Multi-Proteinkomplex, der sich aus mehreren Teilkomplexen zusammensetzt. Er reguliert viele Aspekte des mRNA Metabolismus, wie zum Beispiel die Transkription, die Translation und die Deadenylierung von mRNAs. Eine meiner Arbeit vorangegangene Studie zeigte, dass die mutmaßliche ‚Helikase mit Zink-Finger' (HELZ) mit dem Gerüstprotein CNOT1 des CCR4-NOT Komplexes interagiert. Außerdem wurde HELZ in Interaktion mit folgenden Proteinen gefunden: der RNA Polymerase II, den Histon-Methyltransferasen SMYD2 und SMYD3, dem zytoplasmatischem Poly(A)-Bindeprotein PABPC1 und mit dem Translations-Inhibitor DDX6. Daher geht man davon aus, dass HELZ eine Rolle bei der Transkription und der Translation von mRNAs spielt. HELZ ist innerhalb der Gruppe der Metazoa konserviert, und wird der UPF1- Familie der SF1 Helikasen zugeordnet. RNA-Helikasen können RNA-Sekundärstrukturen und RNA-Protein Bindungen umstrukturieren bzw. reorganisieren. Trotz der potentiellen Eigenschaft des Proteins als Helikase und seiner interessanten Interaktion mit anderen Proteinen, ist sehr wenig bekannt über die Rolle von HELZ in der zytoplasmatischen mRNA Regularion. Im Rahmen meiner Doktorarbeit führe ich erstmalig eine umfassende Charakterisierung der unterschiedlichen Rollen von HELZ im Bereich der mRNA Regulation im Menschen durch. Ich zeige wie HELZ, wenn es an die 3'UTR einer Reporter mRNA gebunden ist, den CCR4-NOT Komplex rekrutiert und dadurch den Abbau der mRNA vom 5'- zum 3'- Ende initiiert. Dabei weisen meine Daten auf ein komplexes Interaktionsmuster mit mehreren Bindestellen zwischen HELZ und dem NOT-Modul-Teilkomplex des CCR4-NOT Komplexes hin. Weiterhin gelang es mir, durch vergleichende Transkriptom-Analyse von HELZ-Null und Wild-Typ Zellen mehrere potentiell von HELZ regulierte mRNA Transkripte zu identifizieren. Zusätzlich konnte ich beweisen, dass die ATPase Aktivität, die sich in der Helikase Domäne von HELZ befindet, eine wichtige Rolle spielt und ich habe herausgefunden neu, dass HELZ als möglicher Regulator bei der Längen-Regulation des 3'-Poly(A)-Schwanzes fungiert. Es ist bekannt, dass der 3'-Poly(A)-Schwanz von eukaryotischen mRNAs sehr dynamisch ist und sowohl die Stabilität als auch die Translations Effizienz von mRNAs beeinflußt. Interessanterweise resultierte die Expression einer katalytisch inaktiven HELZ-Mutante in einer generellen Zunahme der mRNA 3'-Poly(A)-Schwanz Länge. Zusammenfassend beschreiben meine Studien bisher unbekannte Funktionen von

HELZ im mRNA-Stoffwechsel und listen mögliche, von HELZ beeinflusste, Transkripte auf. Also kann HELZ dazu beitragen die Genexpression auf mehreren Ebenen zu beeinflussen: auf Ebene der Transkription, der mRNA-Prozessierung, der Translation und über den Abbau der mRNA.

Summary

The CCR4-NOT complex is a multi-subunit complex in eukaryotes involved in many aspects of mRNA regulation including transcription, translation and mRNA deadenylation. The putative helicase with a zinc finger, HELZ, was reported to interact with the scaffolding protein CNOT1 of the CCR4-NOT complex. In addition, HELZ was found to interact with RNA polymerase II, histone methyl transferases (SMYD2 and 3), the cytoplasmic polyA binding protein PABPC1 and the RNA helicase and translation repressor DDX6. Accordingly, HELZ is assumed to function in both transcription and translation regulation. HELZ is conserved in metazoans as part of the UPF1-like family of SF1 helicases. RNA helicases are the designated enzymes that re-organise and re-arrange RNA structure and RNA-protein interactions. Despite the potential nature of the protein as a helicase and its interesting interaction network, very little is known about HELZ in cytoplasmic mRNA regulation. In this thesis, I provide for the first time a comprehensive characterisation of different roles of HELZ in cytoplasmic mRNA regulation in human cells. I describe how HELZ recruits the CCR4-NOT complex to induce 5'-to-3' mRNA decay and translational repression when it is bound to the 3'UTR of reporter mRNAs. My data reveal an intricate mode of interaction with HELZ containing several binding regions for the NOT module of the CCR4-NOT complex. Furthermore, by comparing RNA-Sequencing data of a HEK293T HELZ-null cell line versus the data of the wild type cell line, I identified several potential transcripts regulated by HELZ. Additionally, I provide evidence that the ATPase activity of the helicase domain of HELZ is relevant in cells, and I uncover a novel role for HELZ as a possible regulator of 3'polyA tail regulation. 3'polyA tails of eukaryotic mRNAs are known to be dynamic and to determine the stability and translational efficiency of mRNAs. Interestingly, the expression of a ATPase-inactive HELZ mutant causes a general elongation of the length of the 3'polyA tail of mRNAs. In summary, my thesis describes previously unidentified roles of HELZ in mRNA metabolism and a list of potential transcripts targeted by HELZ. Altogether, HELZ could contribute to the regulation of gene expression at multiple levels: transcription, mRNA processing, translation and mRNA decay.

1 Introduction

1.1 From genes to proteins in eukaryotes

All life-forms contain the macromolecule deoxyribonucleic acid (DNA). In eukaryotic cells, DNA is stored in the nucleus which is separated by a semi-permeable membrane from the cytoplasm. Genes are portions of DNA that can be expressed in cells. For a gene to be expressed, it is first transcribed into ribonucleic acids (RNAs) in a process termed transcription. There are many different types of RNAs in cells, each with specific functions. Among them are messenger RNAs (mRNAs) that are translated into proteins in a process known as translation, occurring in the cytoplasm. The human genome contains slightly less than 60 000 reported genes: two-thirds code for non-coding RNAs and the remaining third code for protein-coding mRNAs (Frankish et al. 2018). However, not all genes are expressed at the same time in one human cell. Gene expression is tightly regulated answering to several intra- and inter-cellular cues. The study of these processes helps to understand how a healthy cell is organised, how it answers to stimuli and, how we can find remedies to diseases.

In eukaryotes, RNA polymerase II (RNA pol II) transcribes genes into precursor mRNA (pre-mRNA). RNA pol II contains a catalytic core and a regulatory C-terminal domain that brings together different co-factors required for the maturation of the pre-mRNA (Harlen and Churchman 2017). Pre-mRNAs are matured into an mRNA during transcription by three major modifications: the addition of the 7-methylguanylate cap structure at the 5’ end (5’cap), the removal of introns by splicing and, the cleavage from the polymerase followed by the addition at the 3’ end of a non-templated adenosines tail (3’polyA tail) by the polyA polymerase (PAP) (Will and Luhrmann 2011; Neve et al. 2017). Typically, human canonical PAP adds about 200 to 250 adenosines to the 3’end of pre-mRNAs in the nucleus (Jalkanen et al. 2014; Laishram 2014). The maturation steps are controlled and recognized by protein complexes that bind to the mRNA forming a messenger ribonucleoprotein (mRNP) (for example the exon junction complex (EJC) is deposited at properly spliced regions between two exons and the nuclear polyA binding protein (nPABP) recognises and binds to the nascent 3’polyA tail) (Wickramasinghe and Laskey 2015; Gehring et al. 2017). These proteins act as markers for properly processed mRNPs and signal the export of mRNPs from the nucleus to the cytoplasm (Delaleau and Borden 2015; Wickramasinghe and Laskey 2015). In the cytoplasm, the composition of mRNPs is re-organised and mRNAs are

either translated and/or transported to a specific location in the cell through interactions with motor proteins or via passive diffusion (Eliscovich and Singer 2017). mRNAs are translated by ribosomes in an energy-driven process (Pena et al. 2017). The portion of the mRNA that is translated is called the open reading frame (ORF) and it is flanked by untranslated regions at the 5' end (5'UTR) and the 3' end (3'UTR), which act as regulatory platforms for the stability and for the translation of the mRNA. The ribosome consists of two large subunits that are composed of ribosomal RNAs (rRNAs) and ribosomal proteins. The ribosome translates three nucleic acid bases (= one codon) into one amino acid residue of the nascent protein. One mRNA can be translated by several ribosomes simultaneously forming a polyribosome or polysome; this increases translation efficiency. Translation is divided into three steps: initiation, elongation and termination with ribosome recycling (Kapp and Lorsch 2004). All of those steps involve several eukaryotic initiation, elongation and releasing factors to coordinate the ribosome and the nascent protein (Dever and Green 2012; Roux and Topisirovic 2012; Aylett and Ban 2017).

These briefly described processes, from the transcription of genes to the translation of RNA into proteins, are tightly regulated and controlled. Any misregulation can ultimately lead to diseases as these fundamental process are the basis of life (Lee and Young 2013; Wang et al. 2016).

1.2 mRNA turnover

1.2.1 mRNA decay pathways

Eukaryotes have developed several quality control mechanisms to identify and dispose of faulty mRNAs, which can lead to the production of abnormal and potentially toxic proteins (Houseley and Tollervey 2009). In the nucleus, misprocessed mRNAs (for example splice-defective (pre-)mRNAs containing introns) have abnormal mRNPs composition. These mRNPs are recognized and prevented from export and the faulty mRNAs are targeted for degradation by the nuclear exosome, a multi-subunit protein complex (Kilchert et al. 2016; Singh et al. 2018). In the cytoplasm, mRNA quality control pathways are coupled with the translation machinery: the nonsense-mediated decay pathway (NMD) targets, *inter alia*, mRNAs containing premature termination codon, the no-go decay pathway (NGD) targets mRNAs where ribosomes stall during

translation, and the non-stop decay pathway (NSD) targets mRNAs lacking a termination codon (Lykke-Andersen and Bennett 2014; Simms et al. 2017). Nuclear and cytoplasmic quality control mechanisms ensure that the cytoplasm contains properly processed and translatable mRNAs.

Nevertheless, mRNAs are also degraded in the cytoplasm to regulate protein-synthesis (Houseley and Tollervey 2009; Labno et al. 2016). In this case, the protective 5'cap and 3'polyA tail structures have to be removed so that the body of the mRNA can be fully degraded. The first and rate-limiting step is the deadenylation of the 3'polyA tail (Eliseeva et al. 2013; Smith et al. 2014). Deadenylation is catalysed by exoribonucleases that hydrolyse polyAs in the 3'-to-5' direction, hence termed deadenylases (Goldstrohm and Wickens 2008). Based on their catalytic motif, deadenylases are divided into two families: the DEDD family (for residues aspartic acid- glutamic acid- aspartic acid- aspartic acid) and EEP family (exonuclease-endonuclease-phosphatase) (Yan 2014). There are two main cytoplasmic deadenylases complexes in eukaryotic cells: the PAN2/PAN3 complex (polyA nuclease 2/3 complex) and the CCR4-NOT complex (carbon catabolite repression 4 - not on TATA-less complex) (Miller and Reese 2012; Wolf and Passmore 2014; Labno et al. 2016). PAN2, of the PAN2/PAN3 complex, is a DEDD-type deadenylase. In humans, the PAN2/PAN3 complex is thought to shorten 3'polyA tails to about 150 bases which can be followed by subsequent decay (Wolf and Passmore 2014; Webster et al. 2018). The CCR4-NOT complex acts as the major deadenylase complex in humans and contains two deadenylases (from both deadenylase families) (Miller and Reese 2012; Collart and Panasenko 2017). The CCR4-NOT complex is described in detail in the next sub-chapter (1.2.2 The CCR4-NOT complex).

Once the mRNA is deadenylated, it can be degraded either by the 5'-to-3' or by the 3'-to-5' decay pathways (Houseley and Tollervey 2009; Labno et al. 2016). In the 5'-to-3' decay pathway, the deadenylated mRNA is decapped by the DCP1/DCP2 complex (decapping proteins 1/2). The complex consists of the pyrophosphatase DCP2 and several enhancers, activators and regulators like DCP1 or EDC4 (enhancer of decapping 4) (Grudzien-Nogalska and Kiledjian 2017). Once the cap is removed, the mRNA body is fully digested by the 5'-to-3' exoribonuclease 1 (XRN1) (Nagarajan et al. 2013; Labno et al. 2016). In the 3'-to-5' decay pathway, the cytoplasmic exosome degrades the mRNA from the deadenylated 3'end. Eukaryotic exosome complexes are highly conserved. They consist of nine non-catalytic core components arranged in a

hexameric ring capped by a tri-protein ring (Zinder and Lima 2017). In higher eukaryotes, the nuclear and cytoplasmic exosome complex differ in their association with ribonucleases and associated protein complexes (Tomecki et al. 2010; Kilchert et al. 2016). Together, the 5'-to-3' and the 3'-to-5' mRNA decay pathways constantly remove mRNAs to fine-tune protein synthesis in the cell.

1.2.2 The CCR4-NOT complex

The CCR4-NOT complex is essential in all eukaryotes and has been referred to as the "control freak of the eukaryotic cell" as it plays key roles in many steps of mRNA metabolism (Miller and Reese 2012). The function, structure and composition of the complex are well conserved in eukaryotes with some species-specific variations.

The human CCR4-NOT complex is organised into four structurally different modules that docks onto the conserved scaffold protein CNOT1 (CCR4-NOT transcription complex subunit 1). CNOT1 is a large protein with several folded motifs and domains: HEAT repeats (huntingtin, elongation factor 3, protein phosphatase 2A and TOR1), a similar middle domain to the eukaryotic initiation factor 4G domain (MIF4G) and a CNOT9 binding domain (CN9BD) (**Figure 1**) (Collart and Panasenko 2017). There are two deadenylases in the complex: the EEP deadenylase CNOT6, or its paralogue CNOT6L, and the DEDD deadenylase CNOT7, or its paralogue CNOT8 (Bartlam and Yamamoto 2010; Petit et al. 2012). CNOT6/CNOT6L binds to CNOT7/CNOT8, which in turn, binds to the MIF4G domain of CNOT1, thereby constituting the so-called catalytic module (**Figure 1**) (Bartlam and Yamamoto 2010; Petit et al. 2012). Recent studies have shed light into the potential reason for the presence of two deadenylases within the complex. The 3'polyA tail of mRNAs are constantly bound by PABPC1s (cytoplasmic polyA-binding protein 1), which promote mRNA translation and protect the 3'polyA tail in the cytoplasm. The function of CNOT6 appears to be to actively displace PABPC1 from the 3'polyA tail of mRNAs, making way for CNOT7 to deadenylate the unbound 3'polyA tail; both enzymes work together to deadenylate the targeted mRNA (Webster et al. 2018; Yi et al. 2018). The CAF40/CNOT9 module (CCR4-associated factor 40kDa/CNOT9) consists of the protein CNOT9 bound to the CN9BD of CNOT1 (**Figure 1**) (Chen et al. 2014; Mathys et al. 2014). The NOT module comprises CNOT2 and CNOT3 interacting with each other and to the C-terminal domain of CNOT1 (**Figure 1**) (Bhaskar et al. 2013; Boland et al. 2013). The N-terminal region of CNOT1 is bound by CNOT10 and CNOT11

forming the CNOT10/CNOT11 module (Bawankar et al. 2013; Mauxion et al. 2013) (**Figure 1**).

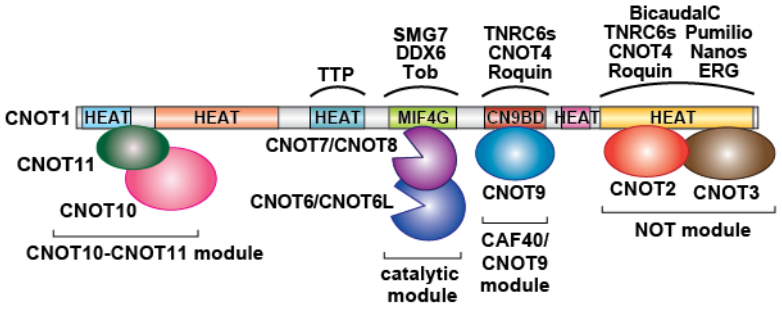

Figure 1 Schematic representation of the human CCR4-NOT complex. CNOT1 is the scaffold protein of the complex and contains several indicated HEAT repeats, a MIF4G domain and a CN9BD. The CCR4-NOT complex is organised into four structurally independent modules, which are marked on the lower part of the scheme. Different RNA-associated proteins interacting with various components of metazoan CCR4-NOT complex are indicated on the upper part of the scheme (Chicoine et al. 2007; Horiuchi et al. 2009; Braun et al. 2011; Van Etten et al. 2012; Fabian et al. 2013; Loh et al. 2013; Chen et al. 2014; Mathys et al. 2014; Raisch et al. 2016; Rambout et al. 2016; Sgromo et al. 2017; Keskeny et al. 2019). TTP - tristetraprolin, DDX6 - DEAD-box helicase 6, SMG7 - suppressor of morphogenetic defects in genitalia 7, Tob - transducer of ErbB-2, TNRC6s - trinucleotide repeats-containing gene 6, ERG - E26-related gene.

So far, the best studied function of the CCR4-NOT complex is as the major deadenylase complex in cells leading to mRNA decay. The complex is recruited to specific mRNAs by interacting with RNA-binding and RNA-associated proteins. For example, the CCR4-NOT complex is a key player in micro RNA (miRNA)-induced repression of mRNAs (**Figure 1**) (Braun et al. 2011; Chen et al. 2014; Mathys et al. 2014). miRNAs are 20-22 nucleotides-long RNAs coded in the genome that recognise mRNAs by base-pairing to repress the expression of the targeted-mRNAs (Fabian and Sonenberg 2012; Iwakawa and Tomari 2015). A miRNA is bound by argonaute proteins and TNRC6 proteins forming a miRNA-induced silencing complex (miRISC) (Fabian and Sonenberg 2012). The CCR4-NOT complex interacts with TNRC6 proteins and is recruited to miRNA-targeted mRNAs: thus, the miRNA-targeted mRNAs are translationally repressed, deadenylated and degraded by the 5'-to-3' mRNA decay pathway (Jonas and Izaurralde 2015). Additionally, the CCR4-NOT complex can be recruited by many specific RNA-binding/associated proteins like the protein TTP, which binds to AU-rich mRNAs (Fabian et al. 2013; Fu and Blackshear 2017), or the Roquin proteins, which bind to a constitutive decay element sequence on specific transcripts (Leppek et al. 2013; Sgromo et al. 2017). In both cases, targeted mRNAs are deadenylated by the CCR4-NOT complex and degraded (**Figure 1**). The

study of the recruitment of the CCR4-NOT complex has revealed a common mode of interaction between SLiMs (short linear motifs) in RNA-binding/associated proteins that mediate the binding to the CCR4-NOT complex (Jonas and Izaurralde 2013). A SLiM is, by definition, a short sequence (on average six amino acids long), located in disordered and fast evolving sequence regions and can fold into a secondary structure upon binding (Davey et al. 2015). SLiM-mediated interactions are present in a multitude of pathways and the network of the CCR4-NOT complex is one of the many that highlights the use of SLiMs (Davey et al. 2015). The typical low affinity character of a SLiM-mediated interaction has the advantage of being easily modulated to not retain and detain the CCR4-NOT complex longer than the complex requires to perform its deadenylase function on targeted mRNAs.

Another known function of the CCR4-NOT complex is its ability to repress mRNA translation independently of its deadenylase activity (Collart 2016). This is well illustrated in the miRNA-induced repression of mRNAs. Indeed, miRNAs induce translational repression of their target via the recruitment of the RNA helicase DDX6 by the CCR4-NOT complex (Chen et al. 2014; Mathys et al. 2014). The CCR4-NOT complex interacts directly with DDX6 in cells, which is a known translational repressor and binds to several decapping factors (Coller and Parker 2005; Chen et al. 2014; Mathys et al. 2014). However, the CCR4-NOT complex likely employ other mechanisms to induce translational repression, independently of DDX6 (Jonas and Izaurralde 2015).

The many other functions of the CCR4-NOT complex are less understood but conserved in eukaryotes. The complex was originally identified as a transcriptional regulator of genes, it is also involved in DNA damage response and can induce protein degradation and ubiquitination (Miller and Reese 2012; Shirai et al. 2014; Gupta et al. 2016; Ukleja et al. 2016; Collart and Panasenko 2017). Since its discovery, the CCR4-NOT complex has been the focus of much research. Ongoing studies are working out to understand how the complex assembles and how it is regulated to accomplish its many functions.

1.3 Helicases

1.3.1 Definition and classification

Helicases are crucial enzymes that are involved in virtually all aspects of nucleic acid metabolism. They are present in archaea, prokaryotes, eukaryotes, and also in viruses. By definition, a helicase is a DNA/RNA-dependent ATPase that unwinds nucleic acids. However, their function now includes among others: remodelling of nucleic acid-protein interactions, (re-)annealing nucleic acids and acting as clamps (Jankowsky 2011; Wu 2012; Cordin and Beggs 2013; Bourgeois et al. 2016; Kanaan et al. 2018). Helicases are divided into six super-families (SF) based on their primary structure: SF1 and SF2 comprise prokaryotic and eukaryotic helicases, whereas SF3, SF4, SF5 and SF6 are not represented in eukaryotes but in prokaryotes, archaea and viruses only (Fairman-Williams et al. 2010). Helicases can be further divided according to their substrate (DNA and/or RNA, single stranded, double stranded, hybrids), their directionality of action (3'-to-5' are "A" type, 5'-to-3' are "B" type) and their structure (ring-shaped or non-ring-shaped) (Fairman-Williams et al. 2010). Non-eukaryotic, ring-forming helicases are included in SF3, SF4, SF5 and SF6 and will not be discussed in this thesis.

The helicase core of SF1 and SF2 helicases are very alike and fold into tandem RecA-like domains (Ye et al. 2004). The RecA-like domain, originally identified in recombinase A of *Escherichia coli* and part of the ATPase family, folds into central beta sheets flanked by alpha helices (Ye et al. 2004). Additionally, SF1 helicases often have other domains that are inserted in-between the primary structure of the helicase core without affecting the secondary structure of the helicase (Fairman-Williams et al. 2010). Both SFs are characterised by twelve motifs and one additional SF1- or SF2-specific motif. All motifs are organised in the same sequential way within the helicase core primary structure (**Figure 2** A) (Fairman-Williams et al. 2010). The sequences of the motifs are different depending if the helicase is SF1 or SF2; but the functions of the motifs are similar. Motifs define the directionality and the binding to nucleic acid (motifs Ia-c, IV, IVa (not present in SF1), V and Vb), and the coordination of the adenine base of the ATP (motifs Q, IIIa (not present in SF2)) (**Figure 2** A) (Fairman-Williams et al. 2010; Sloan and Bohnsack 2018). Motifs III and Va are involved in the interplay between ATP binding and nucleic acid binding sites and are highly conserved within

each superfamily. Motif I (aka Walker A), motif II (aka Walker B) and motif VI are located in and around the cleft between the two RecA-like domains and are the only motifs conserved between both SFs (Fairman-Williams et al. 2010). The reason behind this conservation is that these motifs are crucial for binding and hydrolysis of ATP. Notably, motif II consists of the linear DExx sequence; essential for any helicase as mutating the strictly conserved aspartic acid or glutamic acid residues abolish the ATPase activity of helicases (Fairman-Williams et al. 2010).

Based on phylogenetic sequence alignments, SF1 helicases are further divided into three families, and SF2 are divided into ten families (Fairman-Williams et al. 2010). Most families contain both DNA and RNA helicases, with the exception of the DEAD-box SF2 family (named after motif II: aspartic acid- glutamic acid- alanine- aspartic acid), which are all RNA helicases. Eukaryotic helicases that can bind and act on RNA belong to the UPF1-like SF1 family (up-frameshift suppressor 1-like family contains eleven members) and to the Ski2-like (superkiller2-like), the RIG1-like (retinoic acid-inducible gene 1 protein-like), the DEAD-box and the DEAH/RHA (DEAH amino acid in motif II / RNA helicase A) SF2 families (**Figure 2** B).

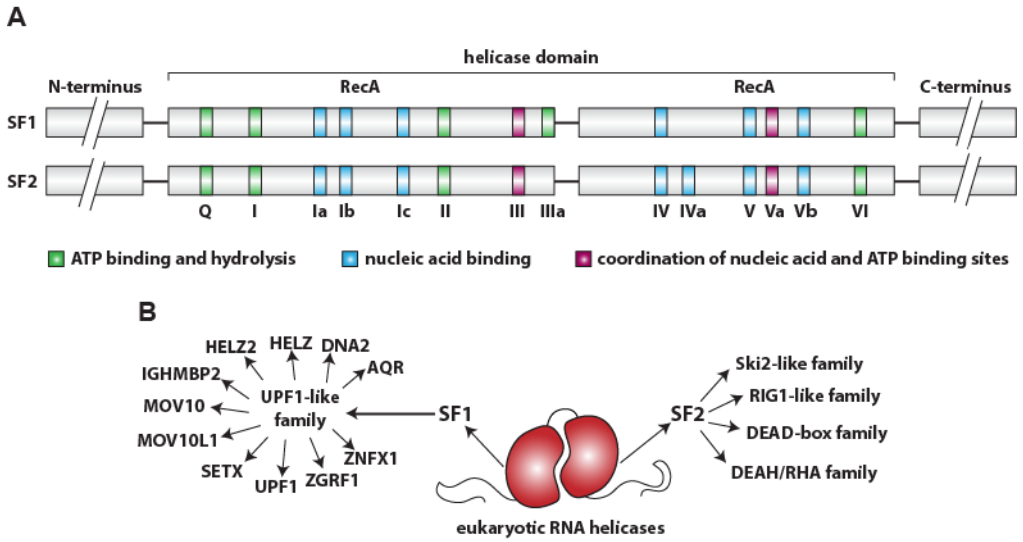

Figure 2 Overview of SF1 and SF2 helicases. **A** Schematic representation of SF1 and SF2 helicases with primary structure arrangement of common and SF-specific motifs. SF1 and SF2 helicase cores are folded in a tandem RecA-like domain. Each domain contains several motifs; names and color-coded functions of the motifs are indicated below the scheme. Green motifs contribute to ATP binding and hydrolysis, blue motifs contribute to nucleic acid binding, and purple motifs contribute to the coordination between nucleic acid and ATP binding. Adapted from (Fairman-Williams et al. 2010). **B** Scheme of an SF1/SF2-RNA helicase folded into two RecA-like domains (in red) with N- and C-terminal extensions (in black). Eukaryotic RNA helicases belong to the UPF1-like SF1 family and the Ski2-like, the RIG-I-like, the DEAD-box and the DEAH/RHA SF2 families. The eleven UPF1-like helicases are named. AQR - aquarius protein, DNA2 - DNA replication ATP-dependent helicase/nuclease 2, HELZ(2) - helicase with a zinc finger (2), IGHMBP2 - immunoglobulin mu-binding protein 2, MOV10(L1) - moloney

leukemia virus 10 (like1), SETX - senataxin protein, ZGRF1 - GRF-type zinc finger domain-containing protein 1, ZNFX1 - NFX1-type zinc finger-containing protein 1.

1.3.2 Mechanism of action of eukaryotic RNA helicases

There are two main mechanisms of action of RNA helicases described so far: the canonical unwinding and the non-canonical duplex unwinding (Jankowsky 2011; Sloan and Bohnsack 2018). In the canonical unwinding, the helicase unwinds RNA strands by hydrolysing ATP and moves forward on the single strand. The second mechanism was only characterised for DEAD-box SF2 helicases: helicases unwind RNA by binding to ATP and are released from the strand by hydrolysing ATP. In both cases, long stable duplex strands are more challenging to unwind and thus, require several rounds of ATP hydrolysis. However, a number of studies characterising the function of helicases have revealed a diverse set of mechanism of action. Indeed, RNA helicases can re-arrange protein and nucleic acid interactions, translocate on nucleic acids without unwinding and, act as clamps, as opposed to a strict RNA-unwinding activity (Jankowsky 2011; Cordin and Beggs 2013; Fiorini et al. 2015; Kanaan et al. 2018). Also, RNA helicases can give directionality to a process: for example, in splicing, the hydrolysis of ATP by a helicase is used to re-arrange RNA and protein interaction to assure the effective splicing of introns and commits the process to the next step of splicing (Ficner et al. 2017).

To accomplish their many functions, RNA helicases often have N- and C-terminal auxiliary extensions, which are used for protein-protein interactions, regulation of the helicase specificity and activity, and recruitment to specific RNPs (Fairman-Williams et al. 2010; Sloan and Bohnsack 2018). Helicases typically have low sequence specificity for their substrate so they can bind in a non-specific way to various RNAs or they rely on their auxiliary RNA-binding domains (like RNA recognition motifs (RRM) domains, hnRNPK homology (KH) domains, zinc fingers (ZnF) domains) or other proteins to bring them to the correct RNP. In other cases, the ATPase activity is regulated by regulatory auxiliary extensions or by protein interactions. For example, for many DEAD-box helicases, the ATPase activity is activated or inhibited when one of their RecA-like domains binds to a MIF4G domain of a MIF4G-containing proteins (Oberer et al. 2005; Bourgeois et al. 2016; Sloan and Bohnsack 2018). The interaction network of a RNA helicase can thus be quite complex with each component contributing to the regulation of the function of the helicase.

1.3.3 UPF1-like SF1 RNA helicases

Up to date, all eukaryotic SF1 helicases with RNA binding ability belong to the UPF1-like family (**Figure 2** B). Some of those helicases have been described to have both RNA and DNA binding abilities. Generally, UPF1-like RNA helicases are very different in terms of domain organisation, apart from a common helicase domain and long N- and C-terminal extensions. Interestingly, half of these helicases have non-related ZnF domains (Fairman-Williams et al. 2010). In humans, loss of the ATPase activity of some SF1 helicases has already been linked to diseases (Maystadt et al. 2004; Guenther et al. 2009; Ronchi et al. 2013).

The study of UPF1-like RNA helicases has previously uncovered new aspects of mRNA regulation. Up to date, five UPF1-like helicases have been shown to function directly in mRNA metabolism; specifically, in transcription, splicing, RNA silencing, and the quality control of mRNAs (Arjan-Odedra et al. 2012; Richard et al. 2013; De et al. 2015; Vourekas et al. 2015; Ariumi 2016; Ficner et al. 2017; TD and Wilkinson 2018). For example, MOV10 acts in the miRNA pathway as part of the miRISC complex and can restrict the propagation of mobile genetic elements and of viral RNAs (Meister et al. 2005; Arjan-Odedra et al. 2012; Gregersen et al. 2014; Ariumi 2016; Warkocki et al. 2018). The helicase AQR is part of the intron binding complex aiding the spliceosome to remodel itself for efficient splicing (De et al. 2015; Ficner et al. 2017). The helicase MOV10L1 is essential for the biogenesis of piwi-interacting RNA (piRNA) (Vourekas et al. 2015; Fu et al. 2016). The piRNA pathway is an animal-specific pathway where small non-coding RNAs are encoded in the genome and are used to stop the propagation of transposable elements (Siomi et al. 2011). Last but not least, the helicase UPF1 plays a central role in the NMD pathway, among other roles. The activity of UPF1 is inhibited by its N- and C-terminal regions. During NMD, UPF1 is released of this inhibition in a step-wise manner to finally recruit an endonuclease or the CCR4-NOT complex to target the NMD-transcript to decay (Fiorini et al. 2015; Dehghani-Tafti and Sanders 2017; Gupta and Li 2018; TD and Wilkinson 2018). However, there is a clear disparity between how well each individual SF1 RNA helicases has been studied so far. Given the importance of characterised UPF1-like helicases in mRNA metabolism, the study of overlooked SF1 helicases promises to deepen our understanding of mRNA regulation.

1.4 The putative helicase HELZ

HELZ proteins are UPF1-like SF1 helicases, paralogues of HELZ2, MOV10 and MOV10L1 (Fairman-Williams et al. 2010). HELZ proteins (along with HELZ2, MOV10 and MOV10L1) are only present in metazoans and not in plants nor fungi. So far, HELZ proteins have only been superficially studied; even though these proteins contain several features that indicate that they might have an interesting role in mRNA metabolism. In human, the *HELZ* gene is located on chromosome 17 and the protein is 1942 amino acid residues long (**Figure 3** A).

HELZ proteins are characterized by a CCCH-type ZnF at the N-terminus, a putative helicase core in the mid region, and a PABPC1 interacting motif 2 (PAM2) followed by a long unstructured C-terminal tail (**Figure 3** A) (Wagner et al. 1999). The helicase core of HELZ has retained all the conserved motifs that defines it as a UPF1-like SF1 helicase. In **Figure 3** B, the motif II of human HELZ, HELZ2, MOV10, MOV10L1 and UPF1 are aligned highlighting the conservation of the DExx sequence, crucial for the ATPase activity of any helicase. However, there has been no study done so far to characterise the ATPase and helicase activity of HELZ; thus, HELZ proteins are referred to as putative helicases. ZnFs are domains that coordinate protein-protein and protein-nucleic acid interactions in almost all cellular pathways (Razin et al. 2012). These domains are frequent features in eukaryotic proteins and are divided in several families. The Cys_2His_2 family is among the largest ZnF family and several structural studies have revealed the conservation of the protruding finger-like fold coordinating a metal ion (zinc for example) (Razin et al. 2012). The ZnF of HELZ is a CCCH-type zinc finger derived from the Cys_2His_2 family, it is the only eukaryotic SF1 helicases with such a ZnF domain (Razin et al. 2012; Fu and Blackshear 2017). The majority of CCCH-type containing proteins that are characterized, are involved in RNA metabolism (Fu and Blackshear 2017); the archetype member of this family, TTP, binds to AU-rich elements located in the 3'UTR of mRNAs through its tandem CCCH-type zinc fingers. Another striking feature present in HELZ proteins is the PAM2 motif. So far, they are the only helicases described to have this motif. PAM2 motifs enable the interaction with MLLE domain-containing proteins, such as PABPC1 (MLLE stands for the amino acids: Methionine- Leucine- Leucine- Glutamic Acid) (Albrecht and Lengauer 2004). The PAM2 motif has been identified and structurally characterised in many other proteins, besides HELZ, that are related to translation regulation and

binding to PABPC1 (for example eIF4G, TNRC6C and PAN3) (Albrecht and Lengauer 2004; Eliseeva et al. 2013). PABPC1 comprises four RRMs in the N-terminal region that bind to the 3'polyA tail of mRNAs and a C-terminal region that contains the MLLE domain (Eliseeva et al. 2013; Smith et al. 2014). *Homo sapiens* (*Hs*) HELZ was reported to interact with PABPC1 via its PAM2 motif in human cell extracts (Hasgall et al. 2011). The rest of the C-terminal tail of *Hs* HELZ is not folded, based on secondary structure predictions, and has low complexity regions with enriched Proline, Serine and Glutamine residues (Zimmermann et al. 2018). Nevertheless, two LxxLAP motifs (Leucine-x-x-Leucine-Alanine-Proline (x being any amino acid)) can be identified (Hasgall et al. 2011) (**Figure 3** A). LxxLAP motifs have been studied in the protein hypoxia-inducible factor (HIF) 1α where it acts as an oxygen sensor: in state of normoxia, the Proline residue of the motif is hydroxylated and this targets the protein for degradation. In hypoxic state, the hydroxylation does not occur, HIF1α binds to HIF1β, and the HIF-dimer induces a hypoxic stress response. However, a previous study has revealed that *Hs* HELZ expression is not regulated in an oxygen-dependent manner in human cells (Hasgall et al. 2011).

There is, so far, only one study reporting the expression pattern of HELZ in a whole organism. The authors used a reporter gene inserted in the mouse *HELZ* gene and visualised the expression of the tag in the developing mouse embryo (Wagner et al. 1999). They observed a wide and dynamic expression pattern in different embryonic stages and developing organs. They could also detect high transcript levels of *HELZ* predominantly in the adult brain and, to a lower extent, in the heart and kidneys (Wagner et al. 1999).

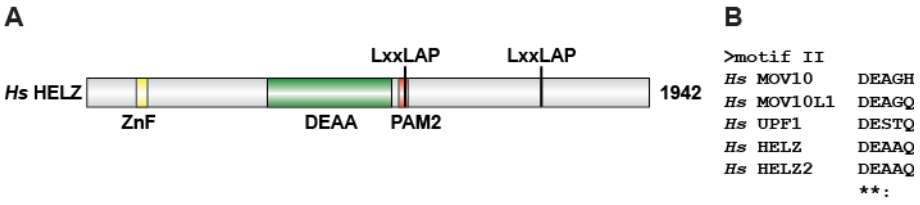

Figure 3 HELZ proteins are putative UPF1-like SF1 helicases. **A** Schematic representation of *Hs* HELZ. The conserved zinc finger (ZnF), helicase domain (DEAA), PAM2 and LxxLAP motifs are highlighted in yellow, green, orange and as black bars, respectively. Predicted unstructured regions are in grey. **B** Alignment of motif II, a key motif for the ATPase activity of helicase, in human MOV10, MOV10L1, UPF1, HELZ, and HELZ2. Stars mark strictly conserved amino acid residues within the motif, colon marks amino acid residues from the same class of amino acid.

In addition to several sequence features pointing to a role in RNA metabolism, several studies on *Hs* HELZ have also revealed that it might function in the regulation of transcription and mRNA translation. Indeed, *Hs* HELZ was reported to interact with RNA polymerase II and SMYDs 2 and 3 proteins (SET and MYND domain-containing proteins) in human cervical cancer cell line (HeLa cells) (Hamamoto et al. 2004; Diehl et al. 2010). SMYDs are histone methyltransferases involved in the differentiation of myocytes and cardiomyocytes for SMYD2, and SMYD3 is involved in cancer and tumorigenesis (Spellmon et al. 2015). The authors propose that *Hs* HELZ forms a complex with RNA pol II and either one of the two SMYDs to regulate the transcription of specific genes (Hamamoto et al. 2004; Diehl et al. 2010). Interestingly, HELZ proteins have been described as predominantly cytoplasmic and do not possess any nuclear localisation signal (Wagner et al. 1999; Hasgall et al. 2011; Ayache et al. 2015). In line with a cytoplasmic function, a report described *Hs* HELZ to be a positive regulator of cellular proliferation and translation initiation in HeLa cells (Hasgall et al. 2011). The authors observed that the overexpression of *Hs* HELZ increased the protein levels of a reporter mRNA without affecting its stability. Accordingly, when the expression of *Hs* HELZ was reduced, the same authors observed a decrease of general protein translation. In another study, Nagai et al. concluded that *Hs* HELZ might act as a tumour suppressor as *HELZ* transcript levels are down-regulated in 28 out of 95 cancer cell lines (predominantly in brain cancer cell lines, cervical cancer cell lines and neuroblastoma cells) and, as exogenous expression of *Hs* HELZ was found to reduce the cellular growth of hepatoma cells (HepKANO) (Nagai et al. 2003). However, it is unclear how *Hs* HELZ exactly contributes to the regulation of transcription and mRNA translation.

Interestingly, *Hs* HELZ was found to interact with CNOT1 of the CCR4-NOT deadenylase complex in a human embryonic kidney cell line (HEK293) in a study to understand how miRNAs induce translational repression (Mathys et al. 2014). The Mid region of CNOT1 (consisting of the third HEAT repeat, the MIF4G domain and the CN9BD in **Figure 1**) was identified to interact with *Hs* HELZ, among other potential interactors (Mathys et al. 2014). In addition, the authors revealed that within the Mid region, the repressive activity was centred at the MIF4G domain of CNOT1. However, *Hs* HELZ was not interacting with the MIF4G domain of CNOT1 in human cells. On the other hand, the authors identified the RNA helicase DDX6 that could readily interact with the MIF4G domain of CNOT1, and structures of this interaction were obtained

(Chen et al. 2014; Mathys et al. 2014). The authors do not pursue further to study the significance of the interaction between HELZ and the CCR4-NOT complex. Yet, this interaction might indicate a role of *Hs* HELZ in mRNA decay.

Lastly, *Hs* HELZ was identified in mass spectrometry analysis as a potential interactor of *Hs* DDX6 in HEK293 cells but this interaction was not confirmed by independent experiments (Ayache et al. 2015). *Hs* HELZ was also reported to associate with P-bodies and CCR4-NOT-enriched stress granules (SG) in HEK293 cells (Youn et al. 2018). P-bodies are eukaryotic membrane-less foci held together via a network of interactions between RNA, RNA-associated proteins and, most components of the 5'-to-3' mRNA decay pathway including the CCR4-NOT complex (Schutz et al. 2017; Luo et al. 2018). P-bodies are thought to be localised storage environments for mRNPs until translation or degradation, and for mRNA decay factors until their use (Coller and Parker 2005; Standart and Weil 2018). SG are also cytoplasmic foci but, in contrast to P-bodies, they are induced by cellular stress and contain proteins regulating translation (Protter and Parker 2016).

In summary, *Hs* HELZ has an interesting interaction network that implies a role in cytoplasmic mRNA metabolism, even though the biological significance and the nature of the detected interactions remain uncharacterised. Moreover, HELZ could have an ATPase activity linked to its helicase domain. Thus, this study aimed to understand the cellular function of HELZ in cytoplasmic mRNA regulation in human cells.

2 Aims

All mRNAs are bound by regulatory proteins forming mRNPs. RNA helicases are enzymes that have the ability to re-organise RNP composition and structure; thus, playing pivotal roles in the formation of mRNP and in determining the fate of mRNAs. However, many helicases are uncharacterized. I chose to study the human putative RNA helicase with a zinc finger HELZ as it has several intriguing features and yet, has been overlooked. Based on protein sequence, metazoan HELZ is a putative UPF1-like SF1 helicase with unique elements: a CCCH-type ZnF domain and a PAM2 motif. These features relate to RNA control and RNA binding. Previous studies have already placed HELZ as a potential regulator of transcription and translation, but mechanistic details are unknown (Nagai et al. 2003; Hamamoto et al. 2004; Diehl et al. 2010; Hasgall et al. 2011). Moreover, given its suggested interaction network (interacting with PABPC1, the scaffolding protein CNOT1 of the CCR4-NOT deadenylase complex, and potentially, DDX6 and P-bodies), it might have a role in cytoplasmic mRNA decay (Hasgall et al. 2011; Mathys et al. 2014; Ayache et al. 2015). Thus, *Hs* HELZ displays the potential of being a regulator of mRNA at multiple levels.

Surprisingly, there is a lack of knowledge about the direct impact of *Hs* HELZ on mRNA. Furthermore, the role of the helicase domain has not been examined and its impact on cellular pathways remains elusive. Therefore, my aim was to investigate a potential role of *Hs* HELZ in cytoplasmic mRNA regulation. I have divided my work into four parts tackling four distinct questions:

1) HELZ and mRNA stability: Does *Hs* HELZ affect mRNAs when bound to 3'UTRs?
2) HELZ and translation: Does *Hs* HELZ play a role in the regulation of mRNA translation?
3) The helicase domain of HELZ: Does *Hs* HELZ require its ATPase activity to function in mRNA regulation?
4) HELZ and its potential mRNA targets: Which cellular processes are affected by *Hs* HELZ?

To answer these questions, I combined biochemical and genetic approaches performed in human cell lines. Henceforth, *Hs* HELZ will be referred to as HELZ.

3 Materials and methods

3.1 Materials

3.1.1 Oligonucleotides

All oligonucleotides were purchased from Sigma-Aldrich in lyophilized form and diluted in autoclaved MilliQ water (Merck) to a final concentration of 100 μM.

Table 1 Oligonucleotides for the design of HELZ constructs

(numbers represent amino acid residues on HELZ protein; the restriction enzyme, if used, are named; "m" stands for mutagenesis primers; "g" stands for guide RNA; "f" and "r" stand for forward and reverse, respectively)

	Sequence (5' to 3')
1_SacII_f	atacatccgcggatatggaagacagaagagctgaaaagt
1_SalI_f	atacatgtcgacatggaagacagaagagctgaaaagt
1050_XbaI_r	acattctagattaatcacccaccacagcaaccagggat
1051_SacII_f	atacatccgcggatcccattgctctgtgctctattggaa
1051_BspTI_f	atacatcttaagcccattgctctgtgctctattggaa
1475_BspTI_f	atacatcttaagcttccttcacatctgaatagc
1475_SacII_f	atacatccgcggatcttccttcacatctgaatagc
1474_XbaI_r	acattctagaaatggggccgggttgtgcaatggct
1643_BspTI_f	atacatcttaaggtagccagcaacccagcatttcca
1643_SacII_f	atacatccgcggatgtagccagcaacccagcatttcca
1642_XbaI_r	acattctagattattcaatgtctctgctgttatcatta
1793_BspTI_f	atacatcttaagcaatcttctttcaacttttcatccc
1793_SacII_f	atacatccgcggatcaatcttctttcaacttttcatccc
1792_XbaI_r	acattctagattagttactgtggtcctgaagct
1942_XbaI_r	acattctagattatttaaaatatgagtaaaagcca
m_E795Q_f	tttacacacattctattagatcaagctgcccaggccatggagtgtga
m_E795Q_r	tcacactccatggcctgggcagcttgatctaatagaatgtgtgtaaa
m_F1107V_f	gtgttgaatccgctggcacctgaagttatcccccgggctctaagactgca
m_F1107V_r	tgcagtcttagagcccgggggataacttcaggtgccagcggattcaacac
m_del1595-1610_f	catcgtgatcaaagtgaaacacgggaactagtacagagtagaagcccaccagcagtccc
m_del1595-1610_r	gggactgctggtgggcttctactctgtactagttcccgtgtttcactttgatcacgatg
m_del1746-1764_f	cagtatcttcttcttcgctccctagcttaagctcagttcaaccttgttctgaagaag
m_del1746-1764_r	cttcttcagaacaaggttgaactgagcttaagctagggagcgaagaagaagatactg
m_del1890-1907_f	gcagtagcccccagagctctgcggggccccctcctgagcaggccaagaagag

m_del1890-1907_r	ctcttcttggcctgctcaggaggggcccgcagagctctgggggctactgc
m_del1915-1942_f	gcccaggccccctcctgagcaggccaagatgcagtacaaacttatcctgaacggtaaaacc
m_del1915-1942_r	ggttttaccgttcaggataagtttgtactgcatcttggcctgctcaggaggggcctgggc
m_1474-GB1_f	ctgagagccattgcacaacccggccccattatgcagtacaaacttatcctgaac
m_1642-GB1_f	aactttaatgataacagcagagacattgaaggcgctatgcagtacaaacttatcctgaacggtaaa
m_1792-GB1_f	cagtgtaaagagcttcaggaccacagtaacggcgctatgcagtacaaacttatcctgaacggtaaa
m_1942-GB1_f	gcaatggcttttactcatattttaaaatgcagtacaaacttatcctgaac
m_GB1-pnEA_r	cggatctcctagggctagctctagatcattccgttacggtgtaggttttg
m_GB1-6xHis_f	taccaaaacctacaccgtaacggaaatgggcagcagccatcatcatcatcactaatctagagctagccctaggagat
m_6xHis-pnEA_r	atctcctagggctagctctagattagtgatgatgatgatgatggctgctgcccatttccgttacggtgtaggttttggta
m_1942-MBP_f	ggagctcaggcgggagctcaggcgaaactgaagaaggtaaactggtaatc
m_1942-MBP_r	gtcgagactgcaggctctagattaagtctgcgcgtctttaagggcttcatc
m_pLIB-2xSTREP_f	gcggaattcaaaggcctacgtcgacatgggcagcggctggagccacccccagttcgagaagggcagcggctggagccacccccagttcgagaagcccattgctctgtgctctattggaa
m_2xSTREP-1051_r	ttccaatagagcacagagcaatgggcttctcgaactgggggtggctccagccgctgcccttctcgaactgggggtggctccagccgctgcccatgtcgacgtaggcctttgaattccgc
g_HELZ_KO_f	caccggcaactagtaacgccctctc
g_HELZ_KO_r	aaacgagagggcgttactagttgcc
HELZ_exon8_f	ggtgttatgaagaggagagt
HELZ_exon8_r	ctgtactagcttaggacaga

Table 2 Quantitative polymerase chain reaction (qPCR) primers

("f" and "r" stand for forward and reverse, respectively)

Target	Sequence (5' to 3')
TMEM35A_f	tcggtgatcctctcaaacgc
TMEM35A_r	gcctggaaggtccactcttt
HLA-DMA_f	ccaatgtggccagatgacct
HLA-DMA_r	ggacaccgggattttcccat
SPARC_f	ctagaggctcagtggtggga
SPARC_r	tccctagagcccctgagaag
BASP1_f	tggatttccaagatccgcgt
BASP1_r	tggacaagctaagtgggctc
TENM1_f	tcgcctgatggaaccctcta
TENM1_r	ccattgctgctggtaatcgc
PCDH10_f	tctagttgacagacctcgcc
PCDH10_r	ggaacatgcagattgctgcg
HELZ_f	tcaagctagcagccaagagg
HELZ_r	gaatgcagcatcgtgtcacc

OSTN_f	acagtcagggaagagaaatcagcc
OSTN_r	agcctctggaatttgaaagccg
MAOA_f	gtggtatgtgaagcagtgcg
MAOA_r	atcttggcagtcaaggtcgg
IFIH1_f	ccaaatggaagttgcccagc
IFIH1_r	aatggttggaactccttgcg

3.1.2 Plasmids

Table 3 HELZ (Uniprot P42694)

Construct	Comment
pT7-MS2-HA-HELZ	
pT7-EGFP-HELZ	
pT7-MS2-HA-HELZ 1 to 1050	HELZ_N
pT7-EGFP-HELZ 1 to 1050	
pT7-MS2-HA-HELZ 1051 to 1942	HELZ_C
pT7-EGFP-HELZ 1051 to 1942	
pLIB-2xSTREP-HELZ 1051 to 1942 -MBP	
pT7-MS2-HA-HELZ_F1107V	Disrupts PAPBC1-binding
pT7-EGFP-HELZ_F1107V	
pT7-MS2-HA-HELZ_E795Q	Disrupts motif II of the helicase domain
pT7-EGFP-HELZ_E795Q	
pT7-MS2-HA-HELZ_E795Q F1107V	Disrupts motif II of the helicase domain and PABPC1-binding
pT7-EGFP-HELZ_E795Q F1107V	
pnEA-NvM-HELZ 1051 to 1474-GB1-6xHis	HELZ_C1
pT7-MS2-HA-HELZ 1 to 1474	HELZ_N+C1
pT7-EGFP-HELZ 1 to 1474	
pT7-MS2-HA-HELZ 1475 to 1942	HELZ_C2
pT7-EGFP-HELZ 1475 to 1942	
pnEA-NvM-HELZ 1475 to 1942-GB1-6xHis	
pnEA-NvM-HELZ 1475 to 1642-GB1-6xHis	HELZ_C2a
pnEA-NvM-HELZ 1475 to 1642 del. 1595 to 1610-GB1-6xHis	HELZ_C2aΔ1
pnEA-NvM-HELZ 1643 to 1792-GB1-6xHis	HELZ_C2b
pnEA-NvM-HELZ 1643 to 1792 del. 1746 to 1764-GB1-6xHis	HELZ_C2bΔ2
pnEA-NvM-HELZ 1793 to 1942-GB1-6xHis	HELZ_C2c
pnEA-NvM-HELZ 1793 to 1942 del. 1890 to 1904-GB1-6xHis	HELZ_C2cΔ3
pnEA-NvM-HELZ 1793 to 1914-GB1-6xHis	HELZ_C2cΔ4
pSUPERpuro-shHELZ	Used to knock-down HELZ in human cells, generated by Dr. Duygu Kuzuoglu-Öztürk
pSpCas9(BB)-2A-PURO-HELZ-sgRNA-V2	Used to generate HEK293T HELZ-null cells

Table 4 Human deadenylation and decapping factors

Construct	Uniprot	Reference
pT7-MS2-HA-CNOT1_isoform1	A5YKK6	(Chen et al. 2014)l
pT7-EGFP-C1-CNOT1_isoform1 1085 to 1605		(Petit et al. 2012)
pSUPERpuro-shCNOT1		(Chen et al. 2014)
pT7-EGFP-C1-CNOT7 D40A E45A	Q9UIV1	(Braun et al. 2011)
pcDNA3.1-lambdaN-HA-C1-PAN3	Q58A45	(Braun et al. 2011)
pCIneo-lambdaN-HA-DCP2	Q8IU60	(Tritschler et al. 2009)
pT7-EGFP-C1-DCP2 E148Q		(Chang et al. 2014a)
pCIneo-lambdaN-HA-DCP1a	Q9NPI6	(Tritschler et al. 2009)
pCIneo-lambdaN-HA-EDC4	Q6P2E9	(Tritschler et al. 2009)
pcDNA3.1-lambdaN-HA-C1-PATL1	Q86TB9	(Braun et al. 2010)
pCIneo-lambdaN-HA-DDX6	P26196	(Tritschler et al. 2009)
pT7-EGFP-C1-DDX6		
pcDNA3.1-lambdaN-HA-C1-XRN1	Q8IZH2	(Braun et al. 2012)

Table 5 Reporters, controls and empty vectors

Construct	Reference
pT7-EGFP-C1	(Tritschler et al. 2009)
pT7-EGFP-MBP	(Lazzaretti et al. 2009)
pT7-MS2-HA-C1	
pEGFP-N3-F-Luc	(Lazzaretti et al. 2009)
pCIneo-R-Luc	(Pillai et al. 2004)
pCIneo-R-Luc-6xMS2bs	(Pillai et al. 2004)
pCIneo-R-Luc-6xMS2bs-A95-MALAT1	(Bhandari et al. 2014)
pcDNA3.1-β-globin-GAPDH	Kind gift from Dr. Jens Lykke-Andersen
pcDNA3.1-β-globin-6xMS2bs	
pcDNA3.1-MS2-HA-TNRC6A 1462 to 1962	Q8NDV7, (Chen et al. 2014)
pSUPERpuro-shcontrol	Kind gift from Prof. Oliver Mühlemann

3.1.3 Buffers, solutions and reagents

All buffers were diluted in autoclaved MilliQ water, except if stated otherwise.

Table 6 Buffers and solutions for the analysis of DNA

Name	Composition
TBE	89 mM Trizma base (Sigma), 88.9 mM Boric acid (Merck), 2.5 mM Ethylenediaminetetraacetic acid (EDTA; Merck)
Agarose gel	For 1%: 3 g of Agarose (Peqlab) dissolved in 300 ml of TBE, 0.003% Ethidium bromide (Roth)
5xDNA dye	20% Ficoll PM 400 (Sigma), 1 mM EDTA (Merck), 0.1% Sodium dodecylsulfate (SDS; Serva), 0.05% Bromophenol Blue (Sigma)

10xin-house Taq buffer	200 mM Trizma base-HCl pH 8.55 (Sigma), 160 mM Ammonium sulfate ($(NH_4)_2SO_4$; Merck), 0.1% Tween (Sigma), 20 mM Magnesium chloride ($MgCl_2$; Merck)
DNA ladder	20% of 100 bp or 1 kbp DNA ladder (New England Biolabs (NEB)), 20% of 5xDNA dye, 60% of 10 mM Trizma-HCl (Sigma), 1 mM EDTA (Merck)

Table 7 Buffers and solutions for the analysis of RNA

Name	Composition
5xDNAse buffer	400 mM HEPES-KOH pH7.5 (Roth), 120 mM $MgCl_2$ (Merck), 10 mM Spermidine (Sigma), 20 mM Dithiothreitol (DTT; Biomol)
Agarose gel	For 1.2%: 3.6 g Agarose Ultra Pure (Invitrogen) dissolved in 270 ml of MilliQ water (Merck) and 30 ml of 10xMOPS
Glyoxal	60% Dimethylsulfoxide (DMSO; VWR), 50% Dionised glyoxal, 12% 10x MOPS, 6% Glycerol 86-89% (Sigma)
RNA dye	95% Deionized formamide, 0.05% SDS (Serva), 0.05% Xylene cyanol FF (Sigma), 0.05% Bromophenol Blue (Sigma)
10xMOPS	200 mM 3-(*N*-morpholino) propanesulfonic acid (MOPS; Roth), 80 mM Sodium acetate (NaOAc; Merck), 10 mM EDTA (Merck), pH is adjusted to 7 with NaOH, light-sensitive
20xSSC	3 M Sodium chloride (NaCl; Roth), 300 mM Trisodiumcitrate dehydrate (Roth)
Church buffer	500 mM Sodium phosphate pH 7, 7% SDS (Serva), 1 mM EDTA (Merck)
Northern blot Washing buffer	40 mM Sodium phosphate pH 7, 1% SDS (Serva), 1 mM EDTA (Merck)
Gradient buffer	75 mM Potassium chloride (KCl; Merck), 1.5 mM $MgCl_2$ (Merck), 10 mM Trizma base-HCl pH 7.4 (Sigma)
Polysome profile lysis buffer	10 mM Trizma-HCl pH 7.4 (Sigma), 10 mM NaCl (Roth), 1.5 mM $MgCl_2$ (Merck), 0.5% Triton X-100 (Merck), 2 mM DTT (Biomol), 1 U/µl RiboLock RNase inhibitor (Thermo Fischer), 0.5% Sodium-deoxycholate (Sigma), 50 µg/ml Cycloheximide (CHX was dissolved in ice-cold Ethanol; Serva)
Sucrose	10% or 50% of D (+)-Sucrose (Fluka) dissolved in Gradient buffer, filtered
Deionized formamide	Mix 5 g of AG 501-X8 mixed-bead resin (Bio-Rad) per 50 ml Formamide (Merck)
Deionized glyoxal	Mix 20 g of AG 501-X8 mixed-bead resin (Bio-Rad) per 20 ml Glyoxal solution (Sigma)
Sodium phosphate pH 7	Titrate 1 M di-Sodium hydrogen phosphate (Na_2HPO_4 ; Roth) with 1 M Sodium dihydrogen phosphate monohydrate (NaH_2PO_4 ; Merck) to pH 7

Table 8 Buffers and solutions for the analysis of proteins

Name	Composition
NET buffer	50 mM Trizma base-HCl pH 7.5 (Sigma), 150 mM NaCl (Roth), 0.1 % Triton-X 100 (Merck), 1 mM EDTA pH 8 (Merck)
2xProtein sample buffer (2xPSB)	100 mM Trizma base-HCl pH 6.8 (Sigma), 4% w/v SDS (Serva), 20% Glycerol (Sigma), 200 mM DTT (Biomol), 0.05% Bromophenol Blue (Sigma)
1xPSB	50% 2xPSB, 200 mM Trizma base (Sigma)
8% Resolving gel	375 mM Trizma base pH 8.7 (Sigma), 26% Rotiphorese Gel 30 (Roth), 0.1% SDS (Serva), 0.5% Ammonium persulfate (10% w/v) (APS 10%; Sigma), 0.15% *N, N, N', N'*-Tetramethylenediamine (TEMED; Sigma). Other percentage calculated accordingly
Stacking gel	100 mM Trizma base pH 6.8 (Sigma), 16% Rotiphorese Gel 30 (Roth), 0.08% SDS (Serva), 0.5% APS 10% (Sigma), 0.12% TEMED (Sigma)
Laemmli buffer	3.5 mM SDS (Serva), 0.19 mM Glycin (Roth), 24.8 mM Trizma base (Sigma) diluted in desalted water
Transfer buffer	20 mM Trizma base (Sigma), 149 mM Glycin (Roth), 0.1% SDS (Serva), 20% Methanol (Roth) diluted in desalted water
Ponceau	0.5% w/v Ponceau S (Sigma), 1% ice-cold Acetic acid (Merck)
PBS	10 mM Na_2HPO_4 (Roth), 1.8 mM Potassium dihydrogen phosphate (KH_2PO_4 ; Merck), 0.137 M NaCl (Roth), 2.7 mM KCl (Merck)
Western blot Blocking buffer	PBS, 0.3% Tween20 (Sigma), 5% w/v Skimmed milk powder (Reform)
Western blot Washing buffer	PBS, 0.3% Tween20 (Sigma)
Detection solution A	2 mM Luminol (light-sensitive; Roth), 100 mM Trizma base-HCl pH 8.6 (Sigma)
Detection solution B	7 mM *p*-Coumaric acid (Sigma) dissolved in DMSO (VWR)
Detection solution mix	90% detection solution A, 10% detection solution B, 0.01% Hydrogen peroxide 35% (Sigma)
Coomassie stain	45% Methanol (Roth), 10% Acetic acid 96% (Merck), 0.1% w/v Brilliant Blue R-250 (Thermo Fischer)
Coomassie destaining	25% Isopropanol (Roth), 10% Acetic acid 96% (Merck)
Pulldown buffer for insect cells	50 mM HEPES pH 7 (Roth), 200 mM NaCl (Roth), 2 mM DTT (Biomol), 0.01 mg/ml DNAse I, cOmplete™ Protease Inhibitor (Roche)
Ni-lysis buffer for *E. coli*	50 mM HEPES pH 7 (Roth), 200 mM NaCl (Roth), 20 mM Imidazole (Roth), 2 mM ß-mercaptoethanol (Sigma), 0.01 mg/ml DNAse I, Lysozyme (Sigma) and cOmpleteTM Protease Inhibitor (Roche)
Pulldown buffer for *E. coli*	50 mM HEPES pH 7 (Roth), 200 mM NaCl (Roth), 2 mM DTT (Biomol), 0.01 mg/ml DNAse I, Lysozyme, cOmplete™ Protease Inhibitor (Roche)
Pulldown elution buffer	50 mM HEPES pH 7 (Roth), 200 mM NaCl (Roth), 2 mM DTT (Biomol), 25 mM D-(+)-Maltose monohydrate (Roth)

3.1.4 Enzymes

All enzymes were used in their corresponding buffer according to the manufacturer's protocol.

Table 9 Enzymes

Types	Enzymes
Polymerases	Taq polymerase (lab-made), *Pfu* DNA Polymerase (Thermo Fischer and Promega), Phusion® High-Fidelity DNA polymerase (Thermo Scientific)
Restriction enzymes	SacII, FastDigest SalI, FastDigest XbaI, FastDigest BspTI, FastDigest DpnI (all from Thermo Fischer)
Other DNA modifying enzymes	RevertAid H Minus Reverse Transcriptase (Thermo Fischer), T4 DNA Ligase (Thermo Fischer), FastAP Thermosensitive Alkaline Phosphatase (Thermo Fischer), TOPO TA cloning (Invitrogen)
Enzymes for RNA study	DNaseI RNase-free (Thermo Fischer), RNase H (NEB), RiboLock RNase Inhibitor (Thermo fischer)
Enzymes for protein study	Lysozyme (Sigma), DNaseI (Roche), RNase A (Qiagen)

3.1.5 Antibodies

All antibodies used for Western blot were diluted in Blocking buffer and stored at -20°C. For antibodies used in immunofluorescence assay, see the sub-chapter 3.2.2.4 Immunofluorescence assay.

Table 10 Antibodies

Antibody	Source	Cat. number	Dilution
Anti-HA-HRP	Roche	12013819001	1:5000
Anti-HA	BioLegend	9015303	1:1000
Anti-GFP	Roche	11814460001	1:3000
Anti-GFP	In-house		1:400
Anti-mouse IgG-Alexa594	Thermo Fischer	A-11005	1:1000
Anti-mouse IgG-HRP	GE Healthcare	NA931V	1:10000
Anti-rabbit IgG-HRP	GE Healthcare	NA934V	1:10000
Anti-tubulin	Sigma Aldrich	T6199	1:10000
Anti-*Hs* HELZ	Abnova	H00009931-M02	1:1000
Anti-*Hs* PABPC1	Abcam	Ab21060	1:5000
Anti-*Hs* CNOT1	In-house		1:1000
Anti-*Hs* CNOT3	Abcam	Ab55681	1:1000
Anti-*Hs* DDX6	Bethyl Laboratories	A300-461Z	1:3000

3.1.6 Cell lines characteristics

Table 11 Human, bacterial and insect cell lines used in this study

Cell line	Use	Comment
HEK293T wt and mutant	Biochemical assays	
HeLa		Coding for a NMD reporter. Kind gift from Prof. Oliver Mühlemann
***E. coli* TOP10**	Molecular cloning	From Thermo Fischer
***E. coli* DH5α**		
***E. coli* DH10EMBACY**		Kind gift from Prof. Imre Berger
***E. coli* BL21 Star™ (DE3)**	Recombinant protein expression	From Invitrogen
***Spodoptera frugiperda 21* (*Sf21*)**	Recombinant protein expression	Kind gift from Prof. Imre Berger

Table 12 Growth conditions and media

Cell line	Growth condition	Media
HEK293T	Adherent on flasks or plates at 37°C, 5% CO_2	DMEM[a] (Gibco)
HeLa		
TOP10	On LB-Agar[b,c] plates at 37°C; In suspension[c] at 37°C	LB media: 0.5% w/v Yeast extract (Roth), 1% w/v Peptone ex casein (Roth), 85 mM NaCl (Roth)
DH5α		
DH10EMBACY	On LB-Agar[b,d] plates at 37°C	
BL21 Star™ (DE3)	On LB-Agar plates[b,c] at 37°C; In suspension in Syrax flasks[c] at 37°C or 20°C	
Sf21	At 27°C on plates At 27°C in suspension in Pyrex flasks	Sf-900™ II (Gibco)

[a] Dulbecco's Modified Eagle Media (4.5 g/l D-glucose) (DMEM) (Gibco) supplemented with 10% Foetal Calf Serum (FCS - Thermo Fischer), 2 mM L-Glutamine (Gibco), and with 1:200 Penicillin/Streptomycin (Gibco)

[b] 10 cm LB-Agar plates are made by dissolving 1.5% w/v of Agar-Agar (Roth) in hot LB and poured directly in 10 cm dishes.

[c] Antibiotic incorporated according to plasmid resistance: ampicillin (AppliChem) 100 μg/ml, kanamycin (Serva) 25 μg/ml, both dissolved in MilliQ water (Merck) and filtered through a 0.22 μm Millex-HV (Merck).

[d] Antibiotic incorporated to 10 cm plates of LB-Agar: kanamycin 25 μg/ml (Serva), tetracyclin 10 μg/ml (Fluka), chloramphenicol 34 μg/ml (dissolved in ethanol 100% (Merck); Sigma), and gentamycin 10 μg/ml (Roth), and I added 2 g of Bluo-Gal (Thermo Fischer, diluted in a 50:50 water: DMSO mix) and 5 mM IPTG (Roth, dissolved in MilliQ water (Merck)) on the plates before seeding the cells.

3.2 Methods

3.2.1 Molecular cloning

3.2.1.1 Polymerase chain reaction

Phusion® HF DNA polymerase (Thermo Fischer) was used to amplify DNA according to the manufacturer's protocol (see **Table 1** for primers). Phusion polymerase is a 5'-to-3' DNA polymerase with 3'-to-5' exonuclease activity and generating blunt-ends. A typical set-up is as follows:

Amount	Reagent
10 μl	5x Phusion HF buffer
1 μl	dNTPs (10 mM each, from Peqlab)
10 ng	Template
1 μl	Forward primer (25 μM)
1 μl	Reverse primer (25 μM)
0.5 μl	Phusion® HF DNA polymerase
Up to 50 μl	MilliQ water

Temperature	Time	
98°C	30 sec.	
98°C	10 sec.	35X
2°C less than the annealing primers	30 sec.	
72°C	45 sec. per kbp	
72°C	10 min.	
4°C	Until use	

After the thermic cycles, 5xDNA dye is added to the amplicon mix (12 μl of dye for 50 μl of amplicon mix), and the mix is separated by size in a 1% Agarose gel at a 100 Volt (V). A DNA ladder is loaded on the gel to estimate the size of the PCR-products. Bands are visualized by UV light (as ethidium bromide is added to the agarose gel and fluoresce under UV) on a Quantum Gel Documentation Imager (Vilber). The appropriate PCR-product is cut-out and purified from the gel using the GeneJET Gel Extraction kit from Thermo Fischer Scientific.

3.2.1.2 Restriction reaction and ligation

Amplicons and 1 μg of plasmid vector were digested with appropriate restriction enzymes (1 μl of each enzyme was used per reaction, see **Table 1** and **Table 9**) at 37°C for 10 minutes when FastDigest enzymes were used, or for one hour when 10 Units of SacII was used. Vectors were further treated with 1 μl of FastAP (see **Table 9**). FastAP is an alkaline phosphatase which cleaves the phosphate group from DNA thus preventing re-ligation of the vector. FastAP was inactivated by incubating the mixture at 75°C for 5 minutes. Digestion products were separated on a 1% agarose gel, visualized as previously described, and the appropriate bands were purified with the GeneJET Gel extraction kit. The digested amplicons and vectors were mixed to have a twofold excess of amplicons or, if the amplicon is bigger than the vector (which is usually the case for full-length HELZ, N and C), to have an equal ratio of both. Ligation of the digested amplicons to the vector was done in 20 μl (1 μl of T4 DNA ligase 5 U/μl (see **Table 9**), 2 μl of T4 DNA Ligase 10x buffer, appropriate amount of vector and amplicon, and MilliQ water to fill up) at 22°C overnight. To test for the correct insertion in the vector, a test restriction was performed to cut out the insert from the plasmid using 0.5 μg of plasmid digested with 0.5 μl of enzyme(s) in a final volume of 20 μl (in appropriate buffer). The products were visualised in an agarose gel as previously described.

3.2.1.3 Site-directed mutagenesis

Site-directed mutagenesis was used to introduce point mutation or to insert/remove sequences to a plasmid of interest. The technique relies on the amplification of the plasmid with a specific complementary primer-pair designed to introduce a desired mutation. I used *Pfu* polymerase to amplify the complete plasmid sequence with a specific primer-pair. The mutagenesis was performed according to the manufacturer's protocol (Stratagene mutagenesis protocol). After amplification, the template plasmid was digested with DpnI, which recognizes and cleaves at $GA^{m6}\downarrow TC$ sites at 37°C for 6 to 8 hours. The template plasmid is methylated as it is purified from *E. coli* whereas the product of the mutagenesis is not methylated as it is *in vitro* amplified. The digestion with DpnI results in the complete digestion of the template plasmid without affecting the mutagenesis product.

3.2.1.4 Transformation

E. coli TOP10 (for plasmids ≥ 7000 bp) or DH5α (for plasmids < 7000 bp) were transformed with ligation products or site-directed mutagenesis products. Typically, 10 µl of product was incubated with 100 µl of heat-shock competent bacteria on ice for 15 minutes; the bacteria were then heat-shocked at 42°C for 40 seconds. The samples were incubated for one hour at 37°C in 600 µl of LB media. The bacteria were then plated on 10 cm LB-Agar dishes with appropriate antibiotics (see **Table 11** and **Table 12**).

3.2.1.5 Plasmid extraction from *E. coli*

Single bacterial colonies were picked from LB-Agar plates and grown in suspension in 3 ml of LB-antibiotic at 37°C for 8 hours or overnight. Plasmid DNA was extracted using the QIAGEN® Plasmid Mini kit. A correct plasmid was determined by restriction reaction analysis and by sequencing. Positive colonies were inoculated in 100 ml of LB-antibiotic flask and grown at 37°C overnight. The plasmid DNA was extracted with the QIAGEN® Plasmid *Plus* Midi kit or the QIAfilter Plasmid Midi kit (for plasmids coding for shRNA and gRNA) and used for transfection. The concentration of the plasmids was measured with the NanoDrop ND-1000 (PeqLab).

3.2.1.6 Sequencing

All inserts of constructs used in this thesis were sequenced. Sequencing reactions were performed with the BigDye™ Terminator v3.1 Cycle Sequencing Kit (Thermo Fischer) according to the manufacturer's protocol by the Genome centre of the Max Planck Institute for Developmental Biology, Tübingen.

3.2.2 Cell culture methods

3.2.2.1 Transfection in human cell lines

I used Lipofectamine® 2000 Transfection Reagent from Thermo Fischer to transfect human cell lines. Briefly, on day one, the cells are seeded in 6-wells plate(s) at a concentration of $0.85x10^6$ cells in 2 ml/well in DMEM without antibiotics to have a confluence of about 60-70% at the time of transfection. The next day, cells were transfected with a mixture of plasmids (see **Table 13**). On day three, the media was replaced by DMEM with antibiotics; and finally, the cells were harvested on the fourth day. For immunoprecipitation assay, I seeded $4x10^6$ cells in 10 cm dish and transfected, the next day, with the reagent TurboFect from Thermo Fischer according

to the manufacturer's protocol. Cells were harvested on the fourth day (see the sub-chapter 3.2.5.1 Immunoprecipitation assay). The amount of transfected plasmids DNA is listed in **Table 13**.

Table 13 Amount (μg per well or plate) of plasmid DNA used for transfection

Reporters	**Amount**
pEGFP-N3-F-Luc	0.2
pCIneo-R-Luc	0.2
pCIneo-R-Luc-6xMS2bs	0.2
pCIneo-R-Luc-6xMS2bs-A95-MALAT1	0.5
pcDNA3.1-ß-globin-GAPDH	0.5
pcDNA3.1-ß-globin-6xMS2bs	0.5
Tethered proteins	**Amount**
pT7-MS2-HA-HELZ	1
pT7-MS2-HA-HELZ 1 to 1050	1.35
pT7-MS2-HA-HELZ 1051 to 1942	2.5
pT7-MS2-HA-HELZ_F1107V	1
pT7-MS2-HA-HELZ_E795Q	1.2
pT7-MS2-HA-HELZ_E795Q+F1107V	1
pT7-MS2-HA-HELZ 1 to 1474	1
pT7-MS2-HA-HELZ 1475-1942	0.25
pT7-MS2-HA-CNOT1_isoform1	1
pcDNA3.1-MS2-HA-TNRC6A 1462 to 1962	0.5
Deadenylation assay	**Amount**
pT7-EGFP-C1-CNOT1_isoform1 1085 to 1605	1
pT7-EGFP-C1-CNOT7 D40A E45A	1
pT7-EGFP-MBP	0.04
Decapping assay	**Amount**
pT7-EGFP-C1-DCP2 E148Q	2
pT7-EGFP-C1	0.15
(Co-)Immunoprecipitation (IP)	**Amount**
pT7-EGFP-HELZ	20/15 (co-IP)
pT7-EGFP-HELZ 1 to 1050	30
pT7-EGFP-HELZ 1051 to 1942	25
pT7-EGFP-HELZ_F1107V	20
pT7-EGFP-HELZ_E795Q	15 (co-IP)
pT7-EGFP-HELZ 1 to 1474	15
pT7-EGFP-HELZ 1475 to 1942	3.5

pT7-EGFP-CNOT1_isoform1	15 (co-IP)
pcDNA3.1-lambdaN-HA-PAN3	10 (co-IP)
pCIneo-lambdaN-HA-DCP2	10 (co-IP)
pCIneo-lambdaN-HA-DCP1a	10 (co-IP)
pCIneo-lambdaN-HA-EDC4	10 (co-IP)
pcDNA3.1-lambdaN-HA-C1-PatL1	10 (co-IP)
pCIneo-lambdaN-HA-DDX6	10 (co-IP)
pcDNA3.1-lambdaN-HA-C1-XRN1	10 (co-IP)
GFP-tagged proteins	**Amount**
pT7-EGFP-HELZ	2
pT7-EGFP-HELZ_F1107V	2
pT7-EGFP-HELZ_E795Q	1.5
pT7-EGFP-HELZ_E795Q+F1107V	1
pT7-EGFP-C1-DDX6	0.2

3.2.2.2 Knock-down of HELZ and of CNOT1 in HeLa cells

Depletion of HELZ or CNOT1 in HeLa cell lines was performed according to a previously described protocol (Brummelkamp et al. 2002; Chen et al. 2014). The pSUPERpuro plasmid encodes for a small hairpin RNA (shRNA) transcribed by RNA polymerase III. The transcript resembles a pri-miRNA-like fold and thus undergo processing by the Drosha complex and the Dicer complex; to finally be incorporated into a miRISC complex and target the ORF of the desired mRNA. The pSUPERpuro plasmid also encodes for the Puromycin N-acetyltransferase gene from *Streptomyces alboniger*, which was used to select positively transfected cells when they were treated with Puromycin. pSUPERpuro-shcontrol was used as a control; the shRNA coded in the plasmid does not target any mRNAs. HeLa cells were transfected with 4 μg of pSUPERpuro shHELZ/shCNOT1/shcontrol using Lipofectamine® 2000 (Thermo Fischer) and selected a day later with media supplemented with 1.5 μg/ml of Puromycin. After 24 hours of selection, cells were re-seeded. On the following day, cells were transfected with the appropriate 2 μg of pSUPERpuro shRNA and various constructs, as described in sub-chapter 3.2.2.1 Transfection. Cells were selected for 48 hours in media with 1.5 μg/ml of Puromycin and the media was changed back to one without Puromycin for 24 hours. Cells were then harvested. The knock-down efficiency was assessed by western blotting. The protein levels of the targeted protein in knock-down cells were compared to various amounts of control knock-down cells (10, 25, 50 and 100%) (see sub-chapter 3.2.5.2 Western blot).

3.2.2.3 Generation of the HEK293T HELZ-null cell line

To generate a stable cell line that is not expressing HELZ, I used the CRISPR/Cas9 genome engineering technology (clustered regularly interspaced short palindromic repeats / CRIPSR-associated protein 9). The CRISPR locus is originally part of a bacteria's adaptive immune system (Lone et al. 2018). The locus encodes for short viral DNAs (or other foreign nucleic acids of attacking pathogen) separated by short motifs termed protospacer adjacent motif (PAM), specific for each bacterium. Cas genes are localised close to the CRISPR locus. When under attack by the same or similar pathogen, the transcribed RNA from the CRIPSR locus guide Cas proteins, which recognise the host-specific PAM sequence, leading to degradation of the DNA of the intruder; thus, defending the bacteria. The system has been adapted to be used in a plethora of cell lines and organisms and modified to be used as a versatile genome editing tool (Jiang and Doudna 2017; Lone et al. 2018). The CRISPR/Cas9 has seen considerable development. Cas9 DNA endonuclease activity is guided by a sgRNA (a single guide RNA comprising at least one CRISPR RNA (crRNA) and one trans-activating crRNA (tracrRNA)) to target DNA. When bound to target DNA, Cas9 complex creates a double strand break in the genomic DNA and, due to imperfect repair mechanisms, this can lead to frameshifts or a mutated mRNA sequence, ultimately resulting in degradation of the mRNA by NMD and/or absence of protein expression in the cells (Jiang and Doudna 2017).

To generate a HEK293T HELZ-null cell line, I followed the protocol of Zhang's lab (Ran et al. 2013). The short guide sequence was designed using CHOPCHOP online tool to minimize off-target effects (Montague et al. 2014; Labun et al. 2016; Tycko et al. 2016) and incorporated in the pSpCas9(BB)-2A-Puro-V2.0 vector (Addgene 62988). The sgRNA targets exon 8 of *HELZ* gene and contains an encoded AGG PAM, specific for *Streptococcus pyogenes* Cas9. The sgRNA is transcribed as one entity, i.e. containing both crRNA and tracrRNA. HEK293T cells are transfected twice with the construct as described in the knock-down protocols (see the sub-chapter 3.2.2.2 Knock-down of HELZ and of CNOT1). 24 hours after the second transfection, cells are seeded in a 96-well plate at 3.5 cells/ml. This low concertation ensured that one cell per well was seeded in most wells of the 96-wells plate and thus grew as a single colony of homogenous genotype in selective media (DMEM with 1.5 μg/ml of Puromycin). Once a single colony has grown to about 70% of the surface of the well, it is transferred to a bigger well until sufficient material could by analysed by western

blot using a specific antibody against HELZ (**Figure 4** A). I had several positive cellular clones that were not expressing HELZ protein based on western blot analysis (**Figure 4** A). The genomic DNA (gDNA) was extracted from three clones using the Wizard® SV Genomic DNA Purification System (Promega). *HELZ* exon 8 region was amplified using specific primers (see **Table 1**), the fragments were separated on an agarose gel and extracted. The resulting fragments were inserted in a TOPO-TA cloning vector according to the manufacture's' protocol (TOPO-TA cloning, Thermo Fischer). At least three TOPO-TA plasmids per cellular clones were sequenced. All of them contained mutations leading to frameshifts in the ORF of *HELZ*. I performed my experiments with the B8 clone, which is hereafter referred to as HEK293T HELZ-null cell line. **Figure 4** B displays the alignment of the targeted region in HELZ-null cell lines with the original genomic sequence in HEK293T wild-type cells. 7 nucleotides were deleted in HEK293T HELZ-null cells genome causing a premature stop codon after the potential translation of 159 amino acid residues. No other ORF was identified by computational analyses that could result in protein production (ORF finder, www.ncbi.nlm.nih.gov).

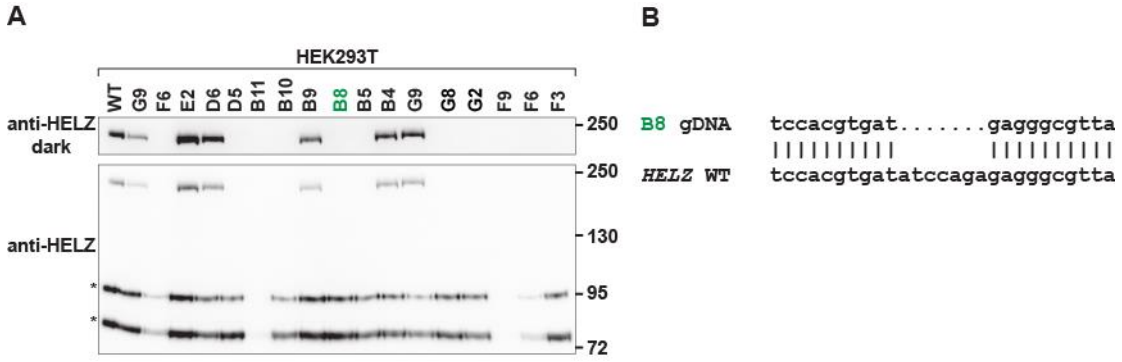

Figure 4 Characterisation of HEK293T HELZ-null cell line. **A** Western blot detecting endogenous HELZ protein in the screened cellular clones that were transfected with pSpCas9(BB) -2A-Puro-HELZ-gRNA-V2. The cellular clonal name of the protein samples are indicated on top of the gel. Clone B8, which is the clone used in this thesis (highlighted in green). In the upper panel (anti-HELZ dark), the signal was uniformly increased to detect weak bands. In the lower panel, non-specific bands detected by the antibody are marked with asterisks. Note that these band are independent of the expression of HELZ protein. **B** Alignment of HEK293T HELZ-null cell (B8) gDNA sequence with wild-type sequence on exon 8 of *HELZ* gene.

3.2.2.4 Immunofluorescence assay

The immunofluorescence assay was used to detect intra-cellular localisation of proteins in fixed cells by microscopy techniques. HeLa cells were seeded at 60 000 cells per well in 24-wells plate on sterilised 12 mm round coverslips (Menzel-Glaser). The next day, pT7-MS2-HA-HELZ and pT7-EGFP-C1-DDX6 were transfected, the latter serving as a P-body marker. Two days after transfection, cells were washed once

with filtered PBS and fixed for 10 minutes with 4% paraformaldehyde (Sigma). After a five-minutes wash with PBS, cells were permeabilized by incubating them 10 minutes with 0.5% Triton X-100 (Merck) diluted in PBS; and washed again three times five minutes with PBS. The primary antibody, anti-HA (**Table 10**), was diluted in PBS containing 10% FCS (Thermo Fischer) and 0.1% Tween20 (Sigma). Samples were incubated with the primary antibody solution for one hour at room temperature and washed three times for five minutes with PBS-0.1% Tween20 (Sigma). Next, the plates were placed in a dark box (to protect the samples from light) for the rest of the procedure. The secondary antibody incubation with anti-mouse IgG coupled with Alexa 594 (Thermo Fischer) diluted in PBS-0.1% Tween20 (**Table 10**) lasted one hour followed by a single five-minutes wash with PBS-0.1% Tween20. The nucleus was stained by incubating the cells for five minutes at room temperature with 1:100 Hoechst reagent (Invitrogen) diluted in PBS, followed by three final washes of five minutes with PBS. The coverslips were mounted on Microscope slides SuperFrost® cut edges (Roth) using Fluoromount-G® (Southern Biotech). Images were taken by the confocal microscope Leica TCS Sp8 with a HC PL APO 63x 1.40 OIL CS2 objective using photodetectors HYD1 488 (for GFP), PMT2 (for Hoechst) and HYD3 594 (for Alexa 594). Images were analysed with ImageJ and processed with Adobe Photoshop.

3.2.3 The tethering assay

The tethering assay allows to study the levels of protein and mRNA of a reporter artificially bound by a specific protein. The assay does not rely on the RNA binding ability of the studied protein, which can be unknown or can complicate the analysis of the function of the protein. The first described tethering assay was performed with the MS2-system and several other systems were quickly developed. In this thesis, I also used the MS2-system which is derived from the bacteriophage MS2 (Bos et al. 2016). A homodimer of the MS2 coat protein binds with a strong affinity (dissociation constant (kd) of 10^{-9} M) to 21 nucleotides forming a stem-loop structure (Bos et al. 2016). The tethering assay takes advantage of the high specificity and affinity of the MS2 coat protein for the MS2 binding sites (MS2bs). A reporter plasmid encodes for a transcript containing MS2bs in the 3'UTR. Another plasmid encodes for the desired protein tagged with the MS2 coat protein. Co-expression of both plasmids enables the MS2-fused protein to directly bind to the reporter with MS2 binding sites in the 3'UTR of the reporter mRNA. As a control, a reporter without the MS2bs is used to identify non-

specific effects of the simple expression of the MS2-fused protein. The reporter systems that I used encode for different ORFs and are referred to as the luciferase reporter and the β-globin reporter systems.

The luciferase reporters code for the *Renilla* luciferase from the sea pansy (*Renilla reniformis*) and the Firefly luciferase from the common eastern firefly (*Photinus pyralis*). There are two *Renilla* luciferase reporters that code for the ORF of the *Renilla* luciferase. One *Renilla* reporter encodes for six MS2bs (R-Luc-6xMS2bs) in the 3'UTR; the other *Renilla* reporter does not contain MS2bs in the 3'UTR (R-Luc). The firefly luciferase reporter (F-Luc) does not contain any MS2bs and is used as a transfection control. Luciferases are enzymes that produce a flash of light when transforming their substrate into a product. This flash of light can be detected by a spectrophotometer and correlates with the amount of luciferase protein translated in the transfected cells. I can thus assess protein levels of the reporters by measuring the luciferase activity. Cells were collected and luciferase assays were performed with the Dual-Luciferase® Reporter Assay system from Promega and measured by the Centro LB 960 Mikroplatten Luminometer (Berthold Technologies). This system uses the advantage that both enzymes (*Renilla* luciferase and Firefly luciferase) have different co-factors to transform their substrate into a product and thus, can be measured from the same sample one after the other. mRNA levels of the reporters were visualised by northern blot and quantified (see sub-chapter 3.2.4.2 Northern blot). *Renilla* luciferase values (protein and mRNA) were normalised to the values obtained for the Firefly luciferase (transfection control).

The β-globin reporters encode for the human β-globin (three exons and two introns). One reporter encodes for a 3'UTR containing 6xMS2bs (β-6xMS2bs). The control reporter codes for the last 480 nucleotides from the glyceraldehyde 3-phopshate dehydrogenase (GAPDH) gene in-frame with the β-globin ORF (β-GAP), and does not encode for MS2bs in its 3'UTR. In the experiments, I co-transfected both reporters and a plasmid encoding for a MS2-tagged protein. I did not detect protein levels but I detected mRNA levels of the β-globin reporters, which were visualised by northern blot and quantified (see sub-chapter 3.2.4.2 Northern blot). The control β-GAP reporter ran slower than the β-6xMS2bs reporter due to its extra nucleotides encoding GAPDH, enabling the separation of the two reporters on a RNA agarose gel.

The values obtained for the β-6xMS2bs mRNA levels were divided by the mRNA levels of the β-GAP to give a fold-change per condition.

The tethered proteins, mainly HELZ and its fragments/mutants, are all fused to MS2-HA at their N-terminus. The MS2-tag is used to tether the proteins. The HA-tag is used to visualise the protein in western blot. MS2-HA is expressed as a control in a separate sample. The ratios of the reporters (for both the luciferase and β-globin reporters) obtained in the control MS2-HA condition were set to 100 to allow an easier comparison of the changes in reporter levels when they were co-expressed with MS2-HA-HELZ-tagged proteins.

3.2.4 RNA-based assays

3.2.4.1 RNA extraction

RNA was extracted from cells with TRIzol™ Reagent (Thermo Fischer) according to the manufacturer's recommendations. Additionally, the suspended RNA samples were treated with 2 U of DNAseI in DNAseI buffer (Thermo Fischer) and in the presence of 20 U of RiboLock RNase Inhibitor (Thermo Fischer) in a total volume of 50 µl for 40 minutes at 37°C. RNase-free MilliQ water was added to a final volume of 200 µl. The solution was mixed with 200 µl of Phenol: Chloroform: Isoamylalcohol (25:24:1) (ITW Reagents) and vortexed for 15 seconds. After spinning the sample for 15 minutes at 4°C at full speed on a table-top centrifuge, the upper-phase was transferred to a new 1.5 ml Eppendorf. The RNA was precipitated by adding 20 µl of sodium acetate 3 M pH 5.2 and 600 µl of ice-cold pure ethanol (Merck), and incubated for at least one hour at -20°C. The RNA was pelleted by centrifugation for 30 minutes at 4°C at full speed in a table-top centrifuge and washed twice with 75% ethanol by centrifugation for 10 minutes at 4°C at full speed in a table-top centrifuge. The pellet was resuspended in 11 µl of RNase-free MilliQ water and incubated 15 minutes at 30°C. 1 µl of the RNA extract was used to measure RNA concentration on the NanoDrop ND-1000 (PeqLab). The ratio of the optical density measured at both 260 nm and 280 nm ($OD_{260/280}$) was used to assess the purity of the RNA: ratios between 1.9 and 2 were defined as pure.

3.2.4.2 Northern blot

Northern blots were performed essentially as described before (Gatfield et al. 2003). Briefly, samples were separated in a 1.2 % (RNA) agarose gel immerged in 1xMOPS buffer (see sub-chapter 3.1.3 Buffers, solutions and reagents) at 40-55 V overnight. Afterwards, the gel was sandwiched between a Rotilabo®-Blotting Paper 0.35 mm thick (Roth, henceforth referred to as Blotting paper) at the bottom and a positively charged nylon membrane (GeneScreen Plus, NEN Life Sciences) on top, both soaked in 10xSSC (see sub-chapter 3.1.3 Buffers, solutions and reagents). Two Blotting papers soaked in 2xSSC were placed on top of the nylon membrane and additional weights were added to allow the samples to blot overnight onto the membrane by osmosis. The transferred RNA on the membrane was crosslinked with UV. The membrane was pre-hybridized in 0.1 g/l of sonicated salmon testis DNA in Church buffer for at least one hour at 65°C in a low speed rotisserie. ^{32}P-labelled (a radioactive phosphorus with an extra neutron, totalling 32 atomic mass units) probes were added to the membrane with the Church buffer and incubated overnight at 65°C at a low speed in a PEQLAB Perfect Blot Hydrisingoven (VWR). The membrane was then washed five times for at least 20 minutes with Northern blot washing buffer (see sub-chapter 3.1.3 Buffers, solutions and reagents), sealed and exposed to a GE Storage Phosphor Screen (Fujifilm). The signal on the screen was detected by the Amersham Typhoon RGB Biomolecular Imager or the STORM 860 Imager from Molecular Dynamics, and quantified with the ImageQuant TL program of GE Healthcare. Images were processed with Adobe Photoshop.

The production of reporter specific ^{32}P-labelled probes was performed by PCR amplification in presence of α-^{32}P-labelled dATP and α-^{32}P-labelled dCTP resulting in the incorporation of the radioactive phosphorus into the final product. The PCR mix and thermic cycles are shown in the following table:

Amount	Reagent
5 µl	10xin-house Taq buffer
1 µl	dNTPs (10 mM dGTC and dTTC, 0.1 mM dATP and dCTP, from Peqlab)
200 ng	Construct specific DNA template
1 µl	Reverse primer (10 µM, see **Table 14**)
Up to 40 µl	MilliQ water

1 µl	In-house Taq polymerase
5 µl	α-^{32}P dATP (10 µCi/µL)
5 µl	α-^{32}P dCTP (10 µCi/µL)

Temperature	**Time**	
94°C	3 min.	
94°C	30 sec.	39X
54°C	30 sec.	
72°C	30 sec.	
72°C	7 min.	
4°C	Until use	

The sequences of the reverse primers are as follows:

Table 14 Reverse primers for ^{32}P-labelled probes

Detection of	**Sequence**
R-Luc reporters	CCCTCGAGTTGTTCATTTTTGAGAACTC
F-Luc	CCGGAATTCTTACAATTTGGACTTTCCGCC
β-globin reporters	TTAGTGATACTTGTGGGCCAGGGC

The PCR product was purified by centrifugation: the volume was topped up with MilliQ water to 200 µl and laid on top of a 1 ml syringe containing at least 700 µl of Sephadex G50 beads (Sigma) on top of 1 cm of Silanized Glass Wool (Sigma). After spinning for 5 minutes at 2 000 rpm in a Multifuge 1L centrifuge Heraeus (Thermo Scientific), the purified PCR product was collected and transferred to the Church buffer with the membrane, or stored at -20°C until use.

3.2.4.3 qPCR

qPCR analysis was performed to validate the mRNA targets identified in the RNA-Seq analysis. To prepare complementary DNA (cDNA) from HEK293T cells and HEK293T HELZ-null cells, total RNA was extracted as described in sub-chapter 3.2.4.1 RNA extraction. The synthesis of the first strand of the cDNA was performed according to RevertAid H Minus Reverse Transcriptase protocol (Thermo Fischer) with random hexamer. The final volume was diluted 1:1 with MilliQ water. The qPCR was performed on either the CFX96 Real-Time System C1000 or the CFX384 Real-Time System C1000 from Bio-Rad in white wells plates (Bio-Rad). The qPCR mix for the 96-wells plate was as follows: 1 µl (or a ten-fold serial dilution down to 10^{-3}) of cDNA, 0.4 µM of

forward and reverse primer (**Table 2**), 10 μl of iTaq™ Universal SYBR® Green Supermix (Bio-Rad) and MilliQ water up to 20 μl. For 384-wells plates, the amounts were scaled down to a final volume of 10 μl. The thermic cycle was as follows: 1) 95°C 2 min, 2) 40 times: 95°C 30 sec, 55°C 30 sec, 73°C 1 min, 3) 73°C 10 min, and 4) heating from 65°C to 95°C by incrementing of +0.5°C every 5 sec. The signal was detected by the CFX Manager™ Software (Bio-rad). Primers were designed with the online tool Primer3 (Koressaar and Remm 2007; Untergasser et al. 2012). All primers were tested for specificity prior to use. qPCR Ct values were analysed with the Pfaffl Method (Pfaffl 2001).

3.2.4.4 Polysome profiling

Polysome profiles were obtained as described before (Kuzuoglu-Ozturk et al. 2016). HEK293T cells and HEK293T HELZ-null cells were seeded at $9x10^6$ cells in 20 ml of media in 14 cm dish. After 24 hours, the media was changed with warm fresh media for at least 2h30 before treating the cells with 50 μg/ml of CHX for 30 minutes. Cells were washed twice with 5 ml of PBS with 50 μg/ml of CHX on ice and scraped in fresh 1.5 ml Eppendorf. After a spin of 1 minute at 4°C at 5000 g in a table-top centrifuge, 200 μl of Polysome profile lysis buffer was used to resuspend the pellet, and the suspended pellet was incubated for 20 minutes on ice. The insoluble fraction was separated from the soluble fraction by spinning the samples for 10 minutes at 4 °C at 10 000 g, the supernatants were transferred to a new 1.5 ml Eppendorf tube. The $OD_{260/280}$ ratios were measured with the NanoDrop ND-1000, and sample volumes were adjusted with Lysis buffer to have an equal OD: this ensure that an equal amount of RNA was loaded on the sucrose gradient. The 10% to 50% sucrose gradient (with 50 μg/ml of CHX) was prepared on the same day in 5 ml Thinwall Polypropylene Tubes for SW 55 Ti rotor (Beckman Coulter) with the Gradient Master (Biocomp). Equal amounts (and volume, as it was adjusted) of samples were placed on top of the sucrose gradients that were then spun at 41 000 rpm for 2 hours at 4°C in the Optima™ L-100XP Utracentrifuge (Beckam Coulter). Afterwards, the gradients were placed in the fractionation instrument (ISCO Dichte Gradient Fraktionator program, Axel Semrau Gmbh). 3M™ Fluorinet Fc-40 (IoLiTec) was pumped pushing the gradient upwards in the detector measuring RNA levels of the samples. The values were exported and analysed in Microsoft Excel.

3.2.4.5 RNA-Sequencing

RNA-Seq was performed for HEK293T cells and HEK293 HELZ-null cell lines. Cells were seeded in 14 cm dish (two dishes per cell line) at a concentration of $7x10^6$ cells in 20 ml of DMEM media. The next day, cells were harvested as described in (Calviello et al. 2016) and stored at -80°C until further use. RNA extraction was performed with the Mini Prep RNAeasy kit (Qiagen). The quality of the RNA was evaluated with the 2100 Bioanalyser RNA Analysis kit (Agilent). cDNA preparations were performed as described in TruSeq® RNA Sample Preparation v2 (Illumina). The steps included purifying and fragmenting the mRNAs, synthesis of the first cDNA strand with random hexamer, digesting the RNA template and synthesising the second cDNA strand, performing end repair to generate blunt end cDNA, adenylating the 3' ends of the blunt fragments, ligating the adapters (A set), enriching the DNA fragments, normalizing the DNA template and pooling the libraries. RNA-Seq libraries were sequenced by the Genome Center of the Max Planck Institute for Developmental Biology, Tübingen with the Illumina Hiseq3000 with cbot instrument. Computational analysis was done by Dr. Eugene Valkov and Ramona Weber (both members of the laboratory) based on the protocol of Prof. Dr. Marcus Landthaler from the Max Delbrück Center for Molecular Medicine in Berlin. Briefly, the reads originated from rRNA were removed with Bowtie2 (Langmead and Salzberg 2012). Remaining reads were mapped onto the human genome using Tophat2 (Kim et al. 2013). Read count analysis was performed using QuasR (Gaidatzis et al. 2015) in order to do the differential expression analysis using edgeR (Robinson et al. 2010; McCarthy et al. 2012).

3.2.5 Protein-based assays

3.2.5.1 Immunoprecipitation assay

The immunoprecipitation assay was used to study the interaction between proteins in cells. A protein of interest is recognized by a specific antibody and the antibody is pulled down by beads. Protein interactors can be analysed by western blot. Immunoprecipitation assays were performed from cellular samples grown in 10 cm dish (see sub-chapter 3.2.2.1 Transfection). Cells were washed once with PBS, scraped in 1 ml of NET buffer supplemented with 10% Glycerol (Sigma) and cOmpleteTM Protease Inhibitor (Roche), and incubated on ice for at least 15 minutes. After spinning the samples for 15 minutes at 4°C at full speed in a table-top centrifuge,

the supernatants were treated with 2 μl of RNaseA 7000 U/ml (Qiagen) for 30 minutes at 4°C and re-spun for 15 minutes at 4°C at full speed on a table-top centrifuge. The input samples (100 μl) were taken from the supernatant and mixed with 100 μl of 2x protein sample buffer. The volume of the inputs (200 μl) was set to 100%: amounts loaded on a SDS-PAGE were calculated based on this number and are indicated in the figure legends. The rest of the supernatants were mixed with 2.5 μl of polyclonal anti-GFP antibody (from rabbit, see **Table 10**) for one hour at 4°C. The anti-GFP antibodies were immunoprecipitated by adding 50 μl/samples of Gammabind G Sepharose® (GE Healthcare), which recognizes the Fc region of IgG derived from mammals. After one hour of incubation at 4°C, the beads were washed three times with NET buffer and suspended in 100 μl of 2x protein sample buffer. The volume of 100 μl of the immunoprecipitated fraction (GFP-IP) was set to 100%: amounts loaded on a SDS-PAGE were calculated based on this number and are indicated in the figure legends.

3.2.5.2 Western blot

The western blot is a common technique used to detect specific proteins in a sample by the use of antibodies. The proteins were denatured by heat by incubating them for 3 to 5 minutes at 95°C prior to loading. Samples were loaded on a SDS-PAGE (polyacrylamide gel electrophoresis) and upon application of an electric current (100 to 200 V), the negatively-charged proteins were separated by size. Polyacrylamide gels are porous and the size of the pores correlates with the percentage of acrylamide in the gel. A protein ladder (Thermo Fischer) with defined proteins sizes ran together with the samples to mark various sizes (in kDa). The second step was to blot/transfer the separated proteins onto a membrane, also using electrophoresis. Effectively, the gel and a nitrocellulose membrane (GVS North America) were held tightly together by the use of Blotting papers and a clamp-like device. This ensemble was immerged in Transfer buffer in a Transfer chamber and an electric current was passed through (typically for 2h30 at 55 V) (**Table 8**). The equal transfer of proteins was visually assessed by Ponceau staining of the membrane, which stained the proteins and was easily washed out with water. The final step of western blot is the detection step. Membranes were incubated in Blocking buffer for one hour to reduce nonspecific binding of antibodies (**Table 8**). Next, membranes were incubated overnight at 4°C with the primary antibody recognising the target protein (**Table 10**). After washing three

times with Washing buffer, the membranes were incubated for one hour at room temperature with the secondary antibody, if necessary. The secondary antibody recognises the IgG of the primary antibody, and is coupled to HRP. After another round of washes, the membranes were incubated for 3 minutes with the Detection solution and the chemiluminescence was detected by the Fusion FX Imager from Vilber Lourmat. Images were processed with Adobe Photoshop.

3.2.5.3 Recombinant protein expression

Sub-complexes of the CCR4-NOT and MBP were purified by a previous PhD candidate in the lab, Tobias Raisch, and a detailed protocol can be found in (Chen et al. 2014) and (Sgromo et al. 2017). The composition of the sub-complexes used are described in **Table 15**.

Table 15 Purified sub-complexes of human CCR4-NOT

Name	Composition
CNOT1_N/10/11	CNOT1_N (residues 1-1000), CNOT10 (residues 25-707), CNOT11 (residues 257-498)
Pentameric	CNOT1 (residues 1093-2371), CNOT7, CNOT9 (residues 19-285), CNOT2 (residues 344-540) and CNOT3 (residues 607-753)
CNOT7+CNOT1	CNOT1 (residues 1099-1317), CNOT7
CNOT9 module	CNOT1 (residues 1356-1588), CNOT9 ARM (19-285)
CD - Connector Domain	CNOT1 (residues 1607-1815)
NOT module	CNOT1 (residues 1829-2361), CNOT2 (353-540), CNOT3 (607-748)

3.2.5.3.1 Recombinant protein expression from Sf21 insect cells

I describe the optimised protocol for the expression of 2xSTREP-HELZ_C-MBP in *Sf21* insect cells. The protocol is derived from (Berger et al. 2013).

The plasmid pLIB-2xSTREP-HELZ 1051 to 1942 (C)-MBP was transformed into *E. coli* DH10EMBACY. This bacterial strain contains the recipient bacmid to create a baculovirus. Transformed bacterial cells were plated on an LB plate containing appropriate antibiotics and coated with 2 mg of BluoGal diluted in DMSO (Goldbio) and 50 µL of 100 mM IPTG (Roth) (see **Table 11** and **Table 12**). IPTG activates the integration of the insert in the bacmid via a Tn7 transposon site. BluoGal allowed blue/white screening of colonies containing positive integration of the insert into the bacmid. Positive colonies were white where LacZa was disrupted by the integration of the insert. Positive colonies were picked and grown in 3 ml LB media at 37°C overnight.

The baculovirus genome was extracted with the QIAGEN® Plasmid Mini kit (replacing the column-based purification steps with ethanol precipitation). Semi-adherent *Sf21* cells were seeded in a 6-wells plate in 3 ml of Sf-900™ II (Gibco). Once the concentration of *Sf21* cells reached 10^6 cells/ml, the baculovirus DNA was transfected in two wells with FuGENE HD Transfection Reagent (Promega) according to the manufacturer's protocol. Positively transfected cells were visualised with an inverted microscope detecting YFP (the yellow fluorescent protein is coded in the baculovirus genome). After 60 hours, the supernatant (two times 3 ml) containing low titre virus (V_0) was collected and kept at 4°C. The transfected cells were collected two days after media-change and ran in a SDS-PAGE stained with Coomassie blue to visualize the expression of the protein. To obtain high titre V_1 baculovirus, 3 ml of V_0 were incubated with 10^6 *Sf21* cells/ml in suspension (in a horizontal rotation of 2.5 cm orbit at 120 rpm) in a final volume of 28 ml. The next day, 22 ml of media were added to the culture. After 24 hours, the cells stopped growing, as assessed by measuring the density of the cells. That day was called the day of proliferation arrest (DPA) and it meant that the cells were infected and produced high titre viruses. At DPA+24 hours, the supernatant containing V_1 was collected. At DPA+72 hours, the cells were lysed by brief sonication and protein expression checked on a SDS-PAGE. To get lysates with HELZ expressed, I infected 50 ml of $2x10^6$ *Sf21* cells/ml with 0.5 ml of V_1 on day 1. The next day, cells had grown to around 3-4 $x10^6$ *Sf21* cells/ml, I diluted the culture back to 2 $x10^6$ *Sf21* cells/ml. On day 3, cells were at DPA and I harvested them at DPA+48 hours by spinning them down gently for 5 minutes. Pellets were stored at -20°C until use. Cells were lysed in 5 ml of Pulldown buffer for insect cells (see sub-chapter 3.2.5.4 In *vitro MBP* pulldown) and sonicated briefly. The lysates were spun for 15 minutes at 12 000 g at 4°C and the supernatants were filtered through a 0.45 μm Millex-HV (Merck) filter and used for pulldowns.

3.2.5.3.2 Recombinant protein expression in E. coli cells

I optimised a protocol to express and extract recombinant HELZ_C fragments with an MBP-tag at the N-terminus and a GB1-6xHis-tag at the C-terminus from bacteria.

Bacterial expression vectors containing the gene of interest were transformed into *E. coli* BL21 Star cells and plated on 10 cm plates with appropriate antibiotic (see **Table 11** and **Table 12**). Several positive colonies (from one plate) were grown

together in LB media at 37°C in suspension at 180 rpm in a Multitron Standard (Infors-HT). For MBP-HELZ_C1 and MBP-HELZ_C2, the transformed colonies were incubated in 6 litres cultures (big batch culture). Whereas for MBP-HELZ_C2 fragments, the transformed colonies were incubated in 100 ml cultures (small batch culture). When the OD_{600} reached 0.4, 0.5 mM of IPTG (Roth) was added to the cultures to induce protein expression. After an hour, the temperature of the incubator was reduced to 20°C. The expression of the recombinant protein was repressed by LacI. IPTG sequester LacI thus alleviating the repression of the recombinant protein. The cultures were incubated overnight. Cells were pelleted by spinning them at 6000g. Pellets were resuspended in Ni-lysis buffer (see **Table 8**) and lysed by sonication or by a French press. From this step on, all handlings were performed at 4°C. Big batch cultures were spun for one hour at 41 000 rpm in the Optima™ L-100XP Utracentrifuge (Beckman Coulter). Small batch cultures were spun for 15 minutes at 9900g for 15 minutes. Both big batch and small batch supernatants were filtered through a 0.45 μm filter Millex-HV (Merck). Big batch cleared and soluble fractions were loaded into an equilibrated 1 ml Ni^{2+} HiTrap™ iMAC HP (GE Healthcare) column. After an initial washing step, proteins were eluted in Ni-lysis buffer supplemented with 500 mM imidazole (Merck) in an Äkta Pure (GE Healthcare). Proteins were detected at 280 nm and the fractions of the single peak were analysed on a SDS-PAGE and pooled. Small batch cleared and soluble fractions were mixed one hour at 4°C with 1 ml of washed and equilibrated Protino® Ni-NTA Agarose beads (Macherey Nagel). After washing by gravity flow, proteins were eluted in 5 ml of Ni-Lysis buffer supplemented with 500 mM imidazole (Merck), and used for pulldowns.

3.2.5.4 *In vitro* MBP pulldown

The amount mixed for the MBP pulldowns are indicated in **Table 16**. For input samples, protein mixes were adjusted with pulldown buffer to a final volume of 20 μl for HELZ_C and 40 μl for HELZ_C fragments (unless otherwise stated in figure legend) (**Table 8**), and mixed with 2x protein sample buffer at a ratio of 1:1. The total input volumes are arbitrarily set to 100%: amounts loaded on a SDS-PAGE were calculated based on this number and indicated in figure legends. For pulldown samples, protein mixes were adjusted to a final volume of 1 ml in a 1.5 ml Eppendorf. For full-length HELZ_C, protein mixes were incubated with 100 μl of washed and equilibrated Amylose Resin beads (BioLabs) for 1 hour at 4°C on a rotating wheel. For HELZ_C

fragments, protein mixes were incubated for 1 hour at 4°C on a rotating wheel, then 50 µl of washed and equilibrated Amylose Resin beads (BioLAbs) were added to the mix and incubated for one more hour at 4°C on a rotating wheel. For both set-up, samples were washed five times with Washing buffer after the incubation step with beads (**Table 8**). Proteins were eluted from the beads with 100 µl of Elution buffer for 30 minutes at 4°C. The supernatants were transferred to new Eppendorfs and the proteins were precipitated with 100 µl of ice-cold TCA 20% on ice for 30 minutes. Samples were spun for 30 minutes at 4°C at full speed in a table-top centrifuge. Pellets were resuspended in 35 µl of 1x protein sample buffer (unless otherwise stated in the figure legend). This final volume was set to 100%: amounts loaded on a SDS-PAGE were calculated based on this number and indicated in figure legends. Samples were analysed by SDS-PAGE. Gels were stained with Coomassie blue overnight, washed with Coomassie destaining solution and scanned. Images were processed with Adobe Photoshop.

Table 16 Amounts of recombinant and purified proteins in pulldown assays

Protein	**Input**	**Pulldown**
MBP	5 µg	20 µg
HELZ_C	20 µl	1000 µl
HELZ_C1	125 µg	500 µg
HELZ_C2	125 µg	500 µg
CNOT1_N/10/11	15 µg	165 µg
pentameric	20 µg with HELZ_C, 30 µg with HELZ_C1 and HELZ_C2	220 µg with HELZ_C, 120 µg with HELZ_C1 and HELZ_C2
CNOT7+CNOT1	12.5 µg	50 µg
CNOT9 module	12.5 µg	50 µg
CD domain	9 µg	36 µg
NOT module	20 µg	80 µg
HELZ_C2a	16 µl	800 µl
HELZ_C2aΔ1	1 µl	45 µl
HELZ_C2b	2 µl	100 µl
HELZ_C2bΔ2	2 µl	100 µl
HELZ_C2c	20 µl	1000 µl
HELZ_C2cΔ3	16 µl	800 µl
HELZ_C2cΔ4	2 µl	100 µl

4 Results

4.1 HELZ and mRNA stability

In the first chapter of the Results section, I present my data describing the effect of HELZ when bound to the 3'UTR of reporter mRNAs by the use of the tethering assay in human cells. HELZ likely has a role in cytoplasmic mRNA regulation as 1) HELZ was reported to interact with PABPC1 in HeLa cell extracts, which binds to the 3'polyA tail of mRNAs (Hasgall et al. 2011). 2) HELZ was reported to interact with the scaffold protein CNOT1 of the CCR4-NOT deadenylase complex in HEK293 cells (Mathys et al. 2014). Deadenylation is the first rate-limiting step of cytoplasmic mRNA decay pathways. The main deadenylase complex in human cells is the CCR4-NOT complex. 3) HELZ was identified by mass spectrometry as a potential interactor of the RNA helicase and translational repressor DDX6 (Ayache et al. 2015). 4) Although it is not known if HELZ binds to RNA, it contains several features that provide RNA-binding ability (via its CCCH-type ZnF and/or UPF1-like helicase domain), or alternatively it might be recruited indirectly by interacting with PAPBC1-bound mRNAs in the cytoplasm (**Figure 5** A). Some results in this section are included in a manuscript, ready for submission.

4.1.1 The tethering of HELZ and HELZ_C reduces the mRNA levels of bound reporters

I performed tethering assays to quantify and study the potential effect of HELZ when bound to the 3'UTR of reporter mRNAs. This assay by-passes the RNA-binding ability of the studied protein by artificially tethering the MS2-tagged fusion protein to a reporter mRNA containing six MS2-binding sites in the 3'UTR (see sub-chapter 3.2.3 The tethering assay). With this assay, I can evaluate how translation (by measuring protein levels) and mRNA stability (by measuring mRNA levels) of a reporter are affected by the binding of HELZ to its 3'UTR. The luciferase reporter system consists of three constructs. First, a plasmid coding for the Firefly luciferase (F-Luc) which is used as a transfection control. Second, two plasmids encoding for the *Renilla* luciferase: one with 6xMS2bs in the 3'UTR (R-Luc-6xMS2bs) and one without MS2bs in the 3'UTR (R-Luc). HELZ was expressed in HEK293T cells with a N-terminal MS2-HA-tag (see sub-chapter 3.2.2.1 Transfection and **Table 13**). The transfection mixture

contained, per sample, the F-Luc plasmid, the MS2-HA-HELZ plasmid (or the control MS2-HA plasmid), and either one of the R-Luc reporters. MS2-HA-HELZ was tethered to the R-Luc-6xMS2bs mRNA reporter when they were co-expressed in cells. The R-Luc reporter was used to detect nonspecific binding of the MS2-HA-tagged protein on the reporter. *Renilla* luciferase values were normalised to the values of the co-transfected F-Luc (R-Luc/F-Luc or R-Luc-6xMS2bs/F-Luc), and the ratios were normalised to the corresponding ratio obtained in presence of the control MS2-HA reporter, which was set, arbitrarily, to 100.

Interestingly, the tethering of MS2-HA-HELZ caused a reduction of the protein levels of R-Luc-6xMS2bs (**Figure 5** B blue bars). This reduction in protein levels can be explained by a similar decrease of the mRNA levels of the R-Luc-6xMS2bs reporter when MS2-HA-HELZ was tethered (**Figure 5** B and C lane 2). The R-Luc reporter (protein and mRNA levels) was not affected by the co-expression of MS2-HA-HELZ: R-Luc values obtained in presence of MS2-HA were comparable to the R-Luc values in presence of MS2-HA-HELZ (**Figure 5** D and E lane 1 vs 2). Thus, HELZ negatively affects mRNA levels as well as protein expression when bound to a reporter mRNA.

To get some insight into the mechanism exploited by HELZ to induce the observed decrease of the levels of the tethered reporter, I asked which part of HELZ is responsible for this effect. I constructed two plasmids coding for either the structured N-terminal region (HELZ_N or N, residues 1-1050) or the C-terminal region (HELZ_C or C, residues 1051-1942), both with a N-terminus MS2-HA-tag to be expressed in human cell lines (**Figure 5** A). HELZ_N contains the CCCH-type ZnF and the putative helicase domain; whereas HELZ_C contains the PAM2 motif and the unstructured C-terminal tail. Surprisingly, the tethering of MS2-HA-HELZ_C decreased the protein levels of the R-Luc-6xMS2bs reporter (**Figure 5** B blue bars). Moreover, the decrease in protein levels was accompanied by a decrease in the mRNA levels of the R-Luc-6xMS2bs reporter, similarly to the tethering of MS2-HA-HELZ full-length on the same reporter (**Figure 5** B and C lane 4). In contrast, the tethering of MS2-HA-HELZ_N affected neither protein nor mRNA levels of the R-Luc-6xMS2bs reporter (**Figure 5** B and C lane 3). The expression of MS2-HA-HELZ_N and MS2-HA-HELZ_C did not affect the control R-Luc reporter levels (**Figure 5** D and E lanes 3 and 4). MS2-HA-HELZ, -N and -C were expressed at similar levels in HEK293T cells (**Figure 5** F). Thus, the C-terminal tail can recapitulate the repression effect of the R-Luc-6xMS2bs reporter that is also observed in the presence of full-length HELZ.

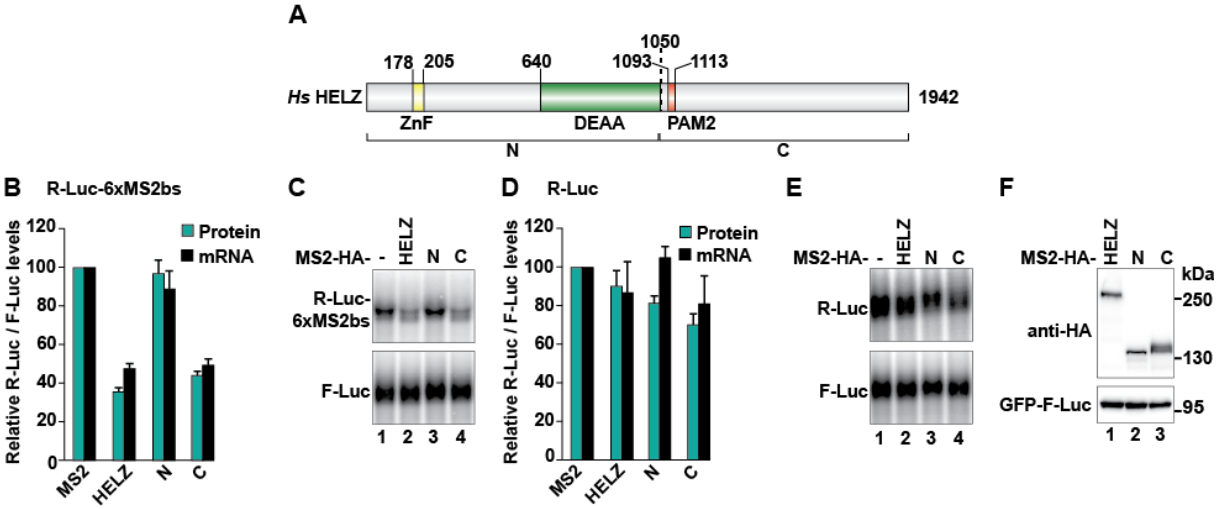

Figure 5 The tethering of HELZ and HELZ_C reduces protein expression and mRNA levels of bound R-Luc-6xMS2bs reporter. **A** Schematic representation of *Hs* HELZ. The conserved zinc finger (ZnF), helicase domain (DEAA) and PABPC1 interacting motif 2 (PAM2) are highlighted in yellow, green, and orange, respectively. Predicted unstructured regions are in grey. Domain boundaries are indicated on top of the graph via their respective amino acid numbers. N and C fragments are indicated as well, the dotted line marks the border in the scheme. **B** Tethering assay of MS2-HA-HELZ full-length and fragments on R-Luc-6xMS2bs reporter in HEK293T cells. The graph shows the quantification of protein (blue bars) and mRNA (black bars) levels of R-Luc-6xMS2bs normalised to F-Luc levels, and set to 100 for MS2-HA; mean values +/- standard deviation (SD) are shown for three independent experiments. **C** Representative northern blot of samples quantified in B. **D** Co-expression of MS2-HA-HELZ and fragments with control R-Luc reporter in HEK293T cells. The graph shows the quantification of protein (blue bars) and mRNA (black bars) levels of R-Luc normalised to the levels of F-Luc, and set to 100 for MS2-HA; mean values +/- SD of four independent experiments are shown. **E** Representative northern blot of samples quantified in D. **F** Expression levels of MS2-HA-HELZ and its fragments assessed by western blot from samples quantified in B. GFP-F-Luc was used as a transfection control.

The luciferase reporter system allows the simultaneous detection of changes at protein and mRNA levels. However, since they code for non-endogenous ORF, I also studied the effect of the tethering of HELZ on a reporter system derived from the human β-globin gene (see sub-chapter 3.2.3 The tethering assay). In this system, two reporter mRNAs were concomitantly transfected together with MS2-HA-HELZ. Specifically, one of the reporter contains 6xMS2bs in its 3'UTR (β-6xMS2bs). The other reporter (β-GAP) does not contain 6xMS2bs in its 3'UTR, and is used as a control. To differentiate both reporters on a northern blot, the control β-GAP plasmid contains an additional sequence of the N-terminal region of the human GAPDH protein, in-frame with the β-globin ORF (see sub-chapter 3.2.3 The tethering assay). Thus, the mRNA of the control β-GAP reporter runs slower on an agarose gel compared to the mRNA of the β-6xMS2bs reporter. β-globin reporters were only detected at mRNA levels. The mRNA

levels of the β-6xMS2bs reporter were divided by the mRNA levels of the β-GAP reporter and. the ratios were normalized to the ratio obtained in presence of the control MS2-HA tether, which was set arbitrarily to 100.

Using this reporter system, I could determine that the tethering of MS2-HA-HELZ and MS2-HA-HELZ_C decreased specifically the mRNA levels of the β-6xMS2bs reporter (**Figure 6** A and B lanes 2 and 4). The tethering of MS2-HA-HELZ_N did not change the mRNA levels of the β-6xMS2bs reporter (**Figure 6** A and B lane 3). The mRNA levels of the β-GAP reporter were not affected when co-transfected with any of the MS2-HA-HELZ constructs (**Figure 6** B). These results are similar to those observed with the luciferase reporter system. Thus, based on results presented in **Figure 5** and **Figure 6**, I conclude that the tethering of HELZ and HELZ_C reduces the mRNA levels of bound reporters, independently of the system used. Furthermore, the ZnF and the putative helicase domain within the N-terminal region of HELZ do not contribute to the effect of HELZ in tethering assays.

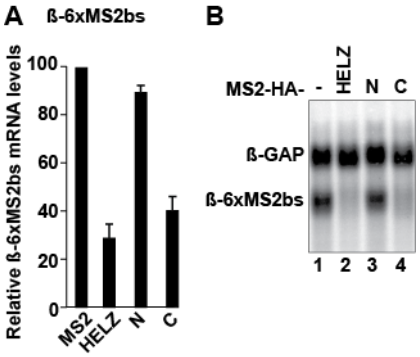

Figure 6 HELZ and HELZ_C reduces mRNA levels of bound ß-6xMS2bs reporter. **A** Tethering assay of MS2-HA-HELZ and fragments using the ß-globin reporters system in HEK293T cells. The graph shows the quantification of the mRNA levels of the ß-6xMS2bs reporter normalized to the mRNA levels of the control ß-GAP reporter and set to 100 for MS2-HA; the mean values +/- SD for four independent experiments are shown. **B** Representative northern blot of samples quantified in A.

4.1.2 PABPC1-binding is dispensable for HELZ to reduce mRNA levels of bound reporters

Since the tethering of HELZ and HELZ_C causes a decrease in mRNA levels of reporters, I set out to identify what was the reason behind this effect and aimed to have a closer look at the C-terminal tail of HELZ. Based on Quick2D analysis (Zimmermann et al. 2018), HELZ_C is unstructured and contains the PAM2 motif, which is conserved among metazoans HELZ. HELZ is the only helicase so far to have an identified PAM2 motif (Albrecht and Lengauer 2004).

PAM2 motifs confer binding to the MLLE domain of PABPC1, and indeed it was shown that HELZ can interact with PABPC1 from human cell extracts (Hasgall et al. 2011). PABPC1 binds to the 3'polyA tail of mRNAs in the cytoplasm and is a key factor in translation and in mRNA stability (Eliseeva et al. 2013). Thus, I tested if PAPBC1 contributed to the repression of reporters induced by HELZ. To do this, I disrupted the binding of HELZ to PABPC1 by introducing a point mutation in the PAM2 motif, and tested the mutant in tethering assay. The phenylalanine at position 1107 in HELZ, conserved in all PAM2 motifs identified, was mutated to a valine (F1107V, **Figure 7** A). Based on published structures of PAM2 bound to PABPC1, this mutation is sufficient to destabilise the aromatic stacking that are important for the binding to the MLLE domain of PABPC1 (Albrecht and Lengauer 2004; Kozlov et al. 2010).

I performed GFP-immunoprecipitation assays in HEK293T cells to visualize and confirm the lack of interaction between HELZ_F1107V and PABPC1. Effectively, GFP-tagged HELZ, GFP-HELZ_F1107V, and GFP-MBP as a negative control, were expressed in different samples in HEK293T cells. The cell lysates were treated with RNaseA prior to the GFP-IP to break any RNA-based interactions. The samples were mixed with GFP antibody and immunoprecipitated with Gammabind G Sepharose beads (see sub-chapter 3.2.5.1 Immunoprecipitation assay). The input fractions, corresponding to samples before the GFP-immunoprecipitation assay, and the GFP-IP samples were analysed by western blot with specific antibodies. In **Figure 7** B, GFP-MBP (lanes 1 and 4), GFP-HELZ (lanes 2 and 5), and GFP-HELZ_F1107V (lanes 3 and 6) are detected in both input and GFP-IP fractions by the use of an anti-GFP antibody. GFP-HELZ interacted with endogenous PABPC1 (**Figure 7** B lane 5), thus confirming a previous report (Hasgall et al. 2011). As expected, introducing the F1107V mutation in the PAM2 motif disrupted this interaction (**Figure 7** B lane 6). Therefore, HELZ_F1107V does not interact with PABPC1 and HELZ binds to PABPC1 only via it PAM2 motif.

Furthermore, HELZ was previously shown to interact with CNOT1 in HEK293 cells (Mathys et al. 2014). I detected this interaction when GFP-HELZ was immunoprecipitated (**Figure 7** B lane 5). Mutating the PAM2 motif did not affect the ability of HELZ to interact with endogenous CNOT1 (**Figure 7** B lane 6).

This indicates that HELZ interacts with CNOT1 independently of PABPC1 and that the point mutation introduced does not affect the overall fold of the protein.

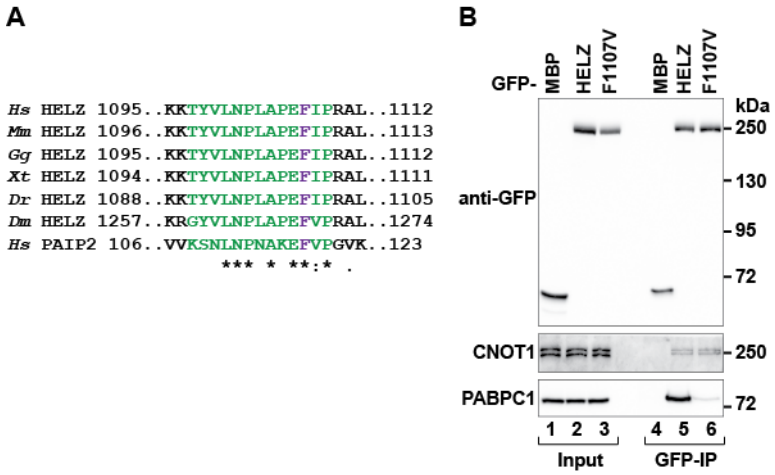

Figure 7 HELZ_F1107V mutant does not interact with PABPC1 in HEK293T cells. **A** Alignment of PAM2 motifs in HELZ proteins from representative species and in the human PABP Interacting Protein 2 (PAIP2). PAM2 motifs are highlighted in green and the characteristic phenylalanine (F) is in purple, amino acid residues numbers are indicated for the first and last residues. Stars mark strictly conserved amino acid residues within the motif, colon and dot mark amino acid residues from the same class of amino acid. *Mm* stands for *Mus musculus*; *Gg*, *Gallus gallus*; *Xt*, *Xenopus tropicalis*; *Dr*, *Danio rerio*; *Dm*, *Drosophila melanogaster*. **B** GFP-immunoprecipitation assay in HEK293T cells displaying the interaction between GFP-HELZ and GFP-HELZ_F1107V with endogenous CNOT1 and PABPC1. GFP-MBP was used as a control. To detect GFP-tagged proteins, 1.2% of input and 20% of the GFP-IP were loaded on a SDS-PAGE. To detect endogenous PABPC1 and CNOT1, 1.2% of input and 35% of the GFP-IP were loaded on a SDS-PAGE. Lysates were treated with RNaseA prior to the assay.

Next, I performed tethering assays with MS2-HA-HELZ_F1107V to test if PABPC1 was contributing to the ability of HELZ to repress the mRNA expression of reporters. MS2-HA-HELZ_F1107V reduced protein levels and mRNA levels of the R-Luc-6xMS2bs reporter (**Figure 8** A and B lane 3), without affecting the levels of the control R-Luc reporter (**Figure 8** C and D lane 3). MS2-HA-HELZ_F1107V also reduced the mRNA levels of the β-6xMS2bs reporter without affecting the control β-GAP reporter (**Figure 8** F and G). In both reporter systems, the reduction of protein and mRNA levels of R-Luc-6xMS2bs and β-6xMS2bs were similar to those observed when MS2-HA-HELZ was tethered. Both proteins were expressed at similar levels (**Figure 8** E). Thus, PABPC1-binding is dispensable for HELZ to reduce the mRNA levels of artificially bound reporters. However, PABPC1 binding might be required to recruit HELZ to endogenous mRNAs targets.

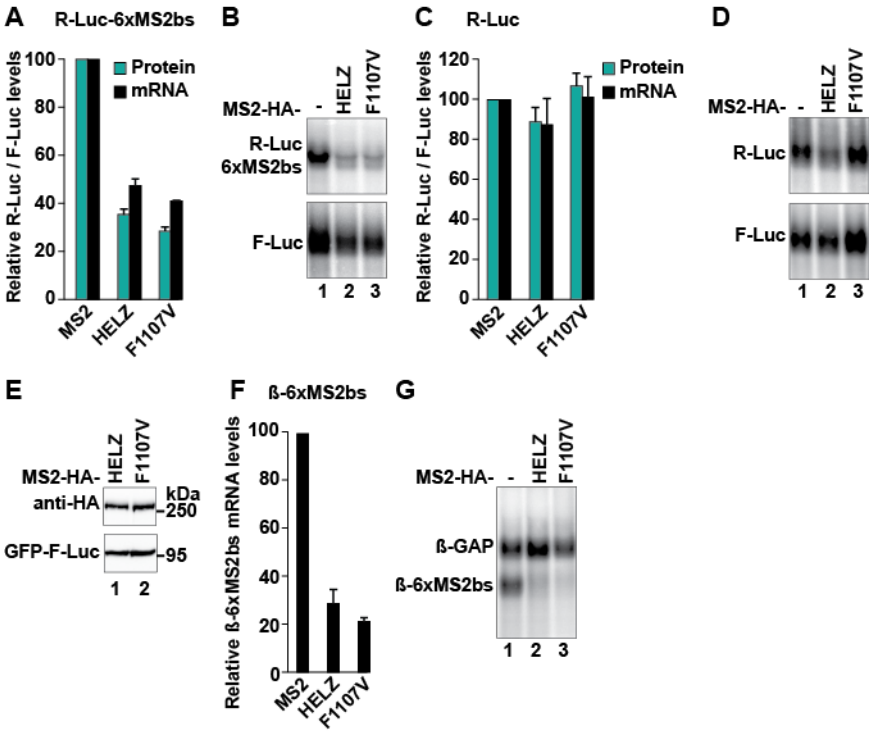

Figure 8 PABPC1-binding is dispensable for HELZ to repress the expression of bound reporters. **A** Tethering assay of MS2-HA-HELZ and MS2-HA-HELZ_F1107V to the R-Luc-6xMS2bs reporter in HEK293T cells. The graph shows the quantification of protein (blue bars) and mRNA (black bars) levels of R-Luc-6xMS2bs normalized to F-Luc levels and set to 100 for MS2-HA; the mean +/- SD are represented for three independent experiments. **B** Representative northern blot of samples shown in A. **C** Co-expression of MS2-HA-HELZ and MS2-HA-HELZ_F1107V with R-Luc control reporter in HEK293T cells. Graph shows the quantification of protein (blue bars) and mRNA (black bars) levels of R-Luc normalized to F-Luc levels and set to 100 for MS2-HA; mean values +/- SD are represented for four independent experiments. **D** Representative northern blot of samples quantified in C. **E** Expression levels of MS2-HA-HELZ and MS2-HA-HELZ_F1107V assessed by western blot in samples from A. GFP-F-Luc served as a transfection control. **F** Tethering assay of MS2-HA-HELZ and MS2-HA-HELZ_F1107V with the ß-globin reporters in HEK293T cells. Graph shows the quantification of mRNA levels of ß-6xMS2bs normalised to the control ß-GAP reporter and set to 100 for MS2-HA; mean values +/- SD are shown for four independent experiments. **G** Representative northern blot of samples shown in E.

4.1.3 HELZ induces reporter mRNA decay via the 5'-to-3' pathway

Reduction of mRNA abundance can be caused by either a decrease in transcription or an increase in the decay rate of the reporter mRNA. HELZ has been found in complexes with both RNA polymerase II and CNOT1 of the CCR4-NOT deadenylase complex, therefore it could play a role in both transcription and mRNA decay (Hamamoto et al. 2004; Diehl et al. 2010; Mathys et al. 2014). However, as demonstrated in the previous experiments, the overexpression of HELZ caused a reduction of mRNA levels only when it was tethered to reporters (R-Luc-6xMS2bs and

β-6xMS2bs); thus, requiring binding to the mRNA of the reporter. Control reporters (R-Luc, F-Luc and β-GAP) mRNA levels were not affected by the overexpression of HELZ (**Figure 5** D and E, **Figure 6** A and B). This indicates that the decrease in mRNA levels observed upon HELZ tethering is likely caused by an increase in the decay of the tethered reporter mRNA. In contrast, an inhibitory effect of HELZ overexpression on transcription should have been detectable independently of tethering; hence would also affect control reporters, which was not the case (also I compared the values obtained in presence of HELZ with the values in presence of the control MS2-HA). Consequently, I decided to analyse the mode of mRNA decay induced by HELZ in more details.

To address this question, HELZ was tethered to a reporter mRNA in cells where the CCR4-NOT complex deadenylase activity was blocked by a dominant negative approach. To block deadenylation, a construct coding for a catalytically inactive mutant of CNOT7 (GFP-CNOT7* (D40E/E45A)) and a construct coding for the middle region of CNOT1 (GFP-CNOT1_M (residues 1085-1605)) were co-expressed in HEK293T cells. The expression of CNOT7* should block the CCR4-NOT deadenylation activity in a dominant negative manner. The expression of CNOT1_M should compete with endogenous CNOT1 for binding to the catalytic subunit, without affecting the other regulatory subunits of the complex. Combined, the co-expression of these constructs should impair the activity of the CCR4-NOT complex in cells. As a control, GFP-MBP was expressed to comparable levels as both GFP-CNOT7* and GFP-CNOT1_M in control samples (**Figure 9** A). As a positive control for the assay, a construct expressing MS2-HA-TNRC6A silencing domain (TNRC6ASD) was co-expressed with the luciferase reporters in both control and deadenylation block conditions. TNRC6ASD is known to trigger 5'-to-3' mRNA decay of miRNA-targeted mRNAs and when it is tethered to reporter mRNAs (Chen et al. 2009; Lazzaretti et al. 2009). In **Figure 9** B and C lane 2, the tethering of MS2-HA-TNRC6ASD induced the degradation of the R-Luc-6xMS2bs in the control GFP-MBP condition. Inhibition of CCR4-NOT-dependent deadenylation by co-expression of GFP-CNOT7* and GFP-CNOT1_M severely impaired TNRC6ASD in its ability to induce mRNA decay of the R-Luc-6xMS2bs reporter (**Figure 9** B and C lane 6). Thus, co-expressing CNOT7* with CNOT1_M effectively blocks deadenylation of the CCR4-NOT complex.

Next, MS2-HA-HELZ was tethered to the R-Luc-6xMS2bs in both control and deadenylation block conditions. Similar to TNRC6ASD, the decrease of mRNA levels

of the R-Luc-6xMS2bs reporter induced by HELZ was impaired when the deadenylase activity of the CCR4-NOT complex was blocked (**Figure 9** B and C lanes 3 vs 6). Furthermore, the same results were obtained with the β-globin reporter system: HELZ-induced decrease of the mRNA levels of the β-6xMS2bs reporter was also impaired when deadenylation was blocked (**Figure 9** D and E lanes 2 vs 4). Thus, not only do these experiments support the finding that HELZ induces mRNA decay of bound reporters, but they also demonstrate that HELZ requires the deadenylase activity of the CCR4-NOT complex to do so (**Figure 9**).

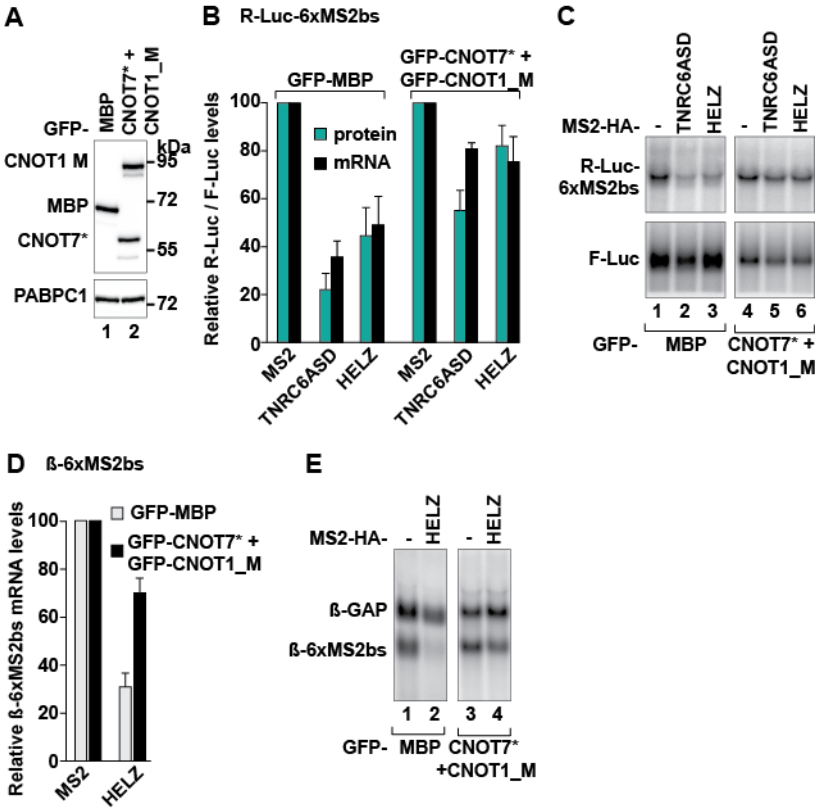

Figure 9 Blocking deadenylation by the CCR4-NOT complex impairs HELZ-induced mRNA degradation of bound reporters**. A** Western blot comparing the expression levels of control samples expressing GFP-MBP and samples co-expressing GFP-CNOT7* and GFP-CNOT1_M, in HEK293T cells. PABPC1 was used as a loading control. **B** Deadenylation block assay using the luciferase reporters system. MS2-HA-TNRC6ASD served as a positive control for the assay. The graph illustrates the quantification of protein (blue bars) and mRNA (black bars) levels of R-Luc-6xMS2bs normalised to F-Luc values, and set to 100 for MS2-HA in the control situation (GFP-MBP) and in deadenylation-block condition (GFP-CNOT7* + GFP-CNOT1_M); mean values +/- SD are shown for three independent experiments. **C** Representative northern blot of samples quantified in B. **D** Similar to B but using the ß-globin reporter system. The graph illustrates the quantification of mRNA levels of ß-6xMS2bs normalized to ß-GAP levels, and set to 100 for MS2-HA in the control situation (grey bars) and in the deadenylation-block condition (black bars); mean values +/- SD are represented for three independent experiments. **E** Representative northern blot of samples quantified in D.

After identifying that HELZ requires the CCR4-NOT complex to induce mRNA decay of reporter mRNA, I tested whether HELZ-induced mRNA decay via the 5'-to-3' mRNA decay pathway. In the 5'-to-3' mRNA decay pathway, deadenylation of the targeted mRNA is followed by decapping by the DCP1/DCP2 complex and subsequent degradation of the mRNA body by the 5'-to-3' exonuclease by XRN1 (Labno et al. 2016). The overexpression of a catalytically inactive mutant of the decapping enzyme DCP2 (GFP-DCP2* (E148Q)) efficiently blocks decapping in HEK293T cells (Wang et al. 2002; Chang et al. 2014a); the expression of GFP alone was used as a control in separate samples (**Figure 10** A). The tethering of MS2-HA-CNOT1 was used as a positive control for decapping-dependent mRNA decay in tethering assays (Kuzuoglu-Ozturk et al. 2016). Block of decapping by the overexpression of GFP-DCP2* impaired the ability of both MS2-HA-HELZ and MS2-HA-CNOT1 to induce reporter mRNA decay (**Figure 10** B and C, GFP-DCP2* condition). Consequently, the mRNA levels of the tethered reporter accumulated in a deadenylated form, seen on the northern blot as a fast-migrating band (**Figure 10** C lanes 5 and 6). Most likely, the deadenylated bands were not degraded by XRN1. In contrast, in the presence of GFP, MS2-HA-HELZ and MS2-HA-CNOT1 induced the degradation of the reporter mRNA (**Figure 10** B and C, GFP condition). Therefore, HELZ-induced mRNA decay in tethering assay requires CCR4-NOT-dependent deadenylation followed by DCP1/DCP2-dependent decapping of the mRNA; hence, HELZ induces 5'-to-3' mRNA decay when tethered to the 3'UTR of a reporter mRNA.

Very little is known about the interaction network of HELZ. Since HELZ induces 5'-to-3' mRNA decay in tethering, I investigated if it could interact with other factors of this pathway. I performed co-GFP-immunoprecipitation assays where GFP-HELZ is co-expressed with various HA-tagged decay factors in HEK293T cells. In these assays, GFP-HELZ efficiently associated with HA-PAN3 of the PAN2/PAN3 deadenylation complex, HA-EDC4 of the decapping complex and HA-PatL1 (**Figure 10** D to F, lanes 4). In contrast, GFP-HELZ could not bind to HA-DDX6, HA-DCP1a, HA-DCP2 nor HA-XRN1 (**Figure 10** F to I, lanes 4). The lack of interaction with DDX6 was unexpected as it was observed in mass spectrometry analysis (Ayache et al. 2015). However, as HELZ interacts with various 5'-to-3' mRNA decay factors, DDX6 might be indirectly associated with HELZ. These results also highlight potential redundant ways of recruiting the 5'-to-3' mRNA decay pathway to targeted mRNAs by HELZ.

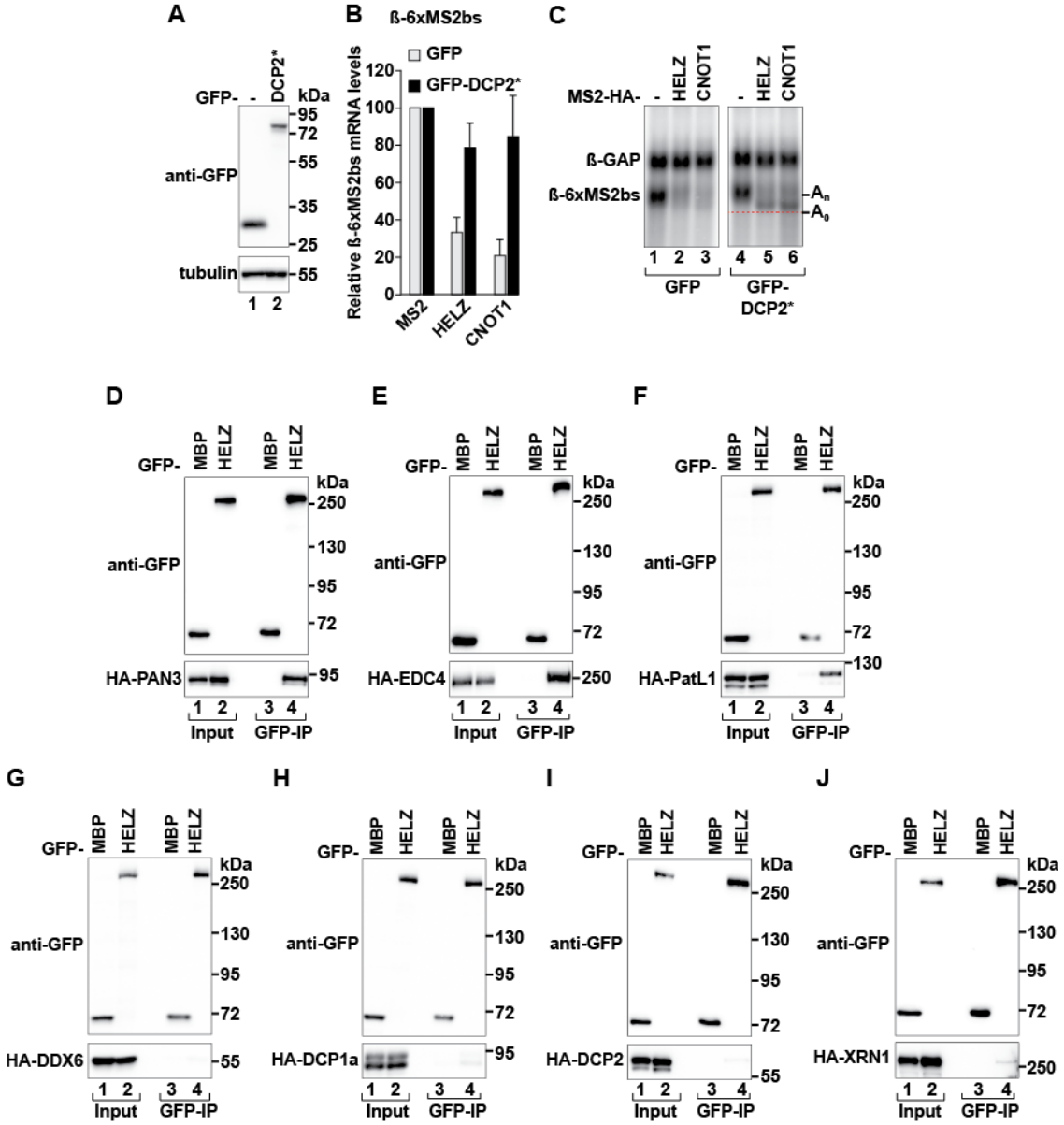

Figure 10 HELZ-induced mRNA decay is decapping-dependent. **A** Western blot comparing the expression level of GFP and GFP-DCP2* (E148Q) in HEK293T cells. Tubulin is used as a loading control. **B** Decapping block assay with the ß-globin reporters in HEK293T cells. GFP was expressed in the control condition and catalytically inactive GFP-DCP2* was used to block decapping. The graph illustrate the quantification of mRNA levels of the ß-6xMS2bs normalized to the ß-GAP values and set to 100 for MS2-HA, in the control (grey bars) and the decapping block (black bars) conditions; mean values +/- SD are represented for three independent experiments. **C** Representative northern blot of samples shown in B. The mRNA of the ß-6xMS2bs reporter with 3'polyA (A_n) and deadenylated (A_0) are marked on the right side of the blot, and a red dotted line indicates A_0 position across the gel. **D to J** Co-GFP-immunoprecipitation assays showing the interaction, or lack thereof, of GFP-HELZ with HA-tagged 5'-to-3' factors: HA-PAN3 of the PAN2/PAN3 complex (**D**), decapping factors HA-EDC4 (**E**), HA-PatL1 (**F**), HA-DDX6 (**G**), HA-DCP1a (**H**), HA-DCP2 (**I**); and HA-XRN1 (**J**). For the detection of GFP-tagged proteins, 2% of input and 20% of GFP-IP were loaded on a SDS-PAGE. For the detection of HA-tagged proteins, 1% of input and 30% of GFP-IP were loaded in a SDS-PAGE. All samples were treated with RNaseA prior to immunoprecipitation experiments.

4.1.4 HELZ interacts with the CCR4-NOT complex

In the aim of shedding light on the molecular detail of how HELZ recruits the 5'-to-3' mRNA decay pathway to target mRNAs, I investigated how HELZ interacts with the CCR4-NOT complex as it is the first and rate-limiting step of the 5'-to-3' mRNA decay pathway. To tackle this challenge, I performed GFP-immunoprecipitation assays in HEK293T cells and *in vitro* pulldown assays with recombinant proteins.

As shown in **Figure 7** B, GFP-HELZ interacts with endogenous CNOT1. Moreover, GFP-HELZ also interacted with endogenous CNOT3 (**Figure 11** A lane 6). To dissect how HELZ interacted with CNOT1 and CNOT3, GFP-HELZ_N and GFP-HELZ_C were expressed in HEK293T cells and immunoprecipitated. Similar to the situation in tethering assay where HELZ_C was recapitulating the effect of HELZ full-length, GFP-HELZ_C associated with both CNOT1 and CNOT3 (**Figure 11** A lane 7). HELZ_N, on the other hand, was not able to immunoprecipitate CNOT1 nor CNOT3 (**Figure 11** A lane 8). HELZ_C also contained the PAM2 motif and thus, interacted with endogenous PABPC1 (**Figure 11** A lanes 6 and 8). Interestingly, when GFP-HELZ and GFP-HELZ_C were tested in co-immunoprecipitation assay with overexpressed HA-tagged CNOT1, the protein levels of HA-CNOT1 were stabilised in the input fraction (**Figure 11** A lanes 1 vs 2 and 4). This effect has been observed previously in our lab and hint to the possibility that an interaction between two proteins may promote protein stabilization. Furthermore, GFP-HELZ_C associated, like GFP-HELZ, with HA-CNOT1 (**Figure 11** B lane 6 and 8). Thus, HELZ and HELZ_C interact with the CCR4-NOT complex in cells and induce mRNA decay when tethered to reporters.

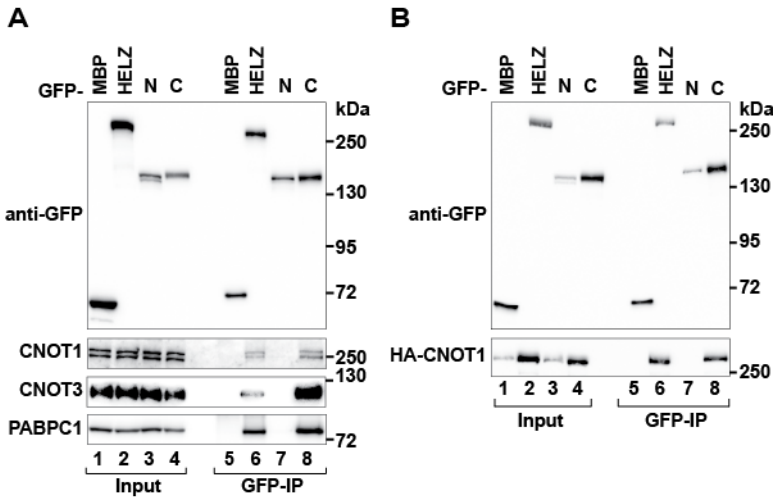

Figure 11 The C-terminal tail of HELZ interacts with the CCR4-NOT complex. **A** and **B** GFP-immunoprecipitation assays illustrating the interaction between GFP-HELZ, GFP-

HELZ_N and GFP-HELZ_C and endogenous CNOT1, CNOT3 and PABPC1 (**A**) or co-expressed HA-CNOT1 (**B**). For the detection of GFP-tagged proteins, 1.2% of input and 20% of GFP-IP were loaded in a SDS-PAGE. For the detection of endogenous proteins, 1.2% of input and 35% of GFP-IP were loaded. For the detection of HA-CNOT1, 1% of input and 30% of GFP-IP were loaded in a SDS-PAGE. All samples were treated with RNaseA prior to immunoprecipitation experiments.

Next, I wanted to further define the details of the interaction between HELZ and the CCR4-NOT complex. To do this, I performed *in vitro* MBP pulldown assays with recombinant HELZ fragments and different combinations of purified sub-complexes of the CCR4-NOT complex. First, I had to optimise the expression of recombinant HELZ and HELZ fragments in both *Sf21* insect cells and *E. coli* (see sub-chapter 3.2.5.3 Recombinant protein expression). Full-length HELZ was possibly too large to be expressed in *E. coli* but I could readily detect the expression of full-length HELZ with a tandem STREP-tag at the N-terminus and a MBP-tag at the C-terminus in lysates of infected *Sf21* insect cell. However, this full-length HELZ-fusion protein could not be purified despite testing several conditions. A possible explanation for this might be that the tandem STREP and MBP tags were masked or that the protein precipitated on the beads. Nevertheless, a construct coding for the full HELZ_C with a tandem STREP-tag at the N-terminus and a MBP-tag at the C-terminus (HELZ_C-MBP) could be expressed in *Sf21* insect cells and could be crudely extracted from the lysates of infected insect cells via its C-terminal MBP-tag. The lysate from *Sf21* cells expressing HELZ_C-MBP was used for *in vitro* pulldown assays to test for an interaction with purified CCR4-NOT complex. The CCR4-NOT is a large multi-subunit complex; thankfully, our lab has a wide expertise in protein complex expression and purification, and has accumulated a large collection of different purified sub-complexes of the deadenylase complex. The sub-complexes used in this thesis were purified by Tobias Raisch, a former PhD candidate in the lab (see **Table 15** for a detailed account of the purified sub-complexes used) (Sgromo et al. 2017).

In the *in vitro* MBP pulldowns, the purified MBP (used as a negative control), and the lysate of insect cells, where HELZ_C-MBP was expressed, were mixed with distinct purified sub-complexes of the CCR4-NOT complex following the protocol described in sub-chapter 3.2.5.4 *In vitro* MBP pulldown. The first two sub-complexes tested, namely CNOT1_N/10/11 and the pentameric sub-complexes (**Figure 12** A), constitute together the major part of the CCR4-NOT complex. Proteins not included are CNOT8 (CNOT7 paralogue) and the two paralogues CNOT6 and CNOT6L (**Figure**

12 A). HELZ_C-MBP pulled down the pentameric sub-complex consisting of CNOT1_M+C (residues 1093-2371), CNOT7, CNOT9 (residues 19-285) and the C-terminal domains of CNOT2 (residues 344-540) and CNOT3 (residues 607-753), whereas MBP alone did not (**Figure 12** B lanes 12 vs 9). In contrast, HELZ_C-MBP did not pull down the CNOT1_N/10/11 complex (**Figure 12** B lane 11) consisting of CNOT1_N (residues 1-1000), CNOT10 (residues 25-707) and CNOT11 (residues 257-498), indicating that these subunits were not involved in HELZ_C binding. This result shows for the first time that the C-terminal tail of HELZ interacts directly with the CCR4-NOT complex and discriminate between different components of the deadenylase complex.

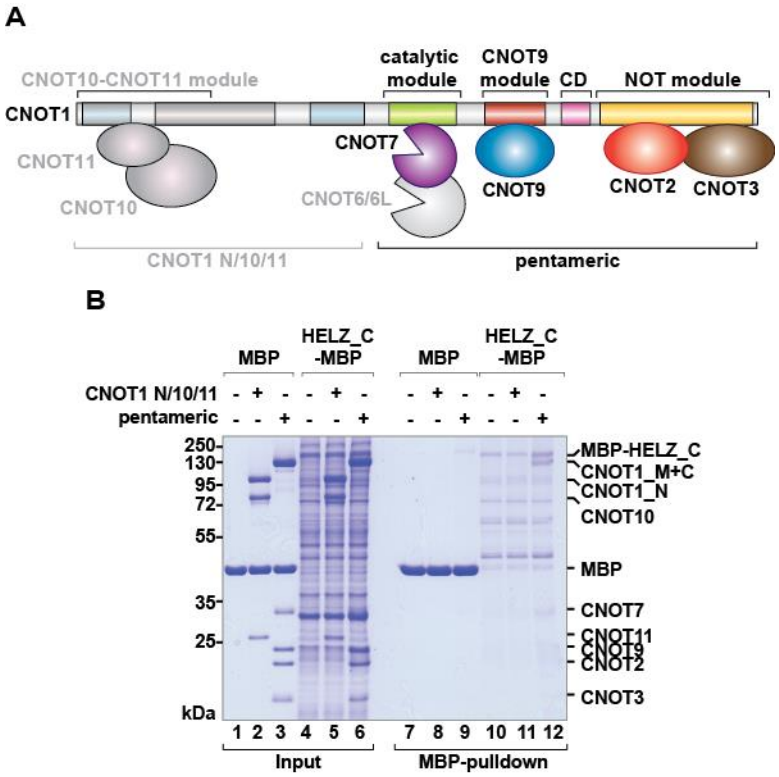

Figure 12 HELZ directly interacts with the CCR4-NOT complex. **A** Schematic representation of the human CCR4-NOT complex. The purified pentameric sub-complex are highlighted in colour. CNOT6 and CNOT6/L were not present in the sub-complexes tested and are thus in grey. The CNOT1_N/10/11 sub-complex is also displayed in grey. **B** *In vitro* MBP pulldown revealing the interaction between recombinant HELZ_C-MBP and one of the two CCR4-NOT sub-complexes. CNOT1_N/10/11 and the pentameric sub-complexes (CNOT1_M+C, CNOT7, CNOT9, CNOT2 and CNOT3) were mixed with purified MBP as a negative control, and with lysates of insect cells expressing HELZ_C-MBP. 37.5% of the input and 55% of the eluate were loaded in the SDS-PAGE, amounts are described in Table 16. The gel was stained with Coomassie blue.

4.1.5 The C-terminal tail of HELZ directly interacts with the NOT module

The C-terminal tail of HELZ is predicted to be unstructured and constitutes about half of the protein length (**Figure 5** F). The purified pentameric complex consists of five proteins arranged in three distinct modules docked on the scaffold protein CNOT1 (**Figure 12** A). Therefore, I performed further mapping experiments to define the binding region of HELZ_C to the CCR4-NOT complex. To this end, I created two constructs that code for two non-overlapping fragments of HELZ_C: HELZ_C1 (residues 1051-1474) and HELZ_C2 (residues 1475-1942) (**Figure 15** A). Since HELZ_C has low sequence conservation among metazoans, I designed the fragments based on sequence alignments of vertebrate HELZ proteins. The designed fragments have similar size in length and the borders are kept in a very low conservation region. Both fragments were expressed in *E. coli* with an MBP-tag at their N-terminus and a GB1-6xHis tag at their C-terminus (MBP-HELZ_C1 and MBP-HELZ_C2) (see sub-chapter 3.2.5.3.2 Recombinant protein expression in E. coli). The bacterial lysates were filtered and the fragments were purified by a Ni^{2+} ion affinity chromatography (see sub-chapter 3.2.5.3.2 Recombinant protein expression in *E. coli*). Both fragments were then used in MBP pulldowns to test their interaction with smaller assemblies of the pentameric sub-complex (for amounts used see **Table 16** and sub-chapter 3.2.5.4 In *vitro MBP* pulldown for details about the protocol).

Despite the instability of the fragments making them prone to degradation, both HELZ_C1 and HELZ_C2 pulled down the pentameric complex (**Figure 13** A and B, lanes 20, respectively). Moreover, not only did both fragments interacted with the pentameric complex but they could both interact with the NOT module (**Figure 13** A and B, lanes 24, for C1 and C2 respectively). The NOT module sub-complex consisted of the SHD domain of CNOT1 (residues 1829-2361), CNOT2 (residues 353-540) and CNOT3 (residues 607-748). In contrast, none of the HELZ fragments tested pulled down CNOT7 bound to the MIF4G domain of CNOT1 (residues 1099-1317), the CNOT9 module comprising of CNOT1 (residues 1356-1588) and CNOT9 (residues 19-285), nor the CD corresponding to residues 1607-1815 of CNOT1 (**Figure 13** A and B, for both lanes 21, 22, 23 respectively for corresponding sub-complexes). Thus, HELZ contains at least two non-overlapping regions directly interacting with the NOT module of the CCR4-NOT complex.

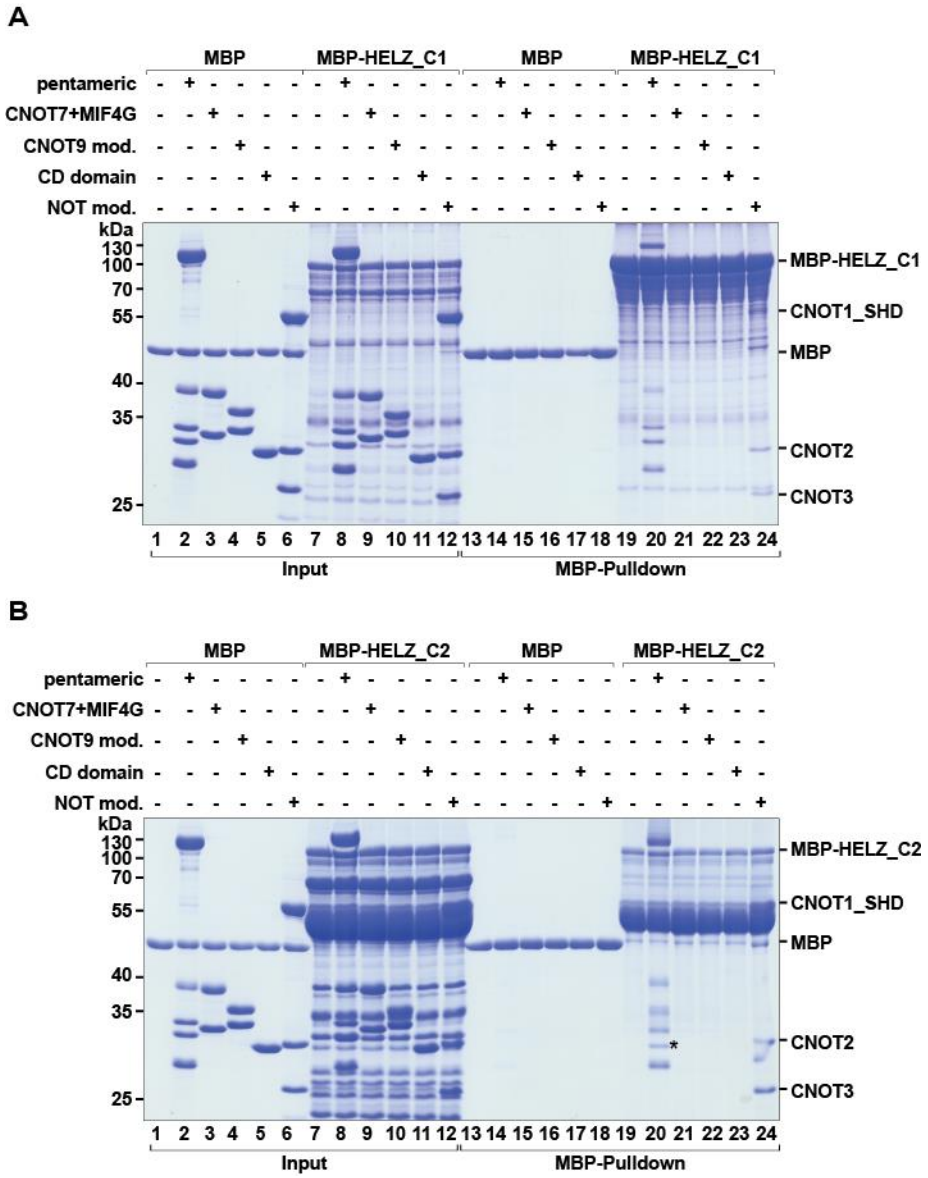

Figure 13 Two non-overlapping fragments of HELZ_C interact with the NOT module of the CCR4-NOT. **A** and **B** *In vitro* MBP pulldowns of two non-overlapping fragments of HELZ_C with the pentameric CCR4-NOT sub-complex and sub-complexes thereof. Purified lysates of bacteria expressing HELZ_C1 (**A**) and HELZ_C2 (**B**) were mixed with the indicated sub-complexes, MBP served as a negative control. The difference in size between the pentameric CNOT3 and the NOT module CNOT3 results from the presence of a non-cleavable tag in the pentameric CNOT3. For clarity, only the NOT module, HELZ fragments and MBP were labelled on the right side of the gel. The asterisk marks a band of HELZ_C2 in complex with the pentameric sub-complex. 33% of input (in a final volume of 60 µl) and 55% of eluate were loaded in a SDS-PAGE, the gels were stained with Coomassie Blue.

4.1.6 HELZ_C2 induces reporter mRNA decay via the CCR4-NOT complex

My results so far revealed that the C-terminal tail of HELZ can induce mRNA decay in tethering assays and that it interacts with the CCR4-NOT complex in cells. These findings support the hypothesis that tethered HELZ recruits the CCR4-NOT

complex to reporter mRNAs to induce their decay. The mode of binding seems to be complex as at least two non-overlapping fragments of the C-terminal tail of HELZ are interacting with the NOT module of the CCR4-NOT complex *in vitro*.

To ascertain that the two binding regions (HELZ_C1 and HELZ_C2) were functionally relevant in cells, I set out to study these two fragments in HEK293T cells. I first tested the interaction between the fragments of GFP-HELZ with components of the CCR4-NOT complex by performing GFP-immunoprecipitation assays in HEK293T cells. However, the transfection of GFP-HELZ_C1 did not result in detectable expression of the fusion protein. This might be due to intrinsic instability of the fragment leading to its fast degradation. To overcome this issue, a fragment comprising the N-terminal part of HELZ, which contains the ZnF and the helicase domain, together with the C1 sequence (HELZ_N+C1) could be expressed in HEK293T cells. Interestingly, GFP-HELZ_N+C1 did not associate with any of the tested CCR4-NOT components (**Figure 14** A lane 7). In contrast, GFP-HELZ_C2 clearly interacted with all tested proteins, namely endogenous CNOT1, CNOT2 and CNOT3 in HEK293T cells (**Figure 14** A lane 8). It is also worth noting that HELZ_C2 was more efficient than HELZ full-length to pull down the CCR4-NOT complex components (**Figure 14** A lanes 6 vs 8). This effect could be due to the complexity of the full-length protein containing the ZnF and the putative helicase domain which could contribute to other HELZ-related CCR4-NOT-independent functions. And hence, the full-length could be less efficient in pulling down the CCR4-NOT complex than the simpler HELZ_C2 fusion protein.

To test if the ability of HELZ fragments to recruit the cellular CCR4-NOT complex is associated with reporter mRNA decay in tethering assay, I performed tethering assays with MS2-HA-HELZ_N+C1 and MS2-HA-HELZ_C2 using the ß-globin reporter system in HEK293T cells. HELZ_N+C1 did not induce reporter mRNA decay when tethered (**Figure 14** C and D lane 3). On the other hand, HELZ_C2 reduced the mRNA levels of the reporter to a similar extent as seen for the tethering of HELZ full-length (**Figure 14** C and D lane 4). All MS2-HA-tagged proteins were expressed at comparable levels (**Figure 14** B). Furthermore, when deadenylation of the CCR4-NOT complex was blocked using the same approach as in **Figure 9** (co-expression of CNOT7* with CNOT1_M), MS2-HA-HELZ_C2 was severely impaired in its ability to induce mRNA decay (**Figure 14** C and D lanes 4 vs 8). These results indicate that HELZ_C2 induces mRNA decay when tethered to a reporter mRNA via the recruitment of the CCR4-NOT complex. Furthermore, as shown for full-length

HELZ, HELZ_C2 requires the deadenylation activity of the CCR4-NOT complex to do so.

The data for HELZ_C1 is harder to interpret: it interacts with the NOT module *in vitro* but it could not induce mRNA decay nor interact in cells with the CCR4-NOT complex, when expressed with the N-terminal region of HELZ. These data could be due to the N-terminal part of HELZ masking the C1 binding region or a regulation of the binding site on HELZ_C1 to the deadenylase complex. Therefore, further studies are required to understand the regulation of HELZ_C1 to the CCR4-NOT complex.

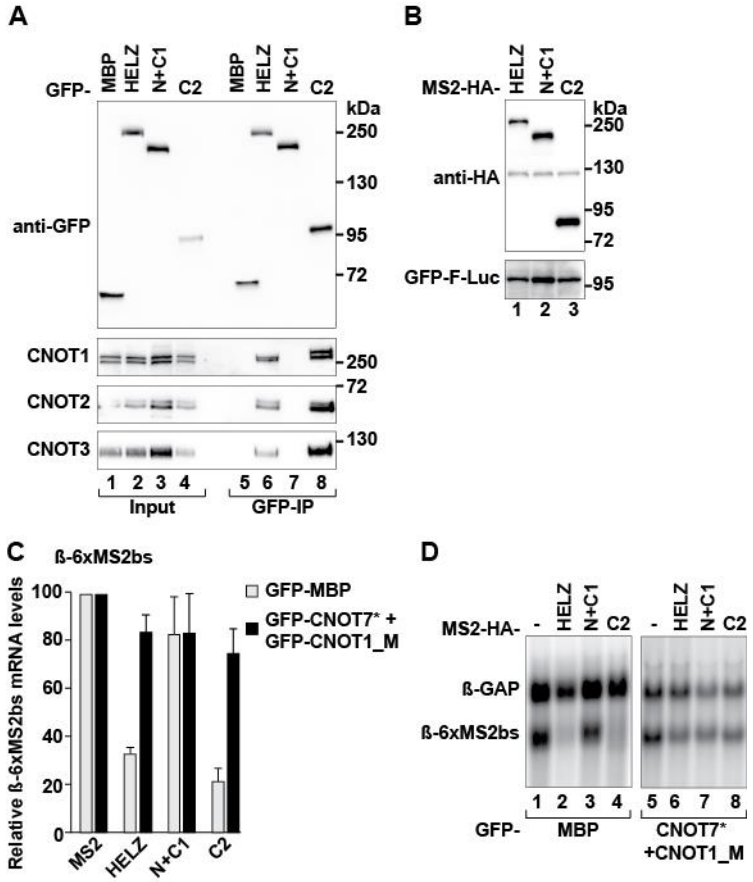

Figure 14 HELZ_C2 induces reporter mRNA decay via the CCR4-NOT complex. **A** GFP-immunoprecipitation assay displaying the interaction of GFP-HELZ proteins (full-length, N+C1 and C2 fragments) with endogenous CNOT1, CNOT2 and CNOT3 in HEK293T cells. GFP-MBP served as a negative control. For the detection of GFP-tagged proteins, 1.2% of input and 20% of GFP-IP were loaded in a SDS-PAGE. For the detection of endogenous proteins, 1.2% of input and 35% of GFP-IP were loaded in a SDS-PAGE. All samples were treated with RNaseA prior to the assay. **B** Expression levels of MS2-HA-HELZ used in panel C and D. GFP-F-Luc was used as a transfection control. **C** Deadenylation block assay with the ß-globin reporters in HEK293T cells. The experiments were performed as described in **Figure 9**. The graph illustrates the quantification of mRNA levels of ß-6xMS2bs normalised to the control ß-GAP reporter and set to 100 for MS2-HA in the control condition (grey bars) and the deadenylation block condition (black bars); the mean values +/- SD are represented for three independent experiments. **D** Representative northern blot of samples quantified in C.

4.1.7 HELZ_C2 contains multiple binding sites for the CCR4-NOT complex

In an attempt to define a precise motif within HELZ_C2 fragment that could be used for future studies, I performed *in vitro* MBP pulldowns with non-overlapping fragments of HELZ_C2. To design the constructs, I made use of the sequence alignment of HELZ_C2 in representative vertebrates (from *Homo sapiens, Mus musculus*, *Gallus gallus*, *Xenopus tropicalis* and, *Danio rerio*) which indicated a conservation of strict sequence identity of 28.5% (calculated with ClustalO and Uniprot, comparable to 31% of strictly conserved sequence identity for HELZ_C in the same vertebrates). Intriguingly, four short stretches ranging from 15 to 27 amino acid residues showed a particularly high level of conservation (>68% sequence identity). Thus, leading me to question if these motifs could be involved in the binding to the CCR4-NOT complex.

I created in total seven constructs (three wild-type and four deletion mutants of the stretches) coding for different HELZ_C2 fragments with an MBP-tag at the N-terminus and a GB1-6xHis-tag at the C-terminus, to be expressed in *E. coli* (**Figure 15** A). HELZ_C2a consists of residues 1474 to 1642 and, HELZ_C2aΔ1 does not contain residues 1595 to 1610 which showed an identity of 82%. HELZ_C2b consists of residues 1643 to 1792 and, HELZ_C2bΔ2 does not contain residues 1746 to 1764 which showed an identity of 89.5%. HELZ_C2c consists of residues 1793 to 1942 and contains two patches: HELZ_C2cΔ3 does not contain residues 1890 to 1904 which showed an identity of 80%, and HELZ_C2cΔ4 does not contain residues 1915 to 1942 which showed an identity of 68%. The lysates of *E. coli* cultures expressing the individual fragments were purified by a Ni^{2+} ion affinity chromatography, the eluted proteins were used in pulldown assays (see sub-chapter 3.2.5.3.2 Recombinant protein expression in *E. coli* for protocol). The eluates were then tested for their ability to interact with the NOT module as it was the only sub-complex interacting with HELZ_C2 (**Figure 13** B).

Surprisingly, all three fragments, C2a, C2b and C2c, interacted, albeit weakly, with the NOT module (**Figure 15** B lanes 20, 24 and 28, respectively). Strikingly, the deletion of stretch 4 in fragment C2c (the last 28 residues in C2c, C2cΔ4) visibly affected the interaction with the NOT module (**Figure 15** B lane 28 vs 32). Deletion of any other conserved stretches did not have a strong effect on the binding to the NOT module (**Figure 15** B lanes 22, 26 and 30 for C2aΔ1, C2aΔ2 and C2cΔ3 respectively).

This demonstrates that HELZ_C2 contains several sites that contact the CCR4-NOT complex. The short sequence of 28 amino acid residues at the C-terminus of HELZ (stretch 4) represents, so far, the shortest sequence identified to contribute to this interaction.

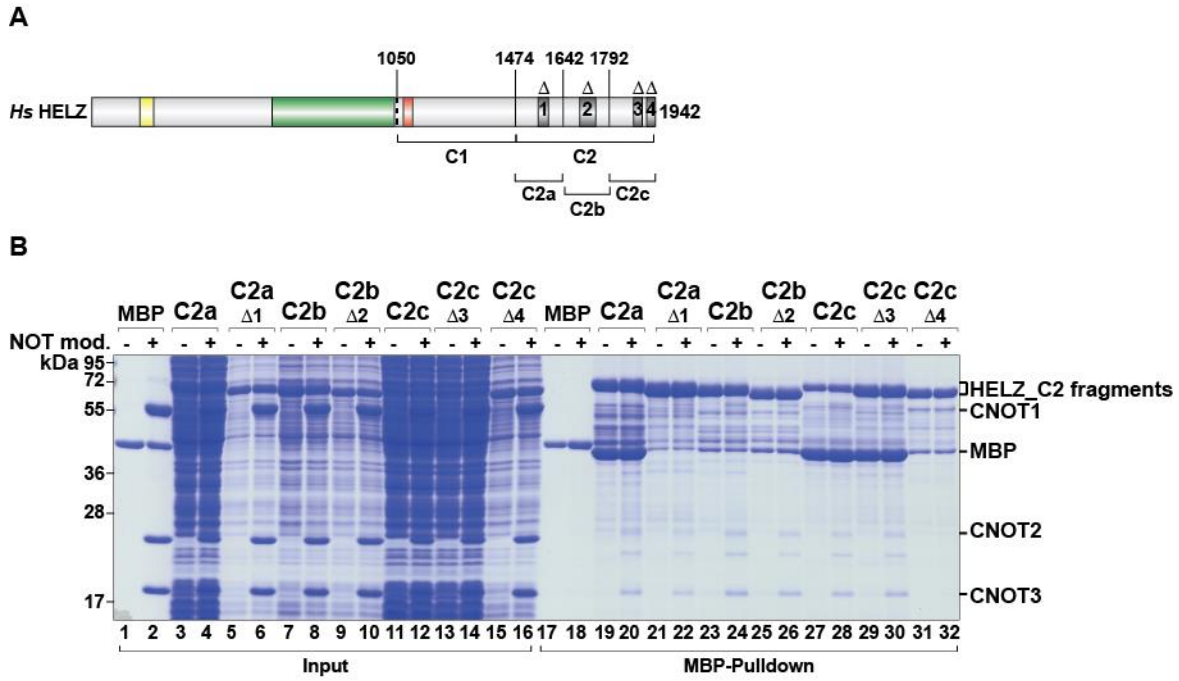

Figure 15 HELZ contains multiple binding sites to interact *in vitro* with the NOT module. **A** Schematic view of HELZ fragments based on **Figure 5** A. Amino acid residues numbers are indicated on top of the scheme. The four conserved stretches are highlighted in dark grey and numbered. Fragments denotations are indicated below the scheme. **B** *In vitro* MBP pulldown displaying the interaction between HELZ_C2 fragments and respective deletion mutants with the NOT module. Purified bacterial lysates expressing MBP-HELZ_C2 fragments were mixed with purified NOT module, MBP served as a negative control. 50% of input and 67% of eluate were loaded in a SDS-PAGE, the gel was stained with Coomassie Blue.

To summarize the results of the first chapter:

- The tethering of HELZ to an mRNA induces decay via the 5'-to-3' decay pathway.
- HELZ interacts with several components of the 5'-to-3' mRNA decay pathway in cells.
- Two fragments of the C-terminal tail of HELZ interact *in vitro* with the CCR4-NOT complex.
- A fragment of HELZ_C, HELZ_C2, contains several regions that directly bind to the NOT module of the CCR4-NOT complex *in vitro*.
- HELZ_C2 can also induce mRNA decay and requires the deadenylation activity of the CCR4-NOT complex to do so.

- The helicase domain, the ZnF and the PAM2 motif do not contribute to the degradation of reporter mRNA when HELZ is tethered. Nevertheless, these features might be crucial to bring HELZ to endogenous target mRNA.

Taken together, these results support a model where HELZ can regulate cytoplasmic mRNA decay of bound mRNAs via its interaction with the CCR4-NOT complex.

4.2 HELZ and mRNA translation

Active mRNA translation is typically associated with an increase in cellular growth and proliferation. HELZ has been described as a negative regulator of cellular growth (Nagai et al. 2003), and work done by another lab described HELZ as a promoter of global translation and of cellular proliferation in HeLa cells (Hasgall et al. 2011). HELZ interacts with key regulators of translation, for example with PABPC1 via its PAM2 motif. Based on a previous report and my results, it also interacts with the CCR4-NOT complex which is known to repress translation and induce mRNA decay of transcripts (Miller and Reese 2012). Given the unclear role of HELZ in translation, I set out to investigate the effect of HELZ on mRNA translation when it is bound to a reporter mRNA. Furthermore, I wanted to determine whether HELZ affects general protein translation in HEK293T cells. Some results in this section are included in a manuscript ready for submission.

4.2.1 HELZ induces translational repression of a reporter lacking a 3'polyA tail

Changes in mRNA levels can directly affect protein levels. Therefore, I made use of a luciferase reporter system that allows me to study the effects at protein level independently of mRNA decay. The reporter encodes for the *Renilla* luciferase ORF followed by a 3'UTR with 6xMS2bs, an internal stretch of 95 adenosines, and a short sequence derived from the mouse long non-coding RNA metastasis associated lung adenocarcinoma transcript 1 (R-Luc-MALAT1, **Figure 16** A). The short sequence from the MALAT1 lnRNA is processed by RNaseP and forms a cloverleaf structure similar to tRNAs (Wilusz et al. 2012). The R-Luc-MALAT1 reporter mRNA does not contain a 3'polyA tail and the internal polyA stretch is protected from deadenylation by the highly structured MALAT1 sequence. Hence, it is refractory to deadenylation and can allow the detection of effects on translation independently of mRNA decay in tethering assay (**Figure 16** A).

Interestingly, the tethering of MS2-HA-HELZ to the R-Luc-MALAT1 reporter in HEK293T cells reduced the protein levels of the *Renilla* luciferase without affecting the mRNA levels of the reporter (**Figure 16** B and C lane 2). Therefore, the tethering of HELZ induces translational repression of bound reporter mRNAs lacking a 3'polyA tail.

To investigate which region of HELZ was responsible for this effect, MS2-HA-HELZ_N and MS2-HA-HELZ_C were tethered on the R-Luc-MALAT1 reporter in HEK293T cells. Intriguingly, neither HELZ_N nor HELZ_C could repress the translation of the R-Luc-MALAT1 reporter, as the R-Luc-MALAT1 values were comparable to the control condition with MS2-HA (**Figure 16** B blue bars). Furthermore, when the mRNA level of the reporter was visualised by northern blot, it revealed that the tethering of MS2-HA-HELZ_C stabilised the mRNA levels of the reporter (**Figure 16** C lane 4). This was unexpected as the tethering of HELZ_C on 3'polyA containing reporters (R-Luc-6xMS2bs, ß-6xMS2bs) induced mRNA decay (**Figure 5** and **Figure 6**). Moreover, protein and mRNA values of the R-Luc-MALAT1 reporter did not correlate when HELZ_C was tethered: the increase in mRNA levels did not lead to a concomitant increase of protein levels (**Figure 16** B). This puzzling result could indicate another level of regulation of HELZ_C of non-3'polyA mRNA, or might be an artefact of the tethering assay. The tethering of MS2-HA-HELZ_N did not to affect the R-Luc-MALAT1 reporter (**Figure 16** B and C lane 3). Thus, these experiments do not lead to a clear conclusion regarding which part of HELZ is responsible for the repression of the R-Luc-MALAT1 reporter mRNA.

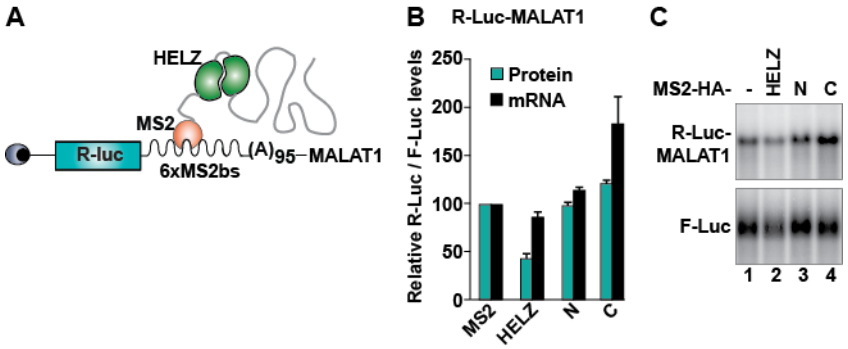

Figure 16 HELZ represses translation of a bound reporter mRNA refractory to deadenylation. **A** Schematic view of the R-Luc-MALAT1 mRNA reporter used to measure translational effects independently of mRNA decay. The ORF encodes for the *Renilla* luciferase and is shown in blue, followed by the 6xMS2bs, a stretch of 95 adenosines $(A)_{95}$ and the MALAT1 sequence in the 3'UTR. MS2-HA protein is represented as a light orange circle bound to one MS2bs and fused to HELZ, which is represented by two folded RecA-like domains of the helicase core (in green) and the unfolded rest of the protein (in grey). **B** Tethering of HELZ and its fragments on the R-Luc-MALAT1 reporter in HEK293T cells. The graph displays the quantification of protein (blue bars) and mRNA (black bars) levels of the R-Luc-MALAT1 reporter normalized to the values of the F-Luc control and

set to 100 for MS2-HA; the mean +/- SD are represented for three independent experiments. **C** Representative northern blot of samples quantified in B.

As mentioned previously in the Introduction, the CCR4-NOT complex is known to repress translation independently of mRNA decay. I posited that HELZ could be in part responsible for the repression effect of the CCR4-NOT complex: HELZ could mask the function of PABPC1 and prevent translation of specific mRNAs. To test this hypothesis, I performed a knock-down of HELZ in HeLa cells and tethered MS2-HA-CNOT1. If HELZ was not responsible for CNOT1-dependent translation repression, I expected that the tethering of MS2-HA-CNOT1 would repress translation independently of the presence of HELZ (in both control and KD cells).

The KD of HELZ was performed with a double transfection of a shRNA against *HELZ* mRNA in HeLa cells. A non-targeting shRNA was used as control (see sub-chapter 3.2.2.2 Knock-down of HELZ and of CNOT1). The efficiency of the KD was assessed by western blot with an antibody against HELZ. Increasing amount of protein samples from control KD cells were loaded and compared to a protein sample from HELZ KD cells. The KD was very efficient as less than 10% of HELZ protein was detected in the KD sample (**Figure 17** A lanes 1 vs 5). The tethering of MS2-HA-CNOT1 induced the translational repression of the R-Luc-MALAT1 reporter, regardless if HELZ was present or not (**Figure 17** B). Hence, CNOT1 does not require HELZ to repress the translation of the R-Luc-MALAT1 reporter. However, the CCR4-NOT complex might nevertheless require HELZ to recruit it to endogenous mRNAs.

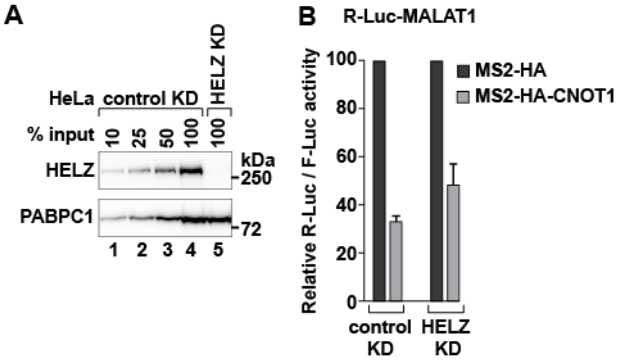

Figure 17 CNOT1 represses translation of the R-Luc-MALAT1 reporter independently of HELZ. **A** KD efficiency of HELZ in HeLa cells assessed by western blot. Different amounts of control KD cell lysates were loaded, indicated as a percentage of input to determine HELZ KD efficiency. PABPC1 was used as a loading control. **B** Tethering of MS2-HA-CNOT1 on the R-Luc-MALAT1 reporter in control HEK293T cells and in HELZ KD cells. The graph displays the quantification of the protein levels of R-Luc-MALAT1 normalised to the values of F-Luc reporter and set to 100 for MS2-HA; mean values +/- SD are shown for three independent experiments. Control KD and HELZ KD are indicated on the x-axis; dark grey bars represent MS2-HA tethering and light grey bars represent MS2-HA-CNOT1 tethering.

4.2.2 CNOT1 and DDX6 contribute to HELZ-induced repression of translation

To test whether HELZ-induced translational repression required the CCR4-NOT complex, a KD of CNOT1 in HeLa cells was performed (**Figure 17** and see 3.2.2.2 Knock-down of HELZ and of CNOT1 for protocol). The rationale of knocking down the scaffold protein CNOT1 of the CCR4-NOT complex was that this should result in the instability of the complex in cells. Transient shRNA treatment caused an effective reduction of CNOT1 protein levels to about 25% as compared to control KD cells (**Figure 18** A). Using the tethering assay, I discovered that knocking-down CNOT1 strongly affected the ability of MS2-HA-HELZ to repress the translation of the R-Luc-MALAT1 reporter (**Figure 18** B). This result demonstrates that HELZ requires the CCR4-NOT complex to induce translational repression of reporter mRNAs.

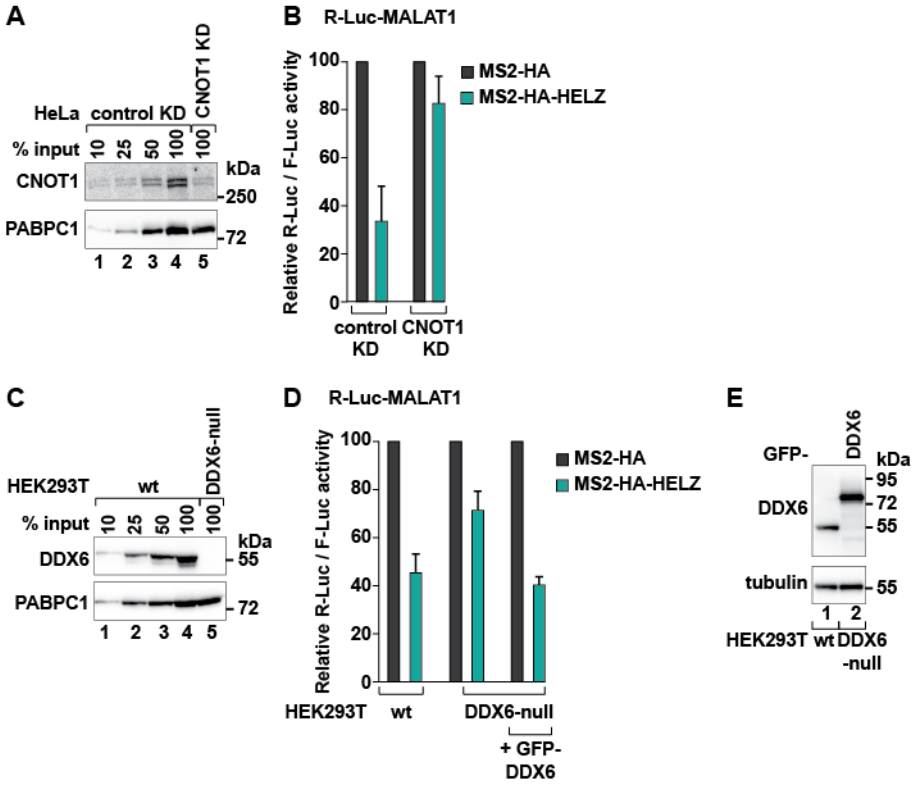

Figure 18 HELZ induces translational repression of the R-Luc-MALAT1 reporter in a CNOT1-dependent manner. **A** KD efficiency of CNOT1 performed in HeLa cells assessed by western blot. Different amounts of control KD cell lysates, indicated as the percentage of total input, were loaded to determine CNOT1 KD efficiency. PABPC1 was detected as a loading control. **B** The tethering of HELZ on the R-Luc-MALAT1 reporter in control and CNOT1 KD HeLa cells. Control KD and CNOT1 KD are indicated on the x-axis; dark grey bars represent values for MS2-HA, green bars represent values for MS2-HA-HELZ. R-Luc-MALAT1 values were normalised to the values of F-Luc control reporter and set to 100 for MS2-HA; mean values +/- SD are shown for four independent experiments. **C** Western blot showing the lack of DDX6 detection in HEK293T DDX6-null cell compared to HEK293T wt cells. Gradual amounts of wt protein samples, indicated as the percentage of input, were loaded to have a visual impression of the lack of DDX6 in the DDX6-null cell line. PABPC1 served as loading control. **D** Tethering of HELZ on the R-Luc-MALAT1

reporter in wt and DDX6-null HEK293T cells. Tethering assay performed in HEK293T wt and in DDX6-null cells where DDX6 levels are restored (+GFP-DDX6) are indicated on the x-axis; dark grey bars represent MS2-HA tethering, green bars represent MS2-HA-HELZ tethering. R-Luc-MALAT1 values were normalised to F-Luc values and set to 100 for MS2-HA; mean values +/- SD are shown for four independent experiments. **E** GFP-DDX6 expressed in DDX6-null cells was expressed at similar levels as DDX6 in wt cells, as assessed by western blot. Tubulin was used as a loading control.

Studies have pointed out that in the context of the miRNA silencing pathway, the DEAD-box RNA helicase DDX6 plays a role in translational repression (Chen et al. 2014; Mathys et al. 2014). In fact, structural studies have shown that DDX6 directly interacts with the CCR4-NOT complex. It was therefore interesting to test if HELZ requires the helicase DDX6 to induce translational repression, even though the two proteins do not interact. To this end, I took advantage of a HEK293T DDX6-null cell line created by a post-doc in the laboratory, Dr. Chung-Te Chang. This cell line does not express DDX6 due to CRISPR/Cas9 mediated gene editing. Loss of DDX6 protein in the null cells was confirmed by western blot analysis (**Figure 18** C lane 5).

HELZ was partially impaired in its ability to repress translation when tethered to the R-Luc-MALAT1 reporter in DDX6-null cells (**Figure 18** D). However, the translational repression activity of HELZ was not completely blocked indicating that, in addition to DDX6, other factors were contributing to HELZ-mediated translational repression. To prove that the reduced effect was indeed caused by the loss of DDX6, I transfected GFP-DDX6 in HEK293T DDX6-null cells to restore DDX6 expression (**Figure 18** E lanes 1 in wt cells vs 2 in DDX6-null cells). The tethering of MS2-HA-HELZ in DDX6-complemented DDX6-null cells restored the repression activity, reaching a similar level measured in wt cells (**Figure 18** D). Thus, the observed negative effect on the activity of HELZ tethering in DDX6-null cells is explained by the lack of DDX6.

Although, I could not identify an interaction between HELZ and DDX6 in GFP-immunoprecipitation assays (**Figure 10** G), DDX6 might be recruited to the R-Luc-MALAT1 reporter indirectly via its interaction with CNOT1, which in turn can be recruited by HELZ. Since HELZ binds to the NOT module *in vitro* and DDX6 binds to the MIF4G domain of CNOT1, it is therefore plausible that both helicases can bind simultaneously to the CCR4-NOT complex.

4.2.3 HELZ does not primarily localise to P-bodies

HELZ was described as a predominantly cytoplasmic protein even though it was also described to regulate transcription in the nucleus (Wagner et al. 1999; Hamamoto et al. 2004; Diehl et al. 2010; Hasgall et al. 2011). In the cytoplasm, P-bodies are membrane-less foci enriched in factors involved in mRNA decay and translational repression such as the CCR4-NOT complex, the RNA helicase DDX6, the decapping factor DCP1 and the exonuclease XRN1 (see sub-chapter 1.2.2 The CCR4-NOT complex). mRNAs are thought to be kept at a translationally repressed state in P-bodies of epithelial cells (Hubstenberger et al. 2017). A prey-bait based study performed in HEK293T cells and validated for some specific targets in HeLa cells identified HELZ to be part of P-bodies (Youn et al. 2018). I thus wanted to test if HELZ localises to P-bodies in HeLa cells (see sub-chapter 3.2.2.4 Immunofluorescence assay).

I performed immunofluorescence assays where MS2-HA-HELZ and GFP-DDX6 were co-expressed in HeLa cells. MS2-HA-HELZ was detected by an antibody coupled to Alexa 594. MS2-HA-HELZ is dispersed in the nucleus and in the cytoplasm of HeLa cells (**Figure 19**). GFP-DDX6 localised to the cytoplasm as a diffuse signal and as puncta which represent P-bodies (**Figure 19**). MS2-HA-HELZ did not co-localised with GFP-DDX6 in P-bodies (**Figure 19**). Nevertheless, I cannot rule out that small amounts of HELZ localise to P-bodies as its signal was dispersed in the cytoplasm. Hence, HELZ could function as a transcriptional regulator in the nucleus and as a mRNA translation and stability regulator in the cytoplasm in HeLa cells.

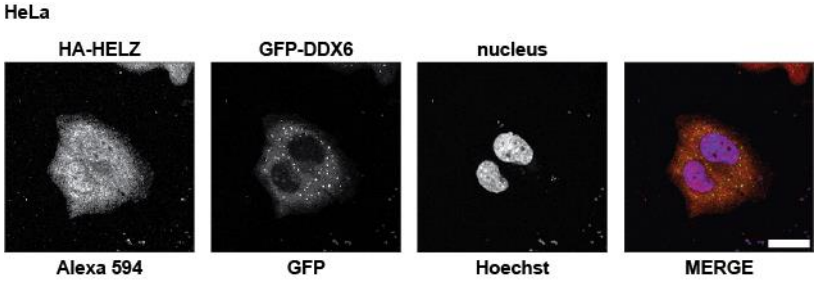

Figure 19 Overexpressed MS2-HA-HELZ is dispersed in HeLa cells. HeLa cells were co-transfected with MS2-HA-HELZ and GFP-DDX6. MS2-HA-HELZ was detected with an antibody coupled to Alexa 594. GFP-DDX6 serves as a P-bodies marker. The nucleus was stained with the Hoechst dye. The last panel shows a merge image of the three signals (HELZ in red, DDX6 in green and the nucleus in blue); the white scale bar represents 10 μm.

4.2.4 HELZ does not affect general mRNA translation in HEK293T cells

Decreasing the expression of HELZ in HeLa cells was shown to reduce general translation assessed by polysome profiling (Hasgall et al. 2011). Hence, the authors conclude that HELZ could act as a general protein promoter. As I have shown that HELZ represses the translation of reporter mRNA when it is bound to mRNA, I wanted to test if and how HELZ affects general protein translation of HEK293T cells.

To study a possible effect of HELZ on general mRNA translation, I generated a cell line that did not express HELZ using the CRISPR/Cas9 genome engineering system. Details of the protocol to generate the HEK293T HELZ-null cell line are given in sub-chapter 3.2.2.3 Generation of the HEK293T HELZ-null cell line. There is no detectable level of HELZ protein in the HELZ-null cell line and the cells were viable with no apparent phenotypes (**Figure 20** A). Furthermore, loss of HELZ expression did not affect the protein levels of PABPC1, CNOT2, CNOT3 or DDX6 (**Figure 20** A). Hence, HELZ is not essential for HEK293T cells growth in culture.

To study the impact of the absence of HELZ on the different steps of general mRNA translation, I performed polysome profiling experiments. Polysome profiling allows a visual representation of global translation at a certain time-point by separating RNPs by density on a sucrose gradient. The most abundant RNAs in cells are ribosomal RNAs (rRNA constitute about 80% of total RNA), which are used as a readout for the visualization of the different translational apparatus along the profile. Effectively, the profile shows several peaks: the 40S peak (corresponding to the small ribosomal subunit), the 60S peak (corresponding to the large ribosomal subunit), the 80S peak (fully assembled ribosome on a mRNA) and several peaks corresponding to two, three, four, five ribosomes on one mRNA molecule (polysomes). Polysome fractions represent actively translating ribosomes in the elongation phase.

In my experiments, the polysome profiles of HEK293T wt and HEK293T HELZ-null cells were nearly identical and could be reproducibly overlaid as illustrated in **Figure 20** B. This indicates that the loss of HELZ expression in HEK293T does not visibly change global translation.

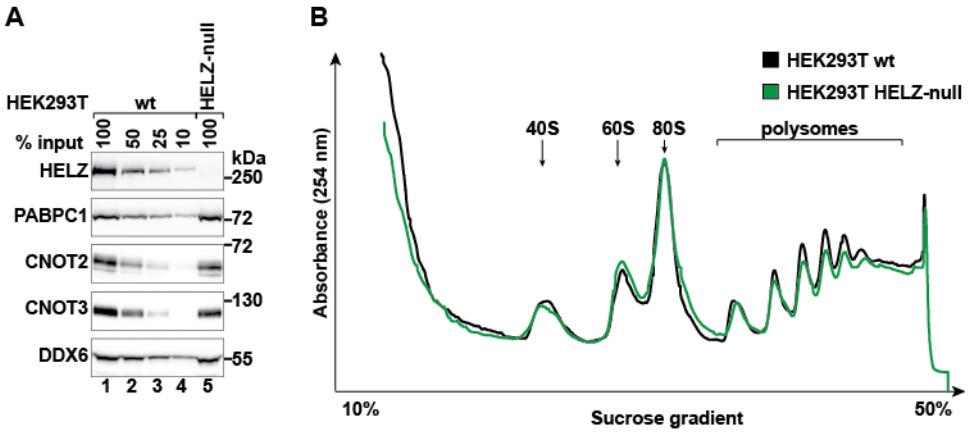

Figure 20 Loss of HELZ in HEK293T does not affect general mRNA translation. **A** Comparative expression of several proteins in HEK293T wt vs HEK293T HELZ-null cells. Decreasing amounts of HEK293T wt cell lysates were loaded as indicated by the percentage of input on top of the gels to visualize the difference in expression between HEK293T wt and HEK293T HELZ-null cells. **B** Polysome profile of HEK293T wt cells (black line) and HEK293T HELZ-null cells (green line). The different peaks on the tracing are indicated on the upper part of the scheme. The profile corresponds to one out of five replicates.

4.3 The helicase domain of HELZ

HELZ has all the motifs that classify it as a UPF1-like SF1 helicase. However, the ATPase and helicase activities of HELZ have not been determined and the protein is thus referred to as a putative helicase. Nevertheless, given that the function of UPF1, MOV10, MOV10L1 and AQR (all UPF1-like helicases) are linked to their helicase and ATPase activities and their studies have brought new aspects of RNA metabolism to light, my third focus was to investigate the helicase function of HELZ and its role in mRNA regulation.

4.3.1 A HELZ ATPase mutant changes the electrophoretic mobility of reporters

The enzymatic characterisation of a helicase typically requires the purification of the entire protein or the helicase domain to be used in *in vitro* assays. These assays allow to define nucleic acid binding preference(s), direction of action and ATPase activity. However, we could not express the helicase domain of HELZ in a bacterial system in the laboratory. Therefore, I opted to get some insight into the importance of the helicase domain using an alternative approach: I introduced a single point mutation in the conserved motif II of the helicase domain of HELZ and analysed the effect of this mutant in human cells. Motif II coordinates the binding and hydrolysis of ATP in all characterized eukaryotic helicases (Fairman-Williams et al. 2010). Therefore, mutating

this motif (DEAA to DQAA - mutant HELZ_E795Q or E795Q) in HELZ protein would disrupt the ATPase activity crucial for the function of the helicase (Pause and Sonenberg 1992). I used the tethering assay to understand what could be the effect of HELZ_E795Q on mRNAs as there is no knowledge of which endogenous mRNA HELZ affects.

HEK293T cells were co-transfected with MS2-HA-HELZ_E795Q, R-Luc-6xMS2bs and F-Luc. In these experiments, the tethering of MS2-HA and MS2-HA-HELZ and its fragments served as controls as they have known effects on the reporter mRNAs. The tethering of HELZ_E795Q decreased the protein and mRNA levels of the R-Luc-6xMS2bs reporter (**Figure 21** A and B lane 6) without affecting the mRNA levels of the control reporter F-Luc (**Figure 21** B). However, the reduction of reporter mRNA levels caused by HELZ_E795Q was weaker compared to full-length HELZ, HELZ_C and HELZ_F1107V tethering. Even though, GFP-HELZ_E795Q could interact with HA-CNOT1 in GFP-immunoprecipitation assay (**Figure 21** D lanes 5 vs 6).

Curiously, when the mRNA levels of the reporters were visualized by northern blot, the residual mRNA levels of the R-Luc-6xMS2bs reporter as well as the mRNA levels of the F-Luc control reporter were changed in their electrophoretic mobility and shifted upwards on the gel when cells were expressing MS2-HA-HELZ_E795Q (**Figure 21** B lane 6). This effect does not depend on tethering as the F-Luc reporter was also affected. Both MS2-HA-HELZ full-length and -HELZ_E795Q were expressed at similar levels (**Figure 21** C and **Figure 23** A). Thus, the expression of HELZ_E795Q was enough to cause a change in the electrophoretic mobility of the reporter mRNA. This effect could be due to an abnormal folding of the mutant protein causing adverse effects on the co-transfected reporters. However, if the fold was severely affected by the point mutation, it would unlikely retain the ability to interact with HA-CNOT1 (**Figure 21** D) and the introduction of this point mutation in other helicases do not affect the fold of the protein (Fairman-Williams et al. 2010). Therefore, I conclude that the observed modification of reporter mRNAs when HELZ helicase mutant is expressed, is most likely caused by the lack of the ATPase activity of HELZ.

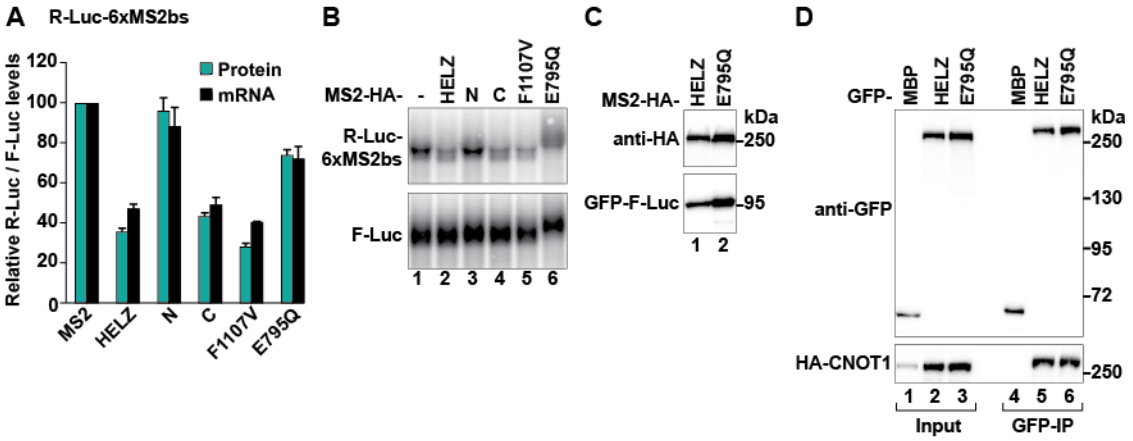

Figure 21 The expression of the ATPase mutant HELZ_E795Q affects the electrophoretic mobility of reporter mRNAs. **A** Tethering of MS2-HA-HELZ and fragments/mutants to the R-Luc-6xMS2bs reporter in HEK293T cells. The graph illustrates protein (blue bars) and mRNA (black bars) levels of R-Luc-6xMS2bs normalised to the values of F-Luc; and set to 100 for MS2-HA. Tethered MS2-HA-tagged proteins are indicated on the x-axis. Mean values +/- SD are presented for three independent experiments. **B** Representative northern blot of samples quantified in A. **C** Expression of MS2-HA-HELZ and MS2-HA-HELZ_E795Q in HEK293T cells as analysed by western blot. GFP-F-Luc served as a loading control. **D** GFP-immunoprecipitation assay displaying the interaction of GFP-HELZ and GFP-HELZ_E795Q with HA-CNOT1 in HEK293T cells. GFP-MBP served as a negative control. For the detection of GFP-tagged proteins, 1.2% of input and 20% of GFP-IP were loaded on a SDS-PAGE. For the detection of HA-CNOT1, 1% of input and 30% of GFP-IP were loaded in a SDS-PAGE. All samples were treated with RNaseA prior to the immunoprecipitation assay.

4.3.2 The expression of HELZ_E795Q leads to the elongation of the 3'polyA tail of reporter mRNAs

The difference in the electrophoretic mobility, observed as an upward shift of reporter mRNAs, when HELZ_E795Q is co-expressed in cells can be explained by an increase in the length of the mRNA reporter which would then run slower in an agarose gel. An increase in length of endogenous mRNAs could be caused, for example, when an intron is retained or if the 3'polyA tail length is increased. Intron retention is typically observed when an mRNA is misprocessed in the nucleus. However, the luciferase reporters do not contain introns. Furthermore, F-Luc and R-Luc-6xMS2bs protein levels were readily detected, meaning that the mRNAs of the reporters are being actively exported and translated in the cytoplasm. Therefore, it was likely that the increase in size indicated by the shift on the northern blot gel was caused by an increase in 3'polyA tail length. If this was true, then the R-Luc-MALAT1 reporter should not be affected as it does not contain a 3'polyA tail (**Figure 16** A). I thus tested if the co-expression of MS2-HA-HELZ_E795Q caused a similar shift of the R-Luc-MALAT1 reporter.

The mRNA of the R-Luc-MALAT1 reporter was not shifted upwards when MS2-HA-HELZ_E795Q was expressed, even though it was still observed for F-Luc (**Figure 22** A and B lane 6). This experiment strongly indicates that the 3'polyA tail of reporters are extended when HELZ_E795Q is expressed.

Interestingly, the tethering of HELZ_E795Q did not lead to the translational repression of the R-Luc-MALAT1 reporter (**Figure 22** A). Indicating that, in addition to CNOT1 (and partially DDX6), the ATPase activity of HELZ is likely involved in the translational repression effect in tethering assays induced by full-length HELZ. Furthermore, the result in **Figure 22** A and B lane 5 also revealed that the interaction between HELZ and PABPC1 was dispensable for HELZ to repress the translation of the R-Luc-MALAT1 reporter. However, as for the experiments in **Figure 8**, the interaction with PABPC1 might be important for the recruitment of HELZ to endogenous mRNAs. This could explain why in tethering assays this interaction is not necessary.

In summary, the ATPase activity of HELZ is most likely important for the translational effect of HELZ when tethered to the R-Luc-MALAT1 reporter. However, the ATPase activity of HELZ does not affect the length of the R-Luc-MALAT1 reporter mRNA.

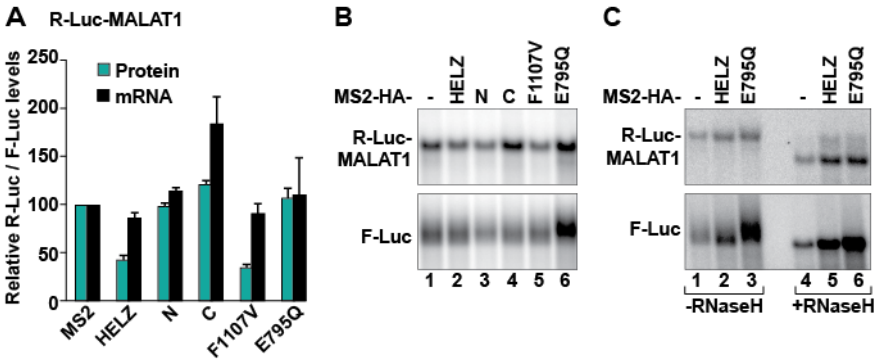

Figure 22 HELZ_E795Q expression leads to an elongation of the 3'polyA tail of reporter mRNAs. **A** Tethering of MS2-HA-HELZ and its fragments/mutants to the R-Luc-MALAT1 reporter in HEK293T cells. The graph illustrates protein (blue bars) and mRNA (black bars) levels of R-Luc-MALAT1 normalised to the values of F-Luc; and set to 100 for MS2-HA. Tethered MS2-HA-tagged proteins are indicated on the x-axis. Mean values +/- SD are represented for three independent experiments. **B** Representative northern blot of samples quantified in A. **C** Representative northern blot of a RNaseH assay with oligo (dT). Samples as in A were treated with (+RNaseH) or without (-RNaseH) RNaseH in the presence of oligo (dT), which digested the 3'polyA tail of mRNAs.

To confirm that the 3'polyA tail was indeed elongated, I performed an RNaseH assay with single-stranded sequences of fifteen deoxythymines (oligo (dT)). RNaseH cleaves the RNA strand in a RNA/DNA duplex and hence, cleaves the 3'polyA tail of

an mRNA when it is annealed to oligo (dT). Effectively, I treated mRNA extracts from samples where MS2-HA-HELZ or MS2-HA-HELZ_E795Q (and MS2-HA as negative control) were co-expressed with R-Luc-MALAT1 and F-Luc. RNA extracts were mixed in the presence of oligo (dT) with or without RNaseH. The digested RNA samples were then purified and visualised by northern blot. As expected, R-Luc-MALAT1 mRNAs ran quicker when the samples were treated with RNaseH regardless of the tethered protein (**Figure 22** C lanes 1 vs 4 for example). In this case, the oligo (dT) annealed to the internal polyA stretch of 95 adenosines monophosphates present in the R-Luc-MALAT1 reporter (**Figure 16** A). The RNaseH recognised this site, digested the RNA strand and removed the polyA stretch and the 3'MALAT1 sequence from the remaining R-Luc ORF with the 6xMS2bs, which is detected by northern blot. Thus, the RNaseH assay did indeed work.

Interestingly, in line with an elongation of the 3'polyA tail length of the control reporter, F-Luc mRNAs were no longer shifted upwards in RNaseH treated samples when co-expressed with MS2-HA-HELZ_E795Q (**Figure 22** C lanes 3 vs 6). In similar experiments were the ß-globin reporters were used and co-expressed with HELZ_E795Q, both reporter mRNAs were no longer shifted upwards in RNaseH treated samples (not shown). Thus, the shift upwards of reporters when co-expressed with the HELZ helicase mutant most likely results from an increase in the 3'polyA tail length of reporter mRNA with a 3'polyA tail.

4.3.3 PABPC1-binding is dispensable for HELZ_E795Q to affect the length of the 3'polyA tail of reporter mRNAs

The 3'polyA tails of mRNAs are bound by PABPC1 in the cytoplasm and HELZ interacts with PABPC1 via its PAM2 motif (**Figure 7** B). I posited that PABPC1-binding might be required for HELZ helicase mutant to affect the 3'polyA tail length of reporter mRNAs. To test this hypothesis, I generated a double mutant disrupting both the ATPase activity and PABPC1 binding in HELZ protein (HELZ_E795Q+F1107V mutant). GFP-F-Luc was used as a negative control. GFP-tagged proteins were transfected with the ß-globin reporters in HEK293T cells. I performed the experiments without artificially tethering the HELZ constructs as I have demonstrated that the extension of the 3'polyA tail of reporter mRNAs does not require the artificial tethering of HELZ_E795Q (**Figure 21** B lane 6 and **Figure 22** B lane 6, F-Luc reporter was not

artificially bound by HELZ_E795Q). Thus, the experiments were performed with GFP-tagged proteins instead of MS2-HA tagged proteins.

As expected, the expression of GFP-F-Luc and GFP-HELZ did not affect the run of the ß-globin reporters in an agarose gel (**Figure 23** A lanes 1 and 2). However, the expression of GFP-HELZ_E795Q and of GFP-HELZ_E795Q+F1107V caused a similar upwards shift of both ß-globin reporters on the northern blot gel (**Figure 23** A lanes 3 and 4). GFP-HELZ and mutants were expressed at comparable levels (**Figure 23** B). Thus, PABPC1 binding is not required to induce the 3'polyA tail elongation of reporter mRNAs.

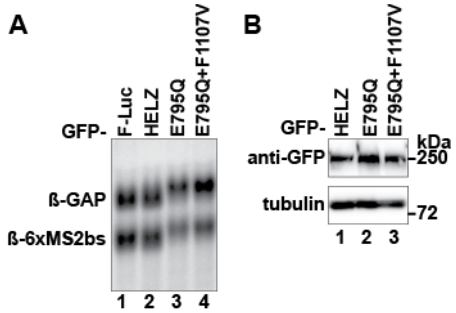

Figure 23 The change in electrophoretic mobility of the reporters mRNA induced by HELZ_E795Q is independent of tethering and of PABPC1 binding. **A** Representative northern blot of the ß-globin reporters co-expressed with GFP-F-Luc, as a control, and GFP-HELZ constructs. **B** Western blot displaying the expression of GFP-HELZ, GFP-HELZ_E795Q, and GFP-HELZ_E795Q+F1107V. Tubulin was used as a loading control.

This preliminary study on the helicase function of HELZ indicated that:

- The expression of HELZ_E795Q in cells results in a change in the electrophoretic mobility of reporter mRNAs as seen as a shift upwards on agarose gels.
- This effect does not require the artificially tethering of the protein on reporter mRNAs.
- This effect is due to an increase in the length of the 3'polyA tail of reporter mRNAs and thus, requires that the reporter mRNA has a 3'polyA tail.
- The interaction with PABPC1 is not necessary for the upward shit of reporter mRNAs to occur.

Taken together, these results strongly indicate that HELZ has an ATPase activity linked to a helicase function. Furthermore, this function of HELZ in cells is likely linked to 3'polyA tail regulation.

4.4 HELZ and its potential mRNA targets

The final chapter of the Result section focuses on my attempt to identify endogenous mRNAs regulated by HELZ in cells. Based on previous results, HELZ has the potential to regulate mRNA on multiple levels. Also, HELZ possess features related to RNA binding. To gain an unbiased insight into which mRNAs are regulated by HELZ, I performed an RNA-Seq experiment and confirmed my findings in independent qPCR analysis.

4.4.1 Identification of differentially expressed RNAs upon the loss of HELZ in cells

I performed an RNA-Seq experiment and compared the transcriptome of HEK293T with HEK293T HELZ-null cells. The protocol used is detailed in sub-chapter 3.2.4.5 RNA-Sequencing. Comparative transcriptome analysis identified 4 923 differentially expressed transcripts with a false discovery rate (FDR) lower than 0.05 (in other words, the identified transcripts are likely at 99.5% to be truly differentially expressed when HELZ expression is removed). FDR was used instead of p-values because FDR are adjusted p-values for false positive discovery (the discovery of transcripts that are not differentially expressed but were found to do so by error) based on the Benjamini-Hochberg method. Out of 4 923 differentially expressed transcripts, 27 were upregulated and 58 were down regulated by more than 2 folds (log2FC>1) when HELZ was not present (**Table 17**). Interestingly, the *HELZ* mRNA was still detected and listed among the downregulated transcripts in HEK293T HELZ-null cell lines (**Figure 24** D). Nevertheless, as assessed by genomic sequencing of the *HELZ* gene, the mutation introduced in the targeted exon 8 caused a frame-shift in the ORF; thereby preventing the production of HELZ protein and likely, targeting the transcript to NMD.

Global Ontology (GO) analysis revealed that the affected transcripts coded for proteins involved in the nervous system, skeletal system, immune system, development, cell adherence and communication, and some act as transcription factors. Although this experiment was performed in cells that have been heavily modified, the data obtained give a list of transcripts that HELZ might regulate. However, further analysis is needed and required to fully comprehend the role of HELZ in these the biological processes.

Table 17 RNA-Seq data

Gene	Log2FC
LINC01419[r]	3,50
GFRA3	2,49
ZNF385C	2,48
TSSC2[r]	2,32
TMEM35	1,92
BARX1	1,91
LINC01399[r]	1,87
HLA-DMA	1,64
SPARC	1,62
BASP1	1,50
RIPPLY3	1,50
CXCL12	1,49
NPIPA5	1,49
PLEKHA4	1,47
TCN2	1,42
UGT3A2	1,27
RHBDL1	1,26
SULT1C4	1,24
INS-IGF2	1,24
IGF2	1,20
UTS2R	1,15
CCDC153	1,13
ADNP2	1,13
LINC01033[r]	1,10
TENM1	1,09
PRUNE2	1,07
SALL3	1,06
HIST3H2BB	-1,01
BMP2	-1,01
HRASLS	-1,01
BCYRN1[r]	-1,02
NKX3-1	-1,03
GHR	-1,04
PSMB9	-1,04
GPR50	-1,04
INA	-1,05
IL32	-1,05
CA2	-1,06
LY6G6D	-1,08
MT1F	-1,13
IFIH1	-1,15
ZNF35	-1,17
COL4A4	-1,18
HIST1H2AI	-1,21
TCF24	-1,24
STYK1	-1,25
CADPS2	-1,25
PCDH10	-1,28
RNF43	-1,31
BSPRY	-1,33
NAALAD2	-1,35
AGPAT4-IT1	-1,36
LDAH	-1,37
SOX2-OT[r]	-1,39
HELZ	-1,44
SATB1	-1,44
RPS17	-1,46
CASP8	-1,47
FAAH2	-1,50
EPCAM	-1,54
CMTR2	-1,59
IGFBP3	-1,64
TLR3	-1,75
ALDH1A2	-1,98
CR2	-2,00
FERMT1	-2,06
RPRM	-2,18
LGALS3	-2,18
EDA2R	-2,20
LHX8	-2,28
SNTB1	-2,32
BST2	-2,33
CELF2	-2,38
OSTN	-2,43
GALNT14	-2,64
ACADL	-2,66
MYD88	-2,72
SP110	-2,87
UNCX	-2,91
MAOA	-3,10
MGST1	-3,82
LINC00693[r]	-6,00
SLC8A1	-6,59
DTX3	-11,53

[r] marks non-coding RNAs.

Briefly, I describe the known function of some proteins of the studied transcripts in **Figure 24** (underlined in **Table 17**). TMEM35 stands for transmembrane protein 35A and has been described to regulate long-term memory and stress factors in mice, and cellular proliferation and migration in osteosarcoma. HLA-DMA, for HLA class II histocompatibility antigen, DM α chain, presents exogenous antigen to the immune system. SPARC (secreted protein acidic rich in cysteine), also known as osteonectin, is an ion binding protein important in extracellular matrix arrangement. BASP1, for brain acid soluble protein 1, is involved in neurite outgrowth. TENM1, teneurin-1, has been described as playing a role in neuronal development. IFIH1 is a RigI-like SF2 helicase (interferon-induced helicase C domain containing protein 1) and is important

in innate immune response by binding to viral RNAs. PCDH10 (protocadherin 10) is important for cell adherence. Osteocrin (OSTN) is involved in dendritic growth and osteogenesis. MAOA, monoamine oxidase A, is an enzyme involved in the production of neuroactive and vasoactive amines in the central nervous system and peripheral tissues.

4.4.2 Validation of the RNA-Seq data by qPCR

To validate the RNA-Seq data, I performed qPCR experiments with the help of Dr Heike Budde. RNA was extracted from three independent samples and cDNA was synthesized according to the protocol in sub-chapter 3.2.4 RNA-based assays. All oligo pairs were designed with the online tool Primer3 and tested for specificity. Specificity was evaluated based on the melting curve of the qPCR profile and by loading the post-qPCR products on an agarose gel. A single peak in the melting curve and a single band in the agarose gel indicated that the primers were specific. Although I aimed to obtain specific primer pairs for all targets, I eventually worked with a few primer pairs that fitted the experimental procedure. Hence, the transcripts that were studied were not chosen for any other reasons than experimental design.

The Ct values of each target were measured by the qPCR along with the Ct values of the reference transcript coding for *actin*. The analysis of the fold change for each targets between HEK293T wt samples and HEK293T HELZ-null samples were calculated according to the Pfaffl method, which takes into account the efficiency of the primer pair (Pfaffl 2001). I tested 21 transcripts (13 up-regulated and 8 down-regulated) by qPCR. I confirmed the effect observed in the RNA-Seq data for the 11 up-regulated and 8 down-regulated transcripts, meaning that up-regulated transcripts in RNA-Seq were up-regulated when detected by qPCR, and *vice versa*. The two up-regulated transcript that were not up-regulated in the qPCR analysis might be indirectly affected by the loss of HELZ or an artefact of the experiment.

To confirm that the changes observed resulted from the absence of HELZ, I performed add-back experiments where the expression of HELZ was restored in HEK293T HELZ-null cells to a similar level or higher levels compared to endogenous HELZ protein levels in HEK293T wt (**Figure 24** A and B). mRNAs were extracted, cDNAs were synthesized and qPCRs were done for nine transcripts (five up-regulated and four down-regulated transcripts) (**Figure 24** C and D). As expected, adding-back increasing amounts of HELZ showed a tendency for up-regulated transcripts to

decrease mRNA abundance to wt levels (**Figure 24** C). For downregulated transcripts, this tendency was less striking as the log2FC values were more dispersed, possibly hinting to a complex mode of regulation of HELZ of these transcripts (**Figure 24** D). Thus, HELZ might directly regulate the expression level of the validated transcripts.

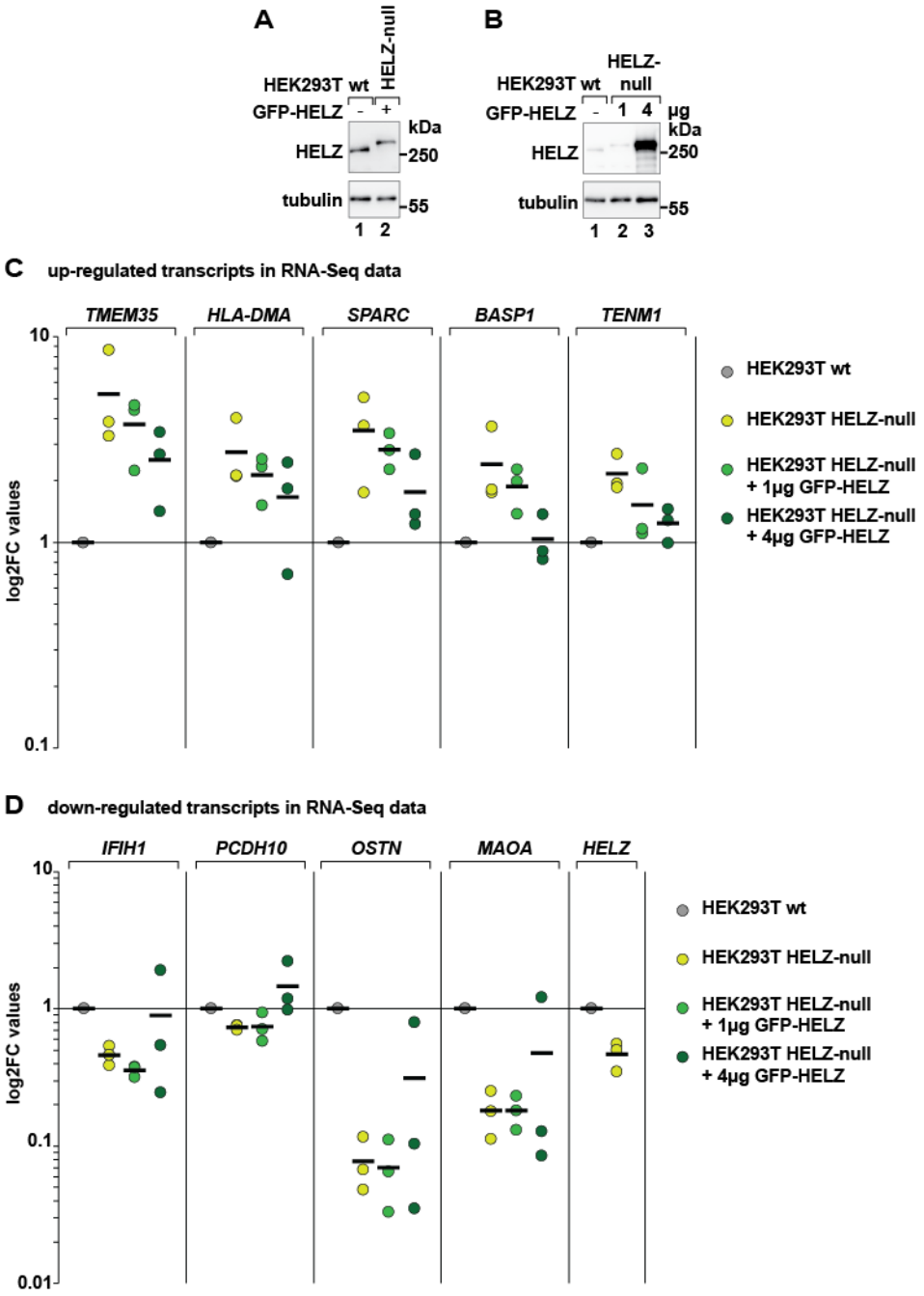

Figure 24 qPCR validation of RNA-Seq results and add-back experiments. **A** Western blot of endogenous HELZ protein levels in HEK293T wt compared to the artificial expression of GFP-HELZ in HEK293T HELZ-null cells. GFP-HELZ runs slower than endogenous HELZ because of the GFP-tag. Tubulin was used as a loading control. **B** Comparison of endogenous HELZ protein expression in HEK293T wt vs GFP-HELZ expression in HEK293T HELZ-null cells. GFP-HELZ was expressed in two different concentrations: 1 µg of transfected (GFP-)HELZ gave a similar level as endogenous HELZ in wt cells (also illustrated in A); in lane 3, 4 µg of GFP-HELZ was transfected.

Tubulin was used as a loading control. **C** qPCR results of up-regulated transcripts in HEK293T HELZ-null cell line identified in RNA-Seq normalised to *actin* transcript levels. The graph represent the log2 fold change of Ct values from HEK293T HELZ-null cell compared to HEK293T wt according to the Pfafll method. The y axis is in logarithmic base 10 scale. Target names are indicated on top of the graph on the x axis. Represented dots are biological triplicate values of normalised log2FC values; values for HEK293T wt were set to 1. Grey dots represent values from HEK293T wt samples, light green dots from HEK293T HELZ-null samples, green dots from HEK293T HELZ-null expressing 1μg of GFP-HELZ, dark green dots from HEK293T HELZ-null expressing 4μg of GFP-HELZ. The black bars in each lanes represent the average log2FC of the triplicate per sample and per transcript. **D** qPCR results of down-regulated transcripts in HEK293T HELZ-null cell line identified in RNA-Seq normalised to *actin* transcript levels. The graph is represented exactly like in C. HELZ is down-regulated in HEK293T HELZ-null cell.

5 Discussion

5.1 HELZ in mRNA regulation

The balance between mRNA biogenesis and degradation dictates the amount of mRNAs present at a certain time point in a cell. The biogenesis of mRNA takes place in the nucleus and involves the transcription of genes and the maturation of pre-mRNAs. Properly processed mRNAs are then exported to the cytoplasm where their message can be translated. Degradation removes faulty, unwanted and nonessential mRNAs from the cytoplasmic pool. Throughout its “life”, an mRNA is always bound by different RNA-binding proteins forming an mRNP. The composition of an mRNP is tightly regulated as it defines the location and the fate of the mRNA. RNA helicases are the designated enzymes that bind to and re-organise RNA structure and RNA-protein interactions in an ATP-dependent manner. In this thesis, I described my work on the putative RNA helicase HELZ which revealed itself to be a regulator of multiple steps of mRNA expression.

5.1.1 HELZ regulates the expression of endogenous transcripts

Little is known about the putative helicase HELZ, especially in regards to its mRNA targets or its biological impact. Thus, I performed an RNA-Seq analysis comparing the transcriptome of HEK293T cells expressing HELZ versus HEK293T HELZ-null cells to gain an insight into the cellular processes regulated by HELZ. The transcriptome analysis revealed close to 5 000 transcripts differentially expressed in HEK293T HELZ-null cells. Almost all transcripts identified were coding for proteins. GO analysis of those proteins revealed that they mostly function in the nervous system (TMEM35, BASP1, TENM1, MAOA, OSTN), the skeletal system (TMEM35, OSTN, SPARC), the immune system (HLA-DMA, IFIH1), regulating cell adherence and cellular communication (SPARC and PCDH10).

However, pleiotropic effects could be responsible for the changes observed in HEK293T HELZ-null cells as transcription factors were also identified in the transcriptome analysis. This could mean that only some targets are regulated directly by HELZ and a cascade of events led to the result in the RNA-Seq data. Moreover, the HEK293T HELZ-null cells are heavily modified cells compared to somatic cells in human and might not reflect the whole range of the transcripts targeted by HELZ or,

vice versa, might highlight false positive targets. Yet, I could reveal a direct causal link between the absence of HELZ and some transcripts identified in the RNA-Seq analyses by performing add-back experiments. Indeed, when HELZ was exogenously expressed in HEK293T HELZ-null cells, the expression level of some transcripts were increased or decreased to reach a similar level as measured in HEK293T wt cells. Thus, HELZ could directly regulate the level of expression of some of those specific transcripts, likely through the combination of transcriptional and post-transcriptional events. Furthermore, this is the first time that a list of potential transcripts regulated by HELZ has been produced, linking the protein to defined biological systems.

Transcriptional control is the first step of regulation of gene expression. Interestingly, HELZ was identified to be part of two complexes regulating transcription: with RNA pol II and histone methyltransferases 2 and 3 (SMYD2/3). The complex SMYD2-HELZ-RNA pol II was proposed to regulate the transcription of genes related to mRNA translation in cardiac cells (Diehl et al. 2010). The complex SMYD3-HELZ-RNA pol II was implied to regulate oncogenes, developmental-related genes and genes involved in cell metabolism in HEK293 cells (Hamamoto et al. 2004). Although not tested, HELZ might exert transcriptional control of the transcripts identified in the RNA-Seq in complex with RNA pol II and SMYD2/3. It would be interesting to understand the mechanism of action and target specificity of these different complexes, and how HELZ contribute to the function of these complexes.

Interestingly, HELZ seem to regulate the expression of targets that overlap with its assumed localisation in mice embryos (Wagner et al. 1999). Wagner and colleagues tagged the *HELZ* gene in mice and observed a dynamic, complex and wide-spread pattern in all stages of the developing embryo (Wagner et al. 1999). Moreover, the authors noted strong signals in the neural tube and notochord (which will later form the nervous system), in the forming skeletal system and in developing organs like the heart myocardium, ears, eyes and limbs (Wagner et al. 1999). In the adult mouse, *HELZ* mRNA seems to be predominantly expressed in the brain and testis, and at lower levels in the heart and kidneys (Wagner et al. 1999; Hasgall et al. 2011). This observation hints to the potential conservation of the mRNA targets of HELZ in mammals.

5.1.2 HELZ induces degradation of reporter mRNAs via the CCR4-NOT complex

Post-transcriptional control of gene expression includes cytoplasmic decay of mRNAs. The shortening and removal of the 3'polyA tail by the CCR4-NOT complex prompt the total decay of the mRNA via the 5'-to-3' or the 3'-to-5' decay pathways. To regulate which mRNAs are to be degraded, the CCR4-NOT complex is recruited by various RNA-binding/associated proteins to specific transcripts. Work done by several labs indicate that the recruitment of the CCR4-NOT complex frequently relies on short linear motifs (SLiMs) embedded in low-complexity region of the RNA-binding/associated protein (Chicoine et al. 2007; Horiuchi et al. 2009; Braun et al. 2011; Van Etten et al. 2012; Fabian et al. 2013; Loh et al. 2013; Chen et al. 2014; Mathys et al. 2014; Raisch et al. 2016; Rambout et al. 2016; Sgromo et al. 2017; Keskeny et al. 2019). For example, human Nanos proteins have SLiMs known as NIMs (NOT1-interacting motif) in their unstructured N-terminal region that mediate the interaction with the NOT module of the CCR4-NOT complex (Raisch et al. 2016). In the case of HELZ, I identified multiple fragments in the C-terminal region directly interacting with the NOT module of the CCR4-NOT complex. HELZ seems to follow this common trend of associating with the deadenylase complex via multiple binding sites located in the unstructured and low-complexity region of the C-terminal tail, though the regions are still too long to pinpoint the presence of defined SLiMs.

Interestingly, both the CCR4-NOT complex deadenylase activity and its assembly were required for HELZ to respectively induce mRNA decay and to repress translation of a reporter mRNA in tethering assay in human cells. I could demonstrate that the 5'-to-3' decay pathway was used by HELZ to degrade reporter mRNAs by inhibiting the CCR4-NOT deadenylase activity and the decapping activity of the DCP1/DCP2 complex. The translational repression effect induced by HELZ, when it is tethered to a reporter mRNA lacking a 3'polyA tail, was impaired when the scaffolding protein CNOT1 was reduced, which would affect the general assembly of the complex. The RNA helicase DDX6 seemed to also contribute partially to the repression of mRNA translation induced by HELZ. I could not detect any interaction between HELZ and DDX6 in human cells, in contrast to what was previously reported (Ayache et al. 2015). However, it is likely that DDX6 is recruited to HELZ-bound reporter mRNAs via the CCR4-NOT complex as both helicases bind to different regions of the CCR4-NOT

complex. Taken together, my data illustrate the functional significance of the interaction between HELZ and the CCR4-NOT complex.

Furthermore, as mentioned earlier, I could confirm a causal link between the absence of HELZ and five transcripts that were up-regulated in add-back experiments, in other words introducing back HELZ in HEK293T HELZ-null cells lowered the expression of mRNA of those five transcripts to similar levels of expression measured in HEK293T cells. These transcripts might be bound by HELZ which could then recruit the CCR4-NOT complex to induce their decay. Hence, it would explain why these transcripts are up-regulated when HELZ is not present in cells as they are degraded in presence of HELZ.

In conclusion, I described in this thesis a physical and functional link between HELZ and the CCR4-NOT complex. Future work determining the molecular details of this interaction and which transcripts are regulated by HELZ through the recruitment of the CCR4-NOT complex will help our understanding of this newly identified function of HELZ in the regulation of post-transcriptional gene expression.

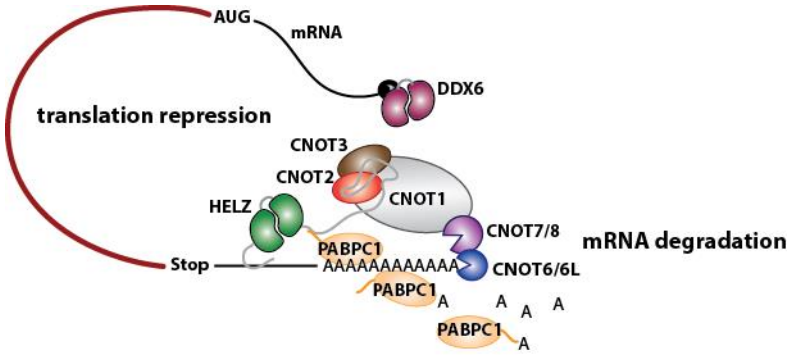

Figure 25 HELZ interacts with the CCR4-NOT complex and commits bound mRNA to degradation. In this model, HELZ (pictured as two green RecA-like domains with N- and C-extensions) is bound to an mRNA and recruits the CCR4-NOT complex. This causes translational repression and concurrent mRNA decay of reporter mRNAs. HELZ interacts with the CCR4-NOT complex by binding to the NOT module via multiple binding sites in its long C-terminal tail. HELZ also interacts with PABPC1 but not with DDX6.

5.1.3 RNA helicases and the CCR4-NOT complex

Interestingly, the CCR4-NOT complex was reported to interact with two other RNA helicases, in line with its function as a global regulator of gene expression. Indeed, the CCR4-NOT complex is associated with the RNA helicase DDX6 and UPF1. DDX6 is part of the DEAD-box SF2 helicase family and consist of a helicase domain and very short N- and C-extensions (Jonas and Izaurralde 2013). DDX6 was not reported to bind specific mRNAs but seems to bridge different complexes, like the

deadenylase and the decapping complexes, to regulate translational repression and mRNA decay (Jonas and Izaurralde 2013; Chen et al. 2014; Mathys et al. 2014). DDX6 interacts with the MIF4G domain of CNOT1 with its C-terminal RecA-like folds of the helicase domain. Structures of the interaction have been solved and revealed a typical MIF4G-RecA-like interaction that could regulate the helicase activity of DDX6 (Chen et al. 2014; Mathys et al. 2014). The interaction between DDX6 and the CCR4-NOT complex is therefore quite different to how HELZ interacts with the deadenylase complex. It is unlikely that the CCR4-NOT complex could regulate the helicase activity HELZ but HELZ could identify mRNA targets via its helicase domain and bring the CCR4-NOT complex to induce mRNA decay.

The other helicase is the SF1 helicase UPF1, from the same helicase family as HELZ. UPF1 protein contains a central helicase core flanked by a CH domain (rich in Cysteine and Histidine amino acids) at the N-terminal part, and an SQ domain (rich in Serine and Glutamine amino acids) at the C-terminal part of the protein (Fiorini et al. 2015; Gupta and Li 2018). The N- and C-terminal domains inhibit the activity of the UPF1 helicase (Franks et al. 2010; Chakrabarti et al. 2011; Chakrabarti et al. 2014). The release of this regulatory inhibition is achieved by binding to UPF2 in the case of the CH domain and by post-translational modifications, mostly phosphorylation, of the SQ domain. These events are occurring during NMD in a step-wise manner leading to the recruitment of either the endonuclease SMG6 or the SMG5-SMG7 heterodimer by phosphorylated UPF1. SMG7 has a strong preference for binding to CNOT8 compared to CNOT7 and by this direct interaction, the CCR4-NOT complex is recruited to NMD targets (Loh et al. 2013). The NMD-targeted mRNAs are then degraded by the 5'-to-3' decay pathway (Loh et al. 2013; Schweingruber et al. 2013). Thus, UPF1 uses the CCR4-NOT complex to degrade NMD-targets and recruits the deadenylase complex by its interaction with SMG7 which is regulated by phosphorylation events.

Similar to UPF1, HELZ might require the CCR4-NOT complex to induce the degradation of specific mRNAs. These mRNAs could be targeted because of defects identified by HELZ. Unlike UPF1, HELZ directly interacts with the CCR4-NOT complex. I have not explored the impact of post-translational modifications on this interaction. However, several large-scale mass spectroscopy analyses have revealed potential phosphorylation sites in the C-terminal tail and one phosphorylation site in the N-terminal region of HELZ (Hornbeck et al. 2015). Post-translational modifications could therefore account for the differences observed in cells and *in vitro* in the binding of

HELZ_C1 fragment to the CCR4-NOT complex. The amino acid residues involved in the interaction with the CCR4-NOT complex in this fragment might be masked or regulated by post-translational modifications in cells.

5.1.4 HELZ: a cell-specific regulator of mRNA translation?

Another major regulation step of gene expression in the cytoplasm is the control of mRNA translation: which mRNAs are to be translated, when should specific proteins be produced and how much proteins need to be expressed in a given cell. mRNA translation is tightly regulated and controlled by several signalling pathways (Roux and Topisirovic 2012). I analysed the effect of HELZ on general mRNA translation using a similar method as Hasgall and colleagues (Hasgall et al. 2011). However, I could not detect any difference between the profiles of HEK293T cells and HE293T HELZ-null cells, which is in contrast to what Hasgall et al. observed (a decrease in the expression of HELZ lead to a reduction of general translation). Thus, I conclude that HELZ does not affect general protein translation in HEK293T cells. There are several differences between my experiment and that of Hasgall and colleagues (Hasgall et al. 2011). One difference is the cell line used in each experiments: HEK293T cells were used in this thesis vs HeLa cells in the report of Hasgall et al.. Another difference is associated with the method used to block the expression of HELZ: I used HELZ-null cells generated by the CRISPR/Cas9 technology, whereas Hasgall and collaborators performed siRNA-based knock-down of gene expression. Both differences could be responsible for the discrepancy between the published results and the results presented in this thesis.

Interestingly, the effect on general mRNA translation is not the only difference that is observed between published results and the results in this thesis. Indeed, the cellular localisation of HELZ has been reporter to be different in different cell lines. Wagner et al. described a general cytoplasmic and extra-cellular localisation of the reporter gene expressed under the control of *HELZ* promoter in different tissues of the mouse embryo (Wagner et al. 1999). Hasgall and colleagues observed a predominant cytoplasmic signal for human HELZ protein in immunofluorescence assay in MCF-7 cells (a breast cancer cell line) (Hasgall et al. 2011). Youn et al. identified HELZ to be part of P-bodies and SG in human embryonic kidney cells (HEK293), both are cytoplasmic foci (Youn et al. 2018). In my immunofluorescence experiments in cervical

cancer cells (HeLa), I observed that MS2-HA-HELZ is expressed as a diffuse cytoplasmic and nuclear protein, excluded from the nucleolus.

In conclusion, combining my data with already published results, it seems that the effect of HELZ on translation and its cellular localisation varies between different cell types. Moreover, given the possible role in transcriptional control, cytoplasmic mRNA decay and translation, the localisation of HELZ might be tightly regulated to achieve different functions in different cells.

5.1.5 HELZ: a potential new player in 3'polyA tail regulation?

HELZ proteins contain a highly conserved helicase core that retains all necessary motifs important for nucleic acid binding, ATP binding, co-ordination of ATP hydrolysis and nucleic acid binding, and ATP hydrolysis. It is part of the UPF1-like SF1 helicase family (Fairman-Williams et al. 2010). The activity of SF1 and SF2 helicases requires ATP hydrolysis; thus, the introduction of a point mutation in the conserved motif II (aka Walker B motif) abolishes the activity of a helicase (see the Introduction sub-chapter 1.3 Helicases). I introduced a point mutation in the motif II of HELZ (HELZ_E795Q) to study a potential function of the putative helicase in human cells. With this approach, I uncovered a surprising and unexpected effect of the helicase mutant on reporter mRNAs. Indeed, the 3'polyA tail of reporter mRNAs were elongated when the HELZ ATPase mutant was co-expressed in cells, and this was true for all reporters with a 3'polyA tail.

The elongation of the 3'polyA tail of reporter mRNAs did not require the tethering of the helicase mutant to reporter mRNAs, suggesting that it might affect endogenous mRNAs. This also suggest that HELZ might bind mRNAs in a non-specific manner, similarly to UPF1 and MOV10 (Gregersen et al. 2014). Further studies are required to clarify if this is the case. Moreover, the expression of an ATPase-dead mutant of UPF1 did not lead to this shift when similar reporter mRNAs and endogenous mRNAs were co-expressed and visualized by northern blot (Mendell et al. 2002; Franks et al. 2010). Indicating that the upward shift is specific for the function of HELZ, which seemingly possess an ATPase activity linked to its helicase domain.

Remarkably, HELZ is the only helicase that interacts with PABPC1 via a PAM2 motif (Albrecht and Lengauer 2004). The interaction between HELZ and PABC1, which binds to the 3'polyA tail of mRNAs, could be used as a signal for HELZ to determine the presence, length or absence of the 3'polyA tail of mRNAs. PABPC1 could modulate

the function of the helicase domain or, *vice versa*, the binding to PABPC1 could be regulated by the activity of the helicase as the PAM2 motif is in close proximity of the helicase domain (**Figure 3** A). However, the upward shift of the reporters observed when HELZ_E795Q was expressed was independent of PABPC1-binding. This suggests that the helicase mutant cannot respond to PABPC1 binding (and thus, the shift upward is independent of PABPC1 binding); nonetheless, this mutant also affected the length of the 3'polyA tail of reporter mRNAs.

Actually, the length of the 3'polyA tails of mRNAs is known to be not stable throughout an mRNA "life". The 3'polyA tails are dynamically regulated in the cytoplasm, after their addition in the nucleus, during embryonic development but also in somatic cells (Norbury 2013; Jalkanen et al. 2014; Nicholson and Pasquinelli 2018). The effect of the length of the 3'polyA tail of an mRNA on stability and/or translatability, is dependent on the transcript, on the composition of the mRNP and on the current cellular state (Jalkanen et al. 2014). The mechanism of cytoplasmic 3'polyA tail length involves deadenylases removing adenosines (As), PAPs adding As, and is described to be coordinated by CPEB proteins (cytoplasmic polyadenylation element-binding protein) that recognises *cis*-elements on mRNAs (Villalba et al. 2011; Fernandez-Miranda and Mendez 2012; Ivshina et al. 2015). However, the 3'polyA tail of mRNAs is also thought to be pruned once it is exported from the nucleus to the cytoplasm, i.e. the nascent 3'polyA tail can be shortened once it is in the cytoplasmic environment (Jalkanen et al. 2014). The balance between the activity of the two types of enzymes regulates 3'polyA tail lengths. PAPs are divided into two families: canonical PAPs (PAPα, β, γ) and non-canonical PAPs, the latter are often referred to as terminal uridylyltransferases (TUT 1 to 7) (Laishram 2014). The canonical PAPα and PAPγ are the major adenylases of pre-mRNAs in the nucleus, PAP β is testis specific (Jalkanen et al. 2014; Laishram 2014). So far, TUT2 (aka hGLD2 or PAPD4), TUT3 (aka hGLD4 or PAPD5), and TUT6 (aka Star-PAP or PAPD2) have been described to polyadenylate mRNAs in the cytoplasm and in the nucleus of human cells (Fernandez-Miranda and Mendez 2012; Laishram 2014). In regards to deadenylases, in addition to the CCR4-NOT deadenylase complex and the PAN2/PAN3 deadenylase complex, the protein PARN is a deadenylase of the DEDD family and can also deadenylate mRNAs in the cytoplasm (Godwin et al. 2013; Yan 2014).

Interestingly, MOV10, a paralogue of HELZ, has been recently identified to restrict the propagation of a retrotransposon by facilitating the uridylation of mRNAs by

TUTs (Warkocki et al. 2018). In a similar way, HELZ might use TUTs to regulate the length of the 3'polyA tail of mRNAs. The regulation of the activity of TUTs by HELZ would then be compromised in the case of the ATPase mutant and could lead to constant polyadenylation of the mRNAs, or uridylation/guanylation events which could protect the mRNA from deadenylases (Chang et al. 2014b; Lim et al. 2018). Another possibility that is supported by my data is that HELZ interacts with two deadenylase complexes in cells (the CCR4-NOT complex and the PAN2/PAN3 complex), making HELZ a likely candidate to regulate their activities. The helicase mutant could inhibit or prevent the mRNA from deadenylation. The regulation of TUTs and/or deadenylases are not exclusive and a combination of both could be responsible for the regulation of the 3'polyA tails of mRNA by HELZ (**Figure 26**). For example, CPEB controls the 3'polyA tail of specific mRNAs by interacting and modulating the activity of both PARN and TUT2/3 (Villalba et al. 2011; Fernandez-Miranda and Mendez 2012).

Taken together, HELZ could act as a sensor of 3'polyA tails and inducing either polyadenylation or deadenylation depending on, as yet, unknown factors (**Figure 26**). HELZ could also rely on the direct recruitment of the CCR4-NOT complex to induce the decay of targeted mRNAs. This model is close to the proposed function of UPF1 that can act as a 3'UTR length sensor (Hogg and Goff 2010).

In conclusion, I provide evidence that HELZ has an ATPase activity linked to its helicase domain and that it could play a role in the regulation of 3'polyA tail length. To clarify the function of HELZ in 3'polyA tail regulation, there is a need to define which transcripts might be affected by the ATPase mutant and to fully characterise the helicase activity *in vitro* (in presence and absence of PABPC1). Moreover, it would be interesting to test if HELZ regulates the activity of deadenylases and/or TUTs and the outcome of this regulation (**Figure 26**).

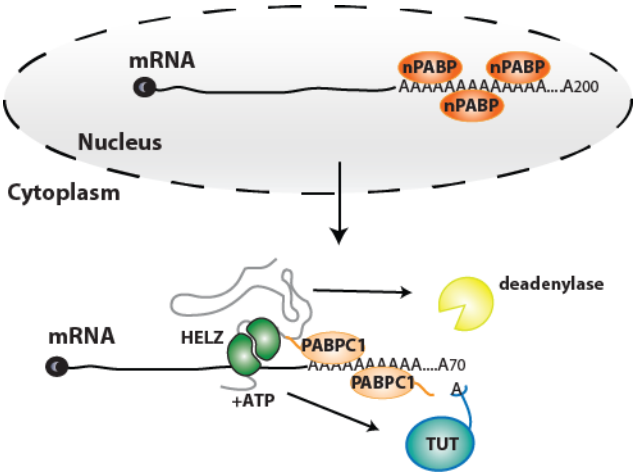

Figure 26 A possible function of HELZ in 3'polyA tail regulation. HELZ is present in the cytoplasm and in the nucleus (not represented). The 3'polyA tail is generated by canonical PAPs in the nucleus and about 200 nucleotides are added to the mRNA. The nascent 3'polyA tail is bound by nPABPs. In the cytoplasm, nPABP are replaced by PABPC1 and the 3'polyA tail can be shorter or longer than the original 3'polyA tail. the function of HELZ might be to sense 3'polyA tail length and to regulate deadenylases and/or TUTs accordingly.

5.2 Concluding remarks

Studying HELZ proved to be challenging but full of intriguing discoveries. Before this work, the putative helicase HELZ was described in transcriptional control and the regulation of mRNA translation. I identified several transcripts that are affected (positively and negatively) by HELZ, giving for the first time a biological context to the protein in cells. I further unveiled new functions of HELZ in post-transcriptional cytoplasmic mRNA regulation. My data indicate that HELZ is an active helicase as interfering with the ATPase activity of the helicase resulted in a lack-of-function phenotype. This phenotype exposed a potential new role of HELZ in 3'polyA tail length regulation. In agreement with this function, HELZ interacts with two deadenylase complexes, the CCR4-NOT and the PAN2/PAN3 complexes, and with the cytoplasmic polyA binding protein PABPC1. Finally, I demonstrate that HELZ interacts directly and functionally with the CCR4-NOT complex in human cells. In the cytoplasm, HELZ requires the activity and a fully assemble CCR4-NOT complex to induce translational repression and decay of mRNA reporters in tethering assay. Based on my data and published work, I draw the outline of a model where HELZ regulates gene expression by modulating transcription, mRNA processing, translation and mRNA decay of specific transcripts.

Acknowledgments

I would like to thank Prof. Dr. Elisa Izaurralde for giving me the opportunity to do my PhD in her lab, for her support and her teachings.
I am grateful to Prof. Dr. Doron Rapaport and Prof. Dr. Remco Sprangers for giving me advices throughout the years. I would like to thank Prof. Dr. Ralf Sommer and Prof. Dr. Ralf-Peter Jansen for seeing me through my PhD. Thanks also to Dr. Sarah Danes and Dr. Dagmar Sigurdardottir for co-ordinating the PhD school/program. A special thanks to Dr. Lara Wohlbold and Dr. Heike Budde for their critical feedback on the thesis.

I would like to thank all past and current members of the department II. All of you helped me in your own way to grow as a scientist and as a person. I would like to mention especially Annamaria Sgromo for our time in the lab but as well as all the trips and escapades that made our time even more enjoyable.

I am grateful to my friends who came down to visit and for not letting time nor space between our friendship.

I thank from the bottom of my heart my family for their constant support, advice and love. Thanks for always being there for me and for charging my batteries.

Last but not least, I would like to thank Julian Harbarth for his support, his love and for our home.

Finally, thank you the reader.

Curriculum Vitae

Name: **Aoife Julia Hanet**

Date of birth: **11th April 1989**

Nationality: **Belgian - Irish**

Work experience:

10.2013-04.2019: **PhD candidate**, Department of Biochemistry, Max Planck Institute for Developmental Biology

Title: Roles of the putative helicase HELZ in mRNA metabolism

09.2015-03.2017: **External and Internal PhD representative** of Max Planck Institute of Developmental Biology; **Head of Collect team** of the Survey Workgroup and Secretary Workgroup of the PhDnet

Organised the PhD retreat 2016 (36 participants, 4 international guest speakers, 3 days long)

Organised and managed teams remotely to collect data for the 2017 PhDnet survey, encompassing all Max Planck institutes world-wide

2003, 2010, 2012: **Interpreter English-French/French–English** and support service for the Worldwide Indoor Soccer Police Association (WISPA)

Week-long events on indoor soccer tournament involving the police/army force of ~40 countries and enabling discussions and exchanges about security topics

Attended: Durban, South Africa (2013), Wavre, Belgium (2010 – 10 years anniversary), and Moscow, Russia (2012)

2006-2008: singlehandedly managed the dining-room in the restaurant "Gavroche" in Wavre, Belgium

07.2006: Student job in Tractabel- engineering, Brussels, Belgium

Education:

09.2011-09.2013: **M.Sc. in Biomedical Sciences** from the University Catholique de Louvain, Brussels, Belgium

*Title: Study of miR-128 and one of its target of unknown function, SZR*D1

Erasmus poster: Effect of FDA approved drugs on Ataxin-3 subcellular localization

09.2007-08.2011: **B.Sc. in Biomedical Sciences** from the University Catholique de Louvain, Brussels, Belgium

Title: Set up of the SNaPshot technique for the detection of SNPs in exon 6 of the interleukin-7 gene in Rheumatoid Arthritis patient

06.2007: Institute Saint-Jean Baptiste, Wavre, Belgium, General education certificate (major: science)

Language skills:

Native: English, French

Other languages: German (B2), Dutch (A2), Japanese (A1)

Remarkable achievements:

Recipient of a **doctoral funding grant** from the selective IMPRS “form Molecules to Organisms”, Tübingen, Germany

Presented own scientific work as a **poster at the international conference EMBO** Protein Synthesis and Translational Control 2015

Work published in peer-reviewed scientific journals:

-Hanet A., Fauser M., Raisch T., Kuzuoğlu-Öztürk D., Chang C.-T., Bhandari D., Igreja* C., Wohlbold* L., Izaurralde E. submitted 04.2019. HELZ directly interacts with CCR4-NOT and causes decay of bound mRNAs

-Wang Z.-J., Hanet A., Weishäupl D., Martins I.M., Sowa A.S., Riess O., Schmidt T. CNS Neuroscience and Therapeutics 2018. Divalproex sodium modulates nuclear localisation of ataxin-3 and prevents cellular toxicity caused by expanded ataxin-3

Reference

- Albrecht, M. and T. Lengauer (2004). "Survey on the PABC recognition motif PAM2." Biochem Biophys Res Commun **316**(1): 129-138.
- Ariumi, Y. (2016). "Guardian of the Human Genome: Host Defense Mechanisms against LINE-1 Retrotransposition." Front Chem **4**: 28.
- Arjan-Odedra, S., C. M. Swanson, N. M. Sherer, S. M. Wolinsky and M. H. Malim (2012). "Endogenous MOV10 inhibits the retrotransposition of endogenous retroelements but not the replication of exogenous retroviruses." Retrovirology **9**: 53.
- Ayache, J., M. Benard, M. Ernoult-Lange, N. Minshall, N. Standart, M. Kress and D. Weil (2015). "P-body assembly requires DDX6 repression complexes rather than decay or Ataxin2/2L complexes." Mol Biol Cell **26**(14): 2579-2595.
- Aylett, C. H. and N. Ban (2017). "Eukaryotic aspects of translation initiation brought into focus." Philos Trans R Soc Lond B Biol Sci **372**(1716).
- Bartlam, M. and T. Yamamoto (2010). "The structural basis for deadenylation by the CCR4-NOT complex." Protein Cell **1**(5): 443-452.
- Bawankar, P., B. Loh, L. Wohlbold, S. Schmidt and E. Izaurralde (2013). "NOT10 and C2orf29/NOT11 form a conserved module of the CCR4-NOT complex that docks onto the NOT1 N-terminal domain." RNA Biol **10**(2): 228-244.
- Berger, I., F. Garzoni, M. Chaillet, M. Haffke, K. Gupta and A. Aubert (2013). "The multiBac protein complex production platform at the EMBL." J Vis Exp(77): e50159.
- Bhandari, D., T. Raisch, O. Weichenrieder, S. Jonas and E. Izaurralde (2014). "Structural basis for the Nanos-mediated recruitment of the CCR4-NOT complex and translational repression." Genes Dev **28**(8): 888-901.
- Bhaskar, V., V. Roudko, J. Basquin, K. Sharma, H. Urlaub, B. Seraphin and E. Conti (2013). "Structure and RNA-binding properties of the Not1-Not2-Not5 module of the yeast Ccr4-Not complex." Nat Struct Mol Biol **20**(11): 1281-1288.
- Boland, A., Y. Chen, T. Raisch, S. Jonas, D. Kuzuoglu-Ozturk, L. Wohlbold, et al. (2013). "Structure and assembly of the NOT module of the human CCR4-NOT complex." Nat Struct Mol Biol **20**(11): 1289-1297.

- Bos, T. J., J. K. Nussbacher, S. Aigner and G. W. Yeo (2016). "Tethered Function Assays as Tools to Elucidate the Molecular Roles of RNA-Binding Proteins." Adv Exp Med Biol **907**: 61-88.
- Bourgeois, C. F., F. Mortreux and D. Auboeuf (2016). "The multiple functions of RNA helicases as drivers and regulators of gene expression." Nat Rev Mol Cell Biol **17**(7): 426-438.
- Braun, J. E., E. Huntzinger, M. Fauser and E. Izaurralde (2011). "GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets." Mol Cell **44**(1): 120-133.
- Braun, J. E., F. Tritschler, G. Haas, C. Igreja, V. Truffault, O. Weichenrieder and E. Izaurralde (2010). "The C-terminal alpha-alpha superhelix of Pat is required for mRNA decapping in metazoa." EMBO J **29**(14): 2368-2380.
- Braun, J. E., V. Truffault, A. Boland, E. Huntzinger, C. T. Chang, G. Haas, et al. (2012). "A direct interaction between DCP1 and XRN1 couples mRNA decapping to 5' exonucleolytic degradation." Nat Struct Mol Biol **19**(12): 1324-1331.
- Brummelkamp, T. R., R. Bernards and R. Agami (2002). "A system for stable expression of short interfering RNAs in mammalian cells." Science **296**(5567): 550-553.
- Calviello, L., N. Mukherjee, E. Wyler, H. Zauber, A. Hirsekorn, M. Selbach, et al. (2016). "Detecting actively translated open reading frames in ribosome profiling data." Nat Methods **13**(2): 165-170.
- Chakrabarti, S., F. Bonneau, S. Schussler, E. Eppinger and E. Conti (2014). "Phospho-dependent and phospho-independent interactions of the helicase UPF1 with the NMD factors SMG5-SMG7 and SMG6." Nucleic Acids Res **42**(14): 9447-9460.
- Chakrabarti, S., U. Jayachandran, F. Bonneau, F. Fiorini, C. Basquin, S. Domcke, et al. (2011). "Molecular mechanisms for the RNA-dependent ATPase activity of Upf1 and its regulation by Upf2." Mol Cell **41**(6): 693-703.
- Chang, C. T., N. Bercovich, B. Loh, S. Jonas and E. Izaurralde (2014a). "The activation of the decapping enzyme DCP2 by DCP1 occurs on the EDC4 scaffold and involves a conserved loop in DCP1." Nucleic Acids Res **42**(8): 5217-5233.

- Chang, H., J. Lim, M. Ha and V. N. Kim (2014b). "TAIL-seq: genome-wide determination of poly(A) tail length and 3' end modifications." Mol Cell **53**(6): 1044-1052.
- Chen, C. Y., D. Zheng, Z. Xia and A. B. Shyu (2009). "Ago-TNRC6 triggers microRNA-mediated decay by promoting two deadenylation steps." Nat Struct Mol Biol **16**(11): 1160-1166.
- Chen, Y., A. Boland, D. Kuzuoglu-Ozturk, P. Bawankar, B. Loh, C. T. Chang, et al. (2014). "A DDX6-CNOT1 complex and W-binding pockets in CNOT9 reveal direct links between miRNA target recognition and silencing." Mol Cell **54**(5): 737-750.
- Chicoine, J., P. Benoit, C. Gamberi, M. Paliouras, M. Simonelig and P. Lasko (2007). "Bicaudal-C recruits CCR4-NOT deadenylase to target mRNAs and regulates oogenesis, cytoskeletal organization, and its own expression." Dev Cell **13**(5): 691-704.
- Collart, M. A. (2016). "The Ccr4-Not complex is a key regulator of eukaryotic gene expression." Wiley Interdiscip Rev RNA **7**(4): 438-454.
- Collart, M. A. and O. O. Panasenko (2017). "The Ccr4-Not Complex: Architecture and Structural Insights." Subcell Biochem **83**: 349-379.
- Coller, J. and R. Parker (2005). "General translational repression by activators of mRNA decapping." Cell **122**(6): 875-886.
- Cordin, O. and J. D. Beggs (2013). "RNA helicases in splicing." RNA Biol **10**(1): 83-95.
- Davey, N. E., M. S. Cyert and A. M. Moses (2015). "Short linear motifs - ex nihilo evolution of protein regulation." Cell Commun Signal **13**: 43.
- De, I., S. Bessonov, R. Hofele, K. dos Santos, C. L. Will, H. Urlaub, et al. (2015). "The RNA helicase Aquarius exhibits structural adaptations mediating its recruitment to spliceosomes." Nat Struct Mol Biol **22**(2): 138-144.
- Dehghani-Tafti, S. and C. M. Sanders (2017). "DNA substrate recognition and processing by the full-length human UPF1 helicase." Nucleic Acids Res **45**(12): 7354-7366.
- Delaleau, M. and K. L. Borden (2015). "Multiple Export Mechanisms for mRNAs." Cells **4**(3): 452-473.
- Dever, T. E. and R. Green (2012). "The elongation, termination, and recycling phases of translation in eukaryotes." Cold Spring Harb Perspect Biol **4**(7): a013706.

- Diehl, F., M. A. Brown, M. J. van Amerongen, T. Novoyatleva, A. Wietelmann, J. Harriss, et al. (2010). "Cardiac deletion of Smyd2 is dispensable for mouse heart development." PLoS One **5**(3): e9748.
- Eliscovich, C. and R. H. Singer (2017). "RNP transport in cell biology: the long and winding road." Curr Opin Cell Biol **45**: 38-46.
- Eliseeva, I. A., D. N. Lyabin and L. P. Ovchinnikov (2013). "Poly(A)-binding proteins: structure, domain organization, and activity regulation." Biochemistry (Mosc) **78**(13): 1377-1391.
- Fabian, M. R., F. Frank, C. Rouya, N. Siddiqui, W. S. Lai, A. Karetnikov, et al. (2013). "Structural basis for the recruitment of the human CCR4-NOT deadenylase complex by tristetraprolin." Nat Struct Mol Biol **20**(6): 735-739.
- Fabian, M. R. and N. Sonenberg (2012). "The mechanics of miRNA-mediated gene silencing: a look under the hood of miRISC." Nat Struct Mol Biol **19**(6): 586-593.
- Fairman-Williams, M. E., U. P. Guenther and E. Jankowsky (2010). "SF1 and SF2 helicases: family matters." Curr Opin Struct Biol **20**(3): 313-324.
- Fernandez-Miranda, G. and R. Mendez (2012). "The CPEB-family of proteins, translational control in senescence and cancer." Ageing Res Rev **11**(4): 460-472.
- Ficner, R., A. Dickmanns and P. Neumann (2017). "Studying structure and function of spliceosomal helicases." Methods **125**: 63-69.
- Fiorini, F., D. Bagchi, H. Le Hir and V. Croquette (2015). "Human Upf1 is a highly processive RNA helicase and translocase with RNP remodelling activities." Nat Commun **6**: 7581.
- Frankish, A., M. Diekhans, A. M. Ferreira, R. Johnson, I. Jungreis, J. Loveland, et al. (2018). "GENCODE reference annotation for the human and mouse genomes." Nucleic Acids Res.
- Franks, T. M., G. Singh and J. Lykke-Andersen (2010). "Upf1 ATPase-dependent mRNP disassembly is required for completion of nonsense- mediated mRNA decay." Cell **143**(6): 938-950.
- Fu, M. and P. J. Blackshear (2017). "RNA-binding proteins in immune regulation: a focus on CCCH zinc finger proteins." Nat Rev Immunol **17**(2): 130-143.
- Fu, Q., R. R. Pandey, N. A. Leu, R. S. Pillai and P. J. Wang (2016). "Mutations in the MOV10L1 ATP Hydrolysis Motif Cause piRNA Biogenesis Failure and Male Sterility in Mice." Biol Reprod **95**(5): 103.

- Gaidatzis, D., A. Lerch, F. Hahne and M. B. Stadler (2015). "QuasR: quantification and annotation of short reads in R." Bioinformatics **31**(7): 1130-1132.
- Gatfield, D., L. Unterholzner, F. D. Ciccarelli, P. Bork and E. Izaurralde (2003). "Nonsense-mediated mRNA decay in Drosophila: at the intersection of the yeast and mammalian pathways." EMBO J **22**(15): 3960-3970.
- Gehring, N. H., E. Wahle and U. Fischer (2017). "Deciphering the mRNP Code: RNA-Bound Determinants of Post-Transcriptional Gene Regulation." Trends Biochem Sci **42**(5): 369-382.
- Godwin, A. R., S. Kojima, C. B. Green and J. Wilusz (2013). "Kiss your tail goodbye: the role of PARN, Nocturnin, and Angel deadenylases in mRNA biology." Biochim Biophys Acta **1829**(6-7): 571-579.
- Goldstrohm, A. C. and M. Wickens (2008). "Multifunctional deadenylase complexes diversify mRNA control." Nat Rev Mol Cell Biol **9**(4): 337-344.
- Gregersen, L. H., M. Schueler, M. Munschauer, G. Mastrobuoni, W. Chen, S. Kempa, et al. (2014). "MOV10 Is a 5' to 3' RNA helicase contributing to UPF1 mRNA target degradation by translocation along 3' UTRs." Mol Cell **54**(4): 573-585.
- Grudzien-Nogalska, E. and M. Kiledjian (2017). "New insights into decapping enzymes and selective mRNA decay." Wiley Interdiscip Rev RNA **8**(1).
- Guenther, U. P., L. Handoko, B. Laggerbauer, S. Jablonka, A. Chari, M. Alzheimer, et al. (2009). "IGHMBP2 is a ribosome-associated helicase inactive in the neuromuscular disorder distal SMA type 1 (DSMA1)." Hum Mol Genet **18**(7): 1288-1300.
- Gupta, I., Z. Villanyi, S. Kassem, C. Hughes, O. O. Panasenko, L. M. Steinmetz and M. A. Collart (2016). "Translational Capacity of a Cell Is Determined during Transcription Elongation via the Ccr4-Not Complex." Cell Rep **15**(8): 1782-1794.
- Gupta, P. and Y. R. Li (2018). "Upf proteins: highly conserved factors involved in nonsense mRNA mediated decay." Mol Biol Rep **45**(1): 39-55.
- Hamamoto, R., Y. Furukawa, M. Morita, Y. Iimura, F. P. Silva, M. Li, et al. (2004). "SMYD3 encodes a histone methyltransferase involved in the proliferation of cancer cells." Nat Cell Biol **6**(8): 731-740.
- Harlen, K. M. and L. S. Churchman (2017). "The code and beyond: transcription regulation by the RNA polymerase II carboxy-terminal domain." Nat Rev Mol Cell Biol **18**(4): 263-273.

- Hasgall, P. A., D. Hoogewijs, M. B. Faza, V. G. Panse, R. H. Wenger and G. Camenisch (2011). "The putative RNA helicase HELZ promotes cell proliferation, translation initiation and ribosomal protein S6 phosphorylation." PLoS One **6**(7): e22107.
- Hogg, J. R. and S. P. Goff (2010). "Upf1 senses 3'UTR length to potentiate mRNA decay." Cell **143**(3): 379-389.
- Horiuchi, M., K. Takeuchi, N. Noda, N. Muroya, T. Suzuki, T. Nakamura, et al. (2009). "Structural basis for the antiproliferative activity of the Tob-hCaf1 complex." J Biol Chem **284**(19): 13244-13255.
- Hornbeck, P. V., B. Zhang, B. Murray, J. M. Kornhauser, V. Latham and E. Skrzypek (2015). "PhosphoSitePlus, 2014: mutations, PTMs and recalibrations." Nucleic Acids Res **43**(Database issue): D512-520.
- Houseley, J. and D. Tollervey (2009). "The many pathways of RNA degradation." Cell **136**(4): 763-776.
- Hubstenberger, A., M. Courel, M. Benard, S. Souquere, M. Ernoult-Lange, R. Chouaib, et al. (2017). "P-Body Purification Reveals the Condensation of Repressed mRNA Regulons." Mol Cell **68**(1): 144-157 e145.
- Ivshina, M., I. M. Alexandrov, A. Vertii, S. Doxsey and J. D. Richter (2015). "CPEB regulation of TAK1 synthesis mediates cytokine production and the inflammatory immune response." Mol Cell Biol **35**(3): 610-618.
- Iwakawa, H. O. and Y. Tomari (2015). "The Functions of MicroRNAs: mRNA Decay and Translational Repression." Trends Cell Biol **25**(11): 651-665.
- Jalkanen, A. L., S. J. Coleman and J. Wilusz (2014). "Determinants and implications of mRNA poly(A) tail size--does this protein make my tail look big?" Semin Cell Dev Biol **34**: 24-32.
- Jankowsky, E. (2011). "RNA helicases at work: binding and rearranging." Trends Biochem Sci **36**(1): 19-29.
- Jiang, F. and J. A. Doudna (2017). "CRISPR-Cas9 Structures and Mechanisms." Annu Rev Biophys **46**: 505-529.
- Jonas, S. and E. Izaurralde (2013). "The role of disordered protein regions in the assembly of decapping complexes and RNP granules." Genes Dev **27**(24): 2628-2641.

- Jonas, S. and E. Izaurralde (2015). "Towards a molecular understanding of microRNA-mediated gene silencing." Nat Rev Genet **16**(7): 421-433.
- Kanaan, J., S. Raj, L. Decourty, C. Saveanu, V. Croquette and H. Le Hir (2018). "UPF1-like helicase grip on nucleic acids dictates processivity." Nat Commun **9**(1): 3752.
- Kapp, L. D. and J. R. Lorsch (2004). "The molecular mechanics of eukaryotic translation." Annu Rev Biochem **73**: 657-704.
- Keskeny, C., T. Raisch, A. Sgromo, C. Igreja, D. Bhandari, O. Weichenrieder and E. Izaurralde (2019). "A conserved CAF40-binding motif in metazoan NOT4 mediates association with the CCR4-NOT complex." Genes Dev.
- Kilchert, C., S. Wittmann and L. Vasiljeva (2016). "The regulation and functions of the nuclear RNA exosome complex." Nat Rev Mol Cell Biol **17**(4): 227-239.
- Kim, D., G. Pertea, C. Trapnell, H. Pimentel, R. Kelley and S. L. Salzberg (2013). "TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions." Genome Biol **14**(4): R36.
- Koressaar, T. and M. Remm (2007). "Enhancements and modifications of primer design program Primer3." Bioinformatics **23**(10): 1289-1291.
- Kozlov, G., M. Menade, A. Rosenauer, L. Nguyen and K. Gehring (2010). "Molecular determinants of PAM2 recognition by the MLLE domain of poly(A)-binding protein." J Mol Biol **397**(2): 397-407.
- Kuzuoglu-Ozturk, D., D. Bhandari, E. Huntzinger, M. Fauser, S. Helms and E. Izaurralde (2016). "miRISC and the CCR4-NOT complex silence mRNA targets independently of 43S ribosomal scanning." EMBO J **35**(11): 1186-1203.
- Labno, A., R. Tomecki and A. Dziembowski (2016). "Cytoplasmic RNA decay pathways - Enzymes and mechanisms." Biochim Biophys Acta **1863**(12): 3125-3147.
- Labun, K., T. G. Montague, J. A. Gagnon, S. B. Thyme and E. Valen (2016). "CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering." Nucleic Acids Res **44**(W1): W272-276.
- Laishram, R. S. (2014). "Poly(A) polymerase (PAP) diversity in gene expression--star-PAP vs canonical PAP." FEBS Lett **588**(14): 2185-2197.
- Langmead, B. and S. L. Salzberg (2012). "Fast gapped-read alignment with Bowtie 2." Nat Methods **9**(4): 357-359.

- Lazzaretti, D., I. Tournier and E. Izaurralde (2009). "The C-terminal domains of human TNRC6A, TNRC6B, and TNRC6C silence bound transcripts independently of Argonaute proteins." RNA **15**(6): 1059-1066.
- Lee, T. I. and R. A. Young (2013). "Transcriptional regulation and its misregulation in disease." Cell **152**(6): 1237-1251.
- Leppek, K., J. Schott, S. Reitter, F. Poetz, M. C. Hammond and G. Stoecklin (2013). "Roquin promotes constitutive mRNA decay via a conserved class of stem-loop recognition motifs." Cell **153**(4): 869-881.
- Lim, J., D. Kim, Y. S. Lee, M. Ha, M. Lee, J. Yeo, et al. (2018). "Mixed tailing by TENT4A and TENT4B shields mRNA from rapid deadenylation." Science **361**(6403): 701-704.
- Loh, B., S. Jonas and E. Izaurralde (2013). "The SMG5-SMG7 heterodimer directly recruits the CCR4-NOT deadenylase complex to mRNAs containing nonsense codons via interaction with POP2." Genes Dev **27**(19): 2125-2138.
- Lone, B. A., S. K. L. Karna, F. Ahmad, N. Shahi and Y. R. Pokharel (2018). "CRISPR/Cas9 System: A Bacterial Tailor for Genomic Engineering." Genet Res Int **2018**: 3797214.
- Luo, Y., Z. Na and S. A. Slavoff (2018). "P-Bodies: Composition, Properties, and Functions." Biochemistry **57**(17): 2424-2431.
- Lykke-Andersen, J. and E. J. Bennett (2014). "Protecting the proteome: Eukaryotic cotranslational quality control pathways." J Cell Biol **204**(4): 467-476.
- Mathys, H., J. Basquin, S. Ozgur, M. Czarnocki-Cieciura, F. Bonneau, A. Aartse, et al. (2014). "Structural and biochemical insights to the role of the CCR4-NOT complex and DDX6 ATPase in microRNA repression." Mol Cell **54**(5): 751-765.
- Mauxion, F., B. Preve and B. Seraphin (2013). "C2ORF29/CNOT11 and CNOT10 form a new module of the CCR4-NOT complex." RNA Biol **10**(2): 267-276.
- Maystadt, I., M. Zarhrate, P. Landrieu, O. Boespflug-Tanguy, S. Sukno, P. Collignon, et al. (2004). "Allelic heterogeneity of SMARD1 at the IGHMBP2 locus." Hum Mutat **23**(5): 525-526.
- McCarthy, D. J., Y. Chen and G. K. Smyth (2012). "Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation." Nucleic Acids Res **40**(10): 4288-4297.

- Meister, G., M. Landthaler, L. Peters, P. Y. Chen, H. Urlaub, R. Luhrmann and T. Tuschl (2005). "Identification of novel argonaute-associated proteins." Curr Biol **15**(23): 2149-2155.
- Mendell, J. T., C. M. ap Rhys and H. C. Dietz (2002). "Separable roles for rent1/hUpf1 in altered splicing and decay of nonsense transcripts." Science **298**(5592): 419-422.
- Miller, J. E. and J. C. Reese (2012). "Ccr4-Not complex: the control freak of eukaryotic cells." Crit Rev Biochem Mol Biol **47**(4): 315-333.
- Montague, T. G., J. M. Cruz, J. A. Gagnon, G. M. Church and E. Valen (2014). "CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing." Nucleic Acids Res **42**(Web Server issue): W401-407.
- Nagai, H., A. Yabe, N. Mine, I. Mikami, H. Fujiwara, Y. Terada, et al. (2003). "Down-regulation in human cancers of DRHC, a novel helicase-like gene from 17q25.1 that inhibits cell growth." Cancer Lett **193**(1): 41-47.
- Nagarajan, V. K., C. I. Jones, S. F. Newbury and P. J. Green (2013). "XRN 5'-->3' exoribonucleases: structure, mechanisms and functions." Biochim Biophys Acta **1829**(6-7): 590-603.
- Neve, J., R. Patel, Z. Wang, A. Louey and A. M. Furger (2017). "Cleavage and polyadenylation: Ending the message expands gene regulation." RNA Biol **14**(7): 865-890.
- Nicholson, A. L. and A. E. Pasquinelli (2018). "Tales of Detailed Poly(A) Tails." Trends Cell Biol.
- Norbury, C. J. (2013). "Cytoplasmic RNA: a case of the tail wagging the dog." Nat Rev Mol Cell Biol **14**(10): 643-653.
- Oberer, M., A. Marintchev and G. Wagner (2005). "Structural basis for the enhancement of eIF4A helicase activity by eIF4G." Genes Dev **19**(18): 2212-2223.
- Pause, A. and N. Sonenberg (1992). "Mutational analysis of a DEAD box RNA helicase: the mammalian translation initiation factor eIF-4A." EMBO J **11**(7): 2643-2654.
- Pena, C., E. Hurt and V. G. Panse (2017). "Eukaryotic ribosome assembly, transport and quality control." Nat Struct Mol Biol **24**(9): 689-699.
- Petit, A. P., L. Wohlbold, P. Bawankar, E. Huntzinger, S. Schmidt, E. Izaurralde and O. Weichenrieder (2012). "The structural basis for the interaction between the

CAF1 nuclease and the NOT1 scaffold of the human CCR4-NOT deadenylase complex." Nucleic Acids Res **40**(21): 11058-11072.

- Pfaffl, M. W. (2001). "A new mathematical model for relative quantification in real-time RT-PCR." Nucleic Acids Res **29**(9): e45.
- Pillai, R. S., C. G. Artus and W. Filipowicz (2004). "Tethering of human Ago proteins to mRNA mimics the miRNA-mediated repression of protein synthesis." RNA **10**(10): 1518-1525.
- Protter, D. S. W. and R. Parker (2016). "Principles and Properties of Stress Granules." Trends Cell Biol **26**(9): 668-679.
- Raisch, T., D. Bhandari, K. Sabath, S. Helms, E. Valkov, O. Weichenrieder and E. Izaurralde (2016). "Distinct modes of recruitment of the CCR4-NOT complex by Drosophila and vertebrate Nanos." EMBO J **35**(9): 974-990.
- Rambout, X., C. Detiffe, J. Bruyr, E. Mariavelle, M. Cherkaoui, S. Brohee, et al. (2016). "The transcription factor ERG recruits CCR4-NOT to control mRNA decay and mitotic progression." Nat Struct Mol Biol **23**(7): 663-672.
- Ran, F. A., P. D. Hsu, J. Wright, V. Agarwala, D. A. Scott and F. Zhang (2013). "Genome engineering using the CRISPR-Cas9 system." Nat Protoc **8**(11): 2281-2308.
- Razin, S. V., V. V. Borunova, O. G. Maksimenko and O. L. Kantidze (2012). "Cys2His2 zinc finger protein family: classification, functions, and major members." Biochemistry (Mosc) **77**(3): 217-226.
- Richard, P., S. Feng and J. L. Manley (2013). "A SUMO-dependent interaction between Senataxin and the exosome, disrupted in the neurodegenerative disease AOA2, targets the exosome to sites of transcription-induced DNA damage." Genes Dev **27**(20): 2227-2232.
- Robinson, M. D., D. J. McCarthy and G. K. Smyth (2010). "edgeR: a Bioconductor package for differential expression analysis of digital gene expression data." Bioinformatics **26**(1): 139-140.
- Ronchi, D., A. Di Fonzo, W. Lin, A. Bordoni, C. Liu, E. Fassone, et al. (2013). "Mutations in DNA2 link progressive myopathy to mitochondrial DNA instability." Am J Hum Genet **92**(2): 293-300.
- Roux, P. P. and I. Topisirovic (2012). "Regulation of mRNA translation by signaling pathways." Cold Spring Harb Perspect Biol **4**(11).

- Schutz, S., E. R. Noldeke and R. Sprangers (2017). "A synergistic network of interactions promotes the formation of in vitro processing bodies and protects mRNA against decapping." Nucleic Acids Res **45**(11): 6911-6922.
- Schweingruber, C., S. C. Rufener, D. Zund, A. Yamashita and O. Muhlemann (2013). "Nonsense-mediated mRNA decay - mechanisms of substrate mRNA recognition and degradation in mammalian cells." Biochim Biophys Acta **1829**(6-7): 612-623.
- Sgromo, A., T. Raisch, P. Bawankar, D. Bhandari, Y. Chen, D. Kuzuoglu-Ozturk, et al. (2017). "A CAF40-binding motif facilitates recruitment of the CCR4-NOT complex to mRNAs targeted by Drosophila Roquin." Nat Commun **8**: 14307.
- Shirai, Y. T., T. Suzuki, M. Morita, A. Takahashi and T. Yamamoto (2014). "Multifunctional roles of the mammalian CCR4-NOT complex in physiological phenomena." Front Genet **5**: 286.
- Simms, C. L., E. N. Thomas and H. S. Zaher (2017). "Ribosome-based quality control of mRNA and nascent peptides." Wiley Interdiscip Rev RNA **8**(1).
- Singh, P., U. Saha, S. Paira and B. Das (2018). "Nuclear mRNA Surveillance Mechanisms: Function and Links to Human Disease." J Mol Biol **430**(14): 1993-2013.
- Siomi, M. C., K. Sato, D. Pezic and A. A. Aravin (2011). "PIWI-interacting small RNAs: the vanguard of genome defence." Nat Rev Mol Cell Biol **12**(4): 246-258.
- Sloan, K. E. and M. T. Bohnsack (2018). "Unravelling the Mechanisms of RNA Helicase Regulation." Trends Biochem Sci **43**(4): 237-250.
- Smith, R. W., T. K. Blee and N. K. Gray (2014). "Poly(A)-binding proteins are required for diverse biological processes in metazoans." Biochem Soc Trans **42**(4): 1229-1237.
- Spellmon, N., J. Holcomb, L. Trescott, N. Sirinupong and Z. Yang (2015). "Structure and function of SET and MYND domain-containing proteins." Int J Mol Sci **16**(1): 1406-1428.
- Standart, N. and D. Weil (2018). "P-Bodies: Cytosolic Droplets for Coordinated mRNA Storage." Trends Genet **34**(8): 612-626.
- TD, M. P. and M. F. Wilkinson (2018). "RNA Decay Factor UPF1 Promotes Protein Decay: A Hidden Talent." Bioessays **40**(1).

- Tomecki, R., M. S. Kristiansen, S. Lykke-Andersen, A. Chlebowski, K. M. Larsen, R. J. Szczesny, et al. (2010). "The human core exosome interacts with differentially localized processive RNases: hDIS3 and hDIS3L." EMBO J **29**(14): 2342-2357.
- Tritschler, F., J. E. Braun, C. Motz, C. Igreja, G. Haas, V. Truffault, et al. (2009). "DCP1 forms asymmetric trimers to assemble into active mRNA decapping complexes in metazoa." Proc Natl Acad Sci U S A **106**(51): 21591-21596.
- Tycko, J., V. E. Myer and P. D. Hsu (2016). "Methods for Optimizing CRISPR-Cas9 Genome Editing Specificity." Mol Cell **63**(3): 355-370.
- Ukleja, M., J. M. Valpuesta, A. Dziembowski and J. Cuellar (2016). "Beyond the known functions of the CCR4-NOT complex in gene expression regulatory mechanisms: New structural insights to unravel CCR4-NOT mRNA processing machinery." Bioessays **38**(10): 1048-1058.
- Untergasser, A., I. Cutcutache, T. Koressaar, J. Ye, B. C. Faircloth, M. Remm and S. G. Rozen (2012). "Primer3--new capabilities and interfaces." Nucleic Acids Res **40**(15): e115.
- Van Etten, J., T. L. Schagat, J. Hrit, C. A. Weidmann, J. Brumbaugh, J. J. Coon and A. C. Goldstrohm (2012). "Human Pumilio proteins recruit multiple deadenylases to efficiently repress messenger RNAs." J Biol Chem **287**(43): 36370-36383.
- Villalba, A., O. Coll and F. Gebauer (2011). "Cytoplasmic polyadenylation and translational control." Curr Opin Genet Dev **21**(4): 452-457.
- Vourekas, A., K. Zheng, Q. Fu, M. Maragkakis, P. Alexiou, J. Ma, et al. (2015). "The RNA helicase MOV10L1 binds piRNA precursors to initiate piRNA processing." Genes Dev **29**(6): 617-629.
- Wagner, D. S., L. Gan and W. H. Klein (1999). "Identification of a differentially expressed RNA helicase by gene trapping." Biochem Biophys Res Commun **262**(3): 677-684.
- Wang, E. T., J. M. Taliaferro, J. A. Lee, I. P. Sudhakaran, W. Rossoll, C. Gross, et al. (2016). "Dysregulation of mRNA Localization and Translation in Genetic Disease." J Neurosci **36**(45): 11418-11426.
- Wang, Z., X. Jiao, A. Carr-Schmid and M. Kiledjian (2002). "The hDcp2 protein is a mammalian mRNA decapping enzyme." Proc Natl Acad Sci U S A **99**(20): 12663-12668.

- Warkocki, Z., P. S. Krawczyk, D. Adamska, K. Bijata, J. L. Garcia-Perez and A. Dziembowski (2018). "Uridylation by TUT4/7 Restricts Retrotransposition of Human LINE-1s." Cell **174**(6): 1537-1548 e1529.
- Webster, M. W., Y. H. Chen, J. A. W. Stowell, N. Alhusaini, T. Sweet, B. R. Graveley, et al. (2018). "mRNA Deadenylation Is Coupled to Translation Rates by the Differential Activities of Ccr4-Not Nucleases." Mol Cell **70**(6): 1089-1100 e1088.
- Wickramasinghe, V. O. and R. A. Laskey (2015). "Control of mammalian gene expression by selective mRNA export." Nat Rev Mol Cell Biol **16**(7): 431-442.
- Will, C. L. and R. Luhrmann (2011). "Spliceosome structure and function." Cold Spring Harb Perspect Biol **3**(7).
- Wilusz, J. E., C. K. JnBaptiste, L. Y. Lu, C. D. Kuhn, L. Joshua-Tor and P. A. Sharp (2012). "A triple helix stabilizes the 3' ends of long noncoding RNAs that lack poly(A) tails." Genes Dev **26**(21): 2392-2407.
- Wolf, J. and L. A. Passmore (2014). "mRNA deadenylation by Pan2-Pan3." Biochem Soc Trans **42**(1): 184-187.
- Wu, Y. (2012). "Unwinding and rewinding: double faces of helicase?" J Nucleic Acids **2012**: 140601.
- Yan, Y. B. (2014). "Deadenylation: enzymes, regulation, and functional implications." Wiley Interdiscip Rev RNA **5**(3): 421-443.
- Ye, J., A. R. Osborne, M. Groll and T. A. Rapoport (2004). "RecA-like motor ATPases--lessons from structures." Biochim Biophys Acta **1659**(1): 1-18.
- Yi, H., J. Park, M. Ha, J. Lim, H. Chang and V. N. Kim (2018). "PABP Cooperates with the CCR4-NOT Complex to Promote mRNA Deadenylation and Block Precocious Decay." Mol Cell **70**(6): 1081-1088 e1085.
- Youn, J. Y., W. H. Dunham, S. J. Hong, J. D. R. Knight, M. Bashkurov, G. I. Chen, et al. (2018). "High-Density Proximity Mapping Reveals the Subcellular Organization of mRNA-Associated Granules and Bodies." Mol Cell **69**(3): 517-532 e511.
- Zimmermann, L., A. Stephens, S. Z. Nam, D. Rau, J. Kubler, M. Lozajic, et al. (2018). "A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core." J Mol Biol **430**(15): 2237-2243.
- Zinder, J. C. and C. D. Lima (2017). "Targeting RNA for processing or destruction by the eukaryotic RNA exosome and its cofactors." Genes Dev **31**(2): 88-100.