Translational control by eukaryotic initiation factor (eIF) 4E and 4G homologous proteins

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Ramona Weber

aus Kirchheim unter Teck

Tübingen 2020

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

Tag der mündlichen Qualifikation:	10.11.2020
Stellvertretender Dekan:	Prof. Dr. József Fortágh
1. Berichterstatter:	Prof. Dr. Ralf-Peter Jansen
2. Berichterstatter:	Prof. Dr. Patrick Müller
3. Berichterstatter:	Prof. Dr. Elmar Wahle

Content

1. Abbreviations

Abbreviation	Meaning
40S	Eukaryotic small ribosomal subunit
4E-BM	eIF4E-binding motif
4E-BP	eIF4E-binding protein
4E-T	eIF4E-transporter
4EHP	eIF4E-homologous protein
4EHP-BM	4EHP-binding motif
60S	Eukaryotic large ribosomal subunit
80S	Eukaryotic ribosome
ABCE1	ATP-binding cassette sub-family E member 1
ADP	Adenosine diphosphate
AGO	Argonaute protein
Ala	Alanine
AR	Androgen receptor
ARE	AU-rich element
Arg	Arginine
ATP	Adenosine triphosphate
Bcd	Bicoid
Brat	Brain tumor
CBC	Cap-binding complex
CCR4	Carbon catabolite repressor 4
CDS	Coding sequence
CHX	Cycloheximide
CPSF	Cleavage and polyadenylation specificity factor
CStF	Cleavage stimulation factor
DAP5	Death associated protein 5
DCP2	Decapping enzyme subunit 2
DDX6	DEAD box protein 6
DDX6-BM	DDX6-binding motif
DEAD-box	Asp-Glu-Ala-Asp sequence motif in helicase domains
DENR	Density-regulated protein

Dm	*Drosophila melanogaster*
DNA	Desoxyribonucleic acid
EBFP	Enhanced blue fluorescent protein
eEF1A	Eukaryotic translation elongation factor 1A
eEF1B	Eukaryotic translation elongation factor 1B
eEF2	Eukaryotic translation elongation factor 2
eEF3	Eukaryotic translation elongation factor 3
eIF1	Eukaryotic translation initiation factor 1
eIF1A	Eukaryotic translation initiation factor 1A
eIF2	Eukaryotic translation initiation factor 2
eIF2A	Eukaryotic translation initiation factor 2A
eIF2B	Eukaryotic translation initiation factor 2B
eIF2D	Eukaryotic translation initiation factor 2D
eIF3	Eukaryotic translation initiation factor 3
eIF4A	Eukaryotic translation initiation factor 4A
eIF4B	Eukaryotic translation initiation factor 4B
eIF4E	Eukaryotic translation initiation factor 4E
eIF4F	Eukaryotic translation initiation factor 4F
eIF4G	Eukaryotic translation initiation factor 4G
eIF4H	Eukaryotic translation initiation factor 4H
eIF5	Eukaryotic translation initiation factor 5
eIF5A	Eukaryotic translation initiation factor 5A
eIF5B	Eukaryotic translation initiation factor 5B
EJC	Exon junction complex
ERAD	Endoplasmic reticulum-associated degradation
eRF1	Eukaryotic translation release factor 1
eRF3	Eukaryotic translation release factor 3
F-LUC	Firefly luciferase
FACS	Fluorescence activated cell sorting
GDP	Guanosine diphosphate
GEF	Guanine nucleotide exchange factor
GFP	Green fluorescent protein

GIGYF1/2	Grb10-interacting GYF domain proteins 1/2
Gly	Glycine
GTP	Guanosine triphosphate
GW182	182 kDa protein containing Gly-Trp repeats
GYF domain	Glycine-tyrosine-phenylalanine domain
HA	Hemagglutinin
HEK293T	Human embryonic kidney 293T cells
Hs	*Homo sapiens*
IP	Immunoprecipitation
Ire1	Inositol-requiring 1
IRES	Internal ribosome entry site
JNK	JUN N-terminal kinase
Leu	Leucine
LTM	Lactimidomycin
Lys	Lysine
m^7GTP	7-methylguanosine-5´-triphosphate
mCherry	Monomeric Cherry (family of red fluorescent proteins)
MCTS-1	Malignant T-cell-amplified sequence 1
Met-tRNA$_i^{met}$	Initiator methionine transfer RNA
MIF4G	Middle domain of eukaryotic initiation factor 4G
miRISC	Micro RNA-induced silencing complex
miRNA	Micro RNA
mNG2	Monomeric NeonGreen2
MNK	MAP kinase-interacting serine/threonine-protein kinase
mORF	Main open reading frame
mRNA	Messenger RNA
mRNP	Messenger ribonucleoprotein particle
mTOR	Mammalian/mechanistic target of rapamycin
NAT1	Novel APOBEC1 target 1
ncRNA	Non-coding RNA
NGD	No-go decay
NMD	Nonsense mediated decay

NOT1	Negative on TATA-less protein 1
NSD	Non-stop decay
ORF	Open reading frame
PABP	Poly(A)-binding protein
PAN2/PAN3	Poly(A) nucleases 2/3
PCBP2	Poly(C)-binding protein 2
PERK	Protein kinase RNA-like ER kinase
PI3K	Phosphatidylinositol-3-kinase
PIC	Preinitiation complex
Poly(A)	Polyadenosine
pre-mRNA	Precursor-messenger RNA
Pro	Proline
PTC	Premature termination codon
QTI	Quantitative translation initiation
R-LUC	*Renilla* luciferase
RBP	RNA-binding protein
RFP	Ribosome footprint
RIDD	Regulated IRE1-dependent decay
RNA	Ribonucleic acid
RNA Pol II	RNA polymerase II
RQC	Ribosome-associated quality control
rRNA	Ribosomal RNA
RT-qPCR	Real-time quantitative polymerase chain reaction
SBMA	Spinal bulbar muscular atrophy
shRNA	Short hairpin RNA
SP	Signal peptide
SRP	Signal recognition particle
TNRC6	Trinucleotide repeat-containing gene 6
tRNA	Transfer RNA
Trp	Tryptophane
TRS	Threonyl-tRNA synthetase
TTC5	Tetratricopeptide protein 5

TTP	Tristetraprolin
Tyr	Tyrosine
uORF	Upstream open reading frame
UPR	Unfolded protein response
UTR	Untranslated region
WT	Wild-type
XBP1u/XBP1s	X box-binding protein 1 unspliced/spliced
XRN1	5´-to-3´ exoribonuclease 1
ZNF598	Zinc finger protein 598

2. Summary

All living systems rely on appropriate protein synthesis and therefore evolved intricate mechanisms to control their proteome. Especially, the complexity of metazoan multicellular processes requires a strict and multilayered regulation of gene expression at the level of transcription, mRNA processing, translation, mRNA stability, and protein modifications and stability. Here, I present work on two research projects that uncovered the hitherto unidentified roles of translation initiation (1) and mRNA decay factors (2) specialized in controlling unconventional mRNA translation and co-translational mRNA stability, respectively.

The initiation step of translation commences with the assembly of the eukaryotic initiation factor 4F (eIF4F) complex on the mRNA 5′ cap structure which facilitates the recruitment of the ribosome to mediate protein synthesis. The eIF4F consists of the cap-binding protein eIF4E, the RNA helicase eIF4A and the scaffolding factor eIF4G. Apart from eIF4G, metazoan cells express other eIF4G-like proteins: eIF4G2 – a.k.a. DAP5/p97/NAT1 – and eIF4G3. Although the different homologs perform non-redundant functions in translation, we lack detailed understanding of their individual contribution to the initiation of protein synthesis. To gain insights into the function of DAP5 in translation initiation, I applied transcriptome (RNA-sequencing) and translatome (ribosome profiling) approaches to CRISPR-Cas9 engineered human HEK293T DAP5-null cells. These approaches allowed the identification of a group of mRNAs with complex 5′ untranslated regions (UTRs) that required the DAP5-eIF4A complex for efficient translation. In detail, the 5′ UTRs of DAP5 targets harbor strong secondary structures and at least one upstream open reading frame (uORF) that normally sequesters ribosomes from initiation at the main ORF (mORF). In this context, DAP5 is critical for mORF translation as it mediates re-initiation after uORF translation. Taken together, my studies on DAP5 demonstrate how cells control translation initiation using 5′ UTR-associated mechanisms.

Regulation of translation initiation can also involve additional cap-binding proteins. The eIF4E-homologous protein (4EHP) competes with eIF4E to bind the mRNA 5′ cap structure. Unlike eIF4E, 4EHP does not promote translation since it is unable to recruit eIF4G and the ribosome. Instead, 4EHP represses translation and induces mRNA decay with the help of 4EHP-binding factors. The GIGYF1 and GIGYF2 proteins specifically associate with 4EHP and act as scaffolding proteins for GYF domain-associated proteins like tristetraprolin (TTP) or ZNF598, the RNA helicase and decapping activator DDX6, and NOT1 – a subunit of the CCR4-NOT deadenylase complex. When recruited to an mRNA, the 4EHP-GIGYF1/2

complexes thus elicit translation repression and mRNA decay. Although the molecular mechanisms have been studied in great detail, we only have sparse information on the pool of mRNAs subjected to 4EHP-GIGYF1/2 mediated repression and decay. Analysis of RNA-sequencing and ribosome profiling datasets of 4EHP- and GIGYF1/2-null cells revealed that the 4EHP-GIGYF1/2 complexes co-translationally target mRNAs with perturbed ribosome elongation or specific nascent peptide chains for decay. These include, among others, mRNAs encoding endoplasmic reticulum (ER)-targeted proteins, α- and β-tubulin subunits and mRNAs with CAG codon repeats encoding poly-glutamine (Q) stretches. In detail, the 4EHP-GIGYF1/2 complexes are recruited to target mRNAs via the GYF domain-associated proteins and upon binding to the mRNA 5′ cap, and DDX6, degrade the bound mRNA. My studies show for the first time how a repressive complex, associated with neurological disorders in animal models and affected humans, specifically minimizes the protein output of a subset of mRNAs. Altogether, I identified a previously unappreciated role of the 4EHP-GIGYF1/2 complexes in maintaining protein homeostasis and preventing excessive production of potentially toxic proteins.

In summary, my studies increase our understanding of how metazoan cells utilize the translation control toolbox to fine-tune their proteome and react to developmental cues, adapt to environmental changes and maintain homeostasis. The identified translational programs are fundamental for establishing a multicellular organization but at the same time allow the organism to evade cancer and neurological diseases.

3. Zusammenfassung

Alle Lebewesen beruhen auf einer für sie angemessenen Proteinsynthese und haben daher aufwendige Mechanismen entwickelt, die es ihnen erlauben ihr Proteom zu kontrollieren. Insbesondere die Komplexität multizellulärer Prozesse in Metazoa basiert auf einer strikten und vielschichtigen Regulation der Genexpression auf den Ebenen der Transkription, mRNA Bearbeitung, Translation, mRNA Stabilität und Proteinmodifikation und -stabilität. In dieser Arbeit stelle ich zwei Forschungsprojekte vor, die bislang unbekannte Funktionen von Faktoren in der Translationsinitiation (1), spezialisiert in der Kontrolle der unkonventionellen mRNA Translation, und zum Abbau von mRNA (2), spezialisiert in co-translationaler mRNA Stabilität, aufgezeigt haben.

Die Translationsinitiation beginnt mit dem Zusammenbau des eukaryotischen Initiationsfaktor 4F (eIF4F) Komplexes auf der mRNA 5´ Cap Struktur, welcher die Rekrutierung des Ribosoms ermöglicht, um Proteine zu synthetisieren. eIF4F besteht aus dem Cap-bindenden Protein eIF4E, der RNA Helikase eIF4A und dem Gerüstfaktor eIF4G. Metazoa-Zellen exprimieren außer eIF4G noch andere eIF4G-ähnliche Proteine: eIF4G2 – auch bekannt als DAP5/p97/NAT1 – und eIF4G3. Obwohl die verschiedenen Homologe bekanntermaßen nicht-redundante Funktionen in der Translation erfüllen, fehlt uns ein detailliertes Verständnis deren individueller Beiträge zur Initiation der Proteinsynthese. Um Einblicke in die Funktion von DAP5 bei der Translationsinitiation zu erhalten, habe ich Transkriptom- (RNA-Sequenzierung) und Translatomansätze (Ribosom Profiling) mittels CRISPR-Cas9 veränderten menschlichen HEK293T DAP5-null Zellen angewandt. Diese Vorgehensweise erlaubte mir eine Gruppe an mRNAs mit komplexen 5´ untranslatierten Regionen (UTRs) zu identifizieren, welche den Komplex aus DAP5-eIF4A für eine effiziente Translation benötigen. Im Detail beherbergen die 5´ UTRs der DAP5 regulierten mRNAs stark ausgeprägte Sekundärstrukturen und mindestens einen vorgelagerten offenen Leserahmen (uORF), der normalerweise Ribosomen abfängt und an der Initiation am Haupt- offenen Leserahmen (mORF) hindert. In diesem Zusammenhang ist DAP5 kritisch für die Translation des mORFs, da es Re-Initiation nach der Translation des uORFs herbeiführt. Zusammengefasst zeigen meine Beobachtungen, wie Zellen 5´ UTR-assoziierte Mechanismen verwenden, um die Translationsinitiation zu kontrollieren.

Die Regulation der Translationsinitiation kann zudem zusätzliche Cap-bindende Proteine einbeziehen. Das eIF4E-homologe Protein (4EHP) konkurriert mit eIF4E um die Bindung der mRNA 5´ Cap Struktur. Im Gegensatz zu eIF4E fördert 4EHP nicht die Translation, da es nicht

im Stande ist eIF4G und das Ribosom zu rekrutieren. Statt dessen hemmt 4EHP die Translation und verursacht, mithilfe der 4EHP-bindenden Faktoren, den mRNA Abbau. Die GIGYF1 und GIGYF2 Proteine binden spezifisch 4EHP und agieren als Gerüstfaktoren für GYF-Domänen-assoziierende Proteine wie Tristetraprolin (TTP) oder ZNF598, die RNA Helikase und Decapping Aktivator DDX6, sowie NOT1 – eine Untereinheit des CCR4-NOT Deadenylase Komplexes. Aus diesem Grund verursachen die 4EHP-GIGYF1/2 Komplexe Translationsrepression und mRNA Abbau, wenn sie zu einer mRNA gebracht werden. Obwohl die molekularen Mechanismen in gutem Detail erforscht sind, haben wir dennoch nur spärliche Informationen darüber welche mRNAs durch 4EHP-GIGYF1/2 reguliert werden. RNA-Sequenzierung und Ribosom Profiling von 4EHP- und GIGYF1/2-null Zellen zeigten auf, dass die 4EHP-GIGYF1/2 Komplexe mRNAs mit gestörter Elongation des Ribosoms oder mit spezifischen freiwerdenden Peptidketten zum Abbau bestimmen. Diese beinhalten unter anderem mRNAs, die für Proteine kodieren, die für das endoplasmatische Retikulum (ER) bestimmt sind, sowie α- und β-tubulin Untereinheiten, als auch mRNAs mit CAG Codon Wiederholungen, die für Poly-Glutamin (Q) Ketten kodieren. Die 4EHP-GIGYF1/2 Komplexe werden über die GYF-Domänen-assoziierenden Proteine an Ziel-mRNAs gebracht und bauen die gebundene mRNA ab, nachdem sie an die mRNA 5′ Cap und DDX6 gebunden haben. Meine Untersuchungen zeigen zum ersten Mal wie ein hemmender Komplex, der mit neurologischen Fehlfunktionen in Tiermodellen und betroffenen Patienten in Verbindung gebracht wurde, spezifisch die Proteinproduktion von einer kleinen Gruppe an mRNAs minimiert. Alles in allem habe ich eine bisher unbeachtete Rolle des 4EHP-GIGYF1/2 Komplexes in der Aufrechterhaltung der Proteinhomöostase und der Verhinderung exzessiver Produktion von potenziell toxischen Proteinen identifiziert.

Zusammenfassend erweitern meine Untersuchungen unser Verständnis wie Metazoa-Zellen die Translations- „Werkzeugkiste" nutzen, um ihr Proteom präzise abzustimmen und auf Signale während der Entwicklung zu reagieren, sich auf Umweltänderungen einzustellen und Homöostase beizubehalten. Die identifizierten Translationsprogramme sind wesentlich um eine multizelluläre Organisation zu ermöglichen und erlauben zugleich dem Organismus Krebserkrankungen und neurologische Leiden zu umgehen.

4. List of publications and author contributions

4.1. Discussed publications and manuscripts

The original manuscripts of the published and unpublished manuscripts listed in this section are attached.

A:

Weber R, Kleemann L, Hirschberg I, Chung MY, Valkov E, Izaurralde E and Igreja C. "DAP5 mediates re-initiation after translation of uORFs on structured 5′ leader sequences." (In preparation).

R.W. designed and conducted the experiments. L.K. performed luciferase assays and generated several of the constructs used in this study. I.H. performed and analyzed FACS measurements together with **R.W.** M.-Y.C. performed RNA-immunoprecipitation assays, luciferase assays and generated several of the constructs used in this study. E.V. contributed to data analysis. E.I. was the principal investigator. **R.W.** and C.I. conceived the project, interpreted the results and wrote the manuscript. All authors, with exception of E.I., corrected the manuscript.

B:

Weber R, Chung MY, Keskeny C, Zinnall U, Landthaler M, Valkov E, Izaurralde E and Igreja C. "4EHP and GIGYF1/2 mediate translation-coupled messenger RNA decay." (In revision).

R.W. designed and conducted the experiments. M.-Y.C. performed RNA-immunoprecipitation and pulldown assays. C.K. generated the ZNF598 KO cell line and helped with ribosome profiling and RNA-sequencing experiments. U.Z. and M.L. helped with ribosome profiling and RNA-sequencing experiments and contributed to data analysis. E.V. contributed to data analysis. E.I. was the principal investigator. **R.W.** and C.I. conceived the project, interpreted the results and wrote the manuscript. All authors, with exception of E.I., corrected the manuscript.

C:

Peter D, Ruscica V, Bawankar P, **Weber R**, Helms S, Valkov E, Igreja C, and Izaurralde E. (2019) "Molecular basis for GIGYF-Me31B complex assembly in 4EHP-mediated translation repression." **Genes and Development**. 33: 1355-1360.

D.P. performed all experiments to determine the structure of the complex. V.R. contributed to complex purification. E.V. contributed to structural data analysis. **R.W.** performed the

complementation assays. P.B, S.H. and V.R. performed pulldown and immunoprecipitation assays. P.B. generated several of the constructs used in this study. C.I coordinated the project. E.I. was the principal investigator. D.P., E.V. and C.I. wrote the manuscript. All authors corrected the manuscript.

<u>D:</u>

Peter D, **Weber R**, Sandmeir F, Wohlbold L, Helms S, Bawankar P, Valkov E, Igreja C, and Izaurralde E. (2017). "GIGYF1/2 proteins use auxiliary sequences to selectively bind to 4EHP and repress target mRNA expression." **Genes and Development**. <u>31</u>: 1147–1161.

D.P. performed all experiments to solve the structures of the complexes together with F.S. ITC measurements were performed by D.P. and F.S. S.H. performed pulldown and competition assays. **R.W.** generated the GIGYF1/2-null cell line, performed tethering, complementation and pulldown assays; purified peptides/proteins and generated many of the constructs used in the pulldown assays. L.W. generated several constructs for expression in human cells and performed immunoprecipitation assays. D.P., F.S and E.V. performed the SAXS experiments and structural analysis. P.B. performed immunoprecipitation assays and generated the ARE reporter and the TTP expression vector. C.I. coordinated the project. E.I. was the principal investigator and supervised the project. D.P., C.I., E.V. and E.I. wrote the manuscript. All authors corrected the manuscript.

4.2. Additional publications (during PhD)

Räsch F, **Weber R**, Izaurralde E and Igreja C. (2020) "4E-T bound mRNAs are stored in a silenced and deadenylated form." **Genes and Development**. <u>34</u>: 847-860.

Hanet A, Räsch F, **Weber R**, Ruscica V, Fauser M, Raisch T, Kuzuoğlu-Öztürk D, Chang CT, Bhandari D, Igreja C, and Wohlbold L. (2019) "HELZ directly interacts with CCR4-NOT and causes decay of bound mRNAs." **Life Science Alliance**. <u>2</u>: No.5.

Chang CT, Muthukumar S, **Weber R**, Levdansky Y, Chen Y, Bhandari D, Wohlbold L, Valkov E, and Izaurralde E. (2019) "XRN1 coordinates mRNA decay through direct interactions with decapping factors and CCR4-NOT." **Nucleic Acids Research**. <u>47</u>: 9282–9295.

Grüner S, **Weber R**, Peter D, Chung MY, Igreja C, Valkov E, and Izaurralde E. (2018). "Structural motifs in eIF4G and 4E-BPs modulate their binding to eIF4E to regulate translation initiation in yeast." **Nucleic Acids Research**. 46: 6893–6908.

Grüner S, Peter D, **Weber R**, Wohlbold L, Chung MY, Weichenrieder O, Valkov E, Igreja C, and Izaurralde E. (2016). "The Structures of eIF4E-eIF4G Complexes Reveal an Extended Interface to Regulate Translation Initiation." **Molecular Cell**. 64: 467–479.

4.3. Additional publications (prior to PhD)

Peter D, **Weber R**, Köne C, Chung MY, Ebertsch L, Truffault V, Weichenrieder O, Igreja C, and Izaurralde E. (2015). "Mextli proteins use both canonical bipartite and novel tripartite binding modes to form eIF4E complexes that display differential sensitivity to 4E-BP regulation." **Genes and Development**. 29: 1835–1849.

Peter D, Igreja C, **Weber R**, Wohlbold L, Weiler C, Ebertsch L, Weichenrieder O, and Izaurralde E. (2015). "Molecular Architecture of 4E-BP Translational Inhibitors Bound to eIF4E." **Molecular Cell**. 57: 1074–1087.

5.   Conference presentations and workshops

September 1-4 2020	Cold Spring Harbor Conference: Translational Control, virtual meeting. "DAP5 promotes translation re-initiation downstream of uORFs" (**Selected talk**).
May 26-31 2020	RNA Conference: RNA2020, Vancouver, Canada, virtual meeting. "4EHP and GIGYF1/2 mediate translation-coupled messenger RNA decay" (**Selected talk**).
September 4–7 2019	EMBO Conference: Protein Synthesis and Translational Control, Heidelberg, Germany. "DAP5 mediates re-initiation after uORF translation on structured 5′ leader sequences" (**Selected talk**).
March 26–29 2019	EMBL-EBI course, Hinxton, Cambridge, UK: "Introduction to RNA-Seq and Functional Interpretation".
February 16–22 2019	Gordon Research Conference: Translation Machinery in Health and Disease, Galveston, TX, USA. "GIGYF1/2 proteins regulate the stability of mRNAs with GC$_3$-rich coding sequences" (**Poster**).
September 4–8 2018	Cold Spring Harbor Conference: Translational Control, Cold Spring Harbor, NY, USA. "GIGYF1/2 proteins regulate the stability of mRNAs with GC$_3$-rich coding sequences" (**Poster**).
August 22–26 2017	Cold Spring Harbor Conference: Eukaryotic mRNA Processing, Cold Spring Harbor, NY, USA. "GIGYF1/2 proteins selectively bind 4EHP and repress translation upon recruitment to target mRNAs" (**Selected talk**).
July 10–13 2017	EMBO Conference: Eukaryotic RNA turnover, Oxford, UK. "GIGYF1/2 proteins use auxiliary sequences to selectively bind 4EHP and repress target mRNA expression" (**Poster,** poster prize).

6. Introduction

Living systems rely on the flow of genetic information from nucleic acid [desoxyribonucleic acid (DNA)] to nucleic acid [DNA and ribonucleic acid (RNA)] and from nucleic acid (RNA) to protein. The genetic information encoded in the DNA is copied via DNA replication and passed into RNA molecules in a process known as transcription. Finally, specialized RNA described as messenger RNA (mRNA) is translated into protein.

6.1. Gene expression and mRNA biogenesis in eukaryotes

To organize the genetic information in the nucleus, eukaryotic cells pack DNA and proteins into macromolecular assemblies called chromatin. In these assemblies, DNA is wrapped around histones with the help of histone-associated proteins which control chromatin storage and accessibility. The flow of the genetic information from DNA into RNA, and ultimately into proteins, known has gene expression and summarized in Figure 1, depends on the accessibility of chromatin to enzymes that perform these activities. The dynamics of the chromatin state − important for gene regulation − is achieved by histone modifying and remodeling enzymes (Swygert and Peterson, 2014).

Transcription of DNA into mRNA is the first step in eukaryotic gene expression, catalyzed by RNA polymerase II (Pol II) which generates a precursor-mRNA (pre-mRNA). Pre-mRNAs undergo extensive co-transcriptional processing in the nucleus to form the mature mRNA. For the majority of eukaryotic mRNAs, the processing events include 5′ end capping, splicing and 3′ end processing by polyadenylation (Hantsche and Cramer, 2016). In the case of replication-dependent histone mRNAs, which are not polyadenylated, 3′ end processing involves mRNA cleavage and formation of a stem-loop structure (Marzluff et al., 2008).

The cap structure consists of an N^7-methylguanosine (m^{7}GTP) moiety on the 5′-terminal nucleotide of an mRNA (m^{7}GTPpppN) and is bound by the nuclear cap-binding complex (CBC). The CBC protects the mRNA from 5′ exonucleolytic degradation, promotes further processing and enhances nuclear export (Ramanathan et al., 2016; Topisirovic et al., 2011). Once the mRNA is exported to the cytoplasm, the CBC is replaced by the eukaryotic initiation factor 4E (eIF4E), the cytoplasmic cap-binding protein which is critical for translation (Sato and Maquat, 2009).

mRNAs undergo extensive chemical modifications not only at the 5′ cap but also internally. In the nucleus, mRNAs are co-transcriptionally modified by methyltransferases that catalyze the formation of N^6-methyladenosine (m^{6}A) or N^5-methylcytosine (m^{5}C) (Chen et al., 2019;

Liu et al., 2014; Ping et al., 2014; Squires et al., 2012; Yang et al., 2017). Pseudouridylation, the most abundant post-transcriptional RNA modification, is catalyzed by the pseudouridine (Ψ) synthase PUS1 (Li et al., 2015). These chemical modifications can influence mRNA stability (Wang et al., 2014), alter mRNA structure (Spitale et al., 2015), affect mRNA translation (Meyer et al., 2015; Zhou et al., 2015), enhance pre-mRNA splicing (Haussmann et al., 2016; Lence et al., 2016), promote mRNA transport (Zheng et al., 2013), cause alternative polyadenylation (Llacer et al., 2015), and target mRNAs to RNA-protein droplets (Ries et al., 2019). However, a comprehensive understanding of the role of mRNA modifications in eukaryotic gene expression is still lacking.

Splicing is the process by which non-coding sequences – or introns – within a pre-mRNA are removed by the action of the spliceosome (Matera and Wang, 2014). It leaves coding sequences – or exons – intact and marks exon-exon boundaries by positioning of the exon junction complex (EJC) shortly upstream of the splicing sites (Woodward et al., 2017).

3′ end processing is induced by a cleavage and polyadenylation signal in the pre-mRNA and causes cleavage and subsequent addition of non-templated adenosine residues to the 3′ tail of the mRNA. These actions are performed by the cleavage and polyadenylation specificity factor (CPSF)/cleavage stimulation factor (CStF) and the poly(A) polymerase (PAP), respectively (Weill et al., 2012). The above-mentioned histone mRNAs are specifically cleaved after a conserved stem-loop structure but do not get polyadenylated downstream of the stem-loop (Marzluff et al., 2008).

After pre-mRNA processing, the mature mRNA associated with RNA-binding proteins (RBPs) is exported to the cytoplasm. Nuclear export is mediated by the nuclear export receptor NXF1-p15 which shuttles the messenger ribonucleoprotein particle [mRNP (mRNA+proteins)] complex through the nuclear pore complex (Stutz and Izaurralde, 2003).
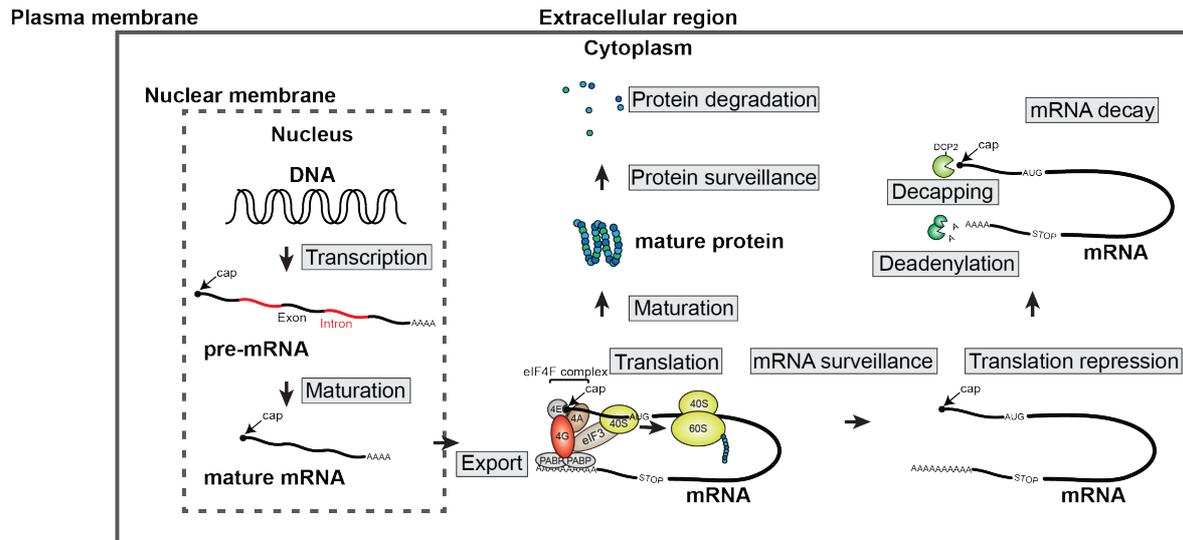
Figure 1: Summary of eukaryotic gene expression. DNA is transcribed into premature-messenger RNA (pre-mRNA) in the nucleus. mRNA capping and polyadenylation occur co-transcriptionally. Pre-mRNA maturation to mature mRNA happens in the nucleus before the mRNA is exported to the cytoplasm. Cytoplasmic mRNA is bound by eIF4F (composed of the cap binding protein eIF4E, the RNA helicase eIF4A and the scaffold factor eIF4G) and circularized. Translation is initiated by recruiting eIF3 and the ribosome. The newly synthetized protein undergoes maturation steps to generate a functional protein which is controlled by protein surveillance mechanisms which can eventually lead to protein degradation. Similarly, cellular surveillance mechanisms act on the mRNA level. These mechanisms induce translation repression of the mRNA which can be directly coupled to mRNA decay by deadenylation enzymes and the decapping factor DCP2.

6.2. Conventional translation of eukaryotic mRNAs

Protein synthesis is the process during which the genetic information on mRNA is decoded. The code, with few exceptions, is universal and resides in the open reading frame (ORF) of the mRNA. Base triplets (codons) of the ORF are designated to encode one amino acid in the peptide chain. Out of the 4^3 (64; 4 nucleotides combined in triplets) possible combinations, 61 codons encode 20 amino acids and three STOP codons terminate elongation of the amino acid chain (Fan et al., 2017). To decode the information of a codon, the cell uses transfer RNAs (tRNAs) which contain anti-codons that base-pair with the mRNA codons. In addition to the physical and specific base-pairing with the mRNA's codons, tRNAs carry specific amino acids loaded by tRNA synthetases (Pang et al., 2014). The decoding process is catalyzed by the ribosome composed of ribosomal RNA (rRNA) and proteins. In detail, the 40S small and the 60S large ribosomal subunits come together on an mRNA and form the core ribosome with A (aminoacyl), P (peptidyl), and E (exit) sites that allow charged tRNAs to enter and recognize the cognate codon, transfer their amino acid to the growing peptide chain, and exit the ribosome uncharged, respectively (Ramakrishnan, 2014).

6.2.1. Initiation

To initiate translation, eukaryotic cells have evolved an intricate chain of events that ensures efficient but regulated use of the cellular resources. In fact, protein synthesis is one of the most energy consuming processes in cells and therefore needs tight regulation (Sonenberg and Hinnebusch, 2009).

The initiation step of translation (summarized in Figure 2) starts off with the assembly of the 43S pre-initiation complex (PIC) consisting of the 40S ribosomal subunit in association with the initiation factors eIF1, eIF1A, the ternary complex (eIF2:GTP:Met-tRNA$_i$) and the multisubunit eIF3 complex. A key event during initiation is the recruitment of the 43S PIC to the cap-proximal 5´ untranslated region (UTR) of the mRNA by the eIF4F, a heterotrimeric complex composed of the cap-binding protein eIF4E, the ATP-dependent RNA helicase eIF4A and the scaffolding factor eIF4G. The latter holds together the eIF4F complex and mediates interactions to other factors involved in translation (Hinnebusch, 2014; Hinnebusch et al., 2016; Jackson et al., 2010).

The interaction of eIF4G on the 5´ end with PABP sitting at the 3´ poly(A) tail, is traditionally recognized to circularize the mRNA (Amrani et al., 2006; Chen and Shyu, 2011; Christensen et al., 1987; Gallie, 1991; Wells et al., 1998). The tight communication between the two ends of the mRNA regulates translation efficiency and mRNA decay. Other models, however, challenge the mRNA circularization model and propose that the close proximity of 5´ and 3´ ends is caused by the inherent property of RNAs to form structures and induce end-to-end proximity. Thus, the evolutionary conserved eIF4G-PABP interaction could be a consequence rather than the cause of mRNA circularization and might serve to reinforce or sense the proximity of 5´ and 3´ ends (Vicens et al., 2018).

eIF4G also binds to eIF3 which directly recruits the 43S PIC to the mRNA. The scanning competent 43S PIC moves in a 5´-to-3´ direction along the 5´ UTR, with the help of eIF4A and accessory factors eIF4B and eIF4H which resolve potentially obstructive mRNA secondary structures.

In eukaryotes, recognition of a start (AUG) codon on the mRNA is facilitated by the presence of the Kozak sequence in its vicinity. In vertebrates, the optimal Kozak sequence is 5´-GCC(A/G)CC**AUGG**-3´ (Kozak, 1987). Recognition of this sequence by the scanning complex, slows down the 43S PIC and induces conformational changes that allow joining of the 60S ribosomal subunit. The accuracy of start codon recognition is guaranteed by eIF1 and eIF1A. These initiation factors stabilize the open conformation of the 43S complex to permit

the passage of the mRNA through the complex (Cheung et al., 2007; Maag et al., 2005; Passmore et al., 2007). Mutations in or depletion of eIF1 lead to the initiation of translation at non-AUG start codons or AUG codons in a poor Kozak context (Hinnebusch, 2014; Pestova et al., 1998; Pestova and Kolupaeva, 2002). Likewise, overexpression of eIF1 restricts translation initiation to AUG codons in a good Kozak context (Ivanov et al., 2010). Dissociation of eIF1 after recognition of the start codon, promotes the hydrolysis of eIF2-bound GTP by eIF2γ. Its GTPase activity is stimulated by eIF5. These events culminate in the release of eIF2-GDP and eIF5 from the scanning complex and allow 60S subunit joining. Assembly of the ribosome unit is assisted by GTP binding and hydrolysis by eIF5B. The release of the GTPase and eIF1A from the 80S initiation complex (joined 40S and 60S subunits), is a prerequisite to form elongation-competent ribosomes. At this step, the eIF4 and eIF3 complexes are still bound to the ribosome and participate in the initial rounds of translation elongation (Kozak, 2001), with a decay half-length of about 12 codons (Bohlen et al., 2020).
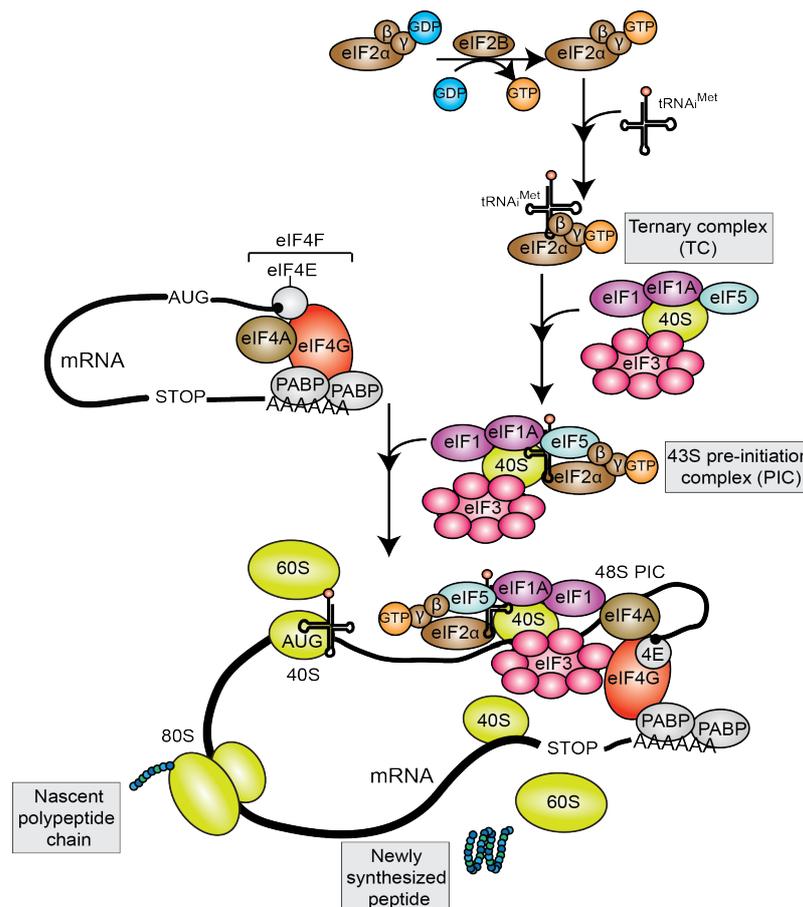
Figure 2: Summary of eukaryotic translation initiation. Eukaryotic translation initiation begins with the assembly of the 43S pre-initiation complex (PIC) which is composed of the 40S ribosomal subunit, the initiation factors eIF1, eIF1A, the ternary complex (TC) – consisting of eIF2:GTP:Met-tRNA$_i$ – and the multisubunit eIF3 complex. The 43S PIC is then recruited to the mRNA by the eIF4F, a heterotrimeric complex composed of the cap-binding protein eIF4E, the ATP-dependent

RNA helicase eIF4A and the scaffolding factor eIF4G. eIF4G bridges the interaction between the eIF4F complex and PABP and mediates the interaction to the eIF3 of the 43S PIC. This chain of events culminates in the recruitment of the 60S ribosomal subunit and assembly of the 80S ribosome upon start codon (AUG) recognition. Subsequently, translation elongation by the 80S ribosome produces the nascent polypeptide chain and the newly synthetized peptide is released during translation termination after stop codon recognition.

6.2.2. Elongation

At the start codon, the ribosome contains the initiator Met-tRNA$_i$ in the P site and empty A and E sites. A new aminoacylated tRNA (aa-tRNA) is delivered to the free A-site by the eukaryotic elongation factor 1A (eEF1A) bound to GTP (Dever and Green, 2012). Upon a stable codon-anticodon interaction between mRNA and tRNA, the 40S ribosomal subunit undergoes a conformational change. The resulting interactions between 18S rRNA and the mRNA-tRNA duplex trigger GTP hydrolysis by eEF1A and dissociation of eEF1A-GDP from the A site. Subsequently, methionine is released from the initiator tRNA by formation of the first peptide bond with the amino acid present on the A-site. The peptidyl transferase activity resides in the 60S ribosomal subunit and requires the concerted action of rRNAs and ribosomal proteins (Ben-Shem et al., 2010). The deacylated initiator tRNA leaves the P site, moves to the E site and is eventually released from the ribosome with the help of eEF2-GTP. eEF2-stimulated hydrolysis of GTP induces the translocation of the peptidyl tRNA from the A to the P site and the net forward movement of the ribosome by one codon on the mRNA (Frank et al., 2007; Taylor et al., 2007). The elongation cycle then repeats until a stop codon is encountered and translation is terminated. The growing (or nascent) peptide emerges from a specialized exit tunnel composed of negatively charged residues of the elongating ribosome (Balchin et al., 2016).

6.2.3. Termination and 60S recycling

Termination of translation is a consequence of the recognition of one of the three stop codons (UAG, UAA, UGA) in the ribosomal A site. This step is catalyzed by the class 1 release factor eRF1, which is recruited to the ribosome by the class 2 release factor eRF3, a GTPase (Kisselev et al., 2003). eRF1 decodes the stop codon by establishing interactions with all three bases of the codons, mimicking a tRNA (Bertram et al., 2000; Blanchet et al., 2015; Bulygin et al., 2011; Kryuchkova et al., 2013). After GTP hydrolysis, eRF3 leaves the termination complex and the ATP-binding cassette sub-family E member 1 (ABCE1) binds to the eRF1-loaded ribosome. With its catalytically essential glycine-glycine-glutamine (GGQ) motif, eRF1

channels the water molecule necessary for the hydrolysis of the last tRNA bound to the polypeptide chain and promotes, stimulated by ABCE1, peptide release (Frolova et al., 1999; Heurgue-Hamard et al., 2005; Preis et al., 2014; Seit-Nebi et al., 2001). In the next step, ABCE1 promotes 60S dissociation thereby coupling translation termination with ribosome recycling (Pisarev et al., 2010; Shao et al., 2016; Shoemaker and Green, 2011; Young et al., 2015).

Termination efficiency is dependent on the identity and context of the stop codon, as measured by the frequency of stop codon readthrough. Of the three stop codons, UGA is the weakest terminator. Furthermore, the +4 position influences the efficiency of stop codon recognition, with UGAC being the least efficient stop codon context. In case of readthrough, UGA is decoded as UGG and incorporates a tryptophane (Trp) residue in the peptide chain ultimately leading to a C-terminally extended peptide. Stop codon readthrough is very prevalent in viruses to expand their proteome but has also been recently validated in mammalian genes (Eswarappa et al., 2014; Firth and Brierley, 2012; Loughran et al., 2018; Yamaguchi et al., 2012). Interestingly, in a species of ciliates where all three stop codons are assigned to amino acids, translation termination is achieved solely by the context in the mRNA. Such examples demonstrate the evolutionary flexibility of the genetic code (Heaphy et al., 2016; Swart et al., 2016).

6.2.4. 40S recycling and the crosstalk between termination and initiation

Following 60S ribosomal subunit dissociation, the tRNA/mRNA-bound 40S ribosomal subunit is recycled by eIF2D or the heterodimeric complex of DENR and MCTS-1 which are homologous to the C- and N-terminal regions of eIF2D, respectively (Skabkin et al., 2010; Young et al., 2018). Interestingly, it has been reported that the initiation factors eIF1, eIF1A, eIF3 and eIF3j also participate in the recycling process of 40S subunits (Pisarev et al., 2007). In addition, eIF3j has been implicated in 60S subunit recycling (Young and Guydosh, 2019).

Apart from their role in ribosome recycling, eIF2D and DENR/MCTS-1 have also been proposed to drive non-canonical translation by activating translation re-initiation during cellular stress and internal ribosome entry site (IRES)-mediated translation initiation on viral mRNAs (Dmitriev et al., 2010; Skabkin et al., 2010). In this context, these proteins recruit Met-tRNA$_i^{Met}$ and other tRNAs, in a GTP-independent manner, to 40S subunits positioned at AUG start codons. In addition, the DENR/MCTS-1 complex was also found to promote main ORF (mORF) translation following translation of very short upstream ORFs (uORFs) characterized

by strong Kozak consensus sequences (stuORFs) and present in cellular transcripts important for cell growth and proliferation (Ahmed et al., 2018; Schleich et al., 2017; Schleich et al., 2014).

Structural insight into the eIF2D and MCTS-1 bound to the 40S ribosome revealed that their C-terminal domains adopt a fold similar to eIF1 and bind to similar locations on the initiation complex. The striking similarity with eIF1 suggests that functionally, eIF2D and MCTS-1 might modulate the access of tRNAs to the ribosome P-site (Hussain et al., 2014; Llacer et al., 2015; Lomakin et al., 2017; Rabl et al., 2011; Weisser et al., 2017). MCTS-1 and the N-terminal region of eIF2D bind the cytosine-cytosine-adenosine (CCA) tail of the initiator tRNA covalently bound to methionine, but not the methionyl moiety itself (Weisser et al., 2017). This observation indicates that these proteins are not specific to the initiator tRNA, and is in agreement with their role in tRNA recycling and recruitment of both initiator and elongator tRNA to 40S subunits under cellular stress (Dmitriev et al., 2010; Skabkin et al., 2010; Young et al., 2018).

Altogether, several lines of evidence point to the tight interconnection of termination and initiation. First, mRNA circularization allows the crosstalk between the mRNA tail and the 5′ end. Second, a set of initiation factors remains bound to the ribosome during elongation, potentiating re-initiation of translation after termination. Third, eIF3j, eIF2D and DENR/MCTS-1 are integral parts of termination and initiation machineries. These factors are recruited to the terminating ribosome and potentially decide if termination is completed or if translation is resumed. The context under which one or the other event is favored remains to be elucidated.

6.3. Mechanisms of translational control

6.3.1. Regulation of translation by eIF4E-binding proteins (4E-BPs)

eIF4E-binding proteins (4E-BPs) are a group of diverse factors that repress translation by interfering with the assembly of the eIF4F complex. 4E-BPs compete with eIF4G for binding to eIF4E and therefore block eIF4F-mediated translation initiation (Mader et al., 1995; Marcotrigiano et al., 1999). 4E-BPs bind eIF4E with their canonical and non-canonical binding motifs separated by a linker sequence on the dorsal and lateral surfaces of eIF4E, respectively. The canonical eIF4E binding motif (4E-BM) is characterized by a conserved YX(R/K)XXLΦXX(R/K), where Y denotes Tyr, X any amino acid, R/K Arg/Lys, L Leu and Φ any hydrophobic amino acid. In contrast, the non-canonical sequences are less conserved

but are generally composed of hydrophobic residues that bind the lateral surface of eIF4E (Peter et al., 2015).

4E-BP activity is regulated by the phosphatidylinositol-3-kinase (PI3K)/AKT and the mammalian target of rapamycin (mTOR) signaling pathway. Non-phosphorylated 4E-BPs are active and tightly bind eIF4E, thus blocking initiation. In contrast, sequential phosphorylation of 4E-BPs by the mTOR kinase lowers the affinity to eIF4E, allows assembly of the eIF4F complex and activation of translation. This way, cells transfer the information from mitogen and nutritional stimuli to the translation machinery and dynamically regulate protein synthesis to ensure appropriate cell growth and proliferation (Pelletier et al., 2015).

6.3.2. Translational control by eIF4E paralogs

Apart from 4E-BPs, other mechanisms can inhibit the recognition of the mRNA 5′ cap by the eIF4F complex. Eukaryotic cells contain alternative cap binding proteins from the eIF4E family, that compete with eIF4E for cap binding. The eIF4E-homologous protein (4EHP, a.k.a. eIF4E2) acts as a translation repressor as it does not interact with eIF4G and thus fails to initiate translation (Hernandez et al., 2005; Joshi et al., 2004; Rom et al., 1998). eIF4E3 has been found to recruit eIF4G and regulate translation of a specific set of mRNAs (Landon et al., 2014).

6.3.2.1. eIF4E homologous protein (4EHP)

Despite high sequence and structural similarity to eIF4E, 4EHP acts as a translation repressor as it does not associate with eIF4G (Hernandez et al., 2005; Joshi et al., 2004; Rom et al., 1998). Although 4EHP has reduced affinity to the cap (Zuberek et al., 2007), it competes with eIF4E for cap binding upon recruitment to specific mRNAs by RBPs. Thus, 4EHP acts as a message-specific rather than general translation repressor. In *Drosophila melanogaster* embryos, 4EHP is specifically recruited to *caudal* and *hunchback* mRNAs by the RBPs Bicoid (Bcd) and Brain tumor (Brat), respectively (Cho et al., 2006; Cho et al., 2005). Similarly, in mouse oocytes, Prep1 recruits 4EHP to *Hoxb4* mRNA to elicit translation repression (Villaescusa et al., 2009). In mammalian cells, 4EHP has multiple binding partners and therefore can post-transcriptionally control the expression of a wide range of genes. 4EHP interaction with the eIF4E-transporter (4E-T), a 4EBP, contributes to the repression of transcripts by micro (mi)RNAs (Chapat et al., 2017; Kubacka et al., 2013; Rasch et al., 2020). 4EHP also forms a complex with the Grb10-interacting GYF (glycine-tyrosine-phenylalanine domain) protein 2 (GIGYF2), a factor originally identified in the regulation of insulin signaling

in mice (Giovannone et al., 2009; Morita et al., 2012). The 4EHP-GIGYF2 repressor complex is recruited to target mRNAs by the zinc finger protein ZNF598 (Morita et al., 2012), to adenosine-uridine (AU)-rich transcripts by tristetraprolin (TTP) (Fu et al., 2016; Tao and Gao, 2015), or to miRNA regulated transcripts by the miRNA-induced silencing complex-associated TNRC6 proteins (Schopp et al., 2017).

Curiously, 4EHP was also identified to interact with the threonyl-tRNA synthetase (TRS) to initiate translation of mRNAs involved in vertebrate development. The formation of the 4EHP-TRS alternative initiation complex was shown to be structurally similar to the eIF4E-eIF4G interaction (Jeong et al., 2019).

6.3.2.2. GIGYF1 and 2

Human cells express two GIGYF paralogs, GIGYF1 and GIGYF2. Together with 4EHP, GIGYF1/2 form translation repressor complexes with specific RBPs, including the ubiquitin ligase ZNF598, TTP, TNRC6 proteins and potentially NOT4 (Fu et al., 2016; Keskeny et al., 2019; Kryszke et al., 2016; Morita et al., 2012). The interaction with the RBPs is mediated by the GYF domain of GIGYF1/2. GYF domains bind to proline-rich motif (PPGΦ, where Φ is any hydrophobic amino acid except for Trp; Figure 3) present in all these RBPs (Ash et al., 2010; Kofler and Freund, 2006).
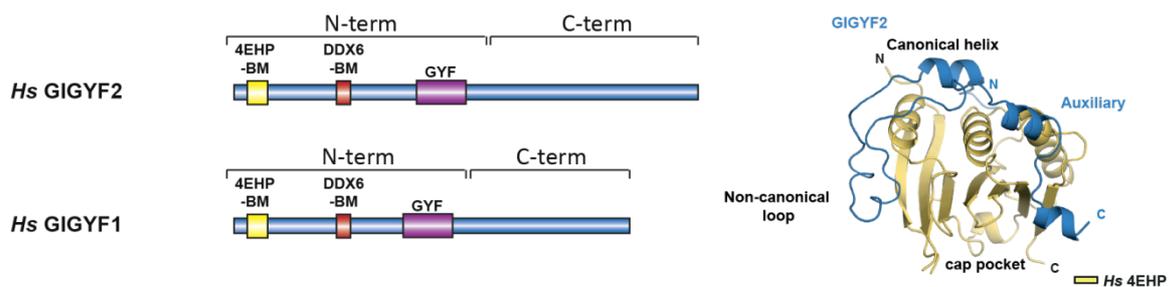
Figure 3: The human GIGYF1 and GIGYF2 proteins and their mode of interaction with 4EHP. N-terminal (N-term) regions of GIGYF1 and GIGYF2 contain the conserved 4EHP-binding motif (4EHP-BM), the DDX6-binding motif (DDX6-BM) and the GYF domain. No known domains are reported for the C-terminal (C-term) regions of GIGYF1 and 2. The 4EHP-BM of GIGYF2 associates with its canonical helix, non-canonical loop sequence and auxiliary sequence on multiple surfaces on 4EHP. 4EHP binds the mRNA 5´ cap in the cap pocket.

TTP binds to the 3´ UTR of mRNAs containing AU-rich elements (AREs), such as *TNFα*. Binding of TTP to the target elicits the turnover of mRNAs involved in inflammatory responses and is expressed in immune cells like macrophages. Dysfunction of TTP is associated with systemic inflammation in animal models while modest overexpression of TTP exerts a protective role against inflammatory diseases in these models (Patial and Blackshear, 2016).

The TNRC6 family of proteins (a.k.a. GW182) is best described for its ability to directly bind argonaute proteins (AGOs), key components of the miRNA-induced silencing complex (miRISC) in metazoans. AGO proteins directly associate with guide miRNAs that target partially complementary sequences in the 3′ UTRs of mRNAs to direct translation repression and decay of the targeted transcript. AGO proteins require additional factors, like the TNRC6 proteins, to mediate mRNA turnover. The TNRC6 proteins recruit translation repressors and decay factors to the miRISC (Jonas and Izaurralde, 2015; Niaz and Hussain, 2018).

ZNF598 (Hel2 in yeast) is an E3 ubiquitin ligase with important functions in mechanisms that surveil the quality of translation elongation to avoid the production of truncated and potentially toxic proteins in the cell. Mechanistically, ZNF598 recognizes and ubiquitinates ribosomes that collided during translation (Garzia et al., 2017; Ikeuchi et al., 2019; Juszkiewicz et al., 2018; Juszkiewicz and Hegde, 2017; Simms et al., 2017b; Sundaramoorthy et al., 2017). Ubiquitinated ribosomes, are signs of damaged translation and elicit Ribosome-associated quality control (RQC) mechanisms that promote recycling of the ribosomal subunits, the degradation of the truncated polypeptide and the decay of the mRNA (see below) (Brandman and Hegde, 2016; Joazeiro, 2019; Simms et al., 2017a).

Another potential GYF domain binding protein is NOT4 (Keskeny et al., 2019), another E3 ubiquitin ligase with functions in cellular protein homeostasis in yeast (Preissler et al., 2015). Although in human cells the function of NOT4 remains largely unknown, the protein has been shown to bind to the CCR4-NOT complex via its CAF40 subunit (Keskeny et al., 2019).

In the repressor complexes, GIGYF1/2 act as large scaffolds that bridge the interaction with additional proteins involved in mRNA turnover. These include the CCR4-NOT deadenylase complex, which removes the poly(A) tail of mRNAs, and the translation repressor and decapping activator DDX6 (Amaya Ramirez et al., 2018; Peter et al., 2019; Ruscica et al., 2019)

In summary, multiple and possibly redundant RBPs recruit the 4EHP-GIGYF1/2 repressor complexes to target mRNAs (see Figure 4 for a simplified summary). However, it is still unclear how and under which conditions the different 4EHP-GIGYF1/2-RBP complexes regulate gene expression. In addition, a comprehensive analysis of mRNAs post-transcriptionally regulated by the different complexes has not been performed to date. As numerous studies associate the loss of GIGYF1/2 or its haploinsufficiency with neurodegeneration and neurological disorders in animal models and affected humans (Giovannone et al., 2009; Iossifov et al., 2014; Krumm et al., 2015; Satterstrom et al., 2020;

Schizophrenia Working Group of the Psychiatric Genomics, 2014; Thyme et al., 2019), a deeper understanding of the mechanism employed by these proteins in the regulation of gene expression and the identification of the targeted transcripts will facilitate our knowledge on these pathologies.
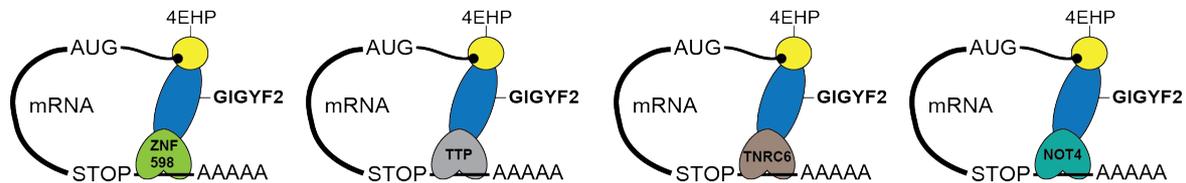
Figure 4: Summary of the multiple repressive complexes formed by human GIGYF2. The 4EHP-GIGYF2 complex forms multiple repressive complexes with ZNF598, tristetraprolin (TTP), TNRC6 and NOT4 proteins to mediate repression of the targeted mRNAs. The possible repressive complexes potentially occur simultaneously in cells and might act redundantly or uniquely on cellular transcripts.

6.3.3. Translational control by eIF4G paralogs

Metazoan cells express multiple eIF4G proteins: eIF4G, eIF4G2 – a.k.a. DAP5/p97/NAT1 – and eIF4G3. eIF4G proteins generally function as scaffolds that bring together multiple actors in the initiation of translation. In their N-termini, eIF4G and eIF4G3 bind PABP and eIF4E. This region is absent in DAP5 (Figure 5). More C-terminal, eIF4G and eIF4G3 have two binding sites (MIF4G and MA3 domains) for the RNA helicases eIF4A and eIF4A2. DAP5 only shares one binding site for eIF4A proteins – the MIF4G domain. The DAP5 MA3 domain is unable to bind eIF4A/4A2 (Liberman et al., 2015). The most C-terminal parts of eIF4G, eIF4G3 and DAP5 each contain a regulatory W2 domain that in eIF4G binds to Mnk kinases (Pyronnet et al., 1999; Shveygert et al., 2010) and in DAP5 to eIF2β (Liberman et al., 2015). Mnk kinases phosphorylate eIF4E thereby altering translation of a subset of mRNAs (Ueda et al., 2004; Uttam et al., 2018).
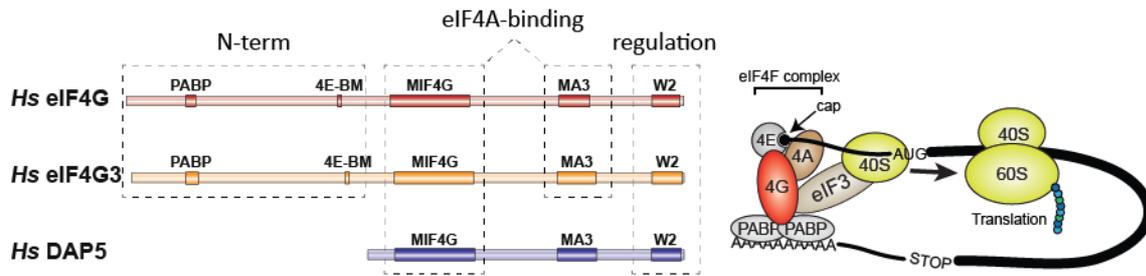
Figure 5: The human eIF4G paralogs and canonical cap-dependent translation initiation. N-terminal (N-term) regions of *Hs* eIF4G and *Hs* eIF4G3 contain poly(A) binding protein (PABP) binding sites and a conserved eIF4E-binding motif (4E-BM). *Hs* DAP5 lacks the N-term region. All eIF4G paralogs share the MIF4G domain which mediates the interaction with the RNA helicase eIF4A. eIF4G and eIF4G3 comprise an additional eIF4A binding site in the MA3 domain. The MA3 domain of DAP5 does not associate with eIF4A. A W2 domain is shared among the eIF4G paralogs but fulfills different regulatory functions. To initiate canonical translation the eIF4F complex – composed of eIF4E, eIF4G and eIF4A – is assembled on the mRNA cap structure. eIF4G mediates the interactions to PABP and eIF3 which directly associates with the 40S small ribosomal subunit. After recognition of the start AUG codon translation of the mRNA begins.

6.3.3.1. DAP5

DAP5 was independently identified by four research groups as p97 (Imataka et al., 1997), death-associated protein 5 (DAP-5) (Levy-Strumpf et al., 1997), eIF4G2 (Shaughnessy et al., 1997) and novel APOBEC-1 target no. 1 (NAT1) (Yamanaka et al., 1997). As DAP5 lacks the N-terminal region of eIF4G proteins, it was initially proposed to act as an eIF4G mimic that sponges translation initiation factors and decreases global translation in cells (Imataka et al., 1997; Yamanaka et al., 1997). However, DAP5 was also required for the initiation of cap-independent translation in subset of mRNAs (Levy-Strumpf et al., 1997). *EIF4G2* (DAP5) mRNA has been shown to be modified by APOBEC-1 which catalyzes the deamination of C to U residues. These modifications on the *EIF4G2* mRNA induce premature stop codons and abolish the expression of the DAP5 protein (Yamanaka et al., 1997). Loss of the DAP5 protein was associated with liver tumor formation in mice.

The function of DAP5 in the initiation of translation has been mainly characterized with bicistronic reporters that drive translation of two coding sequences independently using cap-dependent and cap-independent/IRES-dependent initiation (Henis-Korenblit et al., 2000; Lee and McCormick, 2006; Lewis et al., 2008; Liberman et al., 2009; Marash and Kimchi, 2005; Marash et al., 2008; Nevins et al., 2003; Weingarten-Gabbay et al., 2014; Yoffe et al., 2016). Although these reporter experiments suggest that several cell cycle and apoptosis-associated transcripts, including *EIF4G2* mRNA itself, are translated in a DAP5-dependent manner, little is known about the regulation of translation by DAP5.

Curiously, a more recent study reported that the poly(C)-binding protein 2 (PCBP2) is translated in a DAP5-dependent manner. At the same time PCBP2 binds a C-rich region in the *EIF4G2* 5′ UTR and inhibits the synthesis of the DAP5 protein thus constituting a negative feedback loop (Smirnova et al., 2019).

Interestingly, DAP5 protein synthesis from the *EIF4G2* mRNA always starts at a conserved non-AUG initiation site, for instance GUG in human and mouse and AUU in *Drosophila melanogaster* (Takahashi et al., 2005).

Homozygous loss of DAP5 in mice leads to embryonic lethality because embryos fail to undergo gastrulation (Yamanaka et al., 2000). Although overall translation was unchanged in DAP5-deficient (NAT1$^{-/-}$) mouse and human embryonic stem (ES) cells, they are unable to differentiate, suggesting that DAP5 controls cellular differentiation (Sugiyama et al., 2017; Yamanaka et al., 2000; Yoffe et al., 2016).

In addition to the reported role of DAP5 in IRES-dependent translation initiation, a more recent study finds that DAP5 can bind and utilize the alternative cap-binding protein eIF3d to mediate cap-dependent translation initiation on specific mRNAs (de la Parra et al., 2018).

6.3.3.2.   eIF4G3

eIF4G3 – or eIF4GII – was originally characterized as a functional homolog of eIF4G (Gradi et al., 1998). Like eIF4G, eIF4G3 comes as multiple isoforms arising from multiple promoters, alternative splicing events and usage of non-canonical initiation codons (Coldwell et al., 2012). The expression of eIF4G3 varies widely between cell types but is always expressed at lower levels compared to eIF4G (Coldwell et al., 2012) with highest eIF4G3 expression levels in testis, kidney and brain.

Although eIF4G and eIF4G3 are highly similar and even largely identical in their C-terminal regions, rescue studies have found that eIF4G3 cannot compensate the reduction of translation in eIF4G-depleted cells (Coldwell and Morley, 2006; Ramirez-Valle et al., 2008).

Studies around male-limited infertility in mice revealed that a mutation in the W2 domain of eIF4G3 causes meiotic arrest in mouse spermatocytes and leads to failure in male reproductive function (Hu et al., 2018; Miao et al., 2016; Sun et al., 2010). Since eIF4G3 is located in the nucleus of spermatocytes, eIF4G3 has been suggested to control mRNA metabolism and/or prepare mRNAs for subsequent translation in the cytoplasm (Hu et al., 2018).

6.3.4. The role of the mRNA 5´ UTR in translational control

Recent advances in understanding transcriptome-wide translational control have identified various 5´ UTR-associated and alternative translation events that shape the cellular proteome in higher organisms as a response to environmental changes like starvation, metabolite availability, oxidative stress, unfolded protein response, viral infection or during tumor initiation (Andreev et al., 2015; Ingolia et al., 2011; Ivanov et al., 2018; Pakos-Zebrucka et al., 2016; Sendoel et al., 2017; Starck et al., 2016; Zhang et al., 2015). These alternative translation events rely on *cis*-acting elements present in the 5´ UTR of the transcript and modulate translation initiation at the main ORF (mORF) via various mechanisms. These include the usage of different start codons after leaky scanning of the 43S PIC complex on the 5´ UTR, the use of unconventional initiation factors, re-initiation of translation after upstream ORFs (uORFs), ribosome shunting through structured mRNA sequences that bring distal sequences in close spatial proximity, ribosome frame-shifting and IRES-mediated recruitment of the ribosome that stimulates cap-independent translation initiation. In the following sections, selected mechanisms will be described in greater detail.

6.3.4.1. uORFs

Approximately 40-50% of mammalian mRNAs contain at least one uORF upstream of the mORF. This number, estimated by the presence of conserved start codons upstream of the mORF, or using approaches that capture translating ribosomes on the mRNA (ribosome profiling) (Calvo et al., 2009; Iacono et al., 2005; Ingolia et al., 2011; Matsui et al., 2007), suggests the presence of an additional layer of translational control in mammals.

uORFs are short coding sequences defined by an initiation codon in frame with a STOP codon located upstream of the mORF. In most of the described cases, uORF translation reduces mORF expression by 30-80% (Calvo et al., 2009). However, there are examples where the presence of a uORF enhances mORF translation (Hinnebusch, 2005; Young and Wek, 2016).

In contrast to mORF start sites which are often marked by an AUG, uORFs frequently initiate with alternative, near-cognate start codons, like CUG, GUG, UUG and others (Ingolia et al., 2011; Sendoel et al., 2017). The usage of near-cognate start codons has impaired the identification of uORFs using computational methods. Recent advances with ribosome profiling experiments that capture initiating ribosomes (Ingolia et al., 2011; Lee et al., 2012; Sendoel et al., 2017), have greatly improved our current knowledge on the presence of uORFs in mammalian transcripts. Due to the frequent presence of near-cognate start codons, uORFs

have weak initiation contexts and rely on molecular mechanisms that increase the likelihood of the scanning complex to recognize the start codon. Secondary structures in 5′ UTR sequences, for instance, are proposed to slow down the scanning PIC, increase its dwell time on the mRNA and favor the initiation of translation in non-optimal sequence contexts (Guenther et al., 2018). In addition, alternative initiation factors that preferentially recognize near-cognate codons, may also facilitate initiation at uORFs (Starck et al., 2016).

Known examples of alternative initiation factors priming uORF translation include eIF2A which binds to leucine (Leu)-tRNA independently of GTP and initiates translation at CUG (Leu) start codons in 5′ UTRs (Liang et al., 2014; Starck et al., 2012; Starck et al., 2008; Starck et al., 2016). In yeast, the DEAD-box RNA helicase Ded1p (DDX3 in mammals) plays an important role in translation initiation as it resolves RNA secondary structures present in the 5′ UTR. In the absence of Ded1p, scanning complexes, stalled upstream of the structures, recognize near-cognate start codons. As a result, the ratio of mORF to uORF translation decreases (Guenther et al., 2018).

The best studied example of translational control by uORFs is provided by the *ATF4* mRNA (*GCN4* in yeast). While general protein synthesis is inhibited, *ATF4* mORF is increasingly translated when cells are exposed to stress (Blais et al., 2004). During cellular stress, the α subunit of eIF2 is phosphorylated by one of the four eIF2α kinases: PERK, PKR, GCN2 and HRI. Phosphorylation of eIF2α allows binding of eIF2B to eIF2 but blocks its ability to recycle eIF2-bound GDP into GTP. Since eIF2B is a limiting factor in cells, under stress conditions its availability is highly decreased, leading to global translation shut down (Sonenberg and Hinnebusch, 2009).

Mechanistically, translation of *ATF4* occurs via a combination of re-initiation and leaky scanning. *ATF4* mRNA harbors two uORFs in the 5′ UTR. uORF1 is short and located upstream of the mORF. uORF2 overlaps with the mORF. uORF1 is re-initiation permissive as it allows 40S subunits to resume scanning in the 5′ UTR after translation termination. In conditions of high availability of active eIF2, the 40S subunit quickly reacquires the eIF2 ternary complex and initiates translation at the uORF2 start site. As uORF2 overlaps with *ATF4* mORF, uORF2 translation prevents initiation on the mORF start site. Under stress, however, low levels of active eIF2 delay its binding to scanning 40S subunits, which skip uORF2 start site and initiate translation at the mORF (Gunisova et al., 2018). Although this model of translational control on *ATF4* mRNA according to the cellular conditions is widely accepted, several independent studies have reported a concomitant increase in uORF2 and mORF

translation during cellular stress. Such experimental evidences indicate that regulation of alternative translation under stress conditions is more complicated than originally thought (Andreev et al., 2015; Sidrauski et al., 2015; Starck et al., 2016).

Certain metabolites can also induce a switch in the translation machinery to translate uORFs. Polyamines, for instance, autoregulate the translation of mRNAs encoding metabolic and regulatory proteins in the polyamine biosynthesis pathway, which contain conserved uORFs (Dever and Ivanov, 2018). Under conditions of lower polyamines, most ribosomes bypass the uORFs and translate the main ORF. High polyamine levels, however, competitively inhibit elongation factor eIF5A activity and cause ribosome stalling at problematic peptide motifs, such as poly-proline motifs, during translation elongation. eIF5A suppresses ribosome pausing at sites of non-optimal peptide formation. Under these conditions, scanning ribosomes form a queue behind the stalled ribosome, increasing the frequency of translation initiation at uORFs at the expense of the synthesis of proteins (mORF) required for polyamine synthesis (Ivanov et al., 2018).

As 5′ UTRs have nucleotide composition distinct from coding sequences, the codon usage within a uORF is also largely different to that of a mORF. In detail, in vertebrates 5′ UTR sequences have higher G/C content than their associated coding and 3′ UTR sequences (Zhang et al., 2004). Consequently, although it remains to be experimentally validated, the theoretical coding capacity of uORFs compared to the mORF is biased towards glycine (Gly), alanine (Ala), proline (Pro) and arginine (Arg) amino acids. This biased sequence composition together with G/C-induced secondary structures in the 5′ UTR might slow down or even stall the ribosome during uORF translation. Accordingly, a recent study found that collided ribosomes, so called disomes, are characteristic for uORFs, indicating pervasive ribosome stalling during uORF translation (Tuck et al., 2020). If the stalled ribosomes do not resume translation or are recycled by quality control mechanisms, the likelihood of ribosome frame-shifting is increased, and ribosomes start to translate in a distinct frame. Ribosome frame-shifting is frequent during translation of viral transcripts, as it expands the coding potential of the viral mRNAs. An attractive example is observed in coronaviruses, where pseudoknots present in the viral mRNA slow down ribosomes and promote programmed -1 ribosomal frame-shifting to decode two partially overlapping ORFs (Roberts et al., 2009).

uORF translation also has the potential to extend the coding capacity of mammalian transcriptomes. Multiple micropeptides can be produced and fulfill specific cellular functions. Notably, uORF-derived micropeptides can influence cell growth and interact with or modulate the function of the proteins encoded downstream (Chen et al., 2020). uORF translation

therefore challenges the classical monocistronic assumption about mammalian transcriptomes. However, despite extensive effort, the identification of uORF-derived peptides remains challenging (Slavoff et al., 2013). Likewise, it is unclear how many uORF-derived peptides have a biological function or if the uORFs mostly serve as regulatory units for mORF expression (Somers et al., 2013).

Since there are many seemingly unrelated examples of alternative translation described in literature, it will be important to extract the common principles to generate a unifying picture of how 5′ UTR-associated mechanisms contribute to translational control.

6.3.4.2.　IRES

IRES sequences are RNA elements with specific secondary structures that recruit the ribosome to internal regions of the mRNA bypassing cap-dependent translation initiation. Historically, IRES sequences were first discovered in viruses such as the encephalomyocarditis virus (EMCV) (Jang et al., 1988) or hepatitis C virus (HCV) (Tsukiyama-Kohara et al., 1992). In viruses, the IRES sequences are categorized based on their secondary structure complexity and the protein factor requirement. In fact, IRES sequences with more tightly folded structures tend to require fewer additional factors for translation initiation (Yang and Wang, 2019).

In addition, IRES sequences were also described in cellular mRNAs, often encoding proteins produced during stress response, a condition when cap-dependent translation is inhibited (Komar and Hatzoglou, 2011). However, the efficiency and even existence of many cellular IRES sequences remains controversial, predominantly because most experimental evidence is based on bicistronic reporter assays, the reliability of which has been challenged (Gilbert, 2010; Kozak, 2005).

6.4. mRNA turnover

Precise regulation of gene expression in eukaryotes is controlled by an interplay of mRNA abundance and translation. While mRNA translation is the best predictor for protein abundance, mRNA levels can explain around 40% of the observed variation in protein levels (Schwanhausser et al., 2011). mRNA abundance is influenced by two events: mRNA biogenesis (transcription) and mRNA decay. Most of the 40% of variance in protein levels can be accounted to different transcription rates, while mRNA stability has a considerably smaller role. Nevertheless, control of mRNA stability defines the temporal expression of genes enriched for transcription and signaling factors, chromatin modifiers and genes with functions

in the cell-cycle (Friedel et al., 2009; Hao and Baltimore, 2009; Schwanhausser et al., 2011). In addition, mRNA decay safeguards the quality of the mRNA present in cells (Isken and Maquat, 2007).

Two main decay pathways control mRNA stability. Both pathways begin with the shortening of the poly(A) tail by poly(A)-nucleases 2 and 3 (PAN2/PAN3) and the carbon catabolite repressor 4-negative on TATA (CCR4-NOT) complex. After deadenylation, the mRNA can be further degraded in the 5´-to-3´ or 3´-to-5´ directions. In the 5´-to-3´ decay pathway, the deadenylated mRNA is decapped by DCP2 and degraded by the exonuclease XRN1 (Collart, 2016). 3´-to-5´ decay of deadenylated mRNAs is achieved by the exosome (Mitchell, 2014) and DcpS, a pyrophosphatase that hydrolyses the residual cap structure following exosome activity (Wang and Kiledjian, 2001).

The CCR4-NOT is a multi-subunit complex with roles during transcription initiation, nuclear export, translation regulation and mRNA degradation (Collart, 2016; Miller and Reese, 2012). Cytoplasmic deadenylation, the stepwise removal of adenosine residues from the 3´ end of an mRNA, is however the best studied function of the complex (Wahle and Winkler, 2013). The complex is generally recruited to mRNAs by RBPs that recognize specific motifs located in the 3´ UTR (Bhandari et al., 2014; Chicoine et al., 2007; Leppek et al., 2013; Raisch et al., 2016; Sgromo et al., 2017), but is also involved in bulk mRNA deadenylation (Collart, 2016). The deadenylated mRNA, can subsequently remain silenced in the cell or undergo decay. The fate of the mRNA is then determined by the proteins associated with it.

In yeast, XRN1 mediates general mRNA decay while the exosome is critical for decay of defective transcripts (Frischmeyer et al., 2002; Tsuboi et al., 2012; van Hoof et al., 2002). However, we only start to understand the relative contributions of XRN1 and the exosome as well as their coordination to elicit mRNA decay in metazoans. Studies that assayed the transcriptome-wide function of human XRN1 by RNA-sequencing and ribosome profiling revealed a pervasive accumulation of mRNA and 3´ fragments in the absence of XRN1. These studies indicate that human XRN1 participates in constitutive decay of a subset of cellular mRNAs and in the decay of mRNA fragments following endonucleolytic cleavage presumably to clear nonsense-mediated decay (NMD) targeted mRNAs (Chang et al., 2019).

6.4.1. Translation-associated quality control mechanisms

Eukaryotes evolved surveillance mechanisms to safeguard the integrity of their proteome. In many ways, translation serves as a readout to dictate the fate of an mRNA. Aberrant

translation, slow initiation, elongation or termination rates and increased or reduced translation efficiency have the potential to elicit translation surveillance mechanisms that trigger decay of the ribosome-bound mRNA (Roy and Jacobson, 2013). Translation-coupled decay of the mRNA is executed by 5′-to-3′ or 3′-to-5′ exonucleases, as well as by endonucleases. In the case of 5′-to-3′ decay, XRN1 activity follows the last translating ribosome (Hu et al., 2009; Pelechano et al., 2015). In contrast, exosome-dependent decay of the mRNA requires the removal of ribosomes (Guydosh and Green, 2017). Endonucleases cleave the ribosome-bound mRNAs and generate 3′- and 5′-fragments that will be degraded by XRN1 and the exosome, respectively.

6.4.2. mRNA quality control on defective transcripts

Defective mRNAs as a result of damage or faulty processing are co-translationally degraded to reduce the accumulation of potentially toxic proteins in the cell. Distinct surveillance mechanisms coordinate the elimination of the truncated proteins by the proteasome and ribosome recycling with the degradation of the mRNA (Buskirk and Green, 2017; Joazeiro, 2017). Mechanistically, disruption of translation elongation in the defective transcripts is thought to cause ribosome stalling and collision. These signals of aberrant translation activate specific mRNA decay pathways.

6.4.2.1. NMD

NMD targets transcripts with nonsense mutations that create premature termination codons (PTCs) along the CDS. As conventional STOPs are usually located in the last exon of the mRNA, the proximity of PTCs to EJCs loaded at exon-exon junctions during splicing and not removed during the first round of translation, marks the mRNA as defective. Binding of the essential NMD factor UPF1 to the PTC-located ribosome and the EJC, primes a cascade of events that leads to the degradation of the mRNA (Kurosaki and Maquat, 2016).

Alternatively, in faulty mRNAs where the PTCs occur at the beginning of long exons far away from the next EJC, or the PTCs locate in intron-less genes, NMD relies on a distinct mechanism, known as the faux 3′ UTR model. In these cases, the PTC increases the size of the 3′ UTR of the mRNA. Long 3′ UTRs bind more efficiently to UPF1, which recognizes the paused ribosome and triggers mRNA decay (Amrani et al., 2004; Roy and Jacobson, 2013).

After recognition of the PTC, the mRNA is cleaved by the endonuclease SMG6 at the A site of the PTC stalled ribosome. Subsequent iterative upstream cleavages followed by

exonucleolytic decay complete the destruction of the fragmented transcript (Arribere and Fire, 2018; Lykke-Andersen et al., 2014; Schmidt et al., 2015).

Alternatively, exonucleolytic decay of NMD targets can be initiated by SMG5 and/or SMG7 binding to UPF1, followed by the recruitment of the CCR4-NOT deadenylation machinery and 5´-to-3´ decay (Chen and Shyu, 2003; Lejeune et al., 2003; Loh et al., 2013).

6.4.2.2.   NGD and NSD

No-go decay (NGD) and non-stop decay (NSD) surveillance pathways target mRNAs where the ribosome stalls as a consequence of stable secondary structures or the lack of a stop codon on the transcript (mainly caused by premature polyadenylation or mRNA truncation), respectively. Recognition of the stalled ribosomes in NGD and NSD substrates usually results in endonucleolytic cleavage of the mRNA. NGD targets can also be degraded by XRN1 (D'Orazio et al., 2019; Guydosh and Green, 2014; Pisareva et al., 2011; Shoemaker et al., 2010). The machinery responsible for recognition of the stalled ribosome is composed of the PELO (Dom34 in yeast) and HBSL1 (Hbs1 in yeast) proteins, which are paralogs of eRF1 and eRF3, respectively. Together with ABCE1, these ribosome rescue factors dissociate the stalled ribosome on the defective mRNA (Pisareva et al., 2011).

6.4.3. Co-translational quality control on newly synthesized polypeptides

Most neurodegenerative diseases are associated with abnormal accumulation of specific misfolded and aggregated proteins in the brain tissues. Perturbations in the homeostasis of the endoplasmic reticulum (ER) have been linked with the disease process and ER stress is often observed in pathological conditions like ALS, Parkinson and Alzheimer's disease (Matus et al., 2011).

6.4.3.1.   ER-associated surveillance mechanisms

The secretory pathway produces approximately 36% of the total proteome in eukaryotic cells (Uhlen et al., 2015). The entry point for proteins to the secretory pathway is the ER. A hydrophobic signal sequence, often found in the N-terminus of secreted proteins, is recognized co-translationally by the signal recognition particle (SRP) which leads to a temporary slow-down of translation elongation (Walter and Blobel, 1981). Subsequently, the SRP-ribosome-nascent chain complex is targeted to an ER translocon to facilitate co-translational translocation across the ER membrane or integration into the lipid bilayer in the case of membrane proteins

(Nyathi et al., 2013). When misfolded proteins accumulate in the ER, cells react to the stress and activate the unfolded protein response (UPR). Early UPR responses aim at decreasing protein synthesis at the ER. eIF2α phosphorylation by protein kinase RNA-like ER kinase (PERK) inhibits translation initiation (Teske et al., 2011). Furthermore, activation of regulated inositol-requiring 1 (Ire1)-dependent decay (RIDD) results in the unconventional splicing of X box-binding protein 1 (XBP1u; Hac1 in yeast) mRNA and selective decay of ER-bound mRNAs, as demonstrated in *Drosophila melanogaster* (Hollien and Weissman, 2006). During RIDD, targeted mRNAs are endonucleolytically cleaved and degraded through the combined actions of XRN1 and the exosome (Hollien and Weissman, 2006; Kimmig et al., 2012) and ribosomes are rescued by Dom34 and Hbs1 (Guydosh et al., 2017). The active transcription factor (XBP1s) induces a transcriptional switch that results in the expression of genes involved in protein folding, ER-associated degradation (ERAD), protein quality control and phospholipid synthesis (Hetz, 2012). Furthermore, the IRE1α-JUN N-terminal kinase (JNK) pathway activates autophagy.

Alternatively, cells can attenuate the translocation of specific proteins during ER stress. This so called 'pre-emptive' quality control pathway, co-translationally reroutes misfolded proteins to the cytoplasm for degradation. Selective translocation of misfolded proteins is dependent on the identity of the signal sequence (length, hydrophobicity/charge and amino acid composition) (Kang et al., 2006).

Co-translational mRNA decay also safeguards the integrity of the secretome by detecting unsuccessful targeting of proteins to the ER. In detail, defects in recognition of the signal sequence on the nascent peptide chain or the ER membrane receptor by the SRP leads to the recruitment of the decay machinery to the ribosome-bound mRNA (Karamyshev et al., 2014; Lakshminarayan et al., 2020; Pinarbasi et al., 2018). However, the details how co-translational mRNA decay occurs on such transcripts remains to be determined.

6.4.3.2.   Other surveillance mechanisms

Quality control mechanisms also regulate the production of an optimal cytoplasmic proteome. Co-translational protein folding and mRNA stability are known to be regulated by changes in ribosome dynamics (Hanson and Coller, 2018; Hu et al., 2009; Radhakrishnan and Green, 2016). Nascent-peptide or codon composition, secondary structures, ribosome-associated factors, defective ribosomes or mRNAs can alter ribosome movement on the transcript (Buskirk and Green, 2017; Joazeiro, 2017).

Decreased elongation rates at inter domain regions during protein synthesis are required for proper protein folding as this mechanism allows the association of specific chaperones with the emerging nascent polypeptide (Thanaraj and Argos, 1996). One mechanism to achieve local translation slow-down involves changes in codon usage. Codon optimality affects the speed of the ribosome decoding synonymous codons due to differences in tRNA abundance, wobble interactions and demand (Gardin et al., 2014; Hanson and Coller, 2018; Narula et al., 2019; Wu et al., 2019; Yu et al., 2015).

Codon usage also affects mRNA stability. In detail, mRNAs enriched in slow decoding (non-optimal) codons are less stable than mRNAs enriched in fast decoding (optimal) codons (Presnyak et al., 2015). Slow translating ribosomes on mRNA regions with non-optimal codons are recognized by the decay machinery, like the RNA helicase DDX6 and the CCR4-NOT complex (Buschauer et al., 2020; Radhakrishnan et al., 2016).

Co-translational mRNA decay mechanisms as well counteract the accumulation of excess and potentially cytotoxic proteins. α- and β-TUBULIN mRNA levels are regulated by a feedback mechanism in response to accumulating depolymerized TUBULIN subunits (Cleveland et al., 1981; Gasic et al., 2019; Gay et al., 1989; Pachter et al., 1987). In this case, decay of the mRNA is primed by the interaction of tetratricopeptide protein 5 (TTC5) with the N-terminus of TUBULIN as the protein emerges from the ribosome (Lin et al., 2020).

7. Aims and objectives

Although multiple studies have reported that higher eukaryotes express multiple members of the eIF4E and eIF4G family of proteins (Gradi et al., 1998; Imataka et al., 1997; Levy-Strumpf et al., 1997; Rom et al., 1998; Shaughnessy et al., 1997; Yamanaka et al., 1997), the selective function of each protein in the regulation of translation and mRNA stability is largely unknown. To overcome this knowledge limitation, I aimed at using CRISPR/Cas9 genome editing and high-throughput sequencing techniques (RNA-sequencing and ribosome profiling) to determine transcriptome and translatome changes in cells lacking specific translation-associated proteins. In combination with well-established molecular biology and biochemical applications, the transcriptome and translatome data were used to test various scenarios aiming at identifying and characterizing the molecular principles employed by these proteins in the regulation of gene expression.

7.1. Identification of the DAP5-dependent translatome

Metazoan cells express different paralogs of the eIF4G family of proteins; eIF4G, DAP5 and eIF4G3 (Gradi et al., 1998; Imataka et al., 1997; Levy-Strumpf et al., 1997; Shaughnessy et al., 1997; Yamanaka et al., 1997). eIF4G is the best characterized member of the family and is recognized to initiate translation of many cellular mRNAs important for cell proliferation (Ramirez-Valle et al., 2008). The need of higher eukaryotes to in addition express other eIF4G proteins is still unclear. DAP5, for instance, is a curious example since it cannot bind eIF4E and therefore is unlikely to participate in cap-dependent translation initiation. I hypothesized that the identification of DAP5-dependent mRNAs could help to study and elucidate the function of the protein in translation and its biological role.

7.2. Identification of 4EHP-GIGYF1/2 targets

Despite the recent biochemical and molecular characterization of the 4EHP-GIGYF1/2 repressor complexes (Fu et al., 2016; Kryszke et al., 2016; Peter et al., 2019; Peter et al., 2017; Ruscica et al., 2019; Tao and Gao, 2015), the biological role of these proteins in human cells remains poorly characterized. To elucidate how this repressor complex shapes the cellular proteome, I set out to identify the universe of transcripts post-transcriptionally regulated by the 4EHP-GIGYF1/2 repressor complexes. My experiments were additionally aimed at revealing the potential molecular mechanisms underlying the association of GIGYF1/2 proteins with

different neurodevelopmental and neurological syndromes (Giovannone et al., 2009; Iossifov et al., 2014; Krumm et al., 2015; Satterstrom et al., 2020; Schizophrenia Working Group of the Psychiatric Genomics, 2014; Thyme et al., 2019).

8. Results and discussion

8.1. 5´ leader sequences impose an additional layer of complexity on mammalian translational control and induce DAP5 dependent translation

The results presented and discussed in the following paragraphs are summarized in the unpublished manuscript attached (A: Weber et al., unpublished). The manuscript contains finalized figures and a detailed description of experiments and methods and is intended for publication.

8.1.1. DAP5 controls translation of transcripts with complex 5´ UTRs

To investigate the role of DAP5 in translation, I performed RNA-sequencing and ribosome profiling experiments in HEK293T wild-type and DAP5-null cells. Matched analysis of transcriptome and translatome changes in the null cells identified a group of transcripts encoding signaling proteins with decreased translation efficiency (TE) in the absence of DAP5. Curiously, most of the identified mRNAs showed pervasive translation of upstream open reading frames (uORFs) in long and GC-rich 5´ UTRs. Interestingly, translation of uORFs was mostly initiated at non-AUG but near-cognate start codons, like CUG, GUG, UUG or AUC. Analysis of ribosome profiling data performed in HEK293 cells in the presence of harringtonine – a translational inhibitor specific for initiating ribosomes – (Lee et al., 2012) further confirmed that these near-cognate codons were recognized as translation start sites. Furthermore, I observed that DAP5 target mRNAs displayed qualitative changes in ribosome occupancy as measured by the distribution of ribosome footprints (RFPs) along the target mRNAs. In detail, the RFPs were skewed towards the 5´ UTRs of DAP5 targets in the absence of DAP5. These findings indicate that DAP5 mediates translation of the main ORF (mORF) of the identified targets but is dispensable for translation of uORFs located in the 5´ UTR upstream of the mORF (Figure 6).

As translation of uORFs is detected in as many as 40-50% of mammalian mRNAs (Calvo et al., 2009; Iacono et al., 2005; Ingolia et al., 2011; Matsui et al., 2007), the denomination of 5´ 'untranslated regions' should be revised and referred to as 5´ leaders, to account for the fact that uORF translation is nowhere close to a rare event.
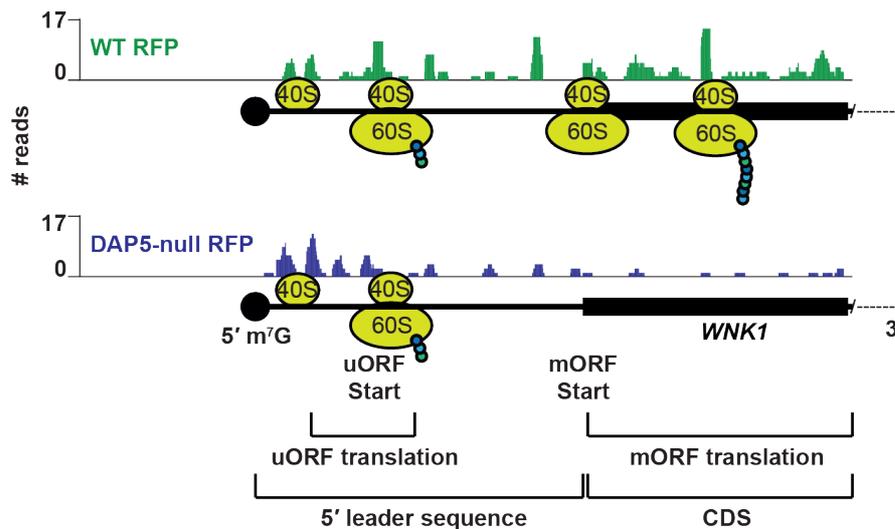
Figure 6: Schematic representation of DAP5-dependent translation. *WNK1* mRNA [exon1 including the 5′ leader and the most 5′ proximal coding sequence (CDS)] and associated ribosome footprints (RFPs) are depicted and represent a model DAP5 target. The 40S small ribosomal subunit scans the *WNK1* mRNA 5′ leader sequence and initiates translation at an upstream open reading frame (uORF) start codon upon 60S ribosomal subunit joining. In control conditions [wild-type (WT)], translation of the main ORF (mORF) occurs in addition to the uORF. Without DAP5 (DAP5-null condition) solely uORF translation occurs and the *WNK1* mRNA mORF is not translated.

Mechanistically, the 5′ leaders of the identified transcripts were sufficient to confer DAP5 dependent translation of a *Renilla* luciferase (R-LUC) ORF cloned downstream in a reporter mRNA. R-LUC protein synthesis from *WNK1*, *ROCK1* and *AKT1* 5′ leader reporters was strongly reduced in DAP5-null cells and could be fully restored by re-expression of full-length DAP5. Importantly, the changes in R-LUC expression were not caused by varying mRNA levels but occurred solely due to altered TE. This observation can be extended to all experiments presented in this section.

8.1.2. eIF4A cooperates with DAP5 for main ORF translation

To study the molecular details of DAP5-dependent translation, I tested the ability of different DAP5 mutants to restore R-LUC expression in null cells. The following mutants or fragments of DAP5 were tested: eIF4A* (unable to recruit the RNA helicase eIF4A), MIF4G (middle domain of eIF4G, eIF4A-binding domain), ΔW2 (lacking the eIF2β interaction domain), an eIF4G ΔN-term (lacking the PABP and eIF4E binding sites and resembling DAP5) and various DAP5 chimeric proteins where the MIF4G, MA3 or W2 domains were replaced by the respective eIF4G domains. None of the constructs fully restored R-LUC protein

synthesis, indicating that all domains of DAP5 are necessary for mORF translation in the context of the *WNK1*, *ROCK1* and *AKT1* 5′ leaders.

RNA-pulldown assays with V5-SBP-tagged DAP5 constructs further revealed that DAP5 efficiently associated with target mRNAs as long as it is able to bind eIF4A. Specifically, the MIF4G domain of DAP5 was sufficient to pull down target mRNAs with similar efficiency as the full-length protein, and the eIF4A* mutant failed to bind to target mRNAs. In agreement with the observation that eIF4A is important for target binding, one third of DAP5 target mRNAs showed Rocaglamide A (RocA) sensitivity, a translation inhibitor that clamps eIF4A onto polypurine RNA sequences (Iwasaki et al., 2016). Interestingly, eIF4G also associated with all tested mRNAs, including DAP5 targets. mRNA binding required the N-terminus of the protein, and thus depended on the interaction with eIF4E and PABP, as eIF4G binds to all capped and polyadenylated mRNAs. These observations suggest that even if main ORF translation is dependent on DAP5, all capped DAP5 target mRNAs are also bound by the eIF4F complex.

In agreement with these observations, closer inspection of the RFP distribution along DAP5-targeted transcripts revealed that the most cap-proximal uORFs were translated in a DAP5-independent manner, as these uORFs were still translated in DAP5-null cells. Cap proximal uORFs are most likely translated in a cap/eIF4F-dependent manner.

8.1.3. DAP5 function requires cap-dependent translation

To determine if DAP5 primes translation using an internal ribosome entry site (IRES) positioned downstream of the cap-proximal uORFs, I performed RNA-pulldown assays in cells where cap/eIF4F-dependent translation was suppressed. Inhibition of eIF4F-dependent translation was achieved by overexpressing a GFP-tagged chimeric 4EBP that competes with eIF4G for eIF4E binding and escapes the regulation by phosphorylation (Peter et al., 2015). Curiously, binding of DAP5 to target mRNAs was strongly inhibited in these conditions, although the protein still interacted with eIF4A. This result demonstrates that DAP5 association with target mRNAs is critically dependent on cap-dependent initiation of translation and excludes the possibility of an IRES mechanism for DAP5 and ribosome recruitment. Thus, DAP5 acts after cap-dependent translation of uORFs.

8.1.4. DAP5 induces re-initiation of translation

Based on these findings, DAP5 might initiate translation downstream of uORFs by distinct mechanisms. The poor initiation context of uORFs in the 5′ leaders of DAP5 targets potentially leads to incomplete recognition of near cognate start sites and decreased initiation of translation. In this context, DAP5 might be able to bind to 43S PICs that scan past the uORFs and – via a process known as leaky scanning – potentially enables translation initiation at the mORF. Alternatively, DAP5 might utilize post-termination complexes and mediate re-initiation of translation after uORF translation. Re-initiation of translation relies on the iterated processes of initiation, elongation and termination while the ribosome is kept on the mRNA and undergoes two or even several rounds of the translation cycle. It is known that re-initiation of translation is dependent on the length of the uORF (Kozak, 2001) and it is assumed that during prolonged ribosome elongation translation initiation factors required for re-initiation are removed from the ribosome.

Interestingly, most of the uORFs observed in DAP5 target 5′ leaders are short (with a median length of 26 amino acids) and compatible with the parameters necessary for re-initiation. To distinguish between these two possible mechanisms of the initiation of translation, I generated a set of *R-LUC* reporters harboring the *WNK1* 5′ leader sequence with uORFs of different lengths. Strikingly, increased uORF length as a result of the removal of STOP codons decreased R-LUC (mORF) expression. These evidences strongly support the notion that DAP5 is involved in re-initiation of translation. In addition, a *WNK1* 5′ leader reporter lacking stop codons, and thus devoid of uORFs, produced N-terminal extended R-LUC proteins from start sites in frame with the mORF but not the short, mORF R-LUC. These results suggest that even if 43S PICs initiate translation at distinct start sites in the *WNK1* 5′ leader via leaky scanning, these complexes are unable to reach the mORF start codon. Thus, as 43S PICs are unable to scan the entire 5′ leaders, DAP5 mediates translation re-initiation after short uORFs.

Altogether, these observations suggest a model where 43S PICs are initially loaded in a cap/eIF4F-dependent manner at the 5′ end of DAP5 targets. In the 5′ leaders of DAP5 targets, the 43S PICs encounter multiple uORFs characterized by near cognate start sites and surrounded or embedded in structured elements. In the presence of this unfavorable initiation context, leaky scanning of the 43S occurs frequently and uORFs are randomly translated. After uORF translation however, DAP5 in complex with eIF4A can bind to the recycling ribosome and keep 40S subunits on the mRNA, which can then be reused for subsequent rounds of

translation initiation until the mORF start site is reached and synthesis of the mORF is initiated. Scanning of the post-termination complexes through the structured 5′ leaders might be favored by eIF4A. In the absence of DAP5, only cap-proximal uORFs are translated by PICs recruited by the eIF4F complex.

In summary, this study reports for the first time a comprehensive set of transcripts that requires the specialized initiation factor DAP5 for translation initiation. These transcripts feature complex 5′ leaders with multiple uORFs that are pervasively translated and permissive for re-initiation at the mORF. Re-initiation after uORF translation is promoted by DAP5 (Figure 7).
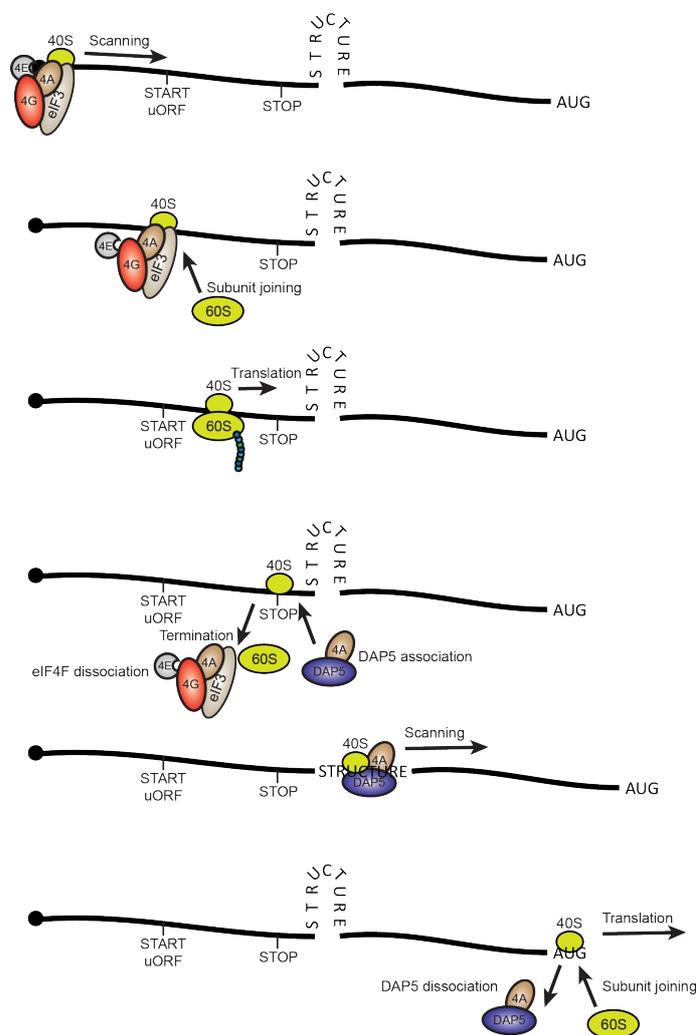
Figure 7: Detailed overview of the different stages in the DAP5-dependent translation re-initiation. DAP5 target mRNAs are first recognized by the cap-dependent translation initiation machinery including eIF4E (4E), eIF4G (4G), eIF4A (4A) and eIF3 which brings the 40S small ribosomal subunit and associated factors to the most 5′ proximal mRNA sequence. In the context of long and structured 5′ leaders the scanning ribosome recognizes uORF start codons and initiates translation. After synthesis of a short peptide encoded within the 5′ leader, translation is terminated and results in the release of the peptide chain and the 60S ribosomal subunit. The DAP5-eIF4A complex allows the 40S ribosome to resume scanning, dissolve the structure and re-initiate translation at the mORF.

8.1.5. DAP5-dependent translation requires termination and ribosome recycling

Canonical translation termination is mediated by eRF1 and ABCE1 which promote peptide chain release and recycling of the 60S ribosomal subunit, respectively. The 40S subunit is then free to resume scanning at the 5′ leader, reacquire the ternary complex and initiate translation at another start site (Gunisova et al., 2018).

To further test the proposed model of re-initiation of translation by DAP5, I interfered with termination of translation and ribosome recycling by overexpressing a dominant negative eRF1 protein or knocking down ABCE1. These experiments showed that although general translation is dependent on these factors, re-initiation of translation by DAP5 in the *WNK1-R-LUC* reporter was highly reduced in conditions that block termination and ribosome recycling. These findings demonstrate that termination after uORF translation is a prerequisite for the re-initiation of translation at downstream ORFs by DAP5.

Several translation factors with roles during the initiation of translation also associate with the termination machinery. These include the non-constitutive eIF3 subunit eIF3J, eIF2D and the related DENR/MCTS-1 complex (Pisarev et al., 2007; Young and Guydosh, 2019; Young et al., 2018). While eIF3J has been found to be involved in 60S recycling (Young and Guydosh, 2019), eIF2D and the DENR/MCTS-1 complex recycle 40S ribosomal subunits (Young et al., 2018). In addition, eIF2D and DENR/MCTS-1 have been implicated in re-initiation after stuORFs (short uORFs) translation (Ahmed et al., 2018; Schleich et al., 2017; Schleich et al., 2014). To test if any of these proteins is involved in DAP5 dependent re-initiation of translation, I conducted shRNA-mediated knockdown and tested the expression of R-LUC in the context of *WNK1* 5′ leader. Double-depletion of eIF2D and DENR/MCTS-1 had no effect on DAP5-dependent expression of R-LUC (*WNK1-R-LUC* reporters). Translation of *R-LUC* in control reporters (without *WNK1* leader or without uORFs) did not vary as well in the absence of these factors. Thus, the function of eIF2D and DENR/MCTS-1 in re-initiation might be restricted to mRNAs with stuORFs characterized by the initiator Met-stop codons (Schleich et al., 2017; Schleich et al., 2014). Likewise, in the absence of eIF3J, R-LUC was still expressed in a DAP5-dependent manner, indicating that eIF3J is not required for the expression of DAP5 target mORFs.

8.1.6. uORF translation is pervasive in the 5′ leaders of DAP5 targets

To detect uORF and mORF translation simultaneously in DAP5 targets, I took advantage of the self-complementing mNeonGreen2 (mNG2) split-fluorophore (Chen et al., 2020; Feng

et al., 2017; Kamiyama et al., 2016). This split-fluorescent approach expresses the mNG2 protein in two fragments: $mNG2_{1-10}$ and $mNG2_{11}$. $mNG2_{1-10}$ originates a non-fluorescent mNG2 due to the lack of the 11th β-strand of the protein; however, upon co-expression with $mNG2_{11}$ (16 amino acids peptide), the two fragments assemble into a functional fluorophore (Chen et al., 2020; Feng et al., 2017; Leonetti et al., 2016). The small size of the 16 amino acids $mNG2_{11}$ was ideal to replace and monitor the expression of one natural uORF in the *WNK1* 5′ leader and further support re-initiation of translation by DAP5. In addition, instead of R-LUC, the mORF encoded the EBFP (enhanced blue fluorescent protein). The split-fluorescent reporters were expressed in cells with a transfection control encoding mCherry. Importantly, the non-overlapping nature of excitation and emission channels of the three fluorophores allowed their simultaneous detection using flow cytometry.

In wild-type cells, all fluorophores – uORF (mNG2), mORF (EBFP) and control ORF (mCherry) – were expressed at the same time. In the majority of cells expressing the mNG2 fluorophore, the EBFP protein was also expressed indicating that uORF and mORF occurred simultaneously. In addition, this observation supports the notion that mORF translation results from re-initiation after each round of uORF translation. In agreement with the observation that uORF translation occurs in the absence of DAP5, null cells still expressed the mNG2 fluorophore. In contrast, EBFP (mORF) expression dropped considerably in DAP5-null cells but was restored by re-expression of DAP5. If the length of $mNG2_{11}$ uORF was long (188 amino acids) and refractory to re-initiation, mORF/EBFP expression was impaired. However, mNG2 fluorescence was not affected by the length of the uORF. These experiments allowed the simultaneous quantification of uORF and mORF translation, revealed that uORFs are translated independently of DAP5 and uORF length, and supported the notion that DAP5 drives mORF via a re-initiation mechanism.

The development of this uORF/mORF reporter system has also far-reaching implications for potential applications outside of this study. First, it is a versatile tool that can be easily adapted to various experimental conditions and biological questions. Second, the readout is universally applicable. Third, it overcomes the challenges of measuring the abundance of the very short peptides encoded in uORFs by western blotting and mass spectrometry-based assays (Slavoff et al., 2013). Lastly, the system is easily scalable and cost-effective and thus can be applied to large-scale studies.

8.2. DAP5 and novel insights into mammalian re-initiation mechanisms

8.2.1. DAP5 re-initiates transcript-specific translation

To date, only few mRNAs have been reported and confirmed to rely on re-initiation for the synthesis of mORF proteins. The data presented here, considerably expands the set of transcripts dependent on re-initiation of translation for protein output.

In literature, *ATF4* (*GCN4* in yeast) mRNA is the best studied example of translation re-initiation (Gunisova et al., 2018). ATF4 synthesis is increased under stress conditions causing eIF2α phosphorylation and block of translation. Since synthesis of ATF4 only occurs in stress conditions, this mRNA might have escaped its identification as a DAP5 target in my experiments. To explore if DAP5 is also involved in re-initiation of translation in cells exposed to stress, I treated wild-type and DAP5-null cells with Thapsigargin, a drug that induces ER stress, causes phosphorylation of eIF2α and results in an increased synthesis of ATF4 (Oslowski and Urano, 2011). However, the increased ATF4 synthesis during cellular stress was unaltered in DAP5-null cells, suggesting that re-initiation of translation on *ATF4* is independent of DAP5. This result points to the possibility that re-initiation events in mammals can also be mediated by eIF4G.

One important difference between the 5´ leaders of *ATF4* and *WNK1* (as a model for DAP5 dependent re-initiation) or other DAP5 target mRNAs is the lack of GC-rich and structured elements in the *ATF4* 5´ leader. These results suggest that DAP5 is not required for re-initiation of translation on non-structured 5´ leaders. In addition, the tendency of DAP5 targets to contain structured elements in the 5´ leader, is in agreement with the fact that binding of DAP5 to mRNA, and subsequently the DAP5 dependent translation (explained in the previous section), is dependent on the RNA helicase eIF4A. These observations support the notion that the DAP5-dependent mechanism of re-initiation of translation is linked to the structured nature of the 5´ leaders of DAP5 targets. However, the precise mechanism of how the structured regions within the 5´ leaders specify DAP5-dependent translation and why other eIF4G paralogs are insufficient to mediate re-initiation in these cases remains unclear.

8.2.2. DAP5 and translation re-initiation complexes

Drugs like lactimidomycin (LTM) and harringtonine interfere with translation elongation by binding to the 60S ribosomal subunit on the 80S ribosome E and A site, respectively. Although cycloheximide (CHX) and LTM share a similar binding mode to the ribosome E site,

their effects on translation inhibition are very different. CHX stalls actively elongating ribosomes, whereas LTM and harringtonine halt ribosomes at sites of initiation without interfering with elongation. Affinity measurements of the drugs for 80S ribosomes suggest that CHX binds with higher affinity to the ribosome E site than LTM, presumably because it is a smaller inhibitor. Thus, CHX has improved accessibility to the ribosome and competes with the tRNA for binding to the E site of the ribosome. In contrast, LTM is unable to dissociate the tRNA from the E site. Consequently, LTM only acts as an efficient translation inhibitor during initiation where the E-site is empty and accessible for the drug (Garreau de Loubresse et al., 2014).

Treatment of cells with LTM and harringtonine followed by ribosome run-off assays reveals a specific block of ribosomes at initiation sites, which can be visualized as quantitative translation initiation (QTI)-peaks after ribosome profiling (Gao et al., 2015; Ingolia et al., 2011; Lee et al., 2012). Therefore, ribosome profiling datasets from LTM and harringtonine-treated cells are classically used to precisely determine translation start sites. Close inspection of these initiation-site profiling data in HEK293 cells (Lee et al., 2012)identified the exact start codons and translation of uORFs in the 5′ leaders of DAP5 targets. Curiously, the AUG start codon leading initiation at the mORF of DAP5 targets is not detected in the same experiments. Such observation has also been reported at some mORF initiation sites in mRNAs that rely on re-initiation mechanisms during translation, such as *ATF4* (Gao et al., 2015). The inability of the initiation-blocking drugs to halt the ribosome at specific re-initiation start sites, suggests that the translation complexes driving re-initiation are molecularly distinct. Although the underlying reasons are unknown, different scenarios can be proposed. One possibility is that re-initiation at the mORF start site is achieved by a fully assembled 80S ribosome that, after termination of uORF translation, does not undergo 60S recycling and continues scanning until the next start codon. Such mechanism have been observed *in vitro* using ribosome toeprint assays (Skabkin et al., 2013) but have not been validated to occur *in vivo* to date. The size of LTM and harringtonine might spatially hinder binding to the 80S ribosome and favor the association of the molecules to the 60S subunit before joining to the 40S, preferentially stalling the ribosome at initiation sites. Alternatively, the translation complexes at these sites can be characterized by an altered composition of initiation factors or structural differences in the arrangement of the E and A sites of the ribosome that limit or change the binding surface of LTM and harringtonine on the 60S subunit. Further studies are necessary to elucidate the structural conditions of the ribosome during re-initiation events and understand the rationale behind these complex observations. However, the highly structured nature of the 5′ leaders in

DAP5 targets might decrease the efficiency of 40S scanning, even in the presence of eIF4A. In this context, an assembled 80S ribosome would melt the structured elements during transit on the mRNA. In addition, it is tempting to speculate that the mechanism of termination at uORFs that are permissive to re-initiation might be different. As some initiation factors remain associated to the ribosome during translation of short uORF it is conceivable that those initiation factors can inhibit 60S and/or 40S recycling and thus promote re-initiation. It will be important to understand the role of DAP5 in this mechanism.

Re-initiation of translation is prominent in bacteria and viruses as it greatly increases the coding capacity of the respective genomes (Powell, 2010; Yoo and RajBhandary, 2008). Together with recent findings (Chen et al., 2020), the data presented here generally challenge the monocistronic assumption in human mRNAs. The widespread and pervasive expression of uORFs in human bi- or multicistronic transcripts also highlights the potential production of uORF-derived peptides with important cellular functions. Indeed, Chen and collaborators (Chen et al., 2020) have shown that uORF-encoded peptides aid the mORF-encoded protein to fulfill its cellular function. Notably, via a re-initiation mechanism, the expression level of the regulatory peptide is at least equal, if not in excess, to the mORF encoded protein. Due to the functional role, mutations in the sequence encoding different uORF peptides resulted in phenotypic changes and decreased cell viability (Chen et al., 2020). Based on these observations, it would be interesting to investigate if the uORF encoded peptides in the 5′ leaders of DAP5 targets similarly perform important cellular functions and if those peptides also tend to associate in cis with the mORF encoded protein. For such studies it will be important to preserve the natural sequence of target 5′ leaders and the mORF and design reporter constructs that are engineered to a minimal degree.

8.3. DAP5 and its role in development and cancer

DAP5 plays important roles during mouse embryonic development. In detail, loss of DAP5 in these animals resulted in embryonic lethality because embryos fail to undergo gastrulation (Yamanaka et al., 2000). Subsequent studies recognized a function for DAP5 in cellular differentiation, but not proliferation (Sugiyama et al., 2017; Yoffe et al., 2016). In agreement with these findings, the DAP5 targets identified in my studies encode proteins related to WNT/FGF-, Ca^{2+}- and insulin-signaling which play a crucial role during differentiation (Basson, 2012). Importantly, some of the mRNAs identified in mice by sequencing polysome-associated mRNAs from cells with and without DAP5 (Sugiyama et al., 2017) overlap with the

transcripts identified in my studies by ribosome profiling (e.g. *WNK1*, *SOS1*, *CREBBP*, etc.). These findings suggest that the function of DAP5 in re-initiation of translation might be conserved between human and mouse. Furthermore, the results of my studies highlight that DAP5-dependent re-initiation of translation is probably an important mechanism during cell differentiation, but is not necessarily required for cell proliferation.

A recent study has described that synthesis of the poly(C)-binding protein 2 (PCBP2) is dependent on DAP5 (Smirnova et al., 2019). *PCBP2* also belongs to the list of transcripts with decreased translation efficiency in the absence of DAP5 compiled in my studies. PCBP2 binds to a C-rich sequence in the 5′ leader of *EIF4G2* (DAP5) and inhibits translation (Smirnova et al., 2019). Thus, DAP5 and PCBP2 regulate one another's expression via a feedback loop. However, it remains to be determined if this regulatory mechanism is constitutive or activated by specific cellular contexts. In a broader sense, it is unclear how the activity of DAP5 is regulated during development in order to promote the differentiation of stem cells (DAP5: off) into specific cell types (DAP5 and re-initiation: on). The abundance and/or activity of proteins such as PCBP2 that dynamically regulate DAP5 levels could provide a working foundation to start investigating the mechanisms underlying changes in DAP5-dependent re-initiation of translation during cellular differentiation. Regulation of the recruitment of DAP5 by eIF4A to target mRNAs, changes in the subcellular localization of the protein (e.g. stress granules) or additional factors that modulate translation termination might also dictate the efficiency of the DAP5 dependent re-initiation.

Although multiple translation initiation factors are regulated post-translationally, increased *EIF4G2* expression, and thus increased DAP5 abundance, is associated with low overall survival in patients with metastatic breast cancer, as shown in the The Cancer Genome Atlas (TCGA) data set (de la Parra et al., 2017). This observation suggests that DAP5-mediated re-initiation of translation might lead to the production of proteins sustaining or promoting tumor formation. Curiously, severe reduction of DAP5 levels in a highly metastatic breast cancer cell line had no effect on primary tumor growth but abrogated metastasis (de la Parra et al., 2017). In addition, mRNAs with long and structured 5′ leaders and coding for proteins involved in FGF/EGF signaling are translationally hyper-activated during tumorigenesis in mouse mammary gland (Nguyen et al., 2018). Silvestrol, a natural product of the family of flavaglines that inactivates the RNA helicase eIF4A by clamping the helicase onto purine-rich sequences, effectively delayed tumorigenesis in this model system (Nguyen et al., 2018). Although the role of DAP5 in this context was not addressed, several of the mRNAs identified in this study are also DAP5 targets, supporting the connection between DAP5 and cancer progression. These

results highlight the vulnerability of cancers to drugs that target translation initiation, and perhaps re-initiation, on cancer-related mRNAs. Strategies targeting DAP5 or its associated proteins and mechanisms might therefore be promising in the development of novel anti-cancer therapeutics.

8.4. Molecular basis for 4EHP-GIGYF1/2 mediated target mRNA decay

The results discussed in this paragraph were published in D: Peter et al., 2017 and C: Peter et al., 2019. The papers and the corresponding detailed experimental results and methods are attached.

In the publication of D: Peter et al., 2017, I generated a CRISPR-Cas9 engineered HEK293T cell line knockout for the GIGYF1 and GIGYF2 paralogs (GIGYF1/2-null cells). This cell line, was used to evaluate the ability of 4EHP to repress the expression of a luciferase reporter mRNA. In detail, λN-HA-tagged 4EHP was tethered to a *Renilla* luciferase (R-LUC) reporter containing five Box B hairpins from the λ phage in the 3′ UTR. To prevent deadenylation and degradation of the reporter in the cytoplasm of HEK293T cells, the 3′ end of the MALAT non-coding RNA (ncRNA) was inserted after the Box B elements and a poly(A) stretch (Wilusz et al., 2012). The internal A_{95} sequence increases mRNA nuclear export and translation in the cytoplasm. This reporter therefore allowed me to solely study the inhibition of translation by 4EHP.

Tethered 4EHP induced a strong reduction in R-LUC activity in control cells. However, in GIGYF1/2-null cells inhibition of translation by 4EHP was greatly impaired. Rescue assays in the null cells revealed that inhibition of translation by 4EHP required its interaction with GIGYF1/2, as only the re-expression of wild-type, but not the mutant of GIGYF2 that fails to interact with 4EHP (GIGYF2 C*), decreased R-LUC activity. In addition, direct binding of wild-type λN-HA-tagged GIGYF2 to the *R-LUC-5BoxB-A95-MALAT* reporter induced similar translational repression. Interestingly, the repressive effect was independent of the ability of GIGYF2 to interact with 4EHP, as the GIGYF2 C* mutant efficiently repressed translation. The ability of 4EHP to bind the mRNA 5′ cap structure was not required for repression of translation, as long as 4EHP was still able to bind the mRNA and recruit GIGYF2.

I also investigated the consequences of the recruitment of the 4EHP-GIGYF1/2 complexes to mRNAs containing AU-rich elements (ARE), instead of the Box B sequences, in the 3′ UTR. AU-rich mRNAs are recognized and destabilized by tristetraprolin (TTP), a known GIGYF1/2-binding partner (Fu et al., 2016; Tao and Gao, 2015). In GIGYF1/2-null cells the expression

of R-LUC from the *R-LUC-ARE-A$_{95}$-MALAT* reporter in the presence of TTP increased two-fold compared to control cells. In this context, the repressive activity of GIGYF2 was dependent on its interaction with 4EHP and the ability of 4EHP to bind the cap. Neither the GIGYF2 C* nor the 4EHP cap* proteins were able to revert the increased reporter expression in GIGYF1/2-null cells.

Subsequent experiments using the R-LUC-ARE-A$_{95}$-MALAT system and published in C: Peter et al., 2019, revealed that the repressive nature of the 4EHP-GIGYF2 complex in addition relies on the binding of GIGYF2 to TTP and the RNA helicase DDX6. The GIGYF2 proteins unable to bind to TTP (GYF*) and DDX6 (WFF*) failed to reduce the expression of the reporter in GIGYF1/2-null cells.

Taken together, these experiments were important to understand the molecular basis of the 4EHP-GIGYF1/2 repressor function and established a working foundation to investigate the role of 4EHP and GIGYF1/2 in translational repression and mRNA decay.

8.5. 4EHP-GIGYF1/2 mediate co-translational mRNA decay

8.5.1. 4EHP and GIGYF1/2 regulate mRNA stability

The results discussed in the following sections are part of the unpublished and attached manuscript (B: Weber et al., unpublished). To investigate the role of 4EHP and GIGYF1/2 in the regulation of translation and mRNA stability, I performed RNA-sequencing and ribosome profiling experiments in GIGYF1/2-null cells (Peter et al., 2017) and 4EHP-null cells (Rasch et al., 2020). Surprisingly, these approaches revealed that very few mRNAs experience changes in translation efficiency in the absence of 4EHP and GIGYF1/2. However, prominent changes in mRNA abundance, and consequently translation, were present in both the cell lines indicating that the primary role of these proteins in human cells is associated with the regulation of mRNA levels. Although the 4EHP-GIGYF1/2 complexes are involved in mRNA decay (Amaya Ramirez et al., 2018; Peter et al., 2019; Ruscica et al., 2019), the number of up- and down-regulated mRNAs was similar in the null cells. Alterations in mRNA levels most likely result from direct and indirect effects following the absence of these proteins that affect both transcription and mRNA stability. To understand the role of the repressor complex in mRNA decay, I studied in more detail the group of transcripts upregulated in both null cell lines. The 4EHP and GIGYF1/2 datasets had a significant overlap of 82 upregulated genes, many of which encoding secreted and membrane-bound proteins. To determine if the stability of several candidate mRNAs was controlled by the 4EHP-GIGYF1/2 complexes, I measured mRNA half-

lives after transcriptional shut-off by Actinomycin D in control and GIGYF1/2-null cells. In agreement with increased levels, the decay rates of several of the transcripts (*DBNDD2*, *CD109*, *ITPR3*, *NPTX1*), but not *β-ACTIN*, were lower in the absence of GIGYF1/2, confirming that these mRNAs are post-transcriptionally regulated by the 4EHP-GIGYF1/2 complexes.

Rescue assays in the null cells with wild-type and mutant 4EHP or GIGYF1/2 proteins, revealed that target mRNA decay requires multiple co-factors of the repressor complex. Wild-type 4EHP restored target mRNA decay in the null cells if co-expressed with GIGYF1/2; yet, 4EHP variants unable to bind the cap or GIGYF1/2 did not. Likewise, disruption of the interactions of GIGYF1/2 with 4EHP, DDX6 and GYF domain binding proteins also compromised the ability of the complex to induce target mRNA decay. These results demonstrate that GIGYF1/2 proteins scaffold 4EHP, DDX6 and GYF domain binding proteins to mediate target mRNA decay.

RNA-immunoprecipitation (RNA-IP) assays with GFP-tagged GIGYF2 coupled with RT-qPCR, showed that the GYF domain of GIGYF2 plays a crucial role in the binding of the protein to target mRNAs. Moreover, 4EHP only associated with the transcripts in complex with GIGYF2.

8.5.2. 4EHP-GIGYF1/2 trigger co-translational mRNA decay

To pinpoint the molecular features of the mRNA required for 4EHP and GIGYF1/2-targeted decay, I generated several reporters containing the coding sequence (CDS) or the 3´ UTR of selected transcripts. These experiments revealed that, at least in the cases of *DBNDD2* and *LGALS3BP* (another 4EHP-GIGYF1/2 target), mRNA decay was determined by the CDS and required active translation. *DBNDD2* and *LGALS3BP* CDS reporters containing early stop codons did not bind and were not destabilized by the 4EHP-GIGYF1/2 complexes. Inhibition of translation by harringtonine and puromycin resulted in a strong reduction of the ability of GIGYF2 to bind to *ENO2* (another 4EHP-GIGYF1/2 target). These experiments indicate that 4EHP and GIGYF1/2 mediate translation-coupled mRNA decay.

8.5.3. Ribosome pausing marks 4EHP and GIGYF1/2 target mRNAs

Closer inspection of the ribosome footprint distribution along the CDS of 4EHP-GIGYF1/2 targeted mRNAs revealed pronounced ribosome pause sites as determined by the local accumulation of footprints in comparison to the median footprint coverage on the transcript.

Prominent examples were *LARGE2*, *CXCL16*, *NCKIPSD*, *IFRD2*, *DBNDD2* and *ENO2*, as illustrated in Figure 8.

To verify that the observed accumulation of footprints resulted from ribosome pauses/stalling, I reanalyzed ribosome profiling data of ribosome run-off experiments following harringtonine or lactimidomicin (LTM) treatment of HEK293 cells (Lee et al., 2012). Curiously, many of the candidate ribosome pause sites observed in the 4EHP and GIGYF1/2 target mRNAs were also detected in these datasets (see also Figure 8). As most of pause sites are not located at translation initiation sites, the most plausible explanation, is that the observed footprints correspond to ribosomes trapped during elongation that are unable to conclude elongation (run-off) following treatment with harringtonine or LTM.

In addition, disome profiling data in mouse embryonic stem cells (Tuck et al., 2020) and human HEK293 cells (Han et al., 2020) identified in *IFRD2* and *LGALS3BP* increased disome occupancy at the corresponding pause sites. As disomes are a sign of ribosome collisions and aberrant translation (Inada, 2020), these analyses confirm that transcripts with stalled ribosomes accumulate in GIGYF1/2 and 4EHP targets.

The GIGYF1/2 interactor and ribosome collision sensor ZNF598 is a major player in ribosome-associated quality control (RQC) (Inada, 2020; Juszkiewicz et al., 2018). To explore if 4EHP-GIGYF1/2 complexes coordinate with ZNF598 the decay of mRNAs marked by aberrant translation, I compared RNA-sequencing and ribosome profiling data from HEK293T ZNF598-null cells with the 4EHP and GIGYF1/2 datasets. ZNF598-null cells also displayed pronounced changes in mRNA abundance. However, among the upregulated mRNAs only a small group was co-regulated in all three datasets. Interestingly, 4EHP and GIGYF1/2 target mRNAs with prominent ribosome pauses (e.g.: *ENO2*, *CXCL16*) also showed increased abundance in the absence of ZNF598. Subsequent measurements of *ENO2* and *CXCL16* mRNA half-lives revealed that these mRNAs are also post-transcriptionally regulated by ZNF598. In summary, these findings show that co-translational decay of some 4EHP-GIGYF1/2 targets also depends on ZNF598.
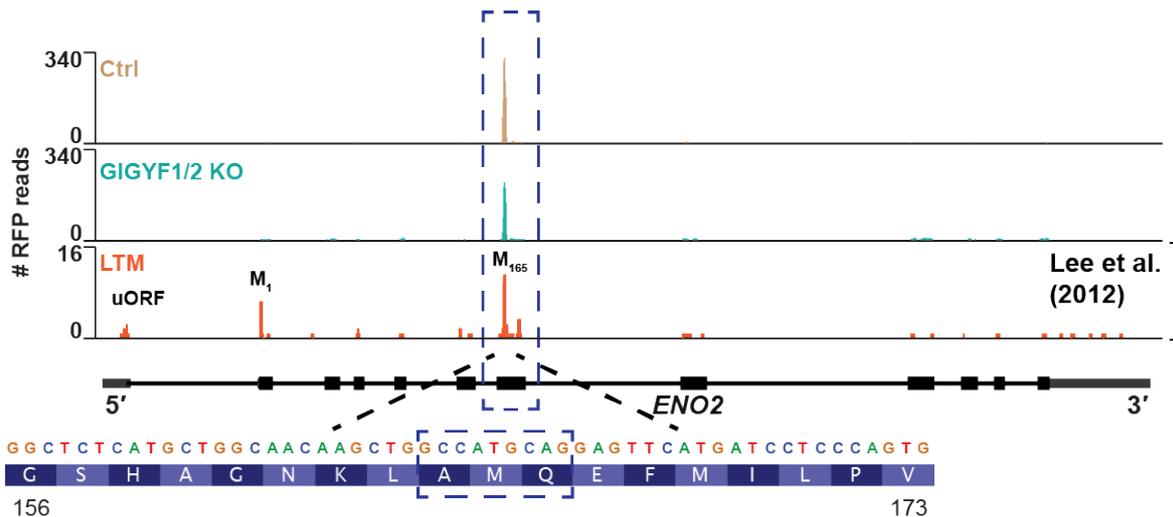
Figure 8: Translational stalling in 4EHP and GIGYF1/2 target mRNAs. *ENO2* mRNA and associated ribosome footprints (RFP) are depicted and represent a model 4EHP and GIGYF1/2 target mRNA. The local accumulation of RFPs in wild-type (Ctrl) and GIGYF1/2-null (KO) cells on *ENO2* mRNA occurs around Met (M_{165}) (see dashed box). Ribosome run-off assays in HEK293 cells treated with the translation inhibitor Lactimidomycin (LTM) that selectively stops ribosomes at sites of initiation but not actively elongating ribosomes (Lee et al., 2012) identify the same ribosome stalling site at M_{165}.

8.5.4. GIGYF1/2-dependent mRNA decay is coupled with co-translational ER targeting

Many of the transcripts regulated by 4EHP and GIGYF1/2 are co-translationally targeted to the ER by the signal recognition particle (SRP), which induces transient elongation pausing after recognition of the signal peptide (SP) (Walter and Blobel, 1981). When the mRNA-associated ribosome reaches the ER, translation resumes and the nascent SP undergoes co-translational translocation into the ER (Nyathi et al., 2013).

To study how 4EHP-GIGYF1/2 mediated mRNA decay is linked with ER-targeting, I used the *LGALS3BP* CDS reporter. LGALS3BP is a secreted protein and is targeted to the ER by its 18 amino acids SP. To test if the exposure of the SP and subsequent ribosome pausing is important for 4EHP-GIGYF1/2 mediated mRNA decay, I introduced STOP codons along the CDS to prevent (*STOP₃₀*) or allow (*STOP₆₀*) the signal peptide to emerge from the ribosome exit tunnel. Strikingly, while the *LGALS3BP-STOP₆₀* mRNA was still regulated by 4EHP-GIGYF1/2 the levels of the *LGALS3BP-STOP₃₀* mRNA were increased in control cells and did not respond to changing GIGYF1/2 levels. These data indicated that decay of *LGALS3BP* mRNA by 4EHP and GIGYF1/2 required the exposure of the LGALS3BP SP and co-translational ER targeting.

Notably, GFP-GIGYF2 also co-purified with HA-tagged LGALS3BP in IP experiments. The binding between the two proteins was mediated by GIGYF2 C-terminus, a region of the protein with previously unknown function. The C-terminus of GIGYF proteins is predicted to mainly form alpha helices but contains no known domains. Additionally, the SP of LGALS3BP, and NPTX1, were sufficient to prompt the interaction with GIGYF2. In detail, SP-R-LUC fusion proteins, but not R-LUC alone, interacted with GIGYF2. Importantly, SP-R-LUC fusion mRNAs were only co-translationally degraded in the presence of 4EHP and GIGYF1/2. Specifically, the abundance of *SP$_{LGALS3BP}$-R-LUC* and *SP$_{NPTX1}$-R-LUC* mRNAs was increased in GIGYF1/2-null cells and was restored to normal levels upon co-expression of 4EHP and GIGYF2. In conclusion, these data show that GIGYF1/2 associate co-translationally with the SP of specific secretome-associated factors to mediate mRNA decay, most likely as a result of changes in elongation speed during co-translational ER targeting.

8.5.5. Regulation of *TUBULIN* mRNA stability by 4EHP-GIGYF1/2 complexes

TUBA4A and *TUBB4A* mRNAs were also negatively regulated by 4EHP and GIGYF1/2. *TUBULIN* transcripts are degraded in a co-translational manner when cells accumulate excess of monomeric TUBULIN subunits (Cleveland et al., 1981; Gasic et al., 2019; Gay et al., 1989; Pachter et al., 1987). To artificially induce TUBULIN autoregulation, GIGYF1/2-null and control cells were treated with the microtubule destabilization agent nocodazole. In the presence of the drug, *TUBULIN* mRNA decay was efficiently induced in control cells; for instance, *TUBA1A* levels dropped to 30% of the original abundance. In GIGYF1/2-null cells though, decay of specific *TUBULIN* mRNAs after nocodazole exposure was impaired. *TUBA1A* mRNA levels, for example, were only reduced to 60% suggesting that GIGYF1/2 proteins participate in *TUBULIN* autoregulation.

The N-terminal end of TUBULINs plays a crucial role in mRNA autoregulation (Pachter et al., 1987; Yen et al., 1988). Activation of mRNA decay in conditions of free subunits excess is primed by the binding of TTC5 to the nascent N-terminal end of TUBULIN (Lin et al., 2020). Closer inspection of disome profiling data in human cells (Han et al., 2020) revealed that ribosome collisions occurred 20 to 30 codons following the start of translation in some *TUBULIN* mRNAs (e.g. *TUBA1A*, *TUBA1B*). Elongation along these mRNAs therefore slows down as soon as the N-terminal end of the corresponding TUBULIN protein emerges from the ribosome exit tunnel and binds to cytosolic factors like TTC5. Thus, recognition of the nascent peptide most likely slows down ribosome transit along the *TUBULIN* mRNA, causing

ribosome collision when a trailing ribosome runs into the leading ribosome. Ribosome collision, as a signal of perturbed translation, triggers the destabilization of the mRNA in a GIGYF1/2-dependent manner. However, the order of events linking TTC5 binding to recruitment of 4EHP-GIGYF1/2 remains unclear.

GIGYF1/2-dependent decay of *DBNDD2* mRNA also relied on the recognition of the nascent peptide by a yet unknown factor. Ribosome stalling on *DBNDD2* mRNA occurred around codon 25. Mutational analysis revealed that the decay of *DBNDD2,* and most likely ribosome stalling, was not the result of impaired decoding at the sequence/peptide motif underlying the pause site. However, a CDS reporter mRNA encoding a DBNDD2 protein with a different N-terminal end (M_1DPN to M_1AAA) was refractory to GIGYF1/2-dependent decay. Thus, the most N-terminal amino acids of DBNDD2 are required for the 4EHP-GIGYF1/2 mediated turnover of the translating mRNA.

8.5.6. A model for the role of GIGYF1/2 in co-translational mRNA decay

The data described here indicate that 4EHP-GIGYF1/2 repressive complexes regulate the turnover of specific transcripts encoding secretome-associated products. mRNA decay is coupled with translation and triggered by altered elongation speed due to ribosome stalling and collision. For some target mRNAs, co-translational decay requires the coordinated action of the ribosome collision sensor ZNF598 and 4EHP-GIGYF1/2. These findings expand the functional landscape of 4EHP-GIGYF1/2 beyond the 3′ UTR-associated mRNA turnover mechanisms operating during inflammation and miRNA-mediated gene silencing (Fu et al., 2016; Kryszke et al., 2016). In agreement with these findings, a recent publication connected 4EHP and GIGYF2 with the co-translational surveillance mechanisms that selectively inhibit protein synthesis from aberrant mRNAs (Hickey et al., 2020). In addition, two very recent studies independently identified EDF1 as a ZNF598-independent sensor of ribosome collisions that recruits 4EHP and GIGYF2 to mRNAs with collided ribosomes to reduce temporary ribosome load (Juszkiewicz et al., 2020; Sinha et al., 2020). It is therefore conceivable that ribosome collisions also serve as marks of perturbed elongation and induce canonical (non-aberrant transcripts) mRNA decay upon co-translational binding of 4EHP and GIGYF1/2.

Ribosome transit along the CDS of an mRNA can be altered by different reasons. Extensive secondary structures impede ribosome translocation on the mRNA and lead to NGD (Doma and Parker, 2006). Expanded sequences of A residues encoding Lys stretches, as observed in prematurely polyadenylated mRNAs or when the ribosome translates into the poly(A) tail of

mRNAs lacking stop codons, result in impaired decoding because of steric problems in the decoding center of the ribosome and strong interactions of the Lys stretch with the negatively charged exit tunnel of the ribosome (Chandrasekaran et al., 2019). Specific peptide sequence combinations also cause local slow-down of peptide bond formation and result in ribosome pausing. For example, Pro-rich motifs stall ribosomes on numerous transcripts due to the decreased reactivity of the Pro amino acid and slow peptide bond formation kinetics. As observed for different 4EHP and GIGYF1/2 target mRNAs, ribosome stalling is primed by factors that recognize the nascent peptide. The SP of secretome-associated mRNAs, for instance, associates with the SRP or GIGYF2, TTC5 binds to the N-terminal end of TUBULINS and yet unidentified factor(s) recognize the N-terminus of DBNND2. Multiple adapter/chaperon proteins thus potentially initiate the recruitment of the decay machinery to specific transcripts via GIGYF1/2 proteins, which scaffold and coordinate subsequent mRNA decay.

Target mRNA decay required the assembly of the 4EHP-GIGYF1/2 complexes and their association with the cap structure. Binding of 4EHP to the cap might stabilize the complex on the mRNA. However, because binding of GIGYF2 to mRNA was equivalent with or without the cap binding protein (see RNA-IP experiments), 4EHP might have other functions in the complex. Although 4EHP has decreased affinity for the cap structure compared to eIF4E (Zuberek et al., 2007), recruitment of the protein by GIGYF1/2 increases its local concentration and the likelihood to replace eIF4E at the cap, repressing translation initiation. Additionally, the lower cap binding affinity of 4EHP might expose the 5′ end of the mRNA to decapping enzymes, facilitating decapping and subsequent mRNA decay by XRN1.

The RNA helicase DDX6 is also critical for GIGYF1/2 mediated mRNA decay. Interestingly, DDX6 is known for its role in co-translational decay of transcripts with poor codon optimality in the CDS. Poor codon optimality is the result of the cumulative usage of codons where the corresponding tRNA is relatively rare in the cells under consideration. The increased usage of rare codons triggers the consumption of the restricted pool of charged tRNAs, stalling of the uncharged ribosomes during elongation and mRNA degradation in a DDX6-dependent manner (Radhakrishnan et al., 2016). The findings described in this thesis indicate that the function of DDX6 in co-translational mRNA is broader and significant for general mRNA decay. This function is most likely associated with the speed of the ribosome during elongation, which, when altered, marks any transcript for decay.

In summary, these studies highlight the GIGYF1/2 proteins as important scaffolds for effector proteins such as 4EHP and DDX6, which induce translation repression and

deadenylation-dependent mRNA decay as a result of ribosome stalling. Recruitment of the 4EHP-GIGYF1/2 complexes to the mRNA is facilitated by recognition of ribosome collisions by ZNF598 or alternatively by co-translational binding of GIGYF1/2 and/or associated factors to the nascent peptide (Figure 9).
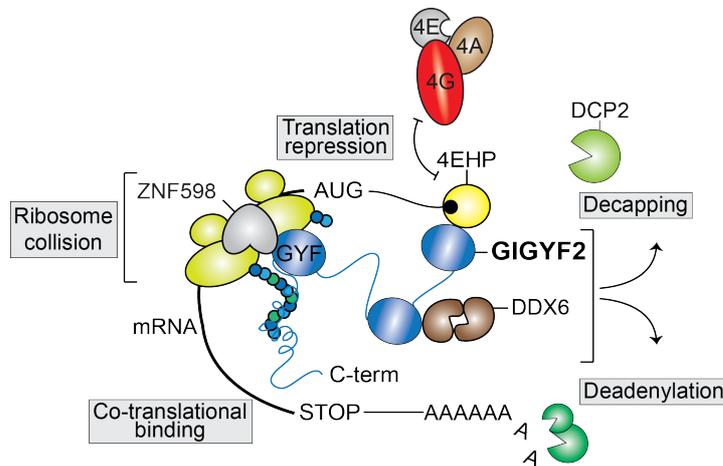

Figure 9: Schematic representation of 4EHP-GIGYF1/2 mediated co-translational mRNA decay. GIGYF1/2 proteins (GIGYF2 is depicted) are co-translationally recruited to the target mRNA by ZNF598, a ribosome collision sensor (Juszkiewicz et al., 2018). GIGYF2 acts as a scaffold that brings the alternative cap binding protein 4EHP to replace eIF4E in complex with eIF4G and eIF4A at the cap thereby inducing translation repression. In addition, the scaffold mediates interactions with the RNA helicase and decapping activator DDX6 and triggers mRNA deadenylation and decapping. The C-terminal proportion of GIGYF2 (C-term) interacts co-translationally with the nascent polypeptide chain.

The observation that GIGYF2 can bind to SPs raises the question if this interaction is mutually exclusive with SRP binding to SPs or if the interactions can occur at the same time. Mutually exclusive interactions with the SP could guarantee that in conditions that compromise SRP binding and ER targeting, GIGYF1/2 would trigger the clearance of the associated mRNA. Such mechanism would prevent the production, accumulation and misfolding of secretome-associated proteins in the cytoplasm. Concomitant binding of SRP and GIGYF1/2 to the SP would signify that target mRNA decay can occur at the ER surface to eliminate transcripts whose nascent peptide fails to co-translationally insert into the ER membrane or simply to reduce the half-life of specific secretome-associated mRNAs and buffer excess protein synthesis. Furthermore, it still remains to be elucidated if GIGYF1/2 proteins have a preference towards specific SPs. If so, it will be important to understand the molecular features subjacent to SP specificity and its regulation under conditions that shape the cellular secretome.

8.6. GIGYF1/2 and ER-associated protein/mRNA quality control in neurodegenerative diseases

Although the precise mechanism is so far unknown, mutations in, loss or haploinsufficiency of GIGYF1/2 are associated with neurodegeneration and neurological disorders like autism and schizophrenia in animal models and affected humans (Giovannone et al., 2009; Iossifov et al., 2014; Krumm et al., 2015; Satterstrom et al., 2020; Schizophrenia Working Group of the Psychiatric Genomics, 2014; Thyme et al., 2019). Interestingly, increased ER stress is a hallmark of many neurodegenerative conditions (Matus et al., 2011). In addition, TUBULIN stability and dynamics are associated with neurodegeneration (Dubey et al., 2015) and mutations in the autoregulatory domain of TUBB4A that interferes with co-translational decay have been described in hereditary dystonia (Hersheson et al., 2013). As regulators of secretome-associated and *TUBULIN* mRNAs, GIGYF1/2 proteins safeguard the integrity of the cellular proteome. By inducing the decay of mRNAs with altered ribosome elongation dynamics, GIGYF1/2 prevent the synthesis of potentially cytotoxic proteins. Altered or compromised GIGYF1/2 function may lead to the accumulation of these unwanted proteins and contribute to disease onset in tissues such as the brain. Additional studies will elucidate if under cellular conditions such as ER stress, GIGYF1/2-dependent mRNA decay of secretome-associated mRNAs has an important role in reducing the accumulation of cytotoxic proteins.

Impaired TUBULIN autoregulation has been linked with defects in cell division (Lin et al., 2020). Further studies will be required to investigate if GIGYF1/2-null cells show defects in cell division, but their contribution to TUBULIN autoregulation suggests that GIGYF1/2 proteins might also safeguard the process of cell division.

Albeit not discussed in the unpublished manuscript (B: Weber et al., unpublished), another group of mRNAs regulated by GIGYF1/2 contain CAG repeats encoding poly-glutamine (Q) stretches in proteins known to have poly-Q length polymorphisms (e.g. *MAGI1*, *NCOR2*, *AR*) (Butland et al., 2007), some of them even associated with cancer and spinal bulbar muscular atrophy (SBMA) as in the case of the androgen receptor (*AR*) transcript (Fan et al., 2014; Sartor et al., 1999). This observation indicates that in the absence of 4EHP and GIGYF1/2, degradation of specific mRNAs containing CAG repeats is inhibited. As a consequence, proteins with poly-Q expansions accumulate in cells with compromised 4EHP and GIGYF1/2 function. Poly-Q stretches are required to establish important functional protein-protein interactions in cells; the expansion of poly-Q stretches however leads to abnormal aggregation of poly-Q proteins and formation of macroscopically visible plaques in tissues such as the brain

(Edwards, 2019). Dysregulation of 4EHP and GIGYF1/2 might therefore result in the accumulation of another type of potentially cytotoxic proteins.

To investigate if the 4EHP-GIGYF1/2 complexes regulate the turnover of mRNAs containing CAG repeats, control and GIGYF1/2 KO cells were transfected with firefly luciferase (F-LUC) reporters containing or lacking a C-terminal extension of 23 Q residues (Figure 10A). F-LUC-23Q activity, but not F-LUC-STOP-23Q or F-LUC-GFP, was increased in GIGYF1/2-null cells compared to control cells. Additionally, re-expression of 4EHP and GIGYF2 in the null cells, reduced F-LUC-23Q activity (Figure 10B). Importantly, increased luciferase activity was explained by increased *F-LUC-23Q* mRNA levels (Figure 10C,D). These results demonstrate that the 4EHP-GIGYF1/2 complexes mediate decay of mRNAs containing CAG repeats. Moreover, mRNA decay is intrinsically associated with translation of the 23 consecutive CAG codons, suggesting again that GIGYF1/2 functions in ribosome-associated surveillance mechanisms. In fact, pathogenic expansions of CAG repeats result in aberrant translation as they cause frequent ribosome frameshifting events (Saffert et al., 2016), altered ribosome dynamics and non-optimal translation from a different reading frame.

Altered ribosome activity in the mRNAs containing CAG repeats targeted by 4EHP and GIGYF1/2 could mark these mRNAs for degradation. However, the repetitive nature of the CAG stretches makes the analysis of our short-read ribosome profiling data unable to pinpoint if ribosome stalling occurs in these mRNAs. Problems in translation elongation could potentially arise from local depletion of the cognate tRNA or interactions of the hydrophobic poly-Q stretch with the ribosomal exit tunnel that lead to ribosome stalling and collisions, which would mark these mRNAs for decay.
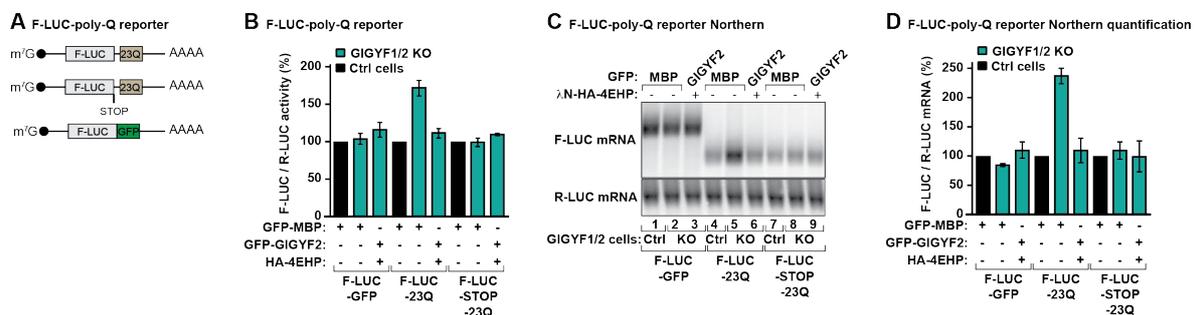
Figure 10: 4EHP-GIGYF1/2 mediate co-translational decay of CAG repeat mRNAs. (A) Schematic representation of the *F-LUC-23Q*, *F-LUC-STOP-23Q* and *F-LUC-GFP* control reporters; glutamine (Q). **(B)** Control (Ctrl) and GIGYF1/2 KO cells were transfected with F-LUC-GFP or F-LUC-23Q or F-LUC-STOP-23Q and R-LUC. F-LUC activity was determined in Ctrl and GIGYF1/2 KO cells, normalized over those of R-LUC and set to 100% in Ctrl cells. **(C)** RNA samples corresponding to (B) were analyzed by northern blotting using probes specific against *F-LUC* and *R-LUC*. **(D)** Quantification of the RNA levels in (C); *F-LUC* levels were normalized to those of *R-LUC* and set to 100% in Ctrl cells.

Interestingly, GIGYF2 proteins itself contain stretches enriched in Q residues and can be classified as a poly-Q protein (Li et al., 2016). This suggests a possible auto-regulatory mechanism for GIGYF1/2 proteins.

Since to date literature does not list studies that address the stability of mRNAs containing CAG repeats in disease, it will be of utmost importance to study how the 4EHP-GIGYF1/2 complexes recognize and induce the decay of CAG repeat mRNAs. Such studies will give important insights into how a cellular decay complex is involved in inducing the turnover of mRNAs containing pathological repeat expansions, and contribute to the understanding and treatment of the associated diseases.

9. Conclusions

Recent advances in understanding control of gene expression in mammalian cells highlight that these processes are more complex than previously anticipated. Control of mRNA transcription was long considered to be central for determining protein abundance, which ultimately defines cellular fate and identity. However, post-transcriptional processes controlling mRNA stability by 3´ UTR-associated mechanisms came more into focus and reveal additional layers of gene regulation. Together with others, the studies described throughout this thesis expand our understanding of the complexity of mammalian gene expression, and highlight the role of the mRNA 5´ UTR/5´ leader and CDS in the control of gene expression.

9.1. 5´ leader-associated translational control in mammals

One part of my doctoral work focused on the functional characterization of the eIF4G homologous protein DAP5 involved in specific processes during translation initiation. My results demonstrate that DAP5 is important for translation initiation on a small cohort of cellular mRNAs encoding signaling proteins that collectively contain small translated uORFs upstream of the mORF. In this context, I was able to show that the 5´ leader sequences of the identified transcripts were sufficient to promote translation in a DAP5-dependent manner. Notably, my experiments show for the first time that DAP5 mediates re-initiation after uORF translation. Re-initiation of translation by DAP5 requires cap-dependent translation of uORFs, eIF4A and conventional translation termination after uORF translation. These results demonstrate how complex mammalian mRNA 5´ leaders buffer the translational output of main ORF proteins. Dysfunctions in DAP5, and thus re-initiation, lead to defects in development and are associated with tumor formation and metastasis (de la Parra et al., 2017; Yamanaka et al., 2000). In conclusion, my studies not only advance our mechanistic understanding of the complexity of mammalian translation initiation but also inform future studies to develop novel cancer therapeutics.

9.2. mRNA turnover as a consequence of perturbed translation

In the other part of my PhD work, I aimed at elucidating the molecular functions of the 4EHP-GIGYF1/2 repressive complexes. Here, the results showed that the predominant cellular function of 4EHP-GIGYF1/2 complexes is the decay of specific mRNA targets, many

encoding secreted and membrane-bound proteins. My results further showed that target mRNA decay was determined by the CDS and was associated with altered ribosome activity, such as pausing and collision. Aberrant translation elongation was recognized by RBPs like ZNF598 that recruit GIGYF1/2, which serve as a hub for recruitment of other factors involved in mRNA repression and decay, like 4EHP, DDX6 and the CCR4-NOT deadenylase complex. Efforts to elucidate why so many of the 4EHP-GIGYF1/2 targets showed ribosome stalling events revealed that specific N-terminal motifs in the corresponding nascent proteins (e.g.: SPs) might induce ribosome pausing and promote co-translational recruitment of GIGYF1/2 proteins. These observations linked GIGYF1/2 to co-translational quality control of ER-targeted and *TUBULIN* mRNAs, among others.

In light of the fact that mutations in or loss of GIGYF1/2 result in neurological disorders and neurodegeneration (Giovannone et al., 2009; Iossifov et al., 2014; Krumm et al., 2015; Satterstrom et al., 2020; Schizophrenia Working Group of the Psychiatric Genomics, 2014; Thyme et al., 2019), and that ER stress is often associated with these diseases (Matus et al., 2011), my findings rationalize for the first time why GIGYF1/2 are associated with disease. In conclusion, my studies elucidate the molecular mechanisms underlying the 4EHP-GIGYF1/2 mediated mRNA repression and decay, pave the way to understand the complexity of human neurological diseases, and help to advance preventive efforts taken to treat neurodegeneration.

10. References

Ahmed, Y.L., Schleich, S., Bohlen, J., Mandel, N., Simon, B., Sinning, I., and Teleman, A.A. (2018). DENR-MCTS1 heterodimerization and tRNA recruitment are required for translation reinitiation. PLoS Biol *16*, e2005160.

Amaya Ramirez, C.C., Hubbe, P., Mandel, N., and Bethune, J. (2018). 4EHP-independent repression of endogenous mRNAs by the RNA-binding protein GIGYF2. Nucleic acids research *46*, 5792-5808.

Amrani, N., Ganesan, R., Kervestin, S., Mangus, D.A., Ghosh, S., and Jacobson, A. (2004). A faux 3'-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. Nature *432*, 112-118.

Amrani, N., Sachs, M.S., and Jacobson, A. (2006). Early nonsense: mRNA decay solves a translational problem. Nature reviews. Molecular cell biology *7*, 415-425.

Andreev, D.E., O'Connor, P.B., Fahey, C., Kenny, E.M., Terenin, I.M., Dmitriev, S.E., Cormican, P., Morris, D.W., Shatsky, I.N., and Baranov, P.V. (2015). Translation of 5' leaders is pervasive in genes resistant to eIF2 repression. eLife *4*, e03971.

Arribere, J.A., and Fire, A.Z. (2018). Nonsense mRNA suppression via nonstop decay. eLife *7*, e33292.

Ash, M.R., Faelber, K., Kosslick, D., Albert, G.I., Roske, Y., Kofler, M., Schuemann, M., Krause, E., and Freund, C. (2010). Conserved beta-hairpin recognition by the GYF domains of Smy2 and GIGYF2 in mRNA surveillance and vesicular transport complexes. Structure *18*, 944-954.

Balchin, D., Hayer-Hartl, M., and Hartl, F.U. (2016). In vivo aspects of protein folding and quality control. Science *353*, aac4354.

Basson, M.A. (2012). Signaling in cell differentiation and morphogenesis. Cold Spring Harb Perspect Biol *4*.

Ben-Shem, A., Jenner, L., Yusupova, G., and Yusupov, M. (2010). Crystal structure of the eukaryotic ribosome. Science *330*, 1203-1209.

Bertram, G., Bell, H.A., Ritchie, D.W., Fullerton, G., and Stansfield, I. (2000). Terminating eukaryote translation: domain 1 of release factor eRF1 functions in stop codon recognition. Rna *6*, 1236-1247.

Bhandari, D., Raisch, T., Weichenrieder, O., Jonas, S., and Izaurralde, E. (2014). Structural basis for the Nanos-mediated recruitment of the CCR4-NOT complex and translational repression. Genes & development *28*, 888-901.

Blais, J.D., Filipenko, V., Bi, M., Harding, H.P., Ron, D., Koumenis, C., Wouters, B.G., and Bell, J.C. (2004). Activating transcription factor 4 is translationally regulated by hypoxic stress. Mol Cell Biol *24*, 7469-7482.

Blanchet, S., Rowe, M., Von der Haar, T., Fabret, C., Demais, S., Howard, M.J., and Namy, O. (2015). New insights into stop codon recognition by eRF1. Nucleic acids research *43*, 3298-3308.

Bohlen, J., Fenzl, K., Kramer, G., Bukau, B., and Teleman, A.A. (2020). Selective 40S Footprinting Reveals Cap-Tethered Ribosome Scanning in Human Cells. Molecular cell *79*, 561-574 e5.

Bulygin, K.N., Khairulina, Y.S., Kolosov, P.M., Ven'yaminova, A.G., Graifer, D.M., Vorobjev, Y.N., Frolova, L.Y., and Karpova, G.G. (2011). Adenine and guanine recognition of stop codon is mediated by different N domain conformations of translation termination factor eRF1. Nucleic acids research *39*, 7134-7146.

Buschauer, R., Matsuo, Y., Sugiyama, T., Chen, Y.H., Alhusaini, N., Sweet, T., Ikeuchi, K., Cheng, J., Matsuki, Y., Nobuta, R.*, et al.* (2020). The Ccr4-Not complex monitors the translating ribosome for codon optimality. Science *368*, eaay6912.

Buskirk, A.R., and Green, R. (2017). Ribosome pausing, arrest and rescue in bacteria and eukaryotes. Philos Trans R Soc Lond B Biol Sci *372*.

Butland, S.L., Devon, R.S., Huang, Y., Mead, C.L., Meynert, A.M., Neal, S.J., Lee, S.S., Wilkinson, A., Yang, G.S., Yuen, M.M.*, et al.* (2007). CAG-encoded polyglutamine length polymorphism in the human genome. BMC Genomics *8*, 126.

Calvo, S.E., Pagliarini, D.J., and Mootha, V.K. (2009). Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. Proceedings of the National Academy of Sciences of the United States of America *106*, 7507-7512.

Chandrasekaran, V., Juszkiewicz, S., Choi, J., Puglisi, J.D., Brown, A., Shao, S., Ramakrishnan, V., and Hegde, R.S. (2019). Mechanism of ribosome stalling during translation of a poly(A) tail. Nature structural & molecular biology *26*, 1132-1140.

Chang, C.T., Muthukumar, S., Weber, R., Levdansky, Y., Chen, Y., Bhandari, D., Igreja, C., Wohlbold, L., Valkov, E., and Izaurralde, E. (2019). A low-complexity region in human XRN1 directly recruits deadenylation and decapping factors in 5'-3' messenger RNA decay. Nucleic acids research *47*, 9282-9295.

Chapat, C., Jafarnejad, S.M., Matta-Camacho, E., Hesketh, G.G., Gelbart, I.A., Attig, J., Gkogkas, C.G., Alain, T., Stern-Ginossar, N., Fabian, M.R.*, et al.* (2017). Cap-binding protein 4EHP effects translation silencing by microRNAs. Proc Natl Acad Sci U S A *114*, 5425-5430.

References

Chen, C.Y., and Shyu, A.B. (2003). Rapid deadenylation triggered by a nonsense codon precedes decay of the RNA body in a mammalian cytoplasmic nonsense-mediated decay pathway. Mol Cell Biol *23*, 4805-4813.

Chen, C.Y., and Shyu, A.B. (2011). Mechanisms of deadenylation-dependent decay. Wiley interdisciplinary reviews. RNA *2*, 167-183.

Chen, J., Brunner, A.D., Cogan, J.Z., Nunez, J.K., Fields, A.P., Adamson, B., Itzhak, D.N., Li, J.Y., Mann, M., Leonetti, M.D.*, et al.* (2020). Pervasive functional translation of noncanonical human open reading frames. Science *367*, 1140-1146.

Chen, X., Li, A., Sun, B.F., Yang, Y., Han, Y.N., Yuan, X., Chen, R.X., Wei, W.S., Liu, Y., Gao, C.C.*, et al.* (2019). 5-methylcytosine promotes pathogenesis of bladder cancer through stabilizing mRNAs. Nat Cell Biol *21*, 978-990.

Cheung, Y.N., Maag, D., Mitchell, S.F., Fekete, C.A., Algire, M.A., Takacs, J.E., Shirokikh, N., Pestova, T., Lorsch, J.R., and Hinnebusch, A.G. (2007). Dissociation of eIF1 from the 40S ribosomal subunit is a key step in start codon selection in vivo. Genes & development *21*, 1217-1230.

Chicoine, J., Benoit, P., Gamberi, C., Paliouras, M., Simonelig, M., and Lasko, P. (2007). Bicaudal-C recruits CCR4-NOT deadenylase to target mRNAs and regulates oogenesis, cytoskeletal organization, and its own expression. Developmental cell *13*, 691-704.

Cho, P.F., Gamberi, C., Cho-Park, Y.A., Cho-Park, I.B., Lasko, P., and Sonenberg, N. (2006). Cap-dependent translational inhibition establishes two opposing morphogen gradients in Drosophila embryos. Current biology : CB *16*, 2035-2041.

Cho, P.F., Poulin, F., Cho-Park, Y.A., Cho-Park, I.B., Chicoine, J.D., Lasko, P., and Sonenberg, N. (2005). A new paradigm for translational control: inhibition via 5'-3' mRNA tethering by Bicoid and the eIF4E cognate 4EHP. Cell *121*, 411-423.

Christensen, A.K., Kahn, L.E., and Bourne, C.M. (1987). Circular polysomes predominate on the rough endoplasmic reticulum of somatotropes and mammotropes in the rat anterior pituitary. Am J Anat *178*, 1-10.

Cleveland, D.W., Lopata, M.A., Sherline, P., and Kirschner, M.W. (1981). Unpolymerized tubulin modulates the level of tubulin mRNAs. Cell *25*, 537-546.

Coldwell, M.J., and Morley, S.J. (2006). Specific isoforms of translation initiation factor 4GI show differences in translational activity. Molecular and cellular biology *26*, 8448-8460.

Coldwell, M.J., Sack, U., Cowan, J.L., Barrett, R.M., Vlasak, M., Sivakumaran, K., and Morley, S.J. (2012). Multiple isoforms of the translation initiation factor eIF4GII are generated via use of alternative promoters, splice sites and a non-canonical initiation codon. The Biochemical journal *448*, 1-11.

Collart, M.A. (2016). The Ccr4-Not complex is a key regulator of eukaryotic gene expression. Wiley interdisciplinary reviews. RNA *7*, 438-454.

D'Orazio, K.N., Wu, C.C., Sinha, N., Loll-Krippleber, R., Brown, G.W., and Green, R. (2019). The endonuclease Cue2 cleaves mRNAs at stalled ribosomes during No Go Decay. eLife *8*, e49117.

de la Parra, C., Alard, A., Ernlund, A., and Schneider, R.J. (2017). Translation initiation factor DAP5 plays an essential role in translational control of breast cancer metastasis. Proceedings of the American Association for Cancer Research Annual Meeting 2017; Abstract nr 4481.

de la Parra, C., Ernlund, A., Alard, A., Ruggles, K., Ueberheide, B., and Schneider, R.J. (2018). A widespread alternate form of cap-dependent mRNA translation initiation. Nat Commun *9*, 3068.

Dever, T.E., and Green, R. (2012). The elongation, termination, and recycling phases of translation in eukaryotes. Cold Spring Harb Perspect Biol *4*, a013706.

Dever, T.E., and Ivanov, I.P. (2018). Roles of polyamines in translation. J Biol Chem *293*, 18719-18729.

Dmitriev, S.E., Terenin, I.M., Andreev, D.E., Ivanov, P.A., Dunaevsky, J.E., Merrick, W.C., and Shatsky, I.N. (2010). GTP-independent tRNA delivery to the ribosomal P-site by a novel eukaryotic translation factor. The Journal of biological chemistry *285*, 26779-26787.

Doma, M.K., and Parker, R. (2006). Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. Nature *440*, 561-564.

Dubey, J., Ratnakaran, N., and Koushika, S.P. (2015). Neurodegeneration and microtubule dynamics: death by a thousand cuts. Front Cell Neurosci *9*, 343.

Edwards, F.A. (2019). A Unifying Hypothesis for Alzheimer's Disease: From Plaques to Neurodegeneration. Trends Neurosci *42*, 310-322.

Eswarappa, S.M., Potdar, A.A., Koch, W.J., Fan, Y., Vasu, K., Lindner, D., Willard, B., Graham, L.M., DiCorleto, P.E., and Fox, P.L. (2014). Programmed translational readthrough generates antiangiogenic VEGF-Ax. Cell *157*, 1605-1618.

Fan, H.C., Ho, L.I., Chi, C.S., Chen, S.J., Peng, G.S., Chan, T.M., Lin, S.Z., and Harn, H.J. (2014). Polyglutamine (PolyQ) diseases: genetics to treatments. Cell Transplant *23*, 441-458.

Fan, Y., Evans, C.R., and Ling, J. (2017). Rewiring protein synthesis: From natural to synthetic amino acids. Biochim Biophys Acta Gen Subj *1861*, 3024-3029.

Feng, S., Sekine, S., Pessino, V., Li, H., Leonetti, M.D., and Huang, B. (2017). Improved split fluorescent proteins for endogenous protein labeling. Nat Commun *8*, 370.

Firth, A.E., and Brierley, I. (2012). Non-canonical translation in RNA viruses. J Gen Virol *93*, 1385-1409.

Frank, J., Gao, H., Sengupta, J., Gao, N., and Taylor, D.J. (2007). The process of mRNA-tRNA translocation. Proceedings of the National Academy of Sciences of the United States of America *104*, 19671-19678.

Friedel, C.C., Dolken, L., Ruzsics, Z., Koszinowski, U.H., and Zimmer, R. (2009). Conserved principles of mammalian transcriptional regulation revealed by RNA half-life. Nucleic acids research *37*, e115.

Frischmeyer, P.A., van Hoof, A., O'Donnell, K., Guerrerio, A.L., Parker, R., and Dietz, H.C. (2002). An mRNA surveillance mechanism that eliminates transcripts lacking termination codons. Science *295*, 2258-2261.

Frolova, L.Y., Tsivkovskii, R.Y., Sivolobova, G.F., Oparina, N.Y., Serpinsky, O.I., Blinov, V.M., Tatkov, S.I., and Kisselev, L.L. (1999). Mutations in the highly conserved GGQ motif of class 1 polypeptide release factors abolish ability of human eRF1 to trigger peptidyl-tRNA hydrolysis. Rna *5*, 1014-1020.

Fu, R., Olsen, M.T., Webb, K., Bennett, E.J., and Lykke-Andersen, J. (2016). Recruitment of the 4EHP-GYF2 cap-binding complex to tetraproline motifs of tristetraprolin promotes repression and degradation of mRNAs with AU-rich elements. RNA *22*, 373-382.

Gallie, D.R. (1991). The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. Genes & development *5*, 2108-2116.

Gao, X., Wan, J., Liu, B., Ma, M., Shen, B., and Qian, S.B. (2015). Quantitative profiling of initiating ribosomes in vivo. Nat Methods *12*, 147-153.

Gardin, J., Yeasmin, R., Yurovsky, A., Cai, Y., Skiena, S., and Futcher, B. (2014). Measurement of average decoding rates of the 61 sense codons in vivo. eLife *3*.

Garreau de Loubresse, N., Prokhorova, I., Holtkamp, W., Rodnina, M.V., Yusupova, G., and Yusupov, M. (2014). Structural basis for the inhibition of the eukaryotic ribosome. Nature *513*, 517-522.

Gasic, I., Boswell, S.A., and Mitchison, T.J. (2019). Tubulin mRNA stability is sensitive to change in microtubule dynamics caused by multiple physiological and toxic cues. PLoS Biol *17*, e3000225.

Gay, D.A., Sisodia, S.S., and Cleveland, D.W. (1989). Autoregulatory control of beta-tubulin mRNA stability is linked to translation elongation. Proceedings of the National Academy of Sciences of the United States of America *86*, 5763-5767.

Gilbert, W.V. (2010). Alternative ways to think about cellular internal ribosome entry. J Biol Chem *285*, 29033-29038.

Giovannone, B., Tsiaras, W.G., de la Monte, S., Klysik, J., Lautier, C., Karashchuk, G., Goldwurm, S., and Smith, R.J. (2009). GIGYF2 gene disruption in mice results in neurodegeneration and altered insulin-like growth factor signaling. Human molecular genetics *18*, 4629-4639.

Gradi, A., Imataka, H., Svitkin, Y.V., Rom, E., Raught, B., Morino, S., and Sonenberg, N. (1998). A novel functional human eukaryotic translation initiation factor 4G. Molecular and cellular biology *18*, 334-342.

Guenther, U.P., Weinberg, D.E., Zubradt, M.M., Tedeschi, F.A., Stawicki, B.N., Zagore, L.L., Brar, G.A., Licatalosi, D.D., Bartel, D.P., Weissman, J.S.*, et al.* (2018). The helicase Ded1p controls use of near-cognate translation initiation codons in 5' UTRs. Nature *559*, 130-134.

Gunisova, S., Hronova, V., Mohammad, M.P., Hinnebusch, A.G., and Valasek, L.S. (2018). Please do not recycle! Translation reinitiation in microbes and higher eukaryotes. FEMS Microbiol Rev *42*, 165-192.

Guydosh, N.R., and Green, R. (2014). Dom34 rescues ribosomes in 3' untranslated regions. Cell *156*, 950-962.

Guydosh, N.R., and Green, R. (2017). Translation of poly(A) tails leads to precise mRNA cleavage. RNA *23*, 749-761.

Guydosh, N.R., Kimmig, P., Walter, P., and Green, R. (2017). Regulated Ire1-dependent mRNA decay requires no-go mRNA degradation to maintain endoplasmic reticulum homeostasis in S. pombe. eLife *6*, e29216.

Han, P., Shichino, Y., Schneider-Poetsch, T., Mito, M., Hashimoto, S., Udagawa, T., Kohno, K., Yoshida, M., Mishima, Y., Inada, T.*, et al.* (2020). Genome-wide Survey of Ribosome Collision. Cell reports *31*, 107610.

Hanson, G., and Coller, J. (2018). Codon optimality, bias and usage in translation and mRNA decay. Nature reviews. Molecular cell biology *19*, 20-30.

Hantsche, M., and Cramer, P. (2016). The Structural Basis of Transcription: 10 Years After the Nobel Prize in Chemistry. Angew Chem Int Ed Engl *55*, 15972-15981.

Hao, S., and Baltimore, D. (2009). The stability of mRNA influences the temporal order of the induction of genes encoding inflammatory molecules. Nat Immunol *10*, 281-288.

Haussmann, I.U., Bodi, Z., Sanchez-Moran, E., Mongan, N.P., Archer, N., Fray, R.G., and Soller, M. (2016). m(6)A potentiates Sxl alternative pre-mRNA splicing for robust Drosophila sex determination. Nature *540*, 301-304.

Heaphy, S.M., Mariotti, M., Gladyshev, V.N., Atkins, J.F., and Baranov, P.V. (2016). Novel Ciliate Genetic Code Variants Including the Reassignment of All Three Stop Codons to Sense Codons in Condylostoma magnum. Mol Biol Evol *33*, 2885-2889.

Henis-Korenblit, S., Strumpf, N.L., Goldstaub, D., and Kimchi, A. (2000). A novel form of DAP5 protein accumulates in apoptotic cells as a result of caspase cleavage and internal ribosome entry site-mediated translation. Molecular and cellular biology *20*, 496-506.

Hernandez, G., Altmann, M., Sierra, J.M., Urlaub, H., Diez del Corral, R., Schwartz, P., and Rivera-Pomar, R. (2005). Functional analysis of seven genes encoding eight translation initiation factor 4E (eIF4E) isoforms in Drosophila. Mech Dev *122*, 529-543.

Hersheson, J., Mencacci, N.E., Davis, M., MacDonald, N., Trabzuni, D., Ryten, M., Pittman, A., Paudel, R., Kara, E., Fawcett, K.*, et al.* (2013). Mutations in the autoregulatory domain of beta-tubulin 4a cause hereditary dystonia. Ann Neurol *73*, 546-553.

Hetz, C. (2012). The unfolded protein response: controlling cell fate decisions under ER stress and beyond. Nature reviews. Molecular cell biology *13*, 89-102.

Heurgue-Hamard, V., Champ, S., Mora, L., Merkulova-Rainon, T., Kisselev, L.L., and Buckingham, R.H. (2005). The glutamine residue of the conserved GGQ motif in Saccharomyces cerevisiae release factor eRF1 is methylated by the product of the YDR140w gene. The Journal of biological chemistry *280*, 2439-2445.

Hickey, K.L., Dickson, K., Cogan, J.Z., Replogle, J.M., Schoof, M., D'Orazio, K.N., Sinha, N.K., Hussmann, J.A., Jost, M., Frost, A.*, et al.* (2020). GIGYF2 and 4EHP Inhibit Translation Initiation of Defective Messenger RNAs to Assist Ribosome-Associated Quality Control. Mol Cell *79*, 950-962 e6.

Hinnebusch, A.G. (2005). Translational regulation of GCN4 and the general amino acid control of yeast. Annu Rev Microbiol *59*, 407-450.

Hinnebusch, A.G. (2014). The scanning mechanism of eukaryotic translation initiation. Annual review of biochemistry *83*, 779-812.

Hinnebusch, A.G., Ivanov, I.P., and Sonenberg, N. (2016). Translational control by 5'-untranslated regions of eukaryotic mRNAs. Science *352*, 1413-1416.

Hollien, J., and Weissman, J.S. (2006). Decay of endoplasmic reticulum-localized mRNAs during the unfolded protein response. Science *313*, 104-107.

Hu, J., Sun, F., and Handel, M.A. (2018). Nuclear localization of EIF4G3 suggests a role for the XY body in translational regulation during spermatogenesis in mice. Biol Reprod *98*, 102-114.

Hu, W., Sweet, T.J., Chamnongpol, S., Baker, K.E., and Coller, J. (2009). Co-translational mRNA decay in Saccharomyces cerevisiae. Nature *461*, 225-229.

Hussain, T., Llacer, J.L., Fernandez, I.S., Munoz, A., Martin-Marcos, P., Savva, C.G., Lorsch, J.R., Hinnebusch, A.G., and Ramakrishnan, V. (2014). Structural changes enable start codon recognition by the eukaryotic translation initiation complex. Cell *159*, 597-607.

Iacono, M., Mignone, F., and Pesole, G. (2005). uAUG and uORFs in human and rodent 5'untranslated mRNAs. Gene *349*, 97-105.

Imataka, H., Olsen, H.S., and Sonenberg, N. (1997). A new translational regulator with homology to eukaryotic translation initiation factor 4G. EMBO J *16*, 817-825.

Inada, T. (2020). Quality controls induced by aberrant translation. Nucleic Acids Res *48*, 1084-1096.

Ingolia, N.T., Lareau, L.F., and Weissman, J.S. (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. Cell *147*, 789-802.

Iossifov, I., O'Roak, B.J., Sanders, S.J., Ronemus, M., Krumm, N., Levy, D., Stessman, H.A., Witherspoon, K.T., Vives, L., Patterson, K.E.*, et al.* (2014). The contribution of de novo coding mutations to autism spectrum disorder. Nature *515*, 216-221.

Isken, O., and Maquat, L.E. (2007). Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function. Genes Dev *21*, 1833-1856.

Ivanov, I.P., Loughran, G., Sachs, M.S., and Atkins, J.F. (2010). Initiation context modulates autoregulation of eukaryotic translation initiation factor 1 (eIF1). Proceedings of the National Academy of Sciences of the United States of America *107*, 18056-18060.

Ivanov, I.P., Shin, B.S., Loughran, G., Tzani, I., Young-Baird, S.K., Cao, C., Atkins, J.F., and Dever, T.E. (2018). Polyamine Control of Translation Elongation Regulates Start Site Selection on Antizyme Inhibitor mRNA via Ribosome Queuing. Molecular cell *70*, 254-264 e256.

Iwasaki, S., Floor, S.N., and Ingolia, N.T. (2016). Rocaglates convert DEAD-box protein eIF4A into a sequence-selective translational repressor. Nature *534*, 558-561.

Jackson, R.J., Hellen, C.U., and Pestova, T.V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. Nat Rev Mol Cell Biol *11*, 113-127.

Jang, S.K., Krausslich, H.G., Nicklin, M.J., Duke, G.M., Palmenberg, A.C., and Wimmer, E. (1988). A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. J Virol *62*, 2636-2643.

References

Jeong, S.J., Park, S., Nguyen, L.T., Hwang, J., Lee, E.Y., Giong, H.K., Lee, J.S., Yoon, I., Lee, J.H., Kim, J.H., *et al.* (2019). A threonyl-tRNA synthetase-mediated translation initiation machinery. Nat Commun *10*, 1357.

Joazeiro, C.A.P. (2017). Ribosomal Stalling During Translation: Providing Substrates for Ribosome-Associated Protein Quality Control. Annu Rev Cell Dev Biol *33*, 343-368.

Joshi, B., Cameron, A., and Jagus, R. (2004). Characterization of mammalian eIF4E-family members. European journal of biochemistry / FEBS *271*, 2189-2203.

Juszkiewicz, S., Chandrasekaran, V., Lin, Z., Kraatz, S., Ramakrishnan, V., and Hegde, R.S. (2018). ZNF598 Is a Quality Control Sensor of Collided Ribosomes. Mol Cell *72*, 469-481 e467.

Juszkiewicz, S., Slodkowicz, G., Lin, Z., Freire-Pritchett, P., Peak-Chew, S.Y., and Hegde, R.S. (2020). Ribosome collisions trigger cis-acting feedback inhibition of translation initiation. Elife *9*, e60038.

Kamiyama, D., Sekine, S., Barsi-Rhyne, B., Hu, J., Chen, B., Gilbert, L.A., Ishikawa, H., Leonetti, M.D., Marshall, W.F., Weissman, J.S., *et al.* (2016). Versatile protein tagging in cells with split fluorescent protein. Nat Commun *7*, 11046.

Kang, S.W., Rane, N.S., Kim, S.J., Garrison, J.L., Taunton, J., and Hegde, R.S. (2006). Substrate-specific translocational attenuation during ER stress defines a pre-emptive quality control pathway. Cell *127*, 999-1013.

Karamyshev, A.L., Patrick, A.E., Karamysheva, Z.N., Griesemer, D.S., Hudson, H., Tjon-Kon-Sang, S., Nilsson, I., Otto, H., Liu, Q., Rospert, S., *et al.* (2014). Inefficient SRP interaction with a nascent chain triggers a mRNA quality control pathway. Cell *156*, 146-157.

Keskeny, C., Raisch, T., Sgromo, A., Igreja, C., Bhandari, D., Weichenrieder, O., and Izaurralde, E. (2019). A conserved CAF40-binding motif in metazoan NOT4 mediates association with the CCR4-NOT complex. Genes Dev *33*, 236-252.

Kimmig, P., Diaz, M., Zheng, J., Williams, C.C., Lang, A., Aragon, T., Li, H., and Walter, P. (2012). The unfolded protein response in fission yeast modulates stability of select mRNAs to maintain protein homeostasis. eLife *1*, e00048.

Kisselev, L., Ehrenberg, M., and Frolova, L. (2003). Termination of translation: interplay of mRNA, rRNAs and release factors? The EMBO journal *22*, 175-182.

Kofler, M.M., and Freund, C. (2006). The GYF domain. FEBS J *273*, 245-256.

Komar, A.A., and Hatzoglou, M. (2011). Cellular IRES-mediated translation: the war of ITAFs in pathophysiological states. Cell Cycle *10*, 229-240.

Kozak, M. (1987). An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic acids research *15*, 8125-8148.

Kozak, M. (2001). Constraints on reinitiation of translation in mammals. Nucleic Acids Res *29*, 5226-5232.

Kozak, M. (2005). A second look at cellular mRNA sequences said to function as internal ribosome entry sites. Nucleic Acids Res *33*, 6593-6602.

Krumm, N., Turner, T.N., Baker, C., Vives, L., Mohajeri, K., Witherspoon, K., Raja, A., Coe, B.P., Stessman, H.A., He, Z.X.*, et al.* (2015). Excess of rare, inherited truncating mutations in autism. Nature genetics *47*, 582-588.

Kryszke, M.H., Adjeriou, B., Liang, F., Chen, H., and Dautry, F. (2016). Post-transcriptional gene silencing activity of human GIGYF2. Biochem Biophys Res Commun *475*, 289-294.

Kryuchkova, P., Grishin, A., Eliseev, B., Karyagina, A., Frolova, L., and Alkalaeva, E. (2013). Two-step model of stop codon recognition by eukaryotic release factor eRF1. Nucleic acids research *41*, 4573-4586.

Kubacka, D., Kamenska, A., Broomhead, H., Minshall, N., Darzynkiewicz, E., and Standart, N. (2013). Investigating the consequences of eIF4E2 (4EHP) interaction with 4E-transporter on its cellular distribution in HeLa cells. PloS one *8*, e72761.

Kurosaki, T., and Maquat, L.E. (2016). Nonsense-mediated mRNA decay in humans at a glance. J Cell Sci *129*, 461-467.

Lakshminarayan, R., Phillips, B.P., Binnian, I.L., Gomez-Navarro, N., Escudero-Urquijo, N., Warren, A.J., and Miller, E.A. (2020). Pre-emptive Quality Control of a Misfolded Membrane Protein by Ribosome-Driven Effects. Current biology : CB *5*, 854-864 e5.

Landon, A.L., Muniandy, P.A., Shetty, A.C., Lehrmann, E., Volpon, L., Houng, S., Zhang, Y., Dai, B., Peroutka, R., Mazan-Mamczarz, K.*, et al.* (2014). MNKs act as a regulatory switch for eIF4E1 and eIF4E3 driven mRNA translation in DLBCL. Nature communications *5*, 5413.

Lee, S., Liu, B., Lee, S., Huang, S.X., Shen, B., and Qian, S.B. (2012). Global mapping of translation initiation sites in mammalian cells at single-nucleotide resolution. Proc Natl Acad Sci U S A *109*, E2424-2432.

Lee, S.H., and McCormick, F. (2006). p97/DAP5 is a ribosome-associated factor that facilitates protein synthesis and cell proliferation by modulating the synthesis of cell cycle proteins. The EMBO journal *25*, 4008-4019.

Lejeune, F., Li, X., and Maquat, L.E. (2003). Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities. Mol Cell *12*, 675-687.

References

Lence, T., Akhtar, J., Bayer, M., Schmid, K., Spindler, L., Ho, C.H., Kreim, N., Andrade-Navarro, M.A., Poeck, B., Helm, M.*, et al.* (2016). m(6)A modulates neuronal functions and sex determination in Drosophila. Nature *540*, 242-247.

Leonetti, M.D., Sekine, S., Kamiyama, D., Weissman, J.S., and Huang, B. (2016). A scalable strategy for high-throughput GFP tagging of endogenous human proteins. Proc Natl Acad Sci U S A *113*, E3501-3508.

Leppek, K., Schott, J., Reitter, S., Poetz, F., Hammond, M.C., and Stoecklin, G. (2013). Roquin promotes constitutive mRNA decay via a conserved class of stem-loop recognition motifs. Cell *153*, 869-881.

Levy-Strumpf, N., Deiss, L.P., Berissi, H., and Kimchi, A. (1997). DAP-5, a novel homolog of eukaryotic translation initiation factor 4G isolated as a putative modulator of gamma interferon-induced programmed cell death. Molecular and cellular biology *17*, 1615-1625.

Lewis, S.M., Cerquozzi, S., Graber, T.E., Ungureanu, N.H., Andrews, M., and Holcik, M. (2008). The eIF4G homolog DAP5/p97 supports the translation of select mRNAs during endoplasmic reticulum stress. Nucleic acids research *36*, 168-178.

Li, C., Nagel, J., Androulakis, S., Lupton, C.J., Song, J., and Buckle, A.M. (2016). PolyQ 2.0: an improved version of PolyQ, a database of human polyglutamine proteins. Database (Oxford) *2016*.

Li, X., Zhu, P., Ma, S., Song, J., Bai, J., Sun, F., and Yi, C. (2015). Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome. Nat Chem Biol *11*, 592-597.

Liang, H., He, S., Yang, J., Jia, X., Wang, P., Chen, X., Zhang, Z., Zou, X., McNutt, M.A., Shen, W.H.*, et al.* (2014). PTENalpha, a PTEN isoform translated through alternative initiation, regulates mitochondrial function and energy metabolism. Cell Metab *19*, 836-848.

Liberman, N., Gandin, V., Svitkin, Y.V., David, M., Virgili, G., Jaramillo, M., Holcik, M., Nagar, B., Kimchi, A., and Sonenberg, N. (2015). DAP5 associates with eIF2beta and eIF4AI to promote Internal Ribosome Entry Site driven translation. Nucleic acids research *43*, 3764-3775.

Liberman, N., Marash, L., and Kimchi, A. (2009). The translation initiation factor DAP5 is a regulator of cell survival during mitosis. Cell cycle *8*, 204-209.

Lin, Z., Gasic, I., Chandrasekaran, V., Peters, N., Shao, S., Mitchison, T.J., and Hegde, R.S. (2020). TTC5 mediates autoregulation of tubulin via mRNA degradation. Science *367*, 100-104.

Liu, J., Yue, Y., Han, D., Wang, X., Fu, Y., Zhang, L., Jia, G., Yu, M., Lu, Z., Deng, X.*, et al.* (2014). A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. Nat Chem Biol *10*, 93-95.

Llacer, J.L., Hussain, T., Marler, L., Aitken, C.E., Thakur, A., Lorsch, J.R., Hinnebusch, A.G., and Ramakrishnan, V. (2015). Conformational Differences between Open and Closed States of the Eukaryotic Translation Initiation Complex. Molecular cell *59*, 399-412.

Loh, B., Jonas, S., and Izaurralde, E. (2013). The SMG5-SMG7 heterodimer directly recruits the CCR4-NOT deadenylase complex to mRNAs containing nonsense codons via interaction with POP2. Genes Dev *27*, 2125-2138.

Lomakin, I.B., Stolboushkina, E.A., Vaidya, A.T., Zhao, C., Garber, M.B., Dmitriev, S.E., and Steitz, T.A. (2017). Crystal Structure of the Human Ribosome in Complex with DENR-MCT-1. Cell reports *20*, 521-528.

Loughran, G., Jungreis, I., Tzani, I., Power, M., Dmitriev, R.I., Ivanov, I.P., Kellis, M., and Atkins, J.F. (2018). Stop codon readthrough generates a C-terminally extended variant of the human vitamin D receptor with reduced calcitriol response. The Journal of biological chemistry *293*, 4434-4444.

Lykke-Andersen, S., Chen, Y., Ardal, B.R., Lilje, B., Waage, J., Sandelin, A., and Jensen, T.H. (2014). Human nonsense-mediated RNA decay initiates widely by endonucleolysis and targets snoRNA host genes. Genes & development *28*, 2498-2517.

Maag, D., Fekete, C.A., Gryczynski, Z., and Lorsch, J.R. (2005). A conformational change in the eukaryotic translation preinitiation complex and release of eIF1 signal recognition of the start codon. Molecular cell *17*, 265-275.

Mader, S., Lee, H., Pause, A., and Sonenberg, N. (1995). The translation initiation factor eIF-4E binds to a common motif shared by the translation factor eIF-4 gamma and the translational repressors 4E-binding proteins. Mol Cell Biol *15*, 4990-4997.

Marash, L., and Kimchi, A. (2005). DAP5 and IRES-mediated translation during programmed cell death. Cell death and differentiation *12*, 554-562.

Marash, L., Liberman, N., Henis-Korenblit, S., Sivan, G., Reem, E., Elroy-Stein, O., and Kimchi, A. (2008). DAP5 promotes cap-independent translation of Bcl-2 and CDK1 to facilitate cell survival during mitosis. Molecular cell *30*, 447-459.

Marcotrigiano, J., Gingras, A.C., Sonenberg, N., and Burley, S.K. (1999). Cap-dependent translation initiation in eukaryotes is regulated by a molecular mimic of eIF4G. Mol Cell *3*, 707-716.

Marzluff, W.F., Wagner, E.J., and Duronio, R.J. (2008). Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. Nat Rev Genet *9*, 843-854.

Matera, A.G., and Wang, Z. (2014). A day in the life of the spliceosome. Nat Rev Mol Cell Biol *15*, 108-121.

References

Matsui, M., Yachie, N., Okada, Y., Saito, R., and Tomita, M. (2007). Bioinformatic analysis of post-transcriptional regulation by uORF in human and mouse. FEBS letters *581*, 4184-4188.

Matus, S., Glimcher, L.H., and Hetz, C. (2011). Protein folding stress in neurodegenerative diseases: a glimpse into the ER. Curr Opin Cell Biol *23*, 239-252.

Meyer, K.D., Patil, D.P., Zhou, J., Zinoviev, A., Skabkin, M.A., Elemento, O., Pestova, T.V., Qian, S.B., and Jaffrey, S.R. (2015). 5' UTR m(6)A Promotes Cap-Independent Translation. Cell *163*, 999-1010.

Miao, H., Miao, C.X., Li, N., and Han, J. (2016). FOXJ2 controls meiosis during spermatogenesis in male mice. Mol Reprod Dev *83*, 684-691.

Miller, J.E., and Reese, J.C. (2012). Ccr4-Not complex: the control freak of eukaryotic cells. Crit Rev Biochem Mol Biol *47*, 315-333.

Mitchell, P. (2014). Exosome substrate targeting: the long and short of it. Biochemical Society transactions *42*, 1129-1134.

Morita, M., Ler, L.W., Fabian, M.R., Siddiqui, N., Mullin, M., Henderson, V.C., Alain, T., Fonseca, B.D., Karashchuk, G., Bennett, C.F.*, et al.* (2012). A novel 4EHP-GIGYF2 translational repressor complex is essential for mammalian development. Molecular and cellular biology *32*, 3585-3593.

Narula, A., Ellis, J., Taliaferro, J.M., and Rissland, O.S. (2019). Coding regions affect mRNA stability in human cells. Rna *25*, 1751-1764.

Nevins, T.A., Harder, Z.M., Korneluk, R.G., and Holcik, M. (2003). Distinct regulation of internal ribosome entry site-mediated translation following cellular stress is mediated by apoptotic fragments of eIF4G translation initiation factor family members eIF4GI and p97/DAP5/NAT1. The Journal of biological chemistry *278*, 3572-3579.

Nguyen, T.M., Kabotyanski, E.B., Dou, Y., Reineke, L.C., Zhang, P., Zhang, X.H., Malovannaya, A., Jung, S.Y., Mo, Q., Roarty, K.P.*, et al.* (2018). FGFR1-Activated Translation of WNT Pathway Components with Structured 5' UTRs Is Vulnerable to Inhibition of EIF4A-Dependent Translation Initiation. Cancer research *78*, 4229-4240.

Nyathi, Y., Wilkinson, B.M., and Pool, M.R. (2013). Co-translational targeting and translocation of proteins to the endoplasmic reticulum. Biochim Biophys Acta *1833*, 2392-2402.

Oslowski, C.M., and Urano, F. (2011). Measuring ER stress and the unfolded protein response using mammalian tissue culture system. Methods Enzymol *490*, 71-92.

Pachter, J.S., Yen, T.J., and Cleveland, D.W. (1987). Autoregulation of tubulin expression is achieved through specific degradation of polysomal tubulin mRNAs. Cell *51*, 283-292.

Pakos-Zebrucka, K., Koryga, I., Mnich, K., Ljujic, M., Samali, A., and Gorman, A.M. (2016). The integrated stress response. EMBO reports *17*, 1374-1395.

Pang, Y.L., Poruri, K., and Martinis, S.A. (2014). tRNA synthetase: tRNA aminoacylation and beyond. Wiley Interdiscip Rev RNA *5*, 461-480.

Passmore, L.A., Schmeing, T.M., Maag, D., Applefield, D.J., Acker, M.G., Algire, M.A., Lorsch, J.R., and Ramakrishnan, V. (2007). The eukaryotic translation initiation factors eIF1 and eIF1A induce an open conformation of the 40S ribosome. Molecular cell *26*, 41-50.

Pelechano, V., Wei, W., and Steinmetz, L.M. (2015). Widespread Co-translational RNA Decay Reveals Ribosome Dynamics. Cell *161*, 1400-1412.

Pelletier, J., Graff, J., Ruggero, D., and Sonenberg, N. (2015). Targeting the eIF4F translation initiation complex: a critical nexus for cancer development. Cancer research *75*, 250-263.

Pestova, T.V., Borukhov, S.I., and Hellen, C.U. (1998). Eukaryotic ribosomes require initiation factors 1 and 1A to locate initiation codons. Nature *394*, 854-859.

Pestova, T.V., and Kolupaeva, V.G. (2002). The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection. Genes & development *16*, 2906-2922.

Peter, D., Igreja, C., Weber, R., Wohlbold, L., Weiler, C., Ebertsch, L., Weichenrieder, O., and Izaurralde, E. (2015). Molecular Architecture of 4E-BP Translational Inhibitors Bound to eIF4E. Mol Cell *57*, 1074-1087.

Peter, D., Ruscica, V., Bawankar, P., Weber, R., Helms, S., Valkov, E., Igreja, C., and Izaurralde, E. (2019). Molecular basis for GIGYF-Me31B complex assembly in 4EHP-mediated translational repression. Genes Dev *33*, 1355-1360.

Peter, D., Weber, R., Sandmeir, F., Wohlbold, L., Helms, S., Bawankar, P., Valkov, E., Igreja, C., and Izaurralde, E. (2017). GIGYF1/2 proteins use auxiliary sequences to selectively bind to 4EHP and repress target mRNA expression. Genes Dev *31*, 1147-1161.

Pinarbasi, E.S., Karamyshev, A.L., Tikhonova, E.B., Wu, I.H., Hudson, H., and Thomas, P.J. (2018). Pathogenic Signal Sequence Mutations in Progranulin Disrupt SRP Interactions Required for mRNA Stability. Cell reports *23*, 2844-2851.

Ping, X.L., Sun, B.F., Wang, L., Xiao, W., Yang, X., Wang, W.J., Adhikari, S., Shi, Y., Lv, Y., Chen, Y.S.*, et al.* (2014). Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. Cell Res *24*, 177-189.

Pisarev, A.V., Hellen, C.U., and Pestova, T.V. (2007). Recycling of eukaryotic posttermination ribosomal complexes. Cell *131*, 286-299.

References

Pisarev, A.V., Skabkin, M.A., Pisareva, V.P., Skabkina, O.V., Rakotondrafara, A.M., Hentze, M.W., Hellen, C.U., and Pestova, T.V. (2010). The role of ABCE1 in eukaryotic posttermination ribosomal recycling. Molecular cell *37*, 196-210.

Pisareva, V.P., Skabkin, M.A., Hellen, C.U., Pestova, T.V., and Pisarev, A.V. (2011). Dissociation by Pelota, Hbs1 and ABCE1 of mammalian vacant 80S ribosomes and stalled elongation complexes. The EMBO journal *30*, 1804-1817.

Powell, M.L. (2010). Translational termination-reinitiation in RNA viruses. Biochem Soc Trans *38*, 1558-1564.

Preis, A., Heuer, A., Barrio-Garcia, C., Hauser, A., Eyler, D.E., Berninghausen, O., Green, R., Becker, T., and Beckmann, R. (2014). Cryoelectron microscopic structures of eukaryotic translation termination complexes containing eRF1-eRF3 or eRF1-ABCE1. Cell reports *8*, 59-65.

Presnyak, V., Alhusaini, N., Chen, Y.H., Martin, S., Morris, N., Kline, N., Olson, S., Weinberg, D., Baker, K.E., Graveley, B.R.*, et al.* (2015). Codon optimality is a major determinant of mRNA stability. Cell *160*, 1111-1124.

Pyronnet, S., Imataka, H., Gingras, A.C., Fukunaga, R., Hunter, T., and Sonenberg, N. (1999). Human eukaryotic translation initiation factor 4G (eIF4G) recruits mnk1 to phosphorylate eIF4E. The EMBO journal *18*, 270-279.

Rabl, J., Leibundgut, M., Ataide, S.F., Haag, A., and Ban, N. (2011). Crystal structure of the eukaryotic 40S ribosomal subunit in complex with initiation factor 1. Science *331*, 730-736.

Radhakrishnan, A., Chen, Y.H., Martin, S., Alhusaini, N., Green, R., and Coller, J. (2016). The DEAD-Box Protein Dhh1p Couples mRNA Decay and Translation by Monitoring Codon Optimality. Cell *167*, 122-132 e129.

Radhakrishnan, A., and Green, R. (2016). Connections Underlying Translation and mRNA Stability. Journal of molecular biology *428*, 3558-3564.

Raisch, T., Bhandari, D., Sabath, K., Helms, S., Valkov, E., Weichenrieder, O., and Izaurralde, E. (2016). Distinct modes of recruitment of the CCR4-NOT complex by Drosophila and vertebrate Nanos. The EMBO journal *35*, 974-990.

Ramakrishnan, V. (2014). The ribosome emerges from a black box. Cell *159*, 979-984.

Ramanathan, A., Robb, G.B., and Chan, S.H. (2016). mRNA capping: biological functions and applications. Nucleic Acids Res *44*, 7511-7526.

Ramirez-Valle, F., Braunstein, S., Zavadil, J., Formenti, S.C., and Schneider, R.J. (2008). eIF4GI links nutrient sensing by mTOR to cell proliferation and inhibition of autophagy. The Journal of cell biology *181*, 293-307.

Rasch, F., Weber, R., Izaurralde, E., and Igreja, C. (2020). 4E-T-bound mRNAs are stored in a silenced and deadenylated form. Genes & development *34*, 847-860.

Ries, R.J., Zaccara, S., Klein, P., Olarerin-George, A., Namkoong, S., Pickering, B.F., Patil, D.P., Kwak, H., Lee, J.H., and Jaffrey, S.R. (2019). m(6)A enhances the phase separation potential of mRNA. Nature *571*, 424-428.

Roberts, L.O., Jopling, C.L., Jackson, R.J., and Willis, A.E. (2009). Viral strategies to subvert the mammalian translation machinery. Prog Mol Biol Transl Sci *90*, 313-367.

Rom, E., Kim, H.C., Gingras, A.C., Marcotrigiano, J., Favre, D., Olsen, H., Burley, S.K., and Sonenberg, N. (1998). Cloning and characterization of 4EHP, a novel mammalian eIF4E-related cap-binding protein. J Biol Chem *273*, 13104-13109.

Roy, B., and Jacobson, A. (2013). The intimate relationships of mRNA decay and translation. Trends Genet *29*, 691-699.

Ruscica, V., Bawankar, P., Peter, D., Helms, S., Igreja, C., and Izaurralde, E. (2019). Direct role for the Drosophila GIGYF protein in 4EHP-mediated mRNA repression. Nucleic acids research *47*, 7035-7048.

Saffert, P., Adamla, F., Schieweck, R., Atkins, J.F., and Ignatova, Z. (2016). An Expanded CAG Repeat in Huntingtin Causes +1 Frameshifting. J Biol Chem *291*, 18505-18513.

Sartor, O., Zheng, Q., and Eastham, J.A. (1999). Androgen receptor gene CAG repeat length varies in a race-specific fashion in men without prostate cancer. Urology *53*, 378-380.

Sato, H., and Maquat, L.E. (2009). Remodeling of the pioneer translation initiation complex involves translation and the karyopherin importin beta. Genes Dev *23*, 2537-2550.

Satterstrom, F.K., Kosmicki, J.A., Wang, J., Breen, M.S., De Rubeis, S., An, J.Y., Peng, M., Collins, R., Grove, J., Klei, L.*, et al.* (2020). Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. Cell *180*, 568-584 e523.

Schizophrenia Working Group of the Psychiatric Genomics, C. (2014). Biological insights from 108 schizophrenia-associated genetic loci. Nature *511*, 421-427.

Schleich, S., Acevedo, J.M., Clemm von Hohenberg, K., and Teleman, A.A. (2017). Identification of transcripts with short stuORFs as targets for DENR*MCTS1-dependent translation in human cells. Sci Rep *7*, 3722.

Schleich, S., Strassburger, K., Janiesch, P.C., Koledachkina, T., Miller, K.K., Haneke, K., Cheng, Y.S., Kuechler, K., Stoecklin, G., Duncan, K.E.*, et al.* (2014). DENR-MCT-1 promotes translation re-initiation downstream of uORFs to control tissue growth. Nature *512*, 208-212.

References

Schmidt, S.A., Foley, P.L., Jeong, D.H., Rymarquis, L.A., Doyle, F., Tenenbaum, S.A., Belasco, J.G., and Green, P.J. (2015). Identification of SMG6 cleavage sites and a preferred RNA cleavage motif by global analysis of endogenous NMD targets in human cells. Nucleic acids research *43*, 309-323.

Schopp, I.M., Amaya Ramirez, C.C., Debeljak, J., Kreibich, E., Skribbe, M., Wild, K., and Bethune, J. (2017). Split-BioID a conditional proteomics approach to monitor the composition of spatiotemporally defined protein complexes. Nature communications *8*, 15690.

Schwanhausser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. Nature *473*, 337-342.

Seit-Nebi, A., Frolova, L., Justesen, J., and Kisselev, L. (2001). Class-1 translation termination factors: invariant GGQ minidomain is essential for release activity and ribosome binding but not for stop codon recognition. Nucleic acids research *29*, 3982-3987.

Sendoel, A., Dunn, J.G., Rodriguez, E.H., Naik, S., Gomez, N.C., Hurwitz, B., Levorse, J., Dill, B.D., Schramek, D., Molina, H.*, et al.* (2017). Translation from unconventional 5' start sites drives tumour initiation. Nature *541*, 494-499.

Sgromo, A., Raisch, T., Bawankar, P., Bhandari, D., Chen, Y., Kuzuoglu-Ozturk, D., Weichenrieder, O., and Izaurralde, E. (2017). A CAF40-binding motif facilitates recruitment of the CCR4-NOT complex to mRNAs targeted by Drosophila Roquin. Nature communications *8*, 14307.

Shao, S., Murray, J., Brown, A., Taunton, J., Ramakrishnan, V., and Hegde, R.S. (2016). Decoding Mammalian Ribosome-mRNA States by Translational GTPase Complexes. Cell *167*, 1229-1240 e1215.

Shaughnessy, J.D., Jr., Jenkins, N.A., and Copeland, N.G. (1997). cDNA cloning, expression analysis, and chromosomal localization of a gene with high homology to wheat eIF-(iso)4F and mammalian eIF-4G. Genomics *39*, 192-197.

Shoemaker, C.J., Eyler, D.E., and Green, R. (2010). Dom34:Hbs1 promotes subunit dissociation and peptidyl-tRNA drop-off to initiate no-go decay. Science *330*, 369-372.

Shoemaker, C.J., and Green, R. (2011). Kinetic analysis reveals the ordered coupling of translation termination and ribosome recycling in yeast. Proceedings of the National Academy of Sciences of the United States of America *108*, E1392-1398.

Shveygert, M., Kaiser, C., Bradrick, S.S., and Gromeier, M. (2010). Regulation of eukaryotic initiation factor 4E (eIF4E) phosphorylation by mitogen-activated protein kinase occurs through modulation of Mnk1-eIF4G interaction. Molecular and cellular biology *30*, 5160-5167.

Sidrauski, C., McGeachy, A.M., Ingolia, N.T., and Walter, P. (2015). The small molecule ISRIB reverses the effects of eIF2alpha phosphorylation on translation and stress granule assembly. eLife *4*, e05033.

Sinha, N.K., Ordureau, A., Best, K., Saba, J.A., Zinshteyn, B., Sundaramoorthy, E., Fulzele, A., Garshott, D.M., Denk, T., Thoms, M.*, et al.* (2020). EDF1 coordinates cellular responses to ribosome collisions. Elife *9*, e58828.

Skabkin, M.A., Skabkina, O.V., Dhote, V., Komar, A.A., Hellen, C.U., and Pestova, T.V. (2010). Activities of Ligatin and MCT-1/DENR in eukaryotic translation initiation and ribosomal recycling. Genes & development *24*, 1787-1801.

Skabkin, M.A., Skabkina, O.V., Hellen, C.U., and Pestova, T.V. (2013). Reinitiation and other unconventional posttermination events during eukaryotic translation. Mol Cell *51*, 249-264.

Slavoff, S.A., Mitchell, A.J., Schwaid, A.G., Cabili, M.N., Ma, J., Levin, J.Z., Karger, A.D., Budnik, B.A., Rinn, J.L., and Saghatelian, A. (2013). Peptidomic discovery of short open reading frame-encoded peptides in human cells. Nat Chem Biol *9*, 59-64.

Smirnova, V.V., Shestakova, E.D., Bikmetov, D.V., Chugunova, A.A., Osterman, I.A., Serebryakova, M.V., Sergeeva, O.V., Zatsepin, T.S., Shatsky, I.N., and Terenin, I.M. (2019). eIF4G2 balances its own mRNA translation via a PCBP2-based feedback loop. RNA *25*, 757-767.

Somers, J., Poyry, T., and Willis, A.E. (2013). A perspective on mammalian upstream open reading frame function. The international journal of biochemistry & cell biology *45*, 1690-1700.

Sonenberg, N., and Hinnebusch, A.G. (2009). Regulation of translation initiation in eukaryotes: mechanisms and biological targets. Cell *136*, 731-745.

Spitale, R.C., Flynn, R.A., Zhang, Q.C., Crisalli, P., Lee, B., Jung, J.W., Kuchelmeister, H.Y., Batista, P.J., Torre, E.A., Kool, E.T.*, et al.* (2015). Structural imprints in vivo decode RNA regulatory mechanisms. Nature *519*, 486-490.

Squires, J.E., Patel, H.R., Nousch, M., Sibbritt, T., Humphreys, D.T., Parker, B.J., Suter, C.M., and Preiss, T. (2012). Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. Nucleic acids research *40*, 5023-5033.

Starck, S.R., Jiang, V., Pavon-Eternod, M., Prasad, S., McCarthy, B., Pan, T., and Shastri, N. (2012). Leucine-tRNA initiates at CUG start codons for protein synthesis and presentation by MHC class I. Science *336*, 1719-1723.

Starck, S.R., Ow, Y., Jiang, V., Tokuyama, M., Rivera, M., Qi, X., Roberts, R.W., and Shastri, N. (2008). A distinct translation initiation mechanism generates cryptic peptides for immune surveillance. PloS one *3*, e3460.

References

Starck, S.R., Tsai, J.C., Chen, K., Shodiya, M., Wang, L., Yahiro, K., Martins-Green, M., Shastri, N., and Walter, P. (2016). Translation from the 5' untranslated region shapes the integrated stress response. Science *351*, aad3867.

Stutz, F., and Izaurralde, E. (2003). The interplay of nuclear mRNP assembly, mRNA surveillance and export. Trends in cell biology *13*, 319-327.

Sugiyama, H., Takahashi, K., Yamamoto, T., Iwasaki, M., Narita, M., Nakamura, M., Rand, T.A., Nakagawa, M., Watanabe, A., and Yamanaka, S. (2017). Nat1 promotes translation of specific proteins that induce differentiation of mouse embryonic stem cells. Proc Natl Acad Sci U S A *114*, 340-345.

Sun, F., Palmer, K., and Handel, M.A. (2010). Mutation of Eif4g3, encoding a eukaryotic translation initiation factor, causes male infertility and meiotic arrest of mouse spermatocytes. Development *137*, 1699-1707.

Swart, E.C., Serra, V., Petroni, G., and Nowacki, M. (2016). Genetic Codes with No Dedicated Stop Codon: Context-Dependent Translation Termination. Cell *166*, 691-702.

Swygert, S.G., and Peterson, C.L. (2014). Chromatin dynamics: interplay between remodeling enzymes and histone modifications. Biochim Biophys Acta *1839*, 728-736.

Takahashi, K., Maruyama, M., Tokuzawa, Y., Murakami, M., Oda, Y., Yoshikane, N., Makabe, K.W., Ichisaka, T., and Yamanaka, S. (2005). Evolutionarily conserved non-AUG translation initiation in NAT1/p97/DAP5 (EIF4G2). Genomics *85*, 360-371.

Tao, X., and Gao, G. (2015). Tristetraprolin Recruits Eukaryotic Initiation Factor 4E2 To Repress Translation of AU-Rich Element-Containing mRNAs. Molecular and cellular biology *35*, 3921-3932.

Taylor, D.J., Nilsson, J., Merrill, A.R., Andersen, G.R., Nissen, P., and Frank, J. (2007). Structures of modified eEF2 80S ribosome complexes reveal the role of GTP hydrolysis in translocation. The EMBO journal *26*, 2421-2431.

Teske, B.F., Wek, S.A., Bunpo, P., Cundiff, J.K., McClintick, J.N., Anthony, T.G., and Wek, R.C. (2011). The eIF2 kinase PERK and the integrated stress response facilitate activation of ATF6 during endoplasmic reticulum stress. Mol Biol Cell *22*, 4390-4405.

Thanaraj, T.A., and Argos, P. (1996). Ribosome-mediated translational pause and protein domain organization. Protein science : a publication of the Protein Society *5*, 1594-1612.

Thyme, S.B., Pieper, L.M., Li, E.H., Pandey, S., Wang, Y., Morris, N.S., Sha, C., Choi, J.W., Herrera, K.J., Soucy, E.R.*, et al.* (2019). Phenotypic Landscape of Schizophrenia-Associated Genes Defines Candidates and Their Shared Functions. Cell *177*, 478-491 e420.

Topisirovic, I., Svitkin, Y.V., Sonenberg, N., and Shatkin, A.J. (2011). Cap and cap-binding proteins in the control of gene expression. Wiley Interdiscip Rev RNA *2*, 277-298.

Tsuboi, T., Kuroha, K., Kudo, K., Makino, S., Inoue, E., Kashima, I., and Inada, T. (2012). Dom34:hbs1 plays a general role in quality-control systems by dissociation of a stalled ribosome at the 3' end of aberrant mRNA. Molecular cell *46*, 518-529.

Tsukiyama-Kohara, K., Iizuka, N., Kohara, M., and Nomoto, A. (1992). Internal ribosome entry site within hepatitis C virus RNA. J Virol *66*, 1476-1483.

Tuck, A.C., Rankova, A., Arpat, A.B., Liechti, L.A., Hess, D., Iesmantavicius, V., Castelo-Szekely, V., Gatfield, D., and Buhler, M. (2020). Mammalian RNA Decay Pathways Are Highly Specialized and Widely Linked to Translation. Mol Cell *77*, 1222-1236 e13.

Ueda, T., Watanabe-Fukunaga, R., Fukuyama, H., Nagata, S., and Fukunaga, R. (2004). Mnk2 and Mnk1 are essential for constitutive and inducible phosphorylation of eukaryotic initiation factor 4E but not for cell growth or development. Mol Cell Biol *24*, 6539-6549.

Uhlen, M., Fagerberg, L., Hallstrom, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, A., Kampf, C., Sjostedt, E., Asplund, A.*, et al.* (2015). Proteomics. Tissue-based map of the human proteome. Science *347*, 1260419.

Uttam, S., Wong, C., Price, T.J., and Khoutorsky, A. (2018). eIF4E-Dependent Translational Control: A Central Mechanism for Regulation of Pain Plasticity. Front Genet *9*, 470.

van Hoof, A., Frischmeyer, P.A., Dietz, H.C., and Parker, R. (2002). Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. Science *295*, 2262-2264.

Vicens, Q., Kieft, J.S., and Rissland, O.S. (2018). Revisiting the Closed-Loop Model and the Nature of mRNA 5'-3' Communication. Molecular cell *72*, 805-812.

Villaescusa, J.C., Buratti, C., Penkov, D., Mathiasen, L., Planaguma, J., Ferretti, E., and Blasi, F. (2009). Cytoplasmic Prep1 interacts with 4EHP inhibiting Hoxb4 translation. PloS one *4*, e5213.

Wahle, E., and Winkler, G.S. (2013). RNA decay machines: deadenylation by the Ccr4-not and Pan2-Pan3 complexes. Biochimica et biophysica acta *1829*, 561-570.

Walter, P., and Blobel, G. (1981). Translocation of proteins across the endoplasmic reticulum III. Signal recognition protein (SRP) causes signal sequence-dependent and site-specific arrest of chain elongation that is released by microsomal membranes. J Cell Biol *91*, 557-561.

Wang, X., Lu, Z., Gomez, A., Hon, G.C., Yue, Y., Han, D., Fu, Y., Parisien, M., Dai, Q., Jia, G.*, et al.* (2014). N6-methyladenosine-dependent regulation of messenger RNA stability. Nature *505*, 117-120.

References

Wang, Z., and Kiledjian, M. (2001). Functional link between the mammalian exosome and mRNA decapping. Cell *107*, 751-762.

Weill, L., Belloc, E., Bava, F.A., and Mendez, R. (2012). Translational control by changes in poly(A) tail length: recycling mRNAs. Nat Struct Mol Biol *19*, 577-585.

Weingarten-Gabbay, S., Khan, D., Liberman, N., Yoffe, Y., Bialik, S., Das, S., Oren, M., and Kimchi, A. (2014). The translation initiation factor DAP5 promotes IRES-driven translation of p53 mRNA. Oncogene *33*, 611-618.

Weisser, M., Schafer, T., Leibundgut, M., Bohringer, D., Aylett, C.H.S., and Ban, N. (2017). Structural and Functional Insights into Human Re-initiation Complexes. Molecular cell *67*, 447-456 e447.

Wells, S.E., Hillner, P.E., Vale, R.D., and Sachs, A.B. (1998). Circularization of mRNA by eukaryotic translation initiation factors. Molecular cell *2*, 135-140.

Wilusz, J.E., JnBaptiste, C.K., Lu, L.Y., Kuhn, C.D., Joshua-Tor, L., and Sharp, P.A. (2012). A triple helix stabilizes the 3' ends of long noncoding RNAs that lack poly(A) tails. Genes Dev *26*, 2392-2407.

Woodward, L.A., Mabin, J.W., Gangras, P., and Singh, G. (2017). The exon junction complex: a lifelong guardian of mRNA fate. Wiley Interdiscip Rev RNA *8*.

Wu, Q., Medina, S.G., Kushawah, G., DeVore, M.L., Castellano, L.A., Hand, J.M., Wright, M., and Bazzini, A.A. (2019). Translation affects mRNA stability in a codon-dependent manner in human cells. eLife *8*, e45396.

Yamaguchi, Y., Hayashi, A., Campagnoni, C.W., Kimura, A., Inuzuka, T., and Baba, H. (2012). L-MPZ, a novel isoform of myelin P0, is produced by stop codon readthrough. The Journal of biological chemistry *287*, 17765-17776.

Yamanaka, S., Poksay, K.S., Arnold, K.S., and Innerarity, T.L. (1997). A novel translational repressor mRNA is edited extensively in livers containing tumors caused by the transgene expression of the apoB mRNA-editing enzyme. Genes & development *11*, 321-333.

Yamanaka, S., Zhang, X.Y., Maeda, M., Miura, K., Wang, S., Farese, R.V., Jr., Iwao, H., and Innerarity, T.L. (2000). Essential role of NAT1/p97/DAP5 in embryonic differentiation and the retinoic acid pathway. EMBO J *19*, 5533-5541.

Yang, X., Yang, Y., Sun, B.F., Chen, Y.S., Xu, J.W., Lai, W.Y., Li, A., Wang, X., Bhattarai, D.P., Xiao, W.*, et al.* (2017). 5-methylcytosine promotes mRNA export - NSUN2 as the methyltransferase and ALYREF as an m(5)C reader. Cell Res *27*, 606-625.

Yang, Y., and Wang, Z. (2019). IRES-mediated cap-independent translation, a path leading to hidden proteome. J Mol Cell Biol *11*, 911-919.

Yen, T.J., Gay, D.A., Pachter, J.S., and Cleveland, D.W. (1988). Autoregulated changes in stability of polyribosome-bound beta-tubulin mRNAs are specified by the first 13 translated nucleotides. Molecular and cellular biology *8*, 1224-1235.

Yoffe, Y., David, M., Kalaora, R., Povodovski, L., Friedlander, G., Feldmesser, E., Ainbinder, E., Saada, A., Bialik, S., and Kimchi, A. (2016). Cap-independent translation by DAP5 controls cell fate decisions in human embryonic stem cells. Genes Dev *30*, 1991-2004.

Yoo, J.H., and RajBhandary, U.L. (2008). Requirements for translation re-initiation in Escherichia coli: roles of initiator tRNA and initiation factors IF2 and IF3. Mol Microbiol *67*, 1012-1026.

Young, D.J., and Guydosh, N.R. (2019). Hcr1/eIF3j Is a 60S Ribosomal Subunit Recycling Accessory Factor In Vivo. Cell reports *28*, 39-50 e34.

Young, D.J., Guydosh, N.R., Zhang, F., Hinnebusch, A.G., and Green, R. (2015). Rli1/ABCE1 Recycles Terminating Ribosomes and Controls Translation Reinitiation in 3'UTRs In Vivo. Cell *162*, 872-884.

Young, D.J., Makeeva, D.S., Zhang, F., Anisimova, A.S., Stolboushkina, E.A., Ghobakhlou, F., Shatsky, I.N., Dmitriev, S.E., Hinnebusch, A.G., and Guydosh, N.R. (2018). Tma64/eIF2D, Tma20/MCT-1, and Tma22/DENR Recycle Post-termination 40S Subunits In Vivo. Molecular cell *71*, 761-774 e765.

Young, S.K., and Wek, R.C. (2016). Upstream Open Reading Frames Differentially Regulate Gene-specific Translation in the Integrated Stress Response. The Journal of biological chemistry *291*, 16927-16935.

Yu, C.H., Dang, Y., Zhou, Z., Wu, C., Zhao, F., Sachs, M.S., and Liu, Y. (2015). Codon Usage Influences the Local Rate of Translation Elongation to Regulate Co-translational Protein Folding. Molecular cell *59*, 744-754.

Zhang, J.Z., Gao, Y., Lu, Q.P., Sa, R.N., and Zhang, H.F. (2015). iTRAQ-based quantitative proteomic analysis of longissimus muscle from growing pigs with dietary supplementation of non-starch polysaccharide enzymes. J Zhejiang Univ Sci B *16*, 465-478.

Zhang, L., Kasif, S., Cantor, C.R., and Broude, N.E. (2004). GC/AT-content spikes as genomic punctuation marks. Proc Natl Acad Sci U S A *101*, 16855-16860.

Zheng, G., Dahl, J.A., Niu, Y., Fedorcsak, P., Huang, C.M., Li, C.J., Vagbo, C.B., Shi, Y., Wang, W.L., Song, S.H.*, et al.* (2013). ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. Molecular cell *49*, 18-29.

Zhou, J., Wan, J., Gao, X., Zhang, X., Jaffrey, S.R., and Qian, S.B. (2015). Dynamic m(6)A mRNA methylation directs translational control of heat shock response. Nature *526*, 591-594.

Zuberek, J., Kubacka, D., Jablonowska, A., Jemielity, J., Stepinski, J., Sonenberg, N., and Darzynkiewicz, E. (2007). Weak binding affinity of human 4EHP for mRNA cap analogs. RNA *13*, 691-697.

11. Original manuscripts of the discussed publications

A: Weber et al., unpublished

The manuscript contains finalized figures and a detailed description of experiments and methods and is intended for publication. Introduction and discussion sections will still be subjected to major editing and thus are not included. For introduction and discussion of the topic please refer to the main text of the thesis.

DAP5 mediates re-initiation after uORF translation on structured 5′ leader sequences

Ramona Weber[1], Leon Kleemann[1,2], Insa Hirschberg[3], Min-Yi Chung[1], Eugene Valkov[1,4], Elisa Izaurralde[1,†], and Cátia Igreja[1,5]

[1]Department of Biochemistry, Max Planck Institute for Developmental Biology, Max-Planck-Ring 5, D-72076 Tübingen, Germany

[2]Present address: Department of Chemistry and Biochemistry, University of Bern, Freiestrasse 3, 3012 Bern, Switzerland

[3]Department of Mouse Genomics, Friedrich Miescher Laboratory, Max-Planck-Ring 9, 72076 Tübingen, Germany

[4]Present address: Center for Cancer Research, National Cancer Institute, Frederick, MD, 21702-1201

[5] To whom correspondence should be addressed.

† deceased 30th April 2018

Tel: +49-7071-601-1360

Fax: +49-7071-601-1353

Email: catia.igreja@tuebingen.mpg.de

Abstract

Initiation of translation usually relies on the assembly of the eukaryotic initiation factor 4F (eIF4F) complex — consisting of the cap-binding protein eIF4E, the scaffold protein eIF4G and the RNA helicase eIF4A — on the mRNA 5′ cap. eIF4G recruits the 43S preinitiation complex (PIC) to the mRNA which scans the 5′ leader sequence in search for a start site. The 5′ leaders of about 50% of mammalian mRNAs are populated by short upstream open reading frames (uORFs). uORFs have been shown to affect translation of the downstream main ORF (mORF). Some decrease mORF translation by sequestering a fraction of the scanning PICs, whereas others permit a new round of translation at the downstream mORF. Yet, the molecular events coordinating uORF and mORF translation initiation remain understudied.

Here, we find that the non-canonical factor DAP5, or eIF4G homologous protein eIF4G2, is selectively required for translation re-initiation at the mORF following uORF translation. Using ribosome profiling, we identified a set of structured mRNAs containing uORFs with decreased mORF translation in DAP5-null human cells. This group of mRNAs was enriched for signaling and regulatory factors, such as kinases and phosphatases. 5′ leader-based reporters which recapitulate DAP5-dependent re-initiation, indicated that uORF translation precedes mORF re-initiation and persists in the absence of DAP5. Moreover, we observed that uORF length modulates mORF translation. DAP5-dependent re-initiation required eukaryotic release factor 1 (eRF1) and ABCE1-dependent termination and 60S recycling at the uORF, and mRNA selection by eIF4A. Thus, DAP5 emerges as a regulator of translation re-initiation in mRNAs with complex 5′ leaders. Our findings give important mechanistic insight into how translation re-initiation is achieved in mammalian cells and how DAP5 shapes the cellular signaling proteome during development and cancer.

Results

DAP5 mediates the synthesis of signaling proteins

To study the function of DAP5 in translation, we determined the translational landscape of CRISPR-Cas9 engineered DAP5-null and wild-type (WT) HEK293T cells (Figure S1a-c). After isolation and sequencing of ribosome-protected fragments (ribosome profiling) and matched transcriptome analysis (RNA-Seq), we determined genome-wide transcriptional and translational changes (Figs. 1a and S1d) (Ingolia et al., 2011; Zhong et al., 2017). The Ribo-Seq and RNA-Seq were reproducible as replicates clustered together (Figs. S1e, f).

In the absence of DAP5, a group of genes, hereafter referred to as DAP5 targets, showed a significant reduction in translation efficiency (TE; ribosome occupancy/mRNA abundance) (Fig. 1a; n=306, red). Although the majority of DAP5 targets were more abundant, the number of ribosomes per mRNA (footprints or occupancy) decreased in the null cells (Table S1). Other translatome-associated differences were found in a small cohort of mRNAs with increased TE in the null cells (n=23, blue; Fig. 1a and Table S2). In addition to translatome-only changes, we also observed pronounced differences in transcript abundance in the null cells (n=3537; Fig. S1d and Table S2). These differences may result from effects on transcription and/or mRNA turnover following the loss of DAP5.

Notably, DAP5 targets included mRNAs encoding proteins involved in cell signaling and regulation, or cellular response to stimulus, such as the serine/threonine-protein kinases WNK1 [With-No-Lysine (K)1] and ROCK1 (Rho-associated protein kinase 1), the RAC-alpha serine/threonine-protein kinase AKT1 or the phosphatidylinositol 3,4,5-triphosphate 5-phosphatase 2 (INPPL1, a.k.a. SHIP2), among others (Table S1). In agreement with reduced TE, WNK1, ROCK1 and INPLL1 protein levels assessed by immunoblotting were diminished in the absence of DAP5 (Fig. 1c). Importantly, decreased protein synthesis was not caused by deficiency in the expression of eIF4E, eIF4G or eIF4A (Fig. S1c) or changes in global translation in the null cells (Fig. S1g). With the exception of an increase in free 40S subunits,

polysome profiles of DAP5-null cells after sucrose density gradient separation were similar to those of WT cells (Fig. S1g-i). However, the association of *WNK1* and *ROCK1* mRNAs with polysomes, but not *GAPDH*, shifted from the heavy (Figs. S1j-l, lanes 16-18) to the light (Figs. S1j-l, lanes 12-15) polysome fractions of the gradient in the absence of DAP5. These results indicate that the translation of a specific subset of transcripts is DAP5-dependent.

DAP5 targets 5′ leaders have unique features

In addition to the differences in TE, we also observed qualitative changes in the pattern of ribosomal occupancies (footprints) in DAP5 target mRNAs. Ribosome occupancy at the annotated (main) coding sequences (CDS) was markedly decreased in the absence of DAP5. Moreover, ribosome footprints were skewed towards the 5′ leaders of these transcripts (Figs. 1d and S2). Estimation of footprint density (RFP) in all 306 mRNAs revealed that despite the reduction of footprints in the CDS, translation was increased on the 5′ leaders in the null cells compared to WT cells (Fig. 1e), as measured by the ratio of footprints within the 5′ leader relative to the footprints at the annotated downstream CDS start codon. Increased translation in the 5′ leaders of the DAP5 target mRNAs occurred at upstream open reading frames (uORFs) as reflected by experimentally determined initiation-site-profiling in cells treated with harringtonine and lactimidomycin (Lee et al., 2012) (Figs.1d, f). In the presence of these drugs, ribosomes accumulate at the start codons but are allowed to complete elongation over the rest of the CDS (Ingolia et al., 2011; Lee et al., 2012). The majority of the DAP5 targets had more than one uORF in the 5′ leader, with median length of 26 codons, and initiating at near-cognate start codons (CUG, GUG, UUG and AUC) instead of the conventional AUG (Figs. 1g-i). For instance, *WNK1*, a regulator of development and WNT signaling (Rodan and Jenny, 2017), revealed increased ribosome occupancy in two GUG (one of which is in frame with the main CDS), one CUG and one UUG uORFs (Fig. 1d). These observations suggest that DAP5 mediates translation of the main CDS but not of the uORFs located upstream in the transcript.

In fact, close inspection of the RFP profiles shows that cap-proximal uORF translation is DAP5-independent (increased RFPs in null cells), whereas downstream uORFs and CDS are translated in a DAP5-dependent manner (decreased RFPs in null cells) (Figs. 1d, e and S2).

Further analysis of the 5′ leader sequences revealed that DAP5 target mRNAs also have increased length, high GC-content and decreased minimum free energy of 5′ leaders (Fig. S3a-c). Visual inspection of the leader sequences revealed that DAP5-independent uORFs tend to concentrate in the regions of the 5′ leaders adjacent to high predicted propensity for structure (Figs. 1d and S2). These findings indicate that the 5′ leaders of DAP5 targets likely form structured elements which define positional information for DAP5-dependent translation.

Target mRNA 5′ leaders are necessary and sufficient to induce DAP5-dependent translation

Given that DAP5 acts in the initiation of translation of specific mRNAs (Liberman et al., 2015; Liberman et al., 2009; Marash and Kimchi, 2005; Sugiyama et al., 2017; Weingarten-Gabbay et al., 2014; Yoffe et al., 2016), we tested if the sequences of *WNK1*, *ROCK1* and *AKT1* 5′ leaders were sufficient to confer DAP5 sensitivity on a *Renilla* luciferase (R-LUC) reporter (Fig. 2a-c). In comparison to the control reporter with a short 5′ leader, the R-LUC luminescence driven by *WNK1*, *ROCK1* and *AKT1* 5′ leaders was reduced in DAP5-null cells to 20%, 30% and 40%, respectively (Figs. 2a-c). Decreased translation of *WNK1*-, *ROCK1*- and *AKT1*-R-LUC reporters was not due to variations in mRNA abundance in the absence of DAP5 (Figs. S3d-i). In addition, re-expression of DAP5 (full length; FL) in the null cells completely restored R-LUC activity (Figs. 2a-d, S3j-l), indicating that the 5′ leaders of *WNK1*, *ROCK1* and *AKT1* are sufficient to stimulate DAP5-dependent translation of the R-LUC CDS.

To elucidate the molecular details of DAP5-dependent translation, we first tested if translation of the R-LUC reporters was influenced by changes in the sequence of DAP5 protein. We measured the activity of *WNK1*-, *ROCK1*- and *AKT1*-R-LUC in the null cells upon transient

expression of DAP5 mutants. Translation was dependent on the interaction with the RNA helicase eIF4A, as R-LUC activity was not restored in the presence of DAP5 carrying mutations in the eIF4A-binding region (eIF4A*; Figs. 2a-d, S4a). However, binding to eIF4A was not sufficient to induce DAP5-dependent translation, as the expression of the eIF4A-interacting domain MIF4G (middle domain of eIF4G, residues 1-475) alone failed to restore R-LUC activity (DAP5 MIF4G; Figs. 2a-d, S3j-l, S4a). The interaction of DAP5 with eIF2β was also important for translation of the R-LUC reporters; R-LUC activity was still reduced in null cells expressing a DAP5 protein lacking the W2 domain (ΔW2, residues 1-722) and unable to associate with the β subunit of the ternary complex (Figs. 2a-d, S3j-l, S4b).

We then asked if overexpression of eIF4G (FL) or N-terminal truncated eIF4G resembling DAP5 (lacking the PABP and eIF4E binding sites; eIF4G ΔN) (Fig. 2d) would suffice to translate the R-LUC reporters in the absence of DAP5. Curiously, none of the proteins was able to reestablish R-LUC activity (Figs. 2a-c), indicating that *WNK1*, *ROCK1* and *AKT1* 5′ leaders drive DAP5-specific translation. Lastly, we also used DAP5 chimeric proteins where the MIF4G, MA3 or W2 domains were swapped with the respective eIF4G domains (Fig. 2d). Relative to the re-expression of DAP5 (FL), the chimeras were unable to fully restore R-LUC luminescence in the null cells, with the MIF4G and W2 chimeras showing the strongest effects on R-LUC translation (Figs. 2a-d). These findings reinforce the notion that all domains of DAP5, and their specific interactions, are necessary for efficient translation of the R-LUC reporters with DAP5 targets 5′ leaders. All DAP5 protein constructs were expressed at similar levels and reporter mRNA levels were not altered between the conditions (Fig. S3d-l).

Only eIF4A-bound DAP5 can interact with mRNA

To investigate the recruitment of DAP5 to target mRNAs, we performed RNA-pulldown assays and RT-qPCR. In contrast to V5-SBP-MBP, V5-SBP-DAP5 efficiently associated with *WNK1* and *ROCK1* mRNAs, but not *GAPDH* (Figs. 2e-h). Interestingly, the

association of DAP5 with the targets was abolished by mutations on the eIF4A-binding region (eIF4A*; Figs. 2e, f), suggesting that eIF4A mediates mRNA binding. In addition, the MIF4G domain of DAP5 was sufficient to pull down *WNK1* and *ROCK1* mRNAs, either alone (MIF4G) or when present in other DAP5 constructs (MA3 chimera and W2 chimera; Figs. 2e, f). The DAP5 MIF4G was also specifically required for mRNA binding, as substitution by the respective domain in eIF4G, which is 15% sequence identical, prevented the association of DAP5 with the target. Consistent with a role of eIF4A in target mRNA recognition, we observed that one third of DAP5 targets (n=102; Fig. S4c; Table S4) showed experimentally determined Rocaglamide A (RocA) sensitivity (Iwasaki et al., 2016). RocA is a translation inhibitor that clamps eIF4A onto polypurine sequences on the mRNA (Iwasaki et al., 2016). RocA-sensitive mRNAs, such as *WNK1*, show decreased RFP density at the CDS and premature uORF translation in the presence of the drug (Figs. S4d, e) (Iwasaki et al., 2016). These findings indicate that DAP5 specifically interacts with transcripts containing structured 5′ leaders when in complex with eIF4A. In the absence of an interaction with the RNA helicase, binding to the target, and thus translation, is compromised (Figs. 2a-c and e, f).

The interaction of DAP5 ΔW2 with *WNK1* and *ROCK1* transcripts was comparable to wild type protein (Figs. 2e, f). This result suggests, that impaired translation of the R-LUC reporters in null cells upon DAP5 ΔW2 expression (Figs. 2a-c) is unrelated to target binding and is most likely associated with the function of the W2 domain in the initiation of translation.

In contrast to DAP5, eIF4G bound strongly to all tested mRNAs, including the DAP5 targets *WNK1* and *ROCK1*; however, its interaction with mRNA was dependent on the N-terminal region, which contains PABP- and eIF4E-binding motifs (Fig. 2d-g). All proteins were expressed at equivalent levels and did not alter mRNA input levels (Figs. 2e-g, input panels, h).

Altogether, our findings show that both eIF4G and DAP5 bind to *WNK1* and *ROCK1* (DAP5 targets) mRNAs. Whereas the interaction of DAP5 with the mRNA is specific and

reliant on eIF4A, eIF4G binds to all capped mRNAs as part of the eIF4F complex. Once in the mRNA, DAP5 mediates the synthesis of WNK1 and ROCK1 proteins (main CDS) but is dispensable for the translation of cap-proximal uORFs (Figs. 1d, e). Translation of the latter is most likely eIF4G- and eIF4F-dependent. Thus, initiation of translation along the structured 5′ leaders of DAP5 targets switches from a DAP5-independent and eIF4F-dependent mechanism to a DAP5- and eIF4A-dependent mechanism.

DAP5-mediated translation is cap-dependent

DAP5 has been proposed to drive IRES-dependent (Henis-Korenblit et al., 2000; Lee and McCormick, 2006; Lewis et al., 2008; Liberman et al., 2009; Marash and Kimchi, 2005; Marash et al., 2008; Nevins et al., 2003; Weingarten-Gabbay et al., 2014; Yoffe et al., 2016) and IRES-independent (de la Parra et al., 2018) translation. Given that the 5′ leaders of DAP5 targets contain structured elements that could represent IRESes, we generated two cap-proximal truncations in *WNK1*-R-LUC mRNA (Δ1 and Δ2; Fig. 3a). These truncations partially or completely removed the structured region of the 5′ leader containing the uORFs translated in a DAP5-independent manner. Both truncations reduced the abundance of the reporter, and consequently R-LUC activity, in WT and DAP5-null cells (Fig. S4f-h), suggesting they might affect mRNA stability and/or transcription. To only assess changes in translation, we determined the protein/mRNA ratios (TE) for the *WNK1*-R-LUC Δ1 and Δ2 reporters. We observed that despite the low mRNA levels, the Δ1 reporter was still translated in WT cells (Fig. 3b). Moreover, in DAP5-null cells the Δ1 mRNA was less translated and also less dependent on DAP5 relative to the reporter with the complete 5′ leader (*WNK1*) (Figs. 3b and S4f). In addition, *WNK1*-R-LUC Δ2 mRNA was not translated. These findings indicate that the cap-proximal region of the 5′ leader is critical for DAP5-mediated translation.

As truncations in the 5′ leader change mRNA structure and remove uORFs, it remained unclear if the observed effects in R-LUC translation resulted from decreased binding of DAP5

to a putative IRES or to the disruption of cap-proximal uORFs translation. To discriminate between these possibilities, we blocked cap/eIF4F-dependent translation via the overexpression of an engineered eIF4E-binding protein (4E-BP) (Peter et al., 2015) and tested binding of DAP5 to *WNK1*- and *ROCK1*-R-LUC mRNAs. As shown in cap-based pulldowns, overexpressed 4E-BP bound to eIF4E and abolished its interaction with eIF4G (Fig. 3c), thus suppressing cap/eIF4F-dependent translation. Notably, binding of DAP5 to the transcripts was suppressed in the presence of 4E-BP (Figs. 3d, e). Overexpressed proteins were pulled down at comparable levels in the different experimental conditions (Fig. 3f). These findings suggest that DAP5-mediated translation is coupled with recruitment of ribosomes by the eIF4F complex and not by an IRES-dependent mechanism. Given that translation of cap-proximal uORFs occurs in the null cells, we conclude that DAP5 likely acts downstream of cap-dependent translation initiation, i.e., DAP5 drives translation of main CDSes in mRNAs where eIF4F-loaded ribosomes translate uORFs.

DAP5 mediates re-initiation after uORF translation

Our data suggests that the uORFs located in the 5′ leaders of DAP5 targets serve an important role in the translation of the main, canonical CDS by DAP5. However, little is known about the functional contribution of uORFs to the regulation of DAP5-dependent translation. To understand if the uORFs in the 5′ leaders of DAP5 targets play a role in the translation of the downstream CDS, we transfected WT and DAP5-null cells with versions of the *WNK1*-R-LUC reporter containing altered uORFs features. *WNK1* contains at least five uORFs, one of which, uORF2, is in frame with the AUG of the downstream CDS. uORF2 is translated in the absence of DAP5, initiates from a GUG start codon (uGUG) and is 22 codons long (Figs. 1d and 4a). To characterize the functionality of uORFs in *WNK1*, we optimized the initiation sequence context of uORF2 by mutating the uGUG to conventional AUG (uORF2+; Fig. 4a). *WNK1*-R-LUC-uORF2+ mimicked the reporter with the natural 5′ leader of *WNK1* (*WNK1*): It

promoted the expression of R-LUC (main CDS) in a DAP5-dependent manner upon expression in WT and DAP5-null HEK293T cells (Figs. 4b, c, lanes 1-3). To confirm that the uAUG was used in the initiation of translation, we removed all STOP codons, and consequently all uORFs, in frame with the AUG of R-LUC (3 in total) (NO STOP; Fig. 4a). In the *WNK1*-R-LUC-NO STOP reporter, the uAUG was the only initiating codon and produced an R-LUC protein with an extended N-terminal region (70 kDa instead of 35 kDa; Fig. 4c, lanes 13-15 vs 1-3). Although the protein levels of the two forms of the luciferase were similar (Fig. 4c), the N-terminal extension reduced R-LUC activity (Fig. 4b). In addition, the expression and activity of the long R-LUC was similar in WT and null cells, suggesting that its translation was not mediated by DAP5 (Figs. 4b, c). These findings indicate that in the absence of uORF translation DAP5 is not required for the translation of R-LUC.

We also generated reporters where only one (ΔSTOP1) or two (ΔSTOP1+2) of the three STOP codons were removed. Single and double deletions of these STOPs increase the size of uORF2 to 188 or 229 codons, respectively (Fig. 4a). Notably, in these settings R-LUC translation was abolished in HEK293T cells (Figs. 4b, c). Thus, DAP5-dependent translation of the main CDS is influenced by uORF length, as only the short uORF2 (22 codons long) was able to drive R-LUC translation. We then tested the maximum uORF2 length supporting DAP5-mediated translation of R-LUC by extending the position of its STOP to 29, 39 or 49 codons downstream of uAUG (Fig. S5a). Although all tested reporters supported DAP5-dependent translation of R-LUC, the expression and activity of the luciferase in WT and null cells upon re-expression of DAP5 was inversely correlated with uORF2 length (Figs. S5b, c). None of the observed differences could be explained by varying mRNA levels (Figs. 4d, e and S5d, e). The fact that DAP5-dependent translation of R-LUC is only primed by the translation of short uORFs supports the notion that DAP5 drives re-initiation of translation.

We also tested the ability of DAP5 to prime re-initiation of translation in the context of the natural 5′ leader (*WNK1*; Fig. 5a). In the natural 5′ leader sequence, uORF5 initiates with a

conventional AUG (uAUG), is translated in a DAP5-dependent manner and is in frame with a UGA STOP located 4 codons downstream (Figs. 1d, 5a). uORF5 translation occurs in the –1 frame. We interfered with termination at the UGA codon and produced reporters encoding uORF5 with distinct lengths: 118, 30, 19 or 9 codons long (Fig. 5a). The *WNK1* reporters with the engineered 5′ leaders, were transfected in WT and null cells and assayed for R-LUC activity and expression (Figs. 5b, c and S5f-h). As observed above, DAP5-dependent translation of R-LUC was regulated by uORF length. In the presence of a long uORF5 (uORF118 and uORF30), R-LUC was weakly translated (Figs. 5b, c, lanes 4-9 vs 1-3). Small uORF5 primed R-LUC translation in a DAP5-depedent manner but were still less efficient than the natural uORF5 which encodes a 4 amino acid micropeptide (Figs. 5b, c, lanes 10-15). Changes in mRNA abundance of the different reporters were not sufficient to explain the variation in the efficiency of R-LUC translation (Figs. 5b, S5f-h). These findings support a model where DAP5 promotes translation of main CDSes after short uORF translation. In addition, the length of uORF supporting DAP5 re-initiation of translation varied with the distance of the uORF to the cap. The maximal uORF5 length supporting DAP5-dependent translation of R-LUC (4-19 codons; Fig. 5a-c and S5f-h) was smaller than uORF2 length (29 codons; Figs. S5a-e). This uORF positional effect might be associated with the loss of initiation factors during the transit of the ribosome along the 5′ leader and after termination of uORF translation. Indeed, re-initiation depends on the time required for uORF translation which is determined by uORF length and elongation rate, and the initiation factors involved in the re-initiation event (Barbosa et al., 2013; Bohlen et al., 2020; Child et al., 1999; Kozak, 2002; Poyry et al., 2004; Roy et al., 2010).

DAP5 utilizes post-termination translation complexes

In vertebrates, uORFs tend to be characterized by sequence contexts that disfavor translation initiation (Chew et al., 2016). In non-optimal initiation contexts, the 43S PICs inefficiently recognize start codons, scan past the uORFs and initiate translation at a

downstream START codon, a mechanism known as leaky scanning (Barbosa et al., 2013). To explore the possibility that DAP5 initiates translation of downstream CDSes using 43S scanning complexes that scan past uORFs, we again made use of the *WNK1* reporter with the natural 5′ leader (*WNK1*; Fig. 5a). We first determined the changes in R-LUC translation if the 5′ leader lacks all STOP codons in frame with the main ORF (ΔSTOP IF; Fig. 5a). When transfected into WT cells, this reporter originated R-LUC proteins with different sizes (short and long), as observed by immunoblotting (Fig. 5d, lane 3). The synthesis of short and long versions of R-LUC indicates that 43S PICs scanning the 5′ leader initiate translation at different upstream start codons in frame with the main CDS, as expected by leaky scanning. In the null cells though, expression of short R-LUC (35 kDa), but not the majority of the long R-LUCs with N-terminal extensions, was diminished (Fig. 5d, lane 4). To understand if DAP5 primes the initiation of translation with PICs that were able to scan until the main AUG or drives re-initiation using post-termination complexes, we removed the STOPs from all frames in the 5′ leader of *WNK1* (ΔSTOP all F; Fig. 5a). In this reporter, translation can initiate at multiple start codons, but does not terminate before the main CDS R-LUC. Thus, re-initiation of translation does not occur, as the 5′ leader is devoid of uORFs. In cells, long R-LUC proteins were expressed independently of DAP5 (Fig. 5d, lanes 5 and 6). Even if engineering of *WNK1* 5′ leader sequence was performed without introducing new start codons or altering the initiation contexts of the different uORFs (see methods section for details), removal of STOP codons changed start codon recognition, as shown by the presence of long R-LUC proteins with distinct sizes (Fig. 5d, lanes 3-6 vs 1, 2). Notably, short R-LUC was not synthesized in WT and null cells.

These results have several implications. First, translation of the main CDS in the context of *WNK1* 5′ leader is DAP5-dependent. Second, initiation at the main AUG only occurs after uORF translation, suggesting that DAP5 is critical for re-initiation of translation at the main CDS. Lastly, in *WNK1* 5′ leader the PICs are unable to scan until the main AUG, supporting

the notion that DAP5 acts on post-termination translation complexes. Accordingly, the distribution of 40S footprints along the 5′ leader and at the main START codon of DAP5 targets is strikingly different from other mRNAs (Fig. 5e), as measured using previously published 40S ribosome footprinting datasets in HeLa cells (Bohlen et al., 2020). For the majority of the mRNAs expressed in human cells, scanning 40S subunits are detected along all the 5′ leader and culminate at the start codon, since initiation is slower than scanning (Fig. 5e) (Bohlen et al., 2020). In DAP5 targets however, 40S footprints are skewed towards the 5′ end of the mRNA and depleted towards the main CDS START (Fig. 5e). The distribution of 40S footprints along the 5′ leaders of DAP5 targets resembles the distribution of 80S RFPs in the absence of DAP5 (Fig. 1e) or initiating ribosomes in cells treated with harringtonine and lactimidomycin (Fig. 1f) (Lee et al., 2012). Thus, although 80S ribosomes are present and initiate translation at the main AUG of DAP5 targets (Figs. 1d, e, 5f), scanning 40S subunits do not frequently reach the main AUG in DAP5 targets (Fig. 5e). The inefficient scanning of 40S subunits might be coupled with the intrinsic structured nature of the 5′ leaders (Figs. S3b, c) and uORF translation (Fig. 1d) in DAP5 targets.

Simultaneous uORF and main CDS translation in the DAP5 targets

The luciferase-based reporters used in the previous experiments suggested that uORF translation is pervasive and necessary for DAP5-dependent translation of the main CDS. However, in these experiments we are unable to detect the synthesis of uORF-derived peptides, and therefore confirm uORF translation. To simultaneously detect and quantify uORF and mORF translation, we adopted a split-fluorescent protein approach using mNeonGreen2 (mNG2) that expresses the yellow-green-colored protein in two fragments: $mNG2_{1-10}$ and $mNG2_{11}$. $mNG2_{1-10}$ originates a non-fluorescent mNG2 due to the lack of 11th β-strand; however, upon co-expression with $mNG2_{11}$ (16-aa protein), the two fragments assemble a functional mNG2 molecule (Chen et al., 2020; Feng et al., 2017; Leonetti et al., 2016). The

Weber et al.

uORF2 (22 aa) of the *WNK1* 5′ leader was replaced with the mNG2$_{11}$ CDS initiating at a uAUG. In addition, the main CDS encoded the EBFP (enhanced blue fluorescent protein) (Fig. 6a). The split-fluorescent reporters were expressed in WT and DAP5-null cells together with a transfection control expressing mCherry. The non-overlapping excitation and emission spectra of the three fluorophores allowed their simultaneous detection by flow cytometry (Fig. S6a-c).

The *WNK1* mNG2$_{11}$+EBFP reporter recapitulated DAP5-dependent translation of the main CDS, as EBFP signal decreased significantly in DAP5-null cells and increased again upon DAP5 re-expression (Figs. 6b, c). Moreover, in WT cells, co-expression of the two mNG2 plasmids generated fluorescent signal in up to 9% of the cells (Fig. S6a titration). Although the complementation efficiency of the split-mNG2 system was low compared to the transfection efficiency in HEK293T cells (close to 50% mCherry-positive cells in WT cells and around 36% in KO cells; Fig. S6c), it clearly showed that uORF translation occurs in the *WNK1* 5′ leader. Expression of the mNG2 plasmids in trans did not generate a yellow-green signal (Fig. S6a). Close inspection of the fluorescent mNG2 output, showed that the majority of WT and mNG2-positive cells were also EBFP-positive (Fig. 6b and c). In contrast, in the absence of DAP5 a significant proportion of cells expressing a functional mNG2 do not express EBFP, as main ORF translation is suppressed in the null cells (Fig. 6b and c). The number of mNG2 and EBFP double positive cells was partially restored upon re-expression of DAP5 in the null cells (Fig. 6b and c). mNG2 (uORF) translation was not dependent on DAP5 expression (Fig. 6d). Moreover, in WT cells increasing mNG2$_{11}$ uORF2 size to 188 aa (*WNK1* ΔSTOP1+mNG2$_{11}$+EBFP; Fig. 6a) still originated the split-mNG2 signal, but suppressed EBFP expression, as re-initiation after long uORF translation is blocked (Fig. 6e). These observations confirm that uORF translation in the 5′ leader of DAP5 targets promotes main ORF translation. Another implication of our experiments using different reporter systems is that uORF2 and main CDS sequences and peptides are not relevant for the re-initiation of translation by DAP5,

excluding the possibility that uORF translated peptides influence CDS expression *in cis*. These experiments do not dismiss however, that uORF-derived peptides are functional in cells.

DAP5-dependent translation requires termination and ribosome recycling

To further test the re-initiation model by DAP5, we interfered with termination of translation by exploiting a dominant negative mutant of the release factor 1, eRF1[AAQ] (Brown et al., 2015; Shao et al., 2016), to cause local translation arrest at STOP codons. eRF1[AAQ] is unable to hydrolyze the peptidyl-tRNA after STOP codon recognition (Frolova et al., 1999). Cells were transfected with the *WNK1*-R-LUC and GFP-F-LUC reporters in the absence or presence of increasing amounts of eRF1[AAQ] and luciferase activities and expression were measured. As expected upon termination inhibition, eRF1[AAQ] expression decreased R-LUC and GFP-F-LUC protein levels in a concentration-dependent manner (Fig. 7b, lanes 1-4). However, the R-LUC/F-LUC activity ratio varied if R-LUC translation was primed or not by DAP5. In the context of *WNK1* 5′ leader (*WNK1*-R-LUC and *WNK1*-R-LUC-uORF2+), increasing levels of eRF1[AAQ] proportionally decreased R-LUC activity (Figs. 7a-c). In contrast, DAP5-independent translation of R-LUC using a reporter containing a short 5′ leader (R-LUC) or an engineered *WNK1* 5′ leader without STOP codons that leads to the synthesis of an N-terminally extended R-LUC, was less affected by the eRF1 mutant. In these cases, R-LUC/F-LUC activity ratios were constant or even increased in the presence of increased levels of the release factor (Figs. 7b, c). In all the conditions the abundance of the R-LUC reporters remained unchanged (Figs. S6d, e). These observations suggest that inhibition of termination after uORF translation impairs DAP5 function in the initiation of translation at the main CDS.

In agreement with the re-initiation model, similar findings were obtained when 60S recycling was impaired in cells expressing the *WNK1*-R-LUC reporters. As expected, shRNA-mediated depletion of ABCE1 (ABCE1 KD; Fig. 7d) decreased the levels of free 60S subunits in cells, as judged in polysome profiles of control (scramble) or ABCE1 shRNA-treated cells

after sucrose density gradient separation (Fig. 7e). In cells with reduced levels of ABCE1, DAP5-dependent translation of R-LUC (*WNK1*-R-LUC and *WNK1*-uORF2+-R-LUC reporters) was pronouncedly decreased compared to DAP5-independent translation of R-LUC (R-LUC and *WNK1*-NO STOP-R-LUC reporters) (Figs. 7f, g). Depletion of ABCE1 did not affect the levels of the different R-LUC transcripts (Figs. S6f, g). These findings indicate that DAP5 acts on post-termination translation complexes following 40S and 60S subunits dissociation.

The non-constitutive eIF3 subunit eIF3J, eIF2D and the related DENR/MCTS-1 complex associate with the termination machinery (Pisarev et al., 2007; Young and Guydosh, 2019; Young et al., 2018). While eIF3J has been found to be an accessory factor in 60S recycling (Young and Guydosh, 2019), eIF2D and the DENR/MCTS-1 complex recycle 40S ribosomal subunits (Young et al., 2018). In addition, eIF2D and DENR/MCTS-1 have been implicated in re-initiation after stuORFs (short uORFs in good Kozak context) translation (Ahmed et al., 2018; Schleich et al., 2017; Schleich et al., 2014). Consistent with a role in recycling, deletion of these factors in yeast promotes re-initiation of translation in 3′ UTRs, or in reporters containing uORFs, by unrecycled ribosomes (Young and Guydosh, 2019; Young et al., 2018). To exploit if altered function of these factors affects re-initiation of translation by DAP5, we depleted the proteins in cells using shRNA-mediated knockdown (Figs. S7b, c). Cells were then transfected with the reporters driving R-LUC (main CDS) synthesis by re-initiation (DAP5)-dependent and independent mechanisms (Figs. 7a, S7a). In the absence of DENR, MCTS-1 and eIF2D, R-LUC expression from all the reporters was slightly reduced (Figs. S7d, e), suggesting that inhibition of 40S recycling after uORF translation does not favor re-initiation of translation by DAP5. eIF3J depletion also did not change the efficiency of re-initiation of R-LUC translation by DAP5 (Figs. S7f, g). Thus, 60S subunit recycling by eIF3J does not influence re-initiation of translation at the 5′ leaders of DAP5 targets.

Material and Methods
DNA constructs

All the constructs were confirmed by sequencing and are listed in Table S1. All the mutants used in this study were generated by site-directed mutagenesis using the QuickChange Site-Directed Mutagenesis kit (Stratagene).

Generation of the DAP5-null cell line

Two sgRNAs targeting DAP5 were designed and cloned into the pSpCas9(BB)-2A-Puro (PX459) vector [a gift from F. Zhang, Addgene plasmid 48139; (Ran et al., 2013)] using the CHOPCHOP (http://chopchop.cbu.uib.no) online tool as previously described (Peter et al., 2017). Clonal cell lines were obtained and confirmed for gene editing as described in Peter et al., 2017. We observed two frameshift mutations in exon 10 (172 bp deletion in exon/intron 10, and a 1 bp insertion) targeted by sgDAP5-a. These mutations caused defective splicing and intron retention, as evidenced by subsequent RNA sequencing (Fig. S1b). Two mutations were detected in exon 12 (1 bp insertion and 12 bp deletion) targeted by sgDAP5-b. The lack of DAP5 protein was further confirmed by western blotting (Figs. 1c, S1c). RNA sequencing revealed that DAP5 transcript levels were severely reduced in the null cells compared to wild-type cells (Fig. S1a), most likely as a result of NMD. The following guide sequences were used: sgDAP5-a: 5′-CACGTACCTTGGCTCGTTCA-3′; sgDAP5-b: 5′-ACACCATTGGGTTCCTCGCA-3′

Ribosome profiling and RNA sequencing

For ribosome profiling and RNA sequencing HEK293T wild-type and DAP5-null cells were plated on 10 cm dishes 24 hours before harvesting (3.2×10^6 WT cells and 3.5×10^6 null cells per plate). Cells were harvested as described in (Calviello et al., 2016). For total RNA sequencing, RNA was extracted using the RNeasy Mini Kit (50) (Qiagen) and processed

Weber et al.

according to the Illumina TruSeq RNA Sample Prep Kit. For ribosome profiling the original protocol (Ingolia et al., 2012) was used in a modified version also described in (Calviello et al., 2016). The ribosome profiling and total RNA sequencing pools were sequenced on an Illumina Hiseq3000 instrument. Reads originating from ribosomal RNA were removed using Bowtie2 (Langmead and Salzberg, 2012). Remaining reads of the RNA sequencing library were mapped onto the human genome using Tophat2 (Kim et al., 2013) which resulted in 15.7-20.5 million mapped reads with an overall read mapping rate >94% for the RNA sequencing experiment. Ribosome profiling reads were subjected to statistical analysis using RiboTaper that aims at identifying actively translating ribosomes based on the characteristic three-nucleotide periodicity (Calviello et al., 2016). Reads of 29 and 30 nucleotides length showed the best three-nucleotide periodicity and where therefore used for subsequent mapping onto the human genome. This resulted in 2.8-3.8 million mapped reads with an overall read mapping rate >95% for the ribosome profiling experiment. Read count analysis was performed using QuasR (Gaidatzis et al., 2015). Differential expression analysis was conducted using edgeR (McCarthy et al., 2012; Robinson et al., 2010). Translation efficiency (TE) was calculated using RiboDiff (Zhong et al., 2017).

Harringtonine and LTM datasets from human HEK293 cells were downloaded from the Sequence Read Archive database (accession: SRA056377). RocA datasets were retrieved from the GEO database (accession number: GSE70211). Ribosomal RNA reads were filtered using Bowtie 2 (Langmead and Salzberg, 2012). The remaining reads were mapped on the hg19 (UCSC) human genome or the mm9 (UCSC) mouse genome with TopHat2 (Kim et al., 2013). No specific filters for read length were applied.

Analysis of GO terms and nucleotide compositions

Upregulated and downregulated gene groups were defined as being significantly deregulated (FDR<0.005) with a logFC>0 and logFC<0, respectively. No cutoff of the logFC

value was applied so that genes with little but significant changes could also be detected. GO analysis was performed with the R based package goseq (Young et al., 2010). For analysis of 5′ leader nucleotide composition, the respective mRNA sequences were fetched using biomaRt (Durinck et al., 2005; Durinck et al., 2009). Analysis of GC content and length of 5′ leader was performed with R based scripts.

RNA structures were calculated using the ViennaRNA package 2.0 (Lorenz et al., 2011). Metagene analysis was performed using the Deeptools suite of functions (Ramirez et al., 2016). For uORF number, size and start codon analysis the accumulation of ribosome footprint on start codons was assessed using the ribosome profiling dataset in HEK293 cells treated with harringtonine (Lee et al., 2012). Identity of the start codon and the corresponding stop codon was manually assigned.

Ribosome footprint density plots for individual sequencing tracks were visualized using the Integrative Genomics Viewer (IGV) visualization tool (Robinson et al., 2011; Thorvaldsdottir et al., 2013).

Transfections, northern and western blotting

In the rescue assays described in Figs. 2, 3b, 4, and 5, 0.64 x 10^6 WT cells or 0.7 x 10^6 DAP5-null cells were transfected, after seeding in 6-well plates, using Lipofectamine 2000 (Invitrogen). The transfection mixtures contained different amounts of the plasmids expressing R-LUC, GFP-F-LUC or V5-SBP-fusion proteins (*WNK1*-R-LUC reporters: 0.5 µg; GFP-F-LUC: 0.25 µg ; V5-SBP-MBP: 0.3 µg; V5-SBP-DAP5 FL 0.8 µg and mutants: 4A*: 3.25 µg, MIF4G: 0.8 µg, ΔW2: 1.2 µg; V5-SBP-eIF4G FL: 3.25 µg and mutants: ΔN: 0.8 µg; V5-SBP-Chimeras: 0.8 µg).

Cells were harvested two days after transfection and firefly and *Renilla* luciferase activities were measured using the Dual-Luciferase reporter assay system (Promega). Total RNA was isolated using TriFast (Peqlab biotechnologies). For northern blotting, total RNA was

separated in 2% glyoxal agarose gels and blotted onto a positively charged nylon membrane (GeneScreen Plus, Perkin Elmer). [^{32}P]-labelled probes specific for each transcript were generated by linear PCR. Hybridizations were carried out in hybridization solution (0.5 M NaP pH=7.0, 7% SDS, 1 mM EDTA pH=8.0) at 65°C overnight. After extensive washes with washing solution (40 mM NaP pH=7.0, 1% SDS, 1 mM EDTA pH=8.0), the membranes were exposed and band intensities were quantified by PhosphoImager.

Western blot was performed using standard methods. In brief, cells were washed with PBS and lysed with sample buffer (100 mM Tris-HCl pH=6.8, 4% SDS, 20% glycerol, 0.2 M DTT) followed by boiling 5 minutes at 95°C and vortexing to shear genomic DNA. After SDS-PAGE, proteins were transferred onto a nitrocellulose membrane (Santa Cruz Biotechnology) by tank transfer. Primary antibodies were incubated overnight at 4°C, secondary antibodies for an hour at room temperature. All western blots were developed with freshly mixed 10A: 1B ECL solutions and 0.01% H$_2$O$_2$ [Solution A: 0.025 % Luminol (Roth) in 0.1 M Tris-HCl pH=8.6; Solution B: 0.11% P-Coumaric acid (Sigma Aldrich) in DMSO]. Antibodies used in this study are listed in Table S5.

Reverse transcription (RT) and quantitative PCR (qPCR)

1 µg of RNA was mixed with 0.66 µg of random hexamer primers (N$_6$) and denatured at 72°C for 5 min. After addition of a reaction mixture containing a final concentration of 1 x RT buffer, 20 U RiboLock RNase Inhibitor (Thermo Scientific) and 1 mM dNTPs, the RNA samples were incubated at 37°C for 5 min. Incubation with RevertAid H Minus Reverse Transcriptase (200 U, Thermo Scientific) was first performed for 10 min at 25°C, and then at 42°C for one hour. The RT reaction was stopped by incubating the samples for 10 min at 70°C. The qPCR was performed with 1x iTaq SYBR Green Supermix (Biorad), 0.4 µM of each primer and 1 µl of the cDNA sample. mRNA levels were determined by qPCR using sequence-specific primers for the indicated transcripts. qPCR primers designed using Primer-BLAST (NCBI) are

listed in Table S3. Normalized transcript expression ratios from three independent experiments were determined using the Livak method (Livak and Schmittgen, 2001).

Polysome profiling

Polysome profiles were performed as described before (Kuzuoglu-Ozturk et al., 2016). HEK293T cells were pretreated with cycloheximide (50 µg/ml) for 30 min. Lysates were prepared in lysis buffer (10 mM Tris-HCl pH=7.4, 10 mM NaCl, 1.5 mM $MgCl_2$, 0.5% Triton X-100, 2 mM DTT, 50 µg/ml cycloheximide) and polysomes separated on a 10-50% sucrose gradient in gradient buffer (10 mM Tris-HCl pH 7.4, 75 mM KCl, 1.5 mM $MgCl_2$). Polysome fractions were collected using the Teledyne Isco Density Gradient Fractionation System. Protein from sucrose fractions was isolated by methanol extraction. In detail, 4x volumes of MetOH were mixed with the sucrose fractions, then mixed with 1x volume of chloroform and then with 3x volumes of water. After centrifugation, the upper phase was removed leaving the lower and inter-phases which were precipitated using 3x volumes of MetOH. Samples were spun down and the dried pellet dissolved in 2x protein sample buffer. Fractions were analyzed by RT-qPCR.

RNA pulldown

For the RNA pulldown, 3 x 10^6 WT HEK293T cells were plated in 10-cm plates and transfected using Lipofectamine 2000 (Invitrogen) with the following plasmids expressing V5-SBP fusions: MBP: 1.5 µg; DAP5 FL 4 µg and mutants: 4A*: 15 µg, MIF4G: 4 µg, ΔW2: 6 µg; eIF4G FL: 15 µg and mutants: ΔN: 4 µg; V5-SBP-Chimeras: 4 µg. A detailed description of the RNA immunoprecipitation procedure can be found in (Kuzuoglu-Ozturk et al., 2016). An aliquot (20% of the total) of the bead suspension was mixed with SDS-PAGE sample buffer for western blotting after centrifugation to pellet the resin. The remaining beads were used for RNA isolation with TriFast (Peqlab Biotechnologies). cDNA of the input and precipitated fractions

Weber et al.

(20% each) was prepared and analyzed using qPCR (5% of the cDNA), as described above. A list of primers used for the qPCR experiments can be found in Table S3.

Pulldown assays

Co-IP assays were performed in the presence of RNase A as described previously (Peter et al., 2015). HEK293T cells were grown in 10 cm dishes and transfected using Lipofectamine 2000 (Invitrogen) according to the manufacturer's recommendations. The transfection mixtures in Figs. S4a and b contained 1.5 µg of V5-SBP-MBP, 4 µg of V5-SBP-DAP5, and 5 µg of GFP-eIF2β. For the cap pulldown transfection mixtures contained 1 µg GFP-MBP or 12 µg GFP-chimeric-4EBP.

Flow cytometry

Cells were seeded (0.6×10^6 WT and 0.7×10^6 DAP5-null HEK293T cells) in 6-well plates 24 hours before transfection. Transfections were carried out with Lipofectamine 2000 (Invitrogen), with the following transfection mixtures: WNK1-mNG2$_{11}$-EBFP (0.35 µg), mNG2$_{1-10}$ (1 µg), mCherry (10 ng), V5-SBP-MBP (0.25 µg) or DAP5 (0.65 µg). 48 hours after transfection, cells were trypsinized, sedimented (5000 rpm for 3 min at room temperature), resuspended in 1% FBS in PBS, and analyzed using the Becton Dickinson FACSMelody™ Cell Sorter and FlowJo software (Becton Dickison). To determine mNG2, EBFP and mCherry positive events, we analyzed untransfected and control transfected cell batches. Cut-offs were applied uniformly for all measured conditions.

Knockdowns

0.64×10^6 Ctrl cells were transfected with 2 µg pSUPER-puro scramble control or ABCE1 shRNA, after seeding in 6-well plates, using Lipofectamine 2000 (Invitrogen). 24 hours after transfection cells were treated with 5 µM puromycin for 24 hours. Selected cells were re-

seeded and re-transfected with DNA mixtures containing 0.5 µg of *WNK1*-R-LUC reporter plasmids.

Figure legends

Figure 1. DAP5 mediates the synthesis of signaling proteins

(a) Comparative analysis of translation efficiency (TE) in wild type (WT) and DAP5-null HEK293T cells. Genes were plotted as a scatter graph according to changes in ribosome occupancy [$\log_2$FC RFP] on the y axis and mRNA abundance [$\log_2$FC mRNA] on the x axis. Each dot represents an individual gene (n_{total}=9870) selected using FPKM >2. Genes with homodirectional changes in TE are indicated in grey. Genes with increased or decreased TE are highlighted in blue (23 genes) and red (306 genes), respectively.

(b) Gene ontology (GO)-terms associated with the genes with decreased TE in cells lacking DAP5. Bar graph shows -$\log_{10} q$ values for each of the overrepresented category. Values and black circles indicate the % of genes within each category.

(c) Immunoblotting demonstrating the loss of DAP5 in the null cells. In agreement with the observed changes in TE [log2FC(TE)], ROCK1 and WNK1 kinases, and INPPL1 phosphatase protein levels decrease in the absence of DAP5. TUBULIN served as loading control. Blots were probed with antibodies recognizing DAP5, ROCK1, WNK1, INPPL1 and TUBULIN.

(d) Ribosome footprints and total mRNA reads distribution along *WNK1* exon 1 including the 5′ leader and the most 5′ proximal coding sequence in WT and DAP5-null cells. Also shown are the ribosome footprint profiles (RFPs) in HEK293 cells treated with harringtonine and lactimidomycin obtained by Lee and co-workers (Lee et al., 2012). The predicted propensity for secondary structure across *WNK1* 5′ leader, determined using the ViennaRNA package 2.0 (Lorenz et al., 2011), is illustrated in orange in the mRNA panel. uORFs position in the 5′ leader is indicated with the corresponding start codons (GUG, CUG, UUG or AUG). Start codons highlighted in green are in frame with the AUG at the main annotated coding sequence of *WNK1*. Gene annotation is depicted below the profiles. DAP5-independent and -dependent translation is indicated with a black dashed line. CDS: coding sequence.

(e, f) Metagene analyses of ribosome density for the 306 5′ leaders of the DAP5 targets in WT (green) and DAP5-null (blue) cells (e), and for 5′ leaders of DAP5 targets and all expressed transcripts expressed in HEK293 cells treated with harringtonine or lactimidomycin (LTM) (Lee et al., 2012). Ribosome densities were determined as the ratio of footprints within the 5′ leader relative to the footprints at the annotated downstream CDS start codon. The black dashed line indicating DAP5-independent (indep.) and DAP5-dependent translation was defined as the position along the 5′ leaders in which RFP density decreases in the absence of DAP5.

(g, h) Variation of uORF number and length in the 5′ leaders of all DAP5 targets (n=306). Boxes indicate the 25th to 75th percentiles; black line inside the box represents the median; error bars show the outliers.

(i) Start codon usage of uORFs preferentially translated in the absence of DAP5.

Figure 2. 5′ leaders determine DAP5-dependent translation of target mRNAs

(a-c) WT (green) and DAP5-null (blue) HEK293T cells were transfected with different *Renilla* luciferase (R-LUC) reporters that contain the 5′ leader sequences of the *WNK1*, *ROCK1* and *AKT1* mRNAs cloned upstream of R-LUC CDS. Cells were also transfected with the normalization and transfection control F-LUC-GFP that contains a short 5′ leader sequence. In addition, the plasmids expressing V5-SBP tagged maltose binding protein (MBP), DAP5 [full length (FL) or the indicated mutants], eIF4G [full length (FL) or the indicated mutants], or DAP5-eIF4G chimeric proteins were also present in the transfection mixture. R-LUC activity was quantified in WT and DAP5-null cells in the presence of the different proteins two days post-transfection, normalized over to that of F-LUC-GFP and set to 100% in WT cells. The mean values +/- SD of three independent experiments are shown. Schematic representations of the R-LUC reporters are presented above each graph. Proteins are as follow: eIF4A*: eIF4A-binding mutant; MIF4G: DAP5 MIF4G domain; ΔW2: deletion of the DAP5 W2 domain; ΔN: deletion of eIF4G N-terminal region; MIF4G chimera: eIF4G MIF4G domain swapped into

Weber et al.

DAP5; MA3 chimera: eIF4G MA3 domain swapped into DAP5; W2 chimera: eIF4G W2 domain swapped into DAP5.

(d) Schematic representation of the DAP5, eIF4G and DAP5-eIF4G chimeras. PABP: poly(A)-binding protein-binding region; 4E-BM: eIF4E-binding motif; MIF4G: middle eIF4G domain; MA3: MA3 domain; W2: W2 domain; eIF4A-BR: eIF4A-binding region. eIF2β-BR: eIF2β-binding region; MNKs-BR: MNK1 and MNK2-binding region. The amino acid positions at the domain/motif boundaries are indicated below the proteins. The MIF4G domains of DAP5 and eIF4G, and the MA3 domain of eIF4G bind to eIF4A. The W2 domain of DAP5 interacts with eIF2β whereas the corresponding domain in eIF4G associates with the MNK kinases 1 and 2.

(e-g) HEK293T cells were transfected with plasmids expressing V5-SBP-tagged proteins: MBP, DAP5 (FL or the indicated mutants), eIF4G (FL or the indicated mutants), or DAP5-eIF4G chimeras. Streptavidin pulldown assays were performed two days post transfection and protein and RNA samples were obtained for each experimental condition. *WNK1*, *ROCK1* and *GAPDH* mRNA levels in input (0.8 %) and pulldown samples (12 %) were quantified by quantitative PCR (qPCR) following reverse transcription. Values were set to 100% for V5-SBP-MBP. The mean values +/- SD of three independent experiments are shown.

(h) Immunoblot depicting the expression and the pulldown efficiency of the V5-SBP tagged proteins used in e-g. Membranes were probed with anti-V5 antibody.

Figure 3. DAP5-dependent translation requires eIF4F-mediated ribosome recruitment

(a) Schematic representations of the WNK1-R-LUC reporters with cap-proximal deletions in the 5′ leader that partially (Δ1) or completely (Δ2) remove the structured region containing the uORFs translated in a DAP5-independent manner. The predicted propensity for secondary structure across *WNK1* 5′ leader, determined using the ViennaRNA package 2.0 (Lorenz et al., 2011), is illustrated in orange. uORFs position in the 5′ leader is indicated with the

corresponding start codons (GUG, CUG, UUG or AUG). Start codons highlighted in green are in frame with the AUG at the main annotated coding sequence of *WNK1*.

(b) WT and DAP5-null cells were transfected with the *WNK1*-R-LUC reporters represented in (a), F-LUC-GFP, and V5-SBP-MBP or V5-SBP-DAP5. Following transfection, luciferase activities (Protein) were measured and mRNA levels determined by northern blotting. R-LUC values were normalized to the transfection control F-LUC-GFP. The graph shows the protein and mRNA ratios in WT and null cells, set to 100% in WT cells expressing the *WNK1*-R-LUC reporter.

(c) m^{7}GTP-cap pulldown assay showing the interaction between eIF4E and eIF4G in the presence or absence of GFP-4EBP. Inputs (1% for eIF4E and 0.3% for eIF4G and GFP-tagged proteins) and bound fractions (1% for eIF4E and 2% for eIF4G and the GFP-tagged proteins) were analysed by western blotting. Membranes were probed with anti-eIF4E, eIF4G and GFP antibodies.

(d, e) Binding of V5-SBP-DAP5, or V5-SBP-MBP as control, to *WNK1* and *ROCK1* mRNAs was determined by RNA-immunoprecipitation in the presence or absence of GFP-4EBP. Proteins were pulled down using streptavidin beads. mRNA levels in input (0.8%) and IP samples (12%) were quantified by RT-qPCR and set to 100% in for V5-SBP-MBP. Bars indicate the mean value; error bars represent SD (n=3).

Figure 4. DAP5 mediates re-initiation following uORF translation

(a) Schematic representations of the *WNK1*-R-LUC reporters with changes in uORF2 initiation context and length. uORF2 GUG is in frame with the R-LUC AUG and is 22 codons long. Three STOP codons can be found downstream and in frame with uORF2 GUG. uORF2+: GUG start codon was substituted by AUG to favour the initiation of translation. ΔSTOP1: first STOP codon in frame with uAUG is removed; uORF is then 188 codons long. ΔSTOP1+2: reporter

lacks the two STOP codons after uAUG; length of uORF2 increases to 229 codons. NO STOP: the three STOP codons are absent; the reporter produces a N-terminally extended R-LUC.

(b-e) WT and DAP5-null cells were transfected with the different *WNK1*-R-LUC reporters shown in a, F-LUC-GFP and V5-SBP-MBP or V5-SBP-DAP5. After transfection, R-LUC activity was measured, normalized to F-LUC-GFP and set to 100% for WT expressing *WNK1*-R-LUC (b). Expression of short and long (N-terminally extended) R-LUC was also evaluated by immunoblotting, together with V5-tagged proteins, F-LUC-GFP and TUBULIN (c) using anti-R-LUC, V5, GFP and TUBULIN antibodies. The abundance of the different *R-LUC* and *F-LUC* reporter mRNAs was assessed by northern blotting and quantified as in b (d, e). Bars indicate the mean value; error bars represent SD (n=3).

Figure 5. uORF length is critical for re-initiation of translation by DAP5

(a) Schematic representations of the WNK1-R-LUC reporters with changes in uORF5 length. uORF5 initiates from an AUG start codon in the –1 frame and encodes a short peptide (4 amino acids). uORF118: UGA STOP codon was removed changing the length of uORF5 to 118 codons. uORF30, uORF19, uORF9: position of the STOP codon was moved to 30, 19 or 9 codons downstream of uAUG, respectively. ΔSTOP IF (in-frame): all STOP codons in frame with uAUG were removed. ΔSTOP all F (frames): the *WNK1* 5′ leader lacks STOP codons.

(b, c) WT and DAP5-null cells were transfected with different *WNK1*-R-LUC reporters, F-LUC-GFP and V5-SBP-MBP or V5-SBP-DAP5. Following transfection, luciferase activities (Protein) were measured and mRNA levels determined by northern blotting. R-LUC values were normalized to the transfection control F-LUC-GFP. The graph shows the protein and mRNA ratios in WT and null cells, set to 100% in WT cells expressing *WNK1*-R-LUC. See also Fig. S5. The immunoblot showing the expression of the different proteins is shown in panel c. TUBULIN served as a loading control.

Weber et al.

(d) Western blot depicting the expression of short and long (N-terminally extended) R-LUC proteins produced in WT and DAP5-null cells expressing the *WNK1*-R-LUC ΔSTOP IF and *WNK1*-R-LUC ΔSTOP all F. F-LUC-GFP served as a transfection control.

(e, f) Metagene analyses of 40S (e) and 80S (f) footprints along the 5′ leaders of DAP5 targets (n=306) and all other transcripts expressed in cells. 40S selective profiling data was performed in HeLa cells by Bohlen and co-workers (Bohlen et al., 2020). 80S footprints were retrieved from the ribosome profiling data generated in this study with HEK293T cells. 40S and 80S densities were determined as the ratio of footprints within the 5′ leader relative to the footprints at the annotated downstream CDS start codon.

Figure 6. Concurrent uORF and main CDS translation in DAP5 targets

(a) Schematic representation of the mNeonGreen2 (mNG2) split-fluorescent protein approach and corresponding reporter constructs. Co-expression of the $mNG2_{1-10}$ and $mNG2_{11}$ fragments originates a functional mNG2 fluorescent molecule (Chen et al., 2020; Feng et al., 2017; Leonetti et al., 2016). $mNG2_{11}$ CDS was inserted in *WNK1* 5′ leader and replaced uORF2. $mNG2_{11}$ translation initiates at a uAUG in frame with the main CDS and produces a 16 aa protein. Main CDS encoded the EBFP fluorophore. ΔSTOP1: The first UAG STOP codon after the uAUG was removed and the $mNG2_{11}$ CDS was inserted next to the a UAG STOP located 188 codons downstream of the uAUG.

(b) WT and DAP5-null cells were transfected with the $mNG2_{1-10}$, *WNK1*-$mNG2_{11}$-EBFP, mCherry, and V5-SBP-MBP or V5-SBP-DAP5 plasmids. Following transfection, cells were collected and analyzed by flow cytometry. The histogram shows the EBFP signal intensity in mNG2-positive cells in the presence (WT, null+DAP5) or absence (DAP5-null) of DAP5.

(c, d) Box plots of the EBFP:mNG2 and mNG2:mCherry (d) ratios quantified by flow cytometry of WT, DAP5-null cells and null cells following V5-SBP-DAP5 re-expression. Cells expressed the $mNG2_{1-10}$, *WNK1*-$mNG2_{11}$-EBFP and mCherry reporters. Boxes represent the

25th to 75th percentiles; black line shows the median and the cross the average; error bars represent outliners. Significance was determined by one-sided Wilcoxon rank-sum test and indicated if p<0.05. null+DAP5: DAP5-null cells re-expressing V5-SBP-DAP5.

(e, f) The box plots indicate the EBFP:mNG2 (e) and mNG2:mCherry (f) quantified by flow cytometry of WT and DAP5-null cells transfected with the mNG2$_{1-10}$, *WNK1*-ΔSTOP1-mNG2$_{11}$-EBFP, mCherry, and V5-SBP-MBP or V5-SBP-DAP5 plasmids. Boxes represent the 25th to 75th percentiles; black line shows the median and the cross the average; error bars represent outliners. Significance determined by one-sided Wilcoxon rank-sum test and indicated if p<0.05. null+DAP5: DAP5-null cells re-expressing V5-SBP-DAP5.

Figure 7. DAP5 uses post-termination translation complexes

(a) Schematic representation of *WNK1*-R-LUC reporters with changes in uORF2 length, as described in Fig. 4.

(b, c) HEK293T cells were transfected with the *WNK1*-R-LUC reporters shown in a. Additionally, the transfection mixtures also contained F-LUC-GFP and increasing concentrations of λN-HA-eRF1^{AAQ}. (b) Immunoblot showing the expression levels of the transfected proteins. The membranes were blotted with anti-R-LUC, HA, GFP, DAP5 and TUBULIN antibodies. (c) R-LUC activity was measured, normalized to F-LUC-GFP and set to 100% in the absence of λN-HA-eRF1^{AAQ} for each reporter. Bars indicate the mean value; error bars represent SD (n=3).

(d) Western blots showing shRNA-mediated depletion of ABCE1 in HEK293T cells. TUBULIN served as a loading control.

(e) UV absorbance profile at 254 nM of scramble shRNA (control) and ABCE1-depleted (ABCE1 shRNA) HEK293T cell extracts after polysome sedimentation in a sucrose gradient. 40S and 60S subunits, 80S monosomes, and polysome peaks are indicated.

Weber et al.

(f, g) HEK293T cells were treated with scramble (Scr) or shRNA targeting *ABCE1* mRNA and transfected with the *WNK1*-R-LUC reporters shown in a. (f) The graph shows relative R-LUC activity in control (Scr) and ABCE1 KD cells. R-LUC activity was normalized to that of F-LUC-GFP and set to 100% in Scr-treated cells for each reporter. (g) Immunoblot illustrating the expression of short and long (N-terminally extended) R-LUC, F-LUC-GFP and TUBULIN in control and ABCE1-depleted cells. Blots were probed with anti-R-LUC, GFP and TUBULIN antibodies.

Supplemental Figure legends

Figure S1, related to Figure 1. Characterization of DAP5-null cells

(a, b) Ribosome footprints and total mRNA reads distribution along *EIF4G2* (DAP5) mRNA in wild type (WT) and DAP5-null cells. Of note, RFP and total RNA counts for *EIF4G2* are drastically reduced in the null cells. Dashed box indicates the position of the CRISPR-Cas9 edited region. In panel b, read counts scale for total RNA in DAP5-null cells was increased to show the presence of reads in intron 9 as a result of genome editing. SNORD97 (small nucleolar RNA, C/D Box 97) is encoded in intron 13 of *EIF4G2*.

(c) Western blot demonstrating the loss of DAP5 expression in the null cells. eIF4F subunits eIF4E, eIF4G and eIF4A expression is not decreased in the absence of DAP5. eIF4G and PABP protein levels are even increased in the null cells. TUBULIN served as a loading control. TUBULIN antibody recognizes an epitope common among the α-TUBULIN subunits.

(d) Comparative analysis of changes in ribosome occupancy [$\log_2$FC RFP] on the y axis and mRNA abundance [$\log_2$FC mRNA] on the x axis in wild type (WT) versus DAP5-null HEK293T cells as described in Figure 1a. Genes with increased (n=1771) and decreased (n=1766) mRNA abundance and thus ribosome occupancy are highlighted in purple and orange, respectively.

(e, f) Multidimensional scaling (MDS) analysis for the Ribo-Seq (e) and RNA-Seq (f) replicate libraries from HEK293T wild type (WT) and DAP5-null cells.

(g) UV absorbance profile at 254 nm of HEK293T WT (green) and DAP5-null cells (blue) cell extracts after polysome sedimentation in a sucrose gradient. Absorbance peaks at 254 nm representing free 40S, 60S, 80S monosomes and polysomes are indicated.

(h, i) Ethidium bromide staining of total RNA extracted from the different sucrose fractions. Due to their high cellular abundance, 28S and 18S rRNAs positions in the gel are readily detected.

(j-l) Abundance profiles for *WNK1* (j), *ROCK1* (k) and *GAPDH* (l) mRNAs across the density gradient in WT (green) and DAP5-null (blue) cells. mRNA abundance was determined by quantitative PCR (qPCR). Bars represent the mean value; error bars represent standard deviations (SD) (n=3).

Figure S2, related to Figure 1. Ribosome densities and mRNA read counts in DAP5 targets

(a-j) Ribosome footprints and total mRNA reads distribution along different DAP5 target mRNAs in wild type (WT) and DAP5-null cells. The predicted propensity for secondary structure across the 5′ leaders, determined using the ViennaRNA package 2.0 (Lorenz et al., 2011), is illustrated in red. Gene annotation is depicted below the profiles. DAP5-independent (indep.) and -dependent translation is indicated with a black dashed line.

Figure S3, related to Figures 1 and 2. DAP5-dependent translation is determined by structured 5′ leaders

(a-c) Histograms show the distribution of $\log_{10}$ length (nts), the GC content (%) and the minimum free energy (ΔG/ nts) in the 5′ leaders of the transcripts with decreased translation efficiency (TE down; red) and in all other mRNAs expressed in HEK293T cells (all/10; blue). The number of mRNAs all other mRNAs expressed in HEK293T cells is divided by 10.

Weber et al.

Statistical significance was calculated with the Wilcoxon rank sum test. Bin width, 0.2 in a, 5% in b and 0.5 in c.

(d-i) WT and DAP5-null cells were transfected with plasmids expressing *WNK1-*, *ROCK1-* or *AKT1-*R-LUC, V5-SBP-MBP, V5-SBP-DAP5 (FL or mutants), V5-SBP-eIF4G (FL or ΔN) or V5-SBP-Chimeras. *R-LUC* mRNA levels were determined by northern blotting, normalized to *F-LUC-GFP* and set to 100% in WT cells. Bars represent the mean value; error bars represent SD (n=3).

(j-l) Immunoblot depicting the expression of the proteins used in Figs. 2a-c.

Figure S4, related to Figure 2. DAP5-dependent translation requires the RNA helicase eIF4A

(a, b) Streptavidin-binding protein (SBP) affinity pulldowns were performed two days post cell transfection with SBP-V5-MBP or V5-SBP-DAP (FL or mutants) and GFP-eIF2β (b). Input (1% for the V5 proteins, 0.3% for eIF4A, eIF4A2 and GFP- eIF2β) and pulldown fractions (1% for the V5 proteins, 2% for eIF4A, eIF4A2 and GFP- eIF2β) were analysed by western blotting with anti-V5, anti-eIF4A, anti-eIF4A2 or anti-GFP antibodies.

(c) Venn diagram showing the number of common (n=102; p=5.1995e^{-97} using a hypergeometric test) and unique genes with decreased TE in DAP5-null cells or in cells treated with 0.003 μM of Rocaglamide A (RocA) (Iwasaki et al., 2016).

(d) Metagene analysis of ribosome density for the 306 5′ leaders of the DAP5 targets in WT (green), DAP5-null (dark blue), and RocA-treated cells (light blue) (Iwasaki et al., 2016). Ribosome densities were determined as the ratio of footprints within the 5′ leader relative to the footprints at the annotated downstream CDS start codon. The black dashed line indicating DAP5-independent (indep.) and DAP5-dependent translation was defined as the position along the 5′ leaders in which RFP density decreases in the absence of DAP5.

Weber et al.

(e) Ribosome footprints and total mRNA reads distribution along *WNK1* exon 1 including the 5′ leader and the most 5′ proximal coding sequence in WT and DAP5-null cells. Also shown are the ribosome footprint profiles (RFPs) in HEK293 cells treated with harringtonine and lactimidomycin obtained by Lee and co-workers (Lee et al., 2012) and in HEK293 cells upon treatment with 0.003 μM of RocA (Iwasaki et al., 2016). The predicted propensity for secondary structure across *WNK1* 5′ leader, determined using the ViennaRNA package 2.0 (Lorenz et al., 2011), is illustrated in red. uORFs position in the 5′ leader is indicated with the corresponding start codons. Start codons highlighted in green are in frame with the AUG at the main annotated coding sequence of *WNK1*. Gene annotation is depicted below the profiles. DAP5-independent and -dependent translation is indicated with a black dashed line. CDS: coding sequence.

(f-i) WT and DAP5-null cells were transfected with plasmids expressing *WNK1*-R-LUC reporters, V5-SBP-MBP or V5-SBP-DAP5. (f) Following transfection, luciferase activities were measured (f) and mRNA levels determined by northern blotting (g, h). R-LUC activity and mRNA levels were normalized to the transfection control F-LUC-GFP and set to 100% in WT cells. Bars represent the mean value; error bars represent SD (n=3). Representative northern blots are shown in h. (i) The immunoblot shows the expression levels of the proteins used in the assay depicted in Fig. 3f. Membranes were incubated with anti-V5, GFP and TUBULIN.

Figure S5, related to Figure 4 and 5. Short uORFs support DAP5-dependent re-initiation

(a) Schematic representations of the *WNK1*-R-LUC reporters with changes in uORF2 initiation context and length. uORF2 GUG is in frame with the R-LUC and is 22 codons long. Three STOP codons can be found downstream and in frame with uORF2 GUG. uORF2+: GUG start codon was substituted by AUG to favour the initiation of translation. ΔSTOP1: first STOP codon in frame with uAUG was removed; uORF is then 188 codons long. uORF49, uORF39,

uORF29: position of the STOP codon was moved to 49, 39 or 29 codons downstream of uAUG, respectively.

(b-e) WT and DAP5-null cells were transfected with different *WNK1*-R-LUC reporters, F-LUC-GFP and V5-SBP-MBP or V5-SBP-DAP5. Following transfection, luciferase activities (b) were measured and mRNA levels determined by northern blotting (c, d). R-LUC values were normalized to the transfection control F-LUC-GFP. The graphs show the protein (b) and mRNA levels in WT and null cells (c, d), set to 100% in WT cells expressing *WNK1*-R-LUC. (e) The immunoblot showing the expression of the different proteins is shown in panel d. TUBULIN served as a loading control.

(f-h) WT and DAP5-null cells were transfected with different *WNK1*-R-LUC reporters, F-LUC-GFP and V5-SBP-MBP or V5-SBP-DAP5. Following transfection, luciferase activities (Protein) were measured and mRNA levels determined by northern blotting (c, d). R-LUC values were normalized to the transfection control F-LUC-GFP. The graphs show the luciferase activity (f) and the mRNA levels (g) in WT and null cells, set to 100% in WT cells expressing *WNK1*-R-LUC. (h) Representative northern blot is present in h. See also Fig. 5.

Figure S6. Detection of uORF translation in DAP5 targets 5′ leaders

(a) Density curves of the mNG2 fluorescence quantified by flow cytometry of control transfected (black trace) HEK293T cells or expressing $mNG2_{1-10}$ (light blue trace), $mNG2_{11}$-BFP (pink trace, and $mNG2_{1-10}$ (increasing amounts; green-yellow traces) + $mNG2_{11}$-BFP. mNG2 expression is plotted on a log scale and represents around 200000 cells.

(b) Density curves of the EBFP fluorescence quantified by flow cytometry of control transfected (black trace) HEK293T cells or expressing $mNG2_{1-10}$ (light blue trace) or $mNG2_{11}$-BFP (pink trace). EBFP expression is plotted on a log scale and represents around 200000 cells.

(c) Density curves of the mCherry fluorescence quantified by flow cytometry of control transfected (black trace) HEK293T cells or WT (green trace) and DAP5-null cells expressing

mCherry and V5-SBP-MBP (blue trace) or V5-SBP-DAP5 (yellow trace). mCherry expression is plotted on a log scale and represents around 200000 cells.

(d, e) WT cells expressing increasing concentrations of eRF1^{AAQ} were transfected with different *WNK1*-R-LUC reporters and F-LUC-GFP. Following transfection, luciferase mRNA levels were determined by northern blotting. R-LUC values were normalized to the transfection control F-LUC-GFP. The graph shows the luciferase mRNA levels in cells and set to 100% in the absence of eRF1^{AAQ}. A representative northern blot is present in e. See also Fig. 7.

(f, g) Scramble (Scr) and ABCE1 shRNA-treated cells were transfected with different *WNK1*-R-LUC reporters and F-LUC-GFP. Following transfection, luciferase mRNA levels were determined by northern blotting. R-LUC values were normalized to the transfection control F-LUC-GFP and set to 100% in control knockdown cells. A representative northern blot is present in g. See also Fig. 7.

Figure S7, related to Figure 7. Canonical translation termination and 60S recycling precedes re-initiation of translation by DAP5

(a) Schematic representation of *WNK1*-R-LUC reporters with changes in uORF2 length, as described in Fig. 4.

(b, c) Western blots showing shRNA-mediated depletion of DENR, MCTS-1 and eIF2D or eIF3J in HEK293T cells. TUBULIN served as a loading control.

(d, e) HEK293T cells were treated with scramble (Scr) or shRNA targeting *DENR*, *MCTS-1* and *eIF2D* mRNAs and transfected with the *WNK1*-R-LUC reporters shown in a. The graph shows relative R-LUC activity in control (Scr) and DENR+MCTS-1+eIF2D KD cells. R-LUC activity was normalized to that of F-LUC-GFP and set to 100% in Scr-treated cells for each reporter. The immunoblot illustrating the expression of short and long (N-terminally extended) R-LUC, F-LUC-GFP and TUBULIN in control and DENR+MCTS-1+eIF2D-depleted cells is depicted in e. Blots were probed with anti-R-LUC, GFP and TUBULIN antibodies.

(f, g) HEK293T cells were treated with scramble (Scr) or shRNA targeting *eIF3J* mRNAs and transfected with the *WNK1*-R-LUC reporters shown in a. The graph shows relative R-LUC activity in control (Scr) and eIF3J KD cells. R-LUC activity was normalized to that of F-LUC-GFP and set to 100% in Scr-treated cells for each reporter. A representative northern blot of the *WNK1*-R-LUC in cells with and without eIF3J is shown in g.

(h) HEK293T wild type and null cells in the presence or absence of Thapsigargin. The immunoblot shows the expression of ATF4, DAP5, eIF2α–P. TUBULIN served as loading control.

Weber et al.

References

Ahmed, Y.L., Schleich, S., Bohlen, J., Mandel, N., Simon, B., Sinning, I., and Teleman, A.A. (2018). DENR-MCTS1 heterodimerization and tRNA recruitment are required for translation reinitiation. PLoS Biol *16*, e2005160.

Barbosa, C., Peixeiro, I., and Romao, L. (2013). Gene expression regulation by upstream open reading frames and human disease. PLoS Genet *9*, e1003529.

Bohlen, J., Fenzl, K., Kramer, G., Bukau, B., and Teleman, A.A. (2020). Selective 40S Footprinting Reveals Cap-Tethered Ribosome Scanning in Human Cells. Mol Cell.

Brown, A., Shao, S., Murray, J., Hegde, R.S., and Ramakrishnan, V. (2015). Structural basis for stop codon recognition in eukaryotes. Nature *524*, 493-496.

Calviello, L., Mukherjee, N., Wyler, E., Zauber, H., Hirsekorn, A., Selbach, M., Landthaler, M., Obermayer, B., and Ohler, U. (2016). Detecting actively translated open reading frames in ribosome profiling data. Nat Methods *13*, 165-170.

Chen, J., Brunner, A.D., Cogan, J.Z., Nunez, J.K., Fields, A.P., Adamson, B., Itzhak, D.N., Li, J.Y., Mann, M., Leonetti, M.D.*, et al.* (2020). Pervasive functional translation of noncanonical human open reading frames. Science *367*, 1140-1146.

Chew, G.L., Pauli, A., and Schier, A.F. (2016). Conservation of uORF repressiveness and sequence features in mouse, human and zebrafish. Nat Commun *7*, 11663.

Child, S.J., Miller, M.K., and Geballe, A.P. (1999). Translational control by an upstream open reading frame in the HER-2/neu transcript. J Biol Chem *274*, 24335-24341.

de la Parra, C., Ernlund, A., Alard, A., Ruggles, K., Ueberheide, B., and Schneider, R.J. (2018). A widespread alternate form of cap-dependent mRNA translation initiation. Nat Commun *9*, 3068.

Durinck, S., Moreau, Y., Kasprzyk, A., Davis, S., De Moor, B., Brazma, A., and Huber, W. (2005). BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. Bioinformatics *21*, 3439-3440.

Weber et al.

Durinck, S., Spellman, P.T., Birney, E., and Huber, W. (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nat Protoc *4*, 1184-1191.

Feng, S., Sekine, S., Pessino, V., Li, H., Leonetti, M.D., and Huang, B. (2017). Improved split fluorescent proteins for endogenous protein labeling. Nat Commun *8*, 370.

Frolova, L.Y., Tsivkovskii, R.Y., Sivolobova, G.F., Oparina, N.Y., Serpinsky, O.I., Blinov, V.M., Tatkov, S.I., and Kisselev, L.L. (1999). Mutations in the highly conserved GGQ motif of class 1 polypeptide release factors abolish ability of human eRF1 to trigger peptidyl-tRNA hydrolysis. RNA *5*, 1014-1020.

Gaidatzis, D., Lerch, A., Hahne, F., and Stadler, M.B. (2015). QuasR: quantification and annotation of short reads in R. Bioinformatics *31*, 1130-1132.

Henis-Korenblit, S., Strumpf, N.L., Goldstaub, D., and Kimchi, A. (2000). A novel form of DAP5 protein accumulates in apoptotic cells as a result of caspase cleavage and internal ribosome entry site-mediated translation. Mol Cell Biol *20*, 496-506.

Ingolia, N.T., Brar, G.A., Rouskin, S., McGeachy, A.M., and Weissman, J.S. (2012). The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. Nat Protoc *7*, 1534-1550.

Ingolia, N.T., Lareau, L.F., and Weissman, J.S. (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. Cell *147*, 789-802.

Iwasaki, S., Floor, S.N., and Ingolia, N.T. (2016). Rocaglates convert DEAD-box protein eIF4A into a sequence-selective translational repressor. Nature *534*, 558-561.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol *14*, R36.

Kozak, M. (2002). Pushing the limits of the scanning mechanism for initiation of translation. Gene *299*, 1-34.

Kuzuoglu-Ozturk, D., Bhandari, D., Huntzinger, E., Fauser, M., Helms, S., and Izaurralde, E. (2016). miRISC and the CCR4-NOT complex silence mRNA targets independently of 43S ribosomal scanning. EMBO J *35*, 1186-1203.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods *9*, 357-359.

Lee, S., Liu, B., Lee, S., Huang, S.X., Shen, B., and Qian, S.B. (2012). Global mapping of translation initiation sites in mammalian cells at single-nucleotide resolution. Proc Natl Acad Sci U S A *109*, E2424-2432.

Lee, S.H., and McCormick, F. (2006). p97/DAP5 is a ribosome-associated factor that facilitates protein synthesis and cell proliferation by modulating the synthesis of cell cycle proteins. EMBO J *25*, 4008-4019.

Leonetti, M.D., Sekine, S., Kamiyama, D., Weissman, J.S., and Huang, B. (2016). A scalable strategy for high-throughput GFP tagging of endogenous human proteins. Proc Natl Acad Sci U S A *113*, E3501-3508.

Lewis, S.M., Cerquozzi, S., Graber, T.E., Ungureanu, N.H., Andrews, M., and Holcik, M. (2008). The eIF4G homolog DAP5/p97 supports the translation of select mRNAs during endoplasmic reticulum stress. Nucleic Acids Res *36*, 168-178.

Liberman, N., Gandin, V., Svitkin, Y.V., David, M., Virgili, G., Jaramillo, M., Holcik, M., Nagar, B., Kimchi, A., and Sonenberg, N. (2015). DAP5 associates with eIF2beta and eIF4AI to promote Internal Ribosome Entry Site driven translation. Nucleic Acids Res *43*, 3764-3775.

Liberman, N., Marash, L., and Kimchi, A. (2009). The translation initiation factor DAP5 is a regulator of cell survival during mitosis. Cell Cycle *8*, 204-209.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods *25*, 402-408.

Lorenz, R., Bernhart, S.H., Honer Zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P.F., and Hofacker, I.L. (2011). ViennaRNA Package 2.0. Algorithms Mol Biol *6*, 26.

Marash, L., and Kimchi, A. (2005). DAP5 and IRES-mediated translation during programmed cell death. Cell Death Differ *12*, 554-562.

Marash, L., Liberman, N., Henis-Korenblit, S., Sivan, G., Reem, E., Elroy-Stein, O., and Kimchi, A. (2008). DAP5 promotes cap-independent translation of Bcl-2 and CDK1 to facilitate cell survival during mitosis. Mol Cell *30*, 447-459.

McCarthy, D.J., Chen, Y., and Smyth, G.K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res *40*, 4288-4297.

Nevins, T.A., Harder, Z.M., Korneluk, R.G., and Holcik, M. (2003). Distinct regulation of internal ribosome entry site-mediated translation following cellular stress is mediated by apoptotic fragments of eIF4G translation initiation factor family members eIF4GI and p97/DAP5/NAT1. J Biol Chem *278*, 3572-3579.

Peter, D., Igreja, C., Weber, R., Wohlbold, L., Weiler, C., Ebertsch, L., Weichenrieder, O., and Izaurralde, E. (2015). Molecular architecture of 4E-BP translational inhibitors bound to eIF4E. Mol Cell *57*, 1074-1087.

Pisarev, A.V., Hellen, C.U., and Pestova, T.V. (2007). Recycling of eukaryotic posttermination ribosomal complexes. Cell *131*, 286-299.

Poyry, T.A., Kaminski, A., and Jackson, R.J. (2004). What determines whether mammalian ribosomes resume scanning after translation of a short upstream open reading frame? Genes Dev *18*, 62-75.

Ramirez, F., Ryan, D.P., Gruning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dundar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res *44*, W160-165.

Weber et al.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat Protoc *8*, 2281-2308.

Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat Biotechnol *29*, 24-26.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139-140.

Rodan, A.R., and Jenny, A. (2017). WNK Kinases in Development and Disease. Curr Top Dev Biol *123*, 1-47.

Roy, B., Vaughn, J.N., Kim, B.H., Zhou, F., Gilchrist, M.A., and Von Arnim, A.G. (2010). The h subunit of eIF3 promotes reinitiation competence during translation of mRNAs harboring upstream open reading frames. RNA *16*, 748-761.

Schleich, S., Acevedo, J.M., Clemm von Hohenberg, K., and Teleman, A.A. (2017). Identification of transcripts with short stuORFs as targets for DENR*MCTS1-dependent translation in human cells. Sci Rep *7*, 3722.

Schleich, S., Strassburger, K., Janiesch, P.C., Koledachkina, T., Miller, K.K., Haneke, K., Cheng, Y.S., Kuechler, K., Stoecklin, G., Duncan, K.E.*, et al.* (2014). DENR-MCT-1 promotes translation re-initiation downstream of uORFs to control tissue growth. Nature *512*, 208-212.

Shao, S., Murray, J., Brown, A., Taunton, J., Ramakrishnan, V., and Hegde, R.S. (2016). Decoding Mammalian Ribosome-mRNA States by Translational GTPase Complexes. Cell *167*, 1229-1240 e1215.

Sugiyama, H., Takahashi, K., Yamamoto, T., Iwasaki, M., Narita, M., Nakamura, M., Rand, T.A., Nakagawa, M., Watanabe, A., and Yamanaka, S. (2017). Nat1 promotes translation of specific proteins that induce differentiation of mouse embryonic stem cells. Proc Natl Acad Sci U S A *114*, 340-345.

Weber et al.

Thorvaldsdottir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform *14*, 178-192.

Weingarten-Gabbay, S., Khan, D., Liberman, N., Yoffe, Y., Bialik, S., Das, S., Oren, M., and Kimchi, A. (2014). The translation initiation factor DAP5 promotes IRES-driven translation of p53 mRNA. Oncogene *33*, 611-618.

Yoffe, Y., David, M., Kalaora, R., Povodovski, L., Friedlander, G., Feldmesser, E., Ainbinder, E., Saada, A., Bialik, S., and Kimchi, A. (2016). Cap-independent translation by DAP5 controls cell fate decisions in human embryonic stem cells. Genes Dev *30*, 1991-2004.

Young, D.J., and Guydosh, N.R. (2019). Hcr1/eIF3j Is a 60S Ribosomal Subunit Recycling Accessory Factor In Vivo. Cell Rep *28*, 39-50 e34.

Young, D.J., Makeeva, D.S., Zhang, F., Anisimova, A.S., Stolboushkina, E.A., Ghobakhlou, F., Shatsky, I.N., Dmitriev, S.E., Hinnebusch, A.G., and Guydosh, N.R. (2018). Tma64/eIF2D, Tma20/MCT-1, and Tma22/DENR Recycle Post-termination 40S Subunits In Vivo. Mol Cell *71*, 761-774 e765.

Young, M.D., Wakefield, M.J., Smyth, G.K., and Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biol *11*, R14.

Zhong, Y., Karaletsos, T., Drewe, P., Sreedharan, V.T., Kuo, D., Singh, K., Wendel, H.G., and Ratsch, G. (2017). RiboDiff: detecting changes of mRNA translation efficiency from ribosome footprints. Bioinformatics *33*, 139-141.

Weber *et al.* Fig. 1

a DAP5-null cells

TE down (red)
TE up (blue)

n=23
n=306
WNK1

log$_2$FC(DAP5-null/WT) RFP (y-axis)
log$_2$FC(DAP5-null/WT) mRNA (x-axis)

b

Gene ontology (y-axis) vs -log$_{10}$(q-value) (x-axis)

- signaling — 55%
- cell communication — 54%
- cell surface receptor signaling pathway — 30%
- cellular response to stimulus — 57%
- regulation of response to stimulus — 37%
- regulation of signaling — 34%
- system development — 40%
- regulation of kinase activity — 13%
- protein kinase activity — 13%

c

DAP5
cells: WT | null   kDa   log$_2$FC(TE)

DAP5 — 100
ROCK1 — 130 — -2.0
WNK1 — 250 — -2.6
INPPL1 — 130 — -1.5
TUBULIN — 55

1   2

d

DAP5-independent | DAP5-dependent

reads (y-axis)

RFP WT (green) — 16
RFP DAP5-null (blue) — 16
RFP Lactimid. (red) — 30 — Lee et al. (2012)
RFP Harring. (red) — 30
RNA (gray) — 100

5' — 5' leader — *WNK1* CDS — 3'

uORF1: GUG
uORF2: GUG
uORF3: CUG
uORF4: UUG
uORF5: AUG
AUG

e DAP5 targets 5' leader

DAP5-indep. | DAP5-dependent

relative RFP density (y-axis)

WT (green)
DAP5-null (blue)

5' — scaled 5' leader — Start main ORF

f DAP5 targets 5' leader

DAP5-indep. | DAP5-dependent

relative RFP density (y-axis)

DAP5 targets + Harringt. (red)
DAP5 targets + LTM (red)
All transcripts + Harringt. (purple)
All transcripts + LTM (blue)

5' — scaled 5' leader — Start main ORF

g uORFs

Number of uORFs (y-axis)

h uORFs

Length of uORFs (amino acids) (y-axis)

i uORFs

- AUG — 16%
- CUG — 24%
- GUG — 17%
- UUG — 13%
- AUC — 8%
- other — 22%

Weber *et al.* Fig. 2

a *WNK1*-R-LUC
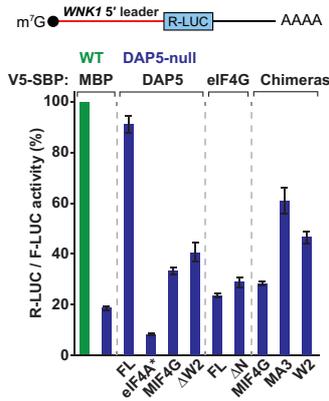
b *ROCK1*-R-LUC
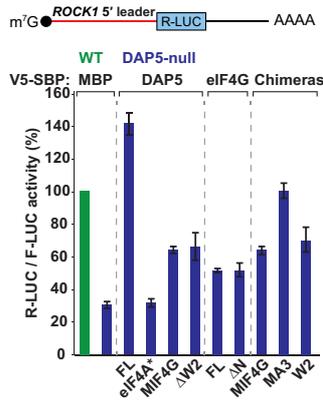
c *AKT1*-R-LUC
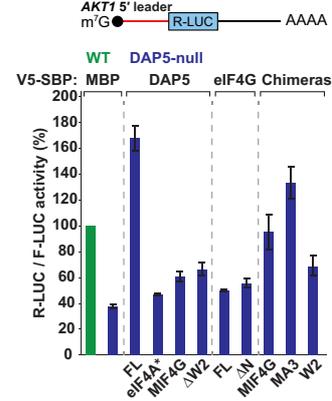
d
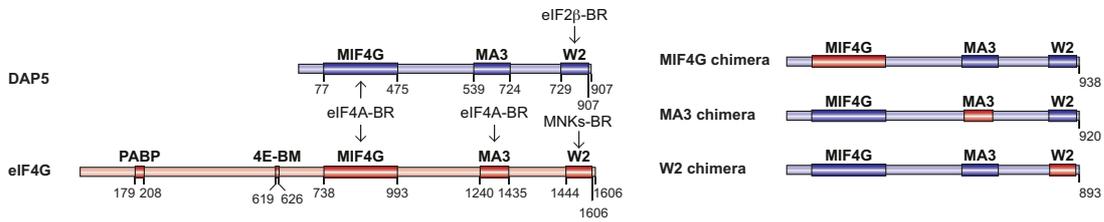
e *WNK1* mRNA
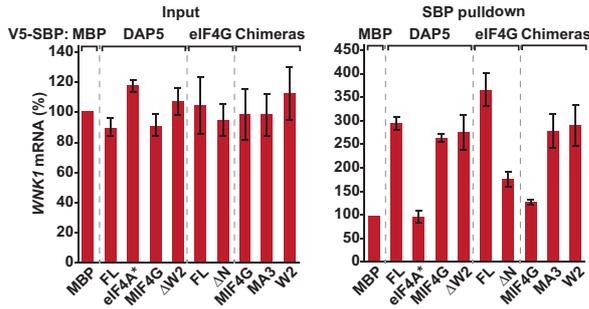
f *ROCK1* mRNA
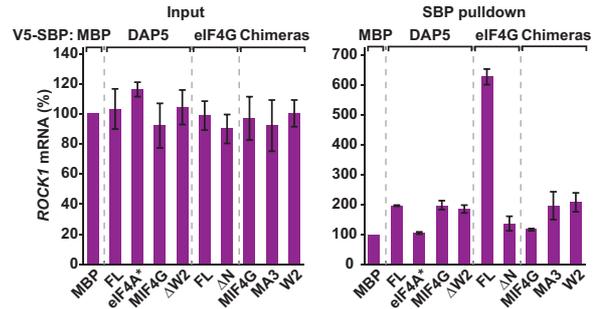
g *GAPDH* mRNA
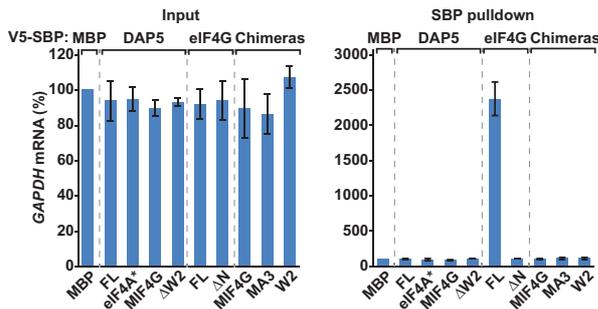
h
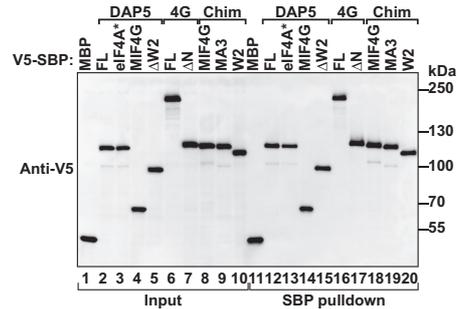
Weber *et al.* Fig. 3

a *WNK1*-R-LUC 5' leader deletions
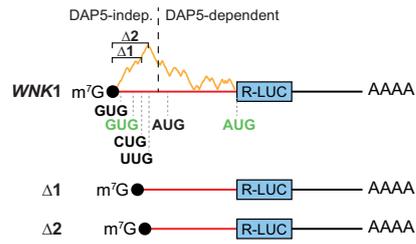
b
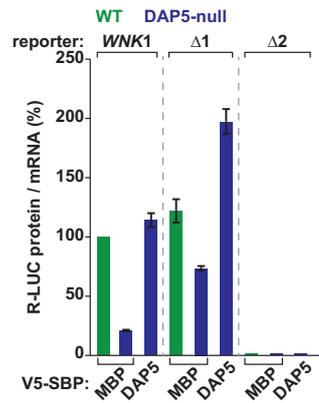
c
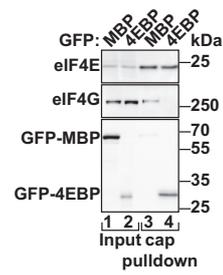
d *WNK1* mRNA
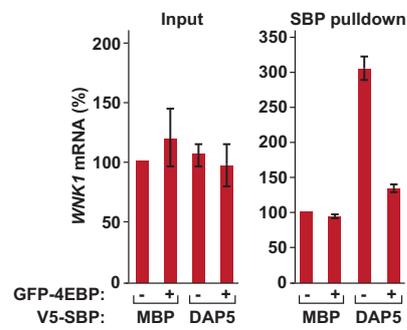
e *ROCK1* mRNA
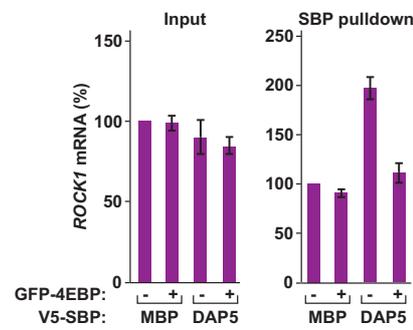
f
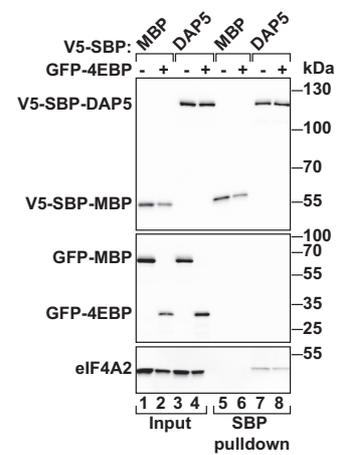
Weber *et al.* Fig. 4

a *WNK1*-R-LUC uORF2 reporters
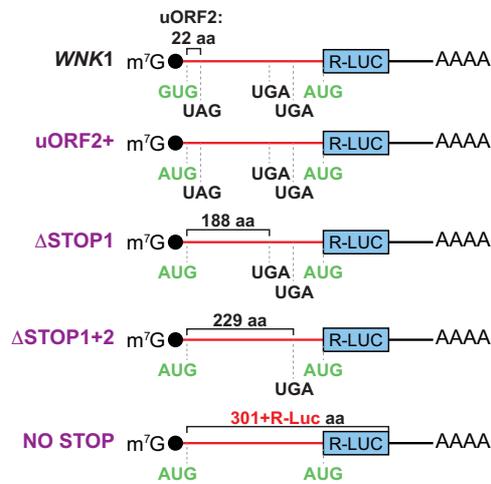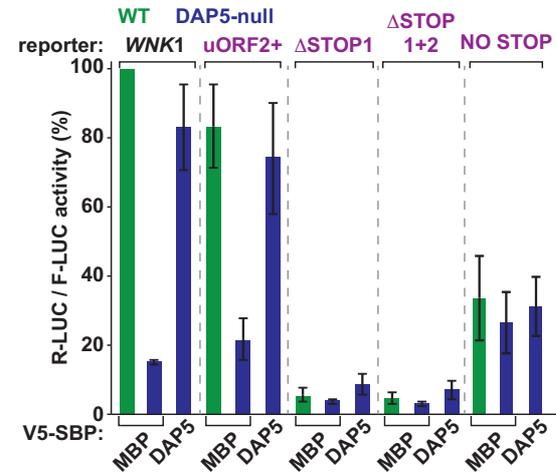
b *WNK1*-R-LUC

c
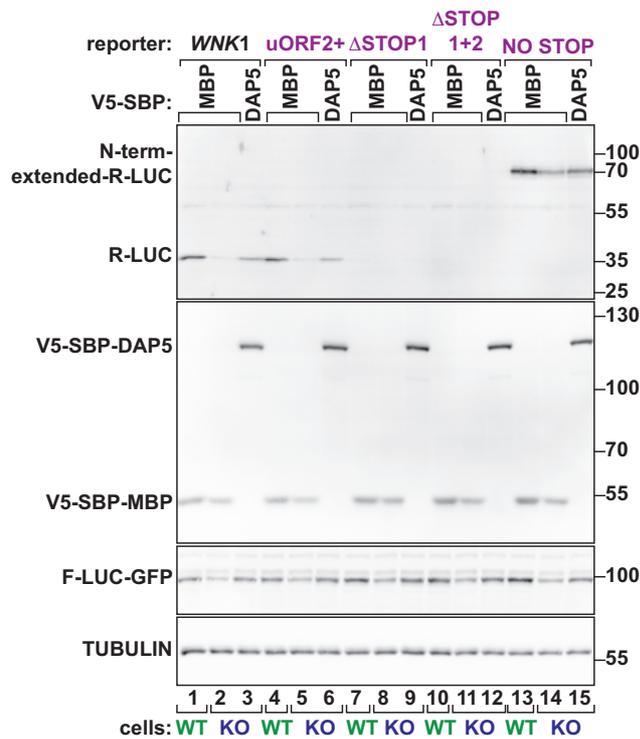
d *WNK1-R-LUC* mRNA
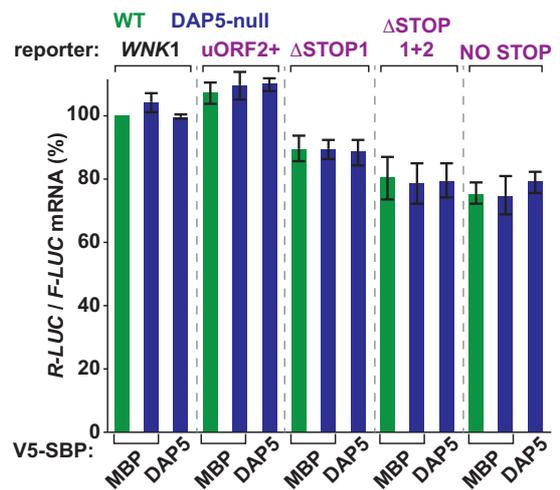
e *WNK1-R-LUC* mRNA
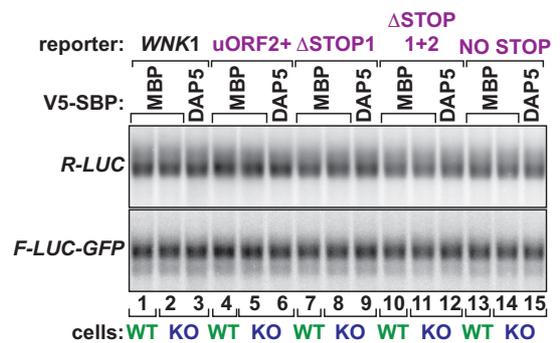
Weber *et al.* Fig. 5

a *WNK1*-R-LUC uORF5 reporters
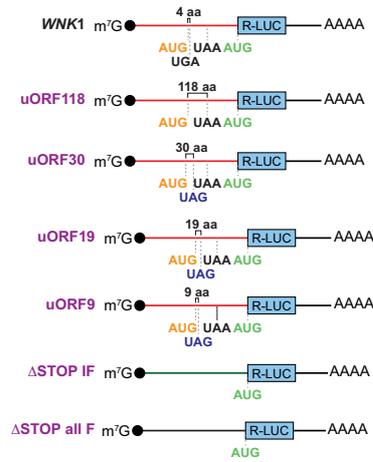
b *WNK1*-R-LUC TE
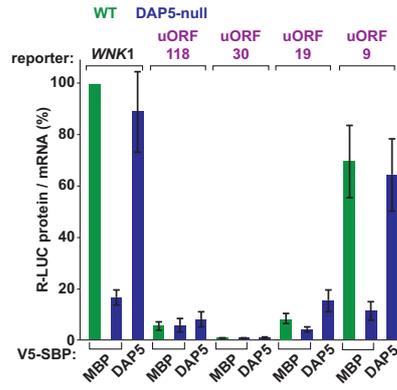
c
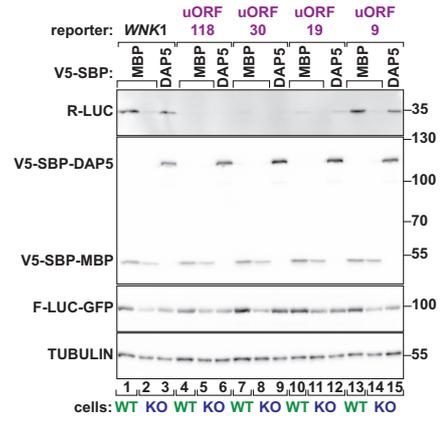
d
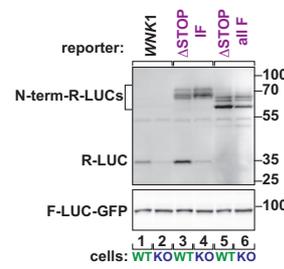
e 40S footprints
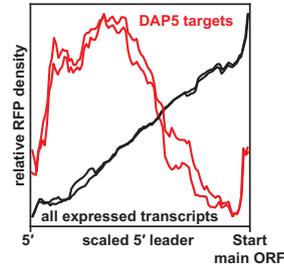
f 80S footprints
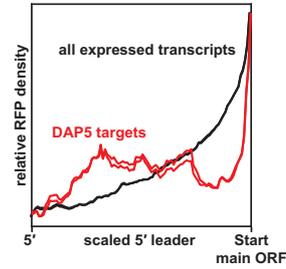
Weber *et al.* Fig. 6

a Fluorescent reporters
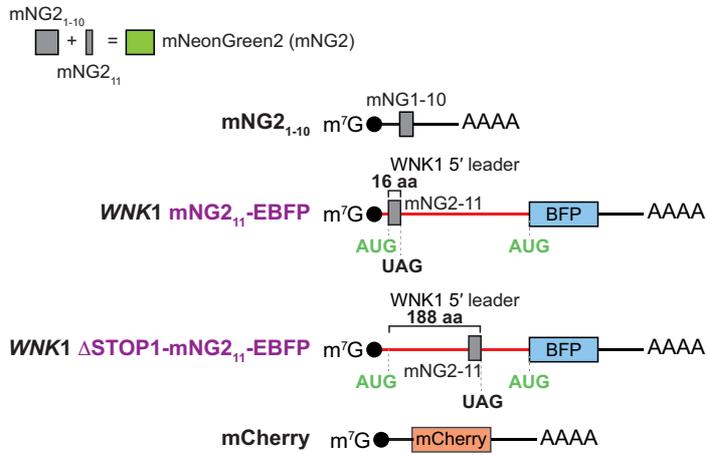
b mNG2-positive cells
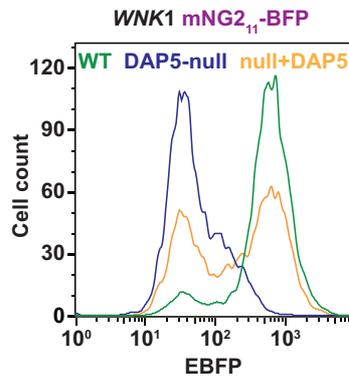
c BFP/mNG2
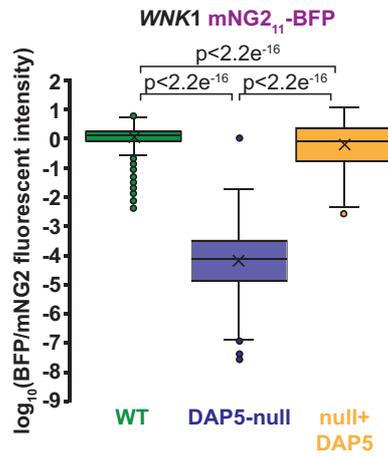
d mNG2/mCherry
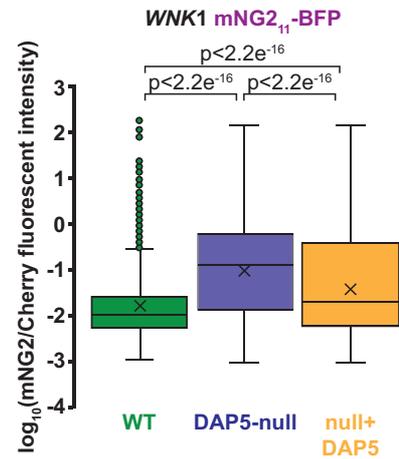
e BFP/mNG2
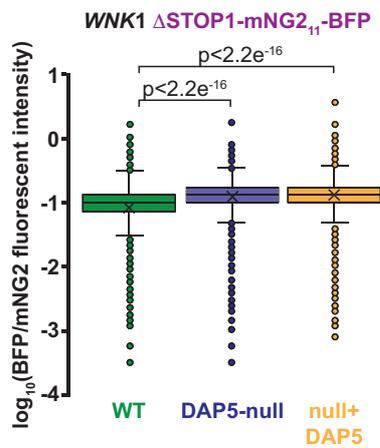
f mNG2/mCherry
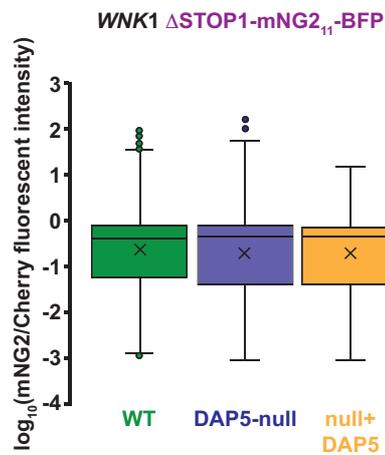
Weber et al. Fig. 7

a *WNK1*-R-LUC reporters
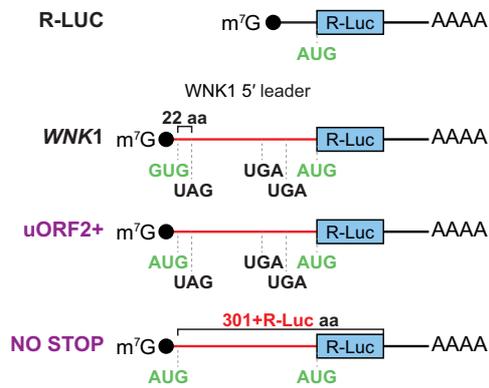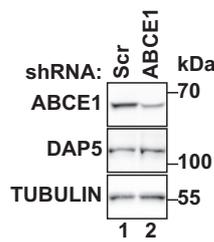
b eRF1^AAQ

c eRF1^AAQ

d ABCE1 KD

e ABCE1 KD
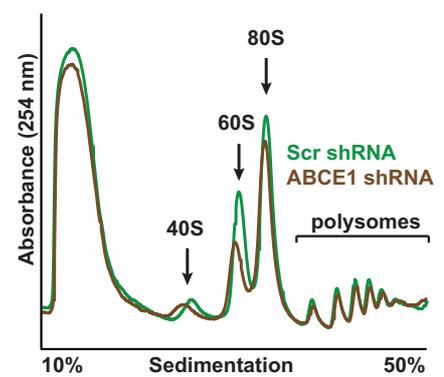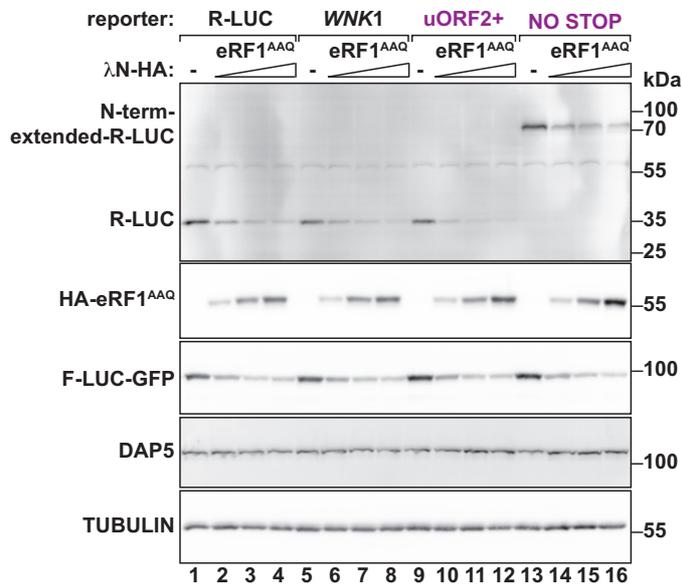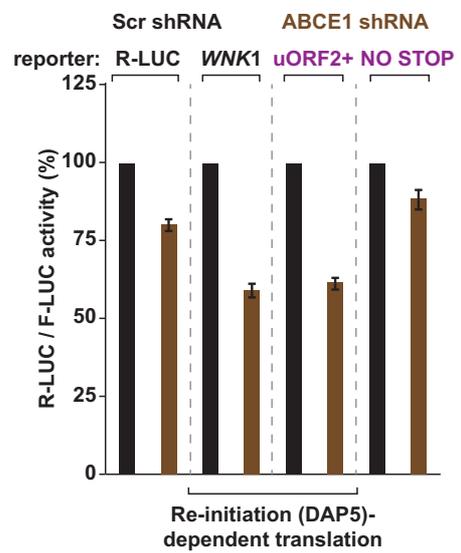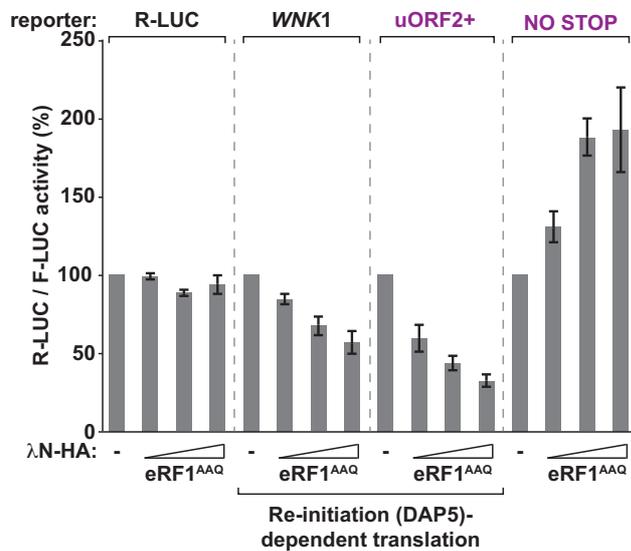
f ABCE1 KD

g ABCE1 KD
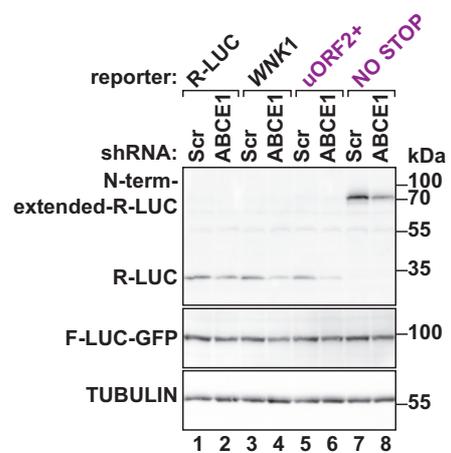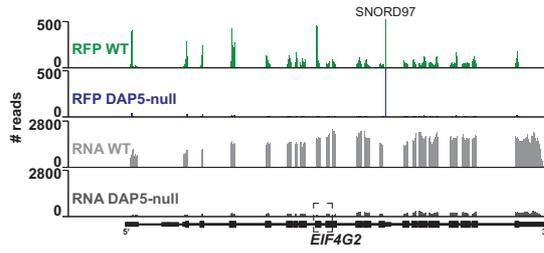
a *EIF4G2* (DAP5)

b total-RNA reads *EIF4G2* (DAP5)

c

d DAP5-null cells

RNA down
RNA up
n=1771
n=1766
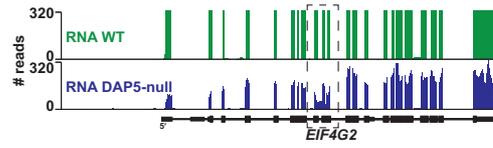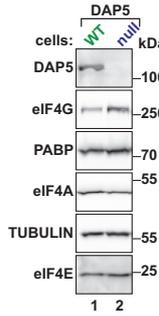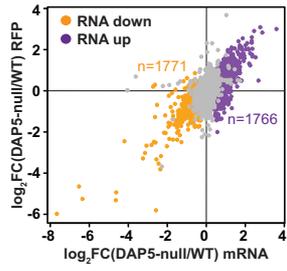
e Ribo-Seq
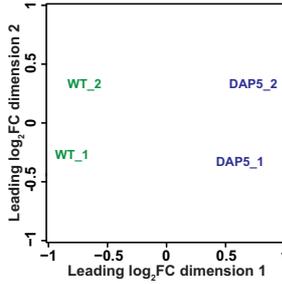
f RNA-Seq
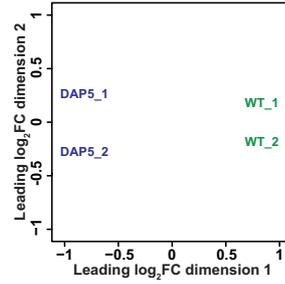
g
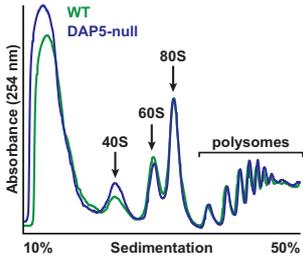
h
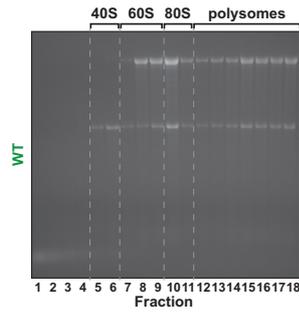
i
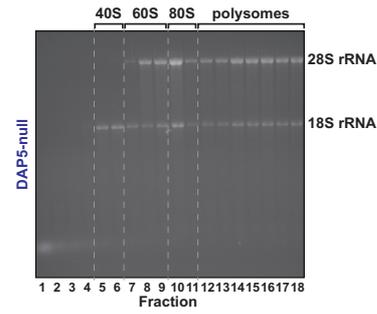
j *WNK1* mRNA
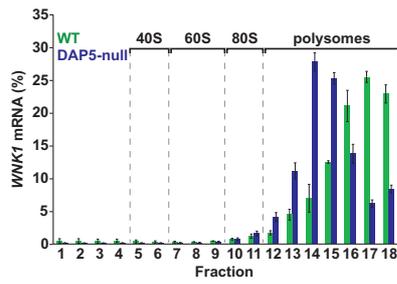
k *ROCK1* mRNA
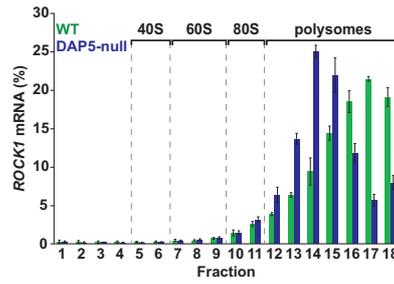
l *GAPDH* mRNA
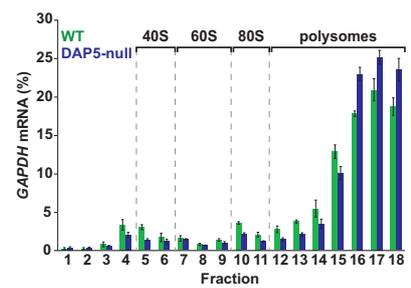
Weber *et al.* Fig. S2

a *ROCK1*
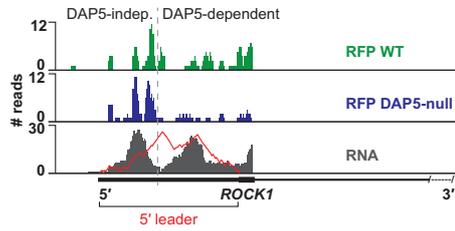
b *AKT1*
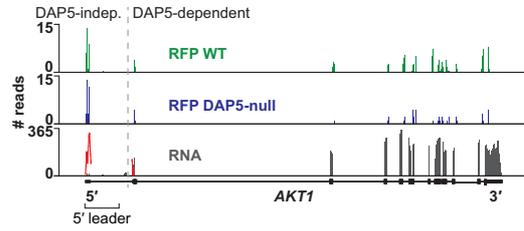
c *EN2*
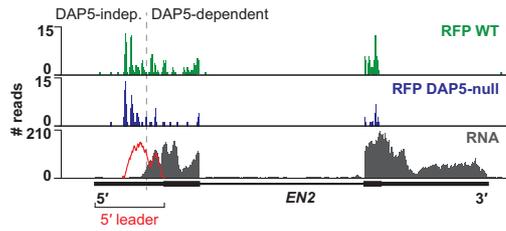
d *GAS1*
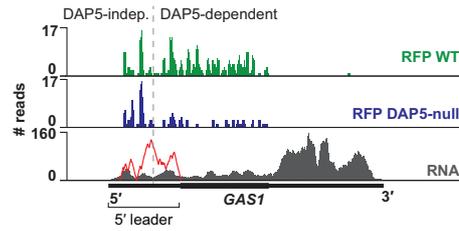
e *TLE4*
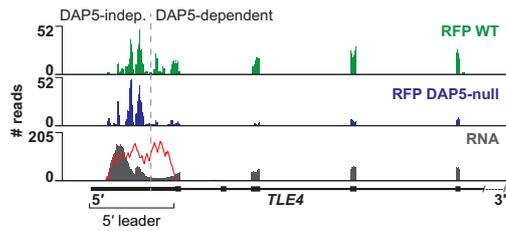
f *TLE3*
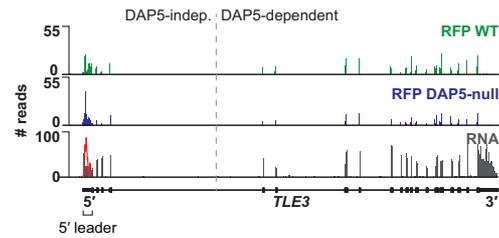
g *INPPL1*
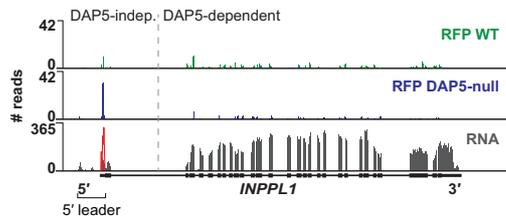
h *CAMSAP2*
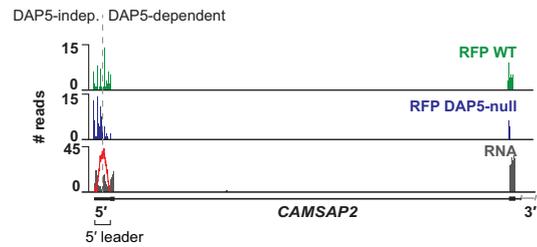
i *ALKBH5*
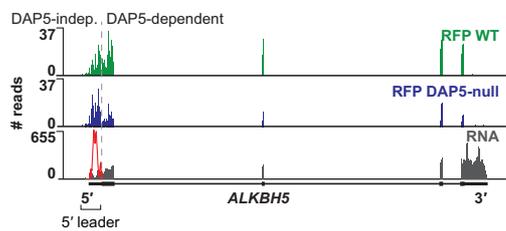
j *C6ORF62*
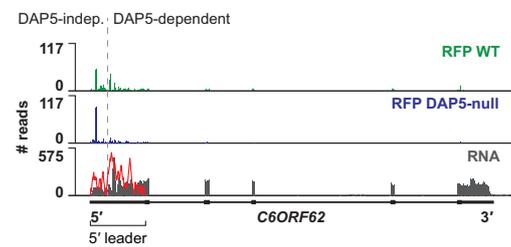
a 5′ leader length
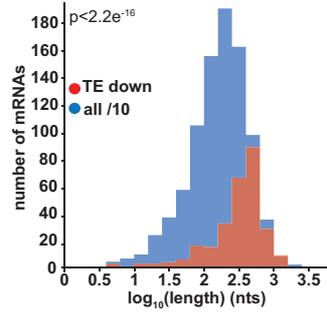
b 5′ leader GC content
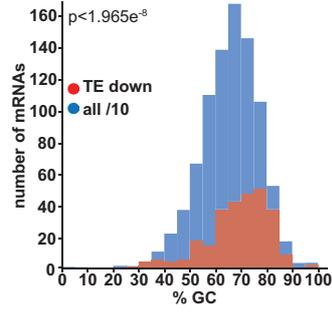
c 5′ leader structure
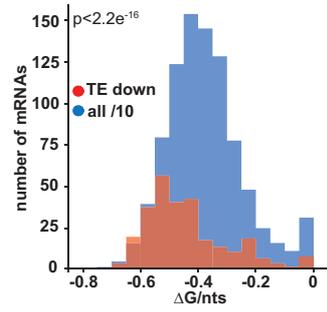
d *WNK1-R-LUC* mRNA
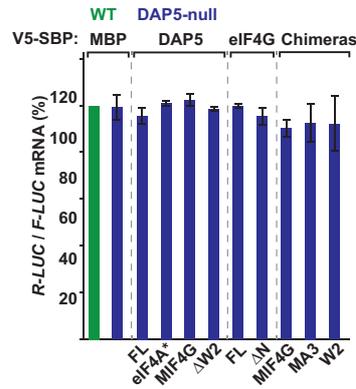
e *ROCK1-R-LUC* mRNA
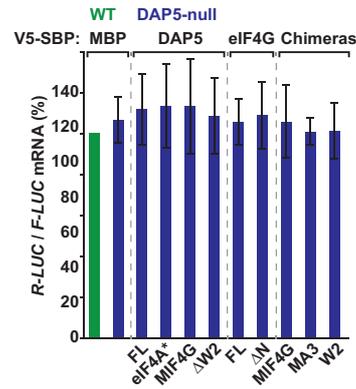
f *AKT1-R-LUC* mRNA
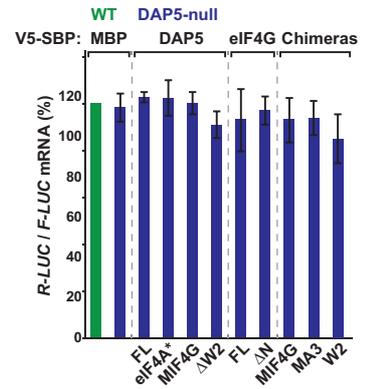
g *WNK1-R-LUC* mRNA
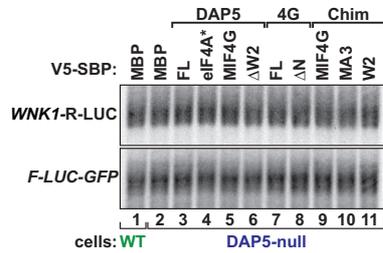
h *ROCK1-R-LUC* mRNA
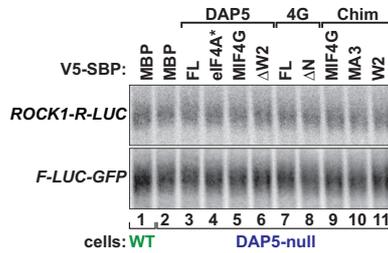
i *AKT1-R-LUC* mRNA
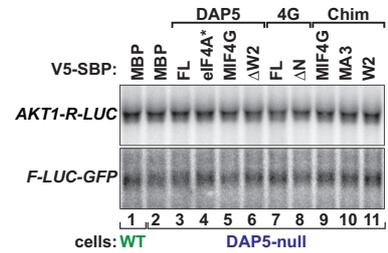
j
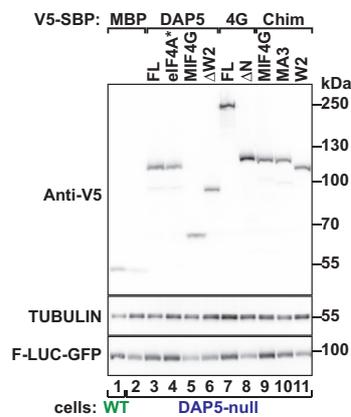
k
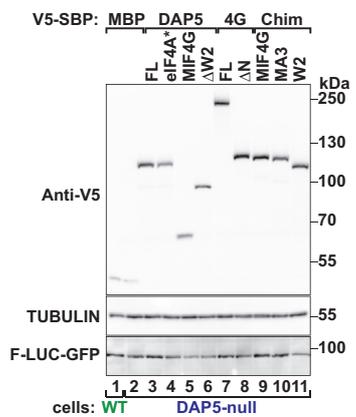
l
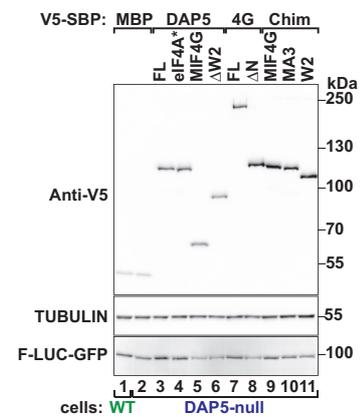
Weber *et al.* Fig. S4

a
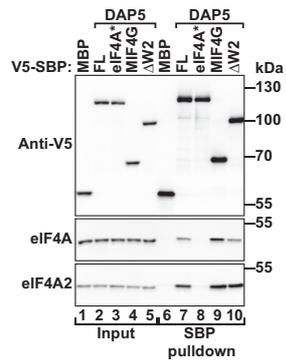
b
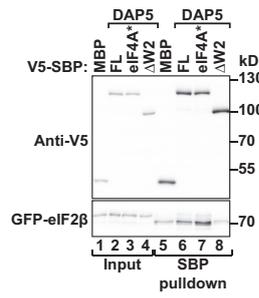
c Target mRNAs
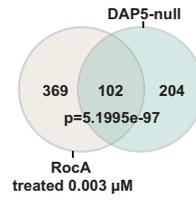
d DAP5 targets 5′ leader
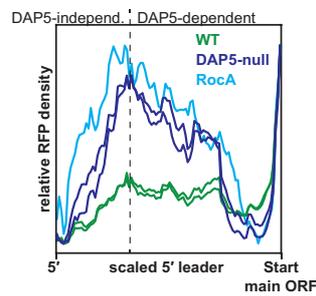
e
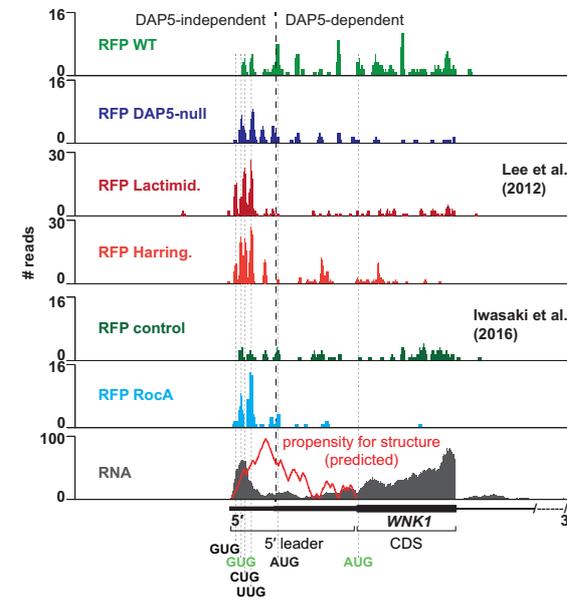
f
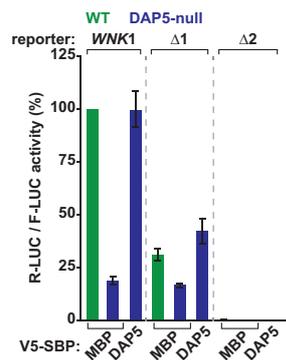
g
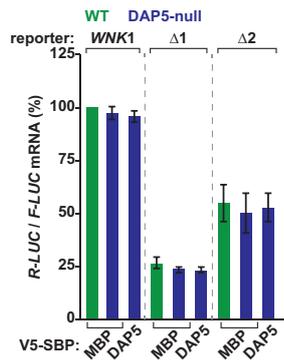
h
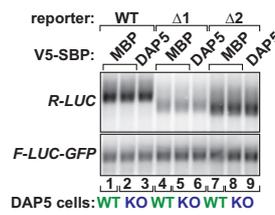
i
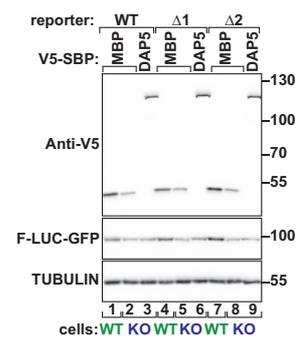
a *WNK1*-R-LUC uORF2 reporters
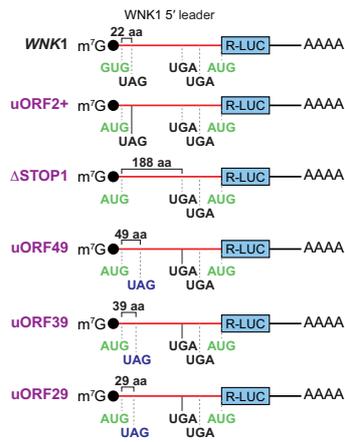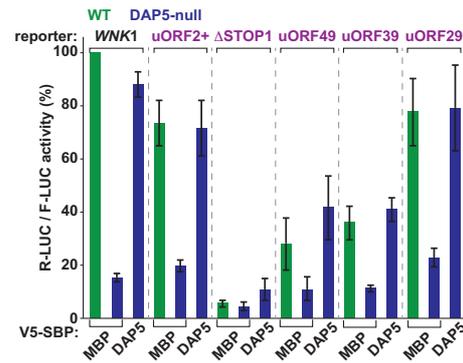
b *WNK1*-R-LUC

c *WNK1-R-LUC* mRNA
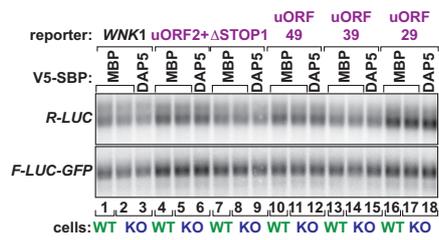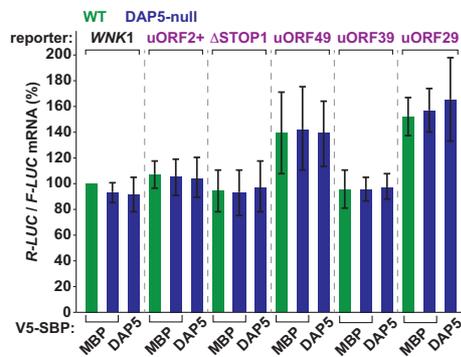
d *WNK1-R-LUC* mRNA

e
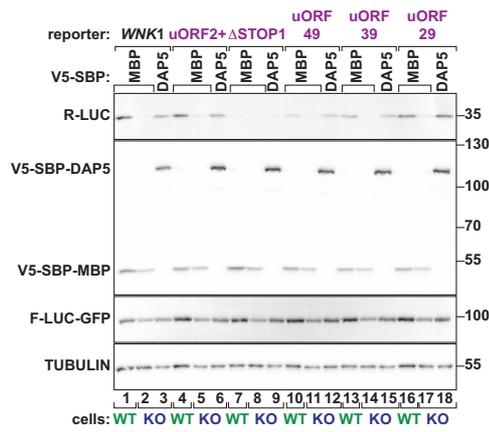
f *WNK1*-R-LUC (uORF5 reporters)
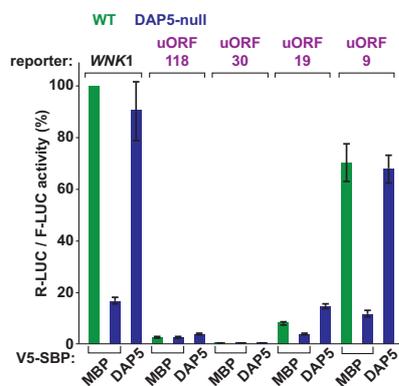
g *WNK1-R-LUC* mRNA (uORF5 reporters)

h

Weber *et al.* Fig. S6

a mNG2

untransf.=0.14%
mNG2$_{1-10}$=1.6%
mNG2$_{11}$-BFP=1.37%
mNG2$_{1-10}$ 0.5 µg + mNG2$_{11}$-BFP=4.15%
mNG2$_{1-10}$ 1 µg + mNG2$_{11}$-BFP=6.41%
mNG2$_{1-10}$ 1.5 µg + mNG2$_{11}$-BFP= 8.9%

b EBFP

untransf.=0%
mNG2$_{1-10}$=0.034%
mNG2$_{11}$-BFP=37%

c mCherry

untransf.=0.13%
WT=47.9%
DAP5-null=35.9%
null+DAP5=40.2%
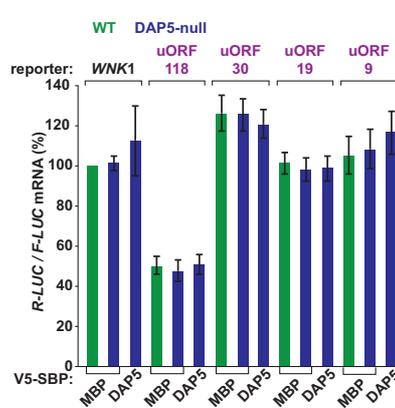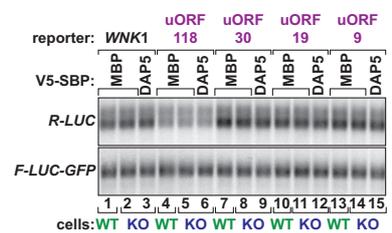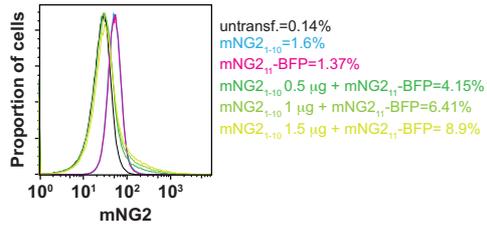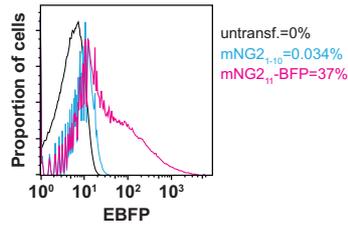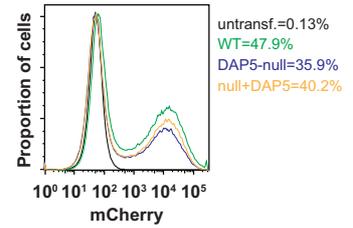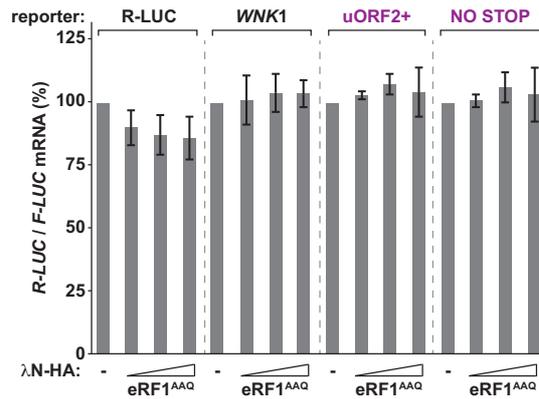
d *WNK1-R-LUC* mRNA
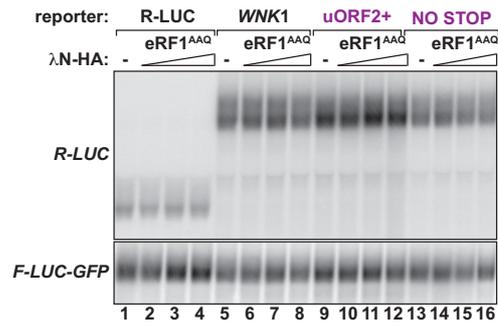
e *WNK1-R-LUC* mRNA

f *WNK1-R-LUC* mRNA
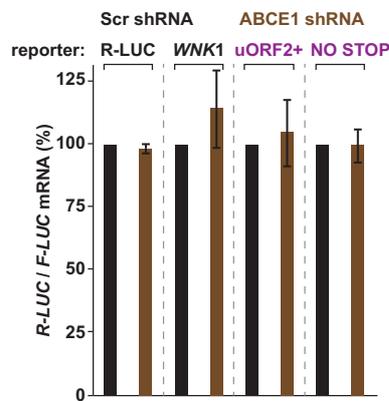
g *WNK1-R-LUC* mRNA
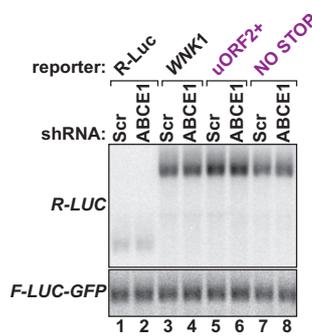
Weber *et al.* Fig. S7

a *WNK1*-R-LUC uORF2 reporters

R-LUC — m⁷G ● — R-Luc — AAAA
AUG

WNK1 5′ leader

*WNK*1 — m⁷G ● — 22 aa — R-Luc — AAAA
GUG UGA AUG
UAG UGA

uORF2+ — m⁷G ● — R-Luc — AAAA
AUG UGA AUG
UAG UGA

NO STOP — m⁷G ● — 301+R-Luc aa — R-Luc — AAAA
AUG AUG

b

shRNA: Scr / DENR+MCTS-1+eIF2D

DENR — 25 kDa
eIF2D — 70
DAP5 — 100
TUBULIN — 55

1 2

c

shRNA: Scr / eIF3J

eIF3J — 35 kDa
DAP5 — 100
eIF3a — 250
TUBULIN — 55

1 2

d

Scr shRNA / DENR+MCTS-1+eIF2D shRNA

reporter: R-LUC / *WNK*1 / uORF2+ / NO STOP

R-LUC / F-LUC activity (%)

Re-initiation (DAP5)-dependent translation

e

reporter: R-LUC / WNK1 / uORF+ / NO STOP

shRNA: Scr / DENR+MCTS-1+eIF2D (repeated)

N-term-extended-R-LUC — 100 / 70 kDa
— 55
R-LUC — 35
F-LUC-GFP — 100
TUBULIN — 55

1 2 3 4 5 6 7 8

f

Scr shRNA / eIF3J shRNA

reporter: *WNK*1 / uORF2+

R-LUC / F-LUC activity (%)

g

shRNA: Scr / eIF3J

WNK1-R-LUC
F-LUC-GFP

1 2

h

cells: WT / KO
Thapsigargin: - + / - +

ATF4 — 55
DAP5 — 100
eIF2α-P — 35
TUBULIN — 55

1 2 3 4

4EHP and GIGYF1/2 mediate translation-coupled messenger RNA decay

Ramona Weber[1], Min-Yi Chung[1], Csilla Keskeny[1,2], Ulrike Zinnall[3], Markus Landthaler[3], Eugene Valkov[1,4], Elisa Izaurralde[1,†] and Cátia Igreja[1,5]

[1]Department of Biochemistry, Max Planck Institute for Developmental Biology, Max-Planck-Ring 5, D-72076 Tübingen, Germany

[2]Present address: Experimental and Clinical Research Center (ECRC), Max Delbrück Center and Charité Berlin, 13125 Berlin, Germany

[3]Berlin Institute for Medical Systems Biology (BIMSB), Max Delbrück Center for Molecular Medicine in the Helmholtz Association (MDC), 10115 Berlin, Germany; IRI Life Sciences, Institute für Biologie, Humboldt Universität zu Berlin, 10115 Berlin, Germany

[4]Present address: Center for Cancer Research, National Cancer Institute, Frederick, MD, 21702-1201

[5]Lead contact

† deceased 30th April 2018

Correspondence: catia.igreja@tuebingen.mpg.de

Weber et al.

Summary (140 words)

Current models of mRNA turnover indicate that cytoplasmic degradation is coupled with translation. However, our understanding of the molecular events that coordinate ribosome transit with the mRNA decay machinery is still limited. Here, we show that 4EHP–GIGYF1/2 complexes trigger co-translational mRNA decay. Human cells lacking these proteins accumulate mRNAs with prominent ribosome pausing. These include among others transcripts encoding secretory and membrane-bound proteins or tubulin subunits. In addition, 4EHP– GIGYF1/2 complexes fail to reduce mRNA levels in the absence of ribosome stalling or upon disruption of their interaction with the cap structure, DDX6 and ZNF598. We further find that co-translational binding of GIGYF1/2 to the mRNA or the nascent peptide chain marks transcripts with perturbed elongation to decay. Our studies reveal how a repressor complex linked to neurological disorders minimizes the protein output of a subset of mRNAs.

Keywords: mRNA decay, translation, ribosome pausing, DDX6, GYF domain, endoplasmic reticulum, signal peptide, nascent chain, tubulin.

Introduction

Regulation of protein synthesis throughout the translation cycle safeguards the production of an optimal proteome. Changes in ribosome dynamics during elongation are required to fine-tune co-translational protein folding and regulate mRNA stability (Hanson and Coller, 2018; Hu et al., 2009; Radhakrishnan and Green, 2016). Factors such as codon or nascent peptide chain composition, secondary structures, ribosome-associated factors, defective ribosomes, and damaged or improperly processed mRNAs influence ribosome movement on the open-reading frame (ORF) [reviewed in (Buskirk and Green, 2017; Joazeiro, 2017)].

Codon optimality is a conserved evolutionary mechanism that affects mRNA stability in a translation-dependent manner (Hanson and Coller, 2018). mRNAs enriched in slow decoding (non-optimal) codons tend to be more unstable than those enriched in fast decoding (optimal) codons (Presnyak et al., 2015). In the unstable transcripts, the slow translating ribosomes are thought to be recognized by proteins of the decay machinery, such as the RNA helicase DDX6 and the CCR4-NOT complex, which then trigger mRNA decay (Buschauer et al., 2020; Radhakrishnan et al., 2016).

Translation-coupled mechanisms also control mRNA stability in response to the accumulation of unwanted or potentially cytotoxic proteins. α- and β-tubulin mRNAs are decayed in response to excess of depolymerized tubulin (Cleveland et al., 1981; Gasic et al., 2019; Gay et al., 1989; Pachter et al., 1987). Binding of tetratricopeptide protein 5 (TTC5) to an N-terminal motif of tubulin activates, by yet unidentified factors, the decay of the ribosome-bound mRNA (Lin et al., 2020). Similarly, quality control checkpoints sense defects in protein targeting to the endoplasmic reticulum (ER) and initiate mRNA degradation. Failure in the interaction of the signal recognition particle (SRP) with the signal sequence of the nascent protein or the receptor at the ER membrane results in the recruitment of the decay machinery to the translating mRNA (Karamyshev et al., 2014; Lakshminarayan et al., 2020; Pinarbasi et

al., 2018). The molecular details of co-translational decay of secretome-associated mRNAs remain unclear.

Damaged or improperly processed mRNAs are also co-translationally degraded. Disruption of elongation in faulty transcripts causes ribosome stalling and collision, decreases translation, and has the potential to induce proteotoxic stress (Simms et al., 2017a). Thus, cells evolved surveillance mechanisms that coordinate the degradation of the truncated protein products and ribosome rescue with mRNA degradation [reviewed in (Brandman and Hegde, 2016; Joazeiro, 2019; Simms et al., 2017a)]. Recognition and ubiquitination of the collided ribosomes by the E3 ubiquitin ligase ZNF598 (Garzia et al., 2017; Ikeuchi et al., 2019; Juszkiewicz et al., 2018; Juszkiewicz and Hegde, 2017; Simms et al., 2017b; Sundaramoorthy et al., 2017) activates mRNA decay (D'Orazio et al., 2019).

ZNF598 binds to the Grb10-interacting GYF (glycine-tyrosine-phenylalanine) domain proteins 1 and 2 (GIGYF1/2) which form a translational repressor complex with the cap-binding eIF4E-homologous protein (4EHP) (Morita et al., 2012). The GYF domain of GIGYF1/2 binds to proteins containing Pro-Pro-Gly-Φ motifs (Φ, hydrophobic amino acid with the exception of tryptophan), such as ZNF598, tristetraprolin (TTP) or the microRNA (miRNA)-induced silencing complex-associated TNRC6 proteins (Fu et al., 2016; Morita et al., 2012; Schopp et al., 2017). These interactions integrate the 4EHP–GIGYF1/2 complexes in miRNA-mediated gene silencing, regulate cytokine production, and control gene expression during embryonic development (Fu et al., 2016; Giovannone et al., 2009; Kryszke et al., 2016; Morita et al., 2012; Schopp et al., 2017; Tollenaere et al., 2019). Together with the CCR4-NOT deadenylase complex and DDX6, 4EHP and GIGYF1/2 repress translation initiation and elicit mRNA decay (Amaya Ramirez et al., 2018; Peter et al., 2019; Ruscica et al., 2019).

Despite the interaction with ZNF598, GIGYF1/2 and 4EHP have not been associated with translational surveillance. Using translatome and transcriptome analysis, we explored the role of 4EHP–GIGYF1/2 complexes in the regulation of translation and mRNA stability. Our

results highlight a role for this repressor complex in co-translational degradation of mRNAs, many of which encoding secreted and membrane-bound proteins. Together with DDX6 and GYF-domain binding proteins, 4EHP and GIGYF1/2 induce decay of mRNAs with disturbed elongation. Our studies indicate that 4EHP and GIGYF1/2 are part of the cellular machinery that selectively reduces the abundance of actively translating mRNAs to fine-tune protein synthesis.

Results

4EHP–GIGYF1/2 complexes regulate the abundance of mRNAs encoding secreted and membrane-bound proteins

To identify mRNAs regulated by 4EHP and GIGYF1/2, we studied genome-wide translational changes using ribosome profiling (Ingolia et al., 2009). Isolation and identification of ribosome-protected fragments coupled to transcriptome analysis were performed in control (Ctrl), CRISPR-Cas9 engineered GIGYF1/2-null (KO) (Peter et al., 2017) and 4EHP-null (Rasch et al., 2020) HEK293T cells (Figures S1A, B). The experiments were reproducible as ribosomal footprints and RNA-Seq library replicates clustered together (Figures S1C, D).

To detect variations in translational efficiency (TE) across experimental conditions, genes were plotted according to changes in mRNA abundance and ribosome occupancy (Figures 1A, B). Only a small subset of mRNAs showed altered TE in the absence of 4EHP (n=24) or GIGYF1/2 (n=7) (Figures 1A, B; S1E, F; Table S1). However, in comparison to control cells 497 and 341 mRNAs exhibited increased abundance in 4EHP- and GIGYF1/2-null cells, respectively (Tables S2, S3). A significant fraction of the mRNAs was commonly upregulated in both null cells (n=82, p=$1.4459e^{-34}$, Figure 1C, Table S4). Although 571 and 569 mRNAs were downregulated in 4EHP- and GIGYF1/2-null cells, there was no significant overlap among the two datasets (n=38, p=0.1963, Figure S1G). Reduction of mRNA levels may be a consequence of indirect effects following the loss of these proteins.

As 4EHP and GIGYF1/2 are negative regulators of mRNA stability (Amaya Ramirez et al., 2018; Kryszke et al., 2016; Ruscica et al., 2019), we focused our analysis on transcripts upregulated in both cell lines (target mRNAs). GIGYF1/2 and 4EHP target mRNAs were overrepresented for genes encoding cell surface and extracellular proteins (Figures 1D, E). In fact, approximately half of the commonly upregulated mRNAs code for endoplasmic reticulum (ER), membrane or secreted proteins (Figure 1F) which are translated at the ER (Hermesh and

Jansen, 2013). Thus, this data suggests that 4EHP–GIGYF1/2 complexes regulate the abundance of a subset of secretome mRNAs.

GIGYF1/2 regulate mRNA stability

We next determined target mRNA decay rates after transcriptional arrest by Actinomycin D using northern blot or RT-qPCR. Decay kinetics in control and GIGYF1/2-null cells fitted to an exponential decay with a single component and $R^2 \geq 0.78$ (Figure 2). We found that *DBNDD2*, *CD109*, *ITPR3* and *NPTX1* mRNAs were stabilized in the absence of GIGYF1/2 as the corresponding half-lives ($t_{1/2}$) mostly doubled in these cells (Figure 2). In contrast, a non-target mRNA such as *β-ACTIN* was degraded similarly in control and null cells (Figure 2F). Therefore, GIGYF1/2 proteins are regulators of mRNA stability, as observed in *Drosophila* and human HeLa cells mainly using reporter assays (Amaya Ramirez et al., 2018; Kryszke et al., 2016; Ruscica et al., 2019).

Multiple GIGYF1/2 co-factors are required for mRNA decay

To obtain insight into the molecular mechanism, we probed if the assembly of the 4EHP–GIGYF1/2–DDX6 complex (Figure S1H) was required for mRNA degradation. We measured mRNA abundance in GIGYF1/2-null cells upon transient co-expression of 4EHP, and wild type (WT) or mutants of the GIGYF paralogs unable to associate with 4EHP (C*), DDX6 (DDX6*) or PPGΦ-containing proteins (GYF domain mutant; GYF*) (Ash et al., 2010; Peter et al., 2019; Peter et al., 2017) (Figures S1I, J). In GIGYF1/2-null cells, target transcript levels increased more than two-fold, but not *β-ACTIN* (Figures 3A, B, S1K). mRNA degradation was restored when 4EHP was co-expressed with WT, but not C*, DDX6* or GYF* GIGYF1/2 mutants (Figures 3A, B, S1K). Moreover, failure in the assembly of the full complex also compromised *DBNDD2* turnover (Figure S1L). Re-expression of GIGYF1/2 alone was not

sufficient to induce mRNA decay in the null cells (data not shown), as 4EHP and GIGYF1/2 protein stabilities are co-regulated (Figures S1A, B) (Morita et al., 2012).

Likewise, the levels of *NPTX1*, *CD109* and *ITPR3* mRNAs in 4EHP-null cells were also at least 2-fold higher (Figure S2A). Target mRNA decay was restored upon co-expression of the 4EHP–GIGYF1/2 complex but not when 4EHP is unable to bind to the cap (cap*) and GIGYF1/2 (S*) (Peter et al., 2017). Wild type and mutant GIGYF1/2 or 4EHP expression levels were similar and did not affect the abundance of *β-ACTIN* mRNA (Figures 3B, S1K, S2A, B).

These results indicate that 4EHP, DDX6 and GYF domain-associated protein(s) bind to GIGYF1/2 to promote target mRNA degradation.

The GYF domain of GIGYF2 mediates mRNA binding

To investigate the recruitment of the complex to target mRNAs, we performed RNA-immunoprecipitation assays (RNA-IP) and RT-qPCR. In contrast to GFP-MBP, GFP-GIGYF2 efficiently associated with *NPTX1*, *CD109*, *DBNDD2* and *ITPR3* (Figure 3C-F, IP graphs). GIGYF2 binding increased with the length of the target coding sequence (CDS), as longer sequences (*CD109*, *ITPR3*) showed higher binding efficiencies than shorter ones (*NPTX1*, *DBNDD2*). The interaction of GIGYF2 C* and DDX6* mutants with the different transcripts was comparable to wild type protein. However, the association of the GIGYF2 GYF* mutant with mRNA was strongly reduced (Figure 3C-F, IP graphs) indicating that target recognition relies on the GYF domain of GIGYF2. All proteins were expressed at equivalent levels and did not alter mRNA steady state levels (Figures 3C-G, input graphs, G).

Similarly, V5-SBP-4EHP bound to mRNA in the presence of GFP-GIGYF2 WT or DDX6*, but not GFP-MBP, GIGYF2 C* or GYF* (Figure S2C-E, pulldown graphs). mRNA degradation, inferred from the steady state mRNA levels, only occurred in cells co-expressing GIGYF2 WT (Figure S2C-E, input graphs). Our results indicate that binding of 4EHP to the

mRNA cap requires GIGYF2 and that recruitment of the complex to the target mRNA is provided by GIGYF2 GYF-domain interacting proteins.

4EHP–GIGYF1/2 trigger co-translational mRNA decay

Next, we dissected the mRNA features required for turnover. We generated reporters containing the CDS or the 3' UTR of *DBNDD2* and *LGALS3BP* to express in control and GIGYF1/2-null cells. To measure protein levels, a hemagglutinin (HA) tag was inserted in frame with each CDS (Figure 4A) whereas the 3' UTRs were preceded by *Renilla* luciferase ORF (R-LUC; Figure S2G).

DBNDD2-HA and LGALS3BP-HA protein levels were increased in GIGYF1/2-null cells (Figures 4B-D). The corresponding transcripts were also more stable in the absence of GIGYF1/2 (Figures 4E, F, I, J), indicating that the CDS is sufficient to recapitulate mRNA decay. In contrast, R-LUC activity and mRNA levels of the 3' UTR reporters were similar in control and null cells (Figures S2H-J). Thus, target mRNA decay is independent of the 3' UTR-associated mechanisms previously associated with the 4EHP–GIGYF1/2 complexes (Fu et al., 2016; Kryszke et al., 2016; Schopp et al., 2017; Tollenaere et al., 2019).

The observation that the CDS determined GIGYF1/2-dependent mRNA decay suggested that turnover occurred co-translationally. We examined the decay rate of intron-less CDS reporters containing a premature STOP three codons downstream of the AUG to prevent the synthesis of an HA-tagged protein (Figures 4A-C). Interestingly, *DBNDD2-STOP₃-HA* and *LGALS3BP-STOP₃-HA* transcripts degraded with similar rates in control and null cells (Figures 4G-J). Hence, GIGYF1/2-mediated mRNA decay requires translation of the CDS.

We then analyzed the association of GIGYF2 and targets with ribosomes using sucrose density gradient separation (Figure S3A). GIGYF2, and its co-factor ZNF598, were mostly observed in the top fractions of the gradient corresponding to free ribonucleoprotein particles (RNPs) and 40S ribosome subunit; Figure S3A, lanes 1-6). However, both proteins also bound

Weber et al.

to polysome-associated ribosomes as part of the signal was detected in polysomes (Figure S3A, lanes 13-16). In addition, GFP-GIGYF2 WT and GYF* mutant co-purified with the ribosomal protein eL22L1 (Figure S3B). Consistent with the idea of ribosome-associated decay, *DBNDD2*, *CD109*, *ITPR3* and *NPTX1* were engaged in translation as the majority of each mRNA was associated with polysomes or 80S monosomes (Figure S3C-F).

We also measured GIGYF2 mRNA recognition in the absence of translation. Inhibition of translation with harringtonine or puromycin did not alter GFP-GIGYF2 expression but profoundly reduced its binding to *ENO2* (Figures S3G-J). *ENO2* is another bona-fide target transcript as decay in null cells is restored upon re-expression of the 4EHP–GIGYF2 complex (Figure S3K).

Altogether, these results suggest that GIGYF2 associates with ribosomes to induce decay of actively translating mRNAs.

Ribosome pausing is evident in 4EHP and GIGYF1/2 target mRNAs

To further investigate the details of co-translational decay by 4EHP and GIGYF1/2, we had a closer look at the ribosome footprint distribution along the CDS of the regulated transcripts. We found pronounced pauses, characterized by the accumulation of unique ribosome footprints greater than the median footprint coverage of the gene, in several of the 4EHP and GIGYF1/2 target mRNAs (Figure 5A, S4, Table S4). There was no preference for location or peptide motif at the pause sites, but several contained Pro and negatively charged amino acids. Proline and acidic residues are known to promote slow peptide bond formation and be enriched at ribosome pause sites (Pavlov et al., 2009; Pelechano and Alepuz, 2017; Schuller et al., 2017; Wohlgemuth et al., 2008).

In *LARGE2* translation was stalled at a Pro-Pro-Asp (P_{57}, P_{58}, D_{59}) motif located in the N-terminal region of the protein (Figure S4A), a previously described strong pause site (Ingolia et al., 2011; Schuller et al., 2017). *LARGE2* has four predicted isoforms with distinct N-terminal

regions. Curiously, with the exception of the canonical version, all other LARGE2 isoforms lack the PPD motif due to large N-terminal truncations or a 30 amino acid deletion that removes residues 30 to 59.

The translational pause present in *CXCL16* at a Pro-Gly-Asn (P_{45}, G_{46}, N_{47}) motif is located after a stretch of 10 hydrophobic residues, 9 of which are Leu (Figure S4B, C). It contains the PG dipeptide which is overrepresented in ribosome stall sequences in bacteria, yeast and humans (Doerfel et al., 2013; Manjunath et al., 2019; Pelechano and Alepuz, 2017; Schuller et al., 2017).

Prominent examples of translation pauses were also observed in the *NCKIPSD*, *ENO2*, *IFRD2*, and *DBNDD2* (Figure 5A; S4D-F). Interestingly, several of the paused ribosomes detected in the 4EHP and GIGYF1/2 target mRNAs were also identified in ribosome run off assays performed in HEK293 cells treated with translational inhibitors by Lee and co-workers (Lee et al., 2012).

Ribosomal run off assays in the presence of lactimidomycin (LTM) followed by ribosome profiling revealed increased footprints at two AUG codons (M_1 and M_{165}) in *ENO2* (Figure S4E). Curiously, M_{165} marks the position of the ribosome pause observed in our profiling experiment. Lactimidomycin associates with ribosomes with an empty E-site (Garreau de Loubresse et al., 2014), a feature of initiating ribosomes or elongating ribosomes with impaired decoding kinetics or slow peptide bond formation (Buschauer et al., 2020; Schmidt et al., 2016). As no alternative start sites or N-terminally truncated protein isoforms have been described for *ENO2*, the LTM footprint at M_{165} most likely represents a ribosome paused during elongation, and not an initiating ribosome.

The ribosome stall observed at the Asp-Asp-Glu motif of *IFRD2* (D_{97}, D_{98}, E_{99}; Figure S4F) was still present after ribosome run off assays in harringtonine-treated cells. Harringtonine is an A-site inhibitor (Garreau de Loubresse et al., 2014) that predominantly marks ribosomes positioned at the start codon (Ingolia et al., 2011). The footprint in the presence of harringtonine

at the DDE motif is not a translation start site since no AUG or near-cognate start codon are present. Thus, the associated ribosome did not conclude elongation (run off) following treatment with the drug. Furthermore, monosome and disome profiling in embryonic stem cells (Tuck et al., 2020) identified in the mouse *Ifrd2* orthologue mRNA increased disome occupancy at an equivalent position of the human transcript (Figure S4G). As disomes are a sign of ribosome collision and delayed elongation, these data confirm that stalled ribosomes accumulate at the DDE motif of *IFRD2*.

The translational stall detected in *DBNDD2* at the Phe-Glu-Asp (F_{23}, E_{24}, D_{25}) peptide was also observed at an equivalent position of the transcript after ribosome run off assays (Figure 5A). Again, the observed footprint most likely identifies a ribosome trapped during elongation since the underlying sequence is incompatible with a translation start site.

These findings indicate that 4EHP and GIGYF1/2 regulate mRNAs with perturbed translation elongation.

4EHP–GIGYF1/2 dependent decay partially relies on ZNF598

The presence of ribosome stalling and collision in GIGYF1/2-regulated transcripts suggests that their decay is coupled to translation surveillance and ZNF598. To verify if that is the case, we applied RNA-Seq and Ribo-Seq to *ZNF598*-null cells (Figure S5A-C). In these cells the main changes occur at the mRNA level, as only a minor fraction of genes (n=7; Table S1) displayed significant changes in translation efficiency (TE) (Figure S5D, E). From the group of genes with increased mRNA abundance in the null cells (n=357), 9.2% and 14% were also upregulated in the absence of GIGYF1/2 ($p=3.089792e^{-8}$) or 4EHP ($p=2.37387e^{-12}$), respectively (Figure S5F, Tables S6 and S7). 4EHP, GIGYF1/2 and ZNF598 commonly upregulated transcripts (n=6) included *ENO2* and *CXCL16*, which display prominent ribosome pauses (Figures S4B, E). Increased transcript levels were the result of enhanced mRNA stability as the half-lives of *ENO2* and *CXCL16* increased in the absence of ZNF598 (Figures S5G, H).

Thus, co-translational decay of a subset of 4EHP and GIGYF1/2 targets is dependent on ZNF598.

Ribosome pausing initiates 4EHP–GIGYF1/2 dependent mRNA decay

To confirm that ribosome pausing in GIGYF1/2 and 4EHP targets induces mRNA decay, we investigated the significance of the translational stall present in *DBNDD2*. We introduced a premature STOP before the pause site in the *DBNDD2-HA* reporter ($STOP_{18}$; Figure 5B) and measured mRNA levels in control and GIGYF1/2-null cells. Consistent with a failure to degrade mRNA in the absence of the translational pause, *DBNDD2-STOP$_{18}$-HA* levels did not vary in cells lacking GIGYF1/2 (Figures 5C, D).

DBNDD2 is a Pro-rich protein with two Pro tripeptides (P_{46}, P_{47}, P_{48} and P_{85}, P_{86}, P_{87}). Although we have not observed ribosome stalling at these positions, poly-Pro motifs are known to interfere with elongation (Gardin et al., 2014; Gutierrez et al., 2013; Ingolia et al., 2011; Pavlov et al., 2009; Wohlgemuth et al., 2008). We generated reporters with STOPs that prevented ($STOP_{39}$ and $STOP_{79}$) or allowed ($STOP_{89}$) the translation of the poly-Pro sequences by the ribosome (Figure 5B). As observed for WT *DBNDD2-HA*, the abundance of the reporters with STOPs after the FED motif was regulated by the 4EHP–GIGYF2 complex (Figures 5B-D). Thus, only the identified ribosome pause is required for GIGYF1/2-dependent degradation of the *DBNDD2-HA* mRNA. All the premature STOPs blocked the synthesis of HA-tagged proteins (Figure 5C) and did not significantly alter the expression of the different reporters (Figure S6A). Since the degradation efficiency of NMD substrates in HEK293 cells is low (Gerbracht et al., 2017), the abundance of *DBNDD2-STOP$_{18}$-HA* was unaltered in UPF1-depleted cells (Figures S6B, C).

To change ribosome occupancy and avoid translational pausing, we substituted the FED motif by triple alanine (Figure 5B). However, the levels of DBNDD2-FED$_{25}$-AAA-HA protein still responded to variations in GIGYF1/2 levels (Figures S6D). This data indicates that

although the ribosome is paused at the FED motif, translational stalling is not a consequence of impaired decoding.

We then explored the possibility that mRNA stability was controlled by factors that recognize the nascent chain. We transfected cells with a version of the reporter originating a protein with a modified N-terminus (M_1AAA instead of M_1DPN). In control cells, *DBNDD2-DPN$_4$-AAA-HA* mRNA was better expressed and more stable than WT reporter (Figures 5E, F). Moreover, its protein levels were not regulated by GIGYF1/2 (Figure S6D). These results show that the first translated codons of *DBNDD2-HA* mRNA are required for decay and suggest that factors that bind to the nascent chain critically control its turnover.

GIGYF1/2-dependent mRNA decay can occur during co-translational ER targeting

Many of the 4EHP–GIGYF1/2 target mRNAs encode signal peptide (SP)-containing proteins (Figure 1F) which undergo SRP-dependent translocation to the ER. Binding of SPR to the SP transiently interferes with translation elongation and leads to stacking and ribosome collision at the 5' end of the mRNA CDS (Arpat et al., 2019; Walter and Blobel, 1981; Wolin and Walter, 1988, 1989).

To study if translational pausing associated with ER-targeting is linked to GIGYF1/2-mediated mRNA decay, we made use of the *LGALS3BP-HA* CDS reporter. LGALS3BP is a secreted protein with 18 amino acids long SP. Consistent with translational pausing during targeting to the ER, ribosome footprints are still detected downstream of the signal sequence (Asp$_{21}$) in run off assays performed in cells treated with harringtonine and LTM (Lee et al., 2012). Likewise, a disome peak is observed at Val$_{60}$ (Han et al., 2020) (Figure 6A).

We introduced STOP codons 30 and 60 residues after the initiating AUG to prevent (*STOP$_{30}$*) or allow (*STOP$_{60}$*) the exposure of the SP from the ribosome tunnel and the targeting of the translating mRNA to the ER (Figure 6B) (Jan et al., 2014; Kowarik et al., 2002). The premature STOPs did not reduce transcript expression (Figure S6E) or trigger NMD, as

LGALS3BP-STOP$_{30}$-HA mRNA levels remain the same in the presence (scramble shRNA) or absence of UPF1 (UPF1 shRNA; Figure S6F).

We observed that *LGALS3BP-HA* and *LGALS3BP-STOP$_{60}$-HA* mRNA levels increased in the absence of GIGYF1/2, and decreased upon re-expression of 4EHP and GIGYF2 (Figures 6C, D). In contrast, *LGALS3BP-STOP$_{30}$-HA* was more abundant in control cells and its levels did not respond to changes in GIGYF1/2 expression (Figures 6C, D, S6E). Our results indicate that GIGYF1/2-dependent mRNA decay requires the synthesis of the SP and co-translational ER targeting of LGALS3BP-HA.

Binding of GIGYF2 to the signal peptide induces co-translational mRNA decay

We next sought to identify the molecular events leading to the recruitment of GIGYF1/2 to ribosomes translating secretome mRNAs. We first observed that GFP-GIGYF2 co-purified with LGALS3BP-HA (Figure 6E). The interaction between the two proteins involved the C-terminal region of GIGYF2 (Figure 6E, lane 8). Moreover, the SP was sufficient to mediate an interaction with GIGYF2. Fusion of the LGALS3BP, or NPTX1, SPs to the N-terminus of R-LUC (Figure 6B) prompted an interaction between the luciferase and GIGYF2 (Figure 6F, lanes 10 and 12). In contrast, R-LUC lacking the SP does not associate with GIGYF2 (Figure 6F, lane 8).

GIGYF2 also efficiently associated with the *LGALS3BP-HA* mRNA (Figures S6G, H). Binding to the mRNA was greatly impaired if a STOP codon was present at the beginning of the CDS (*STOP$_3$*) to prevent the synthesis and exposure of the SP (Figures 4A, S6G, H). These results suggest that GIGYF2 interacts with the nascent SP of translating *LGALS3BP* and *NPTX1* mRNAs.

We then tested if the presence of the SP was sufficient to induce GIGYF1/2-dependent co-translational mRNA decay. We transfected control and null cells with the R-LUC transcripts containing the signal sequence (Figure 6B). By comparison, the abundance of *SP$_{LGALS3BP}$-R-*

LUC and *SP_{NPTX1}-R-LUC* mRNAs increased in GIGYF1/2-null cells (Figures 6G, lanes 5 and 8, H). Re-expression of GIGYF2 and 4EHP was sufficient to decrease transcript levels in the null cells (Figures 6G, lanes 6 and 9, H). In contrast, *R-LUC* abundance did not vary in the absence of GIGYF1/2 (Figures 6G, lanes 1-3, H).

Collectively, these data support a model in which co-translational binding of GIGYF1/2 to the SP elicits decay of a subset of secretome mRNAs. GIGYF1/2-mediated mRNA degradation is most likely favored by changes in elongation during synthesis of the SP.

α- and β-tubulin mRNAs are regulated by GIGYF1/2

α- and β-tubulin mRNAs are co-translationally degraded in cells with excess of soluble subunits (Cleveland et al., 1981; Lin et al., 2020). *TUBA4A* and *TUBB4A* are 4EHP and GIGYF1/2 target mRNAs (Tables S2-S4). Using RT-qPCR, we confirmed that *TUBA4A* and *TUBB4A* abundance, but not *TUBB*, increases in GIGYF1/2-null cells (Figures 7A-C). Consistent with a role of the complex in tubulin mRNA degradation, selective reduction of *TUBA4A* and *TUBB4A* levels was achieved in the null cells by re-expression of 4EHP and GIGYF2 (Figures 7A, B). Furthermore, *TUBA4A* was efficiently bound by GIGYF2 via its GYF domain (Figure 7D).

We also examined if GIGYF1/2 regulate tubulin mRNA abundance in response to the level of free tubulin subunits. Pre-treatment of control cells with the microtubule destabilizing agent nocodazole elicited the decay of several α- and β-tubulin mRNAs, as their levels dropped considerably (Figures 7E-I). Degradation of tubulin mRNAs in the presence of nocodazole was impaired in GIGYF1/2-null cells, with autoregulation of *TUBA4A*, *TUBA1A* and *TUBB* mRNAs being most affected. The abundance of these transcripts mostly doubled in the null cells in the presence of nocodazole (Figures 7E-G). Nocodazole-induced decay of *TUBB4A* and *TUBA1B* transcripts was less sensitive to GIGYF1/2 absence (Figures 7H, I). Thus, GIGYF1/2 participate in tubulin autoregulation.

While tubulin autoregulation occurs co-translationally (Gay et al., 1989; Pachter et al., 1987), it remains unclear if the underlying mechanism is coupled with translational surveillance. Conventional ribosome profiling does not uncover ribosome pausing in tubulin mRNAs (Figures 7J, K). However, analysis of ribosome footprints following ribosome run off assays (Lee et al., 2012), and disome profiling in human HEK293 cells (Han et al., 2020) or mouse ESCs (Tuck et al., 2020) shows that ribosome stalling and collision are frequent in translating *TUBA4A, TUBA1A* and *TUBA1B* mRNAs (Figures 7J, K, S7A-D). Interestingly, the position of the stalled and collided ribosomes occurs 20 to 30 amino acids after the motif at the N-termini of tubulins with a critical role in autoregulation (Figures 7J, K, S7A-D) (Pachter et al., 1987; Yen et al., 1988). These observations suggest that 4EHP and GIGYF1/2 trigger the decay α- and β-tubulin mRNAs with perturbed translation elongation.

Collectively, our work shows that recognition of the nascent peptide chain by surveillance factors such as GIGYF1/2 and detection of ribosome pausing and collisions during translation trigger mRNA decay (Figure S7E).

Discussion

This study shows that 4EHP and GIGYF1/2 are selective regulators of mRNA turnover. Targeted transcripts frequently encode membrane-bound and secreted proteins, implicating 4EHP and GIGYF1/2 in the regulation of a subset of secretome mRNAs. Degradation is coupled to translation and is triggered by changes in ribosome activity during elongation. This function expands the role of the complex as a regulator of gene expression beyond the 3' UTR-directed mechanisms operating during inflammation or miRNA-mediated gene silencing (Fu et al., 2016; Schopp et al., 2017; Tollenaere et al., 2019). Our findings have multiple implications for both translational control and mRNA decay and open future research directions.

Co-translational mRNA decay by GIGYF1/2 proteins

We find that co-translational mRNA degradation requires the coordinated action of several GIGYF1/2 co-factors. Binding and mRNA selection rely on GYF domain interacting proteins, such as ZNF598 which recruits 4EHP and GIGYF1/2 to destabilize transcripts marked by ribosome collisions. An alternate mode of selection relies on the co-translational binding of GIGYF2 (or co-factors) to the nascent peptide chain of the translating mRNA as it emerges from the ribosome exit tunnel. The diversity of mechanisms for target recognition centralized on GIGYF1/2 proteins opens the possibility that mRNA decay is subject to regulation.

GIGYF1/2-directed recruitment of 4EHP to the mRNA cap not only reduces translation initiation but also facilitates the activity of the decay machinery. The reduced cap affinity of 4EHP compared with eIF4E (Chapat et al., 2017; Peter et al., 2017; Rom et al., 1998; Zuberek et al., 2007) exposes the mRNA to decapping (Ruscica et al., 2019). A scenario where the recruitment of deadenylation and decapping factors by GIGYF1/2 occurs co-translationally is in agreement with the ribosomal association and activity of decay factors such as DDX6 (Sweet et al., 2012), the CCR4-NOT complex (Buschauer et al., 2020) and XRN1 (Pelechano et al.,

2015; Tesina et al., 2019; Tuck et al., 2020), and would irreversibly prevent the translation of transcripts with impaired elongation.

Our data further support a role for DDX6 in GIGYF1/2 dependent mRNA decay and suggest that this RNA helicase might also monitor ribosome speed in the targeted transcripts, as demonstrated for mRNAs with poor codon optimality (Radhakrishnan et al., 2016).

GIGYF1/2 mediate mRNA decay in response to disturbed elongation

Here, we present multiple evidences that changes in ribosome dynamics during elongation trigger canonical mRNA degradation by GIGYF1/2. Ribosome pausing and queuing prevailed in 4EHP and GIGYF1/2 target mRNAs. Pause sites included known stalling sequences, such as the Pro-rich motif in *LARGE2*, and were as well associated with factors recognizing the nascent peptide. Co-translational target mRNA decay was in part mediated by ZNF598, a sensor of ribosome collisions, and in the absence of ribosome stalling, the abundance of target-based reporter transcripts was no longer regulated by GIGYF1/2. In addition, degradation of some of the GIGYF1/2 targets, such as α- and β-tubulins and secretome mRNAs, is known to be dependent on translation and ribosome-associated factors (Cleveland et al., 1981; Karamyshev et al., 2014; Lakshminarayan et al., 2020; Lin et al., 2020).

Our findings suggest a model where selective recruitment of 4EHP and GIGYF1/2 to mRNAs with altered elongation promotes translation repression and mRNA degradation. Target selection involves the recognition of stalled ribosomes by specialized co-factors and/or of the nascent peptide chain by surveillance proteins and GIGYF1/2 (Figure S7E). In line with a role in ribosome-coupled mRNA decay, a recent pre-published report suggests that 4EHP and GIGYF2 are components of the surveillance machinery inhibiting the expression of aberrant mRNAs (Hickey et al., 2019).

GIGYF1/2 and disease

Translation-dependent canonical mRNA degradation has been implicated in different cellular events. Autoregulation of tubulin mRNA abundance is crucial for proper cell division (Lin et al., 2020). Moreover, mutations in the autoregulatory domain of TUBB4A which abolish translation-coupled mRNA decay, have been described in hereditary dystonia (Hersheson et al., 2013). As shown in this work, 4EHP and GIGYF1/2 participate in the turnover of different tubulin mRNAs. Additional studies are now required to identify the precise mechanism that link the recruitment of the 4EHP–GIGYF1/2 complex and the binding of TTC5 to the tubulin nascent peptide when cells activate autoregulation (Lin et al., 2020). Furthermore, our findings suggest that in the absence of GIGYF1/2 cells might be more prone to defects in division.

Co-translational mRNA decay likewise guarantees the quality of secretory and membrane proteins. To reduce the accumulation of misfolded and potentially toxic proteins, failure in protein targeting to the ER elicits degradation of the ribosome-bound message (Karamyshev et al., 2014; Lakshminarayan et al., 2020; Pinarbasi et al., 2018). Here, we identify the 4EHP–GIGYF1/2 complex as a member of the quality control system that regulates the turnover of a subset of secretome-associated mRNAs during co-translational assembly in the ER. We propose that 4EHP and GIGYF1/2 could likewise be important to trigger the degradation of specific mRNAs when the nascent chain has reduced ability for ER targeting or folding.

The selective recognition of secretome transcripts appears to rely on the signal peptide (SP). Given the ability to associate with polysomes and ribosomal proteins, one possibility is that like SRP (Voorhees and Hegde, 2015), GIGYF1/2 is co-translationally recruited at the initial stages of membrane/secreted proteins synthesis, scanning translating ribosomes as the SP elongates through the exit tunnel and interferes with the translation cycle. If ER targeting fails, the C-terminal region of GIGYF2 is then posed to bind the exposed SP and elicit decay of the translating mRNA, avoiding the accumulation of misfolded proteins.

Although the causal mechanisms remain unknown, GIGYF1/2 loss and haploinsufficiency are associated with neurodegeneration and neurological disorders in animal models and affected humans (Giovannone et al., 2009; Iossifov et al., 2014; Krumm et al., 2015; Satterstrom et al., 2020; Schizophrenia Working Group of the Psychiatric Genomics, 2014; Thyme et al., 2019). Our work reports a previously unappreciated role of GIGYF1/2 in safeguarding the integrity of the proteome by signalling to conventional decay mRNAs with altered ribosome progression. This function of GIGYF1/2 proteins prevents the synthesis of unwanted or potentially cytotoxic proteins and, if compromised, may contribute to the development of neurological diseases.

Limitations

Our work highlights 4EHP and GIGYF1/2 as regulators of co-translational mRNA decay. To extend our observations to disease-related contexts, identification of the transcripts and ribosome-associated mechanisms regulated by 4EHP and GIGYF1/2 in models of human neurological pathologies and ER-related stress are still necessary. Such studies will greatly increase our knowledge on how translation-coupled mRNA decay tunes the cellular proteome.

Weber et al.

Acknowledgments

Author Contributions

R.W. designed and conducted the experiments assisted by M.-Y.C. and C. K.; U.Z. and M.L. helped with Ribo-Seq and RNA-Seq experiments and contributed to data analysis; E.V. contributed to data analysis. E.I. was the principle investigator. R.W. and C.I. conceived the project, interpreted the results, and wrote the manuscript. All authors, with exception of E.I., read and corrected the manuscript.

The authors declare no competing interests.

Figure legends

Figure 1. The 4EHP–GIGYF1/2 complexes regulate mRNA abundance

(A, B) Genome-wide analysis of ribosome footprints (RFP) and mRNA abundance changes in the GIGYF1/2-null (KO) and 4EHP-null cells relative to control (Ctrl) cells. Logarithmic change in RFP ($\log_2$FC) on the vertical axis is plotted as a scatter graph against the $\log_2$FC of the mRNA abundance. Each dot represents an individual gene (n=9870). Significantly (FDR<0.005) upregulated ($\log_2$FC>0) and downregulated ($\log_2$FC<0) genes are indicated in red and blue, respectively.

(C) Venn diagram showing the number of common and unique genes with increased mRNA abundance in 4EHP and GIGYF1/2 KO cells (n=82; p=$1.4459e^{-34}$ using a hypergeometric test).

(D, E) Gene Ontology analysis on the group of transcripts with increased abundance in GIGYF1/2 or 4EHP KO cells. Bar graphs show $-\log_{10} q$ values for each of the overrepresented category. Values in bracket indicates the % of genes within each category.

(F) Endoplasmic reticulum (ER), membrane or secreted protein products number in 4EHP–GIGYF1/2 target mRNAs. See also Figure S1.

Figure 2. GIGYF1/2 induce mRNA decay

(A-F) Ctrl and GIGYF1/2 KO cells were treated with Actinomycin D (ActD) and harvested at the indicated time points. RNA samples were analyzed by northern blotting (A, B) or RT-qPCR (C-F) and normalized to that of *TUBB* or *18S* rRNA. The value at time zero (before ActD addition) was defined as 100%. Results were plotted as a function of time post ActD addition. Circles represent the mean value; error bars represent the standard deviation (SD) (n=3). The decay curves were fitted to an exponential decay with a single component (dotted lines). R^2 values are indicated for each curve. The half-life of each mRNA in Ctrl and KO cells is represented as the mean +/- SD. *18S* rRNA ethidium bromide staining shows equal loading.

Figure 3. GIGYF1/2 recruit multiple effector proteins to induce mRNA decay

(**A, B**) Ctrl and GIGYF1/2 KO cells were transfected with plasmids expressing λN-HA or λN-HA-4EHP, GFP-MBP, and GFP-GIGYF1/2 (WT or mutants). *DBNDD2* mRNA levels were determined by northern blotting, normalized (Norm.) to *TUBB* and set to 100% in Ctrl cells. *18S* rRNA ethidium bromide staining indicates equal loading. *NPTX1*, *CD109*, *ITPR3* and *β-ACTIN* mRNA levels were quantified by RT-qPCR and normalized to those of *18S* rRNA. Bars represent the mean value; error bars represent SD (n=3). GIGYF1/2 mutants are as follows: C* (4EHP-binding mutant), GYF* (GYF domain mutant) and DDX6* (DDX6-binding mutant).

(**C-F**) GFP-immunoprecipitation (IP) assays were performed two days post cell transfection with GFP-MBP or GFP-GIGYF2 (WT or mutants). mRNA levels in input (0.8%) and IP samples (12%) were quantified by RT-qPCR, normalized over *GAPDH* and set to 100% in the presence of GFP-MBP. Bars indicate the mean value; error bars represent SD (n=3). The length of the CDS of each mRNA is indicated in nucleotides (nt).

(**G**) Immunoblot depicting the expression of the immunoprecipitated GFP proteins. Inputs and immunoprecipitates were 2% and 2.7%, respectively. See also Figure S2.

Figure 4. GIGYF1/2 mediates co-translational mRNA decay

(**A**) Schematic representation of the *DBNDD2-HA* and *LGALS3BP-HA* CDS reporters. Open reading frame (ORF); Hemagglutinin (HA); $STOP_3$ (STOP positioned three codons downstream of AUG); Signal Peptide (SP).

(**B-D**) Ctrl and GIGYF1/2 KO cells were transfected with *DBNDD2-HA* (WT or $STOP_3$) or *LGALS3BP-HA* (WT or $STOP_3$) and *F-LUC-GFP*. Protein samples were analyzed by western blotting using anti-HA and anti-GFP antibodies. (D) DBNDD2-HA protein levels were quantified in Ctrl and KO cells, normalized over to those of F-LUC-GFP and set to 100% in Ctrl cells.

Weber et al.

(E-J) Ctrl and GIGYF1/2 KO cells were transfected with *DBNDD2-HA* (WT or STOP$_3$) (E, G) or *LGALS3BP-HA* (WT or STOP$_3$) (F, H) reporters. Two days post transfection, cells were treated with ActD and harvested at the indicated time points. Panels E to H show representative northern blots. In panels I and J, mRNA levels were quantified as described in Figure 2A. See also Figure S2.

Figure 5. Ribosome pausing triggers GIGYF1/2 dependent mRNA decay

(A) RFP profiles (RFP) of *DBNDD2* in Ctrl, GIGYF1/2 KO, and cells treated with harringtonine and lactimidomycin (LTM) (Lee et al., 2012). Dashed blue box highlights ribosome pausing. Gene annotation, protein sequence and residue numbering are depicted below the profiles.

(B) Schematic representation of the *DBNDD2-HA* reporters. Pause$_{23}$: position of the stall peptide (FED$_{25}$); PPP$_{48}$ and PPP$_{87}$: poly-Pro motifs. STOPs were introduced 18, 39, 79 and 89 codons after the AUG. Alanine substitutions were inserted at the stall site and at the N-terminus of the protein (DPN$_4$).

(C, D) Ctrl and GIGYF1/2 KO cells were transfected with *DBNDD2-HA* (WT or STOPs) and *F-LUC-GFP*. Cells were also co-transfected with GFP-MBP or GFP-GIGYF2 and V5-SBP-4EHP. mRNA levels were determined by northern blotting (C), normalized to *F-LUC-GFP* and set to 100% in Ctrl cells (D). Bars indicate the mean value; error bars represent SD (n=3). The immunoblot showing the expression of the GFP-, HA- and V5-tagged proteins is depicted below the northern blot.

(E) Cells were transfected with *DBNDD2-HA* (WT or DPN$_4$-AAA). Reporter mRNA levels were quantified by RT-PCR, normalized over to those of 18S rRNA and set to 100% for *DBNDD2-HA* WT. Bars indicate the mean value; error bars represent SD (n=3).

Weber et al.

(F) Two days after transfection with *DBNDD2-HA* (WT or DPN$_4$-AAA), cells were treated with ActD and harvested at the indicated time points. mRNA levels were quantified by RT-qPCR as described in Figure 2A. See also Figures S4 and S6.

Figure 6. Binding of GIGYF2 to the signal peptide (SP) triggers mRNA decay

(A) RFP profiles for *LGALS3BP* in Ctrl and GIGYF1/2 KO cells. Also shown are the RFP profiles in the presence of harringtonine and LTM obtained by Lee and co-workers (Lee et al., 2012) and the disome occurrence along the CDS determined by Han et al. (Han et al., 2020). Gene annotation is depicted below the profiles. Met$_1$ (M$_1$); Asp$_{21}$ (D$_{21}$); Val$_{60}$ (V$_{60}$); upstream open reading frame (uORF); signal peptide (SP).

(B) Schematic representation of the *LGALS3BP-HA* and *R-LUC* reporters. STOPs were introduced 30 and 60 codons after the AUG in *LGALS3BP-HA*. The signal sequences of the LGALS3BP and NPTX1 SPs were inserted upstream and in frame with the *Renilla* luciferase (R-LUC) ORF.

(C) *LGALS3BP-HA* (WT and STOPs) mRNA levels were measured by RT-qPCR in samples obtained from Ctrl and GIGYF1/2 KO cells co-transfected with F-LUC-GFP, GFP-MBP or GFP-GIGYF2, and λN-HA-4EHP. mRNA levels were set to 100% in Ctrl cells after normalization to *F-LUC-GFP*. Bars indicate the mean value; error bars represent SD (n=3).

(D) Analysis of GFP and HA proteins expression in Ctrl and GIGYF1/2 KO cells by immunoblotting.

(E, F) The interactions of GFP-GIGYF2 with LGALS3BP-HA, R-LUC, SP$_{LGALS3BP}$-R-LUC or SP$_{NPTX1}$-R-LUC were analyzed by co-IP. Proteins were immunoprecipitated using anti-GFP antibodies. GFP-MBP served as a negative control. Input (0.8% for the GFP proteins, 0.3% for LGALS3BP-HA, and 0.1% for R-LUC) and immunoprecipitated fractions (12% for the GFP proteins and 24% for LGALS3BP-HA and R-LUC) were analyzed by western blotting with anti-GFP, anti-HA or anti-R-LUC antibodies. A schematic representation of the domain

architecture and binding regions of human GIGYF2 is depicted below the immunoblot in panel E. The N-terminal (N-term) region of GIGYF2 contains a 4EHP-binding region (4EHP-BR), a DDX6-binding motif (DBM) and the GYF domain. The C-term is predicted to contain primarily α-helices. The amino acid positions at the domain/motif boundaries are indicated below the protein.

(G, H) *R-LUC*, *SP$_{LGALS3BP}$-R-LUC* or *SP$_{NPTX1}$-R-LUC* abundance was determined after northern blotting in Ctrl and GIGYF1/2 KO cells expressing F-LUC-GFP, GFP-MBP or GFP-GIGYF2, and λN-HA-4EHP. Normalized mRNA levels were set to 100% in control cells. Bars indicate the mean value; error bars represent SD (n=3). The expression of the GFP- and HA-tagged proteins was assessed by immunoblot and is shown below the northern blot.

Figure 7. GIGYF1/2 participate in tubulin co-translational mRNA decay

(A-C) Ctrl and GIGYF1/2 KO cells were transfected with GFP-MBP or GFP-GIGYF2 and λN-HA-4EHP. Tubulin mRNA levels were quantified by RT-qPCR, normalized to those of *18S* rRNA (A, B) or *GAPDH* mRNA (C) and set to 100% in Ctrl cells.

(D) GIGYF2 binding to *TUBA4A* mRNA was determined by RNA-IP as described in Figure 3.

(E-I) Ctrl and GIGYF1/2 KO cells were treated with either DMSO (-) or nocodazole (+) for 3 hours. Tubulin mRNA abundance was quantified by RT-qPCR, normalized to *GAPDH* and set to 100% in the absence of nocodazole. Plotted is the mean ± SD (n=3).

(J, K) RFP profiles of *TUBA4A* and *TUBA1A* in Ctrl, GIGYF1/2 KO, and cells treated with harringtonine or LTM (Lee et al., 2012). The *TUBA1A* profile also shows the distribution of disome footprints along the CDS obtained by Han and co-workers (Han et al., 2020). Dashed squares identify paused ribosomes and disome position. In *TUBA4A* RFP the black horizontal lines indicate footprint peaks that result from non-unique reads with nucleotide sequences common to multiple tubulin subunits. These footprints were not considered as ribosome pauses. Met$_1$ (M$_1$); Trp$_{21}$ (W$_{21}$), Asp$_{21}$ (D$_{21}$).

Weber et al.

STAR Methods

Contact for reagent and resource sharing

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Catia Igreja (catia.igreja@tuebingen.mpg.de).

Experimental model and subject details

Cell lines

All cell lines were cultured at 37°C and 5% CO_2 in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum, 2 mM Glutamine, 1x Penicillin and 1x Streptomycin.

Methods details

DNA constructs

DNA constructs used in this study are listed in the Key Resources Table. All plasmids used in the assays depicted in Figures 3 and S2, the *Renilla* luciferase (R-LUC) and the firefly luciferase (F-LUC)-EGFP reporters were described previously (Lazzaretti et al., 2009; Peter et al., 2017; Pillai et al., 2004). The UPF1 shRNA plasmids were a kind gift from Oliver Mühlemann (Paillusson et al., 2005). To generate the CDS reporters, the sequences of *DBNDD2* or the *LGALS3BP* ORFs were cloned into the NheI-XbaI restriction sites of the pCIneo vector. The C-terminal HA-tag was inserted by site-directed mutagenesis. The *DBNDD2-STOP-HA* reporters with UAA stop codons at various positions (3, 18, 39, 79 and 89 codons downstream of the AUG start site) and the *LGALS3BP-STOP-HA* reporters with UAA stop codons at various positions (3, 30 and 60 codons downstream of the AUG start site) were generated by mutagenesis. To obtain the 3' UTR reporters, the sequences of *DBNDD2* or the *LGALS3BP* 3' UTRs were cloned into the XhoI-NotI and XbaI-NotI restriction sites of the pCIneo-R-LUC vector, respectively. The C-term of GIGYF2 (719-1299) was cloned into the XhoI-BamHI

restriction sites of the pT7-EGFP vector. The eL22L1 (1-122) was cloned into the NheI-XbaI restriction sites of the pCIneo vector; the N-terminal V5-SBP sequence was inserted by mutagenesis. To generate the SP-R-LUC vectors the sequences corresponding to the signal peptides (as annotated by Uniprot) of LGALS3BP (1-18: MTPPRLFWVWLLVAGTQG) and NPTX1 (1-22: MPAGRAARTCALLALCLLGAGA) were cloned upstream of R-LUC ORF in the pCIneo-R-LUC vector by mutagenesis. All the mutants used in this study were generated by site-directed mutagenesis using the QuickChange Site-Directed Mutagenesis kit (Stratagene). All the constructs were confirmed by sequencing.

Generation of the 4EHP-null and ZNF598-null cell lines

sgRNAs targeting 4EHP and ZNF598 were designed using the CHOPCHOP (http://chopchop.cbu.uib.no) online tool (Labun et al., 2016; Labun et al., 2019; Montague et al., 2014) and cloned into the pSpCas9(BB)-2A-Puro (PX459) vector [a gift from F. Zhang, Addgene plasmid 48139; (Ran et al., 2013)]. Clonal cell lines were obtained and confirmed for gene editing as described previously (Peter et al., 2017). Briefly, HEK293T cells were transfected with the sgRNA-Cas9 vectors. Two days post transfection, cells were treated with puromycin (3 µg/ml; Serva Electrophoresis) to select for edited cells. Serial dilutions in 96-well plates were used to obtain single cell clones. Genomic DNA was isolated from single clones using the Wizard SV Genomic DNA Purification System (Promega). The 4EHP locus was PCR amplified and Sanger sequencing of the targeted genomic regions confirmed two frameshift mutations in exon 4 (an 11 nucleotide and a 37 nucleotide deletions) targeted by sg4EHP-a. For sg4EHP-b we did not observe gene editing; the amplified sequence around the target site in exon 2 is wild type. The ZNF598 locus was targeted by sgZNF598-a (exon 3) and sgZNF598-b (exon 4). RNA sequencing shows that in ZNF598 KO cells genome editing resulted in the expression of a ZNF598 transcript that lacks exons 1-4 and retains intron 4 at the 5' end. This transcript has reduced translation efficiency and is subject to degradation. The lack of 4EHP

and ZNF598 expression was confirmed by western blotting (Figures S1A, S5A). See Table S5 for sgRNA sequences.

Ribosome profiling and RNA sequencing

HEK293T (DSMZ, ACC 635) wild type, GIGYF1/2-null (Peter et al., 2017), 4EHP-null (Rasch et al., 2020) and ZNF598-null cells were plated on 10 cm dishes 24 hours before harvesting, as previously described (Calviello et al., 2016). Cells were lysed with lysis buffer (20 mM Tris-HCl pH=7.4, 150 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT, 1% TritonX-100, 0.5% NP40) containing cyclohexamide (100 µg/ml). Lysates were then used for total RNA extraction and ribosome profiling (1/4 of the lysate for each). Total RNA was extracted using the RNeasy Mini Kit (50) (Qiagen) after pre-treating the lysate with 10 U TurboDNase (Thermo Scientific). cDNA libraries were prepared using the TruSeq RNA Sample Prep Kit (Illumina), according to the manufacturers' instructions.

Ribosome profiling was performed according to the original protocol (Ingolia et al., 2012) with the modifications described in (Calviello et al., 2016). Cell lysates were treated with 300 U RNase 1 (Thermo Scientific). Reactions were stopped after 45 min incubation at room temperature by adding 80 U SUPERase Inhibitor (Thermo Scientific). The RNase 1-treated samples were applied to MicroSpin S-400 HR columns (GE Healthcare) to remove free nucleotides and recover the ribosome-protected RNA. RNA extraction was then performed with TriFast FL (Peqlab Biotechnologies) and the RNA Clean & Concentrator Kit (Zymo Research). rRNA was depleted using the Ribo-Zero Gold rRNA Depletion Kit (Illumina, discontinued). Ribosome footprints were excised and extracted from a 17% TBE-Urea gel using 30 and 27 nt RNA oligonucleotides as markers. Ribosome footprints were treated with T4 PNK (NEB) and purified using P:C:I (PanReac AppliChem). 3' and 5' adapters were ligated using T4 RNA ligase 2, truncated K227Q (NEB) and T4 RNA ligase 1 (NEB), respectively. Following adapter ligation, the resulting ribosome footprints were excised and extracted from a 15% TBE-Urea

gel. Adapter-ligated ribosome footprints were reversed transcribed with SuperScript III (Thermo Scientific). cDNA was PCR amplified using Phusion DNA polymerase (Thermo Scientific). PCR amplicons were visualized on a 2.5% low melting agarose (Serva Electrophoresis), excised and purified using the Zymoclean Gel DNA Recovery Kit (Zymo Research). DNA and sample quality were assessed using the Bioanalyzer system (Agilent). The sequences of the oligonucleotides used in this protocol are listed in Table S8.

Two biological replicates were analyzed. The ribosome profiling and total RNA libraries were sequenced using the Hiseq 3000 sequencing system (Illumina). Ribosomal RNA reads were filtered using Bowtie 2 (Langmead and Salzberg, 2012). Remaining reads were mapped on the hg19 (UCSC) human genome with TopHat2 (Kim et al., 2013). For RNA sequencing, 17.0-21.5 million reads were mapped (>87%). Ribosome profiling reads were analyzed for three-nucleotide periodicity using the RiboTaper program to identify actively translating ribosomes (Calviello et al., 2016). Reads corresponding to the lengths of 29 and 30 nucleotides were selected as they showed the most significant three nucleotide periodicity and were then used for subsequent mapping on the human genome with TopHat2. For ribosome profiling, 6.1-9.6 million reads (>95%) of input reads were mapped. Read count analysis was performed with the R/Bioconductor package QuasR (Gaidatzis et al., 2015). Differential expression analysis and multidimensional scaling (MDS) analysis were conducted using edgeR (McCarthy et al., 2012; Robinson et al., 2010) for selected genes with a threshold of 'fragments per kilobase of transcript per million mapped reads' (FPKM) >2. Translation efficiency (TE) was calculated with the RiboDiff program (Zhong et al., 2017).

Harringtonine and LTM datasets from human HEK293 cells were downloaded from the Sequence Read Archive database (accession: SRA056377). The mouse ESC disome and human HEK293 cells disome datasets were retrieved from the GEO database. The respective accession numbers are GSE134020 and GSE145723. Ribosomal RNA reads were filtered using Bowtie 2 (Langmead and Salzberg, 2012). The remaining reads were mapped on the hg19 (UCSC)

human genome or the mm9 (UCSC) mouse genome with TopHat2 (Kim et al., 2013). No specific filters for read length were applied.

Data analysis

Upregulated and downregulated gene groups were defined as being significantly regulated (FDR<0.005) with a $\log_2$FC>0 and $\log_2$FC<0, respectively. No cutoff on the $\log_2$FC value was applied so that genes with little but significant changes could also be detected. GO analysis was performed with the R based package GOseq (Young et al., 2010). The % of genes within each category corresponds to the number of genes belonging to the category and upregulated in the null cells divided by the total number of upregulated genes in HEK293T cells.

UniProt information was retrieved to analyze the presence of a signal peptide or the cellular location of the proteins encoded by 4EHP and GIGYF1/2 target mRNAs. Ribosome footprint density plots for individual sequencing tracks were visualized using the Integrative Genomics Viewer (IGV) visualization tool (Robinson et al., 2011; Thorvaldsdottir et al., 2013).

Ribosome pause scores were determined for each of the common and upregulated mRNAs in 4EHP- and GIGYF1/2-null cells. Maximum (pause site) and median RFP coverage in the CDS of each transcript was retrieved using UCSC annotation and Ribo-Seq in GIGYF1/2-null cells replicate number 1. The pause score refers to the reads at the pause position divided by median reads in the gene. The values are listed in Table S4.

Transfections, northern and western blotting

In the rescue assays described in Figures. 3, 5-7, S2 and S4, 0.64×10^6 Ctrl cells or 0.7×10^6 null cells were transfected, after seeding in 6-well plates, using Lipofectamine 2000 (Invitrogen). The transfection mixtures contained different amounts of the plasmids expressing λN-HA- or V5-SBP-fusion proteins (**λN-HA/V5-SBP-MBP**: 0.25 μg; **4EHP**: 0.25 μg of WT

and cap* mutant, and 0.35 µg of S* mutant) and the GFP-fusion proteins (**MBP**: 0.4 µg, **GIGYF1**: 0.5 µg of WT, C*, GYF* and DDX6* mutants; **GIGYF2**: 1.75 µg of WT, 1.1 µg of C* or 1.35 µg of GYF* and DDX6* mutants). In the experiments shown in Figures 4-6, the transfection mixtures contained plasmids expressing DBNDD2-HA, DBNDD2-STOP$_x$-HA, DBNDD2-FDE$_{25}$-AAA-HA or DBNDD2-DPN$_4$-AAA-HA (0.2 µg), and LGALS3BP-HA, LGALS3BP-STOP$_x$-HA, R-LUC, SP$_{LGALS3BP}$-R-LUC or SP$_{NPTX1}$-R-LUC (0.5 µg). In the assay with the 3' UTR reporters, the transfection mixtures contained 0.5 µg of R-LUC-DBNDD2-3' UTR, R-LUC-LGALS3BP-3' UTR or R-LUC, and 0.25 µg of F-LUC-GFP plasmid DNA.

Cells were harvested two days after transfection and firefly and *Renilla* luciferase activities were measured using the Dual-Luciferase reporter assay system (Promega). Total RNA was isolated using TriFast (Peqlab biotechnologies). For northern blotting, total RNA was separated in 2% glyoxal agarose gels and blotted onto a positively charged nylon membrane (GeneScreen Plus, Perkin Elmer). [^{32}P]-labelled probes specific for each transcript were generated by linear PCR. Hybridizations were carried out in hybridization solution (0.5 M NaP pH=7.0, 7% SDS, 1 mM EDTA pH=8.0) at 65°C overnight. After extensive washes with washing solution (40 mM NaP pH=7.0, 1% SDS, 1 mM EDTA pH=8.0), the membranes were exposed and band intensities were quantified by PhosphoImager. For detection of *DBNDD2* and *LGALS3BP* cellular and reporter mRNAs, complementary and radioactively labelled probes were designed against the CDS of the transcripts. Since the reporter constructs only harbor the CDS but no 5' and 3' sequences, the endogenous mRNAs are expected to run slower on an agarose gel. However, we observed that the signal of transfected reporters is considerably stronger and does not allow the simultaneous detection of cellular and reporter mRNAs.

To test for tubulin mRNA autoregulation, control and GIGYF1/2-null HEK293T cells were grown to 70 % confluency and treated with nocodazole (10 µM, Sigma Aldrich) or DMSO for 3 hours as described previously (Lin et al., 2020). RNA was extracted using TRIzol (Thermo

Scientific), reverse-transcribed and analyzed by quantitative PCR (qPCR; 10% of each RNA sample) as described below.

Western blot was performed using standard methods. In brief, cells were washed with PBS and lysed with sample buffer (100 mM Tris-HCl pH=6.8, 4% SDS, 20% glycerol, 0.2 M DTT) followed by boiling 5 minutes at 95°C and vortexing to shear genomic DNA. After SDS-PAGE, proteins were transferred onto a nitrocellulose membrane (Santa Cruz Biotechnology) by tank transfer. Primary antibodies were incubated overnight at 4°C, secondary antibodies for an hour at room temperature. All western blots were developed with freshly mixed 10A: 1B ECL solutions and 0.01% H_2O_2 [Solution A: 0.025 % Luminol (Roth) in 0.1 M Tris-HCl pH=8.6; Solution B: 0.11% P-Coumaric acid (Sigma Aldrich) in DMSO]. Antibodies used in this study are listed in the Key Resources Table. DBNDD2-HA and LGALS3BP-HA band intensities were quantified using ImageJ (Schneider et al., 2012) and normalized to the band intensities of F-LUC-GFP protein in the same experiment.

Reverse transcription (RT) and quantitative PCR (qPCR)

1 µg of RNA was mixed with 0.66 µg random hexamer primers (N_6) and denatured at 72°C for 5 min. After addition of a reaction mixture containing a final concentration of 1 x RT buffer, 20 U RiboLock RNase Inhibitor (Thermo Scientific) and 1 mM dNTPs, the RNA samples were incubated at 37°C for 5 min. Incubation with RevertAid H Minus Reverse Transcriptase (200 U, Thermo Scientific) was first performed for 10 min at 25°C, and then at 42°C for one hour. The RT reaction was stopped by incubating the samples for 10 min at 70°C. The qPCR was performed with 1x iTaq SYBR Green Supermix (Biorad), 0.4 µM of each primer and 1 µl of the cDNA sample. mRNA levels were determined by qPCR using sequence-specific primers for the indicated transcripts and normalized to *18S* rRNA or *GAPDH* mRNA abundance in the same sample. qPCR primers designed using Primer-BLAST (NCBI) are listed in Table

S8. Normalized transcript expression ratios from three independent experiments were determined using the Livak method (Livak and Schmittgen, 2001).

Half-life experiments

To measure mRNA decay rates, cells were treated with Actinomycin D (10 µg/ml final concentration) two days post transfection or three days after seeding and collected at the indicated time points. mRNA levels determined by northern blotting or qPCR were normalized to the levels of *TUBB* or *18S* rRNA, respectively. Steady state *TUBB* mRNA levels remain unchanged in the absence of GIGYF1/2 (Figure 7C). These values were set to 100 at time point zero. Data points from three independent experiments were plotted and the resulting fitting curves were determined using a one phase exponential decay equation. The R^2 values associated with the fitting of the exponential decay curves were between 0.29 and 0.99. The curves with low R^2 indicate that reduction of mRNA levels over time are not well represented by an exponential decay model whereas high R^2 values indicate that the quantity of mRNA decreases at a rate proportional to its current value. To determine the time required for the decaying quantity to fall to half of its initial value, or half-life, a decay curve was first determined for each replica. The three values were then averaged to have the final half-life value. The three values were also used to determine the error (standard deviation) associated with the measurements. The standard deviation in the half-live values therefore reflects how reproducible the three replicas were.

Polysome profiling

Polysome profiles were performed as described before (Kuzuoglu-Ozturk et al., 2016). HEK293T cells were pretreated with cycloheximide (50 µg/ml) for 30 min. Lysates were prepared in lysis buffer (10 mM Tris-HCl pH=7.4, 10 mM NaCl, 1.5 mM $MgCl_2$, 0.5% Triton X-100, 2 mM DTT, 50 µg/ml cycloheximide) and polysomes separated on a 10-50% sucrose

gradient in gradient buffer (10 mM Tris-HCl pH 7.4, 75 mM KCl, 1.5 mM MgCl$_2$). Polysome fractions were collected using the Teledyne Isco Density Gradient Fractionation System. Protein from sucrose fractions was isolated by methanol extraction. In detail, 4x volumes of MetOH were mixed with the sucrose fractions, then mixed with 1x volume of chloroform and then with 3x volumes of water. After centrifugation, the upper phase was removed leaving the lower and inter-phases which were precipitated using 3x volumes of MetOH. Samples were spun down and the dried pellet dissolved in 2x protein sample buffer. Fractions were analyzed by western blotting.

RNA immunoprecipitation/pulldown

To immunoprecipitate GIGYF2-bound mRNA, 3 x 10^6 HEK293T cells were transfected using Lipofectamine 2000 (Invitrogen) 24 hours after seeding in 10 cm plates. The transfection mixtures contained the plasmid expressing the GFP-fusion proteins (MBP: 2.5 µg, **GIGYF2**: 10 µg of WT, 6 µg of C*, 12 µg of GYF* and 10 µg DDX6* mutants). In Figure S3, the RNA-IP was performed in cells treated with 2 µg/ml harringtonine for 30 min or 200 µg/ml puromycin for 45 min. In Figure S6, GFP-tagged MBP or GIGYF2 were co-expressed with LGALS3BP-HA or LGALS3BP-STOP$_3$-HA (5 µg). To pulldown 4EHP-bound mRNA (Figure S2), cells were co-transfected with the plasmids encoding V5-SBP-fusion proteins (MBP: 1 µg, **4EHP**: 12 µg of WT together with MBP and GIGYF2 C*, or 0.5 µg together with GIGYF2 WT, GYF* and DDX6*) and GFP-fusion proteins (MBP: 1 µg, **GIGYF2**: 5 µg of WT and DDX6*, 4 µg of C* and 8 µg of GYF* mutants). Cells were harvested 48 hours post transfection, washed with ice cold PBS and lysed on ice for 15 minutes in 500 µl of NET buffer [50 mM Tris-HCl pH=7.5, 150 mM NaCl, 0.1% Triton X-100, 1 mM EDTA pH=8.0, 10 % glycerol, supplemented with 1x protease inhibitors (Roche)]. Cell debris was removed by centrifugation at 16,000 g at 4°C. Input samples (5% of the total) were collected for western blotting and RT-qPCR. To immunoprecipitate GFP-GIGYF2, the remaining lysate was then incubated with 3 µl

of anti-GFP antibody (homemade) for an hour, followed by incubation (2 hours) with protein G sepharose resin pre-treated with yeast RNA (250 μg of yeast RNA/100 μl of 50% slurry). For pulldown of SBP-V5-4EHP and associated RNA, cell lysates were immediately incubated with 50 μl of a 50% slurry of streptavidin beads pre-incubated with yeast RNA. Beads were washed 3 times with NET buffer and resuspended in 1 ml of NET buffer without detergent. An aliquot (20% of the total) of the bead suspension, was mixed with SDS-PAGE sample buffer for western blotting after centrifugation to pellet the resin. The remaining beads were used for RNA isolation with TriFast (Peqlab Biotechnologies). cDNA of the input and precipitated fractions (20% each) was prepared and analyzed using qPCR (5% of the cDNA) as described above.

Co-immunoprecipitation (Co-IP) assays

Co-IP assays were performed in the presence of RNase A as described previously (Peter et al., 2015). HEK293T cells were grown in 10 cm dishes and transfected using Lipofectamine 2000 (Invitrogen) according to the manufacturer's recommendations. The transfection mixtures in Figures S1I and J contained 2.5 μg of GFP-MBP or GFP-GIGYF1 WT, 3 μg of GFP-GIGYF1 GYF*, 10 μg of GFP-GIGYF2 WT or 12 μg of GFP-GIGYF2 GYF*. In Figure S3B, GFP-MBP or GFP-GIGYF2 were co-expressed with V5-SBP-eL22L1 (5 μg). The transfection mixture in Figure 6E had 2.5 μg of GFP-MBP, 10 μg of GFP-GIGYF2 WT or C-term, 1.5 μg of GFP-GIGYF2 N-term and 5 μg of LGALS3BP-HA. In Figure 6F, GFP-MBP or GFP-GIGYF2 were transfected with R-LUC (2.5 μg) or SP_x-R-LUC (5 μg). After transfection, cells were treated as described in the RNA-IP section, with the exception that the protein G sepharose resin was not incubated with yeast RNA and the samples were only used for protein analysis.

UPF1 Knockdown

In the reporter assays described in Figures S6B, C and F, 0.64×10^6 Ctrl cells were transfected with 2 μg pSUPER-puro scramble control or UPF1 shRNA, after seeding in 6-well

Weber et al.

plates, using Lipofectamine 2000 (Invitrogen). 24 hours after transfection cells were treated with 5 µM puromycin for 24 hours. Selected cells were re-seeded and re-transfected with the transfection mixtures described above (Transfections section).

Quantification and statistical analyses

Figures 1A and B, S5B and C. Upregulated and downregulated genes were identified using $\log_2$Fold Change (FC) between null and control cells > 0 or < 0, respectively, and False Discovery Rates (FDR) < 0.005.

Figure 1C, S1G and S5D. The hypergeometric test (phyper) in R was applied to estimate the likelihood of list overlap.

Figures 1D and E. The quantitative value represented in the graphs corresponds to -$\log_{10}$(q-value) determined by the GOseq analysis tool (Young et al., 2010).

Figures 2, 4, 5 and S5. Dots represent mean value; error bars represent the standard deviation from three independent experiments. The mRNA decay curves were fitted to an exponential decay with a single component. R^2 values are indicated for each curve.

Figures 3, 4, 5, 6, 7, S2, S3, S4 and S6. The quantitative value that is graphed represents the mean mRNA or protein level values; error bars represent standard deviations from three independent experiments. In the RT-qPCR experiments, normalized transcript expression ratios from three independent experiments were determined using the Livak method (Livak and Schmittgen, 2001).

Data availability

The datasets generated during this study are available at Gene Expression Omnibus (GEO: GSE14484 and GSE149279).

Weber et al.

Supplemental Information

Supplemental Information includes seven figures and five tables (S1-S4 are excel files).

Table S1. Genes with changes in translation efficiency (TE)

Table S2. RNA-Seq and Ribo-Seq analysis in 4EHP-null cells

Table S3. RNA-Seq and Ribo-Seq analysis in GIGYF1/2-null cells

Table S4. Genes commonly upregulated in 4EHP- and GIGYF1/2-null cells

Table S5. RNA-Seq and Ribo-Seq analysis in ZNF598-null cells

Table S6. Genes commonly upregulated in ZNF598- and GIGYF1/2-null cells

Table S7. Genes commonly upregulated in ZNF598- and 4EHP-null cells

Table S8. Primers used in this study

Weber et al.

References

Amaya Ramirez, C.C., Hubbe, P., Mandel, N., and Bethune, J. (2018). 4EHP-independent repression of endogenous mRNAs by the RNA-binding protein GIGYF2. Nucleic Acids Res *46*, 5792-5808.

Arpat, A.B., Liechti, A., De Matos, M., Dreos, R., Janich, P., and Gatfield, D. (2019). Transcriptome-wide sites of collided ribosomes reveal principles of translational pausing. bioRxiv.

Ash, M.R., Faelber, K., Kosslick, D., Albert, G.I., Roske, Y., Kofler, M., Schuemann, M., Krause, E., and Freund, C. (2010). Conserved beta-hairpin recognition by the GYF domains of Smy2 and GIGYF2 in mRNA surveillance and vesicular transport complexes. Structure *18*, 944-954.

Brandman, O., and Hegde, R.S. (2016). Ribosome-associated protein quality control. Nat Struct Mol Biol *23*, 7-15.

Buschauer, R., Matsuo, Y., Sugiyama, T., Chen, Y.H., Alhusaini, N., Sweet, T., Ikeuchi, K., Cheng, J., Matsuki, Y., Nobuta, R.*, et al.* (2020). The Ccr4-Not complex monitors the translating ribosome for codon optimality. Science *368*.

Buskirk, A.R., and Green, R. (2017). Ribosome pausing, arrest and rescue in bacteria and eukaryotes. Philos Trans R Soc Lond B Biol Sci *372*.

Calviello, L., Mukherjee, N., Wyler, E., Zauber, H., Hirsekorn, A., Selbach, M., Landthaler, M., Obermayer, B., and Ohler, U. (2016). Detecting actively translated open reading frames in ribosome profiling data. Nat Methods *13*, 165-170.

Chapat, C., Jafarnejad, S.M., Matta-Camacho, E., Hesketh, G.G., Gelbart, I.A., Attig, J., Gkogkas, C.G., Alain, T., Stern-Ginossar, N., Fabian, M.R.*, et al.* (2017). Cap-binding protein 4EHP effects translation silencing by microRNAs. Proc Natl Acad Sci U S A *114*, 5425-5430.

Cleveland, D.W., Lopata, M.A., Sherline, P., and Kirschner, M.W. (1981). Unpolymerized tubulin modulates the level of tubulin mRNAs. Cell *25*, 537-546.

D'Orazio, K.N., Wu, C.C., Sinha, N., Loll-Krippleber, R., Brown, G.W., and Green, R. (2019). The endonuclease Cue2 cleaves mRNAs at stalled ribosomes during No Go Decay. Elife *8*.

Doerfel, L.K., Wohlgemuth, I., Kothe, C., Peske, F., Urlaub, H., and Rodnina, M.V. (2013). EF-P is essential for rapid synthesis of proteins containing consecutive proline residues. Science *339*, 85-88.

Fu, R., Olsen, M.T., Webb, K., Bennett, E.J., and Lykke-Andersen, J. (2016). Recruitment of the 4EHP-GYF2 cap-binding complex to tetraproline motifs of tristetraprolin promotes repression and degradation of mRNAs with AU-rich elements. RNA *22*, 373-382.

Gaidatzis, D., Lerch, A., Hahne, F., and Stadler, M.B. (2015). QuasR: quantification and annotation of short reads in R. Bioinformatics *31*, 1130-1132.

Gardin, J., Yeasmin, R., Yurovsky, A., Cai, Y., Skiena, S., and Futcher, B. (2014). Measurement of average decoding rates of the 61 sense codons in vivo. Elife *3*.

Garreau de Loubresse, N., Prokhorova, I., Holtkamp, W., Rodnina, M.V., Yusupova, G., and Yusupov, M. (2014). Structural basis for the inhibition of the eukaryotic ribosome. Nature *513*, 517-522.

Garzia, A., Jafarnejad, S.M., Meyer, C., Chapat, C., Gogakos, T., Morozov, P., Amiri, M., Shapiro, M., Molina, H., Tuschl, T*., et al.* (2017). The E3 ubiquitin ligase and RNA-binding protein ZNF598 orchestrates ribosome quality control of premature polyadenylated mRNAs. Nat Commun *8*, 16056.

Gasic, I., Boswell, S.A., and Mitchison, T.J. (2019). Tubulin mRNA stability is sensitive to change in microtubule dynamics caused by multiple physiological and toxic cues. PLoS Biol *17*, e3000225.

Gay, D.A., Sisodia, S.S., and Cleveland, D.W. (1989). Autoregulatory control of beta-tubulin mRNA stability is linked to translation elongation. Proc Natl Acad Sci U S A *86*, 5763-5767.

Gerbracht, J.V., Boehm, V., and Gehring, N.H. (2017). Plasmid transfection influences the readout of nonsense-mediated mRNA decay reporter assays in human cells. Sci Rep *7*, 10616.

Giovannone, B., Tsiaras, W.G., de la Monte, S., Klysik, J., Lautier, C., Karashchuk, G., Goldwurm, S., and Smith, R.J. (2009). GIGYF2 gene disruption in mice results in neurodegeneration and altered insulin-like growth factor signaling. Hum Mol Genet *18*, 4629-4639.

Gutierrez, E., Shin, B.S., Woolstenhulme, C.J., Kim, J.R., Saini, P., Buskirk, A.R., and Dever, T.E. (2013). eIF5A promotes translation of polyproline motifs. Mol Cell *51*, 35-45.

Han, P., Shichino, Y., Schneider-Poetsch, T., Mito, M., Hashimoto, S., Udagawa, T., Kohno, K., Yoshida, M., Mishima, Y., Inada, T.*, et al.* (2020). Genome-wide Survey of Ribosome Collision. Cell Rep *31*, 107610.

Hanson, G., and Coller, J. (2018). Codon optimality, bias and usage in translation and mRNA decay. Nat Rev Mol Cell Biol *19*, 20-30.

Hermesh, O., and Jansen, R.P. (2013). Take the (RN)A-train: localization of mRNA to the endoplasmic reticulum. Biochim Biophys Acta *1833*, 2519-2525.

Hersheson, J., Mencacci, N.E., Davis, M., MacDonald, N., Trabzuni, D., Ryten, M., Pittman, A., Paudel, R., Kara, E., Fawcett, K.*, et al.* (2013). Mutations in the autoregulatory domain of beta-tubulin 4a cause hereditary dystonia. Ann Neurol *73*, 546-553.

Hickey, K.L., Dickson, K., Cogan, J.Z., Replogle, J.M., Schoof, M., D'Orazio, K.N., Sinha, N.K., Frost, A., Green, R., Kostova, K.K.*, et al.* (2019). GIGYF2 and 4EHP inhibit translation initiation of defective messenger RNAs to assist ribosome-associated quality control. BioRxiv.

Hu, W., Sweet, T.J., Chamnongpol, S., Baker, K.E., and Coller, J. (2009). Co-translational mRNA decay in Saccharomyces cerevisiae. Nature *461*, 225-229.

Ikeuchi, K., Tesina, P., Matsuo, Y., Sugiyama, T., Cheng, J., Saeki, Y., Tanaka, K., Becker, T., Beckmann, R., and Inada, T. (2019). Collided ribosomes form a unique structural interface to induce Hel2-driven quality control pathways. EMBO J *38*, e100276.

Ingolia, N.T., Brar, G.A., Rouskin, S., McGeachy, A.M., and Weissman, J.S. (2012). The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. Nat Protoc *7*, 1534-1550.

Ingolia, N.T., Ghaemmaghami, S., Newman, J.R., and Weissman, J.S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science *324*, 218-223.

Ingolia, N.T., Lareau, L.F., and Weissman, J.S. (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. Cell *147*, 789-802.

Iossifov, I., O'Roak, B.J., Sanders, S.J., Ronemus, M., Krumm, N., Levy, D., Stessman, H.A., Witherspoon, K.T., Vives, L., Patterson, K.E.*, et al.* (2014). The contribution of de novo coding mutations to autism spectrum disorder. Nature *515*, 216-221.

Jan, C.H., Williams, C.C., and Weissman, J.S. (2014). Principles of ER cotranslational translocation revealed by proximity-specific ribosome profiling. Science *346*, 1257521.

Joazeiro, C.A.P. (2017). Ribosomal Stalling During Translation: Providing Substrates for Ribosome-Associated Protein Quality Control. Annu Rev Cell Dev Biol *33*, 343-368.

Joazeiro, C.A.P. (2019). Mechanisms and functions of ribosome-associated protein quality control. Nat Rev Mol Cell Biol *20*, 368-383.

Jonas, S., Weichenrieder, O., and Izaurralde, E. (2013). An unusual arrangement of two 14-3-3-like domains in the SMG5-SMG7 heterodimer is required for efficient nonsense-mediated mRNA decay. Genes Dev *27*, 211-225.

Juszkiewicz, S., Chandrasekaran, V., Lin, Z., Kraatz, S., Ramakrishnan, V., and Hegde, R.S. (2018). ZNF598 Is a Quality Control Sensor of Collided Ribosomes. Mol Cell *72*, 469-481 e467.

Juszkiewicz, S., and Hegde, R.S. (2017). Initiation of Quality Control during Poly(A) Translation Requires Site-Specific Ribosome Ubiquitination. Mol Cell *65*, 743-750 e744.

Weber et al.

Karamyshev, A.L., Patrick, A.E., Karamysheva, Z.N., Griesemer, D.S., Hudson, H., Tjon-Kon-Sang, S., Nilsson, I., Otto, H., Liu, Q., Rospert, S.*, et al.* (2014). Inefficient SRP interaction with a nascent chain triggers a mRNA quality control pathway. Cell *156*, 146-157.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol *14*, R36.

Kowarik, M., Kung, S., Martoglio, B., and Helenius, A. (2002). Protein folding during cotranslational translocation in the endoplasmic reticulum. Mol Cell *10*, 769-778.

Krumm, N., Turner, T.N., Baker, C., Vives, L., Mohajeri, K., Witherspoon, K., Raja, A., Coe, B.P., Stessman, H.A., He, Z.X.*, et al.* (2015). Excess of rare, inherited truncating mutations in autism. Nat Genet *47*, 582-588.

Kryszke, M.H., Adjeriou, B., Liang, F., Chen, H., and Dautry, F. (2016). Post-transcriptional gene silencing activity of human GIGYF2. Biochem Biophys Res Commun *475*, 289-294.

Kuzuoglu-Ozturk, D., Bhandari, D., Huntzinger, E., Fauser, M., Helms, S., and Izaurralde, E. (2016). miRISC and the CCR4-NOT complex silence mRNA targets independently of 43S ribosomal scanning. EMBO J *35*, 1186-1203.

Labun, K., Montague, T.G., Gagnon, J.A., Thyme, S.B., and Valen, E. (2016). CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. Nucleic Acids Res *44*, W272-276.

Labun, K., Montague, T.G., Krause, M., Torres Cleuren, Y.N., Tjeldnes, H., and Valen, E. (2019). CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. Nucleic Acids Res *47*, W171-W174.

Lakshminarayan, R., Phillips, B.P., Binnian, I.L., Gomez-Navarro, N., Escudero-Urquijo, N., Warren, A.J., and Miller, E.A. (2020). Pre-emptive Quality Control of a Misfolded Membrane Protein by Ribosome-Driven Effects. Curr Biol.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods *9*, 357-359.

Lazzaretti, D., Tournier, I., and Izaurralde, E. (2009). The C-terminal domains of human TNRC6A, TNRC6B, and TNRC6C silence bound transcripts independently of Argonaute proteins. RNA *15*, 1059-1066.

Lee, S., Liu, B., Lee, S., Huang, S.X., Shen, B., and Qian, S.B. (2012). Global mapping of translation initiation sites in mammalian cells at single-nucleotide resolution. Proc Natl Acad Sci U S A *109*, E2424-2432.

Lin, Z., Gasic, I., Chandrasekaran, V., Peters, N., Shao, S., Mitchison, T.J., and Hegde, R.S. (2020). TTC5 mediates autoregulation of tubulin via mRNA degradation. Science *367*, 100-104.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods *25*, 402-408.

Manjunath, H., Zhang, H., Rehfeld, F., Han, J., Chang, T.C., and Mendell, J.T. (2019). Suppression of Ribosomal Pausing by eIF5A Is Necessary to Maintain the Fidelity of Start Codon Selection. Cell Rep *29*, 3134-3146 e3136.

McCarthy, D.J., Chen, Y., and Smyth, G.K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res *40*, 4288-4297.

Montague, T.G., Cruz, J.M., Gagnon, J.A., Church, G.M., and Valen, E. (2014). CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. Nucleic Acids Res *42*, W401-407.

Morita, M., Ler, L.W., Fabian, M.R., Siddiqui, N., Mullin, M., Henderson, V.C., Alain, T., Fonseca, B.D., Karashchuk, G., Bennett, C.F.*, et al.* (2012). A novel 4EHP-GIGYF2 translational repressor complex is essential for mammalian development. Mol Cell Biol *32*, 3585-3593.

Pachter, J.S., Yen, T.J., and Cleveland, D.W. (1987). Autoregulation of tubulin expression is achieved through specific degradation of polysomal tubulin mRNAs. Cell *51*, 283-292.

Paillusson, A., Hirschi, N., Vallan, C., Azzalin, C.M., and Muhlemann, O. (2005). A GFP-based reporter system to monitor nonsense-mediated mRNA decay. Nucleic Acids Res *33*, e54.

Pavlov, M.Y., Watts, R.E., Tan, Z., Cornish, V.W., Ehrenberg, M., and Forster, A.C. (2009). Slow peptide bond formation by proline and other N-alkylamino acids in translation. Proc Natl Acad Sci U S A *106*, 50-54.

Pelechano, V., and Alepuz, P. (2017). eIF5A facilitates translation termination globally and promotes the elongation of many non polyproline-specific tripeptide sequences. Nucleic Acids Res *45*, 7326-7338.

Pelechano, V., Wei, W., and Steinmetz, L.M. (2015). Widespread Co-translational RNA Decay Reveals Ribosome Dynamics. Cell *161*, 1400-1412.

Peter, D., Igreja, C., Weber, R., Wohlbold, L., Weiler, C., Ebertsch, L., Weichenrieder, O., and Izaurralde, E. (2015). Molecular architecture of 4E-BP translational inhibitors bound to eIF4E. Mol Cell *57*, 1074-1087.

Peter, D., Ruscica, V., Bawankar, P., Weber, R., Helms, S., Valkov, E., Igreja, C., and Izaurralde, E. (2019). Molecular basis for GIGYF-Me31B complex assembly in 4EHP-mediated translational repression. Genes Dev *33*, 1355-1360.

Peter, D., Weber, R., Sandmeir, F., Wohlbold, L., Helms, S., Bawankar, P., Valkov, E., Igreja, C., and Izaurralde, E. (2017). GIGYF1/2 proteins use auxiliary sequences to selectively bind to 4EHP and repress target mRNA expression. Genes Dev *31*, 1147-1161.

Pillai, R.S., Artus, C.G., and Filipowicz, W. (2004). Tethering of human Ago proteins to mRNA mimics the miRNA-mediated repression of protein synthesis. RNA *10*, 1518-1525.

Pinarbasi, E.S., Karamyshev, A.L., Tikhonova, E.B., Wu, I.H., Hudson, H., and Thomas, P.J. (2018). Pathogenic Signal Sequence Mutations in Progranulin Disrupt SRP Interactions Required for mRNA Stability. Cell Rep *23*, 2844-2851.

Presnyak, V., Alhusaini, N., Chen, Y.H., Martin, S., Morris, N., Kline, N., Olson, S., Weinberg, D., Baker, K.E., Graveley, B.R.*, et al.* (2015). Codon optimality is a major determinant of mRNA stability. Cell *160*, 1111-1124.

Radhakrishnan, A., Chen, Y.H., Martin, S., Alhusaini, N., Green, R., and Coller, J. (2016). The DEAD-Box Protein Dhh1p Couples mRNA Decay and Translation by Monitoring Codon Optimality. Cell *167*, 122-132 e129.

Radhakrishnan, A., and Green, R. (2016). Connections Underlying Translation and mRNA Stability. Journal of molecular biology *428*, 3558-3564.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat Protoc *8*, 2281-2308.

Rasch, F., Weber, R., Izaurralde, E., and Igreja, C. (2020). 4E-T-bound mRNAs are stored in a silenced and deadenylated form. Genes Dev.

Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat Biotechnol *29*, 24-26.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139-140.

Rom, E., Kim, H.C., Gingras, A.C., Marcotrigiano, J., Favre, D., Olsen, H., Burley, S.K., and Sonenberg, N. (1998). Cloning and characterization of 4EHP, a novel mammalian eIF4E-related cap-binding protein. J Biol Chem *273*, 13104-13109.

Ruscica, V., Bawankar, P., Peter, D., Helms, S., Igreja, C., and Izaurralde, E. (2019). Direct role for the Drosophila GIGYF protein in 4EHP-mediated mRNA repression. Nucleic Acids Res *47*, 7035-7048.

Satterstrom, F.K., Kosmicki, J.A., Wang, J., Breen, M.S., De Rubeis, S., An, J.Y., Peng, M., Collins, R., Grove, J., Klei, L.*, et al.* (2020). Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. Cell.

Schizophrenia Working Group of the Psychiatric Genomics, C. (2014). Biological insights from 108 schizophrenia-associated genetic loci. Nature *511*, 421-427.

Schmidt, C., Becker, T., Heuer, A., Braunger, K., Shanmuganathan, V., Pech, M., Berninghausen, O., Wilson, D.N., and Beckmann, R. (2016). Structure of the hypusinylated eukaryotic translation factor eIF-5A bound to the ribosome. Nucleic Acids Res *44*, 1944-1951.

Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat Methods *9*, 671-675.

Schopp, I.M., Amaya Ramirez, C.C., Debeljak, J., Kreibich, E., Skribbe, M., Wild, K., and Bethune, J. (2017). Split-BioID a conditional proteomics approach to monitor the composition of spatiotemporally defined protein complexes. Nat Commun *8*, 15690.

Schuller, A.P., Wu, C.C., Dever, T.E., Buskirk, A.R., and Green, R. (2017). eIF5A Functions Globally in Translation Elongation and Termination. Mol Cell *66*, 194-205 e195.

Simms, C.L., Thomas, E.N., and Zaher, H.S. (2017a). Ribosome-based quality control of mRNA and nascent peptides. Wiley Interdiscip Rev RNA *8*.

Simms, C.L., Yan, L.L., and Zaher, H.S. (2017b). Ribosome Collision Is Critical for Quality Control during No-Go Decay. Mol Cell *68*, 361-373 e365.

Sundaramoorthy, E., Leonard, M., Mak, R., Liao, J., Fulzele, A., and Bennett, E.J. (2017). ZNF598 and RACK1 Regulate Mammalian Ribosome-Associated Quality Control Function by Mediating Regulatory 40S Ribosomal Ubiquitylation. Mol Cell *65*, 751-760 e754.

Sweet, T., Kovalak, C., and Coller, J. (2012). The DEAD-box protein Dhh1 promotes decapping by slowing ribosome movement. PLoS Biol *10*, e1001342.

Tesina, P., Heckel, E., Cheng, J., Fromont-Racine, M., Buschauer, R., Kater, L., Beatrix, B., Berninghausen, O., Jacquier, A., Becker, T*., et al.* (2019). Structure of the 80S ribosome-Xrn1 nuclease complex. Nat Struct Mol Biol *26*, 275-280.

Weber et al.

Thorvaldsdottir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform *14*, 178-192.

Thyme, S.B., Pieper, L.M., Li, E.H., Pandey, S., Wang, Y., Morris, N.S., Sha, C., Choi, J.W., Herrera, K.J., Soucy, E.R*., et al.* (2019). Phenotypic Landscape of Schizophrenia-Associated Genes Defines Candidates and Their Shared Functions. Cell *177*, 478-491 e420.

Tollenaere, M.A.X., Tiedje, C., Rasmussen, S., Nielsen, J.C., Vind, A.C., Blasius, M., Batth, T.S., Mailand, N., Olsen, J.V., Gaestel, M*., et al.* (2019). GIGYF1/2-Driven Cooperation between ZNF598 and TTP in Posttranscriptional Regulation of Inflammatory Signaling. Cell Rep *26*, 3511-3521 e3514.

Tuck, A.C., Rankova, A., Arpat, A.B., Liechti, L.A., Hess, D., Iesmantavicius, V., Castelo-Szekely, V., Gatfield, D., and Buhler, M. (2020). Mammalian RNA Decay Pathways Are Highly Specialized and Widely Linked to Translation. Mol Cell.

Voorhees, R.M., and Hegde, R.S. (2015). Structures of the scanning and engaged states of the mammalian SRP-ribosome complex. Elife *4*.

Walter, P., and Blobel, G. (1981). Translocation of proteins across the endoplasmic reticulum III. Signal recognition protein (SRP) causes signal sequence-dependent and site-specific arrest of chain elongation that is released by microsomal membranes. The Journal of cell biology *91*, 557-561.

Wohlgemuth, I., Brenner, S., Beringer, M., and Rodnina, M.V. (2008). Modulation of the rate of peptidyl transfer on the ribosome by the nature of substrates. J Biol Chem *283*, 32229-32235.

Wolin, S.L., and Walter, P. (1988). Ribosome pausing and stacking during translation of a eukaryotic mRNA. EMBO J *7*, 3559-3569.

Wolin, S.L., and Walter, P. (1989). Signal recognition particle mediates a transient elongation arrest of preprolactin in reticulocyte lysate. The Journal of cell biology *109*, 2617-2622.

Yen, T.J., Gay, D.A., Pachter, J.S., and Cleveland, D.W. (1988). Autoregulated changes in stability of polyribosome-bound beta-tubulin mRNAs are specified by the first 13 translated nucleotides. Mol Cell Biol *8*, 1224-1235.

Young, M.D., Wakefield, M.J., Smyth, G.K., and Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biol *11*, R14.

Zhong, Y., Karaletsos, T., Drewe, P., Sreedharan, V.T., Kuo, D., Singh, K., Wendel, H.G., and Ratsch, G. (2017). RiboDiff: detecting changes of mRNA translation efficiency from ribosome footprints. Bioinformatics *33*, 139-141.

Zuberek, J., Kubacka, D., Jablonowska, A., Jemielity, J., Stepinski, J., Sonenberg, N., and Darzynkiewicz, E. (2007). Weak binding affinity of human 4EHP for mRNA cap analogs. RNA *13*, 691-697.

Weber *et al.* Figure 1

A GIGYF1/2 KO cells

B 4EHP KO cells

C Upregulated mRNAs

D GIGYF1/2 KO cells

E 4EHP KO cells

F 4EHP–GIGYF1/2 upregulated mRNAs
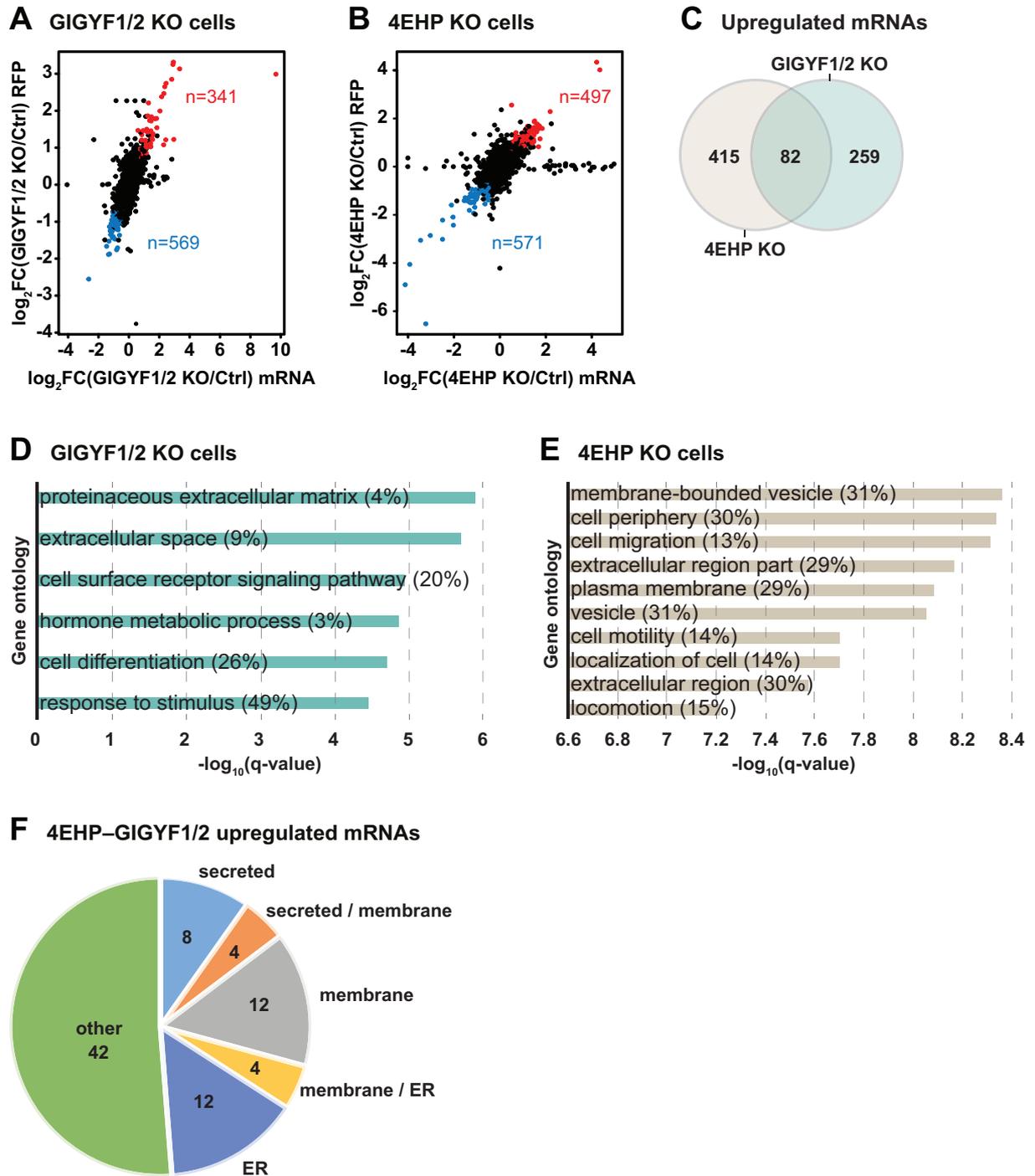
Weber *et al.* Figure 2

A *DBNDD2* mRNA

B *DBNDD2* mRNA

$t_{1/2}$Ctrl = 214 ± 51 min (R^2 = 0.97)
$t_{1/2}$KO = >480 min (R^2 = 0.79)

C *CD109* mRNA

$t_{1/2}$Ctrl = 180 ± 4 min (R^2 = 0.82)
$t_{1/2}$KO = 340 ± 58 min (R^2 = 0.79)

D *ITPR3* mRNA

$t_{1/2}$Ctrl = 219 ± 17 min (R^2 = 0.93)
$t_{1/2}$KO = 422 ± 88 min (R^2 = 0.96)

E *NPTX1* mRNA

$t_{1/2}$Ctrl = 217 ± 15 min (R^2 = 0.78)
$t_{1/2}$KO = 482 ± 121 min (R^2 = 0.96)

F β-*ACTIN* mRNA

$t_{1/2}$Ctrl = 189 ± 79 min (R^2 = 0.87)
$t_{1/2}$KO = 226 ± 104 min (R^2 = 0.92)

A CDS reporters

m⁷G● DBNDD2 ORF [HA] ——AAAA

m⁷G● DBNDD2 ORF [HA] ——AAAA
STOP₃

m⁷G● [SP] LGALS3BP ORF [HA] ——AAAA

m⁷G● [SP] LGALS3BP ORF [HA] ——AAAA
STOP₃

B DBNDD2-HA protein

reporter: WT STOP₃
GIGYF1/2 cells: Ctrl KO Ctrl KO   kDa
DBNDD2-HA   — 35
                                   — 25
F-LUC-GFP   — 100
           1   2   3   4

C LGALS3BP-HA protein

reporter: WT STOP₃
GIGYF1/2 cells: Ctrl KO Ctrl KO   kDa
                                   — 100
LGALS3BP-HA   — 70
F-LUC-GFP   — 100
           1   2   3   4

D HA protein

■ GIGYF1/2 KO
■ Ctrl cells
HA-protein / F-LUC-GFP (%)

E *DBNDD2-HA* mRNA

cells: Ctrtl   GIGYF1/2 KO
DBNDD2-HA mRNA
TUBB mRNA
18S rRNA
0 120 240 360 480 600 0 120 240 360 480 600
Time after ActD treatment (min)

F *LGALS3BP-HA* mRNA

cells: Ctrl   GIGYF1/2 KO
LGALS3BP-HA mRNA
TUBB mRNA
18S rRNA
0 120 240 360 480 600 0 120 240 360 480 600
Time after ActD treatment (min)

G *DBNDD2-STOP₃-HA* mRNA

cells: Ctrl   GIGYF1/2 KO
DBNDD2-STOP₃-HA mRNA
TUBB mRNA
18S rRNA
0 120 240 360 480 600 0 120 240 360 480 600
Time after ActD treatment (min)

H *LGALS3BP-STOP₃-HA* reporter

cells: Ctrl   GIGYF1/2 KO
LGALS3BP-STOP₃-HA mRNA
TUBB mRNA
18S rRNA
0 120 240 360 480 600 0 120 240 360 480 600
Time after ActD treatment (min)

I *DBNDD2-HA* mRNA

Reporter / *TUBB* mRNA (%)
Time after ActD treatment (min)
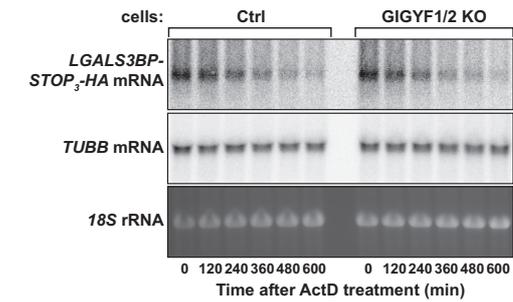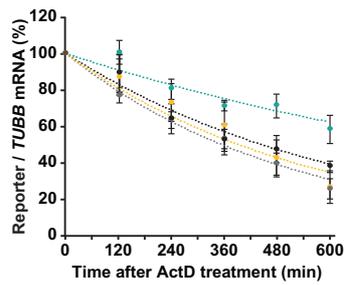
■ Ctrl cells + *DBNDD2-HA* $t_{1/2}$ = 451 ± 67 min (R^2 = 0.97)
■ Ctrl cells + *DBNDD2-STOP₃-HA* $t_{1/2}$ = 358 ± 48 min (R^2 = 0.99)
■ GIGYF1/2 KO + *DBNDD2-HA* $t_{1/2}$ > 600 min (R^2 = 0.91)
■ GIGYF1/2 KO + *DBNDD2-STOP₃-HA* $t_{1/2}$ = 404 ± 105 min (R^2 = 0.95)

J *LGALS3BP-HA* mRNA

Reporter / *TUBB* mRNA (%)
Time after ActD treatment (min)
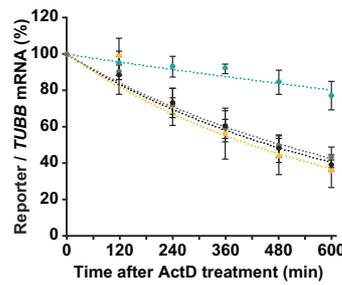
■ Ctrl cells + *LGALS3BP-HA* $t_{1/2}$ = 460 ± 27 min (R^2 = 0.99)
■ Ctrl cells + *LGALS3BP-STOP₃-HA* $t_{1/2}$ = 491 ± 52 min (R^2 = 0.99)
■ GIGYF1/2 KO + *LGALS3BP-HA* $t_{1/2}$ > 600 min (R^2 = 0.90)
■ GIGYF1/2 KO + *LGALS3BP-STOP₃-HA* $t_{1/2}$ = 418 ± 36 min (R^2 = 0.96)
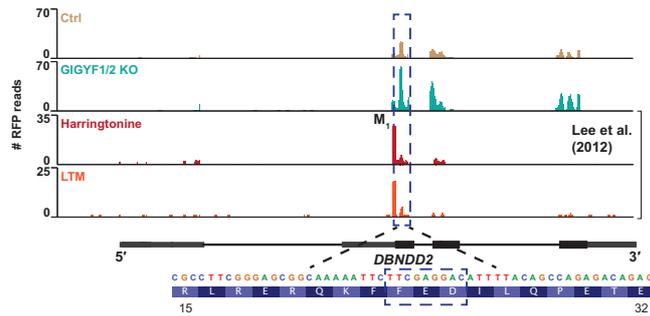
A RFP reads *DBNDD2*

B *DBNDD2-HA* reporters
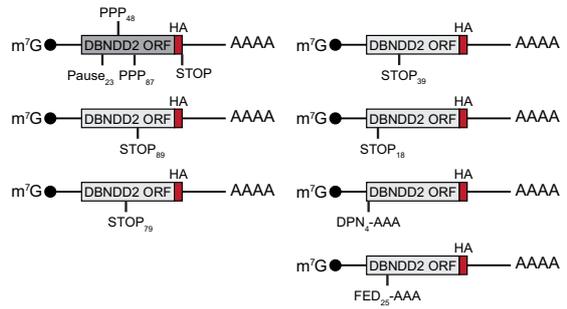
C *DBNDD2-HA*
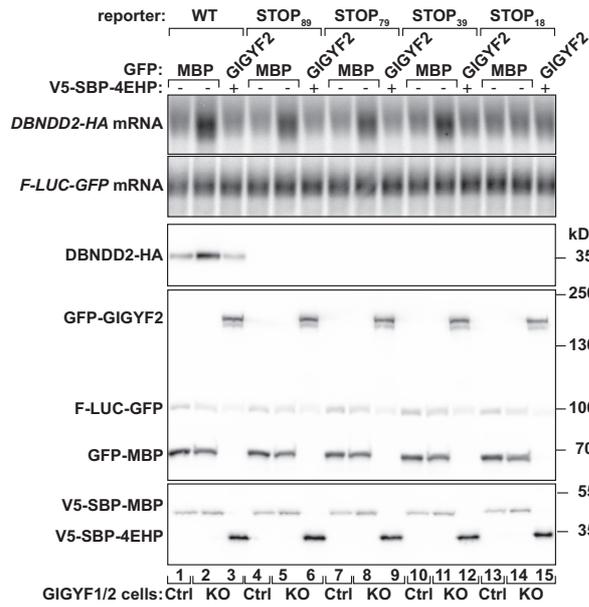
D *DBNDD2-HA* mRNA
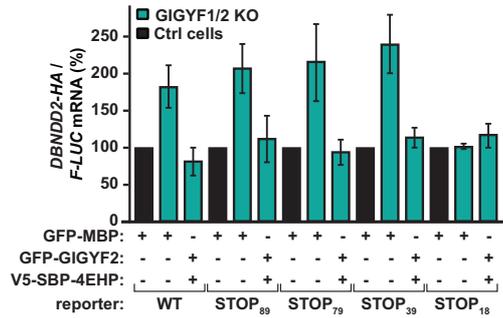
E *DBNDD2-HA* mRNA

F *DBNDD2-HA* mRNA

$t_{1/2}$ WT = 429 ± 39 min (R^2 = 0.98)
$t_{1/2}$ DPN$_4$-AAA> 600 min (R^2 = 0.31)
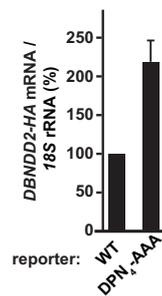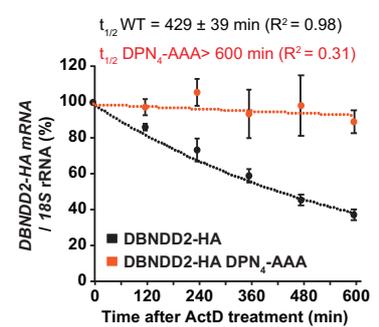
**A** RFP reads *LGALS3BP*
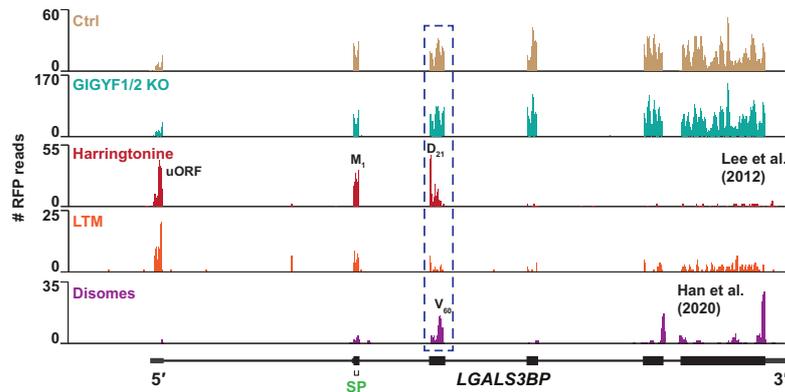
B *LGALS3BP* and *R-LUC* reporters
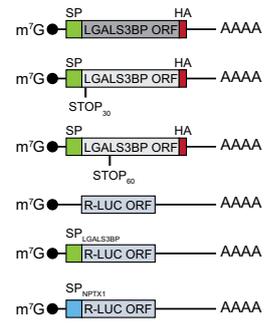
C *LGALS3BP-HA* mRNA
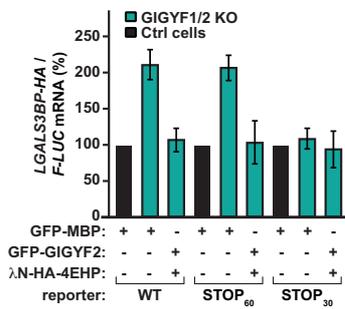
D *LGALS3BP-HA* mRNA
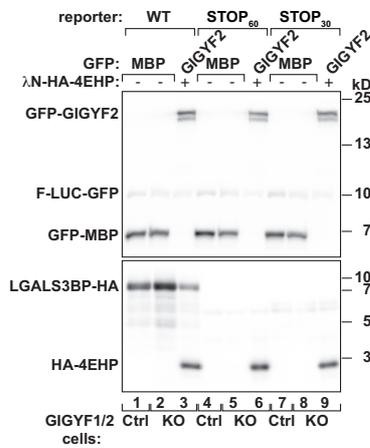
E LGALS3BP-HA
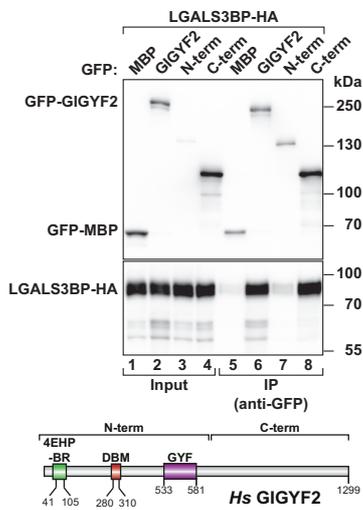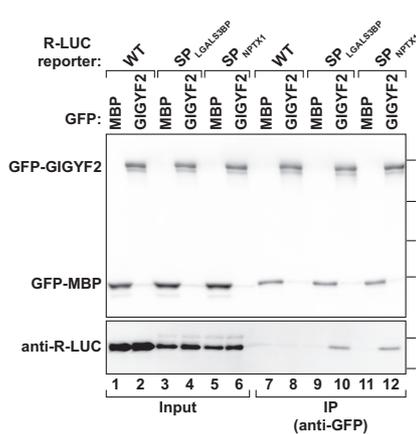
F
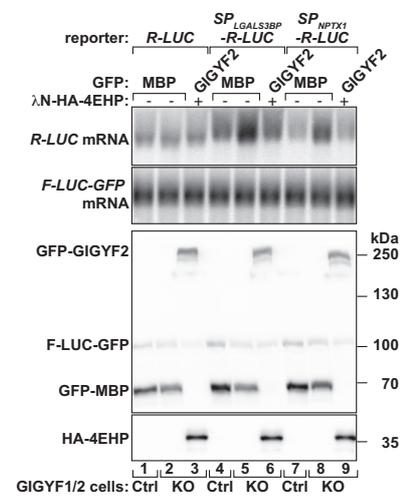
G *R-LUC* mRNA

H *R-LUC* mRNA
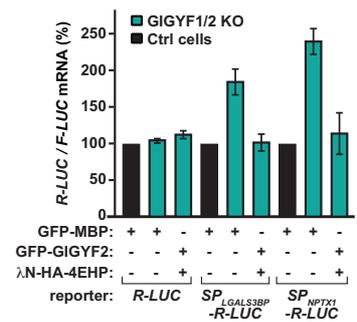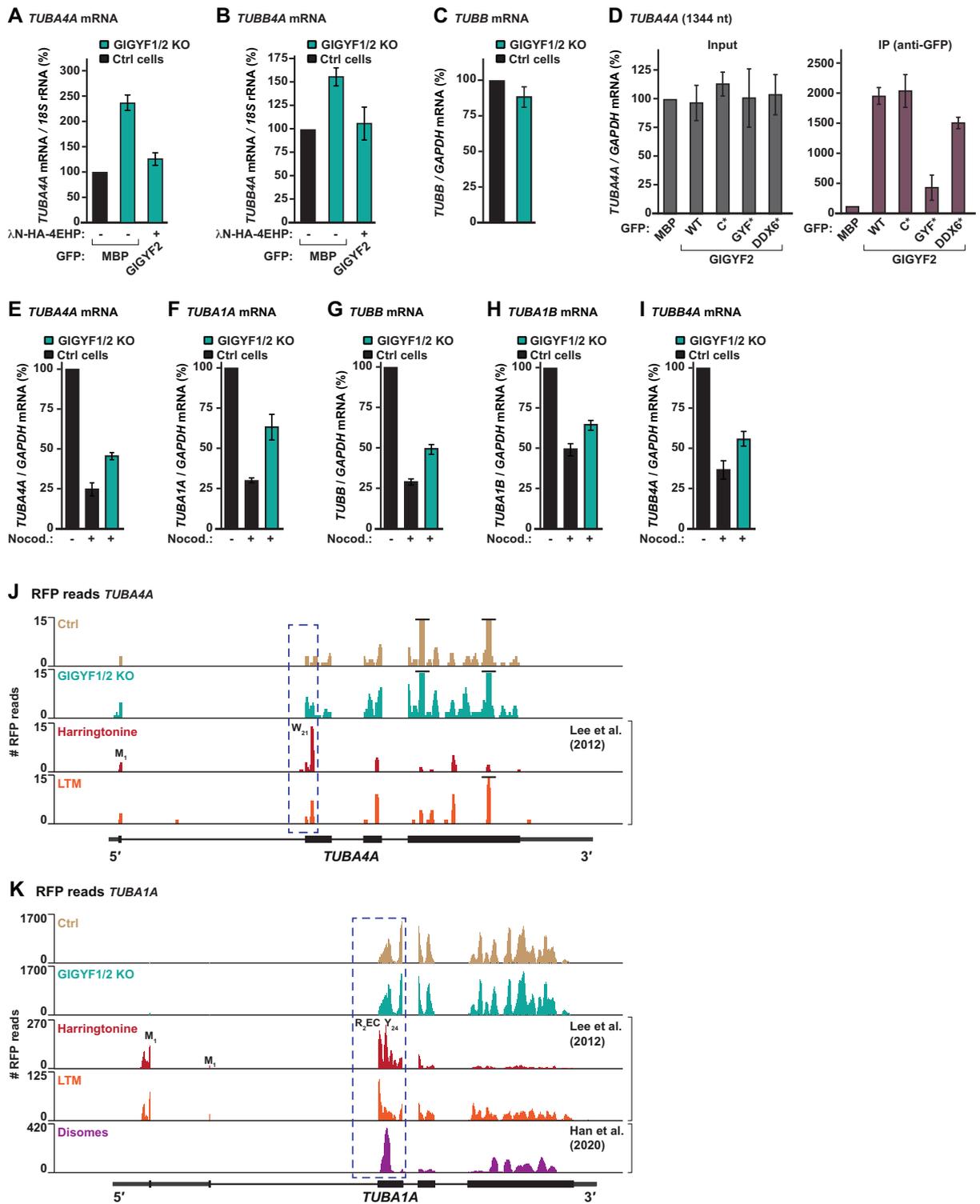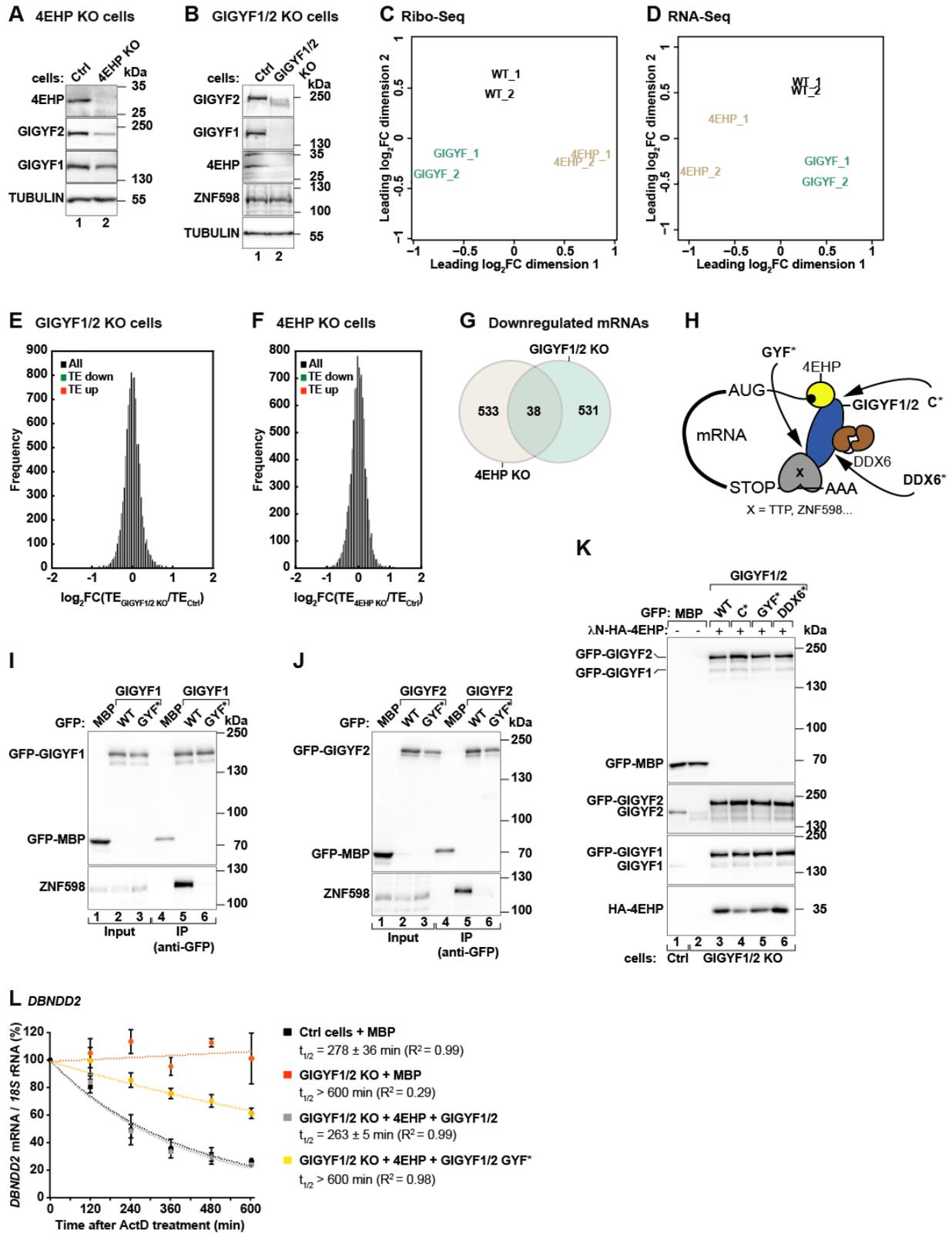
Figure S1, related to Figures 1 and 3. Characterization of 4EHP-null (KO) and GIGYF1/2-null cells

(A, B) Western blots demonstrating loss of endogenous 4EHP in 4EHP KO cells (A) and endogenous GIGYF1 and GIGYF2 in GIGYF1/2 KO cells (B). GIGYF1 and GIGYF2 expression is reduced in 4EHP KO cells and 4EHP is weakly expressed in GIGYF1/2 KO cells. TUBULIN served as loading control. Note that the TUBULIN antibody recognizes an epitope common among the α-tubulin subunits of which TUBA4A and TUBA1A are genes with increased mRNA abundance in 4EHP-null cells.

(C, D) Multidimensional scaling (MDS) analysis for the Ribo-Seq replicate libraries (C) and the RNA-Seq replicate libraries (D) from HEK293T WT, GIGYF1/2-null and 4EHP-null cells.

**(E, F)** Histograms of the number of transcripts (frequency) relative to the $\log_2$FC of translational efficiency (TE) in GIGYF1/2-null (E) and 4EHP-null cells (F). Significantly upregulated (FDR<0.005 and $\log_2$FC>0) transcripts are depicted in orange, significantly downregulated (FDR<0.005 and $\log_2$FC<0) transcripts are shown in green. There were only few genes with changes in TE (Table S1).

(G) Venn diagram of the genes with decreased mRNA abundance (downregulated genes) in GIGYF1/2-null and 4EHP-null cells identifies a non-significant overlap of 38 transcripts.

(H) Schematic representation of the effector complex involved in GIGYF1/2-mediated mRNA decay. GIGYF1/2 canonical (C*) mutant is unable to interact with 4EHP; GIGYF1/2 DDX6* protein does not associate with the RNA helicase DDX6; the GIGYF1/2 GYF* mutant cannot interact with PPGΦ-rich proteins. TTP: tristetraprolin; X: RNA-binding protein.

(I, J) The interaction of GIGYF1 (I) and GIGYF2 (J) WT or GYF domain mutant (GYF*) with ZNF598 was analyzed in co-immunoprecipitation assays using anti-GFP antibodies. GFP-MBP served as a negative control. The input (0.8% for the GFP proteins and 0.3% for ZNF598) and immunoprecipitated fractions (12% for the GFP proteins and 24% for ZNF598) were analyzed by western blotting with anti-GFP and anti-ZNF598 antibodies.

(K) Immunoblot showing the expression of the proteins used in the experiment depicted in Figures 3A, B. Blots were probed with antibodies recognizing GFP, GIGYF1, GIGYF2 or HA. Inputs and immunoprecipitates were 2% and 2.7%, respectively.

(L) Ctrl and GIGYF1/2 KO cells were transfected with plasmids expressing λN-HA or λN-HA-4EHP, GFP-MBP, and GFP-GIGYF1/2 (WT or GYF*). Two days post transfection, cells were treated with ActD and harvested at the indicated time points. *DBNDD2* mRNA levels were quantified by RT-qPCR and normalized to those of *18S* rRNA. Circles represent the mean value; error bars represent SD (n=3). The decay curves were fitted to an exponential decay with a single component (dotted lines). R^2 values are indicated for each curve.
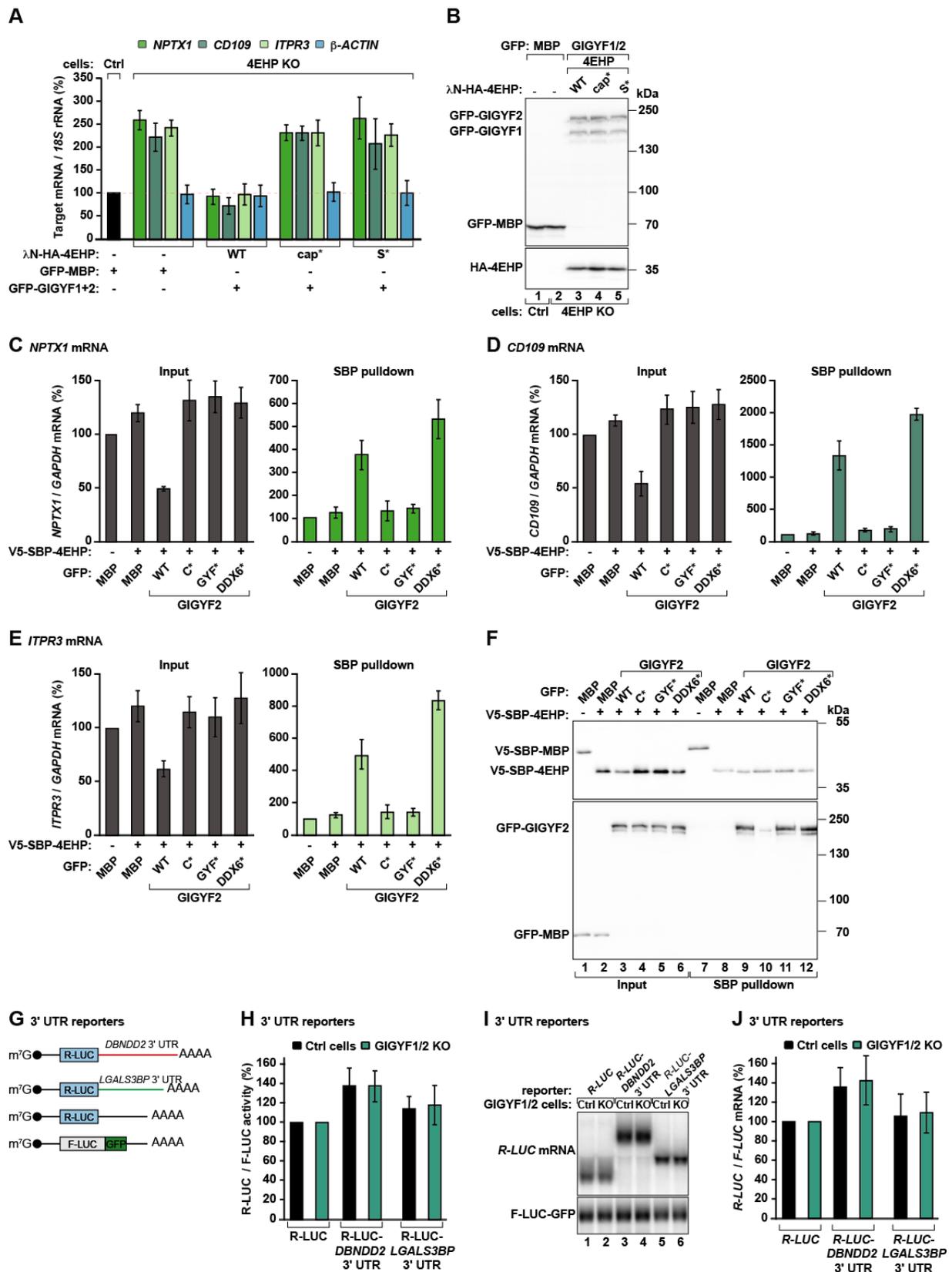
Figure S2, related to Figure 3. Binding of 4EHP to the cap and to GIGYF1/2 is crucial for mRNA decay.

(A) Control and 4EHP-null cells were transfected with plasmids expressing λN-HA alone, wild type (WT) or the indicated λN-HA-4EHP mutants (cap-binding mutant: cap*, GIGYF1/2 specific-binding mutant: S*), GFP-MBP or GFP-GIGYF1 and GFP-GIGYF2. *NPTX1*, *CD109*, *ITPR3* and *β-ACTIN* mRNA levels were determined by RT-qPCR and normalized to that of *18S* rRNA in the presence of the different 4EHP proteins. Bars represent the mean values and error bars denote the SD of three independent experiments.

(B) Western blot showing the expression levels of the proteins used in the experiments shown in (A). Blots were probed with anti-GFP and anti-HA antibodies.

(C-F) HEK293T cells were transfected with plasmids expressing V5-SBP-MBP or V5-SBP-4EHP, GFP-MBP or GFP-GIGYF2 (WT or mutants). Streptavidin binding protein-based pulldowns were performed two days post transfection and protein and RNA samples were collected for each experimental condition. *NPTX1*, *CD109* and *ITPR3* mRNA levels in input (0.8%) and IP samples (12%) were quantified by RT-qPCR, normalized over *GAPDH* and set to 100% in the presence of V5-SBP-MBP. Bars represent the mean value; error bars represent standard deviations from three independent experiments. (F) Western blot showing the expression of the proteins in the inputs (1% for the V5-SBP-tagged proteins and 0.5% for GFP-tagged proteins) and bound fractions (0.9% for the V5-SBP-tagged proteins and 2.7% for GFP-tagged proteins) from the experiments described in C-E.

(G) Schematic representation of the *DBNDD2* and *LGALS3BP* 3′ UTR reporters. *Renilla* luciferase (R-LUC); firefly luciferase (F-LUC); green fluorescent protein (GFP).

(H-J) Ctrl and GIGYF1/2 KO cells were transfected with the R-LUC-*DBNDD2*-3′ UTR, the R-LUC-*LGALS3BP*-3′ UTR or the R-LUC reporters and the transfection control F-LUC-GFP. (H) R-LUC activity was normalized to that of F-LUC-GFP and set to 100% for R-LUC in each cell line. (I) mRNA levels were determined by northern blotting. (J) *R-LUC, R-LUC-DBNDD2-*

3′ UTR and *R-LUC-LGALS3BP-3′* UTR band intensities were normalized to the intensity of

F-LUC-GFP mRNA band and set to 100% for the R-LUC reporter in each cell line.
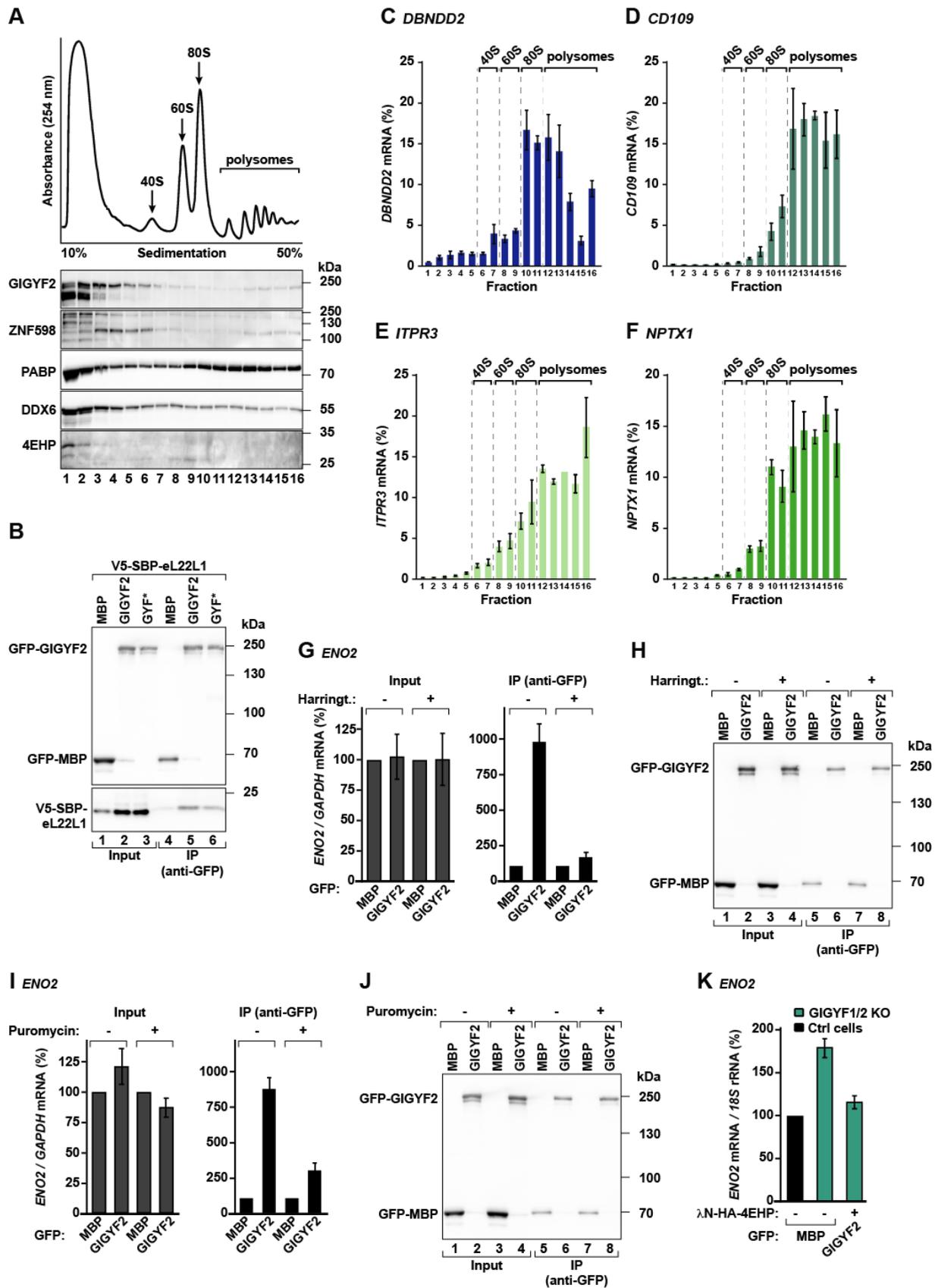
Figure S3, related to Figure 4. GIGYF1/2 promote decay of actively translating mRNAs

(A) UV absorbance profile at 254 nm of HEK293T cell extracts after polysome sedimentation in a sucrose gradient. 40S and 60S subunits, 80S monosomes, and polysome peaks are indicated. The distribution of GIGYF2, ZNF598, PABP, DDX6 and 4EHP across the gradient was analyzed by western blotting and is depicted below the profile.

(B) Immunoblot showing the interaction of GFP-GIGYF2 with V5-SBP-eL22L1 ribosomal protein. Proteins were immunoprecipitated using anti-GFP antibodies. GFP-MBP served as a negative control. The input (0.8% for the GFP proteins and 0.1% for V5-SBP-eL22L1) and immunoprecipitated fractions (12% for the GFP proteins and 24% for V5-SBP-eL22L1) were analyzed by western blotting with anti-GFP and anti-V5 antibodies.

(C-F) Abundance profiles for *DBNDD2*, *CD109*, *ITPR3* and *NPTX1* along the density gradient. mRNA levels were determined by RT-qPCR in samples prepared from total RNA extracted from each sucrose fraction. Bars represent the mean value; error bars denote the standard deviations from three independent experiments.

(G-J) HEK293T transfected with GFP-MBP or GFP-GIGYF2 were incubated with DMSO and the translational inhibitors harringtonine (G, H) or puromycin (I, J). After cell lysis, proteins were immunoprecipitated using anti-GFP antibodies. Protein and RNA samples were obtained for each experimental condition. Input (2%) and immunoprecipitated fractions (2.7%) were analyzed by western blotting. RNA samples were reversed transcribed and *ENO2* expression levels in input (0.8%) and IP samples (12%) were quantified by RT-qPCR, normalized to *GAPDH* and set to 100% in the presence of GFP-MBP. Bars represent the mean value and error bars the standard deviations from three independent experiments.

(K) Control and GIGYF1/2-null cells were transfected with plasmids expressing GFP-MBP or GFP-GIGYF2 and λN-HA-4EHP. RNA samples were collected and *ENO2* mRNA levels were quantified by RT-qPCR. mRNA levels were normalized to that of *18S* rRNA and set to 100% in Ctrl cells.
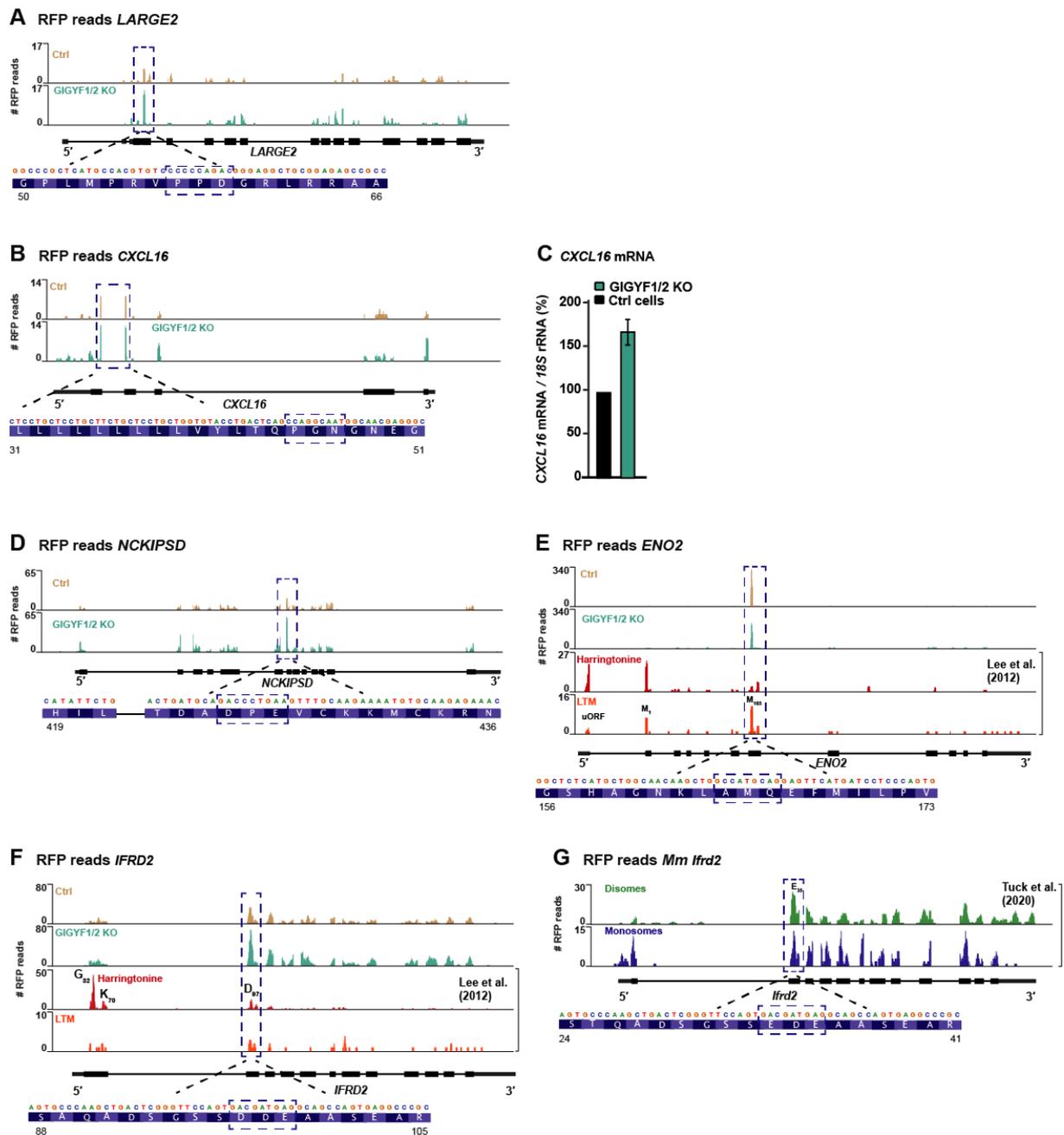
Figure S4, related to Figure 5. Ribosome density profiles in GIGYF1/2–4EHP target mRNAs reveal translational pausing

(A, B) Ribosome density profiles on *LARGE2* and *CXCL16*. The dashed blue box indicates the ribosome pause site. The nucleotide, peptide sequence at the pause site and residue numbering are depicted below the profiles.

(C) *CXCL16* mRNA steady state levels were quantified by RT-qPCR in control (Ctrl) and GIGYF1/2-null (KO) cells. mRNA levels were normalized to that of *18S* rRNA and set to 100% in Ctrl cells.

(D, E) Distribution of RFPs across the CDS of *NCKIPSD* and *ENO2* in Ctrl and GIGYF1/2 KO cells. Translational stalls with increased RFPs are highlighted with a dashed blue box. For *ENO2*, RFP distribution in cells treated with the translational inhibitors harringtonine and lactimidomycin (LTM) obtained by Lee and co-workers (Lee et al., 2012) are also shown. Transcript organization, nucleotide and peptide sequence at the pause site, and residue numbering are depicted below the profiles. Upstream open reading frame (uORF), Met_1 (M_1), Met_{165} (M_{165}).

(F) RFPs distribution along the CDS of *IFRD2* in Ctrl, GIGYF1/2 KO, harringtonine- and LTM-treated cells (Lee et al., 2012). In cells treated with harringtonine, RFPs are observed at different positions of *IFRD2*; one of these corresponds to the paused ribosome at the DDE motif observed in this study. Transcript organization, nucleotide and peptide sequence at the pause site, and residue numbering are depicted below the profiles. Gly_{32} (G_{32}), Lys_{70} (K_{70}), Asp_{97} (D_{97}).

(G) Monosome and disome footprint distribution in *Mus musculus* (*Mm*) *Ifrd2*, as determined by Tuck and co-workers (Tuck et al., 2020). Of note is the occurrence of disomes at an equivalent position of the pause peptide observed in the human orthologue transcript. Transcript organization, nucleotide and peptide sequence at the pause site, and residue numbering are depicted below the profiles. Glu_{35} (E_{35}).
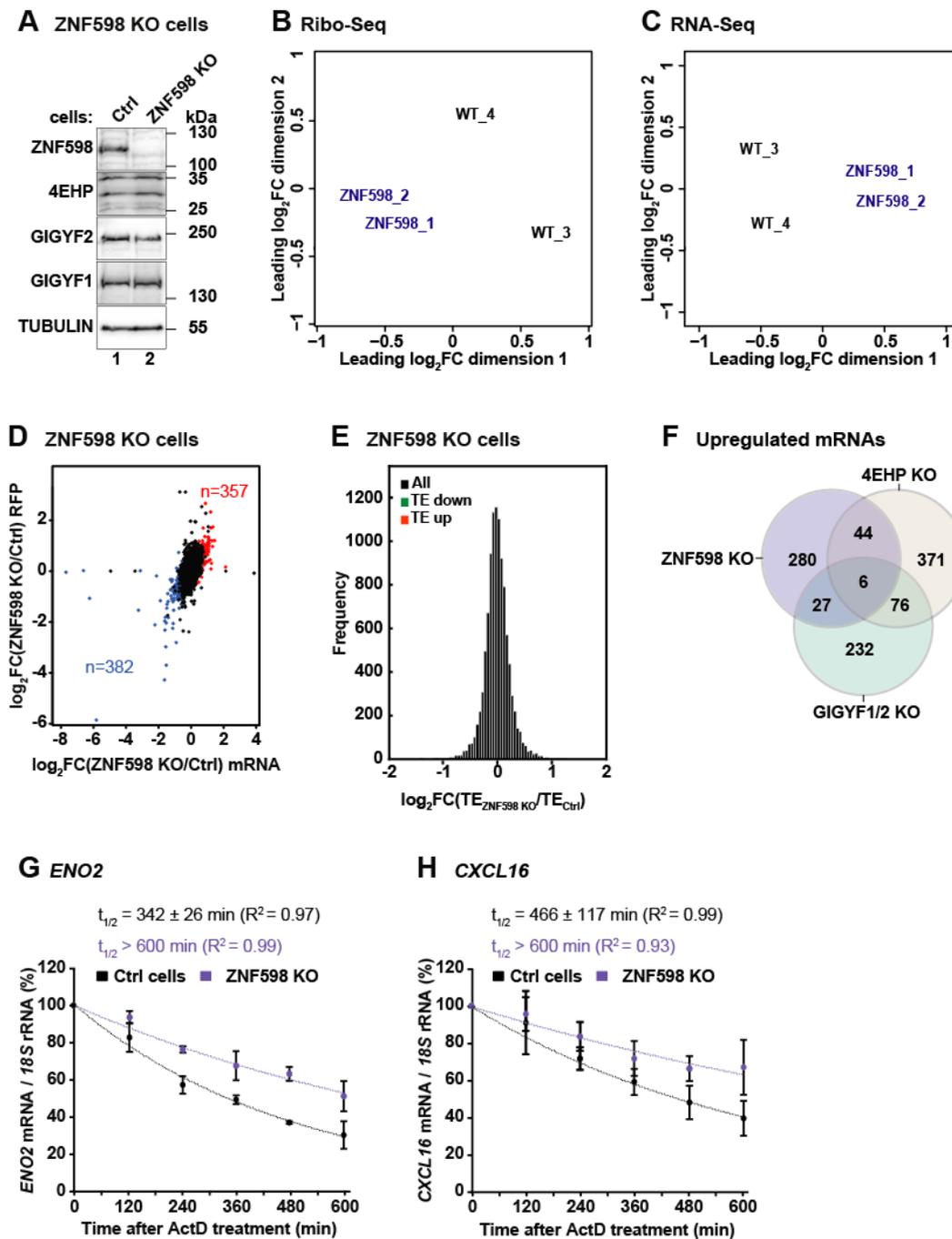
Figure S5. The ubiquitin ligase ZNF598 is only required for co-translational decay of a fraction of GIGYF1/2 targets

(A) Immunoblot showing the lack of ZNF598 expression in the null cells. 4EHP and GIGYF1/2 expression does not vary in ZNF598-null cells. TUBULIN was used as a loading control.

(B, C) Multidimensional scaling (MDS) analysis for the Ribo-Seq replicate libraries (B) and the RNA-Seq replicate libraries (C) from HEK293T WT and ZNF598-null cells.

(D) Genome-wide assessment of changes in RFPs and mRNA abundance in ZNF598 KO cells, depicted on a $\log_2$ scale. Each dot represents an individual gene (n_{total}=10453) with FPKM>2. In ZNF598 KO cells, 357 genes were significantly upregulated (FDR<0.005 and $\log_2$FC>0; red), whereas 382 genes were significantly downregulated (FDR<0.005 and $\log_2$FC<0; blue).

(E) Histogram showing the number of transcripts (frequency) in ZNF598-null with changes in TE ($\log_2$FC) relative to Ctrl cells. Transcripts with increased TE (n=5) in the absence of ZNF598 are shown in orange (FDR<0.005 and $\log_2$FC>0), whereas less translated transcripts (n=2) are depicted in green (FDR<0.005 and $\log_2$FC<0). See also Table S1.

(F) Venn diagram of the genes with increased mRNA abundance (upregulated genes) in ZNF598-null, GIGYF1/2-null and 4EHP-null cells.

(G, H) Ctrl and ZNF598 KO HEK293T cells were treated with Actinomycin D (ActD) and harvested at the indicated time points. *ENO2* and *CXCL16* transcript levels were assessed by RT-qPCR and normalized to that of *18S* rRNA. The normalized value at time zero (before ActD addition) was defined as 100%. Results were plotted as a function of time post ActD addition. Circles represent the mean value; error bars represent the SD from three independent experiments. The decay curves were fitted to an exponential decay with a single component (dotted lines). R^2 values are indicated for each curve. The half-life of each mRNA in Ctrl (black) and null (purple) cells is represented as the mean +/- SD.
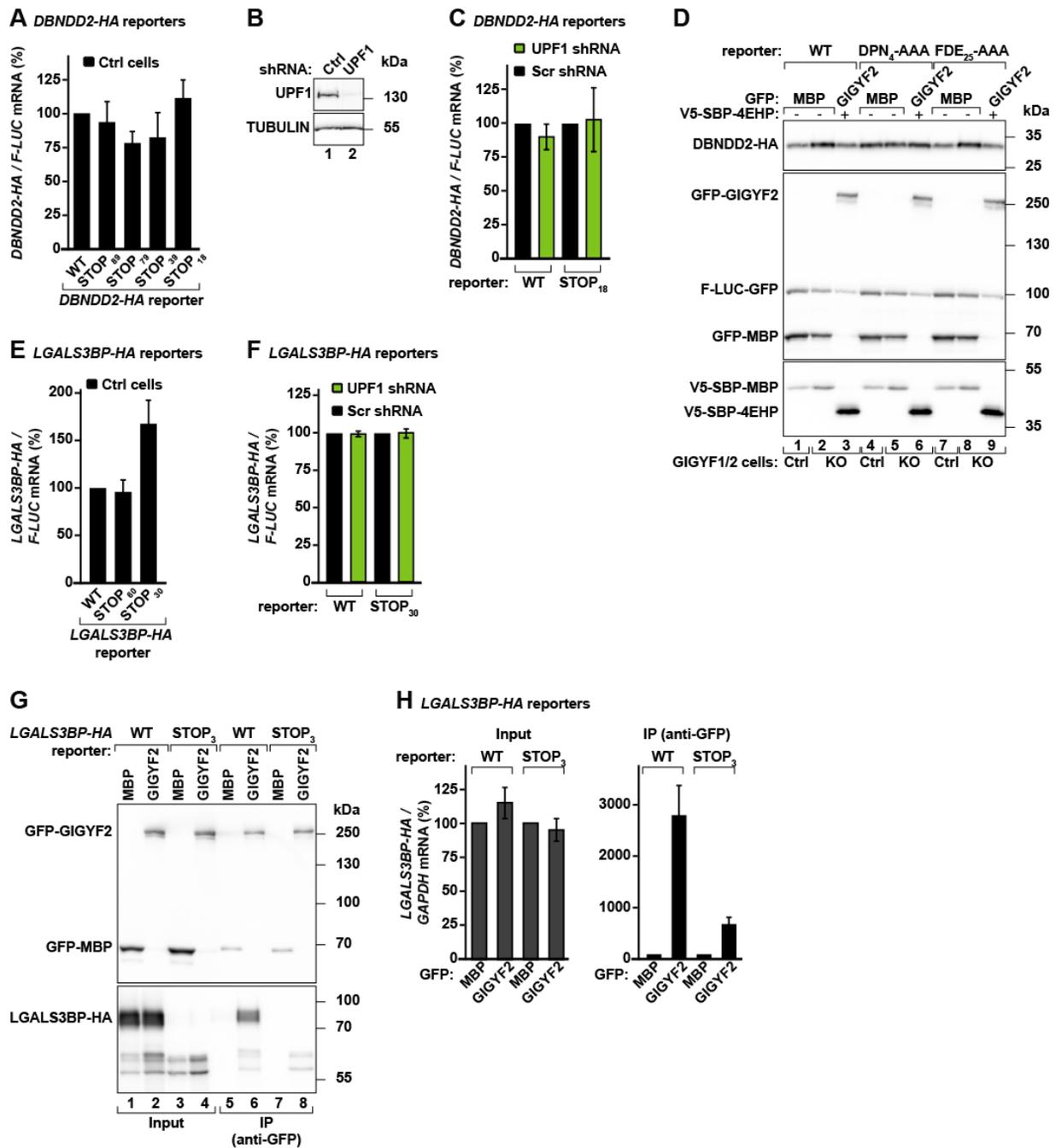
**Figure S6, related to Figures 5 and 6. Wild type and mutant target-based CDS reporters**

(A) Comparison of *DBNDD2-HA* (WT or STOP$_x$) reporter levels in HEK293T cells, as assessed by northern blotting (see Figure 5C). mRNA levels were normalized to *F-LUC-GFP* and set to 100% for the WT reporter. Bars represent the mean values and error bars denote the SD of three independent experiments.

(B) Western blot showing shRNA-mediated depletion of UPF1 in HEK293T cells. TUBULIN served as a loading control.

(C) HEK293T cells were treated with scramble (Scr) or shRNA targeting *UPF1* mRNA and transfected with *DBNDD2-HA* (WT or STOP$_{18}$) and *F-LUC-GFP*. The graph shows *DBNDD2-HA* mRNA abundance in control (Scr) and UPF1 KD cells. mRNA levels were determined by RT-qPCR, normalized to that of *F-LUC-GFP* and set to 100% in Scr-treated cells.

(D) Ctrl and GIGYF1/2-null cells were transfected with WT or mutant *DBNDD2-HA* plasmids. Protein samples were analyzed by western blotting using anti-V5, anti-HA and anti-GFP antibodies.

(E) Quantification of *LGALS3BP-HA* (WT or STOP$_x$) mRNA levels in HEK293T cells. RNA samples were analyzed by RT-qPCR (see Figure 6C). mRNA levels were normalized to *F-LUC-GFP* and set to 100% for the WT reporter. Bars represent the mean values and error bars denote the SD of three independent experiments.

(F) HEK2933T cells treated with Scr or UPF1 shRNAs were transfected with *LGALS3BP-HA* (WT or STOP$_{30}$) and *F-LUC-GFP*. *LGALS3BP-HA* mRNA levels were determined by RT-qPCR, normalized to that of *F-LUC-GFP* and set to 100% in Scr-treated cells.

(G, H) The interaction of GFP-GIGYF2 with LGALS3BP-HA protein (G) and mRNA (H) was analyzed by co-immunoprecipitation in HEK293T cells. Proteins were immunoprecipitated using anti-GFP antibodies. GFP-MBP served as a negative control. Input (0.2% for the GFP proteins and 0.7% for LGALS3BP-HA) and immunoprecipitated fractions (2.7% for the GFP proteins and 5% for LGALS3BP-HA) were analyzed by western blotting with anti-GFP and anti-HA antibodies. In H, RNA samples were obtained for each experimental condition and *LGALS3BP-HA* (WT or STOP$_3$) transcript abundance in input (0.8%) and IP samples (12%) was quantified by RT-qPCR, normalized to *GAPDH* and set to 100% in the presence of GFP-MBP. Bars represent the mean value. Error bars represent standard deviations from three independent experiments.
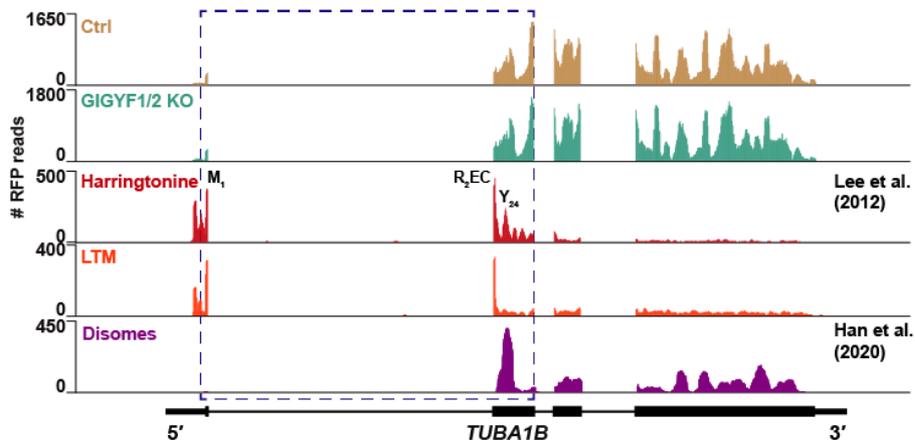
A RFP reads *TUBA1B*

B RFP reads *Mm Tuba4a*
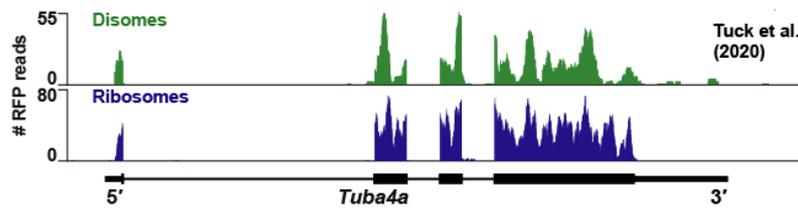
C RFP reads *Mm Tuba1a*
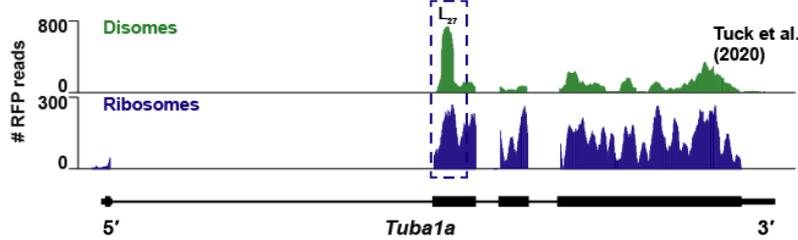
D RFP reads *Mm Tuba1b*

E

15
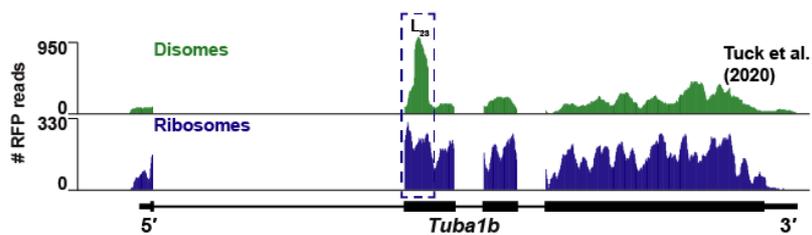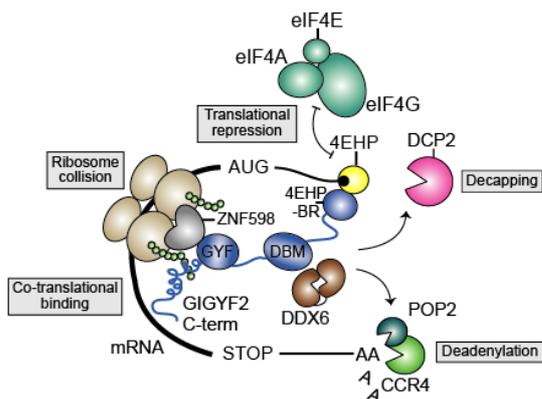
Figure S7, Related to Figure 7. Monosome and disome density profiles in *TUBULIN* mRNAs reveal translational pausing

(A) RFP of *TUBA1B* in Ctrl, GIGYF1/2 KO, and cells treated with harringtonine or lactimidomycin (LTM) as determined by Lee and co-workers (Lee et al., 2012). Harringtonine footprints identify the initiating ribosome at the translation start site (M_1) and paused elongating ribosomes (dashed blue square). The profile also shows the distribution of disome footprints along the CDS in HEK293 cells as obtained by Han and co-workers (Han et al., 2020). Met_1 (M_1); Arg_2 (R_2), Glu (E), Cys (C), Tyr_{24} (Y_{24}). Transcript UTRs, intron and exons are depicted below the profiles.

**(B-D)** Monosome and disome density profiles of *Mm Tuba1a*, *Tuba4a* and *Tuba1b* transcripts based on the data previously obtained by Tuck and co-workers in mouse embryonic stem cells (Tuck et al., 2020). The dashed blue box indicates the occurrence of ribosome collision (disomes) in the first 20-30 codons of tubulins. Transcript UTRs, intron and exons are depicted below the profiles.

(E) Recruitment of 4EHP-GIGYF1/2 complexes to transcripts with perturbed elongation induces translation repression, mRNA deadenylation and decapping. 4EHP, in yellow, competes with eIF4F (eIF4E+eIF4G+eIF4A) for cap-binding, blocking translation and promoting decapping. Binding of 4EHP to the mRNA depends on GIGYF1/2 proteins (in blue), the scaffolds of the repressor complexes. mRNA selection involves the recognition of paused ribosomes by factors such as the E3 ubiquitin ligase ZNF598 (in grey), which binds to the GYF domain of GIGYF1/2. In addition, target selection can be favored by co-translational binding of the C-terminal (C-term) region of GIGYF2 or components of the surveillance machinery to the nascent peptide chain. Recognition of the nascent chain by specific factors, or the synthesis of the nascent peptide itself, might then interfere with ribosome activity, causing ribosome pausing and collisions. Detection of such events coupled to the recruitment of 4EHP and GIGYF1/2, exposes the translating mRNA to degradation. GIGYF1/2 also recruit the RNA

helicase DDX6 which is required for target repression and decay. Altogether, GIGYF1/2 initiate a series of events that irrevocably prevent the translation of mRNAs with impaired elongation. DCP2: decapping enzyme 2; POP2 and CCR4: deadenylases; 4EHP-BR: 4EHP-binding region; DBM: DDX6-binding motif.

Table S8. primers used in this study

		sequence (5' to 3')
qPCR		
DBNDD2	fwd	CCAGCAGCTCCGCCTTC
	rev	GTTGTCCACCCCAGACGAC
CD109	fwd	GGTTGAGGAGCATACTGAAAAT
	rev	TGGCAGTCTAATGCTCACACCC
NPTX1	fwd	TCTGCAGGGATCTTCTCCGTTT
	rev	TCCCAGCTGTGGGAATCCTTTA
ITPR3	fwd	CTGCTGCATTTGTGGACACCTG
	rev	CACTACGCAGGTCAGCGAAGGT
ENO2	fwd	ATGTGTCACTTGTGCTTTGCTC
	rev	ACCCCAGTCATCTTGGGATCTA
CXCL16	fwd	CTCCAGATCTGCCGGTTCATTA
	rev	ATCACCCAGTGTGAAAAGCAGA
TUBA4A	fwd	TGAAACTGGTGCTGGAAAACAC
	rev	CTCCATCAGGAGTGAGGTGAAG
TUBB4A	fwd	CTCGAGGCTTCTGACCTTTGAT
	rev	TTAAAGGTGCGGTTTCCAGAGT
TUBA1A	fwd	CCACAGTCATTGATGAAGTTCG
	rev	GCTGTGGAAAACCAAGAAGC
TUBA1B	fwd	AATTCGCAAGCTGGCTGA
	rev	CGACAGATGTCATAGATGGCC
TUBB	fwd	GAAGCCACAGGTGGCAAATA
	rev	CGTACCACATCCAGGACAGA
GAPDH	fwd	CTCTGCTCCTCCTGTTCGACAG
	rev	TTCCCGTTCTCAGCCTTGACGG
β-ACTIN	fwd	GCAGGAGTATGACGAGTCCGGC
	rev	GTAACAACGCATCTCATATTTG
18S rRNA	fwd	CAGCCACCCGAGATTGAGCA
	rev	TAGTAGCGACGGGCGGTGTG
LGALS3BP-HA	fwd	CTGGGCCTCACCAAGTCTGGCG
	rev	AGCGTAATCTGGAACATCGTAT
DBNDD2-HA	fwd	CCAGCAGCTCCGCCTTC
	rev	AGCGTAATCTGGAACATCGTAT
sgRNA		
sg4EHP-a		TATAGCCACATGGTACGTCC
sg4EHP-b		TGTTTTCTTCATTCTGATCA
sgZNF598		CTACTGCGCCGTGTGCCGCG (Garzia et al., 2017)
sgZNF598		GAAAGGTGTACGCATTGTAC (Garzia et al., 2017)
shRNA		
Scramble		ATTCTCCGAACGTGTCACG (Jonas et al., 2013)
UPF1-I		GAGAATCGCCTACTTCACT (Paillusson et al., 2005)
UPF1-II		GATGCAGTTCCGCTCCATT (Paillusson et al., 2005)
Ribosome profiling		

30 nt RNA marker	AUGUACACGGAGUCGAGCUCAACCCGCAAC-P
27 nt RNA marker	AUGUACACGGAGUCGAGCUCAACCCGC-P
3' adapter	rApp/NNNNT GGA ATT CTC GGG TGC AA GG/3InvdT/
5' adapter (RNA)	GUUCAGAGUUCUACAGUCCGACGAUCNNNN
Reverse transcription primer	GCCTTGGCACCCGAGAATTCCA
Forward primer	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGA
Barcoded reverse primer	CAAGCAGAAGACGGCATACGAGAT_NNNNNN_GTGACTGGAGTTCCTTGGCACCCGAGAATTCCA

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse monoclonal anti-GFP	Roche	Cat. #11814460001
Rabbit polyclonal anti-GFP	In house	
Rabbit polyclonal anti-*Hs*GIGYF1	Bethyl laboratories	Cat. #A304-132A-M
Rabbit polyclonal anti-*Hs*GIGYF2	Bethyl laboratories	Cat. #A303-731A
Rabbit polyclonal anti-*Hs*ZNF598	Bethyl laboratories	Cat. #A305-108A
Rabbit polyclonal anti-*Hs*4EHP	In house	
Mouse monoclonal anti-HA (HRP)	Roche	Cat. #12013819001
Mouse monoclonal anti-Tubulin	Sigma Aldrich	Cat. #T6199
Mouse monoclonal anti-V5	LSBio LifeSpan BioSciences, Inc.	Cat. #LS-C57305
Anti-RENT1 (UPF1)	Bethyl laboratories	Cat. #A301-902A
Anti-R-LUC	Abcam	Cat. #ab185925
Donkey polyclonal anti-rabbit IgG (HRP)	GE Healthcare	Cat. #NA934V
Sheep polyclonal anti-mouse IgG (HRP)	GE Healthcare	Cat. #RPN4201
Chemicals, Peptides, and Recombinant Proteins		
Lipofectamine 2000	Thermo Scientific	Cat. #11668-019
DMEM	In house	
FBS	Thermo Scientific	Cat. #10270-106
Penicillin-Streptomycin	Thermo Scientific	Cat. #15140-122
L-Glutamine	Thermo Scientific	Cat. #25030-024
Agarose	Thermo Scientific	Cat. #16500-500
Low melting Agarose	Serva Electrophoresis	Cat. #11384
Phusion DNA polymerase	Thermo Scientific	Cat. #F-530XL
IgG beads	GE Healthcare	Cat. #17-0885-04
Streptavidin beads	GE Healthcare	Cat. #17-5113-01
TRIzol	Thermo Scientific	Cat. #15596018
TriFast FL	Peqlab Biotechnologies	Cat. #30-2120
P:C:I, stabilized	PanReac AppliChem	Cat. #A0889,0100
RiboLock	Thermo Scientific	Cat. #EO0381

dNTPs	Thermo Scientific	Cat. #R0141, #R0171, #R0161, #R0152
RevertAid H Minus reverse transcriptase	Thermo Scientific	Cat. #EP0451
Actinomycin D	Sigma Aldrich	Cat. #A9415
Puromycin	Serva Electrophoresis	Cat. #33835
Harringtonine	Sigma Aldrich	Cat. #SML1091
Protease Inhibitor (cOmplete, EDTA-free)	Roche	Cat. #05056489001
Yeast RNA	Roche	Cat. #10109223001
Nitrocellulose Transfer Membrane	Santa Cruz Biotechnology	Cat. #sc-3724
Luminol	Roth	Cat. #4203.1
P-Coumaric acid	Sigma Aldrich	Cat. #C9008
GeneScreen Plus nylon Membrane	Perkin Elmer	Cat. #NEF1018001PK
Ribosome profiling materials		
Cycloheximide	Serva Electrophoresis	Cat. #10700
Nocodazole	Sigma Aldrich	Cat #M1404
TurboDNase	Thermo Scientific	Cat. #AM2238
RNA Clean & Concentrator Kit	Zymo Research	Cat. #R1015
Zymoclean Gel DNA Recovery Kit	Zymo Research	Cat. #D4007
MicroSpin S-400 HR Columns	GE Healthcare	Cat. #27514001
RNaseI	Thermo Scientific	Cat. #AM2294
SUPERase Inhibitor	Thermo Scientific	Cat. #AM2694
Ribo-Zero Gold rRNA Removal Kit	Illumina	discontinued
SYBR Gold	Thermo Scientific	Cat. #S11494
SuperScript II reverse transcriptase	Thermo Scientific	Cat. #18064014
SuperScript III reverse transcriptase	Thermo Scientific	Cat. #18080044
T4 Polynucleotide kinase	NEB	Cat. #M0201S
GlycoBlue	Thermo Scientific	Cat. #AM9515
T4 RNA Ligase 2, truncated K227Q	NEB	Cat. #M0351S
5' DNA Adenylation Kit	NEB	Cat. #E2610S
T4 RNA Ligase 1	NEB	Cat. #M0204S
Critical Commercial Assays		

RNeasy Mini Kit	Qiagen	Cat. #74104
Dual-luciferase reporter assay	Promega	Cat. #E1960
iTaq Sybr Green Supermix	Biorad	Cat. #170-8885
Wizard SV Genomic DNA Purification System	Promega	Cat. #A2360
TruSeq RNA sample Prep Kit	Illumina	Cat. #RS-122-2002
Deposited Data		
RNASeq and RiboSeq accession numbers		
Raw and analyzed data	This paper	GEO: GSE144841, GSE149279
Harringtonine and LTM treated HEK293 cells	(Lee et al., 2012)	SRA: SRA056377
Monosome and disome profiling of HEK293 cells	(Han et al., 2020)	GEO: GSE145723
Monosome and disome profiling in mESCs	(Tuck et al., 2020)	GEO: GSE134020
Experimental Models: Cell Lines		
HEK293T	DSMZ	ACC 635
GIGYF1/2 KO	(Peter et al., 2017)	
4EHP KO	(Rasch et al., 2020)	
ZNF598 KO	This study	
Oligonucleotides (See Table S5)		
Recombinant DNA		
pλN-HA-C1-*Hs*eIF4E2 (4EHP)	(Peter et al., 2017)	Uniprot: O60573-1
pλN-HA-C1-*Hs*eIF4E2 W124A (cap*)	(Peter et al., 2017)	
pλN-HA-C1-*Hs*eIF4E2 R103L, E149L (S*)	(Peter et al., 2017)	
pT7-V5-SBP-C1-*Hs*eIF4E2 (4EHP)	(Peter et al., 2017)	
pT7-V5-SBP- C1-*Hs*eIF4E2 W124A (cap*)	(Peter et al., 2017)	
pT7-V5-SBP- C1-*Hs*eIF4E2 R103L, E149L (S*)	(Peter et al., 2017)	
pT7-EGFP-C1-*Hs*GIGYF1	(Peter et al., 2017)	Uniprot: O60573-1
pT7-EGFP-C1-*Hs*GIGYF1 Y39A Y41A M46A L47A (C*)	(Peter et al., 2017)	
pT7-EGFP-C1-*Hs*GIGYF1 Y479A F490A W498A F504A (GYF*)	(Peter et al., 2017)	
pT7-EGFP-C1-*Hs*GIGYF1 W294A, F306A, F312A (DDX6*)	(Peter et al., 2019)	
pT7-EGFP-C1-*Hs*GIGYF2	(Peter et al., 2017)	Uniprot: Q6Y7W6-1
pT7-EGFP-C1-*Hs*GIGYF2 Y41A Y43A M48A L49A (C*)	(Peter et al., 2017)	

pT7-EGFP-C1-*Hs*GIGYF2 Y538A, F549A, W557A, F563A (GYF*)	(Peter et al., 2017)	
pT7-EGFP-C1-*Hs*GIGYF2 W288A, F300A, F306A (DDX6*)	(Peter et al., 2019)	
pT7-EGFP-C1-*Hs*GIGYF2 C-term (719-1299)	This study	
pEGFP-N3-F-Luc-EGFP	(Lazzaretti et al., 2009)	
pCIneo-*Hs*DBNDD2-HA	This study	Uniprot: Q9BQY9-2
pCIneo-*Hs*DBNDD2-STOP$_3$-HA	This study	
pCIneo-*Hs*DBNDD2-STOP$_{18}$-HA	This study	
pCIneo-*Hs*DBNDD2-STOP$_{39}$-HA	This study	
pCIneo-*Hs*DBNDD2-STOP$_{79}$-HA	This study	
pCIneo-*Hs*DBNDD2-STOP$_{89}$-HA	This study	
pCIneo-*Hs*DBNDD2-FED$_{25}$-AAA-HA	This study	
pCIneo-*Hs*DBNDD2-DPN$_4$-AAA-HA	This study	
pCIneo-R-Luc-*Hs*DBNDD2 3' UTR	This study	
pCIneo-*Hs*LGALS3BP-HA	This study	Uniprot: Q08380-1
pCIneo-*Hs*LGALS3BP-STOP$_3$-HA	This study	
pCIneo-*Hs*LGALS3BP-STOP$_{30}$-HA	This study	
pCIneo-*Hs*LGALS3BP-STOP$_{60}$-HA	This study	
pCIneo-R-LUC-*Hs*LGALS3BP 3' UTR	This study	
pCIneo-*Hs* SP$_{LGALS3BP}$-R-LUC	This study	
pCIneo-*Hs* SP$_{NPTX1}$-R-LUC	This study	
pCIneo-R-LUC	(Pillai et al., 2004)	
pSpCas9(BB)-2A-Puro (PX459)	(Ran et al., 2013)	Addgene 48139
pSUPERpuro-BglII-scrambled	(Jonas et al., 2013)	
pSUPERpuro-BglII-HsUPF1-t2	(Paillusson et al., 2005)	
pSUPERpuro-BglII-HsUPF1-t4	(Paillusson et al., 2005)	
Software and Algorithms		
Adobe Illustrator	Adobe	https://www.adobe.com/uk/creativecloud.html
Integrative Genomics Viewer	(Robinson et al., 2011: Thorvaldsdottir et al., 2013)	https://software.broadinstitute.org/software/igv/

ImageJ	(Schneider et al., 2012)	https://imagej.nih.gov/ij/
Bowtie 2	(Langmead and Salzberg, 2012)	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
TopHat 2	(Kim et al., 2013)	https://ccb.jhu.edu/software/tophat/index.shtml
RiboTaper	(Calviello et al., 2016)	https://ohlerlab.mdc-berlin.de/software/RiboTaper_126/
QuasR	(Gaidatzis et al., 2015)	https://bioconductor.org/packages/release/bioc/html/QuasR.html
edgeR	(McCarthy et al., 2012; Robinson et al., 2010)	https://bioconductor.org/packages/release/bioc/html/edgeR.html
RiboDiff	(Zhong et al., 2017)	https://github.com/ratschlab/RiboDiff
goseq	(Young et al., 2010)	https://bioconductor.org/packages/release/bioc/html/goseq.html
biomaRt	Durinck et al., 2005; Durinck et al. 2009)	https://bioconductor.org/packages/release/bioc/html/biomaRt.html
CHOPCHOP	(Labun et al., 2016; Labun et al., 2019; Montague et al., 2014)	http://chopchop.cbu.uib.no

Molecular basis for GIGYF–Me31B complex assembly in 4EHP-mediated translational repression

Daniel Peter,[1,2,4] Vincenzo Ruscica,[1,4] Praveen Bawankar,[1,3,4] Ramona Weber,[1] Sigrun Helms,[1] Eugene Valkov,[1] Cátia Igreja,[1] and Elisa Izaurralde[1,5]

[1]Department of Biochemistry, Max Planck Institute for Developmental Biology, D-72076 Tübingen, Germany; [2]European Molecular Biology Laboratory, 38042 Grenoble Cedex 9, France; [3]Institute of Molecular Biology, 55128 Mainz, Germany

GIGYF (Grb10-interacting GYF [glycine–tyrosine–phenylalanine domain]) proteins coordinate with 4EHP (eIF4E [eukaryotic initiation factor 4E] homologous protein), the DEAD (Asp–Glu–Ala–Asp)-box helicase Me31B/DDX6, and mRNA-binding proteins to elicit transcript-specific repression. However, the underlying molecular mechanism remains unclear. Here, we report that GIGYF contains a motif necessary and sufficient for direct interaction with Me31B/DDX6. A 2.4 Å crystal structure of the GIGYF–Me31B complex reveals that this motif arranges into a coil connected to a β hairpin on binding to conserved hydrophobic patches on the Me31B RecA2 domain. Structure-guided mutants indicate that 4EHP–GIGYF–DDX6 complex assembly is required for tristetraprolin-mediated down-regulation of an AU-rich mRNA, thus revealing the molecular principles of translational repression.

Supplemental material is available for this article.

Initiation of translation by the eukaryotic initiation factor 4E (eIF4E) is regulated by competitor cap-binding proteins of the eIF4E family, such as the eIF4E homologous protein (4EHP; also known as eIF4E2) (Kong and Lasko 2012). 4EHP is responsible for the assembly of translational repressor complexes that inhibit mRNA expression in different biological contexts (Cho et al. 2005, 2006; Villaescusa et al. 2009; Chapat et al. 2017). 4EHP specifically associates with Grb10-interacting GYF (glycine–tyrosine–phenylalanine domain) protein 1 (GIGYF1) and

GIGYF2. These proteins possess an N-terminal 4EHP-binding region (4EHP-BR) and a central compacted GYF domain (Supplemental Fig. S1A; Ash et al. 2010; Peter et al. 2017) that mediates the interaction with ZNF598, tristetraprolin (TTP), or the microRNA (miRNA)-induced silencing complex-associated TNRC6 proteins (Morita et al. 2012; Fu et al. 2016; Schopp et al. 2017). These RNA-associated proteins recruit the 4EHP–GIGYF2 complex to specific mRNAs important for mouse embryonic development, cytokine mRNA expression, or repression of miRNA targets, respectively (Morita et al. 2012; Fu et al. 2016; Tollenaere et al. 2019).

GIGYF proteins do not simply bridge 4EHP to the RNA-associated proteins but rather participate directly in the repression mechanism (Peter et al. 2017). Human GIGYF2 regulates the expression of a subset of mRNAs via the recruitment of the CCR4–NOT complex (Amaya Ramirez et al. 2018). GIGYF proteins also associate with DDX6 (Me31B in *Drosophila melanogaster* [*Dm*] and Dhh1p in yeast) (Amaya Ramirez et al. 2018; Ruscica et al. 2019), an important regulator of gene expression (Ostareck et al. 2014; Wang et al. 2015; Lumb et al. 2017) that acts as translational repressor and enhancer of mRNA decapping (Coller et al. 2001; Radhakrishnan et al. 2016).

DDX6 orthologs are RNA-dependent ATPases of the DEAD (Asp–Glu–Ala–Asp)-box family that feature two globular RecA-like domains (RecA1 and RecA2) connected by a flexible linker. DEAD-box proteins use ATP binding and hydrolysis coupled to RNA binding to promote conformational transitions and remodeling of RNA and/or ribonucleoprotein particles (mRNPs) (Ozgur et al. 2015b). DDX6 has restricted conformational flexibility and limited ATPase activity and requires stimulation by interacting factors (Mathys et al. 2014).

DDX6 assembles in mutually exclusive complexes with P-body components such as EDC3, LSM14A, PatL1, and the eIF4E transporter protein (4E-T) (Jonas and Izaurralde 2013). These proteins use different short linear motifs to associate with two binding pockets in the RecA2 domain of DDX6, referred to here as Phe–Asp–Phe (FDF) and Trp (W) pockets (Tritschler et al. 2008; Sharif et al. 2013; Ozgur et al. 2015a; Brandmann et al. 2018).

To elucidate how GIGYF proteins function together with DDX6 in the regulation of mRNA expression, we determined the crystal structure of an N-terminal conserved motif from *Dm* GIGYF that mediates direct binding to Me31B (Fig. 1A; Supplemental Fig. S1A). This binding motif, characterized by a Pro–Glu–Trp (PEW) sequence and a "split" FDF sequence, binds to Me31B in a unique manner. We further show that recruitment of DDX6 via GIGYF2 is required in human cells for efficient translational repression of an AU-rich reporter mRNA by TTP. Collectively, these data have advanced our understanding of the molecular principles governing the assembly of mRNPs that rely on the 4EHP–GIGYF complex and DDX6 proteins to posttranscriptionally regulate gene expression.

Figure 1. GIGYF proteins contain a conserved Me31B/DDX6-binding motif (MBM). (*A*) Sequence alignment of the MBM of *Dm* and *Homo sapiens* (*Hs*) GIGYF with the CUP homology domain (CHD) of *Dm* and *Hs* 4E-T and *Dm* CUP. Residues with >70% similarity are shown with a light-colored background. Conserved residues are highlighted with a darker background and are printed in white. Secondary structure elements based on the structures presented in this study are indicated *above* the *Dm* GIGYF sequence. Boxed residues highlight the PEW (green) and FDF/IEL (black) motifs. The asterisk indicates the polar residue preceding the FDF motif. (*B*) The binding of HA-*Dm* GIGYF (FL or the indicated proteins) to Me31B was analyzed in coimmunoprecipitation (co-IP) assays using anti-HA antibodies upon S2 cell transfection. HA-MBP served as a negative control. The input (1.5% for the HA proteins and 0.2% for Me31B) and immunoprecipitated (30% for the HA proteins and 45% for Me31B) fractions were analyzed by Western blotting using anti-HA and anti-Me31B antibodies. (*C*) The interaction between GFP-*Dm* Me31B (FL or the indicated RecA domains) and HA-*Dm* GIGYF N-terminal expressed in S2 cells was analyzed in co-IP assays using anti-GFP antibodies. GFP-MBP served as a negative control. Input (3% for the GFP proteins and 1% for the HA proteins) and immunoprecipitated (15% for the GFP proteins and 30% for the HA proteins) fractions were analyzed by Western blotting using anti-GFP and anti-HA antibodies. (*D*) GST pull-down assay showing the interaction between the GST-Me31B RecA2 domain and the MBP-*Dm* GIGYF MBM. GST served as a negative control. The starting material (6.25% for GST proteins and 2% for the MBP proteins) and bound fractions (20%) were analyzed by SDS-PAGE followed by Coomassie blue staining. The size markers (in kilodaltons) are shown at the *right* of each panel.

Results and Discussion

The GIGYF linear motif is necessary and sufficient to directly bind Me31B/DDX6

The GIGYF orthologs contain a short conserved sequence motif with partial similarity to the CUP homology domain (CHD) present in 4E-T proteins (Fig. 1A; Kamenska et al. 2014; Ruscica et al. 2019). Deletion of this Me31B/DDX6-binding motif (MBM) abrogated the interaction of Me31B/DDX6 with transiently expressed and tagged GIGYF (*Dm* GIGYF and *Homo sapiens* [*Hs*] GIGYF1/2) in *Drosophila* and human cells (Fig. 1B; Supplemental Fig. S1B,C; Ruscica et al. 2019). The MBM alone interacted with Me31B/DDX6 as efficiently as full-length (FL) GIGYF or the N-terminal fragment of GIGYF (Fig. 1B; Supplemental Fig. S1B,C), indicating that the MBM is necessary and sufficient for a stable interaction between the proteins.

In coimmunoprecipitation (co-IP) assays, GIGYF proteins associated with the RecA2, but not the RecA1, domain of *Dm* Me31B and human DDX6 (Fig. 1C; Supplemental Fig. S1D,E), as observed previously for other DDX6-interacting factors (Tritschler et al. 2009; Sharif et al. 2013; Ozgur et al. 2015a; Brandmann et al. 2018). The purified recombinant GST-tagged RecA2 domain of Me31B/DDX6 associated with MBP-tagged *Dm* GIGYF and human GIGYF1/2 MBM by pull-down (Fig. 1D; Supplemental Fig. S1F,G). The MBM thus has a crucial role in mediating a direct and stable interaction between GIGYF and DDX6.

The Dm GIGYF MBM interacts with Me31B using a bipartite mode

We hypothesized that the GIGYF MBM binds to the W pocket of DDX6 via the conserved PEW motif because of the apparent sequence similarity to the CHD region of 4E-T (Fig. 1A; Ozgur et al. 2015a). However, the presence of alanine or serine in place of the second phenylalanine in the FDF-like motif (FD<u>A/S</u>) and the absence of an Ile–Glu–Leu (IEL) motif as observed in 4E-T suggest that the binding mode to the conserved hydrophobic FDF pocket of DDX6 may have diverged. To investigate this further, we determined the crystal structure of the *Dm* GIGYF MBM (residues D343–G369) in complex with the RecA2 domain of Me31B (residues E264–V431) to 2.4 Å resolution (Supplemental Table S1).

The RecA2 domain of Me31B adopts a typical α/β-fold characterized by a central six-stranded parallel β sheet covered by α-helical layers on either side (Fig. 2A; Cheng et al. 2005). In the structure of the complex, the GIGYF MBM curves around helices α10 and α11 of Me31B to engage the conserved FDF and W pockets—known binding sites for *Hs* 4E-T, *Hs* and *Saccharomyces cerevisiae* (*Sc*) Edc3, *Sc* Pat1, and *Hs* LSM14A (Fig. 2A; Supplemental Fig. S2A–F; Tritschler et al. 2008; Sharif et al. 2013; Ozgur et al. 2015a; Brandmann et al. 2018). Two distinct structured elements can be identified in the MBM: a short coil running along helix α11 of Me31B and a β hairpin containing a "split" FDF motif (FDx$_4$F) (Fig. 2B–D).

The N-terminal PEW (P347, E348, and W349) peptide trio of the GIGYF MBM initiates a short coil that inserts the aromatic side chain of W349^{GIGYF} (equivalent to W221 in 4E-T) into the hydrophobic pocket formed by residues V283, L310, L311, and F370 between helices α10 and α11 of Me31B (Fig. 2B; Supplemental Fig. S3A). Other DDX6 interactors also feature a large aromatic residue (W91 in *Sc* Edc3, F192 in *Hs* EDC3, F42 in *Sc* Pat1, or F396 in *Hs* LSM14A) inserted at the equivalent pocket of Dhh1/DDX6 (Supplemental Figs. S3, S4A–C). Hydrogen bonding between the side chains of Q306^{Me31B} and K314^{Me31B} and the backbone oxygens of N345^{GIGYF} and A350^{GIGYF} lends additional stability to the interface (Fig. 2B).

The PEW sequence of the GIGYF MBM is then connected via a flexible linker to a β-hairpin structure formed at the FDF pocket of Me31B (Fig. 2A,C,D). The β hairpin serves to orient the FDF motif (F361, D362, and F367^{GIGYF}) to optimally engage Me31B. The F361 and F367^{GIGYF} are in positions structurally equivalent to those observed previously in other FDF or IEL sequences (Supplemental Fig. S5; Tritschler et al. 2008; Sharif et al. 2013; Ozgur et al. 2015a; Brandmann et al. 2018). The
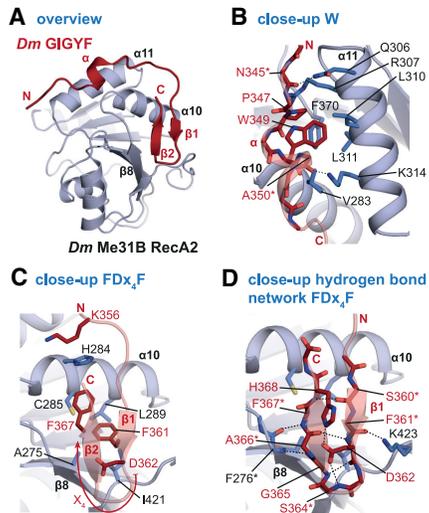
Figure 2. Structure of *Dm* GIGYF MBM bound to Me31B. (*A*) Overview of the structure of the *Dm* GIGYF MBM in complex with the Me31B RecA2 domain. Me31B is colored in light blue, and GIGYF is in red. Selected secondary structure elements are indicated. (*B*) Close-up view on the PEW sequence of GIGYF at the W pocket of Me31B. (*C,D*) Close-up views on the interactions of the FDx$_4$F β hairpin of GIGYF with Me31B. Selected interface residues are shown as sticks. For clarity reasons, all residues labeled with an asterisk are shown without their side chain.

aromatic rings of F361 and F367^{GIGYF} are accommodated in an "edge to face" orientation stabilized by a network of hydrophobic contacts formed by residues A275, H284, C285, L289, and I421^{Me31B} (Fig. 2C). The side chain of D362^{GIGYF} forms hydrogen bonds to the backbone nitrogen of S364^{GIGYF} as well as to the imidazole ring of H368^{GIGYF}, linking the loop region to the C-terminal portion of the hairpin (Fig. 2D). G365 and F367^{GIGYF} participate in backbone-mediated interactions with F276^{Me31B}, thus extending the Me31B β sheet at the tip of the β8 strand (Fig. 2D). The first hairpin strand (β1) is anchored to Me31B by hydrogen bonds between the side chain of K423^{Me31B} and the backbone oxygen of F361^{GIGYF} (Fig. 2D).

The GIGYF MBM contacts two conserved surfaces on *Dm* Me31B in a composite bipartite binding arrangement that combines the salient features of the 4E-T PEW motif with the FDF motif present in EDC3 and LSM14A homologs as well as *Sc* Pat1 (Fig. 1A; Supplemental Figs. S3, S5). Thus, despite the overall conservation of the interface and the mutually exclusive binding, GIGYF exhibits notable structural differences compared with other DDX6 interactors by using a "split" FDx$_4$F motif.

The GIGYF FDx$_4$F motif does not block NOT1 binding to DDX6

NOT1, the scaffold protein of the CCR4–NOT deadenylase complex, interacts with a surface of RecA2 domain adjacent to but not overlapping with the surface engaged by the other DDX6 interactors (Supplemental Fig. S2A; Chen et al. 2014; Mathys et al. 2014). Comparative structural analysis of DDX6-containing complexes indicates that the negatively charged residues preceding the FDF and DW motifs present in *Hs* EDC3 and PatL1 proteins (D203, F204, D205, and F206^{EDC3} and D43, D44, D45,

and W46^{PatL1}) (Supplemental S4A,B) are very likely to induce electrostatic repulsions to the NOT1 residues that face the DDX6 FDF pocket. This will impose a significant unfavorable energetic cost on the assembly of a ternary complex by EDC3/PatL1, DDX6, and NOT1 (Ozgur et al. 2015a). However, in the case of 4E-T, the IEL motif is preceded by a polar residue (Fig. 1A), which permits binding to the DDX6–CNOT1 complex. By analogy, GIGYF can, in principle, assemble into a ternary complex with DDX6–NOT1, as the residues located N-terminally to the FDx$_4$F motif are polar rather than negatively charged (Fig. 1A). We have not validated this structural hypothesis in cells, but an N-terminal region of human GIGYF2 containing the MBM does bind NOT1 in co-IP assays in HeLa cells (Amaya Ramirez et al. 2018), thus providing support to the notion of the existence of a functional CNOT1–DDX6–GIGYF2 ternary complex.

The bipartite binding mode is essential for GIGYF–DDX6 complex assembly

Guided by the structural analysis of the binding interfaces, we next substituted key residues on the W (LK-AA mutant) or FDF (CL-AA mutant) pockets in Me31B/DDX6 (Supplemental Table S2; Supplemental Fig. S4D) and tested binding by co-IP following transient expression in either *Dm* S2 or human cells. The interaction of GIGYF (HA-*Dm* GIGYF or *Hs* GIGYF1/2) with Me31B/DDX6 was strongly impaired by individual or combined pocket mutations (Fig. 3A,B), pointing to a crucial functional role for both binding pockets in stabilizing the association between the proteins.

Conversely, we also analyzed the effect of amino acid substitutions in GIGYF on the interaction with DDX6. Tryptophan substitution by alanine in the PEW motif (W* mutant), of both phenylalanines in the FDx$_4$F motif (FF* mutant), or in combination (WFF* mutant) (Supplemental Table S2) abolished the interaction of *Dm* and human GIGYF with Me31B/DDX6 in cells (Fig. 3C,D; Supplemental Fig. S6A).

Dm HPat and human PatL1 do not contain an FDF motif but rather contain a DW sequence motif that interacts with Me31B/DDX6 (Supplemental Fig. S4A). Interestingly, the mutations in the W and FDF pockets of Me31B/DDX6 also strongly reduced binding to HPat/PatL1, which is consistent with previous observations (Fig. 3A, B; Sharif et al. 2013). However, the disruption of the FDF pocket (CL-AA mutant) did not affect the association of Me31B/DDX6 with 4E-T or LSM14A and only mildly impaired binding to EDC3 (*Dm* and *Hs*) (Fig. 3A; Supplemental Fig. S6B–D). In contrast, the substitutions in the W pocket strongly reduced binding to *Dm* and human 4E-T, EDC3, and LSM14A (Fig. 3A; Supplemental Fig. S6B–D).

Collectively, these binding studies are consistent with a differential contribution of the two binding pockets in DDX6 toward promoting stable interactions with various factors. Reported differences in the binding affinities further support this model: Both *Sc* Pat1 and *Sc* EDC3 are high-affinity binders of *Sc* Dhh1 (K_d of 50 nM and 200 nM, respectively) (Sharif et al. 2013); human DDX6 interactors are rather more diverse in their affinities, with reported K_ds of 230 nM for PatL1, 390 nM for 4E-T, 410 nM for EDC3, and 1.62 µM for LSM14A (Brandmann et al. 2018).
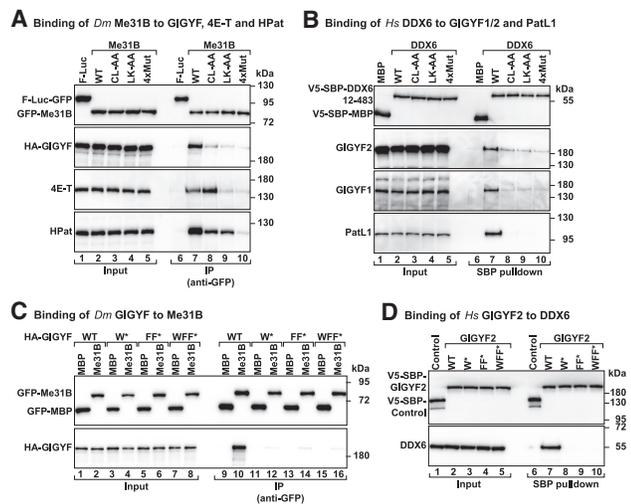
Figure 3. GIGYF proteins use a bipartite binding mode to address DDX6. (*A,B*) Analysis of the interaction of GFP-Me31B with HA-GIGYF, 4E-T, and HPat in S2 cells (*A*) or of V5-SBP-DDX6 with human GIGYF1/2 and PatL1 (*B*). The DDX6 proteins are either wild type (WT) or the indicated mutants. GFP proteins were immunoprecipitated using anti-GFP antibodies, whereas SBP proteins were pulled down using streptavidin-coated beads. Firefly luciferase (F-Luc)-GFP and V5-SBP-MBP served as negative controls. (*A*) The inputs for the *Dm* experiment were 3% for the GFP proteins and 1% for HA-GIGYF, 4E-T, and HPat, whereas bound fractions corresponded to 15% for the GFP proteins, 30% for HA-GIGYF and 4E-T, and 20% for HPat. (*B*) In the pull-down assay with the human proteins, inputs were 1.25% for the V5-SBP proteins and 0.5% for GIGYF1/2 and PatL1, while bound fractions corresponded to 5% for the V5-SBP proteins and 30% for the other proteins. Samples were analyzed by Western blotting using anti-GFP, anti-HA, and anti-V5 antibodies and protein-specific antibodies. (*C*) The interaction between GFP-Me31B and HA-GIGYF—wild type or the indicated mutants (W* [W349A], FF* [F361A, F367A], and WFF*) (Supplemental Table S2)— was analyzed in *Dm* S2 cell lysates using anti-GFP co-IP. GFP-MBP served as a negative control. The input (3% for the GFP proteins and 1% for the HA proteins) and bound fractions (15% for the GFP proteins and 30% for the HA proteins) were analyzed by Western blotting using the indicated antibodies. (*D*) Streptavidin-based pull-down assays showing the association of SBP-V5-*Hs* GIGYF2—wild type or the indicated mutants (W* [W288A], FF* [F300A, F306A], and WFF*) (Supplemental Table S2)—and DDX6. V5-SBP-MBP-F-Luc-EGFP served as a negative control. The input (1.25% for the V5-SBP proteins and 0.5% for DDX6) and bound fractions (8% for the V5-SBP proteins and 30% for DDX6) were analyzed by Western blotting using the indicated antibodies.

GIGYF2–DDX6 interaction contributes to TTP-mediated translational repression

To explore the functional relevance of the GIGYF–DDX6 interaction, we investigated the regulation of mRNA expression by TTP in human cells. TTP represses the expression of AU-rich transcripts via the recruitment of the 4EHP–GIGYF2 complex (Fu et al. 2016; Peter et al. 2017). To test the TTP-mediated repression in a reporter assay, we chose a *Renilla* luciferase (R-Luc) mRNA with two copies of the TNF-α mRNA AU-rich element (ARE) in the 3′ untranslated region (UTR) (Supplemental Fig. S7A). To distinguish the consequences of translational repression from degradation, an internal polyadenosine sequence of 90 nucleotides was "tailed" by a noncoding RNA MALAT1 sequence at the 3′ end, which is generated by RNase P endonucleolytic cleavage, rendering this reporter mRNA resistant to 5′–3′ decay (R-Luc-ARE-

A$_{90}$-MALAT1) (Peter et al. 2017). A plasmid encoding firefly luciferase (F-Luc-GFP) was included as a transfection and normalization control.

To bypass the recruitment of DDX6 via NOT1, we transiently expressed a TTP construct lacking the NOT1-binding motif (ΔCIM) (Fabian et al. 2013). We observed that TTP ΔCIM efficiently repressed the expression of the R-Luc reporter without altering its mRNA abundance in control cells (Fig. 4A; Supplemental Fig. S7B,C). By comparison, TTP-induced translational repression was alleviated in GIGYF1/2-null cells (GIGYF1/2 knockout) even though the observed level of TTP expression was comparable with that in the control cells (Fig. 4, A,B, lane 4 vs. lane 2). In GIGYF1/2-null cells, TTP-mediated translational repression was restored by coexpression of GIGYF2 and its stabilizing partner, 4EHP (Fig. 4A,B). However, the repressive function of TTP could be selectively impaired when 4EHP was coexpressed with the GIGYF2 mutants that do not bind to DDX6 (WFF*) or TTP (GYF*) (Fig. 4A,B; Supplemental Fig. S7D). The repressive function of TTP was critically dependent on the ARE, as none of the factors influenced the expression of a reporter lacking this sequence (R-Luc-A$_{95}$-MALAT1) (Supplemental Fig. S7E,F). Collectively, these data support a model in which the assembly of the 4EHP–GIGYF2–DDX6 complex is a prerequisite for TTP-mediated translational control of AU-rich transcripts.

Concluding remarks

In this study, we showed that GIGYF proteins interact directly with the RNA-dependent ATPase DDX6 via a short motif. This interaction is mutually exclusive with other DDX6-binding partners such as 4E-T, EDC3, LSM14A, and PatL1 and has an important functional role in posttranscriptional regulation (Fig. 4C). We showed that GIGYF2 is a direct link between DDX6 and TTP, which explains at the molecular level why DDX6 is required for ARE mRNA translational repression (Qi et al. 2012). The GIGYF–4EHP complex can also be part of TTP-independent mRNPs via direct mRNA binding (Amaya Ramirez et al. 2018) or the interaction with ZNF598 and TNRC6 proteins (Morita et al. 2012; Schopp et al. 2017) . As the latter are important in ribosome quality control (Garzia et al. 2017; Sundaramoorthy et al. 2017; Juszkiewicz et al. 2018) and miRNA-mediated gene silencing (Chapat et al. 2017), respectively, the control of mRNA expression by the 4EHP–GIGYF–DDX6 complex is relevant for a wide range of cellular transcripts. Furthermore, as all of the components of the complex have been implicated to function in miRNA-mediated translational repression (Chen et al. 2014; Mathys et al. 2014; Chapat et al. 2017; Schopp et al. 2017), the 4EHP–GIGYF–DDX6 complex is likely to have an important role in miRNA-mediated mechanisms.
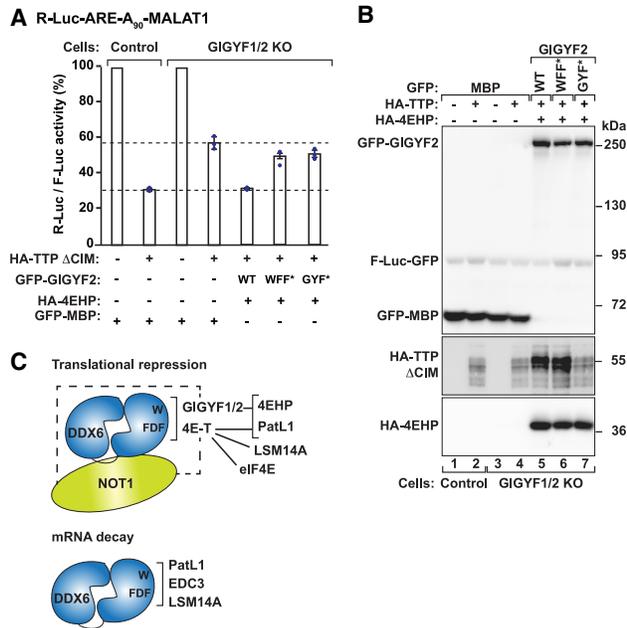
Figure 4. Recruitment of DDX6 by *Hs* GIGYF2 contributes to TTP-mediated translational repression of an ARE mRNA reporter. (*A*) Control or GIGYF1/2-null HEK293T cells (knockout [KO]) were transfected with the R-Luc-ARE-A$_{90}$-MALAT1 reporter and plasmids expressing wild type (WT) or the indicated GIGYF2 mutants, HA-4EHP, and HA-TTP ΔCIM (Fabian et al. 2013). An F-Luc-GFP reporter served as a transfection control. R-Luc activity was normalized to that of the F-Luc transfection control and set to 100% in the absence of TTP for each cell line. Bars represent the mean values, error bars represent standard deviations, and the blue dots represent the individual points from three independent experiments. (*B*) Western blot analysis showing the expression of the proteins used in the complementation assay. Note that TTP ΔCIM is stabilized in GIGYF1/2 knockout cells expressing GIGYF wild type or WFF*. However, repression did not correlate with TTP levels but with the coexpression of a functional repressor complex. (*C*) DDX6 complexes involved in translational repression and mRNA decay. GIGYF1/2, 4E-T, PatL1, EDC3, and LSM14A assemble mutually exclusive complexes with DDX6. NOT1 associates with a different surface of DDX6, but a ternary complex forms only in the presence of 4E-T and possibly GIGYF1/2. Thus, DDX6 serves as a molecular adaptor for the assembly of protein complexes with distinct molecular roles—translational repression versus decapping. Independently of DDX6, 4E-T and GIGYF1/2 bind to NOT1, 4EHP, and PatL1. 4E-T is also an eIF4E- and an LSM14A-binding protein.

Crystallization, data collection, and structure determination

Detailed descriptions of the crystallization conditions and of the structure determination are in the Supplemental Material.

Co-IP assays and Western blotting

Co-IP assays in HEK293T and Schneider S2 cells were performed in the presence of RNase A as described previously (Peter et al. 2015a). All Western blots were developed using the ECL Western blotting detection system (GE Healthcare). The antibodies used in this study are listed in Supplemental Table S3.

Pull-down assays

The in vitro pull-down assays were performed as described previously (Igreja et al. 2014; Peter et al. 2015a,b). The details are in the Supplemental Material.

Complementation assay

HEK293T cells (wild-type or GIGYF1/2-null cells) were seeded in six-well plates (0.6×10^6 cells per well) and transfected using Lipofectamine 2000 (Invitrogen). The transfection mixtures contained 1 μg of the R-Luc reporters and 0.25 μg of the F-Luc control in the presence of 50 ng of λN-HA-TTP ΔCIM, 0.2 μg of GFP-MBP, 1 μg of GFP-GIGYF2 (wild type or mutants), or 0.5 μg of λN-HA-4EHP. F-Luc and R-Luc activities were measured 2 d after transfection using the dual-luciferase reporter assay system (Promega). R-Luc activity was normalized to that of the F-Luc control and set to 100% in the absence of TTP in wild-type and GIGYF1/2-null cells. Total RNA was isolated using TriFast (Peqlab Biotechnologies), and the RNA samples were analyzed by Northern blot as described previously (Behm-Ansmant et al. 2006).

Data availability

Atomic coordinates and structure factors have been deposited in the Protein Data Bank under the accession codes 6S8R (*Dm* Me31B-GIGYF) and 6S8S (*Hs* DDX6-EDC3).

Acknowledgments

References

Amaya Ramirez CC, Hubbe P, Mandel N, Béthune J. 2018. 4EHP-independent repression of endogenous mRNAs by the RNA-binding protein GIGYF2. *Nucleic Acids Res* **46:** 5792–5808. doi:10.1093/nar/gky198

Ash MR, Faelber K, Kosslick D, Albert GI, Roske Y, Kofler M, Schuemann M, Krause E, Freund C. 2010. Conserved β-hairpin recognition by the GYF domains of Smy2 and GIGYF2 in mRNA surveillance and vesicular transport complexes. *Structure* **18:** 944–954. doi:10.1016/j.str.2010.04.020

Behm-Ansmant I, Rehwinkel J, Doerks T, Stark A, Bork P, Izaurralde E. 2006. mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. *Genes Dev* **20:** 1885–1898. doi:10.1101/gad.1424106

Brandmann T, Fakim H, Padamsi Z, Youn J-Y, Gingras A-C, Fabian MR, Jinek M. 2018. Molecular architecture of LSM14 interactions involved in the assembly of mRNA silencing complexes. *EMBO J* **37:** e97869. doi:10.15252/embj.201797869

Chapat C, Jafarnejad SM, Matta-Camacho E, Hesketh GG, Gelbart IA, Attig J, Gkogkas CG, Alain T, Stern-Ginossar N, Fabian MR, et al. 2017. Cap-binding protein 4EHP effects translation silencing by microRNAs. *Proc Natl Acad Sci* **114:** 5425–5430. doi:10.1073/pnas.1701488114

Chen Y, Boland A, Kuzuoğlu-Öztürk D, Bawankar P, Loh B, Chang C-T, Weichenrieder O, Izaurralde E. 2014. A DDX6–CNOT1 complex and W-binding pockets in CNOT9 reveal direct links between miRNA target recognition and silencing. *Mol Cell* **54:** 737–750. doi:10.1016/j.molcel.2014.03.034

Cheng Z, Coller J, Parker R, Song H. 2005. Crystal structure and functional analysis of DEAD-box protein Dhh1p. *RNA* **11:** 1258–1270. doi:10.1261/rna.2920905

Cho PF, Poulin F, Cho-Park YA, Cho-Park IB, Chicoine JD, Lasko P, Sonenberg N. 2005. A new paradigm for translational control: inhibition via 5′–3′ mRNA tethering by Bicoid and the eIF4E cognate 4EHP. *Cell* **121:** 411–423. doi:10.1016/j.cell.2005.02.024

Cho PF, Gamberi C, Cho-Park YA, Cho-Park IB, Lasko P, Sonenberg N. 2006. Cap-dependent translational inhibition establishes two opposing morphogen gradients in *Drosophila* embryos. *Curr Biol* **16:** 2035–2041. doi:10.1016/j.cub.2006.08.093

Coller JM, Tucker M, Sheth U, Valencia-Sanchez MA, Parker R. 2001. The DEAD box helicase, Dhh1p, functions in mRNA decapping and interacts with both the decapping and deadenylase complexes. *RNA* **7:** 1717–1727. doi:10.1017/S135583820101994X

Fabian MR, Frank F, Rouya C, Siddiqui N, Lai WS, Karetnikov A, Blackshear PJ, Nagar B, Sonenberg N. 2013. Structural basis for the recruitment of the human CCR4–NOT deadenylase complex by tristetraprolin. *Nat Struct Mol Biol* **20:** 735–739. doi:10.1038/nsmb.2572

Fu R, Olsen MT, Webb K, Bennett EJ, Lykke-Andersen J. 2016. Recruitment of the 4EHP–GYF2 cap-binding complex to tetraproline motifs of tristetraprolin promotes repression and degradation of mRNAs with AU-rich elements. *RNA* **22:** 373–382. doi:10.1261/rna.054833.115

Garzia A, Jafarnejad SM, Meyer C, Chapat C, Gogakos T, Morozov P, Amiri M, Shapiro M, Molina H, Tuschl T, et al. 2017. The E3 ubiquitin ligase and RNA-binding protein ZNF598 orchestrates ribosome quality control of premature polyadenylated mRNAs. *Nat Commun* **8:** 16056. doi:10.1038/ncomms16056

Igreja C, Peter D, Weiler C, Izaurralde E. 2014. 4E-BPs require non-canonical 4E-binding motifs and a lateral surface of eIF4E to repress translation. *Nat Commun* **5:** 4790. doi:10.1038/ncomms5790

Jonas S, Izaurralde E. 2013. The role of disordered protein regions in the assembly of decapping complexes and RNP granules. *Genes Dev* **27:** 2628–2641. doi:10.1101/gad.227843.113

Juszkiewicz S, Chandrasekaran V, Lin Z, Kraatz S, Ramakrishnan V, Hegde RS. 2018. ZNF598 is a quality control sensor of collided ribosomes. *Mol Cell* **72:** 469–481.e7. doi:10.1016/j.molcel.2018.08.037

Kamenska A, Lu W-T, Kubacka D, Broomhead H, Minshall N, Bushell M, Standart N. 2014. Human 4E-T represses translation of bound mRNAs and enhances microRNA-mediated silencing. *Nucleic Acids Res* **42:** 3298–3313. doi:10.1093/nar/gkt1265

Kong J, Lasko P. 2012. Translational control in cellular and developmental processes. *Nat Rev Genet* **13:** 383–394. doi:10.1038/nrg3184

Lumb JH, Li Q, Popov LM, Ding S, Keith MT, Merrill BD, Greenberg HB, Li JB, Carette JE. 2017. DDX6 represses aberrant activation of interferon-stimulated genes. *Cell Rep* **20:** 819–831. doi:10.1016/j.celrep.2017.06.085

Mathys H, Basquin J, Ozgur S, Czarnocki-Cieciura M, Bonneau F, Aartse A, Dziembowski A, Nowotny M, Conti E, Filipowicz W. 2014. Structural and biochemical insights to the role of the CCR4–NOT complex and DDX6 ATPase in microRNA repression. *Mol Cell* **54:** 751–765. doi:10.1016/j.molcel.2014.03.036

Morita M, Ler LW, Fabian MR, Siddiqui N, Mullin M, Henderson VC, Alain T, Fonseca BD, Karashchuk G, Bennett CF, et al. 2012. A novel 4EHP–GIGYF2 translational repressor complex is essential for mammalian development. *Mol Cell Biol* **32:** 3585–3593. doi:10.1128/MCB.00455-12

Ostareck DH, Naarmann-de Vries IS, Ostareck-Lederer A. 2014. DDX6 and its orthologs as modulators of cellular and viral RNA expression. *Wiley Interdiscip Rev RNA* **5:** 659–678. doi:10.1002/wrna.1237

Ozgur S, Basquin J, Kamenska A, Filipowicz W, Standart N, Conti E. 2015a. Structure of a human 4E-T/DDX6/CNOT1 complex reveals the differ-ent interplay of DDX6-binding proteins with the CCR4–NOT complex. *Cell Rep* **13:** 703–711. doi:10.1016/j.celrep.2015.09.033

Ozgur S, Buchwald G, Falk S, Chakrabarti S, Prabu JR, Conti E. 2015b. The conformational plasticity of eukaryotic RNA-dependent ATPases. *FEBS J* **282:** 850–863. doi:10.1111/febs.13198

Peter D, Igreja C, Weber R, Wohlbold L, Weiler C, Ebertsch L, Weichenrieder O, Izaurralde E. 2015a. Molecular architecture of 4E-BP translational inhibitors bound to eIF4E. *Mol Cell* 1074–1087. doi:10.1016/j.molcel.2015.01.017

Peter D, Weber R, Köne C, Chung MY, Ebertsch L, Truffault V, Weichenrieder O, Igreja C, Izaurralde E. 2015b. Mextli proteins use both canonical bipartite and novel tripartite binding modes to form eIF4E complexes that display differential sensitivity to 4E-BP regulation. *Genes Dev* **29:** 1835–1849. doi:10.1101/gad.269068.115

Peter D, Weber R, Sandmeir F, Wohlbold L, Helms S, Bawankar P, Valkov E, Igreja C, Izaurralde E. 2017. GIGYF1/2 proteins use auxiliary sequences to selectively bind to 4EHP and repress target mRNA expression. *Genes Dev* **31:** 1147–1161. doi:10.1101/gad.299420.117

Qi M-Y, Wang Z-Z, Zhang Z, Shao Q, Zeng A, Li X-Q, Li W-Q, Wang C, Tian F-J, Li Q, et al. 2012. AU-rich-element-dependent translation repression requires the cooperation of tristetraprolin and RCK/P54. *Mol Cell Biol* **32:** 913–928. doi:10.1128/MCB.05340-11

Radhakrishnan A, Chen Y-H, Martin S, Alhusaini N, Green R, Coller J. 2016. The DEAD-box protein Dhh1p couples mRNA decay and translation by monitoring codon optimality. *Cell* **167:** 122–132.e9. doi:10.1016/j.cell.2016.08.053

Ruscica V, Bawankar P, Peter D, Helms S, Igreja C, Izaurralde E. 2019. Direct role for the *Drosophila* GIGYF protein in 4EHP-mediated mRNA repression. *Nucleic Acids Res* **47:** 7035–7048. doi:10.1093/nar/gkz429

Schopp IM, Amaya Ramirez CC, Debeljak J, Kreibich E, Skribbe M, Wild K, Béthune J. 2017. Split-BioID a conditional proteomics approach to monitor the composition of spatiotemporally defined protein complexes. *Nat Commun* **8:** 15690. doi:10.1038/ncomms15690

Sharif H, Ozgur S, Sharma K, Basquin C, Urlaub H, Conti E. 2013. Structural analysis of the yeast Dhh1–Pat1 complex reveals how Dhh1 engages Pat1, Edc3 and RNA in mutually exclusive interactions. *Nucleic Acids Res* **41:** 8377–8390. doi:10.1093/nar/gkt600

Sundaramoorthy E, Leonard M, Mak R, Liao J, Fulzele A, Bennett EJ. 2017. ZNF598 and RACK1 regulate mammalian ribosome-associated quality control function by mediating regulatory 40S ribosomal ubiquitylation. *Mol Cell* **65:** 751–760.e4. doi:10.1016/j.molcel.2016.12.026

Tollenaere MAX, Tiedje C, Rasmussen S, Nielsen JC, Vind AC, Blasius M, Batth TS, Mailand N, Olsen JV, Gaestel M, et al. 2019. GIGYF1/2-driven cooperation between ZNF598 and TTP in posttranscriptional regulation of inflammatory signaling. *Cell Rep* **26:** 3511–3521.e4. doi:10.1016/j.celrep.2019.03.006

Tritschler F, Eulalio A, Helms S, Schmidt S, Coles M, Weichenrieder O, Izaurralde E, Truffault V. 2008. Similar modes of interaction enable Trailer Hitch and EDC3 to associate with DCP1 and Me31B in distinct protein complexes. *Mol Cell Biol* **28:** 6695–6708. doi:10.1128/MCB.00759-08

Tritschler F, Braun JE, Eulalio A, Truffault V, Izaurralde E, Weichenrieder O. 2009. Structural basis for the mutually exclusive anchoring of P body components EDC3 and Tral to the DEAD box protein DDX6/Me31B. *Mol Cell* **33:** 661–668. doi:10.1016/j.molcel.2009.02.014

Villaescusa JC, Buratti C, Penkov D, Mathiasen L, Planagumà J, Ferretti E, Blasi F. 2009. Cytoplasmic Prep1 interacts with 4EHP inhibiting Hoxb4 translation. *PLoS One* **4:** e5213. doi:10.1371/journal.pone.0005213

Wang Y, Arribas-Layton M, Chen Y, Lykke-Andersen J, Sen GL. 2015. DDX6 orchestrates mammalian progenitor function through the mRNA degradation and translation pathways. *Mol Cell* **60:** 118–130. doi:10.1016/j.molcel.2015.08.014

Molecular basis for GIGYF-Me31B complex assembly in 4EHP-mediated translational repression

Daniel Peter[1,2,4], Vincenzo Ruscica[1,4], Praveen Bawankar[1,3,4], Ramona Weber[1], Sigrun Helms[1], Eugene Valkov[1,*], Cátia Igreja[1,*], Elisa Izaurralde[1,†]

[1]Department of Biochemistry, Max Planck Institute for Developmental Biology, Max-Planck-Ring 5, D-72076 Tübingen, Germany

[2]European Molecular Biology Laboratory, 71 avenue des Martyrs, CS 90181, 38042 Grenoble Cedex 9, France.

[3]Institute of Molecular Biology gGMbH, Ackermannweg 4, 55128 Mainz, Germany.

[4]These authors contributed equally to this work.

[*]Corresponding authors: catia.igreja@tuebingen.mpg.de;

eugene.valkov@tuebingen.mpg.de

Tel: +49-7071-601-1370

Fax: +49-7071-601-1353

[†]Deceased April 30, 2018.

Supplemental Information

Supplemental Material

DNA constructs

To generate plasmids for expression in *Escherichia coli*, DNA fragments coding for the human GIGYF1/2 (residues E286–K316^{GIGYF1} and D280–K310^{GIGYF2}) and for *Drosophila melanogaster* (*Dm*) GIGYF (residues D341–G369) were inserted into the XhoI-BamHI (human constructs) and NdeI-NheI (*Drosophila* construct) restriction sites of the pnEA-NpM vector (Diebold et al. 2011), respectively. These constructs express the GIGYF fragments fused N-terminally to a maltose-binding protein (MBP) tag cleavable by the HRV 3C protease. The DNA sequences coding for the RecA2 domain of human DDX6 (residues E303–E472) and of *Dm* Me31B (residues E264–V431) were cloned into the XhoI-BamHI and NdeI-NheI restriction sites of the pnYC-NpG vector (Diebold et al. 2011), respectively. These constructs express the Me31B/DDX6 fragments fused N-terminally to a Glutathione S-transferase (GST) tag cleavable by the HRV 3C protease.

To obtain the plasmid for expression of λN-HA-tagged *Dm* GIGYF in *Dm* S2 cells, the cDNA corresponding to the GIGYF ORF was inserted into the EcoRI and NotI restriction sites of the pAc5.1B-λN-HA vector. The GIGYF N-terminal (residues M1–N640) fragment was amplified by PCR using as template the plasmid containing FL GIGYF and then cloned into the same vector and restriction sites. The cDNA encoding the minimal Me31B-binding region (residues D341–G369) of *Dm* GIGYF was inserted into the EcoRI and NheI restriction sites of the pAc5.1B-λN-HA-GST-V5-His vector. The plasmids required for the expression of GFP- or HA-tagged *Dm* Me31B and *Dm* EDC3 (FL and fragments) were described previously (Tritschler et al. 2007). To obtain the plasmid for expression of F-Luc-GFP in S2 cells, F-Luc cDNA was inserted into the KpnI and XhoI restriction sites of the pAc5.1C vector (Invitrogen). The EGFP cDNA was amplified by PCR, digested with SalI and XhoI enzymes and then cloned into the XhoI restriction site of the pAc5.1C-F-Luc vector. To generate the plasmid expressing

GFP-MBP in S2 cells, the MBP ORF was inserted into the HindIII and NotI restriction sites of the pAc5.1B-EGFP vector.

The ARE reporter, 4EHP, GIGYF2 and the TTP plasmids used in the complementation assay of Fig. 4 were described previously (Peter et al. 2017). Briefly, to generate the R-Luc reporter containing the ARE element (pCIneo-R-Luc-ARE-A_{90}-MALAT1) the sequence of the ARE element present in the 3' UTR of the TNF (Tumor Necrosis Factor)-α mRNA was inserted twice into the 3' UTR of the pCIneo-R-Luc parental plasmid by site-directed mutagenesis. A cDNA containing a stretch of 90 adenines and the mouse MALAT1 3' region sequence (nucleotides 6581-6754, GenBankEF177380.1) were then inserted into the XhoI and NotI restriction sites of the R-Luc-ARE vector. The DNA sequence of the TNF-α ARE is as follows: TTATTTATTATTTATTTATTATTTATTTATTT. To obtain the R-Luc control reporter lacking the ARE sequence [pCIneo-R-Luc-A_{95}-MALAT1; (Kuzuoglu-Ozturk et al. 2016)], the R-Luc ORF was cloned into the NheI and XbaI restriction sites of the pCIneo backbone. A random cDNA stretch was then inserted between XbaI and the SalI restriction sites of the vector to add space between the R-Luc ORF and the polyA stretch. A 95 nucleotide long polyA stretch was inserted by annealing between the SacII and PacI restriction sites of the plasmid. The mouse MALAT1 3' region was inserted downstream of the polyA sequence into the PacI and NotI restriction sites. The XhoI and BamHI restriction sites were used to clone full length 4EHP and GIGYF2 cDNAs into the pλN-HA-C1, pT7-V5-SBP-C1 and pT7-EGFP-C1 vectors. The cDNA encoding TTP ΔCIM (residues M1-313) (Fabian et al. 2013), was inserted between the XhoI and EcoRI restriction sites of the pλN-HA-C1-vector. To obtain the F-Luc-GFP reporter used as a transfection control in the complementation assay, F-Luc cDNA was amplified by PCR using as template the pGL4.12 vector (Promega) and then inserted into the BamHI and NotI restriction sites of the pEGFP-N3 vector (Clontech).

The plasmids required for the expression in human cells of V5-SBP-GIGYF1 and 2 were obtained by insertion of the corresponding ORF cDNA into the XhoI-EcoRI and XhoI-BamHI

restriction sites of the pT7-V5-SBP-C1 vector, respectively. FL GIGYF1 and 2 ORFs were also cloned into the XhoI-Not1 and XhoI-XbaI restriction sites of the pcDNA3.1-MS2-HA-C1 vector, respectively. The N-terminal fragments of GIGYF1 (residues M1–S671) and GIGYF2 (residues M1–T718) were cloned into the pT7-V5-SBP-C1 vector using the restriction sites described for full length proteins. The cDNAs coding for the minimal DDX6-binding regions of GIGYF1 and 2 (residues E286–K316^{GIGYF1} and residues D280–K310^{GIGYF2}) were inserted into the XhoI-XbaI restriction sites of the pCIneo-V5-SBP-MBP vector. The DNA sequences coding for residues M12–P483, RecA1 (residues M12–L306) and RecA2 (residues K307–P483) domains of human DDX6 were cloned into the BglII-BamHI restriction sites of the pT7-EGFP-C1 vector. DDX6 (residues M12–S283) was also subcloned into the pT7-V5-SBP-C1 vector using the same restriction sites. All residue numbers in the *Hs* DDX6 clones used in this study are according to the updated version of the cDNA sequence NM_004397.5. To obtain the plasmid expressing V5-SBP-MBP, SBP (Streptavidin Binding Protein) ORF cDNA was inserted into the NheI and XhoI restriction sites of the pCIneo-V5 vector. MBP cDNA was amplified by PCR, digested with SalI and XhoI and cloned into the XhoI restriction site of the pCIneo-V5-SBP construct. To clone the plasmid expressing V5-SBP-MBP-F-Luc-EGFP in human cells, EGFP-F-Luc cDNA amplified by PCR using pEGFP-N3-F-Luc as template was inserted into the EcoRI and NotI restriction sites of the pCIneo-V5-SBP-MBP vector. To clone the plasmid required for expression of GFP-MBP in human cells, the cDNA corresponding to the MBP ORF was inserted into the XhoI and BamHI restriction sites of the pT7-EGFP-C1 vector.

All the mutants used in this study were generated by site-directed mutagenesis using the QuickChange mutagenesis kit (Stratagene). All the constructs and mutations were confirmed by sequencing and are listed in Supplemental Table S2.

Protein production and purification

The GST-tagged RecA2 domain of *Dm* Me31B used for crystallization was produced in *E. coli* BL21 Star (DE3) cells (Invitrogen) grown in Terrific Broth (TB) medium overnight at 20°C. The cells were lysed by sonication in lysis buffer containing 50 mM sodium phosphate buffer (pH 7.0), 200 mM NaCl and 2 mM DTT supplemented with DNaseI (5 µg/ml), lysozyme (1 mg/ml) and protease inhibitor cocktail (Roche). GST-Me31B (residues E264–V431) was purified from cleared cell lysates using Protino® Glutathione Agarose 4B (Macherey-Nagel), followed by cleavage of the GST tag with HRV 3C protease (home made) overnight at 4°C. The protein was further separated from the cleaved tag using a heparin column (HiTrap Heparin HP 5 ml, GE Healthcare). The protein eluted from the heparin column was pooled and the buffer was exchanged to 10 mM HEPES (pH 7.2), 200 mM NaCl, 2 mM DTT using a VivaSpin 20 centrifugal concentrator (Sartorius; 10,000 MWCO). The protein was stored at -80°C or used directly for crystallization.

For the pulldown assays, the GST-tagged RecA2 domains of human DDX6 (residues E303–E472) and *Dm* Me31B (residues E264–V431) were produced and purified as described above with the difference that the GST tags were not cleaved. The final buffer contained 20 mM sodium phosphate buffer (pH 7.0), 200 mM NaCl and 2 mM DTT.

N-terminally GST tagged HRV 3C protease was produced in *E. coli* BL21 (DE3) Gold cells (Invitrogen) grown overnight at 20°C in Luria-Bertani Broth (LB). The cells were lysed in an EmulsiFlex homogenizer in lysis buffer containing 50 mM Tris-HCl buffer (pH 7.8), 300 mM NaCl and 1 mM DTT supplemented with DNaseI (5 µg/ml), lysozyme (1 mg/ml) and protease inhibitor cocktail (Roche). GST-HRV 3C protease was purified from cleared cell lysates using Protino® Glutathione Agarose 4B (Macherey-Nagel) and dialysed overnight in storage buffer containing 50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1 mM DTT, 10 mM EDTA and 20% of glycerol. The protein was aliquoted and stored at -80°C.

Crystallization

The complex of *Dm* Me31B (residues E264–V431) and GIGYF (residues D342–G368) was reconstituted by incubating the purified Me31B with a synthetic GIGYF peptide (EMC microcollections GmbH) dissolved in the same buffer [2.5 mM; 10 mM HEPES (pH 7.2), 200 mM NaCl, 2 mM DTT]. The final mixture contained 10 mg/mL Me31B (app. 500 μM) with a 1.5x molar excess of the GIGYF peptide (750 μM; app. 2.1 mg/mL). Initial crystals were obtained at 20°C using the sitting-drop vapor diffusion method three days after mixing the Me31B-GIGYF protein solution (10 mg/ml Me31B + 1.5x molar excess of GIGYF; 0.2 μl) with the crystallization solution (0.2 μl) containing 0.1 M sodium acetate (pH 5.0), 0.2 M ammonium chloride and 20% (w/v) PEG 6000. Crystals were optimized by iterative microseeding into drops consisting of 1μl protein solution and 1μl crystallization solution containing 0.1 M sodium acetate (pH 5.0), 0.15 M ammonium chloride and 16% (w/v) PEG 6000. The crystals were grown at 18°C using hanging-drop vapor diffusion.

All crystals were soaked in mother liquor supplemented with 15% (v/v) glycerol for cryoprotection before flash-cooling.

Data collection and structure determination

Data for the Me31B–GIGYF crystals were collected at a wavelength of 0.9996 Å at 100K on a PILATUS 6M detector at the PXII beamline of the Swiss Light Source. Diffraction data were processed with XDS and scaled using XSCALE (Kabsch 2010). The initial phases were obtained by molecular replacement using *phenix.auto*mr (McCoy et al. 2007) using the coordinates of the RecA2 domain of *Hs* DDX6 [PDB 5ANR; (Ozgur et al. 2015)] as a search model (one copy of the model in the asymmetric unit). The initial model of the Me31B RecA2 domain was rebuilt using the *phenix.autobuild* routine (Terwilliger et al. 2008). Analysis of the reflections indicated the presence of significant anisotropy, which complicated further refinement leading to unacceptably high R-factors. The reflections were then analyzed using

the STARANISO server (http://staraniso.globalphasing.org/cgi-bin/staraniso.cgi) and ellipsoidal resolution boundaries were applied during reprocessing (2.4 Å, 2.5 Å and 2.6 Å for a*, b* and c*, respectively). Anisotropic scaling and B-factor sharpening were further utilized to correct for the anisotropy. Following anisotropy correction, the iterative cycles of model building and refinement were carried out with COOT (Emsley et al. 2010) and *phenix.refine* (Afonine et al. 2012), respectively. The peptide chain of GIGYF was then built manually into the difference density in COOT and further refined with *phenix.refine*.

The X-ray diffraction data previously collected for the *Hs* EDC3−DDX6 complex [PDB 2WAX, (Tritschler et al. 2009)] were reprocessed *de novo* with AutoPROC (Vonrhein et al. 2011). The phases were obtained by molecular replacement using *phenix.automr* (Adams et al. 2011) and the coordinates of the RecA2 domain of the *Hs* DDX6 [PDB 5ANR; (Ozgur et al. 2015) as a search model. To analyze in detail the density for the moiety occupying the W binding pocket of DDX6, which was previously modeled by a CAPS (N-cyclohexyl-3-aminopropanesulfonic acid) molecule [PDB 2WAX, (Tritschler et al. 2009)], the EDC3 peptide chains were extended N-terminally into the difference density using COOT and further refined with BUSTER (Smart et al. 2012). In the final refinement rounds of the complex, translation/libration/screw (TLS) parameters were refined for the peptide chain in addition to the individual B-factors.

The stereochemical properties for all structures were verified with MOLPROBITY (Chen et al. 2010), and structural images were prepared with PyMOL (http://www.pymol.org). The data processing and refinement statistics are summarized in Table S1.

Pulldown assays

In the pulldown assays shown in Fig. 1D and Supplemental Fig. S1F and G, purified GST-*Hs* DDX6 (residues E303–E472) or GST-*Dm* Me31B (residues E264–V431; each 2 μM, ca. 50 μg total) were incubated with glutathione-agarose beads (Macherey-Nagel) for 30 min. The

immobilized DDX6 and Me31B proteins were then incubated for 30 min with bacterial lysates expressing GIGYF fragments tagged N-terminally with MBP. Proteins associated with DDX6 or Me31B were eluted with glutathione and analyzed by SDS-PAGE followed by Coomassie Blue staining.

Supplemental Figures

A

**B** Binding of *Hs* GIGYF1 to DDX6
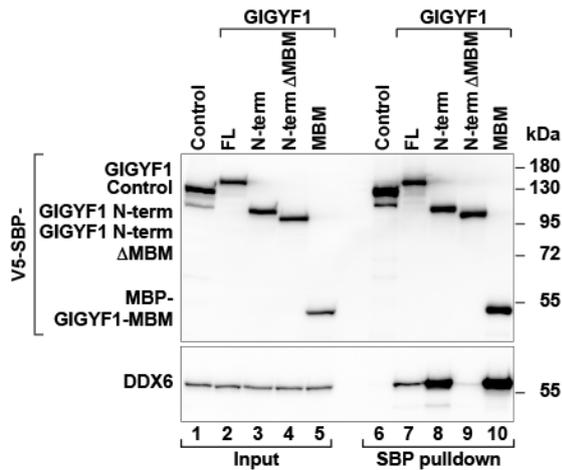
C Binding of *Hs* GIGYF2 to DDX6
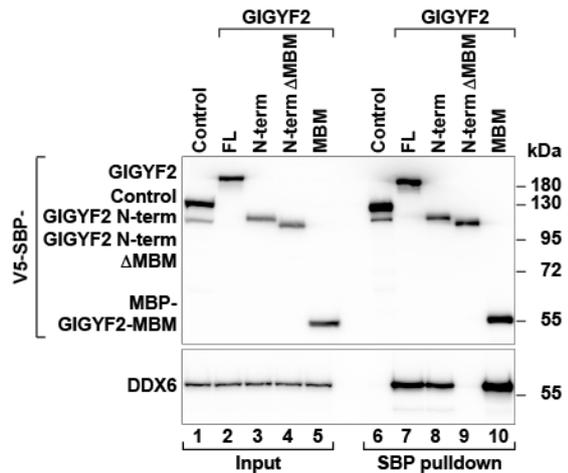
**D** Binding of *Hs* GIGYF1 to DDX6
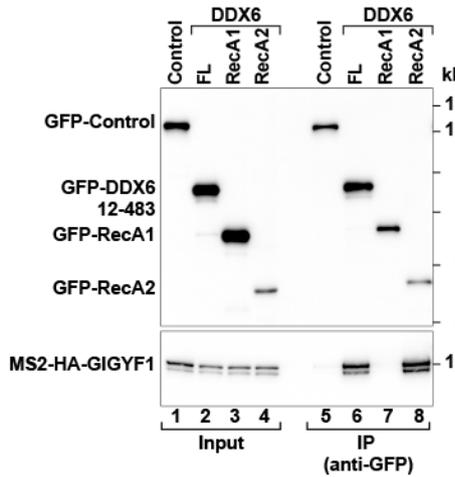
E Binding of *Hs* GIGYF2 to DDX6
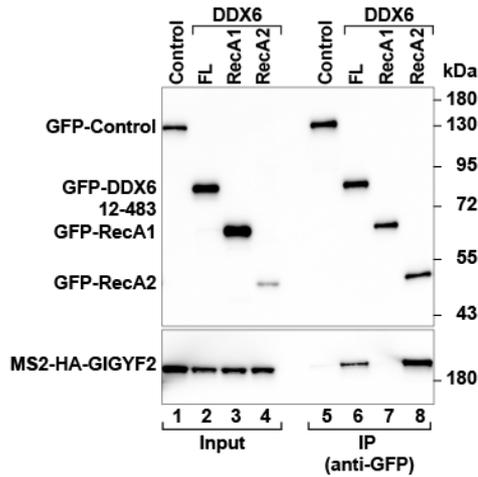
F Binding of *Hs* GIGYF1 to DDX6 *in vitro*
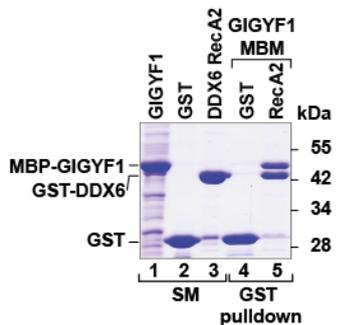
G Binding of *Hs* GIGYF2 to DDX6 *in vitro*
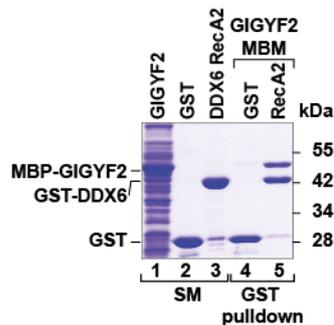
**Supplemental Figure S1.** Human GIGYF1/2 proteins contain an MBM.

(*A*) Schematic representation of *Dm* and human GIGYF. These proteins are divided into N- and C-terminal (term) regions. The N-term contains a 4EHP-binding region (4EHP-BR), a glycine-tyrosine-phenylalanine (GYF) domain and a conserved MBM.

(*B, C*) Immunoprecipitation assays showing the interaction between V5-SBP-GIGYF1 (B) or GIGYF2 (C) (FL or the indicated fragments) and endogenous DDX6. The proteins were pulled down with streptavidin beads. V5-SBP-MBP-F-Luc-EGFP served as a negative control. The input (1.25% for the V5-proteins and 0.5% for DDX6) and bound fractions (8% for the V5-proteins and 30% for DDX6) were analyzed by western blotting using anti-V5 and anti-DDX6 antibodies.

(*D, E*) Western blot analysis showing the interaction between GFP-*Hs* DDX6 and HA-GIGYF1 (D) or GIGYF2 (E) in HEK293T cells after immunoprecipitation with anti-GFP antibodies. V5-SBP-MBP-F-Luc-EGFP served as a negative control. The input (1.5% for GFP-proteins and 15% for HA- proteins) and bound fractions (0.5% for GFP-proteins and 30% for HA-proteins) were analyzed by western blotting using anti-GFP and anti-HA antibodies.

(*F, G*) GST pulldown assays showing the interaction among the purified RecA2 domain of DDX6 and the MBP-MBM of GIGYF1 (F) and GIGYF2 (G). GST served as a negative control. The starting material (SM; 2-4% for MBP-proteins and 6.25% for GST-DDX6 RecA2) and bound fractions (20%) were analyzed by SDS-PAGE followed by Coomassie blue staining.
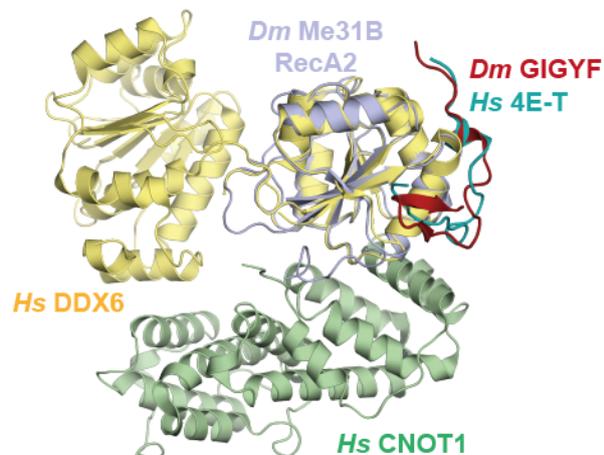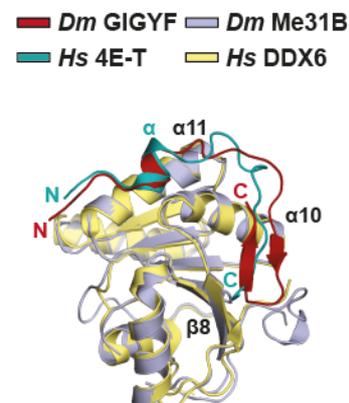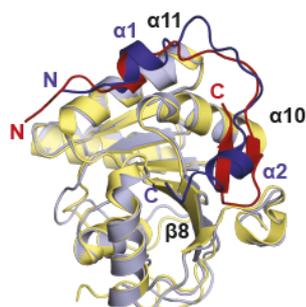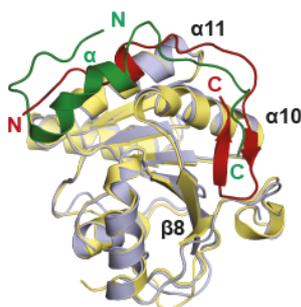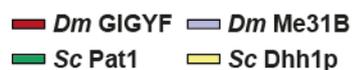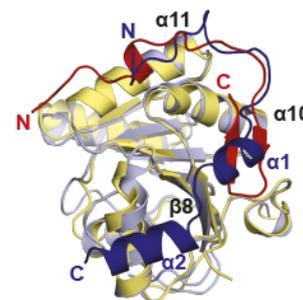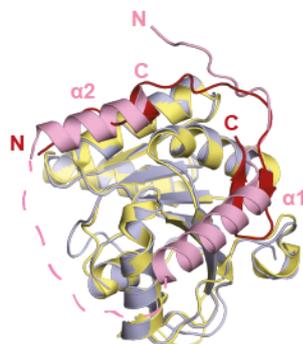
Supplemental Figure S2. Comparison of the *Dm* Me31B–GIGYF complex to other DDX6-complexes. Related to Fig. 2.

(**A, B**) Superposition of the *Dm* Me31B–GIGYF complex with the *Hs* CNOT1–DDX6–4E-T complex [PDB ID: 5ANR; (Ozgur et al. 2015)]. The superposition is shown in the context of the CNOT1–DDX6–4E-T complex (A) and as a close-up view on the RecA2 domain of DDX6/Me31B (B). *Dm* Me31B is colored in light blue, the *Dm* GIGYF peptide in red and *Hs* CNOT1, DDX6 and 4E-T in green, yellow and cyan, respectively. Only the RecA2 domains of DDX6/Me31B were used for the superposition, which align with an RMSD of 0.56 Å over 150 Cα atoms.

(**C, D**) Superposition of the *Dm* Me31B–GIGYF complex and the yeast Dhh1p-complexes with Edc3 (C) and Pat1 (D) [PDB IDs: 4BRU and 4BRW, respectively; (Sharif et al. 2013)]. The views only entail the RecA2 domains of Me31B/Dhh1. Colors for *Dm* Me31B and GIGYF are as above, *Sc* Dhh1 is colored in yellow, Edc3 in purple and Pat1 in green. The RecA2 domains of the complexes were superimposed and align with RMSDs of 0.5 Å over 167 Cα atoms (Edc3-complex) and 0.51 Å over 167 Cα atoms (Pat1-complex).

(**E, F**) Comparison of the *Dm* Me31B–GIGYF with the human DDX6–EDC3 (E) and DDX6–LSM14A [F; PDB ID: 6F9S; (Brandmann et al. 2018)] complexes. A re-analyzed version of the PDB 2WAX (Tritschler et al. 2009) structure was used to overlay the RecA2 domains of the human and *Dm* complexes (RMSD of 0.4 Å over 161 Cα atoms). In the re-analyzed version of the human structure, the CAPS buffer molecule located at the W pocket in the DDX6–EDC3 complex is replaced by an N-terminal extension of the EDC3 peptide. Colors for *Dm* Me31B and GIGYF are as above, *Hs* DDX6 is colored in yellow, EDC3 in dark blue and LSM14A in pink. The DDX6 RecA2 domains of the *Dm* Me31B–GIGYF and *Hs* DDX6–LSM14A complexes were superimposed and align with RMSDs of 0.507 Å over 150 Cα atoms.

Peter329219_Supplemental Fig. S3
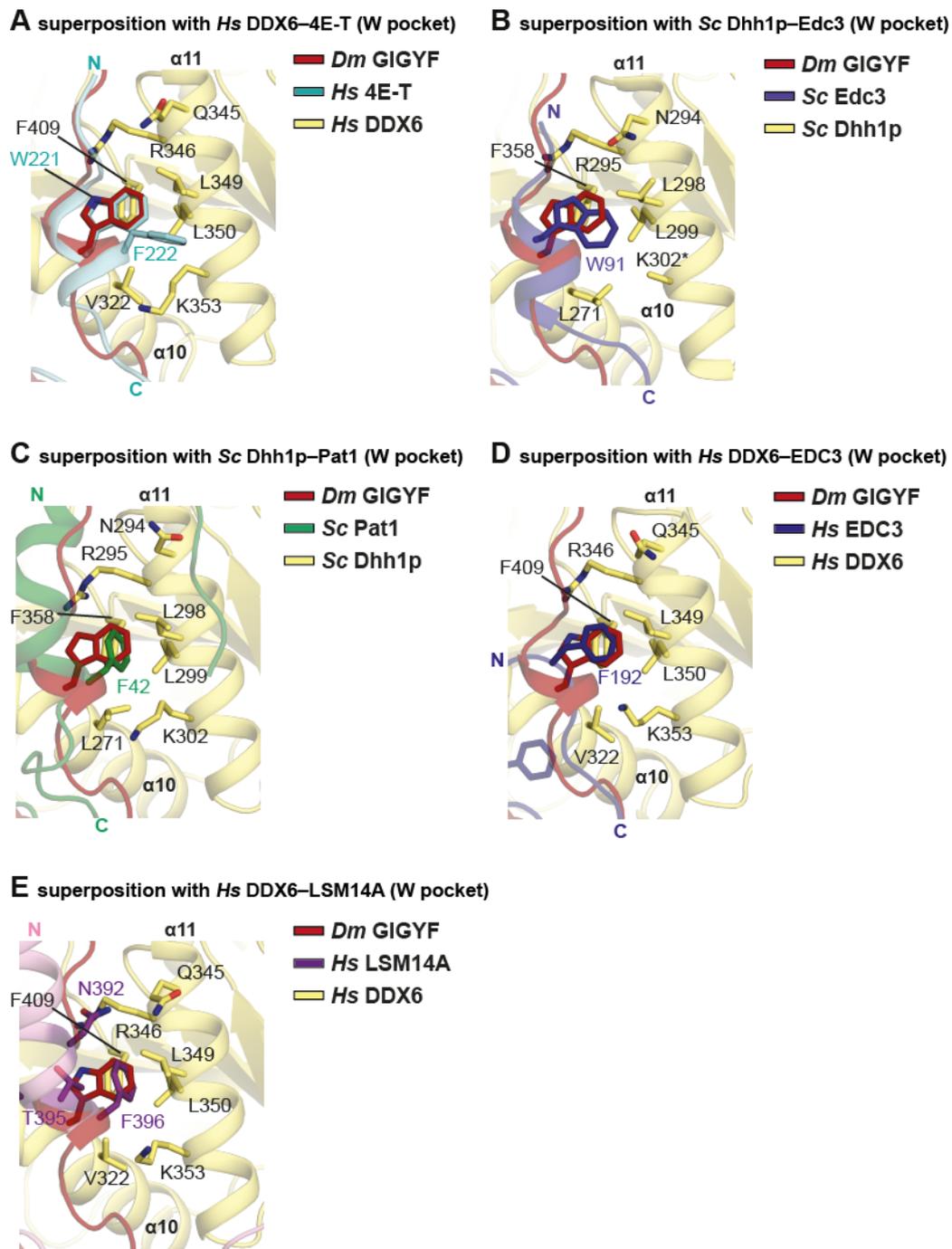
Supplemental Figure S3. The W pockets in the *Dm* Me31B–GIGYF and in the *Hs* DDX6/*Sc* Dhh1p-complexes. Related to Fig. 2.

(A-E) Close-up views on the interactions established by *Hs* 4E-T (A), *Sc* Edc3 (B), *Sc* Pat1 (C), *Hs* EDC3 (D) and *Hs* LSM14A (E) at the W pocket of DDX6-proteins compared to the *Dm* Me31B–GIGYF complex. Colors for the different proteins are as described in Supplemental Fig. S2 with the exception of LSM14A residues which are highlighted in purple. Selected

secondary structure elements are labeled in black for *Hs* DDX6/*Sc* Dhh1p. Me31B was omitted

in the superpositions for clarity reasons. Residue K302 in the Dhh1p–Edc3 complex was built

without a sidechain and is labeled with an asterisk. The PDB IDs are as in Supplemental Fig.

S2.

Supplemental Figure S4. Sequence alignments

(*A-C*) Sequence alignment of the Me31B-binding region of Pat1 (A), EDC3 (B) and LSM14A (C) proteins from *Drosophila melanogaster* (*Dm*), *Homo sapiens* (*Hs*) and *Saccharomyces cerevisiae* (*Sc*). In all aligned sequences, residues with >70% similarity are shown with a light color background and conserved residues are highlighted with a darker background and printed in white. Boxed residues indicate the residues interacting at the W (yellow), FDF (black) or FDK (orange) binding pockets of the RNA helicase. Secondary structure elements are indicated below the sequences for *Sc* Pat1 and *Sc* Edc3 or above the *Hs* EDC3 and LSM14A sequences and are based on the PDB accessions 4BRW, 4BRU, 2WAX and 6F9S, respectively (Tritschler

et al. 2009; Sharif et al. 2013; Brandmann et al. 2018). Red asterisk identifies the negatively charged residue preceding the FDF motif.

(**D**) Sequence alignment of DDX6 homologous proteins from *Drosophila melanogaster* (*Dm*), *Homo sapiens* (*Hs*), *Caenorhabditis elegans* (*Ce*) and *Saccharomyces cerevisiae* (*Sc*). Colors are as described in (A). Red open circles above the alignment indicate the residues mutated in this study. Secondary structure elements are indicated above the sequence for *Dm* Me31B and are based on the structure presented in this study. The numbering of the secondary structure elements takes the N-terminal RecA1 domain of DDX6 into account.

Peter329219_Supplemental Fig. S5
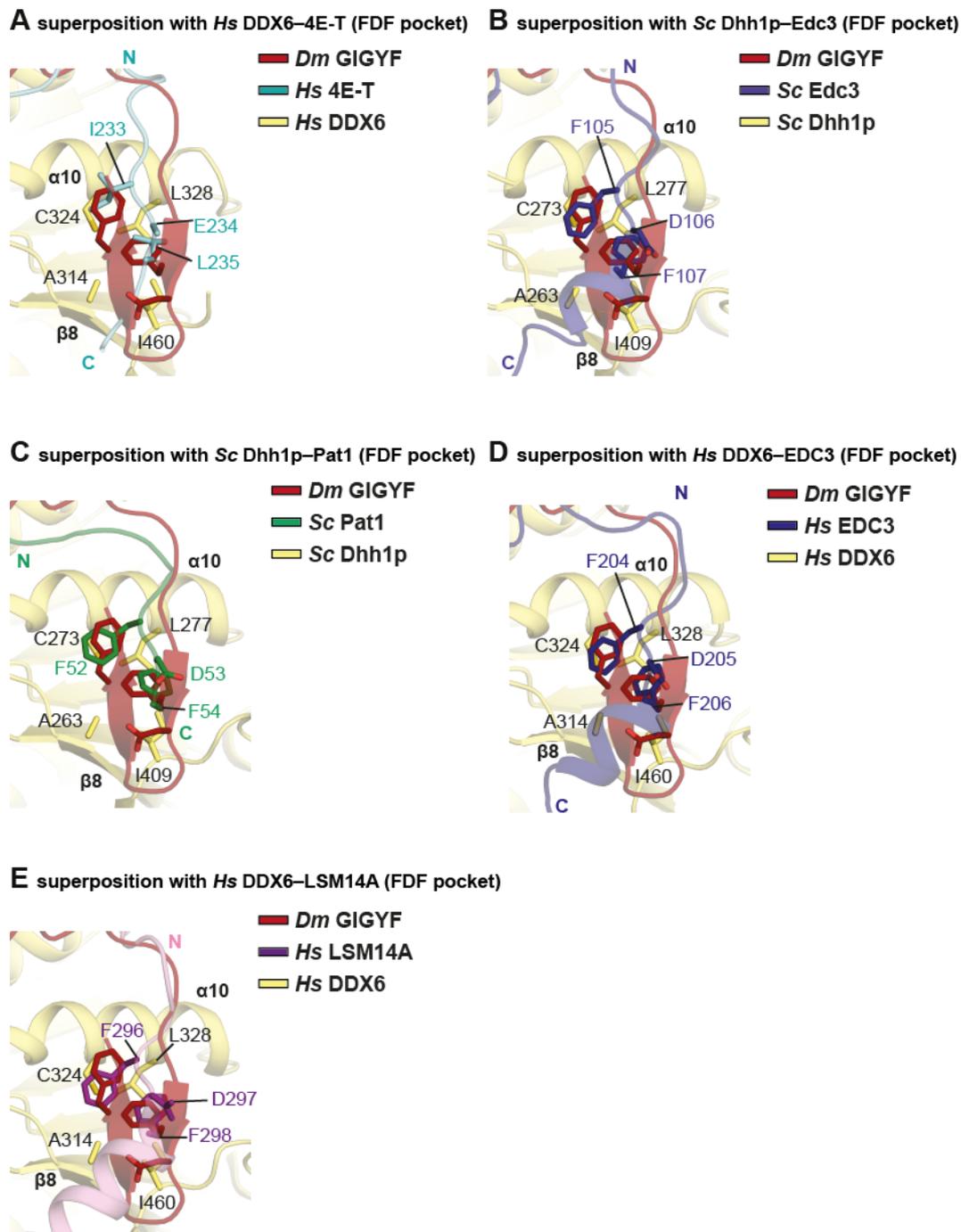
Supplemental Figure S5. The FDF pockets in the *Dm* Me31B–GIGYF and the *Hs* DDX6/*Sc*

Dhh1p-complexes. Related to Fig. 2.

(A-E) Close-up views on the interactions performed by *Hs* 4E-T (A), *Sc* Edc3 (B), *Sc* Pat1 (C),

Hs EDC3 (D) and *Hs* LSM14A (E) at the FDF pocket of the different DDX6-proteins compared

to the *Dm* Me31B–GIGYF complex. PDB IDs and colors for the different proteins are as

described in Supplemental Fig. S2 with the exception of LSM14A residues which are

highlighted in purple. Labels and secondary structure elements are as described in Supplemental

Figure S3. Selected residues are shown as sticks and are as described in Supplemental Fig. S2.

Me31B was omitted in the superpositions for clarity reasons.

Peter329219_Supplemental Fig. S6
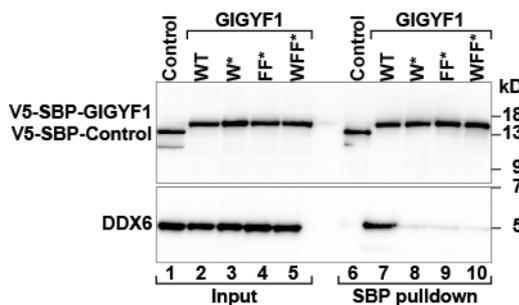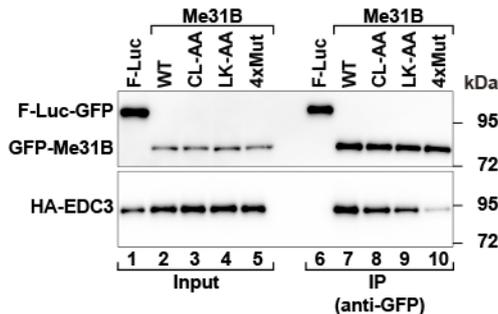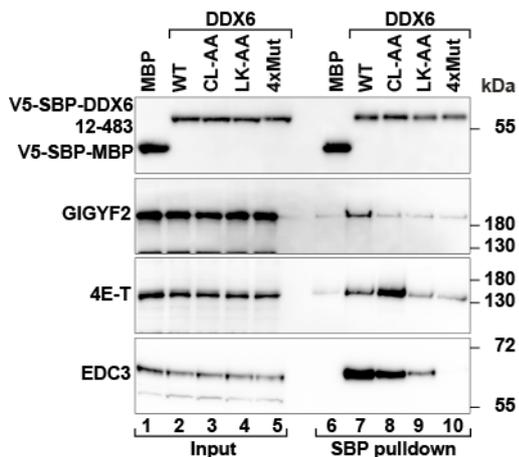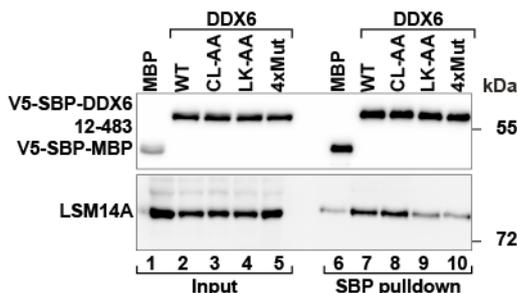
Supplemental Figure S6. The RecA2 domain of Me31B is an interaction hotspot for different proteins. Related to Fig. 3 and 4.

(**A**) Streptavidin-based pulldown assays showing the association of SBP-V5-*Hs* GIGYF1, [WT or the indicated mutants (W*=W294A, FF*=306A, F312A and WFF*; Table S2)] and DDX6. V5-SBP-MBP-F-Luc-EGFP served as a negative control. The input (1.25% for the V5-SBP-proteins; 0.5% for DDX6) and bound fractions (8% for the V5-SBP-proteins; 30% for DDX6) were analyzed by western blotting using the indicated antibodies.

(**B**) Western blot showing the interaction between GFP-*Dm* Me31B (WT or the indicated mutants) and HA-*Dm* EDC3 in Schneider S2 cells. The proteins were immunoprecipitated with anti-GFP antibodies. GFP-F-Luc served as a negative control. The inputs (3% for the GFP-proteins and 1% for HA-EDC3) and immunoprecipitates (15% for the GFP-proteins and 10% for HA-EDC3) were analyzed by western blotting using anti-GFP and anti-HA antibodies.

(***C, D***) The interaction between V5-SBP-*Hs* DDX6 (WT or the indicated mutants) and endogenous GIGYF2, 4E-T, EDC3 (B) and LSM14A (C) was analyzed in HEK293T cells using streptavidin-based pulldowns. The input (1.25% for the V5-SBP-proteins, 0.5% for GIGYF2 and 4E-T, 1.5% for EDC3 and 0.25% for LSM14A) and bound fractions (5% for the V5-SBP-proteins, 30% for GIGYF2, 4E-T, EDC3 and LSM14A) were analyzed by western blotting using the indicated antibodies.
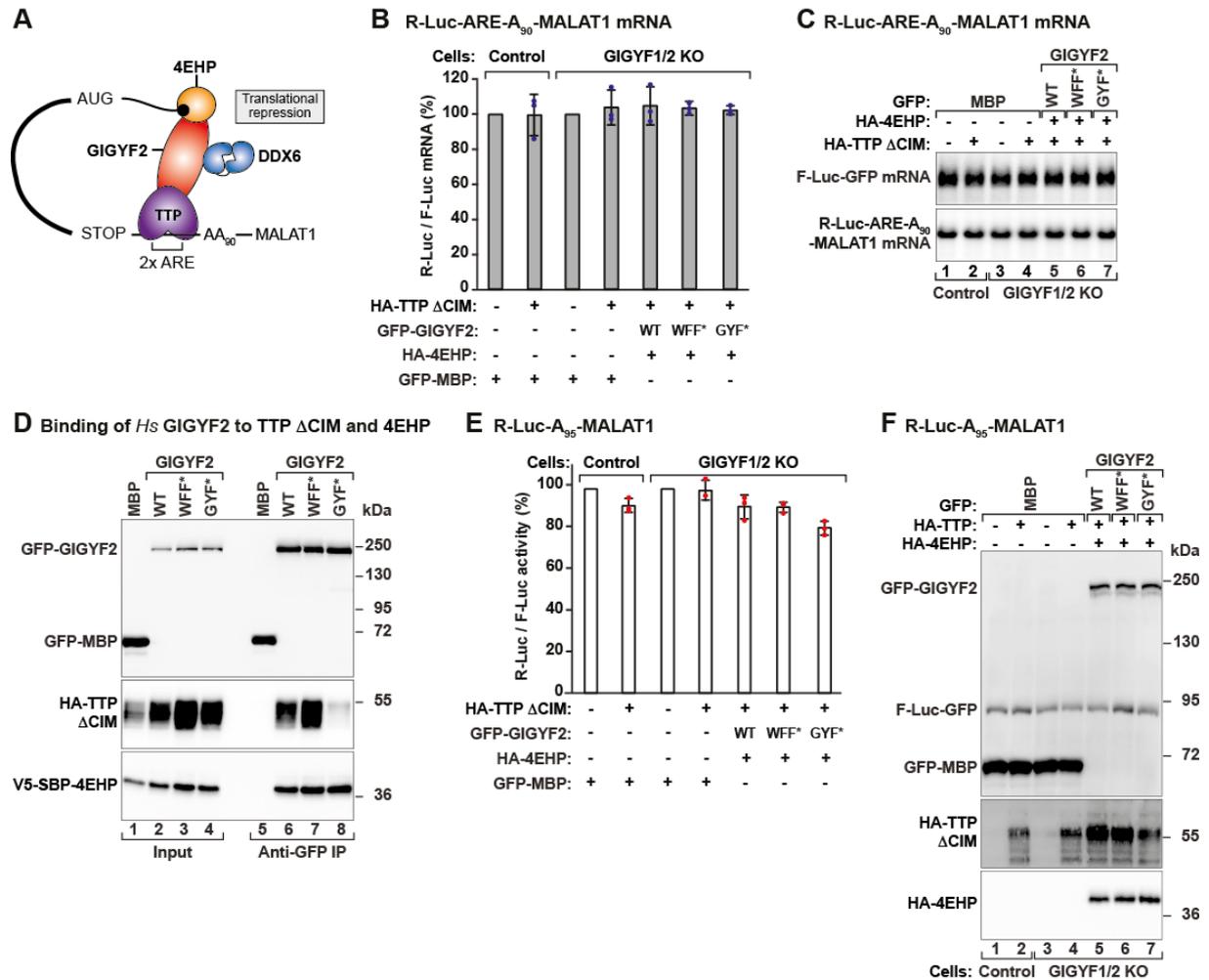
Supplemental Figure S7. Recruitment of DDX6 by *Hs* GIGYF2 contributes to TTP-mediated translational repression of an ARE-containing mRNA reporter. Related to Fig. 4.

(*A*) Schematic representation of the ARE-mRNA and protein complex used in the complementation assay. 4EHP is the cap-binding protein. GIGYF2 binds to 4EHP, DDX6 and TTP. The latter recognizes the two AREs in the 3´ UTR of the mRNA.

(*B, C*) Control or GIGYF1/2-null HEK293T cells (KO) were transfected with the R-Luc-ARE-A_{90}-MALAT1 reporter and plasmids expressing WT or the indicated GIGYF2 mutants, HA-4EHP and a TTP protein lacking the binding site for NOT1 [ΔCIM; (Fabian et al. 2013)]. An F-Luc-GFP reporter served as a transfection control. R-Luc mRNA levels (B) were normalized to that of the F-Luc transfection control and set to 100% in the absence of TTP for each cell line. Bars represent the mean values, error bars represent standard deviations and the blue dots

represent the individual points from three independent experiments. Northern blot analysis of representative RNA samples corresponding to the experiment shown in (B) and in Fig. 4A is depicted in panel C.

(**D**) Immunoprecipitation assay in HEK293T cells depicting the interaction of GFP-*Hs* GIGYF2, WT or the indicated mutants, with HA-TTP ΔCIM or V5-SBP-4EHP. GFP-MBP served as a negative control. Inputs (1% for the V5-SBP-4EHP and 1.25% for GFP-proteins and HA-TTP ΔCIM) and immunoprecipitates (20%) were analyzed by western blot using anti-HA, anti-GFP and anti-V5 antibodies.

(**E, F**) Control or GIGYF1/2-null HEK293T cells (KO) were transfected with the R-Luc-A$_{95}$-MALAT1 reporter and the other plasmids described in B. R-Luc activity (E) was normalized to that of the F-Luc transfection control and set to 100% in the absence of TTP for each cell line. Bars represent the mean values, error bars represent standard deviations and the red dots represent the individual points from three independent experiments. The western blot analysis with the expression of the proteins used in the assay is depicted in panel F.

Table S1. Data collection and refinement statistics

	Dm Me31B–GIGYF complex	*Hs* DDX6-EDC3 complex
Space group	P2₁2₁2₁	C2
Unit Cell		
Dimensions (Å)		
a, b, c	37.4, 42.0, 122.7	172.4, 47.9, 65.8
Angles (°)		
α, β, γ	90, 90, 90	90.0, 96.3, 90.0
Data collection		
Wavelength (Å)	0.999	0.978
Resolution (Å)	42.0-2.40 (2.46-2.40)	42.8-2.21 (2.25-2.21)
R_{sym}	0.188 (1.06)	0.104 (0.467)
Mean $I/\sigma I$	8.9 (2.1)	9.6 (2.2)
Completeness (%)	95.8 (88.5)	95.6 (71.2)
Multiplicity	10.4 (9.7)	2.9 (1.8)
Refinement		
Resolution (Å)	39.7-2.4	25.3-2.21
No. reflections	6800	25943
R_{work}/R_{free}	0.212/0.264	0.181/0.227
No. atoms	1638	3769
Protein	1574	3396
Ligand/ion	28	115
Water	36	258
B-factors (Å²)	26.5	39.9
Protein	26.4	38.8
Ligand/ion	37.0	69.8
Water	22.7	41.2
Ramachandran Plot		
Favored (%)	97.4	96.8
Disallowed (%)	0	0
Root-Mean-Square Deviation		
Bond lengths (Å)	0.002	0.010
Bond angles (°)	0.380	1.060

Values in parentheses are for highest resolution shell.

Supplemental Table S2. Mutants and constructs used in this study

Protein	Name of the construct	Fragments / mutations	Binding site / motif
Hs **GIGYF1** O75420	GIGYF1	Full length	
	N-term	1-671	N-terminal fragment
	N-term ΔMBM	1-671 Δ286-316	Deletion of MBM
	MBM	286-316	MBM
	W*	W294A	W mutant
	FF*	F306A, F312A	FDF mutant
	WFF*	W294A, F306A, F312A	W+FDF mutant
Hs **GIGYF2** (isoform 1) Q6Y7W6-1	GIGYF2	Full length	
	N-term	1-718	N-terminal fragment
	N-term ΔMBM	N-ter Δ280-310	Deletion of MBM
	MBM	280-310	MBM
	W*	W288A	W mutant
	FF*	F300A, F306A	FDF mutant
	WFF*	W288A, F300A, F306A	W+FDF mutant
	GYF*	Y538A, F549A, W557A, F563A	GYF domain mutant
Dm **GIGYF** (isoform 1) Q7KQM6	GIGYF	Full length	
	GIGYF ΔMBM	Full length Δ341-369	Deletion of MBM
	N-term	1-640	N-terminal fragment
	N-term ΔMBM	1-640 Δ341-369	Deletion of MBM
	MBM	341-369	MBM
	W*	W349A	W mutant
	FF*	F361A, F367A	FDF mutant
	WFF*	W349A, F361A, F367A	W+FDF mutant
Hs **DDX6** (P26196; NM_004397.5)	DDX6	12-483	
	RecA2 Δ472-483	303-472	RecA2 domain Δ472-483
	CL-AA	12-483 C324A, L328A	Mutant FDF pocket
	LK-AA	12-483 L349A, K353A	Mutant W pocket
	4xMut	12-483 C324A, L328A, L349A, K353A	Mutant FDF+W pockets
	RecA1	12-306	RecA1 domain
	RecA2	307-483	RecA2 domain
Dm **Me31B** (P23128)	DDX6	Full length	
	RecA1	1-267	RecA1 domain
	RecA2	268-459	RecA2 domain
	RecA2 Δ432-459	264-431	RecA2 domain Δ432-459
	CL-AA	C285A, L289A	Mutant FDF pocket
	LK-AA	L310A, K314A	Mutant W pocket
	4xMut	C285A, L289A, L310A, K314A	Mutant FDF+W pockets
Hs **TTP** P26651	TTP ΔCIM	1-313	Δ314-326, deletion of the CNOT1 interacting motif (CIM)
Hs **4EHP** (isoform 1) O60573-1	4EHP	Full length	

Supplemental Table S3. Antibodies used in this study

Antibody	Source	Catalog Number	Dilution	Monoclonal/ Polyclonal
Anti-HA-HRP (Western blot)	Roche	12 013 819 001	1:5,000	Monoclonal
Anti-HA (Immunoprecipitation)	Covance	MMS-101P	1:1,000	Monoclonal
Anti-GFP	In house		IP	Rabbit polyclonal
Anti-GFP	Roche	11814460001	1:2,000	Monoclonal
Anti-rabbit-HRP	GE Healthcare	NA934V	1:10,000	Polyclonal
Anti-mouse-HRP	GE Healthcare	RPN4201	1:10,000	Polyclonal
Anti-V5	QED Bioscience Inc.	18870	1:5,000	Rabbit polyclonal
Anti-V5	LSBio LifeSpan BioSciences, Inc.	LS-C57305	1:5,000	Monoclonal
Anti- *Dm* Me31B	In house		1:3,000	Rabbit polyclonal
Anti- *Dm* HPat	In house		1:3,000	Rabbit polyclonal
Anti-*Dm* 4E-T	Kind gift from Paul Lasko		1:1,000	Rabbit polyclonal
Anti-*Hs* GYF1	Bethyl laboratories	A304-132A-M	1:1,000	Rabbit polyclonal
Anti-*Hs* GYF2	Bethyl laboratories	A303-731A	1:1,000	Rabbit polyclonal
Anti-*Hs* PatL1	Bethyl laboratories	A303-482 A-M	1:1,000	Rabbit polyclonal
Anti-*Hs* 4E-T	Abcam	ab95030	1:2,000	Rabbit polyclonal
Anti-*Hs* DDX6	Bethyl Laboratories	A300-461A	1:3,000	Rabbit polyclonal
Anti-*Hs* EDC3	Abcam	Ab57780	1:1,000	Monoclonal
Anti-*Hs* LSM14	Abcam	Ab123566	1:1,000	Rabbit polyclonal

References

Adams PD, Afonine PV, Bunkoczi G, Chen VB, Echols N, Headd JJ, Hung LW, Jain S, Kapral GJ, Grosse Kunstleve RW et al. 2011. The Phenix software for automated determination of macromolecular structures. *Methods* **55**: 94-106.

Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH, Adams PD. 2012. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* **68**: 352-367.

Brandmann T, Fakim H, Padamsi Z, Youn JY, Gingras AC, Fabian MR, Jinek M. 2018. Molecular architecture of LSM14 interactions involved in the assembly of mRNA silencing complexes. *EMBO J* **37**.

Chen VB, Arendall WB, 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**: 12-21.

Diebold ML, Fribourg S, Koch M, Metzger T, Romier C. 2011. Deciphering correct strategies for multiprotein complex assembly by co-expression: application to complexes as large as the histone octamer. *J Struct Biol* **175**: 178-188.

Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**: 486-501.

Fabian MR, Frank F, Rouya C, Siddiqui N, Lai WS, Karetnikov A, Blackshear PJ, Nagar B, Sonenberg N. 2013. Structural basis for the recruitment of the human CCR4-NOT deadenylase complex by tristetraprolin. *Nat Struct Mol Biol* **20**: 735-739.

Kabsch W. 2010. Xds. *Acta Crystallogr D Biol Crystallogr* **66**: 125-132.

Kuzuoglu-Ozturk D, Bhandari D, Huntzinger E, Fauser M, Helms S, Izaurralde E. 2016. miRISC and the CCR4-NOT complex silence mRNA targets independently of 43S ribosomal scanning. *EMBO J* **35**: 1186-1203.

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. 2007. Phaser crystallographic software. *J Appl Crystallogr* **40**: 658-674.

Ozgur S, Basquin J, Kamenska A, Filipowicz W, Standart N, Conti E. 2015. Structure of a Human 4E-T/DDX6/CNOT1 Complex Reveals the Different Interplay of DDX6-Binding Proteins with the CCR4-NOT Complex. *Cell Rep* **13**: 703-711.

Peter D, Weber R, Sandmeir F, Wohlbold L, Helms S, Bawankar P, Valkov E, Igreja C, Izaurralde E. 2017. GIGYF1/2 proteins use auxiliary sequences to selectively bind to 4EHP and repress target mRNA expression. *Genes Dev* **31**: 1147-1161.

Sharif H, Ozgur S, Sharma K, Basquin C, Urlaub H, Conti E. 2013. Structural analysis of the yeast Dhh1-Pat1 complex reveals how Dhh1 engages Pat1, Edc3 and RNA in mutually exclusive interactions. *Nucleic Acids Res* **41**: 8377-8390.

Smart OS, Womack TO, Flensburg C, Keller P, Paciorek W, Sharff A, Vonrhein C, Bricogne G. 2012. Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. *Acta Crystallogr D Biol Crystallogr* **68**: 368-380.

Terwilliger TC, Grosse-Kunstleve RW, Afonine PV, Moriarty NW, Zwart PH, Hung LW, Read RJ, Adams PD. 2008. Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr D Biol Crystallogr* **64**: 61-69.

Tritschler F, Braun JE, Eulalio A, Truffault V, Izaurralde E, Weichenrieder O. 2009. Structural basis for the mutually exclusive anchoring of P body components EDC3 and Tral to the DEAD box protein DDX6/Me31B. *Mol Cell* **33**: 661-668.

Tritschler F, Eulalio A, Truffault V, Hartmann MD, Helms S, Schmidt S, Coles M, Izaurralde E, Weichenrieder O. 2007. A divergent Sm fold in EDC3 proteins mediates DCP1 binding and P-body targeting. *Mol Cell Biol* **27**: 8600-8611.

Vonrhein C, Flensburg C, Keller P, Sharff A, Smart O, Paciorek W, Womack T, Bricogne G. 2011. Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr D Biol Crystallogr* **67**: 293-302.

GIGYF1/2 proteins use auxiliary sequences to selectively bind to 4EHP and repress target mRNA expression

Daniel Peter, Ramona Weber, Felix Sandmeir, Lara Wohlbold, Sigrun Helms, Praveen Bawankar, Eugene Valkov, Cátia Igreja, and Elisa Izaurralde

Department of Biochemistry, Max Planck Institute for Developmental Biology, 72076 Tübingen, Germany

The eIF4E homologous protein (4EHP) is thought to repress translation by competing with eIF4E for binding to the 5′ cap structure of specific mRNAs to which it is recruited through interactions with various proteins, including the GRB10-interacting GYF (glycine–tyrosine–phenylalanine domain) proteins 1 and 2 (GIGYF1/2). Despite its similarity to eIF4E, 4EHP does not interact with eIF4G and therefore fails to initiate translation. In contrast to eIF4G, GIGYF1/2 bind selectively to 4EHP but not eIF4E. Here, we present crystal structures of the 4EHP-binding regions of GIGYF1 and GIGYF2 in complex with 4EHP, which reveal the molecular basis for the selectivity of the GIGYF1/2 proteins for 4EHP. Complementation assays in a GIGYF1/2-null cell line using structure-based mutants indicate that 4EHP requires interactions with GIGYF1/2 to down-regulate target mRNA expression. Our studies provide structural insights into the assembly of 4EHP–GIGYF1/2 repressor complexes and reveal that rather than merely facilitating 4EHP recruitment to transcripts, GIGYF1/2 proteins are required for repressive activity.

The initiation of cap-dependent translation involves a series of sequential steps that start with the assembly of the eIF4F on the mRNA 5′ cap structure (Jackson et al. 2010). The eIF4F complex consists of the cap-binding protein eIF4E, the RNA helicase eIF4A, and the scaffold protein eIF4G, which bridges the interaction between the other two subunits in the complex. eIF4G also interacts with eIF3 and mediates the recruitment of the preinitiation complex (PIC; comprising a 40S ribosomal subunit and associated factors) to the mRNA to initiate translation (Jackson et al. 2010).

The assembly of the eIF4F complex is regulated by multiple mechanisms. One major mechanism involves a broad class of eIF4E-binding proteins (4E-BPs) that compete with eIF4G for binding to eIF4E, thereby inhibiting translation initiation (Mader et al. 1995; Marcotrigiano et al. 1999). eIF4G and the 4E-BPs share a conserved, canonical (C) 4E-binding motif with the sequence YX$_4$LΦ (where Y, X, L, and Φ represent Tyr, any amino acid, Leu, and a hydrophobic residue, respectively), which binds to the dorsal surface of eIF4E opposite to the cap-binding pocket (Matsuo et al. 1997; Marcotrigiano et al. 1999; Gross et al. 2003). Both eIF4G and the 4E-

BPs also contain variable noncanonical (NC) 4E-binding motifs that bind to an eIF4E hydrophobic lateral surface, increasing the affinity of the interaction (Kinkelin et al. 2012; Paku et al. 2012; Lukhele et al. 2013; Igreja et al. 2014; Peter et al. 2015a,b; Sekiyama et al. 2015; Grüner et al. 2016). Because eIF4G and 4E-BPs bind to the same surfaces on eIF4E, their binding is mutually exclusive, resulting in translation activation and inhibition, respectively.

An alternative mechanism that inhibits the recruitment of the eIF4F complex involves the recognition of the mRNA 5′ cap by another member of the eIF4E family, the 4E homologous protein (4EHP; also known as eIF4E2) (Rom et al. 1998; Joshi et al. 2004). Despite its sequence and structural similarity to eIF4E (Supplemental Fig. S1A; Rosettani et al. 2007), 4EHP does not interact with eIF4G and thus fails to initiate translation (Rom et al. 1998; Joshi et al. 2004; Hernandez et al. 2005).

4EHP is recruited to specific mRNAs by RNA-binding proteins and thus acts as a sequence-specific rather than a general translational repressor. For example, *Drosophila melanogaster* 4EHP is specifically recruited to and represses translation of *caudal* and *hunchback* mRNAs through interactions with the RNA-binding proteins

Bicoid (Bcd) and Brain tumor (Brat), respectively (Cho et al. 2005, 2006). Mammalian 4EHP has been implicated in post-transcriptional mRNA regulation through its interaction with the nucleocytoplasmic shuttling protein 4E-T (eIF4E transporter), which is a component of P bodies (Kubacka et al. 2013). In mouse oocytes, the homeobox protein Prep1 recruits 4EHP to inhibit the translation of *Hoxb4* mRNA (Villaescusa et al. 2009). 4EHP also forms a translational repressor complex with GIGYF2 (GRB10-interacting GYF [glycine–tyrosine–phenylalanine domain] protein 2 [GYF2]), a protein involved in the insulin signaling pathway (Giovannone et al. 2009; Morita et al. 2012). This repressor complex is recruited to specific mRNAs by the zinc finger protein ZNF598 (Morita et al. 2012). Alternatively, the 4EHP–GYF2 complex is recruited to mRNAs containing AU-rich elements (AREs) in their 3′ untranslated regions (UTRs) by tristetraprolin (TTP) (Tao and Gao 2015; Fu et al. 2016). Thus, through its association with diverse binding partners, 4EHP regulates the translation of mRNAs involved in a broad range of biological process, and disruption of its expression results in perinatal lethality in mice (Morita et al. 2012).

Current models suggest that 4EHP-binding proteins (4EHP-BPs) interact with 4EHP through a canonical 4EHP-binding motif with the sequence $YXYX_4L\Phi$ that is present in GYF1/2 proteins (Fig. 1A; Supplemental Fig. S1B). Although this motif consists of a canonical 4E-binding motif extended by only two N-terminal residues (YX) (Cho et al. 2005; Morita et al. 2012), GYF2 does not bind to eIF4E in vivo (Morita et al. 2012). Conversely, eIF4G, which contains a canonical motif, binds to eIF4E but not 4EHP (Rom et al. 1998; Joshi et al. 2004; Hernandez et al. 2005). In contrast, some 4E-BPs, such as *Homo sapiens* 4E-BP1–3 and 4E-T, which lack the additional YX residues, interact with both eIF4E and 4EHP (Rom et al. 1998; Rosettani et al. 2007; Kubacka et al. 2013). This suggests that the canonical motif is unlikely to be the sole specificity determinant for 4EHP or eIF4E and that the structural basis for this molecular discrimination is unknown. Additionally, structural insights into 4EHP complexes are limited to a complex with the 4E-BP1 canonical motif, which binds to 4EHP in vitro but not in vivo (Rom et al. 1998; Rosettani et al. 2007).

To obtain molecular insights into the assembly of 4EHP repressor complexes, we determined the crystal structures of 4EHP in complex with the binding regions of human 4E-BP1, GYF1, and GYF2. The structures reveal that, in addition to the known canonical motifs that bind to the dorsal surface of 4EHP, 4E-BP1 and GYF1/2 also make contacts with the lateral surface of 4EHP using noncanonical motifs, indicating that lateral binding is a common feature observed in both 4EHP and eIF4E complexes. Remarkably, GYF1/2 proteins, but not 4E-BP1, contain C-terminal auxiliary sequences (A) that extend the interface, contacting 4EHP residues that are not conserved in eIF4E. Our studies reveal the molecular basis for the selectivity of GYF1/2 proteins for 4EHP over eIF4E and provide mechanistic insights into the regulation of cap-dependent translation initiation by 4EHP-repressive complexes.

Results

GYF1/2 proteins bind to the dorsal and lateral surfaces of 4EHP

Given that the canonical motif in 4E-BP1 binds to the dorsal surface of 4EHP in a way similar to how it binds to eIF4E in vitro (Tee et al. 2004; Rosettani et al. 2007), we initially investigated whether the noncanonical sequences in 4E-BP1 could also bind to the lateral surface of 4EHP, as observed in the eIF4E–4E-BP1 complex (Igreja et al. 2014; Peter et al. 2015a). We substituted residues I85 and M101 on the lateral surface of 4EHP with alanine (IM-AA mutant) (Supplemental Fig. S1A; Supplemental Table S1). These residues are structurally equivalent to eIF4E residues I63 and I79, which are required for the noncanonical motifs in 4E-BPs to bind to the lateral surface of eIF4E (Igreja et al. 2014; Peter et al. 2015a). As a control, a 4EHP mutant carrying the W95A substitution on the dorsal surface (mutant W-A [Supplemental Table S1], corresponding to the eIF4E W73A mutant) was designed to disrupt canonical motif binding.

The substitutions in either the dorsal or the lateral surface of V5-SBP-tagged 4EHP disrupted interactions with GFP-tagged 4E-BP1 in HEK293T cells (Fig. 1B, lanes 7,8). Conversely, substitutions in either the canonical (C* mutant) or noncanonical (NC* mutant) motif of 4E-BP1 abolished its interaction with 4EHP (Supplemental Fig. S2A). These results indicate that the interactions between the noncanonical 4E-BP1 sequences and the 4EHP lateral surface are also critical for complex stability.

The interaction between GYF2 and 4EHP requires a canonical 4EHP-binding motif at the N terminus of the protein ($YXYX_4L\Phi$) (Fig. 1A; Supplemental Fig. S1B; Morita et al. 2012). Accordingly, alanine substitutions of the four conserved residues (Y, Y, L, and Φ) in this motif abolished full-length GYF1/2 binding to V5-SBP-4EHP in human cells (Supplemental Fig. S2B,C). However, it is not known whether GYF proteins contain noncanonical sequences. We therefore examined the effects of substitutions at the dorsal and lateral surfaces of 4EHP on interactions with GYF1/2 proteins. The substitutions at either surface reduced but did not abolish 4EHP binding to endogenous GYF2 in human cells (Fig. 1C, lanes 8,9 vs. 7). The interaction was abolished only when the substitutions on both surfaces were combined (Fig. 1C, lane 10). In contrast, the dorsal and lateral substitutions disrupted binding with endogenous GYF1 (Supplemental Fig. S2D). Thus, 4EHP uses its dorsal and lateral surfaces to interact with the GYF1/2 proteins, suggesting that these proteins also contain noncanonical motifs.

GYF1/2 proteins contain noncanonical and auxiliary 4EHP-binding sequences

The noncanonical motifs in 4E-BPs are typically located 12–30 residues C-terminal to the canonical motifs and contain hydrophobic residues (Igreja et al. 2014; Peter et al. 2015a,b). During a search for potential noncanonical motifs in GYF1/2 proteins, we identified a hydrophobic
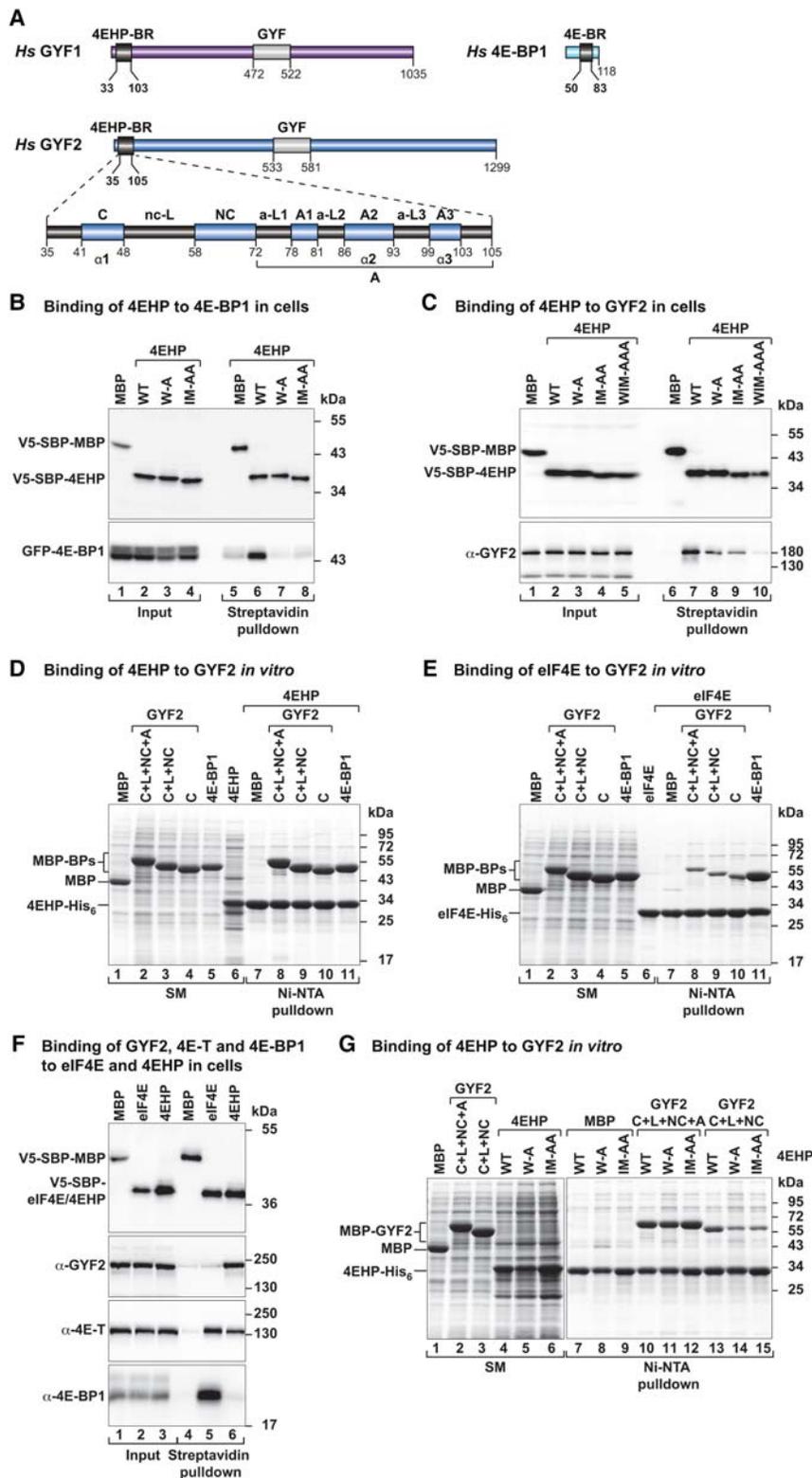
Figure 1. GYF1/2 proteins use canonical, noncanonical, and auxiliary sequences to bind to 4EHP. (*A*) GYF1/2 proteins contain a central GYF domain and an N-terminal 4EHP-binding region (4EHP-BR). The 4EHP-BR includes canonical, noncanonical, and auxiliary motifs (A1–3) connected by linker sequences (nc-L and auxiliary linkers 1–3 [a-L1–3]). The 4E-binding region (4E-BR) of 4E-BP1 contains canonical and noncanonical motifs. (*B*,*C*) Western blots showing the interaction between V5-SBP-4EHP (wild type or the indicated mutants) and GFP-4E-BP1 (full-length) or endogenous GYF2. The proteins were pulled down using strep-tavidin-coated beads. V5-SBP-MBP (malt-ose-binding protein) served as negative control. The inputs (1.5% for the V5-tagged proteins and 1% for the GFP-tagged pro-teins) and bound fractions (3%–5% for the V5-tagged proteins and 20% for GYF2 and GFP-4E-BP1) were analyzed by Western blot-ting using anti-V5, anti-GFP, and anti-GYF2 antibodies. (*D*,*E*) Ni-NTA pull-down assays showing the interactions between GYF2 fragments (C+L+NC+A, C+L+NC, and C) and 4EHP-His$_6$ (M1–F234) (*D*) or eIF4E-His$_6$ (*E*). 4E-BP1 and MBP served as positive and negative controls, respectively. The GYF2 and 4E-BP1 peptides contain an N-terminal MBP tag and a C-terminal GB1 tag. The starting material (SM; 1.3% for MBPs and 6% for 4EHP and purified eIF4E) and bound fractions (7%–10%) were analyzed by SDS-PAGE followed by Coomassie blue staining. (*F*) The interaction between V5-SBP-eIF4E or 4EHP proteins and endogenous GYF2, 4E-T, and 4E-BP1 was analyzed in HEK293T cell lysates using streptavidin pull-downs. The input (1% for 4E-BP1 and 4E-T and 1.5% for V5-SBP-tagged proteins and GYF2) and bound fractions (20% for 4E-BP1 and 4ET, 30% for GYF2, and 5% for the V5-SBP-tagged proteins) were analyzed by Western blotting using the indicated antibodies. (*G*) Ni-NTA pull-down assay showing the inter-action between 4EHP (M1–F234, wild type, or the indicated mutants) and GYF2 frag-ments. MBP served as a negative control. Samples were analyzed as described in *D*. The starting material (2% for the MBP-tagged proteins and 4%–12% for the 4EHP proteins) and bound fractions (10%) were an-alyzed by SDS-PAGE followed by Coomas-sie blue staining.

motif 12 residues C-terminal to the canonical motif con-taining a conserved Phe that we termed the noncanonical motif (Fig. 1A; Supplemental Fig. S1B). The following 30 residues (which we termed auxiliary sequences) are also well conserved and contain several short motifs that

may potentially interact with 4EHP, as was observed pre-viously in the *D. melanogaster* protein Mextli in complex with eIF4E (Peter et al. 2015b).

To more precisely define the GYF1/2 sequences that in-teract with 4EHP, we performed in vitro pull-down assays

using recombinant proteins expressed in *Escherichia coli*. In particular, we tested the binding of 4EHP to GYF1/2 fragments comprising the canonical motif and the noncanonical sequences (i.e., noncanonical linker and motif; L+NC) with and without the auxiliary sequences (fragments C+L+NC and C+L+NC+A) (Supplemental Table S1) as well as a fragment comprising only the canonical motif. 4EHP expressed with a hexahistidine (His$_6$) tag pulled down all three GYF1/2 fragments expressed with an N-terminal maltose-binding protein (MBP) tag (Fig. 1D; Supplemental Fig. S2E), indicating that the canonical motifs are sufficient for 4EHP binding in vitro. The eIF4E-binding region of human 4E-BP1 interacted with 4EHP to a similar extent (Fig. 1D; Supplemental Fig. S2E, lane 11). However, the GYF1/2 fragments bound to eIF4E much less efficiently than 4E-BP1 (Fig. 1E; Supplemental Fig. S2F), indicating that GYF1/2 proteins exhibit selectivity for 4EHP over eIF4E in vitro in the absence of cellular factors.

In contrast to the results obtained in vitro, endogenous GYF1/2 proteins interacted with 4EHP but not eIF4E in cell lysates, as reported previously (Fig. 1F; Supplemental Fig. S2G; Rom et al. 1998; Morita et al. 2012). Similar results were obtained with overexpressed GYF1/2 (Supplemental Fig. S2H,I), suggesting that although GYF1/2 proteins can bind to eIF4E in vitro, their binding affinity may be too low to compete with other 4E-BPs present in cell lysates for binding to eIF4E. As expected, the 4E-T protein interacted with both eIF4E and 4EHP (Fig. 1F; Kubacka et al. 2013). In contrast, although endogenous 4E-BP1 did not bind to 4EHP in cell lysates (Fig. 1F; Supplemental Fig. S2G; Rom et al. 1998), it did bind when it was overexpressed (Fig. 1B). Thus, in cell lysates, the selectivity of GYF1/2 proteins and 4E-BP1 for 4EHP and eIF4E, respectively, is likely to be determined by their affinities and concentrations relative to those of other competing proteins.

GYF1/2 auxiliary sequences increase affinity for 4EHP

Although the GYF1/2 fragments with and without auxiliary sequences associated with wild-type 4EHP in in vitro pull-down assays (Fig. 1G; Supplemental Fig. S2J, lanes 10,13), they were differentially affected by mutations on the 4EHP dorsal and lateral surfaces. The GYF1/2 fragments, including the auxiliary sequences, were insensitive to the mutations (Fig. 1G; Supplemental Fig. S2J, lanes 10–12). In contrast, the binding of the fragments lacking the auxiliary sequences was reduced or abolished by the mutations (Fig. 1G; Supplemental Fig. S2J, lanes 14,15). Thus, the GYF1/2 auxiliary sequences contribute to the stability of the complexes with 4EHP and compensate for the destabilizing effects of the mutations in the dorsal and lateral 4EHP surfaces.

To evaluate the thermodynamic contribution of the GYF1/2 canonical, noncanonical, and auxiliary sequences to the affinity for 4EHP, we performed isothermal titration calorimetry (ITC) experiments. GYF1/2 peptides containing only the canonical motif exhibited dissociation constants (K_Ds) for 4EHP in the high nanomolar range (360 nM^{GYF1} ± 120 nM^{GYF1} and 290 nM^{GYF2} ±

160 nM^{GYF2}) (Supplemental Fig. S3A,B; Supplemental Table S2). Addition of the noncanonical linker and motif (C+L+NC peptides) increased the affinity for 4EHP by 20-fold to 30-fold (K_Ds of 12 nM^{GYF1} ± 2 nM^{GYF1} and 14 nM^{GYF2} ± 1 nM^{GYF2}) (Supplemental Fig. S3C,D; Supplemental Table S2), confirming the importance of the noncanonical sequences for complex formation. Importantly, addition of the auxiliary sequences increased the affinity even further by 30-fold to 40-fold (C+L+NC+A peptides; K_Ds of 0.4 nM^{GYF1} ± 0.2 nM^{GYF1} and 0.3 nM^{GYF2} ± 0.1 nM^{GYF2}) (Supplemental Fig. S3E,F; Supplemental Table S2) relative to that of the peptides lacking the auxiliary sequences.

Collectively, the affinity measurements indicate that the GYF1/2 auxiliary sequences contribute substantially to the affinity for 4EHP by further stabilizing the interactions mediated by the canonical motif and the noncanonical sequences.

The overall architecture of the 4EHP–GYF1/2 and 4EHP–4E-BP1 complexes

To understand the structural basis of complex formation and selectivity, we determined the crystal structures of human 4EHP bound to GYF1 and GYF2 fragments (C+L+NC+A peptides; residues 33–103^{GYF1} and 35–105^{GYF2}) at 2.9 Å and 2.3 Å resolution, respectively (Fig. 2A–E; Table 1; Supplemental Fig. S4A,B). We also determined the crystal structure of 4EHP bound to a peptide comprising the 4E-BP1 canonical motif and noncanonical sequences (C+L+NC) and the corresponding peptide in GYF2 at 1.9 Å and 2.0 Å resolution, respectively (Fig. 2F–H; Table 1; Supplemental Fig. S4C,D).

The cap-binding protein 4EHP adopts a eIF4E-like fold (Fig. 2C–G; Rosettani et al. 2007), and its dorsal and lateral surfaces are very similar to those of eIF4E at the structural and sequence levels (Supplemental Figs. S1A, S4E,F). Upon 4E-BP1 or GYF1/2 binding, no major conformational changes are observed between the 4EHP structures presented in this study compared with a previously determined structure of 4EHP bound to the 4E-BP1 canonical motif (Rosettani et al. 2007).

The 4E-BP1 and GYF1/2 canonical motifs adopt a helical conformation on the dorsal surface of 4EHP, whereas the noncanonical sequences bind to the lateral surface of 4EHP using a binding mode similar to that described for 4E-BP1 in complex with eIF4E (Fig. 2C–J; Supplemental Fig. S4E,F; Peter et al. 2015a; Sekiyama et al. 2015). The binding mode and conformation for the GYF2 fragment comprising the canonical motif and the noncanonical sequences were not influenced by the auxiliary region, as the two GYF2 structures in complex with 4EHP are very similar across common elements irrespective of whether the auxiliary region was present (Supplemental Fig. S4G).

The distinguishing structural feature observed in the 4EHP–GYF1/2 complexes containing the auxiliary sequences is the unprecedented binding mode between these sequences and 4EHP (Fig. 2A–E). The auxiliary sequences extend the binding interface to 2190 Å^2 compared
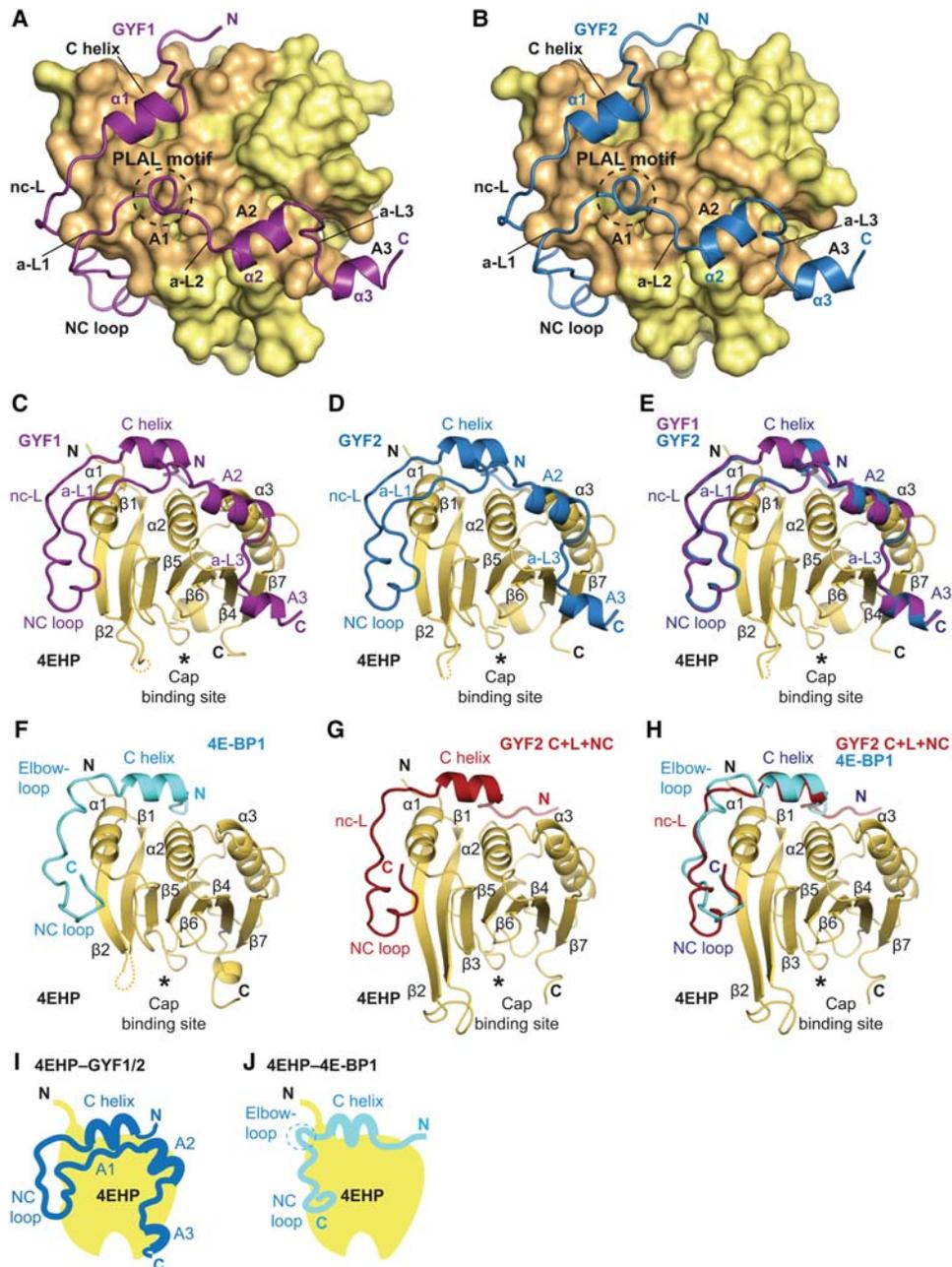
Figure 2. Overall structures of GYF1, GYF2, and 4E-BP1 bound to 4EHP. (*A,B*) Overview of the structures of 4EHP bound to GYF1/2 (C+L +NC+A) fragments. The 4EHP surface is shown in yellow, and surface residues within a radius of 4 Å of the bound GYF1 or GYF2 peptides are colored in orange. The GYF1 and GYF2 peptides are colored in purple and blue, respectively. Selected secondary structure elements in the GYF1/2 peptides are indicated. The invariant PLAL motif of GYF1/2 is circled with a dashed line. (*C,D*) Cartoon representation of the structures of 4EHP bound to GYF1/2. Selected secondary structure elements are labeled in black for 4EHP and in color for GYF1/2. (*E*) Superposition of the structures of 4EHP bound to GYF1 and GYF2. For clarity, the 4EHP molecule from the 4EHP–GYF1 complex was omitted. The structures of the complexes are very similar, and overall root mean square deviations do not exceed 0.32 Å over 227 Cα atoms. (*F*) Structure of 4EHP bound to 4E-BP1. Selected secondary structure elements are labeled in black for 4EHP and in cyan for 4E-BP1. (*G*) Structure of 4EHP bound to the GYF2 C+L+NC fragment. Selected secondary structure elements are labeled in black and red for 4EHP and GYF2, respectively. (*H*) Superposition of the structures of 4EHP bound to the 4E-BP1 and GYF2 C+L+NC peptides. For clarity, the 4EHP molecule from the 4EHP–4E-BP1 complex was omitted. (*I,J*) Schematic representations of 4EHP bound to GYF1/2 and 4E-BP1.

with 1380 Å^2 in their absence. The auxiliary sequences in GYF1/2 can be delineated into three short sequence motifs, termed auxiliary motifs 1–3 (A1–3); connect to the

noncanonical motif by the auxiliary linker 1 (a-L1); and are interconnected by a-L2 and a-L3 (Figs. 1A, 2A,B; Supplemental Fig. S1B).

Table 1. *Data collection and refinement statistics*

	4EHP–GIGYF1 complex (GYF1 C+L+NC+A)	4EHP–GIGYF2 complex (GYF2 C+L+NC+A)	4EHP–GIGYF2 complex (GYF2 C+L+NC)	4EHP-4E-BP1 complex (4E-BP1 C+L+NC)
Space group	P4$_2$	P4$_1$2$_1$2	C2	P2$_1$
Unit cell				
Dimensions				
a, b, c	135.3 Å, 135.3 Å, 60.9 Å	82.6 Å, 82.6 Å, 148.5 Å	152.2 Å, 98.6 Å, 39.3 Å	38.4 Å, 83.4 Å, 70.5 Å
Angles				
α, β, γ	90°, 90°, 90°	90°, 90°, 90°	90°, 99.6°, 90°	90°, 104.3°, 90°
Data collection				
Wavelength	1.000 Å	1.000 Å	1.000 Å	0.999 Å
Resolution	47.8 Å–2.9 Å	45.9 Å–2.3 Å	44.6 Å–2.0 Å	41.7 Å–1.9 Å
R_{sym}	0.111 (0.647)	0.125 (1.14)	0.066 (0.611)	0.145 (1.14)
Mean $I/\sigma I$	10.9 (1.95)	13.2 (2.06)	12.0 (2.09)	8.5 (2.07)
Completeness	99.6% (99.9%)	99.8% (98.3%)	99.8% (99.9%)	99.7% (99.6%)
Multiplicity	3.4 (3.4)	11.2 (10.6)	5.1 (4.8)	6.6 (6.7)
Refinement				
Resolution	47.8 Å–2.9 Å	45.9 Å–2.3 Å	44.6 Å–2.0 Å	41.7 Å–1.9 Å
Number of reflections	24,694	23,497	38,550	33,898
R_{work}/R_{free}	0.204/0.254	0.205/0.242	0.198/0.233	0.226/0.251
Number of atoms	7833	3811	3536	3547
Protein	7833	3741	3386	3305
Ligand/ion	—	—	20	21
Water	—	70	130	221
B-factors	49.4 Å^2	62.0 Å^2	63.9 Å^2	27.8 Å^2
Protein	49.4 Å^2	62.3 Å^2	64.2 Å^2	27.3 Å^2
Ligand/ion	—	—	88.6 Å^2	33.5 Å^2
Water	—	46.3 Å^2	52.7 Å^2	28.7 Å^2
Ramachandran plot				
Favored	95.9%	97.3%	95.5%	97.7%
Disallowed	0%	0%	0%	0%
Root mean square deviation				
Bond lengths	0.003 Å	0.004 Å	0.011 Å	0.003 Å
Bond angles	0.494°	0.540°	1.002°	0.529°

Values in parentheses are for highest-resolution shell.
Ligands: four PO4^{3-} ions in the 4EHP–GIGYF2 (C+L+NC) complex and seven formic acid molecules in the 4EHP–4E-BP1 (C+L+NC) complex.

The GYF1/2 canonical helix stabilizes the interaction with 4EHP and the auxiliary sequences

The canonical helices in GYF1/2 and 4E-BP1 bind to the 4EHP dorsal surface through interactions analogous to those observed for the eIF4G and 4E-BP canonical motifs in complex with eIF4E (Fig. 3A,B; Supplemental Fig. S5A–D; Marcotrigiano et al. 1999; Gross et al. 2003; Kinkelin et al. 2012; Peter et al. 2015a,b; Grüner et al. 2016). The most conserved interactions are mediated by residues corresponding to LΦ in the YXYX$_4$LΦ consensus sequence (M46 and L47^{GYF1}, M48 and L49^{GYF2}, and L59 and M60^{BP1}) and the second Tyr side chain in the canonical 4EHP-binding motif (Y41^{GYF1} and Y43^{GYF2}, corresponding to Y54^{BP1}) (Fig. 3A,B; Supplemental Fig. S5A–D).

The first Tyr in the canonical 4EHP-binding motif (YXYX$_4$LΦ) was suggested to contribute to the binding specificity of 4EHP-BPs (Cho et al. 2005; Villaescusa et al. 2009). Our structural analysis does not support

such a role for this Tyr (Y39^{GYF1} and Y41^{GYF2}). Although its aromatic ring is in contact with P55^{4EHP} (Fig. 3A; Supplemental Fig. S5C,D), this interaction is not unique to 4EHP-BPs, as P55^{4EHP} forms a similar contact with I52^{BP1} at an equivalent position in the motif in the 4E-BP1 complex (IXYX$_4$LΦ^{BP1}) (Fig. 3B).

An important difference between the 4E-BP1 and GYF1/2 canonical motifs is that the latter do not possess an Arg/Lys/Gln residue at position 9 in the extended canonical motif [extended motif: YX(R/K)X$_2$LΦX$_2$(R/K/Q)]. This residue typically contributes to the interaction with eIF4E (Marcotrigiano et al. 1999; Peter et al. 2015a,b). Instead, GYF1/2 proteins contain an aromatic residue at this position (Y50^{GYF1} and F52^{GYF2}) (Fig. 3A; Supplemental Figs. S1B, S5C,D), which establishes hydrophobic contacts with W95^{4EHP} and stabilizes the GYF1/2 auxiliary motifs through intramolecular interactions with the invariant Pro residue in the PLAL motif (P76^{GYF1} and P78^{GYF2}) (see below), thus rationalizing the conservation of both residues.
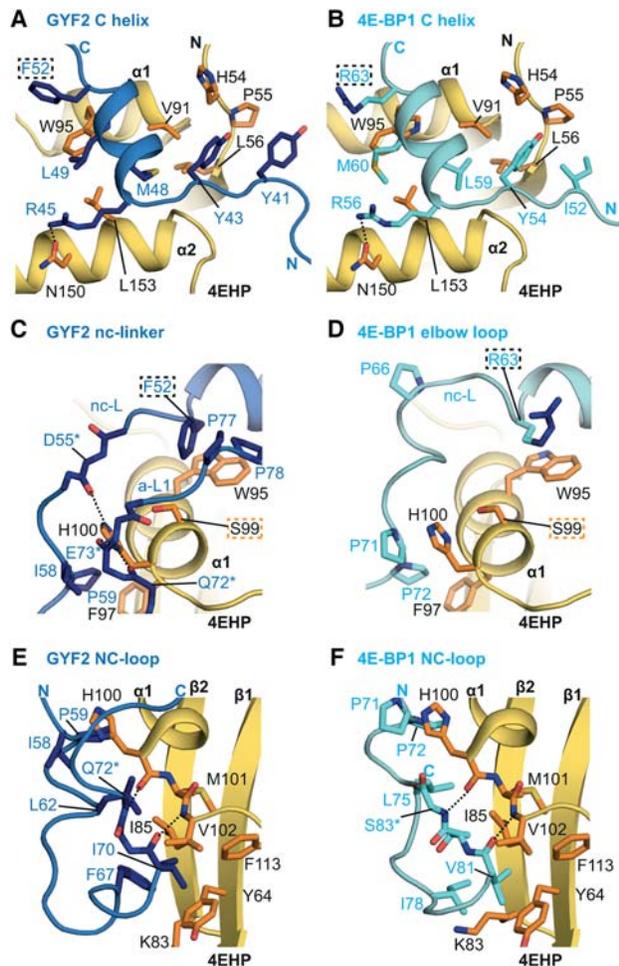
Figure 3. The interactions between the canonical and noncanonical sequences of GYF2 and 4E-BP1 with 4EHP. (*A,B*) Close-up views of the interactions between the 4EHP dorsal surface and the GYF2 and 4E-BP1 canonical helices. 4E-BP1 residue R63 is colored in dark blue after its Cγ atom and is highlighted with a black dashed box. The corresponding residues in GYF2 (F52) are also highlighted by a black dashed box. (*C,D*) Close-up views of the interaction between the 4EHP lateral surface and the GYF2 and 4E-BP1 noncanonical linkers. (*E,F*) Close-up views of the interactions between the 4EHP lateral surface and the GYF2 and 4E-BP1 noncanonical loops. Selected interface residues are shown as sticks. For clarity, all residues labeled with an asterisk are shown without their side chain.

The noncanonical linkers contribute to structural stability in the 4EHP complexes

Following the canonical helix, the noncanonical linkers orient the peptide chains to engage with the 4EHP lateral surface (Fig. 2A–G). In the structures of 4E-BPs bound to eIF4E, these linkers adopt a specific "elbow loop" conformation (Peter et al. 2015a,b). An analogous structural feature is observed in the 4EHP–4E-BP1 complex, where P66^{BP1}, P71^{BP1}, and P72^{BP1} restrict the flexibility of the backbone conformation in the elbow loop (Fig. 3D; Supplemental Fig. S5E,F). Stabilizing hydrophobic contacts, such as between the invariant H100^{4EHP} (cor-

responding to H78^{4E}) and P71^{BP1}, ensure that the overall elbow conformation is almost identical in the eIF4E- and 4EHP-bound complexes (Fig. 3D; Supplemental Fig. S5E,F).

One important difference, however, is that the N77^{4E} side chain is in hydrogen-bonding distance to the R63^{BP1} guanidium group in position 9 of the extended canonical motif and to the T68^{BP1} carbonyl oxygen in the linker region in the eIF4E–4E-BP1 complex (Supplemental Fig. S5F; Peter et al. 2015a). In the 4EHP complex, these contacts cannot be maintained by S99, and, consequently, R63^{BP1} does not contribute to complex stability (Fig. 3D; Supplemental Fig. S5E,F).

Comparison of the GYF1/2 conformations in complex with 4EHP reveals that the linker region is also arranged in a single preferred conformation (Supplemental Fig. S4H) that is distinct from the elbow loop conformation observed in the complexes with 4E-BP1 (Fig. 3, C vs. D). The invariant H100^{4EHP} plays a crucial role in anchoring the GYF1/2 linker to the 4EHP surface through van der Waals contacts with I58^{GYF2} (V56^{GYF1}), while its imidazole ring is also in hydrogen-bonding distance to the D55^{GYF2} (E53^{GYF1}) carbonyl oxygen (Fig. 3C; Supplemental Fig. S5G,H). The principal stabilizing hydrophobic interaction in the GYF1/2 linker region is between the invariant P59^{GYF2} (P57^{GYF1}) and F97^{4EHP}, which is structurally equivalent to the interaction between P72^{BP1} and F97^{4EHP} (Fig. 3C,D; Supplemental Fig. S5E–H).

Noncanonical loops mediate conserved contacts at the 4EHP lateral surface

The noncanonical loops in 4E-BP1 and the GYF1/2 proteins engage a hydrophobic pocket on the lateral surface of 4EHP, which is lined by residues Y64, I85, and M101 (corresponding to eIF4E residues F47, I63, and I79, respectively) (Fig. 3E,F; Supplemental Fig. S5I–L). The conformation of the GYF1/2 noncanonical loops is stabilized by an extensive and conserved network of contacts across the 4EHP lateral surface. The 4E-BP1 and GYF1/2 noncanonical loops differ in conformation and align only at major contact points (Fig. 3, E vs. F).

Strikingly, a carbon–π interaction, through which the conserved aromatic residue Y64^{4EHP} (equivalent to F47^{4E}) makes contacts with I70^{GYF2} (V68^{GYF1}) to anchor this loop at the lateral surface of 4EHP (Fig. 3E; Supplemental Fig. S5K,L), is conserved in the 4EHP–4E-BP1 complex (Y64^{4EHP}–V81^{BP1}) as well as in all eIF4E–4E-BP and eIF4E–eIF4G complex structures (Peter et al. 2015a; Grüner et al. 2016), underscoring the role of this aromatic residue (Y64^{4EHP}, F47^{4E}) in positioning the noncanonical loops at the lateral surface of eIF4E proteins. The invariant F67^{GYF2} (F65^{GYF1}) is critically positioned at the sharp turn of the peptide and stabilizes this conformation via hydrophobic contacts with Y64^{4EHP}, K83^{4EHP}, and I85^{4EHP}. The C-terminal residues in the GYF1/2 and 4E-BP1 noncanonical loops (V68 and Q70^{GYF1}, I70 and Q72^{GYF2}, and V81 and S83^{BP1}) mediate similar backbone interactions with the 4EHP residues H100 and V102 (Fig. 3E,F; Supplemental Fig. S5I–L).

4EHP-specific interactions with GYF1/2 auxiliary motifs close to the cap-binding site

The a-L1 and the invariant PLAL motif at the start of the auxiliary sequences adapt to a composite surface formed between the GYF1/2 canonical helix and the 4EHP surface (Fig. 4A–D). The linker (a-L1) is fixed in position via
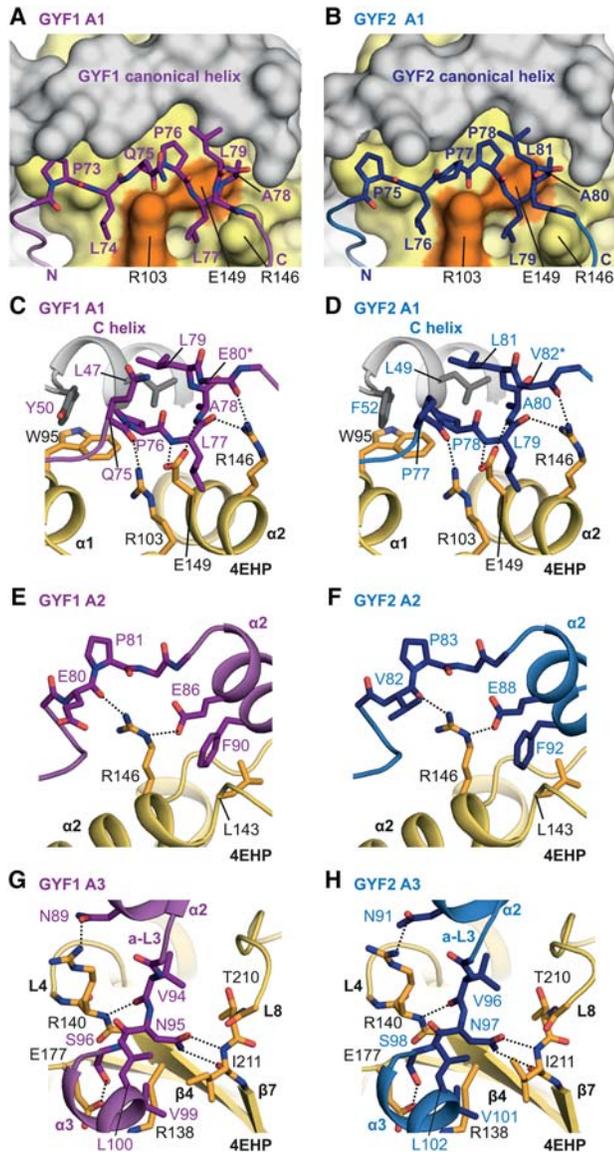

Figure 4. Interaction between the GYF1/2 auxiliary sequences and 4EHP. (A,B) Close-up view of the arrangement of the linker a-L1 and the PLAL motif (A1) in GYF1/2 proteins at the 4EHP dorsal surface. The surface of 4EHP is shown in yellow, and the surfaces of the GYF1/2 canonical helices are shown in gray. The positions of the 4EHP unique residues R103 and E149 are highlighted in orange, and selected GYF1/2 residues are shown as either purple (GYF1) or blue (GYF2) sticks. (C–H) Close-up views of the interactions between 4EHP and the GYF1/2 auxiliary sequences (A1, A2, and A3). Selected GYF1/2 residues and 4EHP interface residues are shown as sticks. The GYF1/2 canonical helices are colored in gray.

$H100^{4EHP}$, which can participate in polar contacts with the carbonyl oxygens of $Q70^{GYF1}$ or $D71^{GYF1}$ ($Q72^{GYF2}$ or $E73^{GYF2}$). The $R103^{4EHP}$ guanidinium group is in hydrogen-bonding distance to the $Q75^{GYF1}$ ($P77^{GYF2}$) carbonyl oxygen, which rationalizes the conservation of the R103 residue in 4EHP but not eIF4E. The invariant $P76^{GYF1}$ ($P78^{GYF2}$) in the PLAL motif coordinates an intramolecular carbon–π interaction with $Y50^{GYF1}$ ($F52^{GYF2}$). Importantly, key interactions of 4EHP with the a-L1 and PLAL motif are mediated by 4EHP-specific residues; e.g., E149, which fixes the orientation of the GYF1/2 chain via hydrogen bonds to $L77^{GYF1}$ ($L79^{GYF2}$) and $A78^{GYF1}$ ($A80^{GYF2}$), as well as $R146^{4EHP}$, which contacts $E80^{GYF1}$ ($V82^{GYF2}$) and $L77^{GYF1}$ ($L79^{GYF2}$) (Fig. 4C,D). Therefore, the a-L1 and the PLAL motif interactions are highly specific and involve residues present only in 4EHP (R103 and E149) and thus would not be possible with eIF4E.

The auxiliary sequences A2 and A3 arrange into two helical elements (auxiliary helices α2 and α3, respectively) (Fig. 2A–E), which are connected by a conserved VNS linker (linker 3 [a-L3]). Helix α2 (A2) shows some conformational heterogeneity across all complex structures (Supplemental Fig. S4H), most likely due to weak contacts at the interface (Fig. 4E–H). In contrast, a-L3 aligns well between the six complex structures. The linker VNS sequence enters a surface groove on 4EHP and interacts closely with 4EHP [e.g., through invariant $S96^{GYF1}$ ($S98^{GYF2}$)], which maintains a backbone hydrogen bond to $R138^{4EHP}$, while its hydroxyl group is in polar contact with the 4EHP-specific residue E177 (Fig. 4G,H). As a consequence of these interactions, helix α3 (A3) is positioned in close proximity to the 4EHP cap-binding pocket (Fig. 2C,D) and is stabilized in this orientation through hydrophobic contacts between $V99^{GYF1}$ ($V101^{GYF2}$) and $L100^{GYF1}$ ($L102^{GYF2}$) and the aliphatic side chain of R138 and I211 in 4EHP (Fig. 4G,H). However, the GYF1/2 peptides containing all of the 4EHP-binding elements did not contribute to 4EHP's affinity for the m^7GpppG cap analog as observed by ITC (Supplemental Table S2; Supplemental Fig. S3G, H), suggesting that additional GYF1/2 sequences may contribute to enhance 4EHP binding to capped mRNAs.

The auxiliary sequences contribute to complex stability in vivo

To assess the biological significance of the interactions mediated by the auxiliary sequences, we substituted 4EHP residues R103 and E149, which interact with the GYF1/2 PLAL motif, with leucine residues, as is observed in eIF4E (4EHP RE-LL mutant). The RE-LL substitutions strongly reduced binding to endogenous GYF2 compared with wild-type 4EHP or the 4EHP dorsal and lateral mutants in human cells (Fig. 5A, lanes 7–10). All of the mutations disrupted binding to endogenous GYF1 (Supplemental Fig. S6A, lanes 7–10). As a control, binding of 4E-BP1 was not affected by the RE-LL substitutions (Fig. 5A, lane 10), indicating that the mutations do not disrupt the 4EHP fold.

We also analyzed the impact of amino acid substitutions in the GYF1/2 proteins on complex formation.
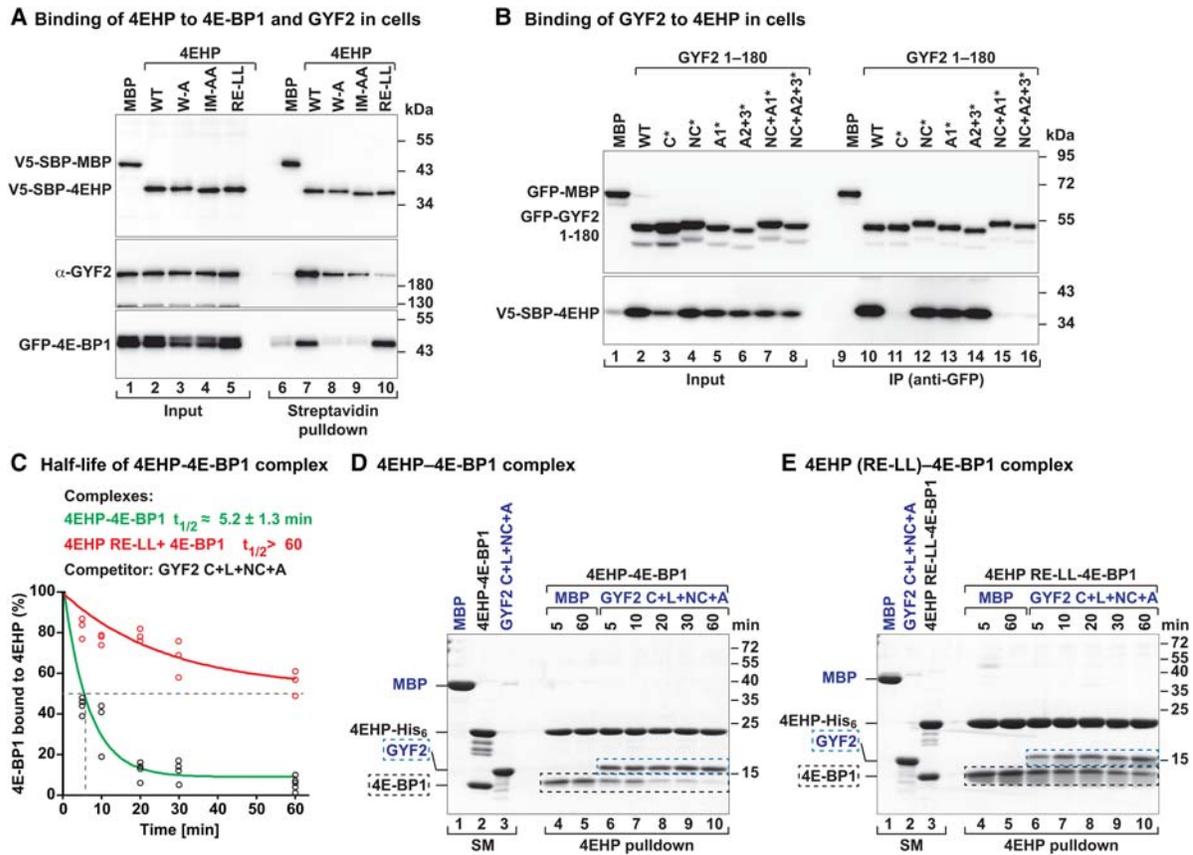
Figure 5. The auxiliary interactions are crucial for the formation of the 4EHP–GYF complex. (*A*) Western blot showing the interaction of endogenous GYF2 or GFP-4E-BP1 with V5-SBP-4EHP (wild type or the indicated mutants). The proteins were pulled down using streptavidin-coated beads. The inputs (1.5% for the V5-tagged proteins and 1% for GYF2 and GFP-4E-BP1) and bound fractions (3% for the V5-tagged proteins and 20% for GYF2 and GFP-4E-BP1) were analyzed by Western blot using the indicated antibodies. (*B*) Interaction of V5-SBP-4EHP with GFP-GYF2 (residues 1–180; either wild type or the indicated mutants). The proteins were immunoprecipitated using anti-GFP antibodies. GFP-MBP served as negative control. The inputs (1.5% for the GFP-tagged proteins and 0.5% for V5-SBP-4EHP) and immunoprecipitates (7.5% for the GFP-tagged proteins and 30% for V5-SBP-4EHP) were analyzed by Western blot using anti-GFP and anti-V5 antibodies. (*C–E*) Purified 4EHP–4E-BP1 complexes containing 4EHP-His$_6$ (wild type or the RE-LL mutant) were incubated in the presence of equimolar amounts of the GYF2 C+L+NC+A peptide C-terminally tagged with GB1 or MBP as a negative control. The proteins bound to 4EHP were pulled down using Ni-NTA beads at the indicated time points and analyzed by SDS-PAGE and Coomassie blue staining. *C* shows the quantification of the amount of 4E-BP1 still associated with 4EHP. $n = 3$. The half-life of the 4EHP–4E-BP1 complex ($t_{1/2}$) in the presence of the competitor protein is represented as the mean ± SD. *D* and *E* show representative SDS-PAGE gels. The positions of the GYF2 and 4E-BP1 peptides are marked by blue and black dashed boxes, respectively. The lanes labeled SM (starting material) show the purified complexes and peptides used in the assay.

Substitutions in the GYF1/2 noncanonical (NC*) or auxiliary (A1* and A2+3*) motifs did not affect binding to overexpressed 4EHP in human cells. This is consistent with structural data showing that R103 and E149 in 4EHP interact via their side chains with the backbone atoms of the GYF1/2 auxiliary sequences (Fig. 5B; Supplemental Fig. S6B, cf. lanes 12–14 and 10). However, binding was disrupted when the mutations in the noncanonical and auxiliary motifs were combined (Fig. 5B; Supplemental S6B, lanes 15, 16 vs. 10). Together with the data showing that the mutations in the canonical motif prevent GYF1/2 binding to 4EHP (Fig. 5B; Supplemental Fig. S6B), this indicates that the canonical motif is necessary but not sufficient for 4EHP-binding in vivo.

We further assessed the relevance of R103 and E149 toward complex stability in competition assays using preassembled 4EHP–4E-BP1 complexes containing either the wild type or the RE–LL 4EHP mutant. These complexes were challenged with an equimolar amount of the GYF2 fragment comprising all 4EHP-interacting elements. The amount of 4E-BP1 that remained bound to 4EHP was determined over time (Fig. 5C–E). The GYF2 fragment displaced 50% of 4E-BP1 from the preassembled 4EHP–4E-BP1 complexes in 5 min ± 1.2 min. Under the same conditions, the GYF2 fragment displaced only 40% of 4E-BP1 bound to the 4EHP RE-LL mutant after 60 min of incubation (Fig. 5C–E). Collectively, the competition experiments together with the observation that the GYF1/2 proteins do not associate with the 4EHP RE-LL

mutant in cell lysates (Fig. 5A; Supplemental Fig. S6A) indicate that the auxiliary sequences afford GYF2 a competitive advantage over 4E-BP1 for binding to 4EHP.

The canonical and auxiliary regions promote complex self-association in solution

In the asymmetric unit in the 4EHP–GYF1/2 crystals, the GYF1/2 canonical and auxiliary motifs are part of a large interface (1008 Å^2) connecting two neighboring complexes (Supplemental Fig. S7A–C), suggesting dimerization. We analyzed the solution properties of these complexes by small-angle X-ray scattering (SAXS). The SAXS parameters for the 4EHP–GYF2 complex were indeed consistent with those of a dimer (Supplemental Fig. S7A,D,E). We also observed that the self-association of the complexes was concentration-dependent (Supplemental Table S3). However, the SAXS measurements of complexes lacking the GYF2 auxiliary sequences were consistent with a monomeric state (Supplemental Fig. S7F,G; Supplemental Table S3).

The putative dimer interface is stabilized by residues that are 4EHP- and GYF-specific, including R202, M161, and Q159 in 4EHP and E46 and E47 in GYF2 (Supplemental Figs. S1A,B, S7A,B). We designed mutations in GYF2 and 4EHP to disrupt dimerization (dimerization mutant, D*) (Supplemental Table S1). These mutations did not affect complex assembly, as the mutated proteins still copurified as a complex (Supplemental Fig. S7H) and retained the same affinity for their partner as the wild-type proteins (Supplemental Fig. S3I,J; Supplemental Table S2). Importantly, however, the SAXS profile of the mutated complex demonstrated the best fit to a monomeric state (Supplemental Fig. S7I; Supplemental Table S3), indicating that the mutations effectively disrupt dimerization and validate the observed interface.

A 4EHP-specific residue reduces 4E-BP1 binding

To probe the molecular basis for the binding preference of 4E-BP1 for eIF4E over 4EHP observed in cell lysates, we measured the binding affinity of the 4E-BP1 peptide for 4EHP and eIF4E using ITC. The affinity of the 4E-BP1 peptide for 4EHP was 10-fold lower compared with eIF4E (K_D = 55 nM ± 14 nM and K_D = 5 nM ± 2 nM, respectively) (Supplemental Fig. S3K,L; Supplemental Table S2) and 100-fold lower compared with the GYF1/2 peptides for 4EHP (Supplemental Table S2).

A possible explanation for the lower affinity of 4E-BP1 for 4EHP compared with eIF4E is that R63$^{\text{BP1}}$ directly interacts with N77 in the eIF4E complex, but this residue is replaced by a Ser (S99) in 4EHP, which breaks this critical contact (Supplemental Fig. S5E,F). Interestingly, a 4EHP mutant in which Ser99 was substituted with Asn (4EHP S99N) showed a 10-fold gain in affinity for 4E-BP1 to a level comparable with eIF4E (4 nM ± 1 nM and 5 nM ± 2 nM, respectively) (Supplemental Fig. S3M; Supplemental Table S2). The S99N 4EHP mutant also bound to endogenous 4E-BP1 (Supplemental Fig. S7J, lane 8 vs. 7), although not to the same level as observed for eIF4E. Binding of

GYF2 was not affected by the S99N mutation because GYF proteins have an aromatic residue at the equivalent R63$^{\text{BP1}}$ position (Y50$^{\text{GYF1}}$ and F52$^{\text{GYF2}}$), which mediates hydrophobic contacts with 4EHP.

The affinity measurements indicate that endogenous 4E-BP1 is unlikely to effectively compete with GYF1/2 proteins for 4EHP binding under equilibrium conditions in cell lysates. This is consistent with the in vivo data that show that 4E-BP1 bound to 4EHP only when overexpressed (Fig. 1B). Furthermore, a single amino acid substitution is responsible for the different affinities of eIF4E and 4EHP for 4E-BP1.

4EHP requires interaction with GYF1/2 proteins to down-regulate mRNA expression

To assess the functional relevance of the 4EHP–GYF1/2 complex in repressing mRNA targets, we tethered λN-HA-tagged 4EHP to an R-Luc reporter containing five binding sites for the λN tag (BoxB hairpins) in the 3′ UTR. To uncouple the effects on translation from the effects on mRNA stability, the reporter contained an internal polyadenosine stretch of 95 residues followed by the 3′ end of the noncoding RNA MALAT1, which is generated through endonucleolytic cleavage by RNase P and is thus not polyadenylated (Wilusz et al. 2012). An F-Luc-GFP reporter served as a transfection control. The λN-HA-4EHP protein repressed the expression of the R-Luc-5BoxB-A$_{95}$-MALAT1 reporter relative to the λN-HA peptide (Fig. 6A) without causing corresponding changes in mRNA levels (Supplemental Fig. S8A,B). The levels of a reporter lacking the BoxB hairpins were not affected (Supplemental Fig. S8C–E), indicating that 4EHP recruitment is prerequisite for repression.

We used CRISPR–Cas9 gene editing to generate a GYF1/2-null HEK293T cell line in which the GYF1/2 levels were reduced below 10% of their control levels, whereas the expression of endogenous 4EHP was not affected (Fig. 6B, lane 4 vs. 1). In this cell line, the repression of the R-Luc mRNA reporter by tethered 4EHP was impaired (Fig. 6A) even though 4EHP was expressed at levels comparable with those observed in control cells (Fig. 6C, lane 2 vs. 4). The 4EHP-mediated repression was restored by transient expression of wild-type GFP-tagged GYF2 but not by the GYF2 canonical mutant (C*) that does not interact with 4EHP (Fig. 6A) despite comparable expression levels (Fig. 6C, lanes 5,6). Thus, 4EHP requires interactions with GYF1/2 proteins for full repressive activity.

In agreement with this conclusion, 4EHP activity in tethering assays correlated with GYF1/2 binding and was independent of cap binding. Indeed, 4EHP mutants with impaired GYF1/2 binding (+/−) exhibited reduced repressive activity, and repression was abolished by combined mutations that disrupt binding to GYF1/2 (Fig. 6D,E; Supplemental Fig. S8F,G). Unexpectedly, however, a 4EHP mutant that does not bind to the cap (cap*) (Supplemental Fig. S8H) but still binds to GYF1/2 (Supplemental Fig. S8I) repressed the expression of the reporter mRNA in a GYF1/2-dependent manner (Fig. 6D; Supplemental Fig. S9A,B). The 4EHP mutants did not repress
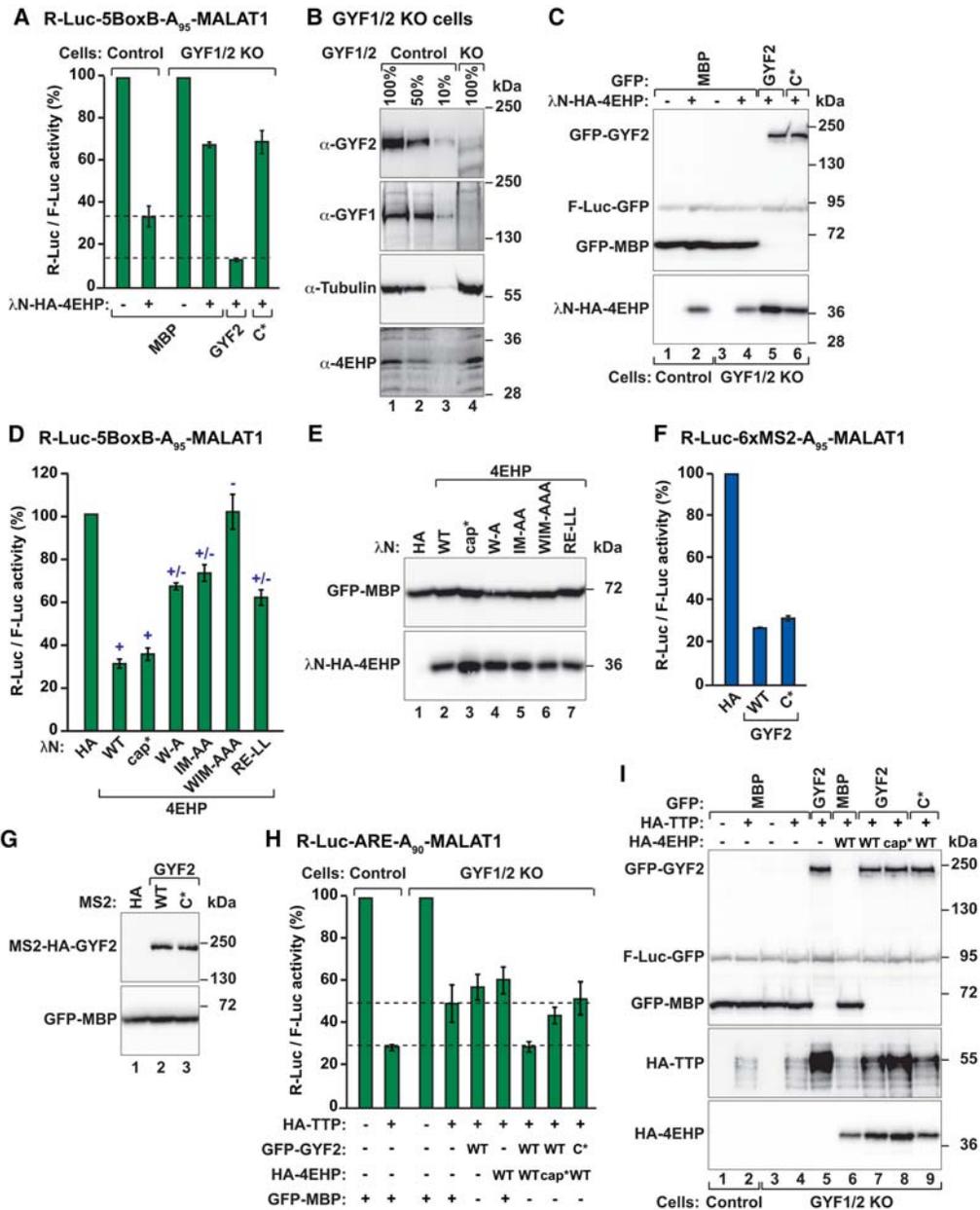
Figure 6. 4EHP requires interaction with GYF1/2 proteins to repress translation. (*A*) A complementation assay using the R-Luc-5BoxB-A$_{95}$-MALAT1 reporter and λN-HA-4EHP (either wild type or the indicated mutants) was performed in control and GYF1/2-null HEK293T cells expressing GFP-MBP or GFP-GYF2 (wild type or canonical mutant). A plasmid expressing F-Luc-GFP served as the transfection control. R-Luc activity was normalized to that of the F-Luc transfection control and set to 100% in cells expressing the λN-HA peptide. Bars represent the mean values, and error bars represent standard deviations from three independent experiments. (*B*) Western blot analysis showing that full-length GYF1/2 levels were strongly reduced relative to control levels in the GYF1/2-null cell line. (*C*) Western blot analysis showing the expression of the λN-HA-4EHP and GFP-GYF2 proteins used in the assay shown in *A*. (*D*) Tethering assay using the R-Luc-5BoxB-A$_{95}$-MALAT1 reporter and λN-HA-4EHP (wild type or mutants) in HEK293T cells. Samples were analyzed as described in *A*. (+) Binding to the GYF1/2 proteins; (+/−) reduced binding to the GYF1/2 proteins; (−) no binding to the GYF1/2 proteins. (*E*) Western blot showing the equivalent expression of the λN-HA-4EHP proteins used in the assay shown in *D*. (*F*) Tethering assay using the R-Luc-6xMS2-A$_{95}$-MALAT1 reporter and MS2-HA-GYF2 (wild type or canonical mutant) in HEK293T cells. The cells were also cotransfected with GFP-MBP and F-Luc-GFP as transfection controls. R-Luc activity was normalized to that of the F-Luc transfection control and set to 100% in cells expressing MS2-HA. Samples were analyzed as described in *A*. (*G*) Western blot analysis showing the equivalent expression of the MS2-HA-GYF2 proteins. (*H*) Control HEK293T cells or cells depleted of GYF1/2 (KO) were transfected with the R-Luc-ARE-A$_{90}$-MALAT1 reporter and plasmids expressing the indicated proteins. The F-Luc-GFP reporter served as a transfection control. R-Luc activity was normalized to that of the F-Luc transfection control and set to 100% in the absence of TTP for each cell line. (*I*) Western blot showing the expression of the proteins in the experiment shown in *H*. Note that TTP is stabilized in GYF1/2-null cells expressing GYF2. However, repression did not correlate with TTP levels but with the coexpression of wild-type GYF2 and 4EHP.

the expression of a control mRNA lacking the BoxB-binding sites (Supplemental Fig. S9C–E).

Our results suggest that the GYF1/2 proteins confer repressive activity to the 4EHP–GYF1/2 complexes. Accordingly, GYF2 repressed mRNA reporter expression in tethering assays. GYF2 activity was independent of 4EHP because the GYF2 canonical mutant still repressed the expression of the R-Luc-6xMS2-A$_{95}$-MALAT1 reporter as efficiently as wild-type GYF2 (Fig. 6F,G). GYF2 did not repress a reporter lacking the MS2-binding sites (Supplemental Fig. S9F,G).

Finally, to investigate 4EHP activity without artificial tethering, we used an R-Luc reporter that included two copies of the ARE present in the 3′ UTR of the TNF-α mRNA (R-Luc-ARE-A$_{90}$-MALAT1). This reporter was repressed in control cells expressing TTP (Fig. 6H). TTP-induced repression was relieved in GYF1/2-null cells (Fig. 6H), although TTP was expressed at levels comparable with those observed in control cells (Fig. 6I, lane 4 vs. 2). In the GYF1/2-null cell line, repression was restored only when GYF2 and 4EHP were coexpressed but not when each protein was expressed individually (Fig. 6H,I). No restoration was observed when 4EHP was coexpressed with a GYF2 canonical mutant (C*). In contrast to the observations in tethering assays, a 4EHP mutant that does not bind the cap (cap*) was impaired in restoring TTP-mediated repression, although it was expressed at levels comparable with wild type (Fig. 6H,I). Collectively, these results indicate that the assembly of the 4EHP–GYF2 complex is required for full repression of target mRNA expression.

Discussion

GYF1/2 proteins are able to discriminate between 4EHP and eIF4E, but the molecular basis for this discrimination remains unknown (Morita et al. 2012). Here, we show that the 4EHP-binding region of GYF1/2 proteins comprises canonical and noncanonical motifs connected by a linker, which recognize the dorsal and lateral surfaces of 4EHP, respectively, in a manner similar to that observed for the diverse 4E-BPs and eIF4G bound to eIF4E (Kinkelin et al. 2012; Peter et al. 2015a,b; Sekiyama et al. 2015; Grüner et al. 2016). Thus, dorsal and lateral binding is conserved and widespread among eIF4E family proteins. Given this common binding interface, GYF1/2 proteins achieve their remarkable selectivity for 4EHP by virtue of unique auxiliary sequences (C-terminal to the noncanonical motif) that contact a surface on 4EHP, which is more divergent among eIF4E paralogs. In particular, the 4EHP-specific residues R103, R140, and E149 interact with GYF1/2 auxiliary sequences and are important for complex formation in vivo. These interactions stabilize the complex assembly by increasing the affinity of the interaction and may have evolved to ensure that in vivo GYF1/2 proteins efficiently compete for 4EHP binding with other potential binding partners such as 4E-BP1, which is more abundant than GYF1/2 but has lower affinity for 4EHP (Hein et al. 2015; this study).

Intriguingly, as a consequence of the 4EHP-specific interactions by the auxiliary sequences, the helix α3 (A3) of GYF1/2 is oriented in close structural proximity to the cap-binding pocket of 4EHP. However, the GYF2 peptide containing all of the 4EHP-binding elements did not increase the affinity of 4EHP for an m^{7}GpppG cap analog (Supplemental Table S2). In the GYF1/2 proteins, there is a long stretch of Gly/Arg-rich sequence immediately following the auxiliary motifs. Given that such low-complexity Gly/Arg-rich regions often confer nonspecific RNA-binding properties to the proteins that contain them (Thandapani et al. 2013), it is tempting to speculate that the GYF1/2 proteins may play a role in stabilizing 4EHP bound to capped transcripts. Furthermore, because 4EHP has a reduced affinity for the cap structure compared with eIF4E (Rom et al. 1998; Zuberek et al. 2007), it is possible that the auxiliary sequence-mediated dimerization observed in this study may have some as yet undefined functions (e.g., increasing local concentration of repressor complexes on the mRNA), but this hypothesis needs to be tested in future studies. The affinity of 4EHP for capped mRNAs may also be stimulated by post-translational modifications such as ISG15 modification (Okumura et al. 2007) and monoubiquitinylation/diubiquitinylation (von Stechow et al. 2015), but whether these are synergistic with the GYF1/2 proteins is currently not known.

The mechanism of repressor complex assembly is likely to be divergent among the 4EHP interactors

Our study provides mechanistic insights into the assembly of 4EHP repressor complexes and raises the question of whether the binding mode is conserved among other 4EHP-BPs. However, it is important to note that although the *D. melanogaster* Brat protein sequence contains a canonical 4EHP-binding motif, the motif is buried within the hydrophobic core of a folded domain (the NHL domain) and therefore is unlikely to participate in interaction with 4EHP (Cho et al. 2006). Furthermore, the canonical motifs in *D. melanogaster* Bcd and the mammalian Prep1 proteins contain internal proline residues and are unlikely to adopt helical conformations, which are crucial for stable binding to 4EHP (Cho et al. 2006; Villaescusa et al. 2009). Thus, either the interaction with Bcd and Prep1 is indirect or the mode of binding has diverged. Only 4E-T features canonical and noncanonical motifs that bind directly to eIF4E and are likely to bind 4EHP in a similar manner (Kubacka et al. 2013; Peter et al. 2015a). Although the 4E-T orthologs do not contain motifs with similarity to the GYF1/2 auxiliary motifs, our structural data indicate that the precise sequence composition of these motifs may not be critical for interactions because the 4EHP-specific residues principally stabilize the complex via contacts with the auxiliary sequence's backbone. However, it remains to be seen whether 4E-T may indeed contain auxiliary sequences and what their mode of binding to 4EHP is.

4EHP–GYF1/2 complex assembly is required for post-transcriptional mRNA regulation

Because 4EHP has a reduced affinity for the cap structure compared with eIF4E (Rom et al. 1998; Zuberek et al. 2007), it has been proposed that it is recruited to specific mRNAs through interactions with proteins that are bound (directly or indirectly) to the mRNA, thus increasing its local concentration and competing with eIF4E in *cis* for binding to the 5′ cap (Cho et al. 2005, 2006; Villaescusa et al. 2009; Morita et al. 2012). According to this model, 4EHP should repress translation independently of GYF1/2 proteins when directly tethered to the 3′ UTR of an mRNA reporter. Unexpectedly, however, we observed not only that 4EHP loses its repressive activity in GYF1/2-null cells but also that its interaction with GYF1/2 proteins is in fact required for full repression. Thus, rather than merely facilitating 4EHP recruitment to an mRNA (e.g., by bridging the interaction between 4EHP and the zinc finger proteins ZNF598 and TTP), the GYF1/2 proteins act directly in the repression. In agreement with this conclusion, GYF1/2 repressed target transcripts in tethering assays independently of 4EHP binding. However, it is also evident that regulation of endogenous transcripts is dependent on the 4EHP–GYF1/2 complex assembly. Indeed, the TTP-mediated repression of an ARE-containing reporter in the GYF1/2-null cell line was restored only when 4EHP was coexpressed with GYF2 that was competent for binding to the 4EHP. Intriguingly, the cap binding by 4EHP was necessary for full repression in this context.

In summary, our studies reveal the structural basis for the assembly of a translational repressor complex consisting of 4EHP and its specific binding partners, the GYF1/2 proteins. We show that the GYF1/2 proteins directly contribute to the repressive activity of 4EHP, thus uncovering an unexpected facet of a mechanism that regulates mRNA expression.

Materials and methods

DNA constructs

The DNA constructs used in this study are described in the Supplemental Material and are listed in Supplemental Table S1. All of the constructs and mutations were confirmed by sequencing.

Protein expression and purification

All of the recombinant proteins were expressed in *E. coli* BL21 Star (DE3) cells (Thermo Fisher Scientific) grown in LB medium overnight at 20°C. The cells were lysed by sonication in lysis buffer containing 50 mM HEPES (pH 7.2), 200 mM (4EHP–4E-BP1) or 300 mM (4EHP–GYF1/2) NaCl, and 2 mM DTT supplemented with 5 µg/mL DNase I, 1 mg/mL lysozyme, and protease inhibitor cocktail (Roche). To purify the complexes containing 4EHP bound to GYF1, GYF2, or 4E-BP1 for crystallization and SAXS, His$_6$-tagged 4EHP (residues A52–F234) was coexpressed with MBP-tagged GYF1 (residues K33–M103), GYF2 (residues A35–T105), or 4E-BP1 (residues T50–S83). The complexes were purified from cleared cell lysates using an amylose resin (New England Biolabs) followed by cleavage of the MBP and His$_6$ tags

with HRV3C protease overnight at 4°C. After cleavage of the tags, the complexes were separated from the MBP and His$_6$ tags using a heparin column (5 mL of HiTrap Heparin HP, GE Healthcare) and further purified on a Superdex 75 column (GE Healthcare) in a buffer consisting of 10 mM HEPES (pH 7.2), 200 mM NaCl, and 2 mM DTT. The complexes were stored at −80°C or used directly for crystallization and SAXS. The 4EHP complexes used in the competition assays were expressed and purified as described above with the difference that the C-terminal His$_6$ tag was not removed from the 4EHP and that the copurified 4E-BP1 peptide contained a C-terminal GB1 tag. The complexes were stored in a buffer containing 20 mM Na-phosphate (pH 7.0), 200 mM NaCl, and 5% (w/v) glycerol.

For the pull-down assays, eIF4E (full length) was expressed with a C-terminal His$_6$ tag, purified from cleared cell lysates using a nickel column (5 mL of HisTrap HP, GE Healthcare), and further purified on a heparin column (5 mL of HiTrap Heparin HP, GE Healthcare) followed by size exclusion chromatography (Superdex 75 column, GE Healthcare) without removing the C-terminal His$_6$ tag. The purified eIF4E-His$_6$ was stored at −80°C in a buffer consisting of 20 mM Na-phosphate (pH 7.0) and 200 mM NaCl.

Pull-downs, competition assays, coimmunoprecipitation, and Western blotting

The in vitro pull-down and competition assays were performed as described previously (Igreja et al. 2014; Peter et al. 2015a,b). All coimmunoprecipitation and pull-down assays in HEK293T cell lysates were performed in the presence of RNase A as described previously (Peter et al. 2015a). All of the Western blots were developed using the ECL Western blotting detection system (GE Healthcare). The antibodies used in this study are listed in Supplemental Table S4. A detailed description of these assays is included in the Supplemental Material.

ITC measurements and SAXS

The ITC measurements and SAXS experiments are described in the Supplemental Material.

Crystallization, data collection, and structure determination

A detailed description of the crystallization conditions and the structure determination process is included in the Supplemental Material. All diffraction data sets were recorded on a Pilatus 6M detector at the PXII beamline of the Swiss Light Source at a temperature of 100 K. The diffraction data and refinement statistics are summarized in Table 1.

Tethering and complementation assays

A detailed description of the procedure to generate the GYF1/2-null cell line is included in the Supplemental Material. For the complementation assays, HEK293T cells (wild-type or GYF1/2-null cells) were seeded in six-well plates (0.6 × 10^6 cells per well) and transfected using Lipofectamine 2000 (Invitrogen). The tethering reporters have been described previously (Kuzuoğlu-Öztürk et al. 2016). The transfection mixtures contained 0.25 µg of pEGFP-N3-F-Luc transfection control reporter, 0.5 µg of pCIneo-R-Luc-5BoxB-A$_{95}$-MALAT1 (or pCIneo-R-Luc-A$_{95}$-MALAT1 without BoxB), and 0.3 and 0.7 µg of the plasmids expressing the λN-HA and λN-HA-tagged 4EHP proteins, respectively. Cells were also cotransfected with plasmids expressing GFP-tagged proteins (0.25 µg of GFP-MBP, 1.8 µg of GFP-GYF2 wild type, and 1.2 µg of GFP-GYF2 C* mutant).

For the 4EHP tethering assays in HEK293T cells, wild-type or mutant λN-HA-4EHP proteins (0.3 µg for wild type and the cap* mutant, 0.8 µg for W-A, 1 µg for IM-AA, 1.2 µg for WIM-AAA, and 0.4 µg for RE-LL proteins) were cotransfected with the same amounts of reporter plasmids as described for the complementation assay. In the tethering assay with the GYF2 protein, the transfection mixture contained 0.25 µg of pEGFP-N3-F-Luc transfection control reporter; 0.5 µg of pCIneo-R-Luc-6xMS2-A$_{95}$-MALAT1 or pCIneo-R-Luc-A$_{95}$-MALAT1; 0.3 µg and 1 µg of the plasmids expressing the MS2-HA and MS2-HA-tagged GYF2 proteins, respectively; and 0.25 µg of GFP-MBP.

For the assay with the pCIneo-R-Luc-ARE-A$_{90}$-MALAT1 reporter, wild-type and GYF1/2-null HEK293T cells were transfected with 1 µg of the ARE reporter and 0.25 µg of the pEGFP-N3-F-Luc transfection control reporter in the presence or absence of plasmids expressing 50 ng of λN-HA-TTPΔNIM, 0.2 µg of GFP-MBP, 1 µg of GFP-GYF2 (wild type or canonical mutant [C*]), and 0.5 µg of λN-HA-4EHP.

Firefly and *Renilla* luciferase activities were measured 2 d after transfection using the dual-luciferase reporter assay system (Promega).

Accession numbers

Atomic coordinates and structure factors have been deposited in the Protein Data Bank under accession codes 5NVK (4EHP–GYF1 C+L+NC+A), 5NVL (4EHP–GYF2 C+L+NC+A), 5NVM (4EHP–GYF2 C+L+NC), and 5NVN (4EHP–4E-BP1 C+L+NC).

References

Cho PF, Poulin F, Cho-Park YA, Cho-Park IB, Chicoine JD, Lasko P, Sonenberg N. 2005. A new paradigm for translational control: inhibition via 5′–3′ mRNA tethering by Bicoid and the eIF4E cognate 4EHP. *Cell* 121: 411–423.

Cho PF, Gamberi C, Cho-Park YA, Cho-Park IB, Lasko P, Sonenberg N. 2006. Cap-dependent translational inhibition establishes two opposing morphogen gradients in *Drosophila* embryos. *Curr Biol* 16: 2035–2041.

Fu R, Olsen MT, Webb K, Bennett EJ, Lykke-Andersen J. 2016. Recruitment of the 4EHP–GYF2 cap-binding complex to tetraproline motifs of tristetraprolin promotes repression and degradation of mRNAs with AU-rich elements. *RNA* 22: 373–382.

Giovannone B, Tsiaras WG, de la Monte S, Klysik J, Lautier C, Karashchuk G, Goldwurm S, Smith RJ. 2009. GIGYF2 gene disruption in mice results in neurodegeneration and altered insulin-like growth factor signaling. *Hum Mol Genet* 18: 4629–4639.

Gross JD, Moerke NJ, von der Haar T, Lugovskoy AA, Sachs AB, McCarthy JE, Wagner G. 2003. Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E. *Cell* 115: 739–750.

Gruner S, Peter D, Weber R, Wohlbold L, Chung MY, Weichenrieder O, Valkov E, Igreja C, Izaurralde E. 2016. The structures of eIF4E–eIF4G complexes reveal an extended interface to regulate translation initiation. *Mol Cell* 64: 467–479.

Hein MY, Hubner NC, Poser I, Cox J, Nagaraj N, Toyoda Y, Gak IA, Weisswange I, Mansfeld J, Buchholz F, et al. 2015. A human interactome in three quantitative dimensions organized by stoichiometries and abundances. *Cell* 163: 712–723.

Hernandez G, Altmann M, Sierra JM, Urlaub H, Diez del Corral R, Schwartz P, Rivera-Pomar R. 2005. Functional analysis of seven genes encoding eight translation initiation factor 4E (eIF4E) isoforms in *Drosophila*. *Mech Dev* 122: 529–543.

Igreja C, Peter D, Weiler C, Izaurralde E. 2014. 4E-BPs require non-canonical 4E-binding motifs and a lateral surface of eIF4E to repress translation. *Nat Commun* 5: 4790.

Jackson RJ, Hellen CU, Pestova TV. 2010. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol* 11: 113–127.

Joshi B, Cameron A, Jagus R. 2004. Characterization of mammalian eIF4E-family members. *Eur J Biochem* 271: 2189–2203.

Kinkelin K, Veith K, Grunwald M, Bono F. 2012. Crystal structure of a minimal eIF4E–Cup complex reveals a general mechanism of eIF4E regulation in translational repression. *RNA* 18: 1624–1634.

Kubacka D, Kamenska A, Broomhead H, Minshall N, Darzynkiewicz E, Standart N. 2013. Investigating the consequences of eIF4E2 (4EHP) interaction with 4E-transporter on its cellular distribution in HeLa cells. *PLoS One* 8: e72761.

Kuzuoglu-Ozturk D, Bhandari D, Huntzinger E, Fauser M, Helms S, Izaurralde E. 2016. miRISC and the CCR4–NOT complex silence mRNA targets independently of 43S ribosomal scanning. *EMBO J* 35: 1186–1203.

Lukhele S, Bah A, Lin H, Sonenberg N, Forman-Kay JD. 2013. Interaction of the eukaryotic initiation factor 4E with 4E-BP2 at a dynamic bipartite interface. *Structure* 21: 2186–2196.

Mader S, Lee H, Pause A, Sonenberg N. 1995. The translation initiation factor eIF-4E binds to a common motif shared by the translation factor eIF-4γ and the translational repressors 4E-binding proteins. *Mol Cell Biol* 15: 4990–4997.

Marcotrigiano J, Gingras AC, Sonenberg N, Burley SK. 1999. Cap-dependent translation initiation in eukaryotes is regulated by a molecular mimic of eIF4G. *Mol Cell* 3: 707–716.

Matsuo H, Li H, McGuire AM, Fletcher CM, Gingras AC, Sonenberg N, Wagner G. 1997. Structure of translation factor eIF4E

bound to m7GDP and interaction with 4E-binding protein. *Nat Struct Biol* **4:** 717–724.

Morita M, Ler LW, Fabian MR, Siddiqui N, Mullin M, Henderson VC, Alain T, Fonseca BD, Karashchuk G, Bennett CF, et al. 2012. A novel 4EHP–GIGYF2 translational repressor complex is essential for mammalian development. *Mol Cell Biol* **32:** 3585–3593.

Okumura F, Zou W, Zhang DE. 2007. ISG15 modification of the eIF4E cognate 4EHP enhances cap structure-binding activity of 4EHP. *Genes Dev* **21:** 255–260.

Paku KS, Umenaga Y, Usui T, Fukuyo A, Mizuno A, In Y, Ishida T, Tomoo K. 2012. A conserved motif within the flexible C-terminus of the translational regulator 4E-BP is required for tight binding to the mRNA cap-binding protein eIF4E. *Biochem J* **441:** 237–245.

Peter D, Igreja C, Weber R, Wohlbold L, Weiler C, Ebertsch L, Weichenrieder O, Izaurralde E. 2015a. Molecular architecture of 4E-BP translational inhibitors bound to eIF4E. *Mol Cell* **64:** 467–479.

Peter D, Weber R, Kone C, Chung MY, Ebertsch L, Truffault V, Weichenrieder O, Igreja C, Izaurralde E. 2015b. Mextli proteins use both canonical bipartite and novel tripartite binding modes to form eIF4E complexes that display differential sensitivity to 4E-BP regulation. *Genes Dev* **29:** 1835–1849.

Rom E, Kim HC, Gingras AC, Marcotrigiano J, Favre D, Olsen H, Burley SK, Sonenberg N. 1998. Cloning and characterization of 4EHP, a novel mammalian eIF4E-related cap-binding protein. *J Biol Chem* **273:** 13104–13109.

Rosettani P, Knapp S, Vismara MG, Rusconi L, Cameron AD. 2007. Structures of the human eIF4E homologous protein, h4EHP, in its m7GTP-bound and unliganded forms. *J Mol Biol* **368:** 691–705.

Sekiyama N, Arthanari H, Papadopoulos E, Rodriguez-Mias RA, Wagner G, Leger-Abraham M. 2015. Molecular mechanism of the dual activity of 4EGI-1: dissociating eIF4G from eIF4E but stabilizing the binding of unphosphorylated 4E-BP1. *Proc Natl Acad Sci* **112:** E4036–E4045.

Tao X, Gao G. 2015. Tristetraprolin recruits eukaryotic initiation factor 4E2 to repress translation of AU-rich element-containing mRNAs. *Mol Cell Biol* **35:** 3921–3932.

Tee AR, Tee JA, Blenis J. 2004. Characterizing the interaction of the mammalian eIF4E-related protein 4EHP with 4E-BP1. *FEBS Lett* **564:** 58–62.

Thandapani P, O'Connor TR, Bailey TL, Richard S. 2013. Defining the RGG/RG motif. *Mol Cell* **50:** 613–623.

Villaescusa JC, Buratti C, Penkov D, Mathiasen L, Planaguma J, Ferretti E, Blasi F. 2009. Cytoplasmic Prep1 interacts with 4EHP inhibiting Hoxb4 translation. *PLoS One* **4:** e5213.

von Stechow L, Typas D, Carreras Puigvert J, Oort L, Siddappa R, Pines A, Vrieling H, van de Water B, Mullenders LH, Danen EH. 2015. The E3 ubiquitin ligase ARIH1 protects against genotoxic stress by initiating a 4EHP-mediated mRNA translation arrest. *Mol Cell Biol* **35:** 1254–1268.

Wilusz JE, JnBaptiste CK, Lu LY, Kuhn CD, Joshua-Tor L, Sharp PA. 2012. A triple helix stabilizes the 3′ ends of long noncoding RNAs that lack poly(A) tails. *Genes Dev* **26:** 2392–2407.

Zuberek J, Kubacka D, Jablonowska A, Jemielity J, Stepinski J, Sonenberg N, Darzynkiewicz E. 2007. Weak binding affinity of human 4EHP for mRNA cap analogs. *RNA* **13:** 691–697.

Supplemental Materials

GIGYF1/2 proteins use auxiliary sequences to selectively bind to 4EHP and repress target mRNA expression

Daniel Peter, Ramona Weber, Felix Sandmeir, Lara Wohlbold, Sigrun Helms, Praveen Bawankar, Eugene Valkov, Cátia Igreja and Elisa Izaurralde

DNA constructs

The plasmids used for the expression of human eIF4E, 4E-BP1 and eIF4G1 (full-length or fragments) in *Escherichia coli* or in human cells have been previously described (Peter et al. 2015a; Grüner et al. 2016). The plasmids for the expression of 4EHP fragments (M1–F234) and (A52–F234) in *E. coli* were obtained by inserting the corresponding cDNA fragments either into the pnYC-NpH (between the XhoI and NheI restriction sites) or the pnYC-CvH (between the XhoI and BamHI restriction sites) vectors that include N- and C-terminal His$_6$ tags (Diebold et al. 2011), respectively. DNA fragments encoding for GYF1 [residues K33–K52 (C), K33–D71 (C+L+NC) and K33–M103 C+L+NC+A)] and GYF2 [residues K35–K54 (C), K35–Q72 (C+L+NC) and K35–T105 (C+L+NC+A)] were inserted into the NdeI-NheI and NdeI-XbaI restriction sites in the pnEA-NpM vector (Diebold et al. 2011), respectively. These constructs express GYF fragments that are N-terminally fused to an MBP-tag, which is cleavable by the HRV3C protease. A DNA fragment encoding the B1 domain of immunoglobulin-binding protein G (GB1; Cheng and Patel 2004) was inserted C-terminally to the GYF fragments by site-directed mutagenesis using the QuikChange mutagenesis kit (Stratagene).

The plasmids for the expression of V5-streptavidin binding protein (SBP)-tagged and λN-hemagglutinin (HA)-tagged 4EHP in human cells were obtained by inserting the full-length 4EHP cDNA into the XhoI and BamHI sites in the pT7-V5-SBP and pλN-HA-C1 vectors (Kuzuoglu-Ozturk et al. 2016), respectively. The plasmids for the expression of full-length GYF1 and GYF2, which are N-terminally fused to GFP or MS2-HA were obtained by inserting the GYF1 cDNA (XhoI-EcoRI, obtained from the Kazusa DNA Research Institute; sj03926) or the GYF2 cDNA (XhoI-BamHI) into the corresponding sites of the pT7-EGFP-C1 and pT7-MS2-HA-C1 vectors. cDNA fragments encoding for GYF1 (residues M1–C177) and GYF2 (residues M1–P180) were introduced into the XhoI and BamHI restriction sites in

the pT7-EGFP-C1 vector. The cDNA encoding 4E-T (eIF4E-Transporter protein; EIF4ENIF1) was inserted into the HindIII and BamHI restriction sites in the pT7-EGFP-C1 vector. The cDNA encoding TTP (Tristetraprolin, residues M1–P313; Fabian et al. 2013) was inserted between the XhoI and EcoRI restriction sites of the pλN-HA-C1 vector. To generate a reporter containing the ARE-element (pCIneo-R-Luc-ARE-A_{90}-MALAT1), the sequence of the ARE element present in the 3' UTR of the TNF (Tumor Necrosis Factor)-α mRNA was inserted twice into the 3' UTR of the pCIneo-R-Luc parental plasmid by site-directed mutagenesis. A cDNA containing a stretch of 90 A and the MALAT1 sequence was then inserted into the XhoI and NotI restriction sites of the R-Luc-ARE vector. The DNA sequence of the TNF-α ARE is as follows: TTATTTATTATTTATTTATTATTTATTTATTT. All of the mutants used in this study were generated by site-directed mutagenesis using the QuikChange mutagenesis kit (Stratagene). All of the constructs and mutations were confirmed by sequencing and are listed in Supplemental Table S1.

Pulldown and competition assays

In the pulldown assays shown in Fig. 1 and Supplemental Fig. S2 and S7, bacterial lysates expressing recombinant 4EHP-His_6 (residues M1–F234, wild-type and mutants) or purified eIF4E-His_6 (2 μM; 50 μg) were incubated with Ni-NTA beads for 30 min. The immobilized 4EHP and eIF4E proteins were then incubated for 30 min with bacterial lysates expressing GYF1, GYF2 or 4E-BP1 fragments (wild-type and mutants) that were N-terminally tagged with MBP and C-terminally tagged with GB1. Proteins associated with 4EHP or eIF4E were eluted with imidazole and analyzed by SDS-PAGE followed by Coomassie staining.

For the competition assays, purified 4EHP–4EBP1 complexes (2 μM) containing 4EHP (residues A52–F234)-His_6 and 4E-BP1 C+L+NC (residues R50–S83; with a C-terminal GB1

tag) were incubated with Ni-NTA beads for 30 min in 50 mM Na-phosphate (pH 7.0) and 200 mM NaCl. The immobilized complexes were then incubated with equimolar amount of purified, GB1-tagged competitor peptides or with MBP as a negative control. After the specified time points, the beads were pelleted and washed three times in the same buffer. Proteins bound to the Ni-NTA beads were eluted with the same buffer containing 500 mM imidazole and analyzed by SDS-PAGE followed by Coomassie staining. The amount of 4E-BP1 bound to 4EHP was quantified using the ImageJ software and normalized to 4EHP levels present at each time point. These values were set to 100 in the presence of MBP. Data points from at least three independent experiments were plotted and the resulting fitting curves were determined using the Levenberg-Marquardt algorithm for single exponential decay functions. The R^2 values associated with the fitting of the exponential decay curves were between 0.82 and 0.96.

ITC analysis

For the ITC measurements, the GB1-stabilized GYF1/2 peptides (wild-type and mutants) and 4E-BP1 peptides were purified as previously described for the other 4E-BPs (Igreja et al. 2014; Peter et al. 2015a,b). The 4EHP protein (residues A52–F234; wild-type and mutants) used in the ITC measurements was expressed with an N-terminal His$_6$ tag and purified from cleared cell lysates using a nickel column (HisTrap HP 5 ml, GE Healthcare). The His$_6$ tag was cleaved by HRV3C protease overnight at 4°C. The protein was further purified using a heparin column (HiTrap Heparin HP 5 ml, GE Healthcare) and a final purification on a Superdex 75 column (GE Healthcare). The 4EHP-GYF2 complex (GYF2 residues K35–T105) used for measuring the affinity for m^{7}GpppG cap analog was purified as described for the 4EHP–GYF2 complex used for crystallization. All of the proteins used in the ITC

measurements were stored at -80°C in a buffer consisting of 20 mM Na-phosphate (pH 7.0) and 200 mM NaCl.

The ITC experiments were performed using a VP-ITC microcalorimeter (MicroCal) at 20°C as described previously (Igreja et al. 2014; Peter et al. 2015a,b). A solution containing either 4EHP (residues A52–F234, wild-type, S99N mutant and dimerization mutant, 1-5 µM) or eIF4E (residues K36–V217, 5 µM) in a calorimetric cell was titrated with tenfold concentrated solutions of GB1-stabilized peptides that were dissolved in the same buffer (20 mM Na-phosphate (pH 7.0) and 150 mM NaCl). The following peptides were used: GYF1 (C, residues K33–K52, 50 µM; C+L+NC, residues K33–D71, 20 µM; C+L+NC+A wild type or dimerization mutant, residues K33–M103, 10 µM), GYF2 (C, residues K35–K54, 50 µM; C+L+NC, residues K35–Q72, 20 µM; C+L+NC+A wild type or dimerization mutant, residues K35–T105, 10 µM) and 4E-BP1 (C+L+NC, residues T50–S83, 50 µM). The affinity for the m^7GpppG cap analog was measured in a buffer containing 20 mM HEPES (pH 7.5) and 200 mM NaCl by titrating a solution of m^7GpppG (400 µM; New England Biolabs) into a solution of 4EHP (residues A52–F234, 40 µM) or 4EHP in complex with GIGYF2 (residues K35–T105, 40 µM) diluted in the same buffer.

The titration experiments consisted of an initial injection of 2 µl followed by 28 injections of 10 µl at 240 s intervals. Each binding experiment was repeated three times. Correction for dilution heating and mixing was achieved by subtracting the final baseline, which consisted of small peaks of similar size. The thermodynamic parameters were estimated using a one-site binding model (Origin version 7.0), whereby the data points for the first injection were removed from the analysis (Mizoue and Tellinghuisen 2004). Because the protein concentration used in these measurements is low (1 µM for 4EHP in the calorimetric cell), dimerization of the 4EHP–GYF1/2 complexes is unlikely to occur and thus it does not contribute to the measured binding constants. Accordingly, GYF2 and 4EHP dimerization

mutants still display a binding affinity similar to that observed for the complexes containing the wild type proteins (Supplemental Table S2 and Fig. S7).

Crystallization

Crystals of 4EHP (residues A52–F234) in complex with GYF1 (residues K33–M103) were obtained at 18°C using the hanging-drop vapor diffusion method two days after mixing the protein solution (16 mg/ml; 1 µl) with the crystallization solution (1 µl) containing 20% PEG 3350 in 0.2 M potassium nitrate. Crystals of 4EHP (residues A52–F234) bound to GYF2 (residues A35–T105) were obtained at 18°C using the hanging-drop vapor diffusion method one day after mixing the protein solution (16 mg/ml, 1 µl) with the crystallization solution (1 µl) containing 0.1 M sodium citrate (pH 5.0), 0.1 M magnesium chloride and 12% PEG 4000.. Crystals of 4EHP (residues A52–F234) in complex with GYF2 (residues A35–Q72) were obtained at 18°C using the hanging-drop vapor diffusion method. Crystals grew in one day after mixing the protein solution (16 mg/ml, 1 µl) with the crystallization solution (1 µl) containing in 0.1 M sodium acetate (pH 4.6) and 0.6 M diammonium phosphate. All of the crystals containing GYF peptides were soaked in mother liquor supplemented with 10–15% glycerol for cryoprotection before flash-freezing in liquid nitrogen.

Crystals of 4EHP (residues A52–F234) in complex with 4E-BP1 (residues T50–S83) were obtained at 18°C using the sitting-drop vapor diffusion method. The crystals grew three days after mixing the protein solution (16.5 mg/ml; 0.1 µl) with the crystallization solution (0.1 µl) containing 0.1 M sodium acetate (pH 4.6) and 1.7 M sodium formate. The crystals were cryoprotected in mother liquor supplemented with 3.5 M sodium formate and flash-frozen in liquid nitrogen.

Data collection and structure determination

The data for all the crystals were collected at 100K on a PILATUS 6M detector at the PXII beamline at the Swiss Light Source. Diffraction data were processed with XDS and scaled using XSCALE (Kabsch 2010). The phases were obtained by molecular replacement using PHASER (McCoy et al., 2007). For the 4EHP–GYF2 (residues K35–T105; C+L+NC+A) complex, the structure of human 4EHP (PDB 2JGB; Rosettani et al. 2007) was used as a search model with an asymmetric unit containing two copies of the model. To solve the structures of the 4EHP–GYF2 (residues K35–Q72; C+L+NC) and 4EHP–4E-BP1 (residues T50–S83; C+L+NC) complexes, two copies of 4EHP from the 4EHP–GYF2 (C+L+NC+A) complex were used as a search model. In the case of the 4EHP–GYF1 (residues K33–M103; C+L+NC+A) complex, four copies of the 4EHP–GYF2 (C+L+NC+A) complex were used as a search model. To minimize model bias, the molecular replacement solutions were used to rebuild the initial models using the PHENIX AutoBuild wizard (Terwilliger et al., 2008). To complete the structure, iterative cycles of model building and refinement were performed with COOT (Emsley et al. 2010) and PHENIX (Afonine et al. 2012), respectively. The GYF2 (C+L+NC+A and C+L+NC) and 4E-BP1 (C+L+NC) peptide chains were manually built into the difference density in COOT and further refined with PHENIX. In the final refinement rounds for the 4EHP–GYF1 (C+L+NC+A) and 4EHP–GYF2 (C+L+NC+A) complexes, translation/libration/screw (TLS) parameters were refined for the peptide chains in addition to the individual B-factors; in the case of the GYF1 complex, non-crystallographic symmetry (NCS) torsional restraints were also used in refinement.

The stereochemical properties for all of the structures were verified with MOLPROBITY (Chen et al. 2010), and structural images were prepared with PyMOL (http://www.pymol.org). The diffraction data and refinement statistics are summarized in Table 1.

Small-angle X-ray scattering (SAXS)

SAXS experiments were conducted at the SWING beamline at the SOLEIL synchrotron. Data collection for the 4EHP–GYF2 complexes was performed in-line with size exclusion chromatography (Superdex 200 Increase 5/150 GL, GE Healthcare) using an Agilent HPLC system in a buffer containing 20 mM HEPES pH 7.5, 200 mM NaCl, and 1 mM TCEP [Tris(2-carboxyethyl)phosphine]. The scattering data were collected at 1 s exposures using an Aviex charge-coupled device detector at a sample-detector distance of 1798 mm and a wavelength of 1.033 Å. Data reduction to absolute units, frame averaging and buffer subtraction were performed using the FOXTROT software (Xenocs, France). Theoretical scattering curves and fitting to the experimental SAXS data was performed using the FoXS software (Schneidman-Duhovny et al. 2013). To ensure protein stability during SAXS data collection, all the 4EHP–GYF2 complexes were measured with a 1.5x molar excess of m^7GpppG cap analog (New England Biolabs) in the protein samples. Therefore, the coordinates of the structures used during the fitting procedures were adjusted such that the 4EHP cap-binding loops were fixed in the bound conformation including the cap analog, which was based on the structure of the m^7GTP-bound 4EHP (PDB 2JGB; Rosettani et al. 2007).

Generation of GYF1/2-null cell line

Two sgRNAs targeting GYF1 and two sgRNAs targeting GYF2 were designed using the DNA 2.0 (ATUM) or CHOPCHOP (http://chopchop.cbu.uib.no) online tools and cloned into the pSpCas9(BB)-2A-Puro (PX459) vector [a gift from F. Zhang, Addgene plasmid 48139; (Ran et al., 2013)]. HEK293T cells were transfected with the sgRNA-Cas9 vectors and selected with puromycin (3 µg ml^{-1}) to obtain stable GYF1/2 knockout cells. To obtain clonal cell lines, single cells were distributed in 96-well plates using serial dilutions. Genomic

DNAs from single clones were isolated using the Wizard SV Genomic DNA Purification System (Promega) and the targeted GYF1 and GYF2 loci were PCR amplified and sequenced to confirm gene editing. For GYF1 we observed two frameshift mutations in exon 7 (4 bp deletion and a 8 bp deletion together with a C->T mutation) targeted by sgGYF1-a and two indels in exon 16 (insertion of 43 bp or 93 bp) produced by sgGYF1-b. These mutations changed the GYF1 reading frame after the respective targeted site and introduced premature STOP codons. One frameshift mutation (16 bp deletion in the first exon removing the start codon) was detected for the GYF2 locus (targeted by sgGYF2-a and sgGYF2-b). This deletion was caused by sgGYF2-b. In contrast, sgGYF2-a did not target the genomic locus as the sequence around this target site is wild-type. The knockouts of GYF1/2 were further confirmed by western blotting. For the GYF2 gene we observe low levels of truncated protein fragments that are consistent with translation initiation at internal AUGs (Figure 6B, lane 4). Taking the GYF2 sequence and the position of the mutations into account, these truncated forms lack the 4EHP-binding region and the expression levels are approximately 10% of wild-type levels. The following guide sequences were used: sgGYF1-a: 5'-GCCAGCGGTCGCCGTCTCGC-3'; sgGYF1-b: GACAAGGACCGGCTCATCGT-3'; sgGYF2-a: 5'- ATTTTGAAAACTCACCATTC-3'; sgGYF2-b: 5'-AATACGGAAAAGAATGGCAG

Supplemental Table S1. Mutants and constructs used in this study.

Protein	Name of the construct	Fragments / mutations	Binding site / motif
Hs **4EHP** (isoform 1) O60573-1	4EHP	Full-length (1–245)	
	4EHP ΔC-term	1–234	Δ235–245
	4EHP truncated	52–234	Δ1–51 & 235–245
	W-A	W95A	Dorsal surface
	IM-AA	I85A, M101A	Lateral surface
	WIM-AAA	W95A, I85A, M101A	Dorsal + lateral surface
	RE-LL	R103L, E149L	Auxiliary surface
	WRE-ALL	W95A, R103L, E149L	Dorsal + auxiliary surface
	S99N	S99N	Dorsal surface
	Cap mutant (cap*)	W124A	Cap-binding pocket
	D* (dimer mutant)	Q159S, M161D, R202E	Dimer interface
Hs **eIF4E** (isoform 1) P06730-1	4E	Full-length (1–217)	
	4E trunc	36–217	
	Cap mutant (cap*)	W102A	Cap-binding pocket
Hs **GIGYF1** O75420	GYF1	Full-length (1–1035)	
	C+L+NC+A	33–103	Complete 4EHP-binding region
	C+L+NC	33–71	Bipartite 4EHP-binding region
	C	33–52	Canonical 4EHP-binding region
	C*	Y39A, Y41A, M46A, L47A	Canonical
	NC*	L60D, F65D, V68D	Non-canonical
	A1*	P76D, L77A	Auxiliary site 1
	A2*	E86A, N95F	Auxiliary site 2
	NC+A1*	L60D, F65D, V68D, P76D, L77A	Non-canonical + auxiliary site 1
	NC+A2+3*	L60D, F65D, V68D, E86A, N95F	Non-canonical + auxiliary site 2
	1-177		N-terminus
	D* (dimer mutant)	E44A, E45F, Q87A	Dimer interface
Hs **GIGYF2** (isoform 1) Q6Y7W6-1	GYF2	Full-length (1–1299)	
	C+L+NC+A	35–105	Complete 4EHP-binding region
	C+L+NC	35–72	Bipartite 4EHP-binding region
	C	35–54	Canonical 4EHP-binding region
	C*	Y41A, Y43A, M48A, L49A	Canonical motif
	NC*	L62D, F67D, I70D	Non-canonical
	A1*	P78D, L79A	Auxiliary site 1
	A2*	E88A, N96F	Auxiliary site 2
	NC+A1*	L62D, F67D, I70D, P78D, L79A	Non-canonical + auxiliary site 1
	NC+A2+3*	L62D, F67D, I70D, E88A, N97F	Non-canonical + auxiliary site 2
	1-180		N-terminus
	D* (dimer mutant)	E46A, E47F, Q89A	Dimer interface
Hs **4E-BP1** Q13541	4E-BP1	Full-length (1–118)	
	4E-BP1 C+L+NC	50–83	eIF4E-binding region
	C*	Y54A, L59A	Canonical motif
	NC*	L75A, V81A	Non-canonical
	C+NC*	Y54A, L59A, L75A, V81A	Canonical+ non-canonical
Hs **4E-T** Q9NRA8	4E-T	Full-length (1–985)	
Hs **TTP** (1–326) P26651	TTP ΔNIM	1–313	Δ314-326, deletion of the NOT1 interacting motif (NIM)

Supplemental Table S2. Thermodynamic parameters for the interaction of 4EHP and eIF4E with the indicated peptides.

GYF peptides *vs* 4EHP					
GYF protein	**K_D (nM)**	**ΔH (kcal mol^{-1})**	**$-T\Delta S$ (kcal mol^{-1})**	**ΔG (kcal mol^{-1})**	**Molar ratio**
GYF1 C (33-52)	360 ± 120	-23 ± 3	14.6	-8.7	1.00 ± 0.01
GYF1 C+L+NC (33-71)	12 ± 2	-30 ± 1	19.3	-10.6	1.00 ± 0.01
GYF1 C+L+NC+A (33-103)	0.4 ± 0.2	-37 ± 4	24.4	-12.6	1.00 ± 0.02
GYF2 C (35-54)	290 ± 160	-22 ± 2	13.5	-8.8	1.00 ± 0.01
GYF2 C+L+NC (35-72)	14 ± 1	-23 ± 2	12.3	-10.6	1.00 ± 0.01
GYF2 C+L+NC+A (35-105)	0.3 ± 0.1	-32 ± 1	19.1	-12.8	1.00 ± 0.01
GYF peptides *vs* 4EHP dimerization mutants					
GYF protein	**K_D (nM)**	**ΔH (kcal mol^{-1})**	**$-T\Delta S$ (kcal mol^{-1})**	**ΔG (kcal mol^{-1})**	**Molar ratio**
GYF1 C+L+NC+A (33-103) D*	0.4 ± 0.3	-34 ± 1	21.6	-12.7	1.01 ± 0.01
GYF2 C+L+NC+A (35-105) D*	0.5 ± 0.3	-30.8 ± 0.5	18.1	-12.7	1.01 ± 0.02
4EBP1 C+L+NC *vs* eIF4E or 4EHP					
4E molecule	**K_D (nM)**	**ΔH (kcal mol^{-1})**	**$-T\Delta S$ (kcal mol^{-1})**	**ΔG (kcal mol^{-1})**	**Molar ratio**
eIF4E	5 ± 2	-18 ± 1	6.4	-11.2	1.00 ± 0.01
4EHP	55 ± 14	-16.4 ± 0.8	6.6	-9.8	1.01 ± 0.01
4EHP S99N	4 ± 1	-21 ± 2	9.7	-11.3	1.01 ± 0.01
m^{7}GpppG cap analog *vs* 4EHP or 4EHP-GIGYF2 complex					
Protein	**K_D (µM)**	**ΔH (kcal mol^{-1})**	**$-T\Delta S$ (kcal mol^{-1})**	**ΔG (kcal mol^{-1})**	**Molar ratio**
4EHP	4 ± 1	-7.3 ± 0.4	0.1 ± 0.5	-7.2	1.01 ± 0.01
4EHP–GYF2 (35–105) complex	6 ± 3	-9 ± 2	1 ± 3	-7.2	1.01 ± 0.01

Note that the presence of the auxiliary sequences increases the entropic penalty (-TΔS) of the interaction between GYF1/2 and 4EHP compared to that of the peptides lacking these sequences [Δ(-TΔS)^{GYF1} = 5.1 kcal/mol^{-1}, Δ(-TΔS)^{GYF2} = 6.8 kcal/mol^{-1}]. One explanation for the increase in the entropic penalty is a higher disorder-to-order transition for the binding of the GYF1/2 C+L+NC+A peptides compared to the C+L+NC peptides. This is supported by the crystal structures in which the auxiliary sequences fold into two α-helices in complex with 4EHP.

Supplemental Table S3. Experimental and theoretical SAXS parameters for different 4EHP–GYF complexes.

Experimental parameters					
4EHP bound to:	**R_g (Guinier) [Å]**	**R_g (real space) [Å]**	**D_{max} [Å]**	**Exp. I(0) [x10^{-2}]**	**Concentration [mg/ml]**
GYF2 C+L+NC	21.1	21.2	71.1	4	10
GYF2 C+L+NC+A	26.1	26.2	90.3	7.9	10
GYF2 C+L+NC+A	25.8	25.8	89.3	3.6	5
GYF2 C+L+NC+A	24.7	24.8	83.7	1.7	2.5
GYF2 C+L+NC+A	23.7	23.7	80.2	0.7	1.25
4EHP D mutant bound to:	**R_g (Guinier) [Å]**	**R_g (real space) [Å]**	**D_{max} [Å]**	**Exp. I(0) [x10^{-2}]**	**Concentration [mg/ml]**
GYF2 C+L+NC+A D* (dim. mutant)	20.6	20.6	73.7	1.3	5
Theoretical parameters					
4EHP–GYF2 C+L+NC+A					
Single complex			Symmetric dimer		
R_g (Guinier) [Å]	**D_{max} [Å]**		**R_g (Guinier) [Å]**	**D_{max} [Å]**	
17.7	64		24.7	85	
4EHP–GYF2 C+L+NC					
Single complex			Symmetric dimer		
R_g (Guinier) [Å]	**D_{max} [Å]**		**R_g (Guinier) [Å]**	**D_{max} [Å]**	
17.1	64		26.1	90	

Supplemental Table S4. Antibodies used in this study.

| Antibody	Source	Catalog Number	Dilution	Monoclonal/ Polyclonal
Anti-HA-HRP (Western blot)	Roche	12 013 819 001	1:5,000	Monoclonal
Anti-HA (Immunoprecipitation)	Biolegend	MMS-101P	1:1,000	Monoclonal
Anti-*Hs* GYF2	Bethyl laboratories	A303-731A	1:1,000	Rabbit polyclonal
Anti-*Hs* GYF1	Bethyl laboratories	A304-132A-M	1:1,000	Rabbit polyclonal
Anti-*Hs* 4E-T	Abcam	ab95030	1:2,000	Rabbit polyclonal
Anti-*Hs* 4EHP	In house		1:200	Rabbit polyclonal
Anti-*Hs* eIF4E	Bethyl laboratories	A301-154A	1:2,000	Rabbit polyclonal
Anti-*Hs* 4E-BP1	Cell Signaling Technology	9452	1:1,000	Rabbit polyclonal
Anti-GFP	In house		IP	Rabbit polyclonal
Anti-GFP	Roche	11814460001	1:2,000	Monoclonal
Anti-rabbit-HRP	GE Healthcare	NA934V	1:10,000	Polyclonal
Anti-mouse-HRP	GE Healthcare	RPN4201	1:10,000	Polyclonal
Anti-V5	QED Bioscience Inc.	18870	1:5,000	Rabbit polyclonal
Anti-V5	LSBio LifeSpan BioSciences, Inc.	LS-C57305	1:5,000	Monoclonal
Anti-tubulin	Sigma Aldrich	T6199	1:10,000	Monoclonal

**Figure S1**

A

B
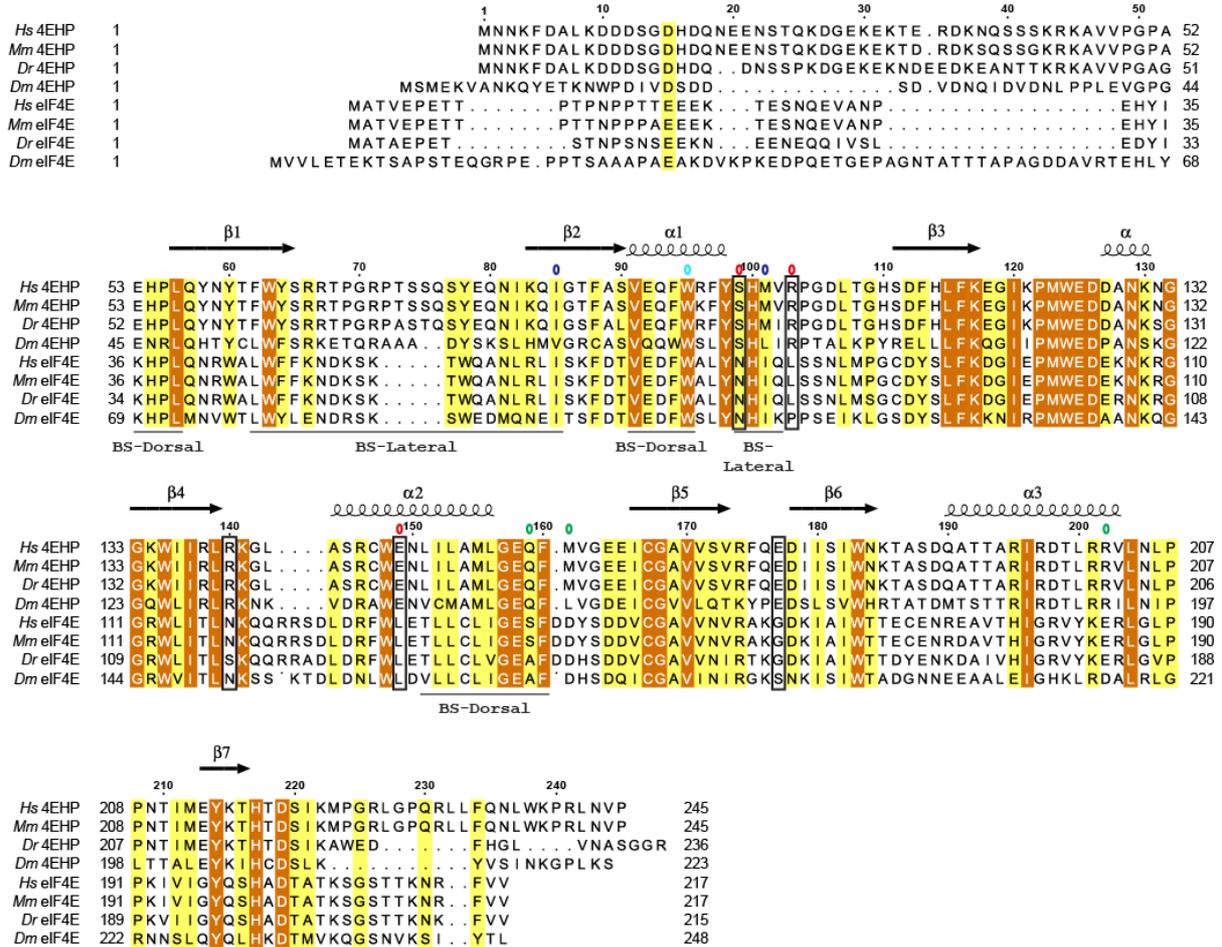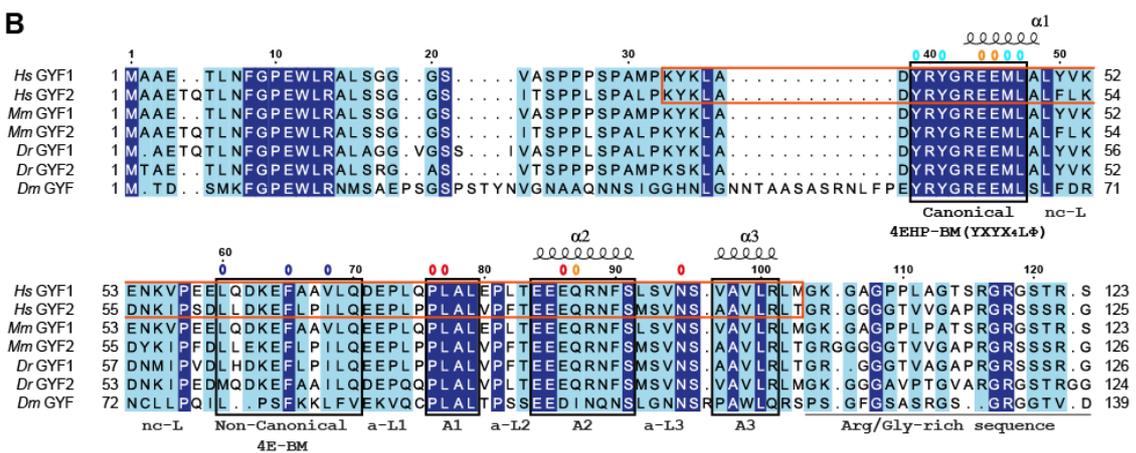
Supplemental Figure S1. Sequence alignments. In all aligned sequences, residues with >70% similarity are shown with a light color background and conserved residues are highlighted with a darker background and printed in white. Secondary structure elements are indicated above the sequences for 4EHP and GYF1 and are based on the structures presented

in this study. (*A*) Sequence alignment of 4EHP and eIF4E orthologous proteins from *Homo sapiens* (*Hs*), *Mus musculus* (*Mm*), *Danio rerio* (*Dr*) and *Drosophila melanogaster* (*Dm*). Residues highlighted in black boxes are specific for 4EHP and are relevant for the interactions described in this study. The dorsal and lateral binding surfaces (BS) of 4EHP are indicated by a line below the sequences. Residues that were mutated in this study are indicated by open circles colored as follows: cyan (dorsal surface), blue (lateral surface), red (4EHP specific residues) and green (dimerization). (*B*) Sequence alignment of GYF proteins. The canonical (C), non-canonical (NC) and auxiliary (A1, A2, A3) sequences are boxed in black. The GYF1/2 sequences visible in the crystal structures are indicated with a red box. Only a short stretch of the Arg/Gly-rich sequence adjacent to the auxiliary motif is shown and underlined. The species are as in A. Open circles above the alignment indicate the residues mutated in this study and are colored as follows: cyan (canonical), blue (non-canonical), red (auxiliary) and orange (dimerization).

Figure S2
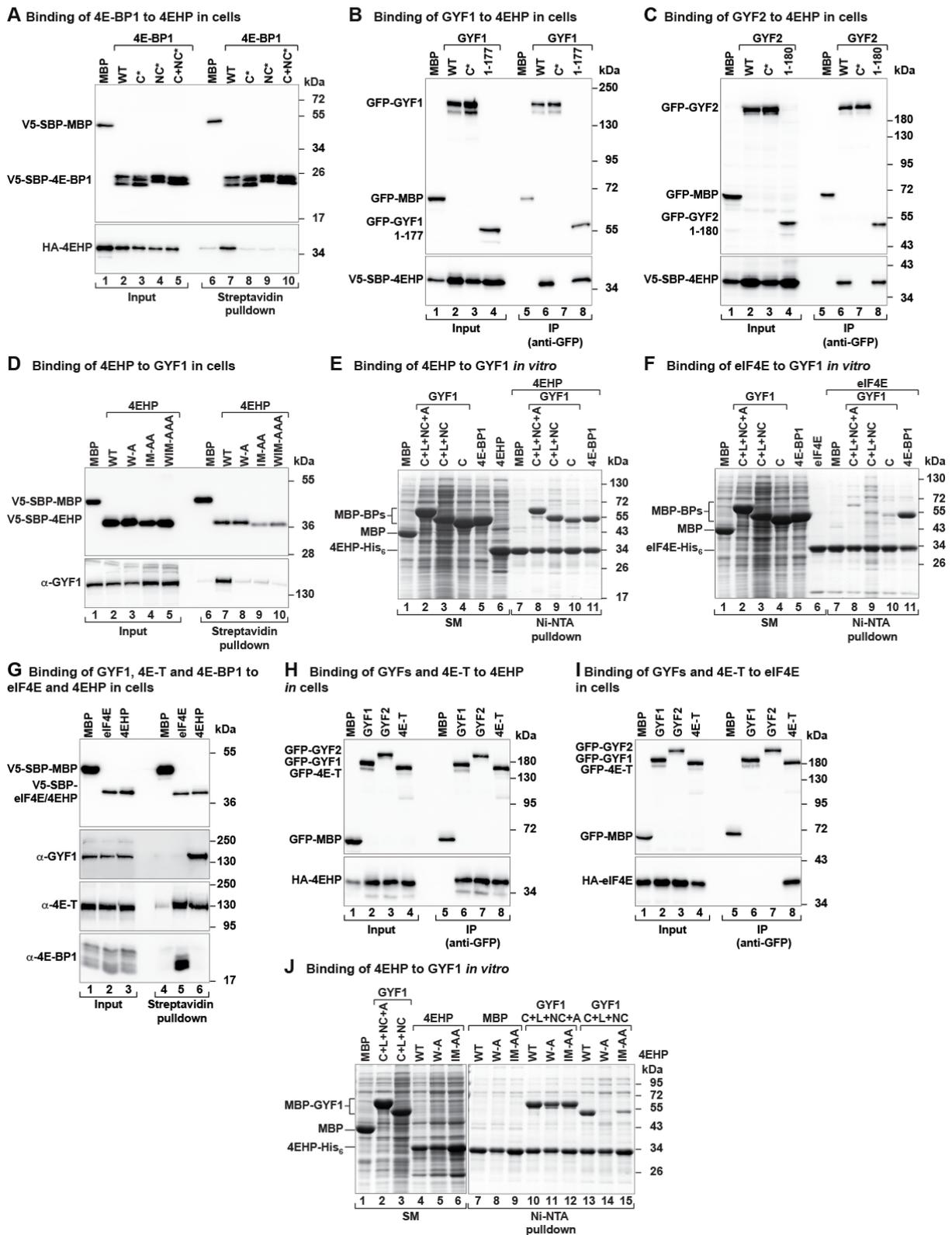
Supplemental Figure S2. Interaction of GYF1, GYF2 and 4E-BP1 with 4EHP. (*A*) The interaction of HA-4EHP with V5-SBP-4E-BP1 (wild-type or the indicated mutants) was tested in HEK293T cell lysates. The proteins were pulled down using streptavidin-coated

beads. V5-SBP-MBP served as negative control. The inputs (1.5%) and bound fractions (3% for the V5-proteins and 5% for HA-4EHP) were analyzed by western blotting using anti-HA and anti-V5 antibodies. (*B*) The interaction of GFP-GYF1 [either full-length, canonical mutant (C*) or N-terminal fragment (residues 1–177)] with V5-SBP-4EHP was analyzed by immunoprecipitation assay in HEK293T cells using anti-GFP antibodies. GFP-MBP served as negative control. The input samples (1.5%) and the immunoprecipitates (10%) were analyzed by western blotting using anti-V5 and anti-GFP antibodies. GYF1 residues 1–177 bound to 4EHP to a similar extent as the full-length protein, indicating that this protein fragment contains the principal 4EHP-binding region of the protein. (*C*) The interaction of GFP-GYF2 [either full-length, canonical mutant (C*) or N-terminal fragment (residues 1–180)] with V5-SBP-4EHP was analyzed as described in B. GYF2 residues 1–180 bound to 4EHP to a similar extent as the full-length protein. (*D*) Western blot showing the interaction of V5-SBP-4EHP (wild-type or the indicated mutants) with endogenous GYF1. The proteins were pulled down using streptavidin-coated beads. V5-SBP-MBP served as negative control. The inputs (1.5% for the V5-tagged proteins and 3% for GYF1) and bound fractions (3% for V5-tagged proteins and 35% for GYF1) were analyzed by western blotting using anti-V5 and anti-GYF1 antibodies. (*E, F*) Ni-NTA pulldown assays showing the interactions of GYF1 fragments (C+L+NC+A, C+L+NC and C) with 4EHP-His$_6$ (E) or eIF4E-His$_6$ (F). The eIF4E-binding region of 4E-BP1 (C+L+NC) binds similarly to both 4EHP and eIF4E, whereas GYF1 associates preferentially with 4EHP. The GYF1 and 4E-BP1 peptides contain an N-terminal MBP-tag and a C-terminal GB1 tag. The starting material (4% for the GYF1 fragments, 6% for 4EHP and recombinant eIF4E) and bound fractions (10% and 15% in panels E and F, respectively) were analyzed by SDS-PAGE followed by Coomassie blue staining. MBP served as a negative control. (*G*) The interaction of V5-SBP-tagged eIF4E or 4EHP proteins with endogenous GYF1, 4E-T and 4E-BP1 was analyzed in HEK293T cell

17

lysates. The proteins were pulled down using streptavidin-coated beads. Inputs (1.5%) and bound fractions (30% for 4E-BP1, GYF1 and 4E-T and 5% for the V5-SBP-tagged proteins) were analyzed by western blotting using anti-V5, anti-4E-BP1, anti-4E-T and anti-GYF1 antibodies. (*H, I*) Western blot analysis showing the interaction of GFP-tagged GYF1, GYF2 and 4E-T with HA-4EHP (H) or HA-eIF4E (I) in HEK293T cells. The proteins were immunoprecipitated using anti-GFP antibodies. The inputs (0.75% for GFP-tagged proteins and 0.5% for the HA-tagged proteins) and immunoprecipitates (15% for GFP-tagged proteins and 25% for HA-tagged proteins) were analyzed using anti-GFP and anti-HA antibodies, respectively. (*J*) Ni-NTA pulldown assay showing the interaction of 4EHP-His$_6$ (wild-type, W-A and IM-AA mutants) with GYF1 fragments with or without the auxiliary region (C+L+NC+A vs. C+L+NC). MBP served as a negative control. The starting material (4%) and bound fractions (9%) were analyzed by SDS-PAGE followed by Coomassie blue staining.

Figure S3

A GYF1 C + 4EHP

B GYF2 C + 4EHP

C GYF1 C+L+NC + 4EHP

D GYF2 C+L+NC + 4EHP

E GYF1 C+L+NC+A + 4EHP

F GYF2 C+L+NC+A + 4EHP

G 4EHP + m⁷GpppG cap

H 4EHP-GYF2 C+L+NC+A complex + m⁷GpppG cap

I GYF1 C+L+NC+A + 4EHP (dimerization mutants)

J GYF2 C+L+NC+A + 4EHP (dimerization mutants)

K 4E-BP1 C+L+NC + 4EHP

L 4E-BP1 C+L+NC + eIF4E

M 4E-BP1 C+L+NC + 4EHP (S99N mutant)
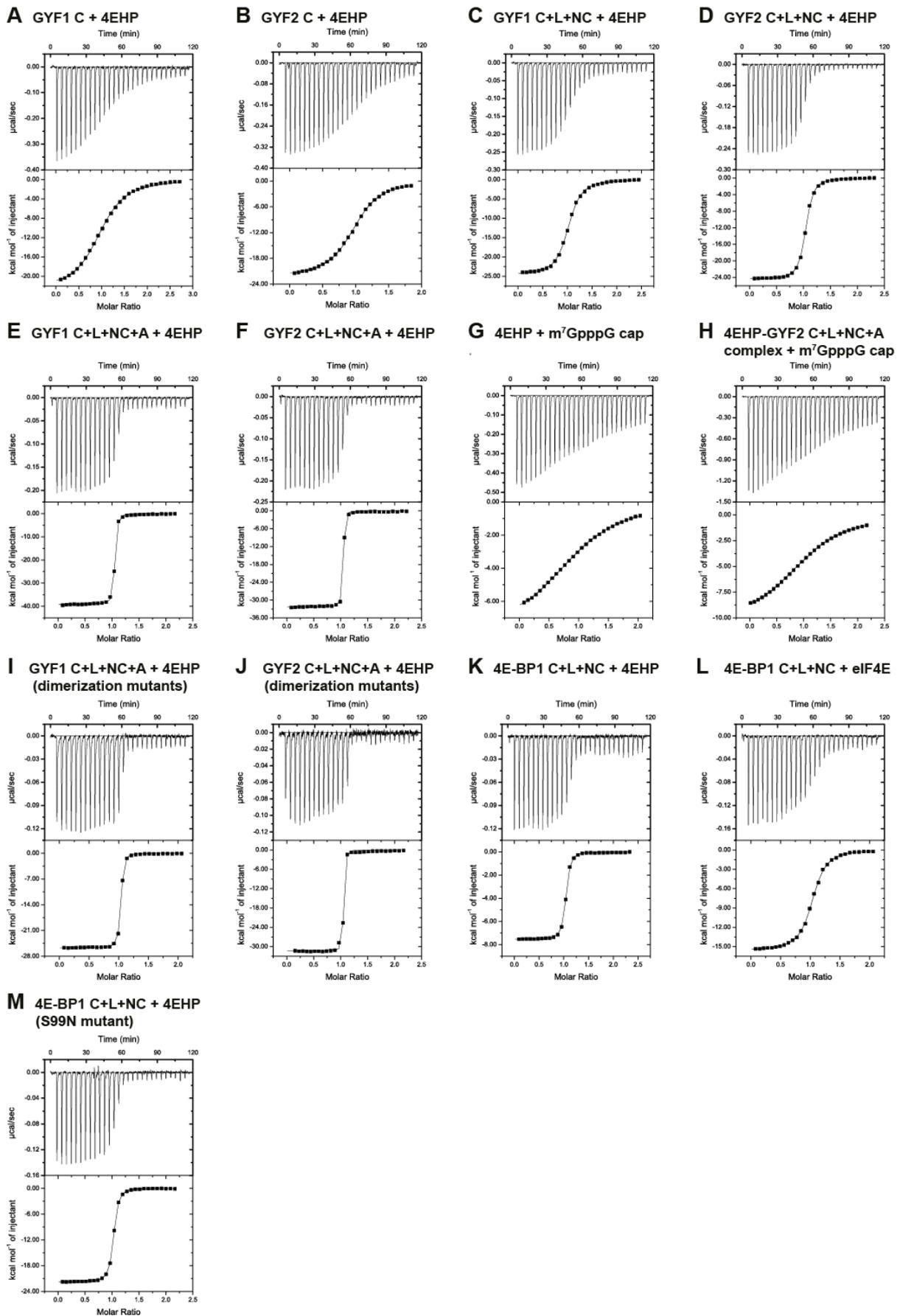
Supplemental Figure S3. Calorimetric titration data for the interaction of 4EHP with peptides derived from GYF1, GYF2 and 4E-BP1 or with m^{7}GpppG cap analog. (*A–F*) Isothermal titration calorimetry (ITC) profiles for the interaction of 4EHP (residues 52–234) with the following peptides: (A) GYF1 C; (B) GYF2 C; (C) GYF1 C+L+NC; (D) GYF2 C+L+NC; (E) GYF1 C+L+NC+A; (F) GYF2 C+L+NC+A. (*G*) ITC profile for the binding of 4EHP (residues 52-234) to m^{7}GpppG cap analog. (*H*) ITC profile for the binding of 4EHP-GYF2 C+L+NC+A complex to m^{7}GpppG cap analog. (*I, J*) ITC profiles for the interaction of 4EHP dimerization mutant with GYF1 and GYF2 (C+L+NC+A) dimerization mutant peptides. (*K - M*) ITC profiles for the interaction of 4E-BP1 (residues 50–83, C+L+NC) with the following proteins: (K) wild-type 4EHP (residues 52–234); (L) wild-type eIF4E (residues 36–217); (M) 4EHP (residues 52–234) S99N mutant. The thermodynamic parameters are shown in Table S2. Upper panels show raw data in (μcal sec^{-1}) and lower panels represent the integration of heat changes associated with each injection (kcal mol^{-1} of injectant). Data was fit using a one-site binding model.

# Figure S4

A 4EHP–GYF1 (C+L+NC+A) - ASU
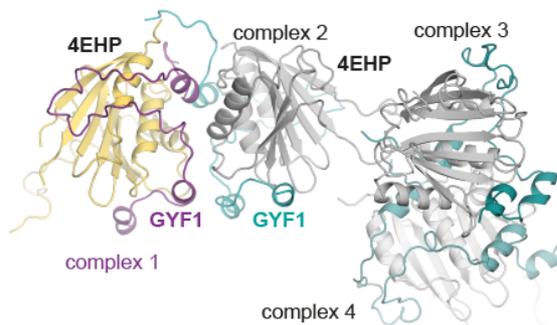
B 4EHP–GYF2 (C+L+NC+A) - ASU
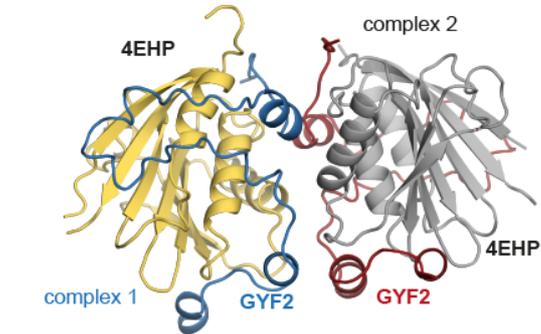
**C** 4EHP–4E-BP1 (C+L+NC) - ASU
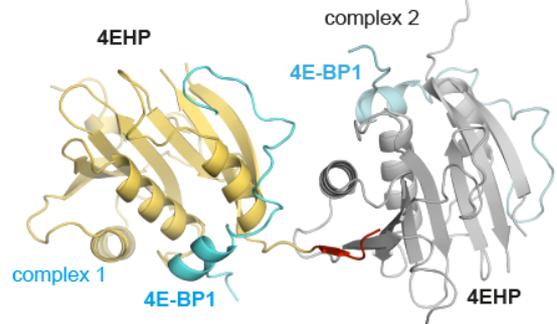
D 4EHP–GYF2 (C+L+NC) - ASU
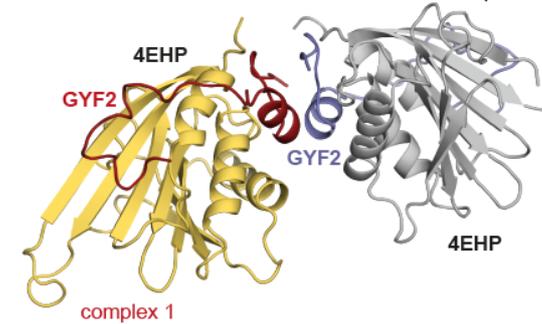
E 4EHP–4E-BP1 vs. eIF4E–4E-BP1
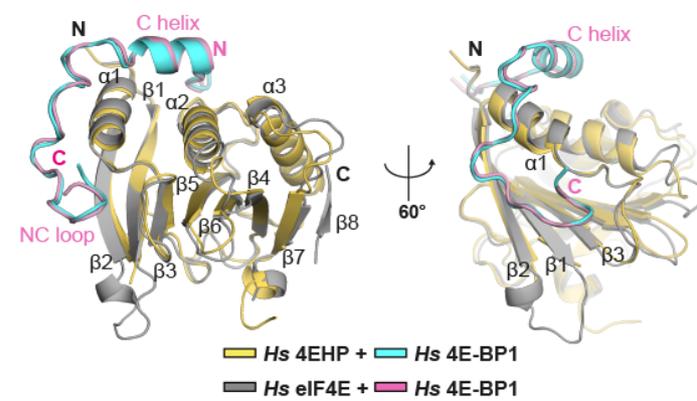
F 4E-BP1–eIF4E/4EHP
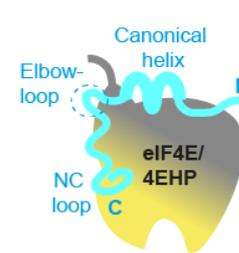
G 4EHP–GYF2 (C+L+NC+A) vs. 4EHP–GYF2 (C+L+NC)

H

Hs 4EHP + *Hs* 4E-BP1
Hs eIF4E + *Hs* 4E-BP1

Hs 4EHP + { *Hs* GYF2 C+L+NC+A
              Hs GYF2 C+L+NC

Hs 4EHP
GYF1
GYF2
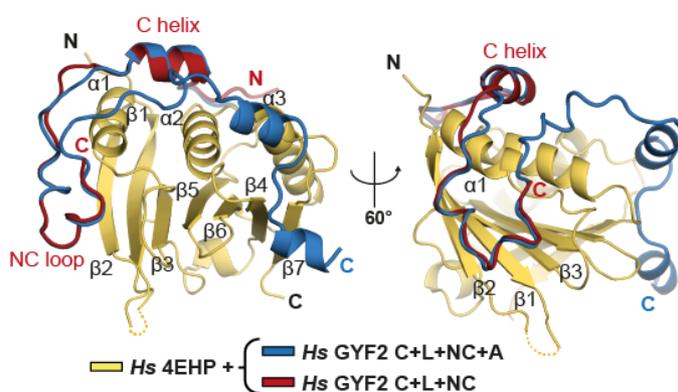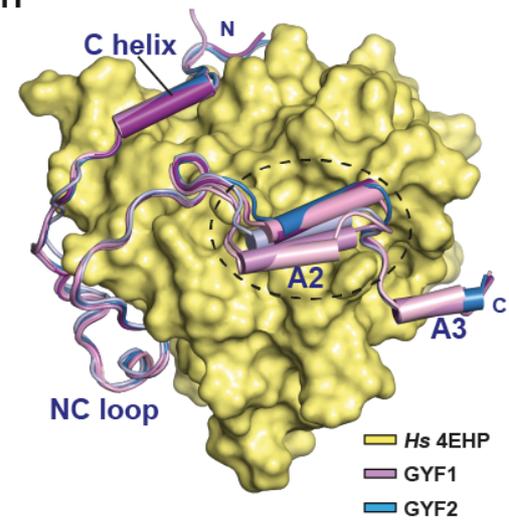
Supplemental Figure S4. Structures of 4EHP bound to GYF1, GYF2 and 4EBP1. (*A*)

Cartoon representation showing the asymmetric unit (ASU) of the 4EHP–GYF1 crystal

form. The ASU contains four 4EHP–GYF1 complexes. In complex 1, which was used for representation, GYF1 is colored in purple and 4EHP is colored in yellow; in the other complexes of the ASU GYF1 is colored in cyan and 4EHP in grey. (*B*) Cartoon representation showing the ASU of the 4EHP–GYF2 crystal form. The ASU contains two 4EHP–GYF2 complexes. In complex 1, GYF2 is colored in blue and 4EHP in yellow. In complex 2, GYF2 is colored in red and 4EHP in grey. (*C*) ASU of the 4EHP–4E-BP1 (C+L+NC) crystal form. There are two 4EHP–4E-BP1complexes in the ASU. In complex 1, 4E-BP1 is colored in cyan and 4EHP is colored in yellow. In complex 2, 4EHP is colored in grey. The N-terminal portion of the 4EHP molecule from complex 1 is colored in red and contains residues from the expression tag, which mediate contacts to complex 2. (*D*) Cartoon representation showing the ASU of the 4EHP–GYF2 (C+L+NC) crystal structure. The ASU contains two 4EHP–GYF2 complexes. In complex 1, GYF2 is colored in red and 4EHP yellow. In complex 2, GYF2 is colored in blue and 4EHP in grey. The structural arrangement of the two complexes that lack the GYF2 auxiliary sequences appears similar to the dimeric arrangement of the complexes containing the auxiliary sequences (panel B). (*E*) Superposition of the structure of 4E-BP1 (cyan) bound to 4EHP (yellow) to the structure of 4E-BP1 (magenta) bound to eIF4E (grey; PDB: 4UED, Peter et al. 2015a). Selected secondary structural elements in 4EHP are label in black. The structures superpose with an RMSD of 0.41 Å over 194 Cα atoms. (*F*) Schematic representation of eIF4E and 4EHP bound to 4E-BP1. (*G*) Superposition of the structure of 4EHP bound to GYF2 C+L+NC+A (blue) with the structure of 4EHP bound to GYF2 C+L+NC (red) peptides. Selected secondary structural elements in 4EHP are label in black. The structures superpose with an RMSD of 0.38 Å over 207 Cα atoms. (*H*) Overlay of all complex structures of 4EHP bound to GYF1 and GYF2 peptides to illustrate the conformational flexibility of helix α2 (A2), which is circled with a black dashed line. The surface of 4EHP is shown in pale yellow and

the GYF peptides are colored in purple and blue for GYF1 and GYF2, respectively. Helical secondary elements (canonical helix, helices α2 and α3) are represented as cylinders.

Figure S5
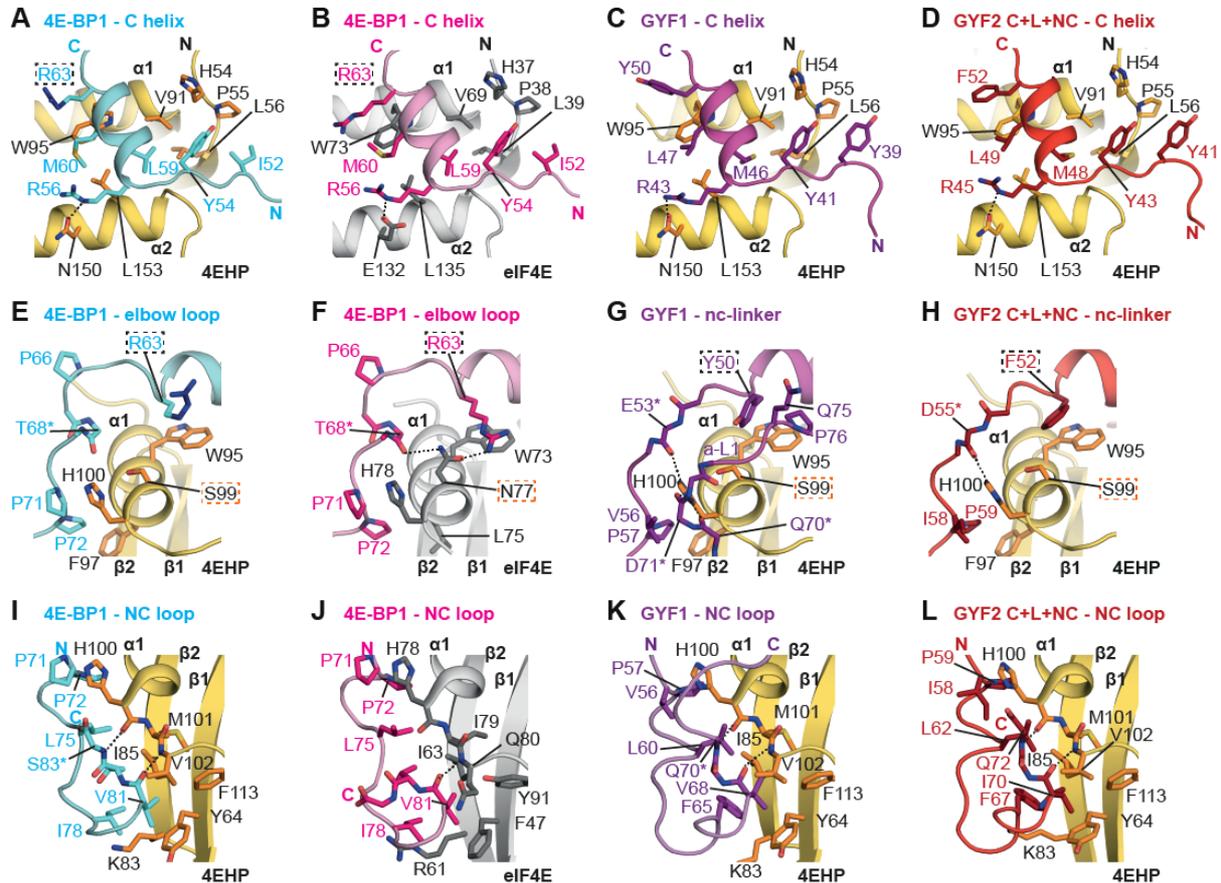
Supplemental Figure S5. Interactions of GYF1, GYF2 and 4E-BP1 with 4EHP and eIF4E. (*A–D*) Close-up views of the interaction between the dorsal surface of 4EHP (A,C,D) or eIF4E (B) and the canonical helices of 4E-BP1 (A, B; Peter et al. 2015a), GYF1 C+L+NC+A (C) and GYF2 C+L+NC (D). Selected residues are shown as sticks. Selected secondary structure elements are labeled in black for 4EHP or eIF4E and in color for the interacting partners. Residue R63[4E-BP1] is colored in dark blue following the Cγ atom in A and highlighted by a black dashed box in A and B. (*E–H*) Close-up views of the interaction between 4EHP (E,G,H) or eIF4E (F) and the non-canonical linkers of 4E-BP1 (E, F; Peter et al. 2015a), GYF1 C+L+NC+A (G) and GYF2 C+L+NC (H). Selected residues are shown as

sticks. Residue R63^{4E-BP1} is colored in dark blue following the Cγ atom in E and highlighted by a black dashed box in E and F. The corresponding residues in GYF1 (Y50) or GYF2 (F52) are also highlighted by a black dashed box. For visual clarity, only backbone atoms are shown for the residues labeled with an asterisk. The residues N77 in eIF4E and S99 in 4EHP are highlighted with orange dashed boxes. (*I–L*) Close-up views of the interaction between the lateral surface of 4EHP (I,K,L) or eIF4E (J) and the non-canonical loops of 4E-BP1 (I, J; Peter et al. 2015a), GYF1 C+L+NC+A (K) and GYF2 C+L+NC (L).

## Figure S6
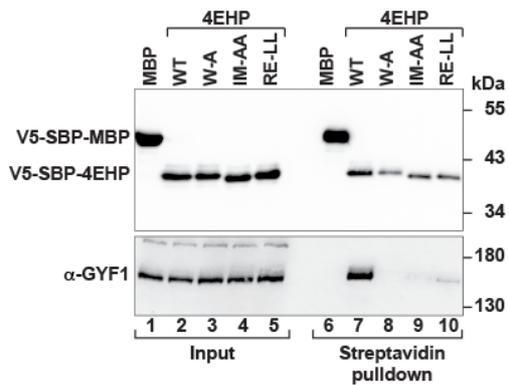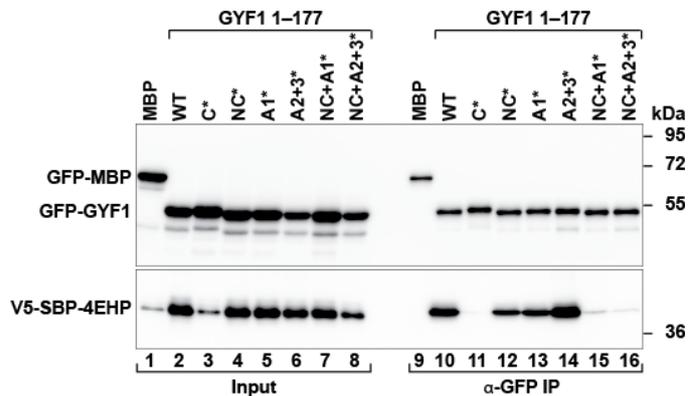
Supplemental Figure S6. Validation of the interfaces observed in the 4EHP–GYF1/2 and 4EHP–4E-BP1 complex structures. (*A*) Western blot analysis showing the interaction of endogenous GYF1 with V5-SBP-tagged 4EHP (WT or the indicated mutants). The proteins were pulled down using streptavidin-coated beads. V5-SBP-MBP served as negative control. The inputs (2.5%) and immunoprecipitates (3% for the V5-tagged proteins and 20% for GYF1) were analyzed by western blotting using anti-V5 and anti-GYF1 antibodies. (*B*) Interaction of GFP-GYF1 N-terminal fragment (residues 1–177; either wild-type or the indicated mutants) with V5-SBP-tagged full-length 4EHP. The proteins were immunoprecipitated from HEK293T cell lysates using anti-GFP antibodies. GFP-MBP served as negative control. The inputs (1.5% for the GFP-tagged proteins and 0.5% for V5-

SBP-4EHP) and immunoprecipitates (7.5% for the GFP-tagged proteins and 30% for V5-SBP-4EHP) were analyzed by western blotting using anti-GFP and anti-V5 antibodies.

Figure S7
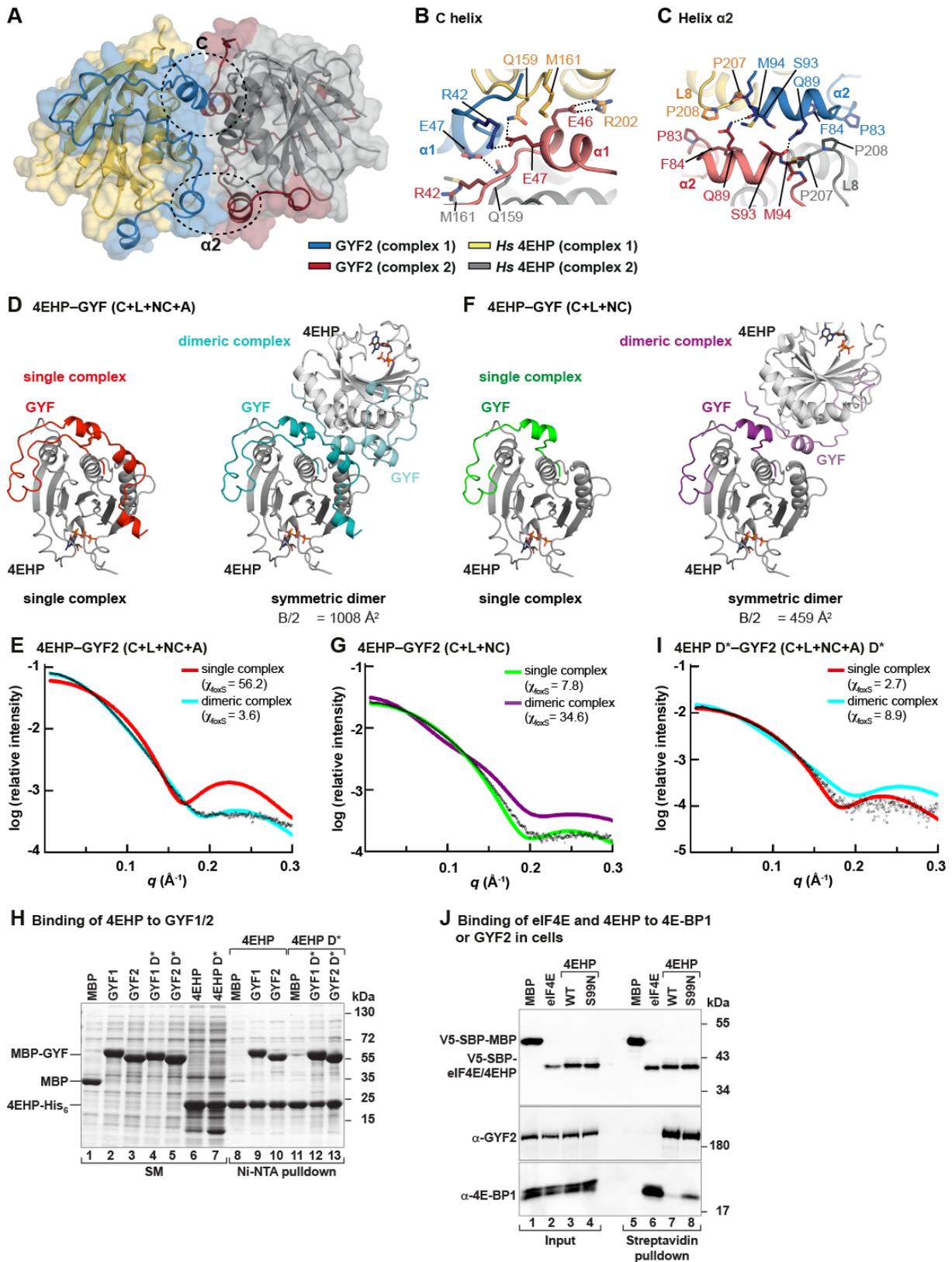
Supplemental Figure S7. The 4EHP–GYF1/2 complexes form dimers in solution. (*A*)

Representation of the dimeric arrangement of the 4EHP–GYF2 complexes in the asymmetric

unit of the crystal. In this dimeric arrangement, the two 4EHP–GYF2 complexes adopt a two-fold rotational symmetry with their dorsal surfaces facing each other. The interface of this arrangement involves: I) the canonical helix of GYF2, which is in contact with the neighboring 4EHP molecule and is extended by the canonical helix of the GYF2 molecule of the neighboring complex, and II) the auxiliary region of GYF2, where helix α2 of one complex faces, in an antiparallel fashion, helix α2 of the GYF2 and loop 8 of the 4EHP present in the neighboring complex. The contacts between the two complexes are highlighted with dashed black circles. In complex 1, GYF2 is colored in blue and 4EHP in yellow. In complex 2, GYF2 is colored in red and 4EHP in grey. (*B, C*) Close-up views on the dimeric interface involving the canonical helices (C helix, panel B) and the auxiliary helix 2 (helix α2, panel C). Selected residues are shown as sticks and colored as in panel A. GYF2 residues E46 and E47 within the canonical motifs of interacting GYF2 molecules contact the guanidinium group of R202[4EHP] and the side chain of Q159[4EHP] on neighboring 4EHP molecules. M161[4EHP] contacts the aliphatic portion of E46[GYF2] in the canonical helix of the neighboring complex. Pro residues in 4EHP loop L8 (P207 and P208) are facing residues proximal to the helix α2 of GYF2 from the neighboring complex (P83, F84, Q89[GYF2]). (*D, F*) Crystallographic models of the 4EHP–GYF2 complexes. The radius of gyration (R_G) and the maximum particle size (D_{max}) were calculated using Scatter and are summarized in Table S3. In the case of the dimeric assemblies, the dimer interface (B/2) was calculated using PISA from the CCP4 package and is indicated below the structures. 4EHP is shown in grey. The GYF2 (C+L+NC+A) peptide is colored in red and cyan in the single and dimeric arrangements, respectively. The GYF2 (C+L+NC) peptide is colored in green and purple in the single and dimeric arrangements, respectively. (*E, G, I*) Small-angle X-ray scattering profiles comparing single and dimeric arrangements of the 4EHP–GYF2 complexes with the experimental data. The data are plotted with the logarithmic scattering intensity on the y-axis

and the scattering angle q on the x-axis. Experimental scattering data of the complexes in solution are shown as open black circles and the fits for the single and dimeric arrangements are shown as a line colored as indicated on the right. The goodness-of-fit χ values, calculated using FoXS, are indicated for each fit. (*H*) Ni-NTA pulldown assay showing the interaction of 4EHP-His$_6$ (M1–F234, wild-type and dimerization mutant) with MBP-tagged GYF1 and GYF2 proteins [(wild-type and dimerization mutant (D*)]. MBP served as a negative control. The input (07% for MBP-tagged proteins and 2% for 4EHP) and bound fractions (9%) samples were analyzed by SDS-PAGE followed by Coomassie blue staining. (*J*) Analysis of the interaction of V5-SBP-tagged eIF4E, 4EHP and the 4EHP S99N mutant with endogenous GYF2 and 4E-BP1 proteins in HEK293T cell lysates. The proteins were pulled down using streptavidin-coated beads. V5-SBP-MBP served as negative control. Input samples (1% for 4E-BP1 and GYF2 and 1.5% for the V5-SBP-tagged proteins) and bound fractions (20% for 4E-BP1 and GYF2 and 5% for V5-SBP-tagged proteins) were analyzed by western blotting using anti-V5, anti-GYF2 and anti-4E-BP1 antibodies.

Figure S8

A R-Luc-5BoxB-A_{95}-MALAT1 mRNA

B

C R-Luc-A_{95}-MALAT1

D

E

F R-Luc-5BoxB-A_{95}-MALAT1 mRNA

G

H Binding of eIF4E and 4EHP to cap in cells

I Binding of 4EHP to GYF2 in cells
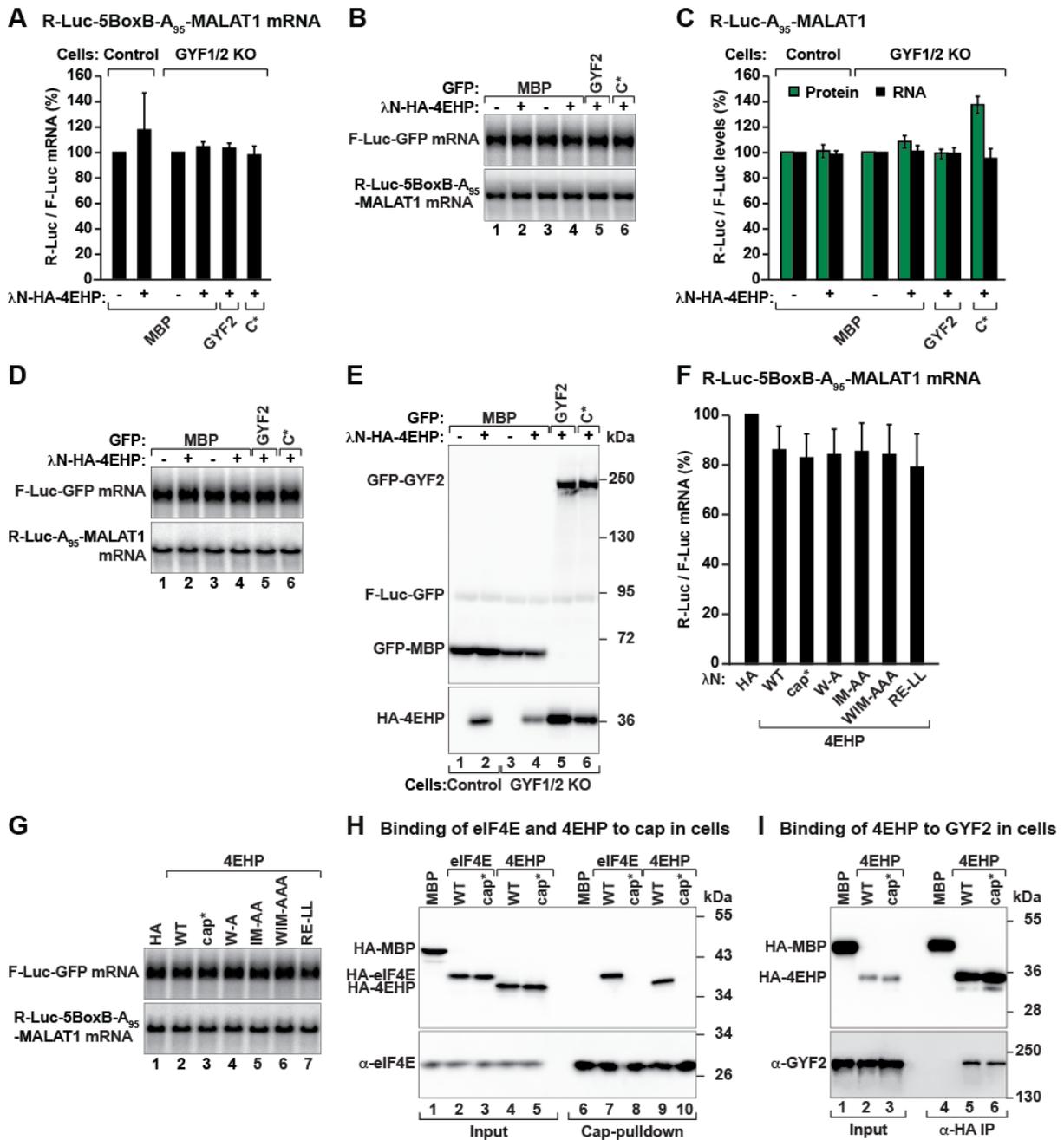
Supplemental Figure S8. 4EHP requires GYF1/2 proteins to repress the expression of bound mRNAs (*A*) A complementation assay using the R-Luc-5BoxB-A_{95}-MALAT1 reporter and λN-HA-4EHP (either wild-type or the indicated mutants) was performed in control and GYF1/2-null HEK293T cells expressing GFP-MBP or GFP-GYF2 (wild-type or canonical mutant, C*). A plasmid expressing F-Luc-GFP served as the transfection control. R-Luc activity and mRNA levels were normalized to those of the F-Luc transfection control

and set to 100% in cells expressing the λN-HA peptide. Normalized R-Luc activities are shown in Figure 6A. The panel shows the corresponding normalize R-Luc-5BoxB-A$_{95}$-MALAT1 mRNA levels. Bars represent the mean values and error bars represent standard deviations from three independent experiments. (*B*) Northern blot of representative RNA samples corresponding to the experiment shown in A and Fig. 6A. (*C*) Complementation assay in WT and GYF1/2-null cells using the R-Luc-A$_{95}$-MALAT1 reporter lacking the BoxB hairpins. A plasmid expressing F-Luc-GFP was used as a transfection control. R-Luc activity and mRNA levels were normalized to those of the F-Luc transfection control and set to 100% in cells expressing the λN-HA peptide. (*D*) Northern blot of representative RNA samples corresponding to the experiment shown in C. (*E*) Western blot analysis showing the equivalent expression of the λN-HA-4EHP and GYF2 proteins used in the complementation shown in C and D. (*F*) Normalized R-Luc-5BoxB-A$_{95}$-MALAT1 mRNA levels corresponding to the experiment shown in Fig. 6D,E. (*G*) Northern blot of representative RNA samples corresponding to the experiment shown in F. (*H*) Lysates from HEK293T cells expressing HA-tagged eIF4E or 4EHP (wild-type or cap mutant, cap*) were pulled down with m^{7}GTP-sepharose beads. Endogenous eIF4E served as positive control. Inputs (0.75% for the HA-tagged proteins and 1% for endogenous eIF4E) and bound fractions (15% for the HA-tagged proteins and 5% for endogenous eIF4E) were analyzed by Western blot using anti-HA and anti-eIF4E antibodies. (*I*) Interaction of HA-tagged 4EHP (wild-type or cap mutant, cap*) with endogenous GYF2 in HEK293T cells. HA-tagged MBP served as a negative control. Inputs (0.37% for the HA-tagged proteins and 0.75% for endogenous GYF2) and immunoprecipitates (15% for the HA-tagged proteins and 20% for endogenous GYF2) were analyzed by Western blot using anti-HA and anti-GYF2 antibodies.

Figure S9

A R-Luc-5BoxB-A_{95}-MALAT1

B

C R-Luc-A_{95}-MALAT1

D

E

F R-Luc-A_{95}-MALAT1

G
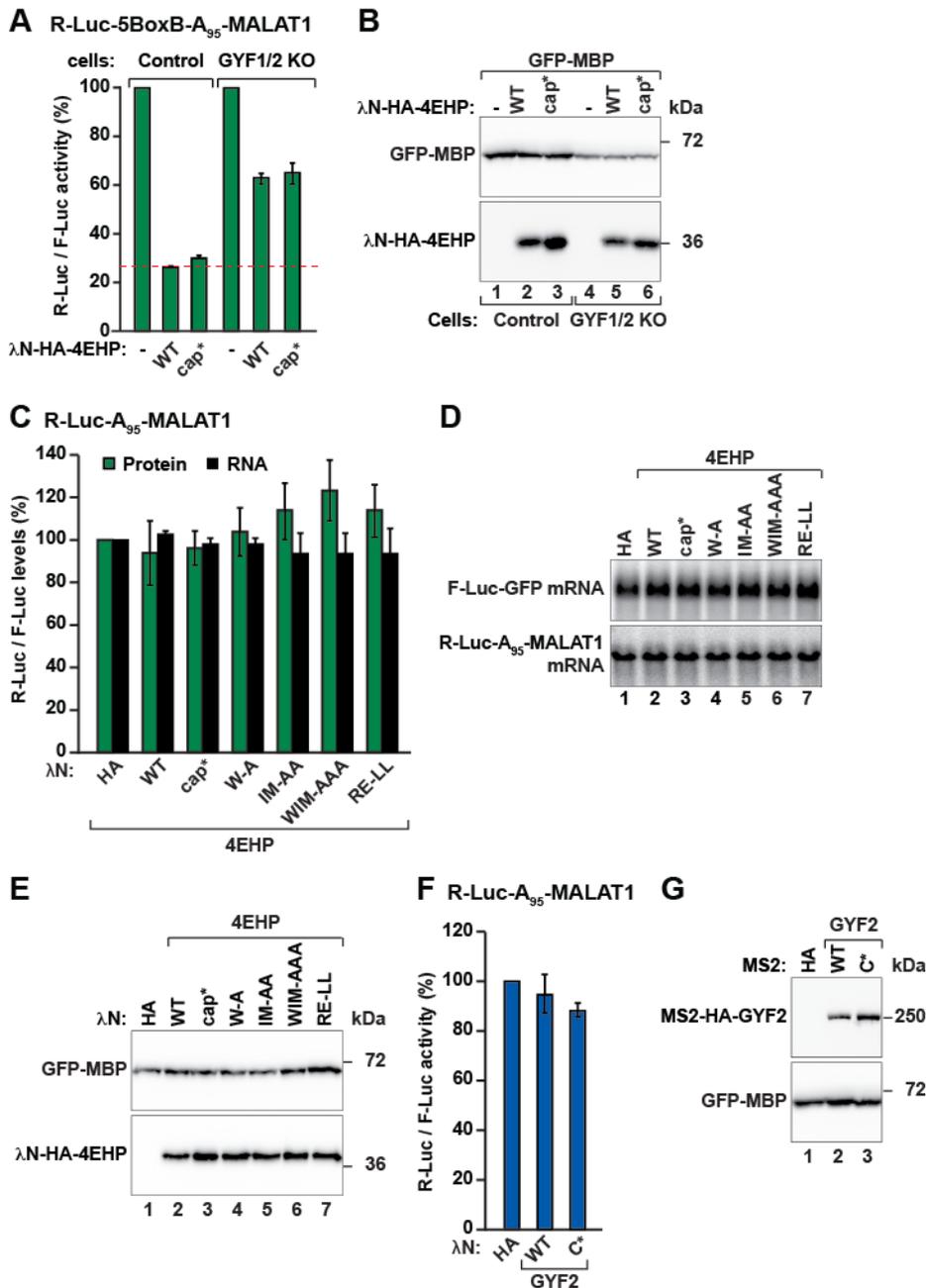
Supplemental Figure S9. 4EHP requires GYF1/2 proteins to repress translation of bound mRNAs (*A*) A complementation assay using the R-Luc-5BoxB-A_{95}-MALAT1 reporter and λN-HA-4EHP [either wild-type or cap mutant (cap*, W124A)] was performed in control and GYF1/2-null HEK293T cells expressing GFP-MBP or GFP-GYF2 (wild-type or canonical mutant). A plasmid expressing F-Luc-GFP served as the transfection control. For each cell type, R-Luc activity was normalized to that of the F-Luc transfection control and set

to 100% in cells expressing the λN-HA peptide. Samples were analyzed as described in Fig. 6A. (*B*) Western blot showing similar expression of the proteins used in A. (*C*) Tethering assay using the R-Luc-A$_{95}$-MALAT1 reporter and λN-HA-4EHP (wild-type or mutants) in HEK293T cells. Samples were analyzed as described in Fig. 6A. (*D*) Northern blot of representative RNA samples corresponding to the experiment shown in C. (*E*) Western blot showing the equivalent expression of the λN-HA-4EHP proteins used in the tethering assay shown in D.

(*E*) Tethering assay using the R-Luc-A$_{95}$-MALAT1 reporter lacking MS2 binding sites and MS2-HA-GYF2 (wild-type or canonical mutant). Samples were analyzed as described in Fig. 6F. The corresponding assay with the reporter containing the MS2 binding sites is shown in Fig. 6F. (*G*) Western blot analysis showing the expression of the GFP-GYF2 proteins used in the tethering assays shown in F.

Supplemental references

Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH, Adams PD. 2012. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* **68:** 352–367.

Chen VB, Arendall WB, 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66:** 12–21.

Cheng Y, Patel DJ. 2004. An efficient system for small protein expression and refolding. *Biochem Biophys Res Commun* **317:** 401–405.

Diebold ML, Fribourg S, Koch M, Metzger T, Romier C. 2011. Deciphering correct strategies for multiprotein complex assembly by co-expression: application to complexes as large as the histone octamer. *J Struct Biol* **175:** 178–188.

Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66:** 486–501.

Fabian MR, Frank F, Rouya C, Siddiqui N, Lai WS, Karetnikov A, Blackshear PJ, Nagar B, Sonenberg N. 2013. Structural basis for the recruitment of the human CCR4-NOT deadenylase complex by tristetraprolin. *Nat Struct Mol Biol* **20**: 735-739.

Kabsch W. 2010. Xds. *Acta Crystallogr D Biol Crystallogr* **66:** 125–132.

Kuzuoglu-Ozturk D, Bhandari D, Huntzinger E, Fauser M, Helms S, Izaurralde E. 2016. miRISC and the CCR4-NOT complex silence mRNA targets independently of 43S ribosomal scanning. *EMBO J* **35:** 1186–1203.

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. 2007. Phaser crystallographic software. *J Appl Crystallogr* **40:** 658–674.

Mizoue LS, Tellinghuisen J. 2004. The role of backlash in the "first injection anomaly" in isothermal titration calorimetry. *Anal Biochem* **326:** 125–127.

Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. 2013. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* **8:** 2281–2308.

Schneidman-Duhovny D, Hammel M, Tainer JA, Sali A. 2013. Accurate SAXS profile computation and its assessment by contrast variation experiments. *Biophys J* **105:** 962–974.

Terwilliger TC, Grosse-Kunstleve RW, Afonine PV, Moriarty NW, Zwart PH, Hung LW, Read RJ, Adams PD. 2008. Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr D Biol Crystallogr* **64:** 61–69.