The role of the RNA-associated proteins NOT4 and Bam in mRNA decay

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Csilla Keskeny

aus Budapest, Ungarn

Tübingen

2019

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

Tag der mündlichen Qualifikation:	01.10.2019
Dekan:	Professor Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Professor Dr. Ralf-Peter Jansen
2. Berichterstatter:	Professor Dr. Thorsten Stafforst

The work described in this thesis was conducted in the laboratory of Professor Dr. Elisa Izaurralde in the Department of Biochemistry at the Max Planck Institute for Developmental Biology, Tübingen, Germany, from September 2014 to January 2019. It was further supervised by Professor Dr. Ralf-Peter Jansen at the Eberhard Karls University Tübingen, Germany, and was supported by a fellowship of the Max Planck Society within the framework of the International Max Planck Research School “From Molecules to Organisms”. I declare that this thesis is the product of my own work. The parts that have been published or where other sources have been used were cited accordingly. Work that was carried out by my colleagues was also indicated accordingly.

Table of Contents

1. Abbreviations

ABCE1	ABC-family ATPase 1
AGO	Argonaute protein
AREs	AU (adenylate-uridylate)-rich elements
Bam	Bag-of-marbles protein
Bgcn	Benign gonial cell neoplasm
BTG2	B-cell translocation gene 2
CAF1/40/130	CCR4-associated factor 1/40/130
CBM	CAF40-binding motif of Roquin, Bam and NOT4
CC	Coiled coil
CCCH	Zinc finger motif containing Cys-Cys-Cys-His, which coordinate a zinc ion
CCR4	Carbon catabolite repressor 4
CCR4-NOT	Carbon catabolite repressor 4; Negative on TATA-less
CN9BD	CAF40/CNOT9-binding domain of NOT1
DCP2	Decapping enzyme subunit 2
DDX6	DEAD-box helicase 6
DEAD-box	Helicase domain characterized by a Asp-Glu-Ala-Asp tetrad
DEDD	Exonuclease domain characterized by a catalytic Asp-Glu-Asp-Asp tetrad
Dm, D. melanogaster	*Drosophila melanogaster*
DNA	Deoxyribonucleic acid
dsRNA	Double-stranded RNA
EEP	Exonuclease-endonuclease-phosphatase domain
eIF	Eukaryotic translation initiation factor
EJC	Exon junction complex
eRFs	Eukaryotic release factors
GSC	Germline stem cell
GW182	182kDa protein containing glycine-tryptophan repeats
HEAT	Helical hairpin repeats
HEK293T	Human embryonic kidney 293T cells
HeLa	Henrietta Lacks; immortal cell line derived from cervical cancer cells
Hs	*Homo sapiens*; human
kDa	kilodalton

LRR	Leucin-rich repeat domain
LTN1	E3 ubiquitin-protein ligase listerin 1
Mei-P26	Meiotic-P26
Met-tRNA$_i^{met}$	Initiator methionine transfer RNA
MIF4G	Middle domain of eukaryotic initiation factor 4G
miRISC	MicroRNA-induced silencing complex
miRNA	MicroRNA
mRNA	Messenger RNA
mRNP	Messenger-ribonucleoprotein particle
NAC	Nascent polypeptide-associated complex
NGD	No-go decay
NMD	Nonsense-mediated decay
NOT1	Negative on TATA-less protein 1
NSD	Nonstop decay
ORF	Open reading frame
PABP	Poly(A)-binding protein
PAN2/3	Poly(A) nuclease 2/3
PAP	Poly(A) polymerase
PIC	Pre-initiation complex
Poly(A)	Poly-adenosine stretch
pre-mRNA	Precursor messenger RNA
PTM	Post-translational modification
RBP	RNA-binding protein
RING	Really interesting new gene; E3 ubiquitin ligase domain type
RNA	Ribonucleic acid
RNA pol II	RNA polymerase II
RNP	Ribonucleoprotein
RRM	RNA-recognition motif
rRNA	Ribosomal RNA
Sc, S. cerevisiae	*Saccharomyces cerevisiae,* baker's yeast / budding yeast
SHD	NOT1 superfamily homology domain
shRNA	Short hairpin RNA
SLiM	Short linear motif
Sp, S. pombe	*Schizosaccharomyces pombe*, fission yeast

Sxl	Sex-lethal
TNRC6	Trinucleotide repeat-containing gene 6 protein, human homolog of GW182
Tob	Transducer of erbB-2 protein
TRAMP	Trf4/Air2/Mtr4p polyadenylation
tRNA	Transfer RNA
TTP	Tristetraprolin
Tut	Tumorous testis protein
UBC	Ubiquitin conjugating protein
UTR	Untranslated region
XRN1	5’-to-3’ exoribonuclease 1
ZnF	Zinc finger domain
ZNF598	Zinc finger protein 598

2.1 Summary

The multi-subunit CCR4-NOT complex is an important regulator of gene expression in eukaryotes. It affects most steps of the messenger RNA (mRNA) lifecycle, but as the major deadenylase in cells, its most studied function is the enzymatic removal of the 3'-poly(A) tail of mRNAs. In the past years, several studies were carried out to shed light on how the CCR4-NOT complex is directed to specific mRNA targets, and to identify proteins that regulate this process. The main aim of my doctoral work was to study metazoan-specific aspects in the assembly of the CCR4-NOT complex and to understand molecular details of its recruitment to mRNA by mRNA-associated proteins.

The NOT4 E3 ubiquitin ligase is highly conserved in eukaryotes and has been shown to associate with the CCR4-NOT complex in budding yeast. Biochemical and structural studies also demonstrated that the C-terminal region of yeast NOT4 interacts with the C-terminal part of NOT1. In contrast, different studies indicated that NOT4 is not a stable component of the CCR4-NOT complex in human (*Hs*) and *Drosophila melanogaster* (*Dm*) cells. During my doctoral studies, I examined the interactions of human NOT4 with components of the CCR4-NOT complex. My studies indicate that NOT4 directly interacts with the NOT1 and CAF40 subunits of the deadenylase complex *in vitro*. The interaction of *Hs* NOT4 with CAF40 is mediated by a conserved CAF40-binding motif (CBM). In addition, *Hs* NOT4 elicits 5'-to-3' decay of bound mRNAs through its interaction with CCR4-NOT. Importantly, depletion of CAF40 abolishes the interaction between NOT4 and the CCR4-NOT complex and impairs the ability of NOT4 to elicit decay of bound mRNAs.

In a collaborative effort among different members of the laboratory, we have also identified a CBM in the *D. melanogaster* protein Bag-of-marbles (Bam). As in the case of NOT4, Bam CBM allows the recruitment of the CCR4-NOT complex via CAF40. Thus, in the absence of CAF40, Bam does not interact with the CCR4-NOT complex.

The findings described in this thesis support a model of targeted deadenylation where RNA-associated proteins recruit the CCR4-NOT complex to specific mRNAs, resulting in the deadenylation, translational repression and degradation of a wide range of transcripts. NOT4 and Bam bind to the same surface of CAF40 as a previously characterized RNA-associated protein, Roquin, which highlights CAF40 as an important platform for the regulated recruitment of the CCR4-NOT complex.

2.2 Zusammenfassung

Der CCR4-NOT Multi-Proteinkomplex ist ein wichtiger Regulationsfaktor der Genexpression in Eukaryonten. Aufgrund seiner bedeutenden Rolle als zelluläre mRNA Deadenylase ist der enzymatischen Abbau des 3‘-poly(A) Schwanzes von mRNAs durch CCR4-NOT am besten untersucht; darüber hinaus beeinflusst der Komplex aber auch die meisten anderen Bereiche des mRNA Metabolismus. In den vergangenen Jahren wurden einige Studien durchgeführt die untersuchten wie der CCR4-NOT Komplex zu bestimmten mRNA Molekülen rekrutiert wird und welche Proteine bei der Regulation dieses Prozesses beteiligt sind. Der Fokus meiner Doktorarbeit war die Untersuchung der Metazoa-spezifischen Aspekte der Zusammensetzung des CCR4-NOT Komplexes in Metazoa, und die Analyse der molekularen Zusammenhänge seiner Rekrutierung zu mRNAs durch mRNA-assoziierte Proteine.

Die E3 Ubiquitin Ligase NOT4 ist in Eukaryonten hochkonserviert und in Bäckerhefe an den CCR4-NOT Komplex gebunden. Biochemische und strukturbiologische Untersuchungen haben ergeben, dass dabei die C-terminale Region des Hefe NOT4 Proteins mit dem C-Terminus von NOT1 interagiert. Im Gegensatz dazu ergaben Studien in humanen (*Hs*) und *Drosophila melanogaster* (*Dm*) Zellen, dass NOT4 in diesen Organismen keinen stabilen Bestandteil des CCR4-NOT Komplexes bildet. Im Rahmen meiner Doktorarbeit habe ich die Interaktion von humanem NOT4 mit Komponenten des CCR4-NOT Komplexes untersucht. Die Ergebnisse meiner Arbeit zeigen, dass NOT4 *in vitro* direkt mit den Untereinheiten NOT1 und CAF40 des Deadenylase Komplexes interagiert. Die Bindung von *Hs* NOT4 an CAF40 erfolgt über ein konserviertes CAF40-Bindemotiv (CBM). Des Weiteren induziert *Hs* NOT4 durch seine Interaktion mit dem CCR4-NOT Komplex den 5‘-3‘ Abbau gebundener mRNAs. Eine Reduktion der Menge an zellulärem CAF40 blockiert die Bindung von NOT4 an CCR4-NOT und vermindert dementsprechend den NOT4-vermittelten Abbau gebundener mRNAs.

In Zusammenarbeit mit Kollegen gelang es außerdem, auch in dem *D. melanogaster* Protein „Bag-of-marbles“ (Bam) ein CBM zu identifizieren. Wie bei NOT4 ermöglicht das CBM in Bam über die Bindung an CAF40 eine Rekrutierung des CCR4-NOT Komplexes. Dementsprechend kann Bam in Abwesenheit von CAF40 nicht mit dem CCR4-NOT Komplex interagieren.

Die Ergebnisse meiner Arbeit lassen sich in einem Model zusammenfassen, in dem RNA-assoziierte Proteine den CCR4-NOT Komplex zu bestimmten mRNAs rekrutieren und

dadurch die Deadenylierung, Translationshemmung und den Abbau zahlreicher Transkripte regulieren. Interessanterweise binden NOT4 und Bam an dieselbe Oberfläche von CAF40 wie das zuvor untersuchte RNA-assoziierte Protein Roquin, was auf eine wichtige Rolle von CAF40 bei der geregelten Rekrutierung des CCR4-NOT Komplexes hindeutet.

3. List of publications

1. A conserved CAF40-binding motif in metazoan NOT4 mediates association with the CCR4-NOT complex

Csilla Keskeny, Tobias Raisch, Annamaria Sgromo, Catia Igreja, Dipankar Bhandari, Oliver Weichenrieder and Elisa Izaurralde (2019)

Genes and Development 33(3-4):236-52

2. Drosophila Bag-of-marbles directly interacts with the CAF40 subunit of the CCR4-NOT complex to elicit repression of mRNA targets

Annamaria Sgromo, Tobias Raisch, Charlotte Backhaus, Csilla Keskeny, Vikram Alva, Oliver Weichenrieder and Elisa Izaurralde (2018)

RNA 24(3):381-95

4. Introduction

4.1 Gene expression in eukaryotes

In eukaryotic cells, the genetic information encoded in the DNA is tightly packaged in the nucleus in chromatin complexes, which mainly consist of histone proteins with the DNA strands wound around them. Chromatin structure is established and regulated by many proteins, including histones, histone modifying enzymes and chromatin remodeling complexes. The information in the DNA is stored or accessed as needed, and these often dynamic changes are enabled by epigenetic modifications such as DNA methylation and histone modifications result in changes to the chromatin structure, rendering the DNA more accessible or less so for enzymes.

Gene expression is a highly controlled multi-step process (Figure 1). The transcription of DNA into messenger RNAs (mRNAs), a mediator molecule for protein production, is carried out by RNA polymerase II. The product of transcription is a precursor-mRNA (pre-mRNA), which undergoes extensive processing in the nucleus to become mature mRNA (Moore 2005).

The processing steps that pre-mRNAs undergo can greatly impact the information contained in the mRNA molecule. The first step is the co-transcriptional addition of a 7-methyl guanosine cap structure to the 5’ end of the emerging pre-mRNA. Proteins of the cap-binding complex assemble on the cap and contribute to further processing and nuclear export of the mature mRNA. The cap also plays a crucial role for mRNA translation in the cytosol (Ramanathan et al. 2016).

Splicing is another processing event, which removes intronic sequences (noncoding intervening sequences in genes) from the pre-mRNA. Alternative splicing mechanisms allow the cell to produce several different proteins from the same gene. mRNA splicing is performed by the spliceosome, which consists of small nuclear RNAs (snRNAs) and proteins. These snRNAs base-pair with consensus RNA sequences to recognize splice sites. After splicing, proteins of the exon junction complex (EJC) mark the site of each splicing event (Shi 2017; Baralle and Baralle 2018).

As RNA polymerase II reaches the end of a gene, signals marking the 3’ end of the emerging pre-mRNA are recognized by RNA-binding proteins and RNA-processing enzymes. The RNA molecule is cleaved, and a stretch of approximately 200-250 adenosines is added to the

3’ end of the RNA by poly(A) polymerase (PAP) in a template-free manner in higher eukaryotes. Exceptions to this are histone mRNAs, which terminate in a highly conserved stem-loop structure (Dominski and Marzluff 2007). For polyadenylated mRNAs, despite the higher number of initially added adenosines, median tail length in higher eukaryotes is around 50-100 adenosines due to "pruning" of the poly(A) tail (Chang et al. 2014; Subtelny et al. 2014; Lima et al. 2017). The cap structure and the poly(A) tail protect the mRNA from exonucleolytic decay and promote mRNA translation. Poly(A)-binding proteins (PABPs) assemble onto the newly added poly-A tail, which together with the cap-binding complex and the exon junction complexes signal that processing of the mRNA is complete (Moore and Proudfoot 2009).

Mature mRNAs are exported from the nucleus through channels in the nuclear membrane, the nuclear pores. Export of most mRNAs in complex with proteins (messenger ribonucleoproteins, mRNPs) requires the binding of the nuclear export receptor NXF1/NXT1, which shuttles the mRNA through the nuclear pore via direct interactions with nucleoporins (Carmody and Wente 2009).

Some mRNAs are positioned at intracellular locations to establish spatially restricted protein synthesis, which is a key mechanism of post-transcriptional regulation. mRNA localization is especially important for some cell types where protein production needs to be spatially controlled (e.g. in highly polarized cells such as neurons) (Jansen 2001; Martin and Ephrussi 2009).

Apart from the coding sequences, mRNAs contain 5’ and 3’ untranslated regions (UTRs). These UTRs contain sequences with regulatory roles that influence mRNA localization, stability and translation efficiency (Fabian et al. 2010; Leppek et al. 2018).

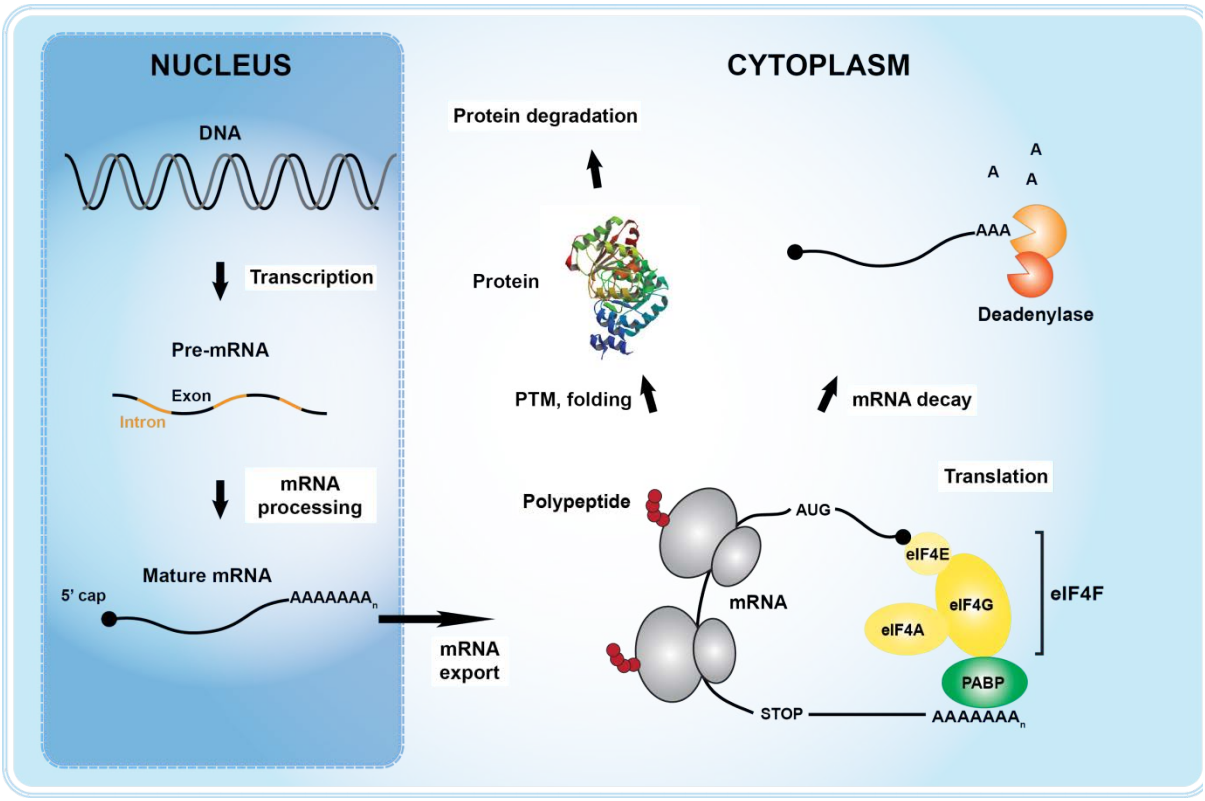


Figure 1. Schematic overview of eukaryotic gene expression. DNA is transcribed in the nucleus into pre-mRNA, which undergoes processing and is subsequently exported to the cytoplasm. The mRNA is translated by ribosomes into a protein, which is folded and post-translationally modified (PTM) to be fully functional. The template mRNA is degraded.

Once the mature mRNA is exported to the cytosol, ribosomes decode the mRNA with the help of transfer RNAs (tRNAs) in a process called translation. Ribosomes are multiprotein complexes that consist of two subunits, the small 40S and the large 60S subunits. Each subunit contains ribosomal RNAs (rRNAs) and multiple ribosomal proteins. Ribosomes interact with tRNAs that act as adaptors between the mRNA and the amino acids they deliver to synthesize the new polypeptide. Initiation of cap-dependent translation and recruitment of the small ribosomal subunit requires the assembly of the cap-binding complex eIF4F that consists of eukaryotic initiation factors (eIFs): the cap-binding protein eIF4E, the DEAD-box RNA helicase eIF4A, and the scaffolding protein eIF4G (Jackson et al. 2010). The eIF4F cap-binding complex recognizes the 5’ cap of the mRNA and initiates the assembly of the translation machinery by recruiting the 43S preinitiation complex (43S PIC) to the mRNA through interactions with the eIF3 complex (Wells et al. 1998; Jackson et al. 2010). The 43S PIC contains the 40S ribosomal subunit and initiation factors eIF1, eIF1A, eIF5, eIF3, and the eIF2-methionyl-tRNA ($tRNA_i^{Met}$)-GTP ternary

complex (Hinnebusch 2017). In addition, eIF4G binds to PABP, which was generally thought to lead to the formation of a closed-loop, circularized structure that is essential for efficient translation. However, recent observations challenge this model, and suggest there may be other explanations for the proximity and communication between the 5' and the 3' ends of the mRNA than the cap-eIF4E-eIF4G-PABP-poly(A) interaction network (Vicens et al. 2018).

Upon recruitment of the 43S PIC by the eIF4F complex, the 48S ribosomal translation initiation complex is formed, and the 5'-UTR is scanned until an AUG translation start codon is found. Binding to the start codon triggers the GTP-dependent dissociation of the initiation factors and subsequent recruitment of the 60S ribosomal subunit and assembly of the 80S ribosome so translation elongation can take place (Bhat et al. 2015).

During the elongation phase of translation, tRNAs with anticodon complementarity to the mRNA codons translocate between the tRNA binding sites on the 80S ribosome, adding an amino acid to the nascent polypeptide chain in each cycle (Dever et al. 2018). This phase is dependent on eukaryotic elongation factors. The tRNA binding sites are the A (aminoacyl or acceptor) site, the P (peptidyl) site and the E (exit) site. The process starts with the initiator $tRNA_i^{Met}$ bound to the P site of a ribosome positioned at an AUG start codon. Each new peptide chain elongation cycle begins with tRNA selection, where an aminoacyl-tRNA harboring the anticodon corresponding to the nucleotide triplets of the mRNA binds to the A site on the ribosome. Next, peptide-bond formation occurs, the tRNAs are positioned in a 'hybrid' state with respect to the ribosome subunits, and the growing peptide chain is transferred to the A site tRNA. Subsequently, the tRNA is repositioned from the hybrid state to a 'classical' state, creating an open A site for the next incoming aminoacyl-tRNA (Schuller and Green 2018). Finally, the ribosome then translocates to decode the next codon on the mRNA, and the deacylated tRNA moves from the P site to the E site (Steitz 2008). Recent findings established that codon optimality contributes greatly to translation elongation speed and mRNA stability (Presnyak et al. 2015).

Translation termination occurs when a stop codon is located in the A site of the ribosome, resulting in the recruitment of termination factors (eRFs). These factors promote the release of the nascent peptide and recycling of the ribosome. The ABC-family ATPase ABCE1 is also involved in this process by inducing dissociation of the ribosome into the 60S subunit and the

mRNA- and tRNA-bound 40S ribosomal subunit (Pisarev et al. 2010; Pisareva et al. 2011; Shoemaker and Green 2011).

Aberrant mRNA substrates cause problems during translation elongation, resulting in translationally stalled ribosomes, which engage the co-translational quality control machinery. Ribosome-associated quality control (RQC) mechanisms sense arrested ribosomes, which collide on the mRNA substrate, and rescue them by releasing the mRNA and nascent peptide for degradation and recycling the ribosomal subunits (Joazeiro 2017). One important effector protein of RQC in human cells is the ubiquitin ligase ZNF598 (also known as Hel2 in yeast), which serves as a direct sensor of two collided ribosomes by recognizing the 40S di-ribosome interface. Ubiquitination of ribosomal proteins by ZNF598 is essential for eliciting quality control pathways to terminally arrest the translation of aberrant transcripts, resolve the defective translational complexes and recycle or degrade their components (Juszkiewicz et al. 2018).

Ribosomal stalling leads to the endonucleolytic cleavage and subsequent 5'-to-3' or 3'-to-5' mRNA degradation. There are several pathways for the degradation of defective mRNA: nonsense-mediated mRNA decay (NMD) targets mRNAs with premature stop codons, nonstop mRNA decay (NSD) degrades mRNAs lacking stop codons, and no-go decay (NGD) acts on mRNAs containing secondary structure elements (e.g. stem-loops) or otherwise inhibiting efficient elongation. These pathways are conserved in eukaryotes, and they rely on protein synthesis and ribosome stalling to reveal mRNA defects (Lykke-Andersen and Bennett 2014).

4.2 mRNA turnover

mRNA turnover is an important aspect of gene expression control. Measurement of mRNA half-lives of almost 20 000 genes in mouse embryonic stem cells indicated that the median estimated half-life for all mRNAs was 7.1 hours (Sharov et al. 2008). Stable mRNAs with half-lives over 12 hours were coding for proteins involved in metabolism and protein biosynthesis, as well as genes encoding for proteins of the extracellular matrix and the cytoskeleton. The mRNAs with half-lives shorter than 2 hours code for proteins involved in transcriptional regulation (such as transcription factors), cell cycle, apoptosis, signal transduction and development. In this study, only 54 mRNAs had half-lives shorter than one hour (Sharov et al. 2008). In contrast, the determination of mRNA half-lives in primary human T-lymphocytes registered hundreds of mRNAs with half-lives shorter than 30 min (Raghavan et al. 2002). These

short-lived transcripts also encode regulatory proteins, including cytokines, cell surface receptors and transcription factors, and their rapid degradation may be an important mechanism to control levels of the encoded regulatory proteins (Raghavan et al. 2002).

Measurements of poly(A) tail length using poly(A) tail sequencing (TAIL-seq) in HeLa and NIH3T3 cells suggested that the length of the poly(A) tail slightly correlated with mRNA half-life, but not with translation efficiency (Chang et al. 2014). Moreover, poly(A)-tail length profiling by sequencing (PAL-seq) showed that poly(A) tail length negatively correlated with translation rates in yeast and in mouse NIH3T3 cells (Subtelny et al. 2014). In *C. elegans*, abundant transcripts had short poly(A) tails as determined using an adapted TAIL-seq method (Lima et al. 2017). These transcripts were also enriched for optimal codons and had high ribosome occupancy, indicating higher translational efficiency (Lima et al. 2017). These findings challenge the model that longer poly(A) tails promote mRNA stability and translation efficiency, which was generally accepted before (Goldstrohm and Wickens 2008; Wahle and Winkler 2013).

Two main mRNA decay pathways control eukaryotic mRNA levels: the 5'-to-3' and the 3'-to-5'pathway (Figure 2). Both require the shortening of the poly(A) tail, a process known as deadenylation. A biphasic model of deadenylation has been proposed, where poly(A) shortening is initiated by the PAN2/PAN3 (poly(A)-nuclease 2 / poly(A)-nuclease 3) complex, where PAN2 is the catalytic subunit. PAN2/PAN3 most likely acts globally on all transcripts, and it has been shown to trim long poly(A) tails (>∼150 nucleotides [nt] in HeLa cells). However, depletion of PAN2/PAN3 from HeLa cells had little effect on steady-state mRNA levels or on the half-lives of individual tested transcripts. Therefore, the PAN2/PAN3 complex seems to have minimal impact on mRNA stability in human cells (Yi et al. 2018). After the initial shortening of the poly(A) tail, deadenylation is carried out by the CCR4-NOT (carbon catabolite repressor 4 - negative on TATA) complex, which is responsible for the shortening of mRNAs with ∼150 nt long poly(A) tails (Yi et al. 2018). Depletion of the catalytic subunits of the CCR4-NOT complex increased global mRNA level and stabilized all tested transcripts (Yi et al. 2018).

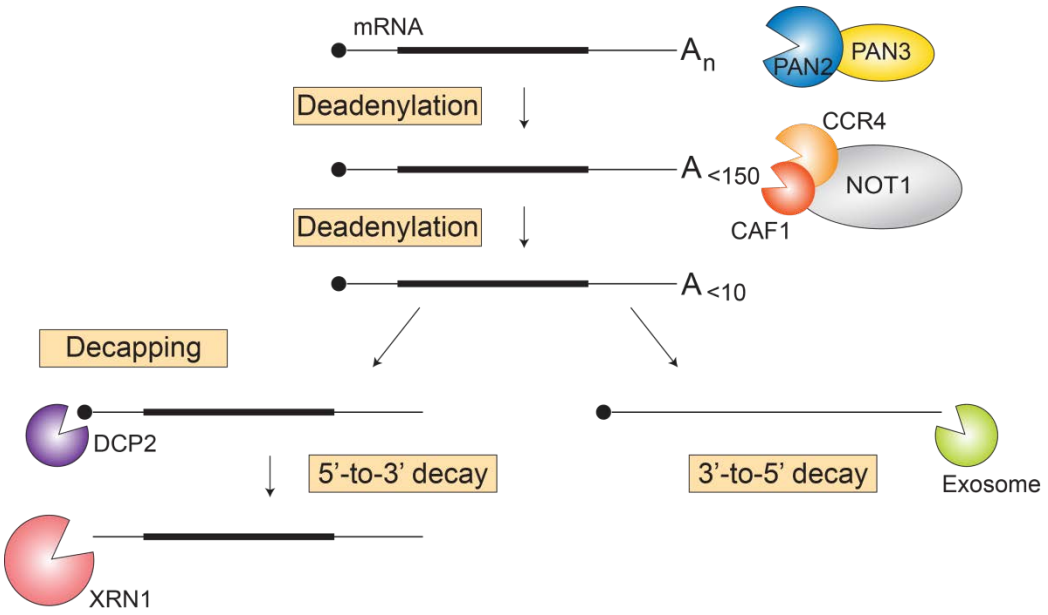


Figure 2. mRNA decay pathways. mRNA degradation starts with the shortening of the 3' poly(A) tail by cytoplasmic deadenylases: the PAN2-PAN3 and the CCR4-NOT complex. Next, the deadenylated mRNA undergoes either 5'-to-3' or 3'-to-5' decay. In 5'-to-3' decay, the cap structure is removed by the decapping complex, and the mRNA is degraded by XRN1. In 3'-to-5' decay, the mRNA is degraded by the exosome complex.

Additionally, the 3'-to-5' exoribonuclease PARN (poly(A)-specific ribonuclease) is responsible for the deadenylation of a small subset of mRNAs (Yamashita et al. 2005).

After deadenylation, the two pathways of mRNA decay diverge: the mRNA either undergoes decapping and subsequent degradation from the 5' end by XRN1 (5'-to-3' exoribonuclease 1), or it is degraded from the 3'end by the exosome complex (3'-to-5' decay). In 5'-to-3' decay, the DCP2 decapping enzyme in complex with activating factors irreversibly removes the 5' cap, exposing the 5' end to degradation by XRN1 (Fig. 2). 3'-to-5' decay is carried out by the exosome, a conserved multi-subunit protein complex, which plays a role in the processing and degradation of almost all RNA classes. The core exosome complex is regulated by the SKI complex in the cytoplasm or the TRAMP (Trf4/Air2/Mtr4p polyadenylation) complex in the nucleus. The exosome is able to target specific transcripts for degradation, based on RNA sequence elements and through cooperation with its cofactors (Zinder and Lima 2017).

Although the scaling from mRNA to protein levels is not linear, there is a positive correlation between them. In *Saccharomyces cerevisiae* (*S. cerevisiae*), close to 80% of the contribution to the control of protein expression is estimated to come from mRNA abundance (Li et al. 2017). Since protein synthesis is a highly energy-consuming process, it is beneficial for

cells to regulate gene expression prior to this step. However, mRNA levels alone are not sufficient to predict protein levels under all conditions. Protein levels are also influenced by translation rates, protein half-life and protein export (Liu et al. 2016).

4.3 Protein degradation

The control of protein turnover is essential for maintaining cellular homeostasis. The major pathway of regulated protein degradation is through the ubiquitin-proteasome system (Metzger et al. 2014). The 26S proteasome is a multi-catalytic ATP-dependent protease complex, which consists of two sub-complexes: the 20S core particle, containing the catalytic chamber, and the 19S regulatory particle, which (together with other regulators) acts as a gatekeeper for substrates entering the proteolytic chamber (Livneh et al. 2016).

Protein ubiquitination is carried out by several enzymes: an ATP-dependent ubiquitin-activating enzyme (E1), an ubiquitin-conjugating enzyme (E2), and an ubiquitin protein ligase (E3). E3s interact with both E2 enzymes and the substrate, and confer substrate specificity to ubiquitination. The human genome encodes more than 600 different RING-type ubiquitin ligases (Deshaies and Joazeiro 2009).

There are two main types of E3 ubiquitin ligases: HECT (Homologous to E6-AP Carboxy Terminus)-type and RING (Really Interesting New Gene)-type. In contrast to HECT-type ubiquitin ligases, RING-type ligases mediate the transfer of ubiquitin directly from E2 to the substrate, and do not serve as catalytic intermediates. The RING domain usually coordinates two Zn^{2+} ions in a cross-brace manner. A common feature of RING-type E3s is homo- and heterodimerization (Metzger et al. 2014).

While targeting proteins for degradation is the best-characterized function of ubiquitin, this post-translational modification has additional regulatory functions that are independent from the proteasome, such as receptor internalization, vesicle sorting, DNA repair and gene silencing (Johnson 2002; Sun and Chen 2004).

4.4 The CCR4-NOT complex

4.4.1 Functions of CCR4-NOT

The CCR4-NOT complex is a key regulator of eukaryotic gene expression. Its importance has been established in most stages of the mRNA lifecycle, from transcription

initiation, through nuclear export to translation regulation and mRNA decay (Miller and Reese 2012; Collart 2016). It has been proposed that the complex is essential for integration of environmental information to regulate gene expression (Collart et al. 2013).

In the nucleus, CCR4-NOT has been shown to impact RNA polymerase (Pol) I and Pol II transcription both negatively and positively, through mechanisms that are not clear yet (Chen et al. 2018). The association between CCR4-NOT and Pol II is thought to be independent of RNA-binding (Kruk et al. 2011). Interaction with the exosome complex also links CCR4-NOT to nuclear mRNA quality control (Azzouz et al. 2009). In addition, the complex interacts with the mRNA nuclear export machinery (Kerr et al. 2011).

The role of CCR4-NOT as a deadenylase is the most studied cellular function of the complex (Wahle and Winkler 2013). Deadenylation is the initial step in mRNA decay and is therefore essential for post-transcriptional gene expression control. Recruitment of the deadenylase complex may occur through a general mechanism via PABPC1, which interacts with both PAN3 (Uchida et al. 2004) and CCR4-NOT associated BTG/TOB (B-cell translocation gene/transducer of erbB-2) proteins (Funakoshi et al. 2007; Mauxion et al. 2008) in metazoans. The CCR4-NOT complex can also be specifically recruited to mRNAs by RNA-binding proteins (Wahle and Winkler 2013).

It has been demonstrated using reporter mRNAs lacking 3’ poly(A) tails that the CCR4-NOT complex is able to repress translation in the absence of deadenylation (Cooke et al. 2010; Chekulaeva et al. 2011; Bawankar et al. 2013; Zekri et al. 2013). Interaction between NOT1 and the translational repressor and decapping activator DDX6 may play a role in this (Chen et al. 2014b; Mathys et al. 2014), as well as interaction between the 4E-transporter (4ET) and DDX6 (Kamenska et al. 2016). In addition, the CCR4-NOT complex is involved in the miRNA pathway, as it is required for miRNA-mediated mRNA decay (Behm-Ansmant et al. 2006). The complex is recruited to miRNA targets via direct interaction with GW182/TNRC6 (Glycine-tryptophan repeat-containing 182kDa protein/Trinucleotide repeat-containing gene 6) proteins (Braun et al. 2011; Chekulaeva et al. 2011; Fabian et al. 2011).

The CCR4-NOT complex has also been implicated in co-translational quality control through the activity of the NOT4 ubiquitin ligase subunit, which will be further discussed in chapter 4.5.3 (Dimitrova et al. 2009; Panasenko 2014).

4.4.2 Recruitment of CCR4-NOT to mRNA targets

Cellular mRNA is always bound by multiple proteins, forming messenger ribonucleoprotein particles (mRNPs) with dynamically changing protein compositions (Hieronymus and Silver 2004). Many RNA-associated proteins have been shown to interact with the CCR4-NOT complex, inducing target-specific translational repression and deadenylation. These proteins often bind to specific elements in the mRNA 3'-UTRs and control gene expression in a wide variety of biological processes (Wahle and Winkler 2013).

The RNA-binding protein tristetraprolin (TTP) directly interacts with the CCR4-NOT complex to induce degradation of mRNAs containing AU-rich elements (AREs) in their 3'-UTRs (Lykke-Andersen and Wagner 2005; Sandler et al. 2011). Its targets include many cytokine-coding mRNAs, the expression levels of which need to be tightly controlled in cells during immune response and therefore have a fast turnover rate (Barreau et al. 2005; Fu et al. 2016).

Nanos protein, essential in early embryonic development in *D. melanogaster*, also directly associates with CCR4-NOT to regulate mRNA stability. Nanos represses the translation of *hunchback* mRNA specifically at the posterior pole of the embryo to establish an anterior-posterior gradient of the Hunchback protein, which is essential for head and thorax development (Tsuda et al. 2003). In many metazoans, including vertebrates, Nanos proteins are essential for germ cell development (Suzuki et al. 2012). The recruitment of the CCR4-NOT complex is essential for vertebrate as well as fly Nanos proteins to induce mRNA decay (Bhandari et al. 2014; Raisch et al. 2016).

Roquin, an ubiquitin ligase with RNA-binding domains, recognizes stem-loop structures in target mRNAs, as shown for the mRNA coding for tumor necrosis factor α (TNF-α). The unstructured C-terminal part of the protein can recruit the CCR4-NOT complex to degrade its target transcripts (Leppek et al. 2013).

GW182/TNRC6 proteins are key factors in microRNA (miRNA) mediated gene silencing. These proteins interact with the miRNA-induced silencing complex (miRISC) and recruit effector complexes, including CCR4-NOT and PAN2-PAN3, to induce translational repression and degradation of target mRNAs (Braun et al. 2011; Chekulaeva et al. 2011; Fabian et al. 2011).

4.4.3 Architecture of the CCR4-NOT complex

The CCR4-NOT complex is assembled around the NOT1 scaffolding protein, and consists of several structurally and functionally distinct modules (Figure 3). The N-terminal part of NOT1 binds NOT10 and NOT11 in metazoans and CAF130 (130 kDa CCR4-associated factor) in yeast (Bawankar et al. 2013). The precise molecular function of this module is not known.

The MIF4G domain of NOT1 binds the deadenylase module, comprising CAF1 (CCR4-associated factor 1) and CCR4 (carbon catabolite repressor protein 4), which are both polyadenosine-specific exonucleases (Petit et al. 2012). CCR4 belongs to the exonuclease-endonuclease-phosphatase (EEP) family of enzymes and binds to CAF1 via a leucin-rich repeat (LRR) in its N-terminal part. CAF1 is a member of the DEDD nuclease family and binds to the MIF4G domain of NOT1 (Petit et al. 2012). In *Schizosaccharomyces pombe (S. pombe)*, CCR4 acts as the major deadenylase, while CAF1 deadenylates only mRNAs not bound by the poly(A)-binding protein Pab1 (Webster et al. 2018). Human CAF1 and CCR4 both catalyze deadenylation, compensating for each other, but, similar to yeast CAF1, human CAF1 cannot deadenylate PABPC1-protected poly(A) tails effectively (Yi et al. 2018).

The NOT1 MIF4G domain is followed by a helical bundle domain that associates with CAF40 (Bawankar et al. 2013; Chen et al. 2014b). This domain of NOT1 is named the CN9BD (CAF40/CNOT9 binding domain), and it folds into a defined domain arranged as a three-helix bundle that is stabilized by CAF40 binding (Chen et al. 2014b; Mathys et al. 2014). CAF40 is a constitutive component of the CCR4-NOT complex (Chen et al. 2001). *Hs* CAF40 has been shown to homodimerize in isolation, however, this is mutually exclusive with binding to NOT1 (Chen et al. 2014b). CAF40 is a crescent-shaped protein that consists of tandem Armadillo repeats formed by alpha helices. The concave, positively charged surface is suggested to be important for nucleic acid binding (Garces et al. 2007). The convex surface of the Armadillo repeats contains conserved tandem tryptophan (Trp) binding pockets, which serve as a protein binding platform, and have been shown to mediate interaction with GW182 and TTP (Chen et al. 2014b; Mathys et al. 2014).

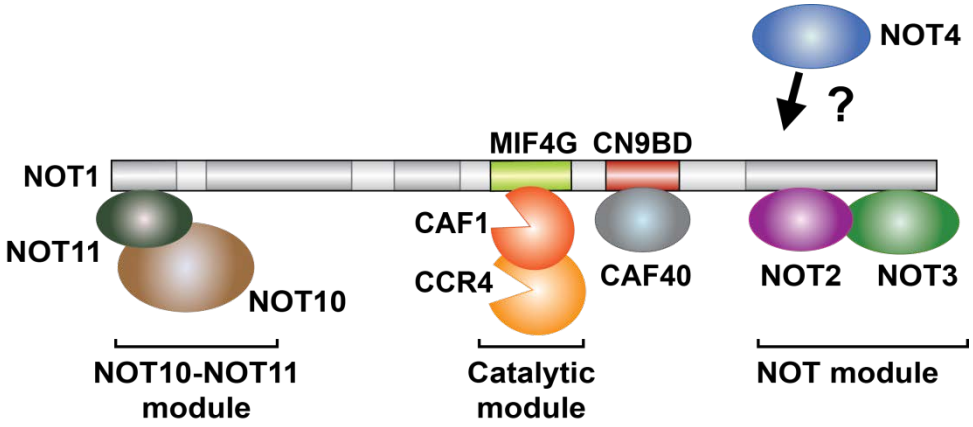


Figure 3. The metazoan CCR4-NOT complex. The NOT1 protein serves as a binding scaffold for the other subunits of the complex. The N-terminal part of NOT1 interacts with NOT10 and NOT11. The central MIF4G domain of NOT1 binds CAF1 and CCR4 to form the catalytic module of the complex. The MIF4G domain is followed by an α-helical bundle, CN9BD, that interacts with CAF40, and a NOT1 superfamily homology domain (SHD) that forms the NOT module together with NOT2 and NOT3. NOT4 is not a stable component of metazoan CCR4-NOT complex.

The NOT module consists of the NOT1 C-terminal domain, NOT2 and NOT3 proteins (Boland et al. 2013). In yeast, the NOT module includes NOT5 (a NOT3 ortholog), and NOT4 as core components (Collart 2016). The NOT module is essential for the activity and stability of the CCR4-NOT complex (Boland et al. 2013). In addition, this module provides binding sites for many RNA-binding proteins (e.g. Nanos, Roquin and Bicaudal-C) that recruit the CCR4-NOT complex to target mRNAs (Chicoine et al. 2007; Leppek et al. 2013; Bhandari et al. 2014; Raisch et al. 2016; Sgromo et al. 2017).

To date, two cryoelectron microscopy structures of the full yeast CCR4-NOT have been published, and both show an overall L-shaped complex (Nasertorabi et al. 2011; Ukleja et al. 2016). However, due to the low resolution (max. 20 Å) of these structures, further studies are needed to create a complete structural model of the CCR4-NOT complex.

4.4.3.1 CAF40

Caf40 was first described in *S. pombe* as a key factor required for nitrogen starvation-induced sexual development (Okazaki et al. 1998). In yeast, deletion of *caf40* elicited mutant phenotypes (such as temperature sensitivity, caffeine sensitivity and sporulation defects) shared by defects in other CCR4-NOT genes (Chen et al. 2001).

The physiological importance of CAF40 in higher eukaryotes is not well understood yet, but several studies have linked CAF40 to development and oncogenesis. CAF40 is shown to be a transcriptional cofactor involved in retinoic-acid induced differentiation of mouse teratocarcinoma cells as well as retinoic acid-controlled lung development (Hiroi et al. 2002). In addition, CAF40 is frequently upregulated in breast cancer, and it has also been linked to the Akt signaling pathway. Phosphorylation of the threonine/serine protein kinase Akt at Ser 473 fully activates the Akt protein, promoting cell proliferation and survival (Nicholson and Anderson 2002). Knockdown of CAF40 reduces Ser 473 phosphorylation of Akt in breast cancer cells. This indicates involvement of CAF40 in activating the Akt signaling pathway and suggests that CAF40 has oncogenic activity (Ajiro et al. 2009). Whole exome sequencing of 20 metastatic melanomas has also revealed a recurrent somatic point mutation at position P131L in the RQCD1 gene coding for CAF40, suggesting that this mutation is selected for during tumor development (Wong et al. 2014). This residue is located on the concave, predominantly hydrophobic surface of the armadillo repeat domain of CAF40, which interacts with the Roquin protein (Sgromo et al. 2017). So far, it has not been investigated whether the P131L mutation affects the CAF40-Roquin interaction. The importance of this mutation in the development and progression of melanoma also requires further studies (Wong et al. 2014).

GW182 proteins, which interact with AGO (Argonaute) protein and are required for miRNA-mediated gene silencing (Fabian and Sonenberg 2012; Braun et al. 2013), bind to CAF40 and NOT1 through tryptophan-containing motifs (Braun et al. 2011; Chekulaeva et al. 2011; Fabian et al. 2011). CAF40 also interacts with the zinc-finger protein tristetraprolin (TTP), which recruits the CCR4-NOT complex to degrade inflammatory mRNAs (Bulbrook et al. 2018).

4.4.3.2 NOT4 as a constitutive or facultative component of CCR4-NOT

Not4 is a stable component of the CCR4-NOT complex in *S. cerevisiae* (Chen et al. 2001). However, it is not an integral subunit of the human and *D. melanogaster* CCR4-NOT complex, as demonstrated by gel filtration and mass spectrometry experiments (Lau et al. 2009; Temme et al. 2010). Thus, NOT4 has been proposed to be a non-constitutive subunit of the CCR4-NOT complex in higher eukaryotes.

It was demonstrated in yeast two hybrid assays that yeast, as well as human NOT1 can interact with human NOT4 (Albert et al. 2000). In yeast, the interaction between the C-terminal part of yeast Not4 and a C-terminal fragment of Not1 has been characterized by X-ray crystallography (Bhaskar et al. 2015). The crystal structure revealed that a low complexity region of Not4 is responsible for the binding (labelled as NOT1 binding site in *Sc* Not4, Figure 4). Interestingly, this low complexity region of yeast Not4 can be only partially aligned with metazoan NOT4, which could explain the weaker association of NOT4 with the CCR4-NOT complex in human and *D. melanogaster* (Bhaskar et al. 2015).

4.5 NOT4

Similar to most components of the CCR4-NOT complex, NOT4 is a conserved protein among eukaryotes (Fig. 4). This is supported by the finding that human NOT4 can complement the phenotype of a yeast Not4 knock-out (Albert et al. 2000). The N-terminal part of NOT4 is highly conserved from yeast to humans, as opposed to the C-terminal region.

The N-terminal part of NOT4 contains a RING domain, which confers ubiquitin ligase activity to the protein (Albert et al. 2002; Mulder et al. 2007). NOT4 harbors two putative RNA-binding domains, a C3H1-type Zn-finger, which is generally implicated in binding AU-rich elements (AREs) in the mRNA 3'-UTR (Lai et al. 2000; Blackshear 2002; Hudson et al. 2004), and an RNA recognition motif (RRM). RRMs are found in over 50% of RNA-binding proteins (Lunde et al. 2007). In yeast, it has recently been shown that the RRM and the C3H1-type Zn finger domain collaborate with the RING domain to control proteasome activity and global proteostasis (Chen et al. 2018). The N-terminal region of NOT4 also contains a positively charged region with coiled coil propensity. In metazoan species, however, these domains remain rather poorly characterized, and their function, their putative RNA-binding properties or RNA targets are currently unknown.

The C-terminal part of NOT4 is predicted to be mostly unstructured. In *S. cerevisiae*, this region was shown to contain an interaction site with Not1 (Bhaskar et al. 2015). In contrast, it has been proposed that NOT4 is a regulatory subunit in metazoans, which is only able to associate with the CCR4-NOT complex in specific cells or under specific conditions (Temme et al. 2010).

In yeast, the Not4 deletion strain (*not4Δ*) displays a slow growth phenotype, as well as sensitivity to many environmental stressors, including high temperature (37 °C), hygromycin B (which leads to errors during protein synthesis) (Mulder et al. 2007; Halter et al. 2014), and hydroxyurea (Lau et al. 2010). The formation of the proteasome is also impaired in *not4Δ* cells: the 19S regulatory particle is defective and unable to deubiquitinate substrates, while the 20S core particles show increased catalytic activity (Panasenko and Collart 2011; Halter et al. 2014).

Heterozygous $NOT4^{+/-}$ mice have defects in vertebrae formation and immunological impairments. NOT4 KO mice die after birth during the pre-weaning stage (before approximately 3-4 weeks of age) (International Mouse Phenotyping Consortium).

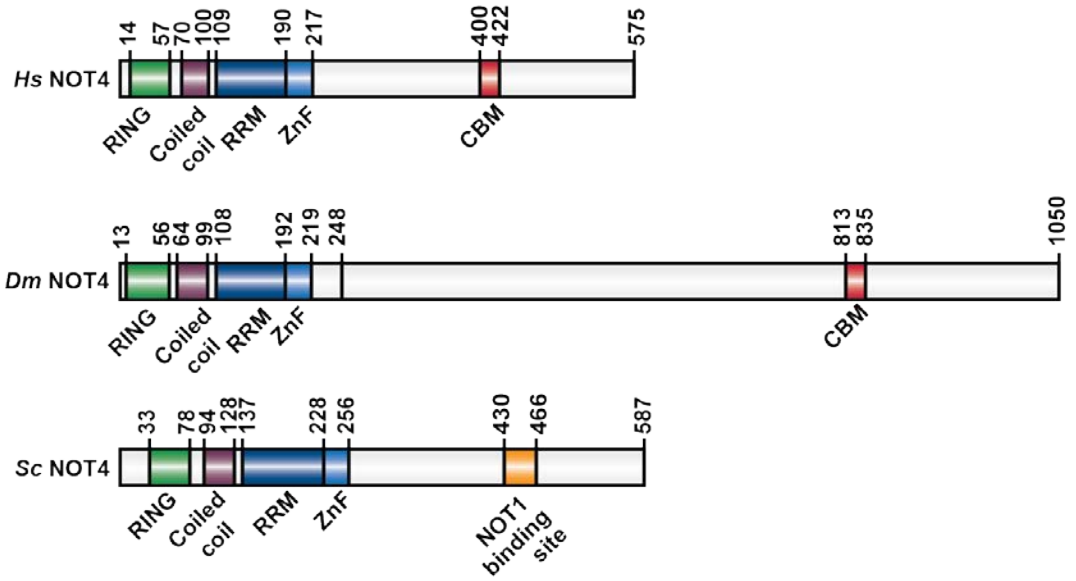


Figure 4. Domain composition of NOT4, adapted from Keskeny et al., 2019. The conserved N-terminus of NOT4 contains a RING-type ubiquitin ligase domain (RING), a positively charged linker with coiled coil propensity (Coiled coil), an RNA-recognition motif (RRM) and a C3H1-type zinc finger domain (ZnF). The CAF40-binding motif (CBM) is shown in red. The CBM is conserved in metazoan NOT4, but not in yeast. *S. cerevisiae* Not4 contains a previously characterized binding site for *S. cerevisiae* Not1, shown in yellow (Bhaskar et al. 2015).

4.5.1 The role of NOT4 in mRNA deadenylation

Despite its physical association with the CCR4-NOT complex in yeast and the potential to bind RNA via the RRM and C3H1 domains, the role of NOT4 in mRNA deadenylation remains unknown. In *D. melanogaster*, NOT4 is dispensable for general mRNA deadenylation (Temme et al. 2010), but the protein has been suggested to play a role in *cyclin B* mRNA decay (Kadyrova et al. 2007). The RNA-associated proteins Nanos and Pumilio are required for repression of the maternal *cyclin B* mRNA in the germline of the fly embryo (Kadyrova et al. 2007). Interaction screens showed that NOT4 interacts with Nanos, and this interaction was proposed to mediate the degradation of *cyclin B* mRNA via recruitment of the CCR4-NOT complex (Kadyrova et al. 2007). However, direct interactions between Nanos and the NOT module of the CCR4-NOT complex have since been demonstrated in human and *D. melanogaster* cells (Bhandari et al. 2014; Raisch et al. 2016), therefore it is possible that NOT4 is dispensable for *cyclin B* mRNA degradation.

4.5.2 NOT4 as an ubiquitin ligase

4.5.2.1 The RING domain of NOT4

Even though several studies have addressed the function of the NOT4 RING domain and ubiquitin ligase activity, the significance of such an activity is not quite clear. The NOT4 RING domain has been shown in hetero-species complementation experiments performed in yeast to be essential for the survival of *not4Δ not5Δ* on 5-fluoro-orotic acid (FOA) media. In these experiments, expression of human NOT4 lacking the conserved N-terminus could not suppress synthetic lethality, while expressing full length NOT4 allowed cell survival (Albert et al. 2002). On the other hand, the complementation of *not4*Δ strains with RING-finger mutants in *S. cerevisiae* showed that not all processes involving yeast Not4 require its E3 ligase activity (Mulder et al. 2007).

The highly conserved RING domain of NOT4 is a special C4C4-type (binding two zinc atoms using its cysteine residues), as opposed to the more common C3HC4-type (containing zinc-coordinating cysteine and histidine residues). Use of nuclear magnetic resonance (NMR) spectroscopy to determine the RING domain protein structure of human NOT4 revealed it consists of an α-helix and three long loops that are stabilized by the coordination of zinc ions. The two zinc ions are bound through cysteine residues in a cross-brace manner, similarly to C3HC4 RING domains, and the overall folding also resembles that of the C3HC4 RING fingers, suggesting similar function (Hanzawa et al. 2001).

Yeast two hybrid experiments showed both human and yeast Not4 are able to interact through their conserved RING domain with the Ubc ubiquitin conjugating enzymes. In addition, direct interaction between human NOT4 and the human E2 enzyme UBCH5B was confirmed by GST pull-down assays (Albert et al. 2002). Another study confirmed that the RING domain of human NOT4 interacts with UBCH5B, but not with several other ubiquitin conjugases. Interestingly, the ubiquitin ligase activity of human NOT4 proved to be dependent on the specific and selective interaction with the UBCH5B ubiquitin conjugating enzyme (Winkler et al. 2004). UBCH5B has two orthologs in yeast: Ubc4 and Ubc5, which exhibit partially distinct functions *in vivo* (Seufert and Jentsch 1990; Chuang and Madura 2005).

An NMR structure of the human UBCH5B/NOT4 complex (Dominguez et al. 2004) identified amino acids in UBCH5B that are involved in binding to the RING domain of NOT4. It was also investigated whether the RRM domain of NOT4 contributes to this interaction, but this

was not the case (Dominguez et al. 2004). The structure of the yeast NOT4 RING domain with Ubc4 has also been determined, in this case by X-ray crystallography (Bhaskar et al. 2015). This structure is very similar to the earlier NMR structures of the human NOT4 ortholog (Hanzawa et al. 2001) and of the yeast Ubc4 in isolation (Cook et al. 1993). The yeast Not4 RING contains two α-helices and three loop regions, as well as the two zinc ions, which are coordinated in a cross-brace fashion by cysteine residues. This structure shows localized differences compared to the previously proposed model of human NOT4-UBCH5B (Dominguez et al. 2004; Bhaskar et al. 2015). Structural alignment of known yeast E2 proteins with Ubc4 revealed subtle differences that explain how the specificity of Not4 is driven toward Ubc4/5 and not toward other Ubc E2 proteins (Bhaskar et al. 2015).

4.5.2.2 Substrates of the NOT4 ubiquitin ligase

In yeast and in human cells, NOT4 has been shown to ubiquitinate a wide array of substrates that cannot be tied to a single cellular process or pathway. Some substrates are polyubiquitinated and degraded by the proteasome, while for others, ubiquitination by NOT4 has regulatory consequences. The following examples serve to illustrate the complexity of the possible NOT4 functions.

There are studies indicating that Not4 contributes to transcriptional control. For example, yeast Not4 regulates *in vivo* H3K4 (histone H3 lysine 4) trimethylation levels and, consequently, gene expression levels through the polyubiquitination and proteasomal degradation of the Jhd2 histone demethylase (Laribee et al. 2007; Mersman et al. 2009). Furthermore, human NOT4 promotes the proteasomal degradation of RNA polymerase II-association factor 1 (PAF1), a subunit of the PAF complex that participates in various steps of the transcriptional process (Sun et al. 2015).

There are the several known substrates of Not4 important for oxidative stress response, suggesting a role for Not4 in the regulation or coordination of these processes. One such example is the Not4-dependent polyubiquitination and degradation of cyclin-C upon H_2O_2 exposure, thereby promoting the expression of genes important for stress response (Cooper et al. 2012). Not4 also regulates the levels of the Yap1, a transcription factor that is crucial for oxidative stress tolerance. Following its recruitment to target genes, Yap1 is degraded in a Not4-dependent

manner. This is proposed to be a mechanism by which some transcription factors are inactivated rapidly when their function is no longer necessary (Gulshan et al. 2012).

It has been recently shown that yeast Not4 plays a role in removing the transcriptionally arrested RNA polymerase II (RNAPII) elongation complex and maintaining genomic integrity under genotoxic stress conditions. This is achieved by indirectly promoting the ubiquitination and proteasome-mediated degradation of the Rbp1 subunit of RNAPII (Jiang et al. 2019).

Yeast Not4 also plays a role in the functional assembly of the proteasome, specifically in the assembly of the 19S regulatory particle (Panasenko and Collart 2011). In addition, Not4 is required for the appropriate interaction between the proteasome and Ecm29, a protein that is important for proteasome stabilization (Panasenko and Collart 2011; De La Mota-Peynado et al. 2013).

4.5.3 The role of NOT4 in co-translational quality control

Several substrates of the NOT4 ubiquitin ligase indicate a role of NOT4 in co-translational quality control, although the connection is not well understood. For example, *S. cerevisiae* Not4 is involved in the ubiquitination of ribosome-associated substrates (Panasenko et al. 2009; Panasenko and Collart 2012). Yeast Not4 has also been suggested to be involved in the ubiquitination and subsequent degradation of polybasic-containing proteins derived from non-stop mRNAs, since protein levels of a non-stop reporter were increased in the absence of Not4 (Dimitrova et al. 2009). On the other hand, Not4 was found to target polypeptides derived from no-go mRNAs containing stalling sequences distant from the 3'-end rather than those derived from non-stop mRNAs (Matsuda et al. 2014). Recent studies have however uncovered Ltn1 (E3 ubiquitin-protein ligase listerin 1) as the main ubiquitin ligase that co-translationally ubiquitinates proteins derived from non-stop and no-go mRNAs (Wilson et al. 2007; Bengtson and Joazeiro 2010; Shao et al. 2013; Matsuda et al. 2014). Ltn1 forms a complex with the RQC factors Tae2, Rqc1 and Cdc48 (Brandman et al. 2012; Defenouillère et al. 2013) and physically associates with the targeted polypeptide and with 80S ribosomes or the 60S ribosomal subunit (Bengtson and Joazeiro 2010; Brandman et al. 2012). Thus, the function of NOT4 in RQC is likely indirect and more complex than ubiquitination of the nascent polypeptide.

The yeast EGD (enhancer of Gal4p DNA binding) complex (also called nascent polypeptide-associated complex, NAC in mammals), which is a ribosome-associated chaperone,

is among the substrates of Not4. Ubiquitination of EGD by Not4 supposedly regulates the association of EGD with the ribosome and with the proteasome (Panasenko et al. 2009).

Yeast Not4 may play a role in RQC through the monoubiquitination of small ribosomal protein eS7A (also known as Rps7A) (Panasenko and Collart 2012; Ikeuchi et al. 2019). This regulatory modification is important for cell viability. Ubiquitinated eS7A associates with actively translating ribosomes (Panasenko and Collart 2012). Importantly, the monoubiquitinated eS7A is further ubiquitinated by Hel2 (Ikeuchi et al. 2019). Ubiquitination of ribosomal proteins by Hel2 (or ZNF598 in humans) is essential for RQC (Juszkiewicz et al. 2018). If Not4 is absent and Hel2 is not able to ubiquitinate its other ribosomal substrate, the uS10 ribosomal protein, no-go decay and RQC are completely inhibited. The molecular details of this process, especially regarding the recruitment of different factors, require further studies (Ikeuchi et al. 2019).

Not4 has been associated with the translational repression of mRNAs undergoing transient ribosome stalling under nutrient withdrawal in yeast (Preissler et al. 2015). Stalled ribosomes lead to the recruitment of Not4 to polysomes (Halter et al. 2014) but the precise mechanism of translational repression remains unclear. Some studies suggest that yeast Not4 can regulate the stability and translation of no-go mRNA reporters (Halter et al. 2014; Preissler et al. 2015; Collart 2016). Whether the destabilization and translational repression of these reporters involves the CCR4-NOT-mediated decay or deadenylation-independent translational repression is currently unclear.

Additional roles in co-translational quality control have been proposed for NOT4 in higher eukaryotes. In response to mitochondrial damage, NOT4 ubiquitinates the ABCE1 ribosome-splitting factor (Wu et al. 2018). Polyubiquitinated ABCE1 is recognized by ubiquitin-binding autophagy receptors and recruited to the mitochondrial outer membrane in order to induce mitophagy (Nguyen et al. 2016). As such, NOT4 serves as a molecular signal for mitophagy initiation. RQC defects and mitochondrial dysfunction have both been linked to neurodegeneration, thus NOT4 may play a role in the pathogenesis of these diseases (Wu et al. 2018).

4.6 Bam (Bag-of-marbles)

Bam is a fly-specific protein, conserved in *D. melanogaster* and other dipteran species. It contains several predicted alpha-helices (Fig. 5). Bam plays an important role in balancing between germline stem cell (GSC) renewal and differentiation in *D. melanogaster* ovaries (Li et al. 2009).

In females, Bam plays a key role in inducing GCSs differentiation (Chen and McKearin 2003; Song et al. 2004). In accordance with this, it is exclusively expressed in late cystoblasts and 2-, 4-, and 8-cell cysts (McKearin and Ohlstein 1995). Loss of Bam leads to uncontrolled stem cell proliferation and germ cell tumors (McKearin and Ohlstein 1995), while ectopic expression of Bam results in differentiation and stem cell loss (Ohlstein and McKearin 1997). In males, Bam is needed to initiate spermatocyte differentiation and is present only in 4-, 8-, and 16-cell cysts (Insco et al. 2009).

Bam physically interacts with and requires the function of Bgcn (benign gonial cell neoplasm) and Sxl (Sex-lethal) in GSC differentiation in females (Flores et al. 2015). Bam controls GSC differentiation at the post-transcriptional level by repressing *nanos* and *E-cadherin* mRNAs. This requires the assembly of a complex containing the Bgcn, Tut (tumorous testis), Sxl and Mei-P26 (meiotic-P26) proteins (Li et al. 2013; Chen et al. 2014a). The Bam-Bgcn complex has been shown to bind to the *nanos* 3'-UTR, repress its translation and promote cystoblast formation (Li et al. 2009).

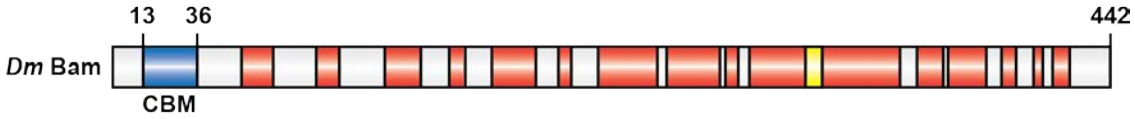


Figure 5. Schematic representation of Bam, adapted from Sgromo et al., 2018. Bam contains several predicted alpha-helices (red) and a beta-strand (yellow). The CAF40-binding motif is shown in blue.

In addition, Bam acts as a positive regulator of hematopoietic progenitor cell maintenance in *D. melanogaster* (Tokusumi et al. 2011). Bam has also been shown to genetically interact with members of the Insulin-like receptor pathway and to positively regulate the expression of the Retinoblastoma-family protein Rbf, which is a crucial regulator of cell proliferation in the posterior signaling center (a hematopoietic cell population in the larval lymph gland) (Tokusumi et al. 2015). In addition, Bam, together with mir-7, genetically interact with

components of the Hedgehog pathway in the maintenance of hematopoietic stem-like progenitors (Tokusumi et al. 2018).

Together with the *Dm* CCR4 homolog Twin (independently of other CCR4-NOT components) and Bgcn, Bam has been shown to promote GSC progeny differentiation (Fu et al. 2015). The Bam-Bgcn complex also binds to the *mei-P26* 3'-UTR to repress translation of *mei-P26* mRNA, which leads to terminal differentiation of progenitor cells in males (Insco et al. 2012). Furthermore, Bam and Bgcn bind the 3'-UTR of *E-cadherin* in S2 cells to repress its translation. Physical interaction between Bam and the translation initiation factor eIF4A inactivates Bam function and promotes the self-renewal of GSCs (Shen et al. 2009).

The ability of Bam to repress its mRNA targets has been linked to its interaction with the CCR4 deadenylase (Fu et al. 2015). This interaction was suggested to be independent of the CCR4-NOT complex, with CCR4 acting as an isolated deadenylase, since binding to both Bam and the CAF1 deadenylase was disrupted by the same mutations in CCR4 (Fu et al. 2015). However, since these mutated residues are located in the hydrophobic core of the leucine-rich repeat (LRR) domain of CCR4 (Basquin et al. 2012), they may have led to the misfolding of this domain and thus might have caused the loss of CCR4-NOT binding non-specifically. Therefore, it is not clear whether Bam requires the assembled CCR4-NOT complex or whether the free CCR4 deadenylase is sufficient to repress the translation of its mRNA targets.

5. Aims and objectives

The CCR4-NOT complex is the major mRNA deadenylase in eukaryotic cells and consequently a central node for mRNA regulation. Many aspects of this regulation are still not completely understood. One way to regulate mRNA decay is the active recruitment of the CCR4-NOT complex to specific mRNA targets (Wahle and Winkler 2013). In this model, RNA-binding proteins, which target specific mRNAs, bind the CCR4-NOT complex, bringing it in close proximity to the mRNA poly(A) tail and thus inducing deadenylation. The main scientific aim and motivation for my doctoral work was to study metazoan-specific aspects in the assembly of the CCR4-NOT complex, and to understand molecular details underlying the recruitment of CCR4-NOT by mRNA-associated proteins. I studied two RNA-associated proteins, NOT4 and Bam, their interactions with the CCR4-NOT complex in higher eukaryotes, and their role in post-transcriptional regulation.

NOT4 is a conserved protein with RNA-binding potential that possesses ubiquitin ligase activity. This combination highlights NOT4 as one of the proteins that could serve as a bridge between the mRNA and protein degradation pathways (Cano et al. 2010). Although the structure and function of the NOT4 protein has been studied extensively in yeast, there have been few studies on this protein in human cells. When I started my doctoral studies, it was still unclear how NOT4 associates with the CCR4-NOT complex in metazoans, but it was clear to be different from the interaction in yeast. It was also unknown whether human NOT4 could elicit mRNA degradation when tethered to a reporter mRNA and whether the ubiquitin ligase activity would play a role downstream of RNA binding. To address these questions, I investigated the interaction between NOT4 and CCR4-NOT in human cells and *in vitro*, using recombinant proteins, and I studied the role of NOT4 post-transcriptional regulation using mRNA reporters.

I was also involved in a collaborative project on the fly-specific Bag-of-marbles (Bam) protein. Bam is a germline stem cell differentiation factor that was proposed to interact with the CCR4 deadenylase, independently of the CCR4-NOT complex (Fu et al. 2015). To test this possibility, I contributed to investigating the interaction of Bam with components of the CCR4-NOT complex in order to better understand how Bam post-transcriptionally regulates mRNA targets.

Unveiling the network of RNA-associated proteins that bind to different modules of the CCR4-NOT complex and deciphering the molecular mechanisms of CCR4-NOT recruitment to

specific mRNA targets is key to understanding details of post-transcriptional gene expression control.

6.1 Synopsis of publication 1

A conserved CAF40-binding motif in metazoan NOT4 mediates association with the CCR4–NOT complex

Csilla Keskeny, Tobias Raisch, Annamaria Sgromo, Catia Igreja, Dipankar Bhandari, Oliver Weichenrieder and Elisa Izaurralde

Genes Dev. 2019 Feb 1;33(3-4):236-252. doi: 10.1101/gad.320952.118.

The CCR4-NOT complex is an important post-transcriptional regulator of gene expression in eukaryotes. The NOT4 E3 ubiquitin ligase is one of the core subunits of this complex in yeast, whereas in human and *D. melanogaster* cells, NOT4 is not stably associated with the complex. In this manuscript, we show that, *in vitro*, human NOT4 nevertheless directly interacts with the NOT1 and CAF40 subunits of the complex in a manner that differs from yeast. In particular, we identified a conserved sequence motif in the C-terminal region of human and *D. melanogaster* NOT4 that interacts with CAF40 (CAF40-binding motif [CBM]). Crystal structures of the CBM–CAF40 complex show that NOT4 targets the same surface of CAF40 as the previously characterized RNA-associated proteins Roquin and Bag-of-marbles, albeit in a different orientation. Consequently, CAF40 emerges as an important mediator for the recruitment of the CCR4-NOT complex. These data lead us to conclude that, in metazoan cells, the interaction between NOT4 and the CCR4-NOT complex is likely a transient and possibly regulated event, and we demonstrated that NOT4 is able to induce decay of bound reporter mRNAs through the 5'-to-3' mRNA decay pathway. Depletion of CAF40 from human cells or structure-guided mutagenesis to disrupt the NOT4–CAF40 interaction impairs NOT4-mediated degradation of reporter mRNAs. In conclusion, our results reveal the molecular details for the association of metazoan NOT4 with the CCR4–NOT complex and highlight the NOT4 ubiquitin ligase as a conserved but non-constitutive cofactor of the CCR4–NOT complex.

My contribution: I performed and analyzed all of the experiments carried out in human cells, and I did the *in vitro* pull-down experiments in Figure 3. Together with my colleagues, I drafted the initial manuscript and took a leading role through all stages of the submission, revision and publication process.

6.2 Synopsis of publication 2

Drosophila Bag-of-marbles directly interacts with the CAF40 subunit of the CCR4–NOT complex to elicit repression of mRNA targets

Annamaria Sgromo, Tobias Raisch, Charlotte Backhaus, **Csilla Keskeny**, Vikram Alva, Oliver Weichenrieder and Elisa Izaurralde

RNA. 2018 Mar;24(3):381-395. doi: 10.1261/rna.064584.117.

Drosophila melanogaster Bag-of-marbles (Bam) is a key factor that determines germline stem cell (GSC) fate. Bam represses the expression of mRNAs encoding stem cell maintenance factors, thereby promoting GSC differentiation. Bam was shown to interact with a putative RNA helicase Bgcn (Benign gonial cell neoplasm) and with the CCR4 deadenylase, a catalytic subunit of the CCR4–NOT complex. Furthermore, Bam was proposed to compete with CAF1 for binding to CCR4, and to displace CCR4 from the CCR4–NOT complex. In this manuscript, we therefore investigated the interaction between Bam and the CCR4–NOT complex by using purified recombinant proteins. We found that Bam does not directly interact with CCR4, but rather binds to the CAF40 subunit of the CCR4-NOT complex. This interaction is mediated by a conserved N-terminal CAF40-binding motif (CBM) in Bam. The crystal structure of the Bam CBM bound to CAF40 shows that the CBM peptide folds into an α-helix when bound to the CAF40 armadillo repeats. Previously, the Roquin protein of *D. melanogaster* had also been shown to contain a CBM and to interact with the same concave surface of CAF40, using a similar binding mode. The two peptides cannot bind to CAF40 simultaneously, and *in vitro* competition assays suggested they may compete for the binding site on CAF40 if simultaneously present in cells. Additionally, we demonstrated that Bam mediates mRNA degradation and translational repression when tethered to reporter mRNAs, and this function was dependent on Bam's interaction with CAF40. In conclusion, Bam recruits the CCR4-NOT complex through direct interaction with CAF40 with the potential to regulate the expression of bound mRNA targets.

My contribution: To study the molecular details of the interaction of Bam with the CCR4-NOT complex, I generated a CAF40-null human cell line using CRISPR-Cas9 genome editing. This cell line was used to demonstrate that Bam is unable to bind to the deadenylase

complex in cells depleted of CAF40. The data characterizing the cell line at the genomic and protein level was produced by me. I was also involved in correcting the manuscript drafted by my colleagues.

7. Discussion

The CCR4-NOT complex is a major regulator of mRNA metabolism in eukaryotes (Collart 2016). It affects many stages of the mRNA lifecycle, including transcription, translation and mRNA decay (Winkler et al. 2006; Funakoshi et al. 2007; Cooke et al. 2010; Kruk et al. 2011). As the principal cytoplasmic deadenylase, the CCR4-NOT complex is essential for bulk mRNA degradation. In addition, RNA-associated proteins are able to recruit CCR4-NOT for the translational repression and degradation of specific subsets of target mRNAs (Wahle and Winkler 2013). In this work, I studied two RNA-associated proteins, NOT4 and Bam, that bind the CAF40 subunit of the CCR4-NOT complex thereby possibly recruiting the complex to target mRNAs. Both factors interact with the same surface as the previously identified CAF40-interaction partner Roquin (Sgromo et al. 2017). Importantly, I showed that in the absence of CAF40 in human cells, the interactions between NOT4 and CCR4-NOT and between Bam and CCR4-NOT are abolished. Studying the function of NOT4 in mRNA regulation, I found that NOT4 promotes the degradation of an mRNA reporter when tethered to its 3'-UTR. Furthermore, I provide evidence that the interaction between human NOT4 and the CCR4-NOT complex is likely a regulated event. The work described in this doctoral thesis, together with other studies, identifies CAF40 as a binding platform for CCR4-NOT recruitment factors. My findings contribute to understanding the molecular and structural details of CCR4-NOT recruitment by RNA-associated proteins and clarify the function of the mammalian NOT4 ubiquitin ligase as a non-constitutive subunit of the CCR4-NOT complex.

7.1 The transient nature of the interaction between NOT4 and the CCR4-NOT complex in human cells

Using recombinant proteins produced in *E. coli*, I showed that full-length human NOT4 as well as its C-terminal fragment directly interact with the CCR4-NOT complex. Co-immunoprecipitation assays show in human and *D. melanogaster* S2 (unpublished data) cells, however, only the C-terminal fragment of NOT4 displayed a stable interaction with the CCR4-NOT complex. Full length *Hs* NOT4 did not efficiently pull down the CCR4-NOT complex in cells. As demonstrated using mass spectrometry and size-exclusion chromatography in human and *D. melanogaster* cells (Lau et al. 2009; Temme et al. 2010), metazoan NOT4 does not constitutively reside in the CCR4-NOT complex. In contrast, *S. cerevisiae* and *S. pombe* NOT4

co-purifies with the CCR4-NOT complex, indicating that NOT4 is an integral component of the complex in yeast (Chen et al. 2001; Stowell et al. 2016). Binding of the C-terminal fragment of NOT4 to the deadenylase complex in not an artefact of unspecific binding due to e.g. aggregation, as this fragment was unable to bind other proteins, such as PABP and EDC4.

These experiments suggest that the interaction between metazoan NOT4 and the CCR4-NOT complex is transient and possibly regulated by the N-terminal region of NOT4. Control of NOT4 association to the deadenylase complex might involve post-translational modifications and/or the binding of regulatory proteins to NOT4, two regulatory mechanisms that are absent in the *in vitro* pull-down assays with recombinant proteins expressed in *E. coli*.

Further experiments aimed at understanding the regulatory role of individual N-terminal domains of *Hs* NOT4 indicated the RRM domain and the region with coiled coil propensity affect the association of *Hs* NOT4 with the CCR4-NOT complex. Deletion of the RRM and the coiled coil domain from the full length protein restored its ability to interact with NOT1 to a level comparable to that of the C-terminal fragment. Furthermore, addition of the RRM domain to the C-terminal fragment inhibited the interaction between NOT4 and NOT1.

It is possible that that C-terminal part of NOT4 is released for interaction with the CCR4-NOT complex as a result of a structural reorganization of the protein. Unless bound to an mRNA, the highly positively charged coiled coil linker and the RRM may sequester the negatively charged CBM. This could also provide a means for the regulation of the NOT4 ubiquitin ligase activity based on the availability of the N-terminal domains.

7.2 The RNA-binding function of NOT4

Tethered NOT4 induces mRNA decay through the 5’-to-3’ decay pathway, suggesting that NOT4 could mediate the recruitment of the CCR4-NOT complex to endogenous mRNAs. NOT4 has been shown to be part of the mRNA interactome, as determined in human HeLa and HEK (human embryonic kidney) cell lines using photoreactive nucleotide-enhanced or conventional UV crosslinking followed by oligo(dT) purification and proteomic analysis (Baltz et al. 2012; Castello et al. 2012). However, to date, it remains to be investigated if NOT4 binds mRNA *in vivo*, and if so, which putative RNA-binding domain is responsible for this function, and also what subset of mRNAs is recognized by NOT4. So far, it is not clear whether the potential mRNA-binding activity of NOT4 has any specificity. In general, the RRM is considered as a non-specific RNA-binding domain (Lunde et al. 2007; Lukong et al. 2008),

while the C3H1-type zinc finger (for example in TTP or Tis11d proteins) has been shown to bind AU-rich elements in the mRNA 3'-UTR (Lai et al. 2000; Blackshear 2002; Hudson et al. 2004). Furthermore, the combination of different RNA-binding domains has been suggested to affect binding specificity and affinity (Lunde et al. 2007). For example, both Dicer and RNase III proteins contain an endonuclease catalytic domain followed by a double-stranded RNA (dsRNA)-binding domain. While both proteins recognize dsRNA, Dicer contains additional domains that allow it to specifically bind dsRNAs of the RNA interference pathway (Macrae et al. 2006; Lunde et al. 2007).

7.3 NOT4 as an RNA-binding ubiquitin ligase

The functional relevance of the NOT4 ubiquitin ligase activity in the context of the CCR4-NOT complex is currently unknown. My results show that the NOT4 RING domain is dispensable for NOT4-mediated decay of a bound reporter mRNA. However, in the past years, there has been an increasing interest in studying ubiquitin ligases which, similar to NOT4, contain predicted RNA-binding domains. These proteins are suggested to play a role in the regulation of mRNA metabolism, even though the importance of the ubiquitin ligase activity in this process is not clear in most cases (Cano et al. 2010). In addition to NOT4, there are at least 14 more ubiquitin ligases with different types of RNA-binding domains in the human proteome, including RRMs, KH (heterogeneous nuclear K-homology) domains, PAZ (Piwi/Argonaute/Zwille) domains or zinc finger domains (Cano et al. 2010). The first link between these features was established when the turnover of cytokine mRNAs containing AU-rich elements was shown to be dependent on ubiquitination (Laroia et al. 2002). Roquin, an E3 ligase that contains a RING domain, a unique ROQ domain and a C3H1-type zinc finger domain, has been shown to regulate mRNA stability and to interact with the CCR4-NOT complex (Vinuesa et al. 2005; Sgromo et al. 2017). It has been suggested that ubiquitin ligases with RNA-binding ability could serve as a bridge between the mRNA and protein degradation pathways (Cano et al. 2010). In the case of Not4, the importance of the ubiquitin ligase activity has been demonstrated in no-go decay in yeast (Ikeuchi et al. 2019), but so far this has not been connected to the protein's RNA-binding potential.

7.4 Recruitment of the CCR4-NOT complex

As shown in this work, recruitment of the CCR4-NOT complex via direct interactions with CAF40 is required for NOT4 and Bam to degrade reporter mRNAs. A growing number of RNA-binding proteins interact with the CCR4-NOT deadenylase complex, which contains conserved binding surfaces to interact with different RNA-associated proteins. Known examples are TTP (Bulbrook et al. 2018), SMG7 (Loh et al. 2013), or BTG2/TOB (Mauxion et al. 2008). These interactions mediate the recruitment of the CCR4-NOT complex to specific mRNAs and have the potential to regulate deadenylation.

Binding of the CCR4-NOT complex is often characterized by functionally relevant peptide motifs embedded in low complexity and unstructured regions, known as short linear motifs (SLiMs) (Van Roey et al. 2014). SLiMs bind other proteins with high specificity and relatively low affinity. Another feature of SLiMs is evolutionary plasticity, meaning that while these motifs may diverge in sequence between orthologs, the interactions they perform are overall maintained (Tompa 2012; Van Roey et al. 2014). Based on these features, the CAF40-binding motifs of NOT4, Bam and Roquin can also be regarded as SLiMs. Despite the structural and functional similarities of CBMs in NOT4, Bam and Roquin, these three proteins do not share sequence identity. Interestingly, structures of the CBMs in complex with CAF40 showed that the orientation of the NOT4 CBM is in an inverted position compared to the Bam and Roquin CBMs (Figure 7). Taken together, these data suggest the CBMs of NOT4, Bam and Roquin evolved independently in order to bind CAF40.

Even though SLiMs are generally thought not to be conserved across species due to evolutionary plasticity of single motif instances (Van Roey et al. 2014), the CBM of NOT4 is conserved among metazoans, suggesting the interaction between NOT4 and CAF40 is also conserved.

In many cases, RNA-binding regions and interaction sites for the CCR4-NOT complex are spatially separated within the proteins. Nanos, Roquin and NOT4 contain highly conserved RNA-binding domains, while the SLiMs responsible for CCR4-NOT binding are located in long unstructured regions of these proteins. The long linker regions may allow the proteins to link the mRNA to the CCR4-NOT complex or any other interactor. In other cases, different proteins are responsible for RNA binding and for CCR4-NOT recruitment. Such is the case of the miRISC, where AGO proteins interact with RNA, while GW182 bridges AGO and the mRNA to the

CCR4-NOT complex. In the case of Bam, it is not clear yet whether the RNA-binding function resides in the protein itself or if Bam associates with additional proteins that mediate mRNA binding.

Importantly, RNA-associated proteins may interact with a single subunit of the CCR4-NOT complex or with multiple ones. Bam is an example of an RNA-associated protein that uses a single subunit, CAF40, to interact with the CCR4-NOT complex. The interaction between Bam and CAF40 proved to be necessary and sufficient to recruit the CCR4-NOT complex to a bound reporter mRNA. This finding is supported by the fact that single point mutations abolishing the interaction between Bam and CAF40, or CAF40 depletion blocked CCR4-NOT mediated mRNA degradation upon tethering of Bam.

In contrast to Bam, and similar to *Dm* Roquin, *Dm* Nanos, TTP and GW182, metazoan NOT4 contains multiple motifs in its mostly unstructured C-terminal region that provide interactions with both NOT1 and CAF40. The loss of CAF40 is sufficient to abolish the interaction between NOT4 and the CCR4-NOT complex in co-immunoprecipitation assays. NOT4 appears to have multiple NOT1-binding sites in its C-terminal region, however, these regions could not be precisely mapped in *in vitro* using pull-down assays. As a result, the individual contribution to the interaction from each subunit could not be clearly dissected. Based on my results, I propose a model where CAF40 is the principal mediator for the interaction of NOT4 with the CCR4-NOT complex. This differs from what is known about *Sc* NOT4, where NOT1 appears to be the protein's sole interactor in the CCR4-NOT complex (see comparison of proposed binding modes in Fig. 6).

Sequence alignments of NOT4 proteins revealed that the CBM is conserved in metazoans but could not be identified in fungi. Consequently, it is possible that in yeast, the only contact between NOT4 and the CCR4-NOT complex is the known interaction (Bhaskar et al. 2015) between NOT4 and NOT1. In human and *D. melanogaster*, binding to CAF40 through the NOT4 CBM is essential for a functional interaction between NOT4 and the complex. Furthermore, human NOT4 can also directly interact with NOT1, as demonstrated *in vitro* using pull-down assays with recombinant proteins. Therefore, the binding mode between metazoan NOT4 and CCR4-NOT differs from yeast, as it involves two components of the deadenylase complex: CAF40 and NOT1. This functional difference in the binding mode of NOT4 to the CCR4-NOT complex in metazoans may also contribute to the fact that NOT4 is not a constitutive

deadenylase complex subunit. As mentioned above, the negatively charged CBM in metazoans may also interact with the positively charged coiled coil linker in the N-terminus of NOT4, which could prevent and thus regulate interaction with the CCR4-NOT complex. The stable interaction and absence of a CBM in yeast could support this theory of structural reorganization as a means of regulating binding of NOT4 to CCR4-NOT in higher eukaryotes.

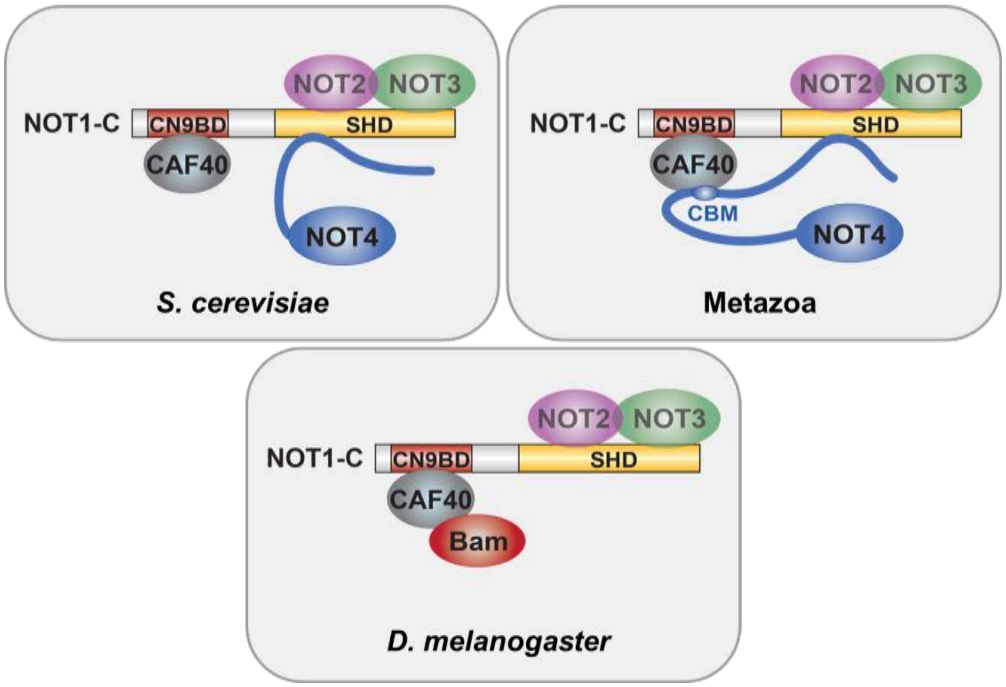


Figure 6. CAF40 mediates association of metazoan NOT4 and Bam with the CCR4-NOT complex. In yeast, Not4 has been shown to bind the C-terminal SHD domain of NOT1. In metazoans, NOT4 associates with the CCR4-NOT complex via binding to both NOT1 and CAF40, a previously unknown interaction partner. NOT4 binds to CAF40 through its CBM in the C-terminal, mainly unstructured part of the protein. *D. melanogaster* Bam also interacts with the CAF40 subunit of the CCR4-NOT complex. NOT4 and Bam both bind to the same surface of CAF40 as Roquin (Sgromo et al. 2017).

7.5 CAF40 is a binding platform for RNA-associated proteins

The results provided in my thesis show that NOT4 and Bam are RNA-associated proteins that recruit the CCR4-NOT complex to mediate degradation and translational repression of reporter mRNAs.

The convex and concave surfaces of CAF40 both provide binding sites for RNA-associated proteins like Roquin, TTP and GW182. Interestingly, NOT4 and Bam both bind to the

same concave surface of the armadillo-repeat domains of CAF40 as Roquin (Fig. 7). This indicates that binding of NOT4, Bam and Roquin to CAF40 is mutually exclusive, hence there might be competition for binding, provided the proteins are all present in the same cells under the same physiological conditions. Protein abundance, localization, additional binding partners and post-translational modifications may also affect competition for CAF40 binding. ITC (isothermal titration calorimetry) experiments to measure binding affinities of the Roquin and Bam CBMs to CAF40 showed that Bam has a much stronger affinity for binding to CAF40 compared to Roquin, which was also supported by *in vitro* competition assays. These results suggest Bam has a competitive advantage over Roquin for binding to CAF40, despite their highly similar binding modes. Bam is a fly-specific protein, and NOT4 and Roquin have *D. melanogaster* orthologs, so competition for binding CAF40 is, in theory, possible between these three proteins in flies. In human cells, NOT4 and Roquin may compete for binding to CAF40, depending on their expression under the same conditions. However, it is not clear whether these proteins are expressed in the same cell types *in vivo*, and whether they compete with each other and with additional, so far unknown interactors of CAF40 under physiological conditions. Therefore, the functional relevance of the shared interaction site on CAF40 requires further studies.

CAF40 and its CBM-binding surface are highly conserved, which suggests there could be more CBM-containing proteins that interact with CAF40. An *in silico* search based on the Bam and *Dm* Roquin CBM profile identified several proteins with potential CBMs, however, testing these proteins for interaction with CAF40 *in vitro* yielded no positive results. This indicates that if there are CBMs in other proteins, they evolved independently from Bam and Roquin, and thus have no sequence conservation.

Apart from Roquin, NOT4 and Bam, CAF40 has been shown to interact with other RNA-associated proteins, such as GW182 and TTP, through surfaces that are distinct from the CBM-binding surface. It is tempting to speculate that GW182/TTP – CAF40 – Bam/NOT4 may form a larger complex together with CCR4-NOT. This could be important for example to regulate degradation of specific mRNA targets. Identifying transcripts regulated by CAF40, and how these targets are affected by the presence or absence of different mRNA-associated proteins that bind to CAF40 would be interesting to explore in the future.

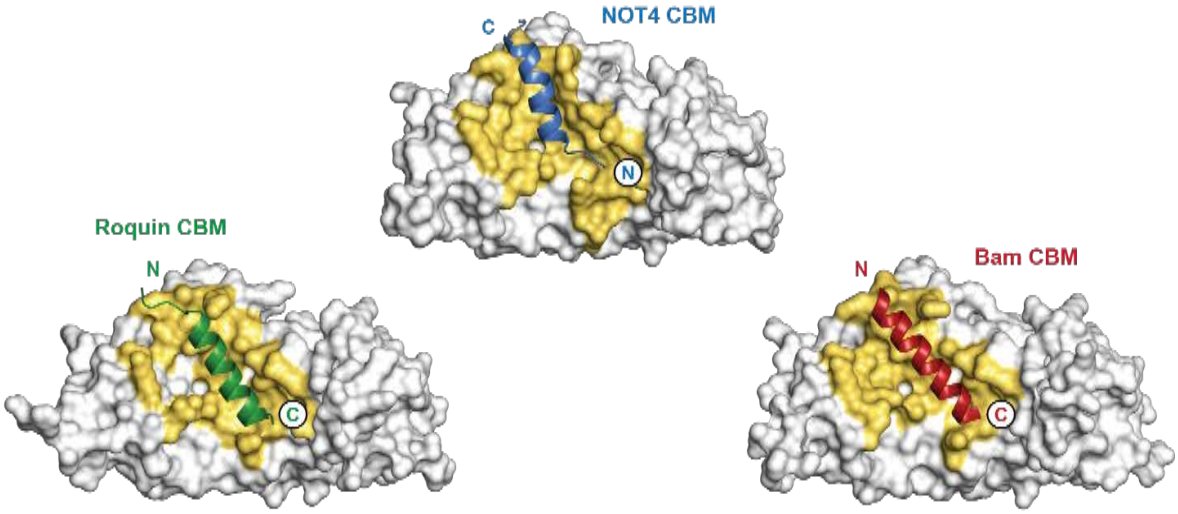


Figure 7. Crystal structure of CBMs bound to CAF40. Surface representation of *Hs* CAF40 in complex with the CBMs of *Dm* NOT4 (Keskeny et al. 2019, Protein Data Bank (PDB) ID: 6hon), *Dm* Roquin (Sgromo et al. 2017, PDB ID: 5lsw) and *Dm* Bam (Sgromo et al. 2018, PDB ID: 5onb).The CBMs are shown in cartoon representation, excluding structurally variable flanks. Interacting residues of CAF40 are colored in yellow. The N- and C-terminal ends of the CBMs are marked. Note the inverted orientation of the NOT4 CBM vs. Roquin and Bam. Adapted from Keskeny et al, 2019.

7.6 Conclusion

My doctoral work contributed to the field of post-transcriptional control, and specifically mRNA decay, by shedding light on mechanisms of recruitment of the CCR4-NOT complex through its CAF40 subunit. The data shown here fit a model of targeted deadenylation where the non-enzymatic subunits of the CCR4-NOT complex serve as binding sites for RNA-associated proteins that recruit the complex to specific mRNA targets, leading to the deadenylation, translational repression and eventual degradation of these mRNAs. These RNA-associated proteins typically contain one or several conserved RNA-binding domains, and spatially separated interaction sites for the CCR4-NOT complex. The binding of NOT4, Bam and Roquin to CAF40 is the first known example of a binding surface on the CCR4-NOT complex which is shared by several proteins with a similar binding mode. It is tempting to speculate that competition for binding CAF40 may be a new regulatory mechanism of CCR4-NOT recruitment, which requires further studies. The interaction network of the CCR4-NOT complex and associated decay factors, as well as interplay between the complex and the downstream mRNA decay machinery are important directions for future research.

8. References

Ajiro M, Katagiri T, Ueda K, Nakagawa H, Fukukawa C, Lin ML, Park JH, Nishidate T, Daigo Y, Nakamura Y. 2009. Involvement of RQCD1 overexpression, a novel cancer-testis antigen, in the Akt pathway in breast cancer cells. *Int J Oncol* **35**: 673-681.

Albert TK, Hanzawa H, Legtenberg YI, de Ruwe MJ, van den Heuvel FA, Collart MA, Boelens R, Timmers HT. 2002. Identification of a ubiquitin-protein ligase subunit within the CCR4-NOT transcription repressor complex. *EMBO J* **21**: 355-364.

Albert TK, Lemaire M, van Berkum NL, Gentz R, Collart MA, Timmers HTM. 2000. Isolation and characterization of human orthologs of yeast CCR4–NOT complex subunits. *Nucleic Acids Research* **28**: 809-817.

Azzouz N, Panasenko OO, Colau G, Collart MA. 2009. The CCR4-NOT complex physically and functionally interacts with TRAMP and the nuclear exosome. *PLoS One* **4**: e6760.

Baltz Alexander G, Munschauer M, Schwanhäusser B, Vasile A, Murakawa Y, Schueler M, Youngs N, Penfold-Brown D, Drew K, Milek M et al. 2012. The mRNA-Bound Proteome and Its Global Occupancy Profile on Protein-Coding Transcripts. *Molecular Cell* **46**: 674-690.

Baralle M, Baralle FE. 2018. The splicing code. *Biosystems* **164**: 39-48.

Barreau C, Paillard L, Osborne HB. 2005. AU-rich elements and associated factors: are there unifying principles? *Nucleic Acids Research* **33**: 7138-7150.

Basquin J, Roudko VV, Rode M, Basquin C, Seraphin B, Conti E. 2012. Architecture of the nuclease module of the yeast Ccr4-not complex: the Not1-Caf1-Ccr4 interaction. *Mol Cell* **48**: 207-218.

Bawankar P, Loh B, Wohlbold L, Schmidt S, Izaurralde E. 2013. NOT10 and C2orf29/NOT11 form a conserved module of the CCR4-NOT complex that docks onto the NOT1 N-terminal domain. *RNA Biol* **10**: 228-244.

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.

Bengtson MH, Joazeiro CA. 2010. Role of a ribosome-associated E3 ubiquitin ligase in protein quality control. *Nature* **467**: 470-473.

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

Bhaskar V, Basquin J, Conti E. 2015. Architecture of the ubiquitylation module of the yeast Ccr4-Not complex. *Structure* **23**: 921-928.

Bhat M, Robichaud N, Hulea L, Sonenberg N, Pelletier J, Topisirovic I. 2015. Targeting the translation machinery in cancer. *Nat Rev Drug Discov* **14**: 261-278.

Blackshear PJ. 2002. Tristetraprolin and other CCCH tandem zinc-finger proteins in the regulation of mRNA turnover. *Biochem Soc Trans* **30**: 945-952.

Boland A, Chen Y, Raisch T, Jonas S, Kuzuoglu-Ozturk D, Wohlbold L, Weichenrieder O, Izaurralde E. 2013. Structure and assembly of the NOT module of the human CCR4-NOT complex. *Nat Struct Mol Biol* **20**: 1289-1297.

Brandman O, Stewart-Ornstein J, Wong D, Larson A, Williams CC, Li GW, Zhou S, King D, Shen PS, Weibezahn J et al. 2012. A ribosome-bound quality control complex triggers degradation of nascent peptides and signals translation stress. *Cell* **151**: 1042-1054.

Braun Joerg E, Huntzinger E, Fauser M, Izaurralde E. 2011. GW182 Proteins Directly Recruit Cytoplasmic Deadenylase Complexes to miRNA Targets. *Molecular Cell* **44**: 120-133.

Braun JE, Huntzinger E, Izaurralde E. 2013. The role of GW182 proteins in miRNA-mediated gene silencing. *Adv Exp Med Biol* **768**: 147-163.

Bulbrook D, Brazier H, Mahajan P, Kliszczak M, Fedorov O, Marchese FP, Aubareda A, Chalk R, Picaud S, Strain-Damerell C et al. 2018. Tryptophan-Mediated Interactions between Tristetraprolin and the CNOT9 Subunit Are Required for CCR4-NOT Deadenylase Complex Recruitment. *J Mol Biol* **430**: 722-736.

Cano F, Miranda-Saavedra D, Lehner PJ. 2010. RNA-binding E3 ubiquitin ligases: novel players in nucleic acid regulation. *Biochem Soc Trans* **38**: 1621-1626.

Carmody SR, Wente SR. 2009. mRNA nuclear export at a glance. *Journal of Cell Science* **122**: 1933-1937.

Castello A, Fischer B, Eichelbaum K, Horos R, Beckmann Benedikt M, Strein C, Davey Norman E, Humphreys David T, Preiss T, Steinmetz Lars M et al. 2012. Insights into RNA Biology from an Atlas of Mammalian mRNA-Binding Proteins. *Cell* **149**: 1393-1406.

Chang H, Lim J, Ha M, Kim VN. 2014. TAIL-seq: genome-wide determination of poly(A) tail length and 3' end modifications. *Mol Cell* **53**: 1044-1052.

Chekulaeva M, Mathys H, Zipprich JT, Attig J, Colic M, Parker R, Filipowicz W. 2011. miRNA repression involves GW182-mediated recruitment of CCR4-NOT through conserved W-containing motifs. *Nat Struct Mol Biol* **18**: 1218-1226.

Chen D, McKearin D. 2003. Dpp signaling silences bam transcription directly to establish asymmetric divisions of germline stem cells. *Curr Biol* **13**: 1786-1791.

Chen D, Wu C, Zhao S, Geng Q, Gao Y, Li X, Zhang Y, Wang Z. 2014a. Three RNA binding proteins form a complex to promote differentiation of germline stem cell lineage in Drosophila. *PLoS Genet* **10**: e1004797.

Chen H, Sirupangi T, Wu ZH, Johnson DL, Laribee RN. 2018. The conserved RNA recognition motif and C3H1 domain of the Not4 ubiquitin ligase regulate in vivo ligase function. *Sci Rep* **8**: 8163.

Chen J, Rappsilber J, Chiang YC, Russell P, Mann M, Denis CL. 2001. Purification and characterization of the 1.0 MDa CCR4-NOT complex identifies two novel components of the complex. *J Mol Biol* **314**: 683-694.

Chen Y, Boland A, Kuzuoglu-Ozturk D, Bawankar P, Loh B, Chang CT, Weichenrieder O, Izaurralde E. 2014b. A DDX6-CNOT1 complex and W-binding pockets in CNOT9 reveal direct links between miRNA target recognition and silencing. *Mol Cell* **54**: 737-750.

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

Chuang SM, Madura K. 2005. Saccharomyces cerevisiae Ub-conjugating enzyme Ubc4 binds the proteasome in the presence of translationally damaged proteins. *Genetics* **171**: 1477-1484.

Collart MA. 2016. The Ccr4-Not complex is a key regulator of eukaryotic gene expression. *Wiley Interdiscip Rev RNA* **7**: 438-454.

Collart MA, Panasenko OO, Nikolaev SI. 2013. The Not3/5 subunit of the Ccr4-Not complex: a central regulator of gene expression that integrates signals between the cytoplasm and the nucleus in eukaryotic cells. *Cell Signal* **25**: 743-751.

Cook WJ, Jeffrey LC, Xu Y, Chau V. 1993. Tertiary structures of class I ubiquitin-conjugating enzymes are highly conserved: crystal structure of yeast Ubc4. *Biochemistry* **32**: 13809-13817.

Cooke A, Prigge A, Wickens M. 2010. Translational repression by deadenylases. *J Biol Chem* **285**: 28506-28513.

Cooper KF, Scarnati MS, Krasley E, Mallory MJ, Jin C, Law MJ, Strich R. 2012. Oxidative-stress-induced nuclear to cytoplasmic relocalization is required for Not4-dependent cyclin C destruction. *J Cell Sci* **125**: 1015-1026.

De La Mota-Peynado A, Lee SY, Pierce BM, Wani P, Singh CR, Roelofs J. 2013. The proteasome-associated protein Ecm29 inhibits proteasomal ATPase activity and in vivo protein degradation by the proteasome. *J Biol Chem* **288**: 29467-29481.

Defenouillère Q, Yao Y, Mouaikel J, Namane A, Galopier A, Decourty L, Doyen A, Malabat C, Saveanu C, Jacquier A et al. 2013. Cdc48-associated complex bound to 60S particles is required for the clearance of aberrant translation products. *Proceedings of the National Academy of Sciences* **110**: 5046-5051.

Deshaies RJ, Joazeiro CA. 2009. RING domain E3 ubiquitin ligases. *Annu Rev Biochem* **78**: 399-434.

Dever TE, Dinman JD, Green R. 2018. Translation Elongation and Recoding in Eukaryotes. *Cold Spring Harb Perspect Biol* **10**.

Dimitrova LN, Kuroha K, Tatematsu T, Inada T. 2009. Nascent peptide-dependent translation arrest leads to Not4p-mediated protein degradation by the proteasome. *J Biol Chem* **284**: 10343-10352.

Dominguez C, Bonvin AM, Winkler GS, van Schaik FM, Timmers HT, Boelens R. 2004. Structural model of the UbcH5B/CNOT4 complex revealed by combining NMR, mutagenesis, and docking approaches. *Structure* **12**: 633-644.

Dominski Z, Marzluff WF. 2007. Formation of the 3' end of histone mRNA: getting closer to the end. *Gene* **396**: 373-390.

Fabian MR, Cieplak MK, Frank F, Morita M, Green J, Srikumar T, Nagar B, Yamamoto T, Raught B, Duchaine TF et al. 2011. miRNA-mediated deadenylation is orchestrated by GW182 through two conserved motifs that interact with CCR4–NOT. *Nature Structural &Amp; Molecular Biology* **18**: 1211.

Fabian MR, Sonenberg N. 2012. The mechanics of miRNA-mediated gene silencing: a look under the hood of miRISC. *Nature Structural &Amp; Molecular Biology* **19**: 586.

Fabian MR, Sonenberg N, Filipowicz W. 2010. Regulation of mRNA translation and stability by microRNAs. *Annu Rev Biochem* **79**: 351-379.

Flores HA, Bubnell JE, Aquadro CF, Barbash DA. 2015. The Drosophila bag of marbles Gene Interacts Genetically with Wolbachia and Shows Female-Specific Effects of Divergence. *PLoS Genet* **11**: e1005453.

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.

Fu Z, Geng C, Wang H, Yang Z, Weng C, Li H, Deng L, Liu L, Liu N, Ni J et al. 2015. Twin Promotes the Maintenance and Differentiation of Germline Stem Cell Lineage through Modulation of Multiple Pathways. *Cell Reports* **13**: 1366-1379.

Funakoshi Y, Doi Y, Hosoda N, Uchida N, Osawa M, Shimada I, Tsujimoto M, Suzuki T, Katada T, Hoshino S. 2007. Mechanism of mRNA deadenylation: evidence for a molecular interplay between translation termination factor eRF3 and mRNA deadenylases. *Genes Dev* **21**: 3135-3148.

Garces RG, Gillon W, Pai EF. 2007. Atomic model of human Rcd-1 reveals an armadillo-like-repeat protein with in vitro nucleic acid binding properties. *Protein Sci* **16**: 176-188.

Goldstrohm AC, Wickens M. 2008. Multifunctional deadenylase complexes diversify mRNA control. *Nat Rev Mol Cell Biol* **9**: 337-344.

Gulshan K, Thommandru B, Moye-Rowley WS. 2012. Proteolytic degradation of the Yap1 transcription factor is regulated by subcellular localization and the E3 ubiquitin ligase Not4. *J Biol Chem* **287**: 26796-26805.

Halter D, Collart MA, Panasenko OO. 2014. The Not4 E3 ligase and CCR4 deadenylase play distinct roles in protein quality control. *PLoS One* **9**: e86218.

Hanzawa H, de Ruwe MJ, Albert TK, van Der Vliet PC, Timmers HT, Boelens R. 2001. The structure of the C4C4 ring finger of human NOT4 reveals features distinct from those of C3HC4 RING fingers. *J Biol Chem* **276**: 10185-10190.

Hieronymus H, Silver PA. 2004. A systems view of mRNP biology. *Genes Dev* **18**: 2845-2860.

Hinnebusch AG. 2017. Structural Insights into the Mechanism of Scanning and Start Codon Recognition in Eukaryotic Translation Initiation. *Trends in Biochemical Sciences* **42**: 589-611.

Hiroi N, Ito T, Yamamoto H, Ochiya T, Jinno S, Okayama H. 2002. Mammalian Rcd1 is a novel transcriptional cofactor that mediates retinoic acid-induced cell differentiation. *The EMBO Journal* **21**: 5235.

Hudson BP, Martinez-Yamout MA, Dyson HJ, Wright PE. 2004. Recognition of the mRNA AU-rich element by the zinc finger domain of TIS11d. *Nat Struct Mol Biol* **11**: 257-264.

Ikeuchi K, Tesina P, Matsuo Y, Sugiyama T, Cheng J, Saeki Y, Tanaka K, Becker T, Beckmann R, Inada T. 2019. Collided ribosomes form a unique structural interface to induce Hel2-driven quality control pathways. *EMBO J* **38**.

Insco Megan L, Bailey Alexis S, Kim J, Olivares Gonzalo H, Wapinski Orly L, Tam Cheuk H, Fuller Margaret T. 2012. A Self-Limiting Switch Based on Translational Control Regulates the Transition from Proliferation to Differentiation in an Adult Stem Cell Lineage. *Cell Stem Cell* **11**: 689-700.

Insco ML, Leon A, Tam CH, McKearin DM, Fuller MT. 2009. Accumulation of a differentiation regulator specifies transit amplifying division number in an adult stem cell lineage. *Proceedings of the National Academy of Sciences* **106**: 22311.

Jackson RJ, Hellen CUT, Pestova TV. 2010. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nature Reviews Molecular Cell Biology* **11**: 113.

Jansen RP. 2001. mRNA localization: message on the move. *Nat Rev Mol Cell Biol* **2**: 247-256.

Jiang H, Wolgast M, Beebe LM, Reese JC. 2019. Ccr4-Not maintains genomic integrity by controlling the ubiquitylation and degradation of arrested RNAPII. *Genes Dev* **33**: 705-717.

Joazeiro CAP. 2017. Ribosomal Stalling During Translation: Providing Substrates for Ribosome-Associated Protein Quality Control. *Annual Review of Cell and Developmental Biology* **33**: 343-368.

Johnson ES. 2002. Ubiquitin branches out. *Nat Cell Biol* **4**: E295-298.

Juszkiewicz S, Chandrasekaran V, Lin Z, Kraatz S, Ramakrishnan V, Hegde RS. 2018. ZNF598 Is a Quality Control Sensor of Collided Ribosomes. *Molecular Cell* **72**: 469-481.e467.

Kadyrova LY, Habara Y, Lee TH, Wharton RP. 2007. Translational control of maternal Cyclin B mRNA by Nanos in the Drosophila germline. *Development* **134**: 1519-1527.

Kamenska A, Simpson C, Vindry C, Broomhead H, Benard M, Ernoult-Lange M, Lee BP, Harries LW, Weil D, Standart N. 2016. The DDX6-4E-T interaction mediates translational repression and P-body assembly. *Nucleic Acids Res* **44**: 6318-6334.

Kerr SC, Azzouz N, Fuchs SM, Collart MA, Strahl BD, Corbett AH, Laribee RN. 2011. The Ccr4-Not complex interacts with the mRNA export machinery. *PLoS One* **6**: e18302.

Kruk JA, Dutta A, Fu J, Gilmour DS, Reese JC. 2011. The multifunctional Ccr4-Not complex directly promotes transcription elongation. *Genes Dev* **25**: 581-593.

Lai WS, Carballo E, Thorn JM, Kennington EA, Blackshear PJ. 2000. Interactions of CCCH zinc finger proteins with mRNA. Binding of tristetraprolin-related zinc finger proteins to Au-rich elements and destabilization of mRNA. *J Biol Chem* **275**: 17827-17837.

Laribee RN, Shibata Y, Mersman DP, Collins SR, Kemmeren P, Roguev A, Weissman JS, Briggs SD, Krogan NJ, Strahl BD. 2007. CCR4/NOT complex associates with the proteasome and regulates histone methylation. *Proc Natl Acad Sci U S A* **104**: 5836-5841.

Laroia G, Sarkar B, Schneider RJ. 2002. Ubiquitin-dependent mechanism regulates rapid turnover of AU-rich cytokine mRNAs. *Proc Natl Acad Sci U S A* **99**: 1842-1846.

Lau NC, Kolkman A, van Schaik FM, Mulder KW, Pijnappel WW, Heck AJ, Timmers HT. 2009. Human Ccr4-Not complexes contain variable deadenylase subunits. *Biochem J* **422**: 443-453.

Lau NC, Mulder KW, Brenkman AB, Mohammed S, van den Broek NJ, Heck AJ, Timmers HT. 2010. Phosphorylation of Not4p functions parallel to BUR2 to regulate resistance to cellular stresses in Saccharomyces cerevisiae. *PLoS One* **5**: e9864.

Leppek K, Das R, Barna M. 2018. Functional 5' UTR mRNA structures in eukaryotic translation regulation and how to find them. *Nat Rev Mol Cell Biol* **19**: 158-174.

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

Li JJ, Chew GL, Biggin MD. 2017. Quantitating translational control: mRNA abundance-dependent and independent contributions and the mRNA sequences that specify them. *Nucleic Acids Res* **45**: 11821-11836.

Li Y, Minor NT, Park JK, McKearin DM, Maines JZ. 2009. Bam and Bgcn antagonize Nanos-dependent germ-line stem cell maintenance. *Proc Natl Acad Sci U S A* **106**: 9304-9309.

Li Y, Zhang Q, Carreira-Rosario A, Maines JZ, McKearin DM, Buszczak M. 2013. Mei-p26 cooperates with Bam, Bgcn and Sxl to promote early germline development in the Drosophila ovary. *PLoS One* **8**: e58301.

Lima SA, Chipman LB, Nicholson AL, Chen YH, Yee BA, Yeo GW, Coller J, Pasquinelli AE. 2017. Short poly(A) tails are a conserved feature of highly expressed genes. *Nat Struct Mol Biol* **24**: 1057-1063.

Liu Y, Beyer A, Aebersold R. 2016. On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell* **165**: 535-550.

Livneh I, Cohen-Kaplan V, Cohen-Rosenzweig C, Avni N, Ciechanover A. 2016. The life cycle of the 26S proteasome: from birth, through regulation and function, and onto its death. *Cell Research* **26**: 869.

Loh B, Jonas S, 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.

Lukong KE, Chang KW, Khandjian EW, Richard S. 2008. RNA-binding proteins in human genetic disease. *Trends Genet* **24**: 416-425.

Lunde BM, Moore C, Varani G. 2007. RNA-binding proteins: modular design for efficient function. *Nat Rev Mol Cell Biol* **8**: 479-490.

Lykke-Andersen J, Bennett EJ. 2014. Protecting the proteome: Eukaryotic cotranslational quality control pathways. *J Cell Biol* **204**: 467-476.

Lykke-Andersen J, Wagner E. 2005. Recruitment and activation of mRNA decay enzymes by two ARE-mediated decay activation domains in the proteins TTP and BRF-1. *Genes Dev* **19**: 351-361.

Macrae IJ, Zhou K, Li F, Repic A, Brooks AN, Cande WZ, Adams PD, Doudna JA. 2006. Structural basis for double-stranded RNA processing by Dicer. *Science* **311**: 195-198.

Martin KC, Ephrussi A. 2009. mRNA localization: gene expression in the spatial dimension. *Cell* **136**: 719-730.

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.

Matsuda R, Ikeuchi K, Nomura S, Inada T. 2014. Protein quality control systems associated with no-go and nonstop mRNA surveillance in yeast. *Genes Cells* **19**: 1-12.

Mauxion F, Faux C, Seraphin B. 2008. The BTG2 protein is a general activator of mRNA deadenylation. *EMBO J* **27**: 1039-1048.

McKearin D, Ohlstein B. 1995. A role for the Drosophila bag-of-marbles protein in the differentiation of cystoblasts from germline stem cells. *Development* **121**: 2937-2947.

Mersman DP, Du HN, Fingerman IM, South PF, Briggs SD. 2009. Polyubiquitination of the demethylase Jhd2 controls histone methylation and gene expression. *Genes Dev* **23**: 951-962.

Metzger MB, Pruneda JN, Klevit RE, Weissman AM. 2014. RING-type E3 ligases: Master manipulators of E2 ubiquitin-conjugating enzymes and ubiquitination. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1843**: 47-60.

Miller JE, Reese JC. 2012. Ccr4-Not complex: the control freak of eukaryotic cells. *Critical Reviews in Biochemistry and Molecular Biology* **47**: 315-333.

Moore MJ. 2005. From birth to death: the complex lives of eukaryotic mRNAs. *Science* **309**: 1514-1518.

Moore MJ, Proudfoot NJ. 2009. Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell* **136**: 688-700.

Mulder KW, Inagaki A, Cameroni E, Mousson F, Winkler GS, De Virgilio C, Collart MA, Timmers HT. 2007. Modulation of Ubc4p/Ubc5p-mediated stress responses by the

RING-finger-dependent ubiquitin-protein ligase Not4p in Saccharomyces cerevisiae. *Genetics* **176**: 181-192.

Nasertorabi F, Batisse C, Diepholz M, Suck D, Bottcher B. 2011. Insights into the structure of the CCR4-NOT complex by electron microscopy. *FEBS Lett* **585**: 2182-2186.

Nguyen TN, Padman BS, Lazarou M. 2016. Deciphering the Molecular Signals of PINK1/Parkin Mitophagy. *Trends Cell Biol* **26**: 733-744.

Nicholson KM, Anderson NG. 2002. The protein kinase B/Akt signalling pathway in human malignancy. *Cell Signal* **14**: 381-395.

Ohlstein B, McKearin D. 1997. Ectopic expression of the Drosophila Bam protein eliminates oogenic germline stem cells. *Development* **124**: 3651-3662.

Okazaki N, Okazaki K, Watanabe Y, Kato-Hayashi M, Yamamoto M, Okayama H. 1998. Novel factor highly conserved among eukaryotes controls sexual development in fission yeast. *Mol Cell Biol* **18**: 887-895.

Panasenko O, Collart MA. 2012. Presence of Not5 and ubiquitinated Rps7A in polysome fractions depends upon the Not4 E3 ligase. *Molecular microbiology* **83**: 640-653.

Panasenko OO. 2014. The role of the E3 ligase Not4 in cotranslational quality control. *Front Genet* **5**: 141.

Panasenko OO, Collart MA. 2011. Not4 E3 ligase contributes to proteasome assembly and functional integrity in part through Ecm29. *Mol Cell Biol* **31**: 1610-1623.

Panasenko OO, David FP, Collart MA. 2009. Ribosome association and stability of the nascent polypeptide-associated complex is dependent upon its own ubiquitination. *Genetics* **181**: 447-460.

Petit AP, Wohlbold L, Bawankar P, Huntzinger E, Schmidt S, Izaurralde E, Weichenrieder O. 2012. The structural basis for the interaction between the CAF1 nuclease and the NOT1 scaffold of the human CCR4-NOT deadenylase complex. *Nucleic Acids Res* **40**: 11058-11072.

Pisarev AV, Skabkin MA, Pisareva VP, Skabkina OV, Rakotondrafara AM, Hentze MW, Hellen CUT, Pestova TV. 2010. The Role of ABCE1 in Eukaryotic Posttermination Ribosomal Recycling. *Molecular Cell* **37**: 196-210.

Pisareva VP, Skabkin MA, Hellen CUT, Pestova TV, Pisarev AV. 2011. Dissociation by Pelota, Hbs1 and ABCE1 of mammalian vacant 80S ribosomes and stalled elongation complexes. *The EMBO Journal* **30**: 1804.

Preissler S, Reuther J, Koch M, Scior A, Bruderek M, Frickey T, Deuerling E. 2015. Not4-dependent translational repression is important for cellular protein homeostasis in yeast. *EMBO J* **34**: 1905-1924.

Presnyak V, Alhusaini N, Chen YH, Martin S, Morris N, Kline N, Olson S, Weinberg D, Baker KE, Graveley BR et al. 2015. Codon optimality is a major determinant of mRNA stability. *Cell* **160**: 1111-1124.

Raghavan A, Ogilvie RL, Reilly C, Abelson ML, Raghavan S, Vasdewani J, Krathwohl M, Bohjanen PR. 2002. Genome-wide analysis of mRNA decay in resting and activated primary human T lymphocytes. *Nucleic Acids Research* **30**: 5529-5538.

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

Ramanathan A, Robb GB, Chan SH. 2016. mRNA capping: biological functions and applications. *Nucleic Acids Res* **44**: 7511-7526.

Sandler H, Kreth J, Timmers HT, Stoecklin G. 2011. Not1 mediates recruitment of the deadenylase Caf1 to mRNAs targeted for degradation by tristetraprolin. *Nucleic Acids Res* **39**: 4373-4386.

Schuller AP, Green R. 2018. Roadblocks and resolutions in eukaryotic translation. *Nat Rev Mol Cell Biol* **19**: 526-541.

Seufert W, Jentsch S. 1990. Ubiquitin-conjugating enzymes UBC4 and UBC5 mediate selective degradation of short-lived and abnormal proteins. *EMBO J* **9**: 543-550.

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

Shao S, von der Malsburg K, Hegde RS. 2013. Listerin-dependent nascent protein ubiquitination relies on ribosome subunit dissociation. *Mol Cell* **50**: 637-648.

Sharov AA, Sharova LV, Shaik N, Nedorezov T, Piao Y, Ko MSH. 2008. Database for mRNA Half-Life of 19 977 Genes Obtained by DNA Microarray Analysis of Pluripotent and Differentiating Mouse Embryonic Stem Cells. *DNA Research* **16**: 45-58.

Shen R, Weng C, Yu J, Xie T. 2009. eIF4A controls germline stem cell self-renewal by directly inhibiting BAM function in the Drosophila ovary. *Proceedings of the National Academy of Sciences* **106**: 11623.

Shi Y. 2017. Mechanistic insights into precursor messenger RNA splicing by the spliceosome. *Nat Rev Mol Cell Biol* **18**: 655-670.

Shoemaker CJ, Green R. 2011. Kinetic analysis reveals the ordered coupling of translation termination and ribosome recycling in yeast. *Proceedings of the National Academy of Sciences* **108**: E1392.

Song X, Wong MD, Kawase E, Xi R, Ding BC, McCarthy JJ, Xie T. 2004. Bmp signals from niche cells directly repress transcription of a differentiation-promoting gene, bag of marbles, in germline stem cells in the Drosophila ovary. *Development* **131**: 1353-1364.

Steitz TA. 2008. A structural understanding of the dynamic ribosome machine. *Nature Reviews Molecular Cell Biology* **9**: 242.

Stowell JAW, Webster MW, Kogel A, Wolf J, Shelley KL, Passmore LA. 2016. Reconstitution of Targeted Deadenylation by the Ccr4-Not Complex and the YTH Domain Protein Mmi1. *Cell Rep* **17**: 1978-1989.

Subtelny AO, Eichhorn SW, Chen GR, Sive H, Bartel DP. 2014. Poly(A)-tail profiling reveals an embryonic switch in translational control. *Nature* **508**: 66.

Sun HY, Kim N, Hwang CS, Yoo JY. 2015. Protein Degradation of RNA Polymerase II-Association Factor 1(PAF1) Is Controlled by CNOT4 and 26S Proteasome. *PLoS One* **10**: e0125599.

Sun L, Chen ZJ. 2004. The novel functions of ubiquitination in signaling. *Curr Opin Cell Biol* **16**: 119-126.

Suzuki A, Saba R, Miyoshi K, Morita Y, Saga Y. 2012. Interaction between NANOS2 and the CCR4-NOT deadenylation complex is essential for male germ cell development in mouse. *PLoS One* **7**: e33558.

Temme C, Zhang L, Kremmer E, Ihling C, Chartier A, Sinz A, Simonelig M, Wahle E. 2010. Subunits of the Drosophila CCR4-NOT complex and their roles in mRNA deadenylation. *RNA* **16**: 1356-1370.

Tokusumi T, Tokusumi Y, Hopkins DW, Schulz RA. 2015. Bag of Marbles controls the size and organization of the Drosophila hematopoietic niche through

interactions with the Insulin-like growth factor pathway and Retinoblastoma-family protein. *Development* **142**: 2261.

Tokusumi T, Tokusumi Y, Hopkins DW, Shoue DA, Corona L, Schulz RA. 2011. Germ line differentiation factor Bag of Marbles is a regulator of hematopoietic progenitor maintenance during Drosophila hematopoiesis. *Development* **138**: 3879.

Tokusumi T, Tokusumi Y, Schulz RA. 2018. The mir-7 and bag of marbles genes regulate Hedgehog pathway signaling in blood cell progenitors in Drosophila larval lymph glands. *genesis* **56**: e23210.

Tompa P. 2012. Intrinsically disordered proteins: a 10-year recap. *Trends Biochem Sci* **37**: 509-516.

Tsuda M, Sasaoka Y, Kiso M, Abe K, Haraguchi S, Kobayashi S, Saga Y. 2003. Conserved role of nanos proteins in germ cell development. *Science* **301**: 1239-1241.

Uchida N, Hoshino S, Katada T. 2004. Identification of a human cytoplasmic poly(A) nuclease complex stimulated by poly(A)-binding protein. *J Biol Chem* **279**: 1383-1391.

Ukleja M, Cuellar J, Siwaszek A, Kasprzak JM, Czarnocki-Cieciura M, Bujnicki JM, Dziembowski A, Valpuesta JM. 2016. The architecture of the Schizosaccharomyces pombe CCR4-NOT complex. *Nat Commun* **7**: 10433.

Van Roey K, Uyar B, Weatheritt RJ, Dinkel H, Seiler M, Budd A, Gibson TJ, Davey NE. 2014. Short linear motifs: ubiquitous and functionally diverse protein interaction modules directing cell regulation. *Chem Rev* **114**: 6733-6778.

Vicens Q, Kieft JS, Rissland OS. 2018. Revisiting the Closed-Loop Model and the Nature of mRNA 5'-3' Communication. *Mol Cell* **72**: 805-812.

Vinuesa CG, Cook MC, Angelucci C, Athanasopoulos V, Rui L, Hill KM, Yu D, Domaschenz H, Whittle B, Lambe T et al. 2005. A RING-type ubiquitin ligase family member required to repress follicular helper T cells and autoimmunity. *Nature* **435**: 452-458.

Wahle E, Winkler GS. 2013. RNA decay machines: deadenylation by the Ccr4-not and Pan2-Pan3 complexes. *Biochim Biophys Acta* **1829**: 561-570.

Webster MW, Chen YH, Stowell JAW, Alhusaini N, Sweet T, Graveley BR, Coller J, Passmore LA. 2018. mRNA Deadenylation Is Coupled to Translation Rates by the Differential Activities of Ccr4-Not Nucleases. *Mol Cell* **70**: 1089-1100 e1088.

Wells SE, Hillner PE, Vale RD, Sachs AB. 1998. Circularization of mRNA by Eukaryotic Translation Initiation Factors. *Molecular Cell* **2**: 135-140.

Wilson MA, Meaux S, van Hoof A. 2007. A genomic screen in yeast reveals novel aspects of nonstop mRNA metabolism. *Genetics* **177**: 773-784.

Winkler GS, Albert TK, Dominguez C, Legtenberg YI, Boelens R, Timmers HT. 2004. An altered-specificity ubiquitin-conjugating enzyme/ubiquitin-protein ligase pair. *J Mol Biol* **337**: 157-165.

Winkler GS, Mulder KW, Bardwell VJ, Kalkhoven E, Timmers HT. 2006. Human Ccr4-Not complex is a ligand-dependent repressor of nuclear receptor-mediated transcription. *EMBO J* **25**: 3089-3099.

Wong SQ, Behren A, Mar VJ, Woods K, Li J, Martin C, Sheppard KE, Wolfe R, Kelly J, Cebon J et al. 2014. Whole exome sequencing identifies a recurrent RQCD1 P131L mutation in cutaneous melanoma. *Oncotarget* **6**: 1115-1127.

Wu Z, Wang Y, Lim J, Liu B, Li Y, Vartak R, Stankiewicz T, Montgomery S, Lu B. 2018. Ubiquitination of ABCE1 by NOT4 in Response to Mitochondrial Damage Links Co-

translational Quality Control to PINK1-Directed Mitophagy. *Cell Metab* **28**: 130-144 e137.

Yamashita A, Chang T-C, Yamashita Y, Zhu W, Zhong Z, Chen C-YA, Shyu A-B. 2005. Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. *Nature Structural &Amp; Molecular Biology* **12**: 1054.

Yi H, Park J, Ha M, Lim J, Chang H, Kim VN. 2018. PABP Cooperates with the CCR4-NOT Complex to Promote mRNA Deadenylation and Block Precocious Decay. *Mol Cell* **70**: 1081-1088 e1085.

Zekri L, Kuzuoglu-Ozturk D, Izaurralde E. 2013. GW182 proteins cause PABP dissociation from silenced miRNA targets in the absence of deadenylation. *EMBO J* **32**: 1052-1065.

Zinder JC, Lima CD. 2017. Targeting RNA for processing or destruction by the eukaryotic RNA exosome and its cofactors. *Genes Dev* **31**: 88-100.

9. Acknowledgements

I want to express my deepest gratitude to my supervisor Prof. Dr. Elisa Izaurralde, who made all of this work possible.

I would like to thank:

Dr. Dipankar Bhandari and Dr. Catia Igreja for the constant support and guidance throughout my PhD and Dr. Lara Wohlbold for advice and comments on this thesis.

Dr. Oliver Weichenrieder for the discussions and collaboration, as well as commenting on the manuscript of this thesis.

Prof. Dr. Ralf J. Sommer for supporting our lab through this difficult time.

Prof. Dr. Ralf-Peter Jansen at the Eberhard Karls University Tübingen for supervising my PhD project, and Dr. Silke Wiesner at Universität Regensburg for being a member of my thesis advisory committee.

Prof. Dr. Thorsten Stafforst at the Eberhard Karls University Tübingen for agreeing to evaluate my thesis.

Prof. Dr. Doron Rapaport and Dr. Martin Bayer for agreeing to join my PhD defense committee.

Dr. Sarah Danes for her help throughout my PhD as IMPRS Program Coordinator.

Maria Gölz and Sibylle Patheiger for their help with all the administrative work over the years.

Dr. Heike Budde, Sigrun Helms and Maria Fauser for introducing me to the lab and for the technical support.

Dr. Tobias Raisch and Dr. Annamaria Sgromo for collaborations on the NOT4 and Bam projects.

Catrin Weiler, Gabriele Wagner, Nadine Weiss, Max Widmann, Ida Axtmann, Marie Palaj, Marion Weber and Yvonne Wörn for their technical help in the lab.

I would like to thank all the former and current members of Department II for the good times inside and outside the lab: Dr. Vincenzo Ruscica, Felix Räsch, Dr. Aoife Hanet, Dr. Maria Fernandez, Ramona Weber, Simone Larivera, Michelle Noble, Dr. Elena Khazina, Dr. Ying Chen, Sowndarya Muthukumar, Dr. Eugene Valkov, Dr. Praveen Bawankar, Dr. Duygu Kuzuoglu-Öztürk, Dr. Chung-Te Chang, Min-Yi Chung, Dr. Moritz Thran, Dr. Stefan Grüner,

Prof. Dr. Stefanie Jonas, Dr. Daniel Peter, Dr. Steffen Schmidt, Yevgen Levdansky, Ayshwarya Seenivasan, Stefanie Becker and Alina Stein.

Especially, I want to thank Gergely Jokhel for his endless support and patience throughout these years.

Finally, I wish to express my love and gratitude to my family and friends for always being on my side and encouraging me. I am most grateful to my grandmother, Dr. Sánta Istvánné Turza Rózsa, who always supported my aspiration to become a scientist.

10. Curriculum vitae

Csilla Keskeny

Work experience

09/2014-10/2019	**PhD in Biochemistry;** Max Planck Institute for Developmental Biology, Tübingen, Germany. Supervisor: Prof. Elisa Izaurralde. Title of PhD thesis: ”The role of the RNA-associated proteins NOT4 and Bam in mRNA decay”.

Education

09/2012-06/2014	**MSc in Biology;** Molecular, Immune, and Microbiology specialization; University of Szeged, Hungary. Supervisor: Dr. Mónika Kiricsi. Title of thesis: ”Molecular events in tumor cells treated with silver nanoparticles“.
09/2008-06/2011	**Professional Translator of English and Hungarian,** Specialized in Sciences; University of Szeged, Hungary
09/2007-06/2012	**BSc in Biology;** Cell and Molecular Biology specialization; University of Szeged, Hungary. Supervisor: Dr. Mónika Kiricsi. Title of thesis: ”Epigenetic study of tumor-associated myofibroblasts”.

Scholarships received

09/2014-09/2018	International Max Planck Research School (IMPRS) PhD fellowship

Publications and presentations

Publications	Keskeny C, Raisch T, Sgromo A, Igreja C, Bhandari D, Weichenrieder O, Izaurralde E. **A conserved CAF40-binding motif in metazoan NOT4 mediates association with the CCR4-NOT complex.** Genes Dev. 2019;33(3-4):236-52. Sgromo A, Raisch T, Backhaus C, Keskeny C, Alva V, Weichenrieder O, Izaurralde E. **Drosophila Bag-of-marbles directly interacts with the CAF40 subunit of the CCR4-NOT complex to elicit repression of mRNA targets.** RNA. 2018;24(3):381-95. Kovacs D, Igaz N, Keskeny C, Belteky P, Toth T, Gaspar R, et al. **Silver nanoparticles defeat p53-positive and p53-negative osteosarcoma cells by triggering mitochondrial stress and apoptosis.** Sci Rep. 2016;6:27902.

Presentations (selection)	Oral presentation at the CSHL Translational Control meeting (4-8 September 2018, Cold Spring Harbor, USA): “Human NOT4 directly interacts with CAF40 and NOT1 to associate with the CCR4-NOT complex.”
	Poster presentation at the EMBL Symposium The Complex Life of mRNA (5-8 October 2016, Heidelberg, Germany): “Metazoan NOT4 is recruited to the CCR4-NOT complex via interactions with CAF40 and NOT1.”

Tübingen, 21.07.2019

11. Appendix: Original manuscripts of discussed publications

1. Keskeny C, Raisch T, Sgromo A, Igreja C, Bhandari D, Weichenrieder O, Izaurralde E. **A conserved CAF40-binding motif in metazoan NOT4 mediates association with the CCR4-NOT complex.** Genes Dev. 2019;33(3-4):236-52.

2. Sgromo A, Raisch T, Backhaus C, Keskeny C, Alva V, Weichenrieder O, Izaurralde E. **Drosophila Bag-of-marbles directly interacts with the CAF40 subunit of the CCR4-NOT complex to elicit repression of mRNA targets.** RNA. 2018;24(3):381-95.

A conserved CAF40-binding motif in metazoan NOT4 mediates association with the CCR4–NOT complex

Csilla Keskeny, Tobias Raisch,[1] Annamaria Sgromo,[2] Cátia Igreja, Dipankar Bhandari, Oliver Weichenrieder, and Elisa Izaurralde[3]

Department of Biochemistry, Max Planck Institute for Developmental Biology, D-72076 Tübingen, Germany

The multisubunit CCR4–NOT mRNA deadenylase complex plays important roles in the posttranscriptional regulation of gene expression. The NOT4 E3 ubiquitin ligase is a stable component of the CCR4–NOT complex in yeast but does not copurify with the human or *Drosophila melanogaster* complex. Here we show that the C-terminal regions of human and *D. melanogaster* NOT4 contain a conserved sequence motif that directly binds the CAF40 subunit of the CCR4–NOT complex (CAF40-binding motif [CBM]). In addition, nonconserved sequences flanking the CBM also contact other subunits of the complex. Crystal structures of the CBM–CAF40 complex reveal a mutually exclusive binding surface for NOT4 and Roquin or Bag of marbles mRNA regulatory proteins. Furthermore, CAF40 depletion or structure-guided mutagenesis to disrupt the NOT4–CAF40 interaction impairs the ability of NOT4 to elicit decay of tethered reporter mRNAs in cells. Together with additional sequence analyses, our results reveal the molecular basis for the association of metazoan NOT4 with the CCR4–NOT complex and show that it deviates substantially from yeast. They mark the NOT4 ubiquitin ligase as an ancient but nonconstitutive cofactor of the CCR4–NOT deadenylase with potential recruitment and/or effector functions.

[*Keywords*: deadenylation; mRNA decay; translational repression; ubiquitination]

Supplemental material is available for this article.

Received September 20, 2018; revised version accepted December 10, 2018.

The CCR4–NOT complex plays a central role in the posttranscriptional regulation of gene expression by catalyzing the removal of the mRNA poly(A) tail, thereby repressing translation and promoting mRNA degradation (Wahle and Winkler 2013; Collart 2016). In addition, the CCR4–NOT complex has the ability to repress translation independently of deadenylation (Cooke et al. 2010; Chekulaeva et al. 2011; Bawankar et al. 2013).

The CCR4–NOT complex (Fig. 1A) is a multisubunit complex (Chen et al. 2001; Lau et al. 2009; Temme et al. 2010) that assembles on the NOT1 scaffold protein, which consists of several α-helical domains that serve to dock the other subunits of the complex (Bawankar et al. 2013). Deadenylation is carried out by two interacting deadenylases; namely, CCR4 and CAF1. They dock onto a central α-helical domain in NOT1 (labeled "MIF4G"), forming the "catalytic module" of the complex (Basquin et al. 2012; Petit et al. 2012). The C-terminal end of NOT1 contains the NOT1 superfamily homology domain (SHD), which is another α-helical domain that interacts with the NOT2–NOT3 heterodimer to form the "NOT module" of the complex (Bhaskar et al. 2013; Boland et al. 2013). The catalytic module and the NOT module are connected by the CAF40-binding domain of NOT1 (labeled "CN9BD") and a connector domain (labeled "MIF4G-C") of unknown function (Chen et al. 2014; Mathys et al. 2014; Raisch et al. 2018). Both the NOT module and the CAF40 subunit of the CCR4–NOT complex have been reported as important peptide-docking sites for the recruitment of the complex by mRNA-associated proteins. The NOT module provides binding sites for Bicaudal-C (Chicoine et al. 2007), Nanos (Bhandari et al. 2014; Raisch et al. 2016), and Roquin (Sgromo et al. 2017). CAF40 is known to be contacted by Roquin (Sgromo et al. 2017), Bag of marbles (Bam) (Sgromo et al. 2018), and TTP (Bulbrook et al. 2018) as well as the GW182/TNRC6 family of proteins that mediates microRNA-mediated mRNA repression and decay (Chen et al. 2014; Mathys et al. 2014).

Present addresses: [1]Department of Structural Biochemistry, Max Planck Institute of Molecular Physiology, D-44227 Dortmund, Germany; [2]IMBA (Institute of Molecular Biotechnology) GmbH, 1030 Vienna, Austria

[3]Deceased April 30, 2018.

Corresponding authors: oliver.weichenrieder@tuebingen.mpg.de, dipankar.bhandari@tuebingen.mpg.de

Article published online ahead of print. Article and publication date are online at http://www.genesdev.org/cgi/doi/10.1101/gad.320952.118. Freely available online through the *Genes & Development* Open Access option.

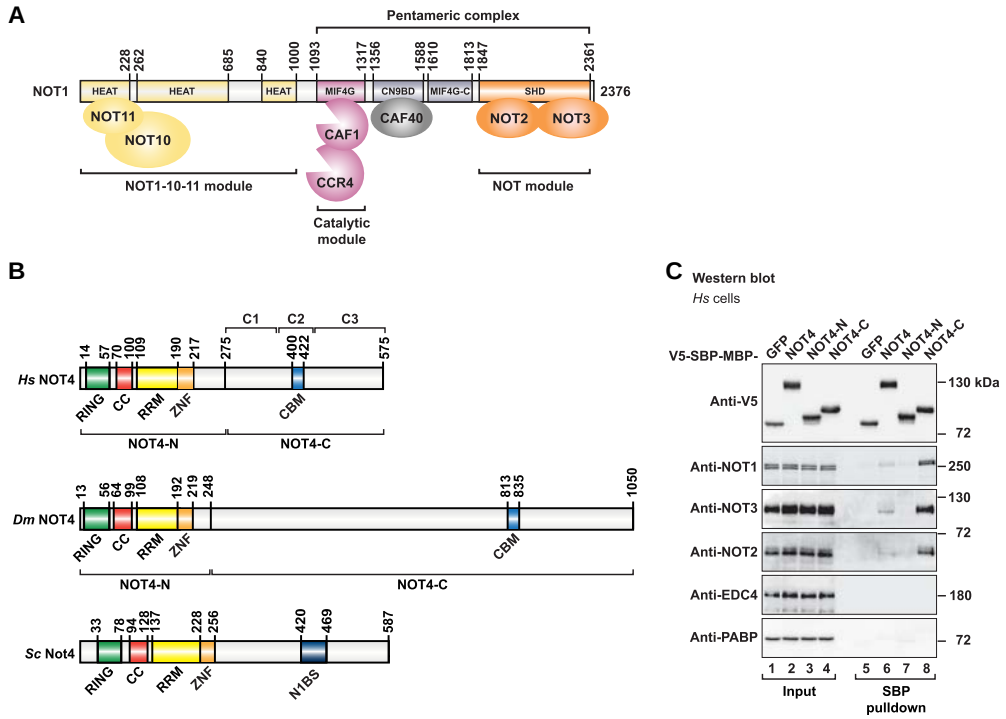


Figure 1. Human NOT4 interacts with the CCR4–NOT complex. (*A*) Composition of the human CCR4–NOT complex. The NOT1 scaffold protein contains N-terminal α-helical domains (classified as HEAT repeat domains) that interact with NOT10 and NOT11 to form the NOT1–10–11 module. NOT1 furthermore contains a central HEAT repeat domain (MIF4G) that binds CAF1 and CCR4 to form the catalytic module, an α-helical bundle that interacts with CAF40 (CN9BD), a connector domain (MIF4G-C), and a NOT1 SHD that forms the NOT module together with NOT2 and NOT3. A "pentameric" complex lacking CCR4 and the NOT1–10–11 module can be assembled from recombinant human CCR4–NOT proteins (Sgromo et al. 2017). (*B*) Domain composition of NOT4 proteins. The conserved N-terminal region of NOT4 (NOT4-N) comprises a RING-type E3 ubiquitin ligase domain (RING), a positively charged linker with coiled-coil propensity (CC), an RNA recognition motif (RRM) domain, and a C3H1-type zinc finger domain (ZNF). The nonconserved C-terminal region of NOT4 (NOT4-C) was found to interact with the CCR4–NOT complex. To map the interactions, *Homo sapiens* (*Hs*) NOT4-C was subdivided into three regions: C1 (residues P275–S376), C2 (residues E377–Q428), and C3 (residues P429–A575). A CAF40-binding motif (CBM) was identified in the C2 region. The CBM is conserved in metazoan NOT4, including *Dm* NOT4, but is not conserved in yeasts. Instead, *Sc* Not4-C harbors a previously characterized binding site for *Sc* Not1 (N1BS) (Bhaskar et al. 2015). (*C*) SBP pull-down of endogenous human NOT proteins with V5-SBP-MBP (V5-streptavidin-binding peptide-maltose-binding protein)-tagged *Hs* NOT4 from HEK293T cell lysates. V5-SBP-GFP-MBP served as negative control. Input samples correspond to 3% of the total lysate for V5-tagged proteins and 2% of the total lysate for NOT proteins. Pull-down samples correspond to 3% of the total pull-down for V5-tagged proteins and 35% of the total pull-down for NOT proteins. The mRNA decapping factor EDC4 and the poly(A)-binding protein PABP served as negative controls.

The N-terminal portion of NOT1 is less well conserved than its C-terminal portion (Basquin et al. 2012) and serves to dock NOT10 and NOT11 as additional subunits of the complex in metazoan species (Bawankar et al. 2013; Mauxion et al. 2013). The CCR4–NOT complexes of *Saccharomyces cerevisiae* (*Sc*) and *Schizosaccharomyces pombe* lack NOT10 and NOT11 proteins. Furthermore, these CCR4–NOT complexes are special because they contain NOT4 as an integral component (Bai et al. 1999; Chen et al. 2001; Nasertorabi et al. 2011; Stowell et al. 2016; Ukleja et al. 2016).

NOT4 (Fig. 1B) is an evolutionarily conserved E3 ubiquitin ligase that contains a RING domain, a linker region with coiled-coil propensity (CC), an RNA recognition motif (RRM) domain, and a C3H1-type zinc finger domain (ZNF). Together, they define the conserved N-terminal region of NOT4 (NOT4-N) (Fig. 1B). The C-terminal region of NOT4 (NOT4-C) (Fig. 1B) is predicted to be unstructured, and its sequence and length are not conserved among NOT4 proteins (The UniProt Consortium 2018). NOT4 causes the ubiquitination of diverse proteins in yeast and also human cells, targeting them for proteasomal degradation or resulting in regulatory changes. Ubiquitination targets include the nascent polypeptide-associated complex NAC (Panasenko et al. 2006), the histone demethylase JHD2 (Mersman et al. 2009), the transcription factor YAP1 (Gulshan et al. 2012), the master regulator of meiosis Mei2 (Simonetti et al. 2017), the cyclin C subunit of the Mediator complex (Cooper et al. 2012), the small ribosomal protein RPS7A (Panasenko and Collart 2012), and the cotranslational quality control factor ABCE1 (Wu et al. 2018). NOT4 has been implicated in cotranslational mRNA quality control and translational repression in the context of stalled ribosomes, such as in the "No-Go"

mRNA decay pathway (Dimitrova et al. 2009; Matsuda et al. 2014; Panasenko 2014; Preissler et al. 2015; Wu et al. 2018).

A crystal structure demonstrated how *Sc* Not4 interacts with the SHD of *Sc* Not1 via an elongated polypeptide from the C-terminal region of *Sc* Not4 (Bhaskar et al. 2015). Using yeast two-hybrid assays, the human NOT4 and NOT1 proteins (*Homo sapiens* [*Hs*] NOT4 and *Hs* NOT1) were also shown to interact via the C-terminal portion of *Hs* NOT1 (Albert et al. 2002). However, the *Sc* Not1-binding sequence of *Sc* Not4 is only partially conserved, at best (Bhaskar et al. 2015), and NOT4 was not detected in mass spectrometric analyses of the native human and *Drosophila melanogaster* (*Dm*) CCR4–NOT complexes (Lau et al. 2009; Temme et al. 2010). This raised the question of whether NOT4 should be regarded as a component or cofactor of the CCR4–NOT complex in metazoans and how it would be recruited to the complex in species other than *S. cerevisiae*.

We therefore performed pull-down experiments from human cell extracts and with purified bacterially expressed proteins to identify and map interactions of human NOT4 with the CCR4–NOT complex. Assisted by alignments of metazoan NOT4 proteins, we uncovered a 23-amino-acid peptide motif in NOT4-C that binds to the CAF40 subunit of the CCR4–NOT complex and hence was termed the NOT4 CAF40-binding motif (CBM). Crystal structures of the CBM–CAF40 complex identified critical contacts required in human and *Dm* S2 cells for an efficient recruitment of NOT4 to the CCR4–NOT complex and for NOT4-mediated mRNA deadenylation and decay via the CCR4–NOT complex. Consequently, NOT4 emerges as a nonconstitutive cofactor of the CCR4–NOT complex in metazoans with a conserved mode of interaction via the CAF40 subunit.

Results

Hs *NOT4-C shows a stable interaction with the CCR4–NOT complex*

To investigate whether and how *Hs* NOT4 interacts with the CCR4–NOT complex in HEK293T cells, we expressed *Hs* NOT4 with a V5-SBP-MBP (V5-streptavidin-binding peptide-maltose-binding protein) tag in HEK293T cells and performed SBP pull-down assays. In agreement with previous reports (Lau et al. 2009; Temme et al. 2010), full-length *Hs* NOT4 failed to pull down the endogenous CCR4–NOT complex efficiently (Fig. 1C, lane 6). *Hs* NOT4-C, however, showed a stable interaction with the CCR4–NOT complex, as indicated by the detection of endogenous NOT1, NOT2, and NOT3 subunits in the pull-down fraction (Fig. 1C, lane 8). This is consistent with previous yeast two-hybrid experiments (Albert et al. 2002). In contrast, *Hs* NOT4-N did not interact with the CCR4–NOT complex (Fig. 1C, lane 7). The lack of an efficient interaction with the full-length protein remains unexplained but hints at a possible regulation of NOT4-C binding by NOT4-N. Additional SBP pull-down experiments showed that it is the presence of the positively charged CC linker and of the RRM domain in NOT4-N that prevents NOT4-C from interacting with the CCR4–NOT complex (Supplemental Fig. S1A,B).

Tethered Hs *NOT4 causes deadenylation-dependent mRNA decay*

To address the relevance of an interaction between NOT4 and the CCR4–NOT complex with a functional assay, we investigated the molecular consequences resulting from the presence of NOT4 in the context of an mRNA. Usually, the recruitment of the CCR4–NOT complex to an mRNA target promotes its deadenylation-dependent decay (Wahle and Winkler 2013). It is unknown, however, whether NOT4 can bind to an mRNA despite the presence of putative and conserved RNA-binding domains in NOT4-N (CC-RRM-ZNF) (Fig. 1B). In the absence of known mRNA targets, we therefore used a tethering assay to direct NOT4 toward the 3′ untranslated region (UTR) of defined reporter mRNAs and tested them for NOT4-dependent deadenylation and decay.

In a first series of experiments, we tethered MS2-tagged *Hs* NOT4 to a β-globin mRNA reporter containing six MS2-binding sites in the 3′ UTR (β-globin-6xMS2bs) (Lykke-Andersen et al. 2000). Even though full-length *Hs* NOT4 did not associate with the CCR4–NOT complex in SBP pull-down assays (Fig. 1C, lane 6), we found that tethered MS2-HA-*Hs* NOT4 caused a substantial reduction of the β-globin-6xMS2bs mRNA level compared with the negative control MS2-HA (Fig. 2A,B). Tethering *Hs* NOT4-C also reduced mRNA levels, whereas tethering *Hs* NOT4-N had no effect (Fig. 2A,B). All *Hs* NOT4 fragments were expressed at comparable levels (Fig. 2C), and none of them affected the expression of the control β-globin mRNA lacking MS2-binding sites (Fig. 2B). Regarding full-length *Hs* NOT4, the discrepancy with the SBP pull-down assay (Fig. 1C) might be rationalized by a higher sensitivity of the tethering assay for weak and possibly transient interactions or, alternatively, by conformational changes of *Hs* NOT4 in the presence of RNA that promote the availability of NOT4-C to the deadenylase complex.

In a second series of experiments, we also tethered MS2-tagged *Hs* NOT4 to another reporter mRNA that encoded Renilla luciferase (R-Luc-6xMS2bs) instead of β-globin, allowing for the quantification of protein abundance in addition to mRNA levels. In agreement with the β-globin mRNA reporter, we observed a clear reduction of both Renilla luciferase mRNA and Renilla luciferase protein levels in the case of tethered *Hs* NOT4 or *Hs* NOT4-C (Supplemental Fig. S2A,B). By comparison, a Renilla luciferase mRNA lacking MS2-binding sites was not affected by the expression of MS2-tagged full-length *Hs* NOT4 or its fragments (Supplemental Fig. S2C,D).

To verify whether the reduction of reporter mRNA levels upon tethering *Hs* NOT4 or *Hs* NOT4-C resulted from deadenylation-dependent decay, we overexpressed a GFP-tagged catalytically inactive mutant of the human mRNA decapping enzyme DCP2 (GFP-DCP2 mut; E148Q). The presence of this mutant is known to impair mRNA decapping in a dominant-negative manner and hence

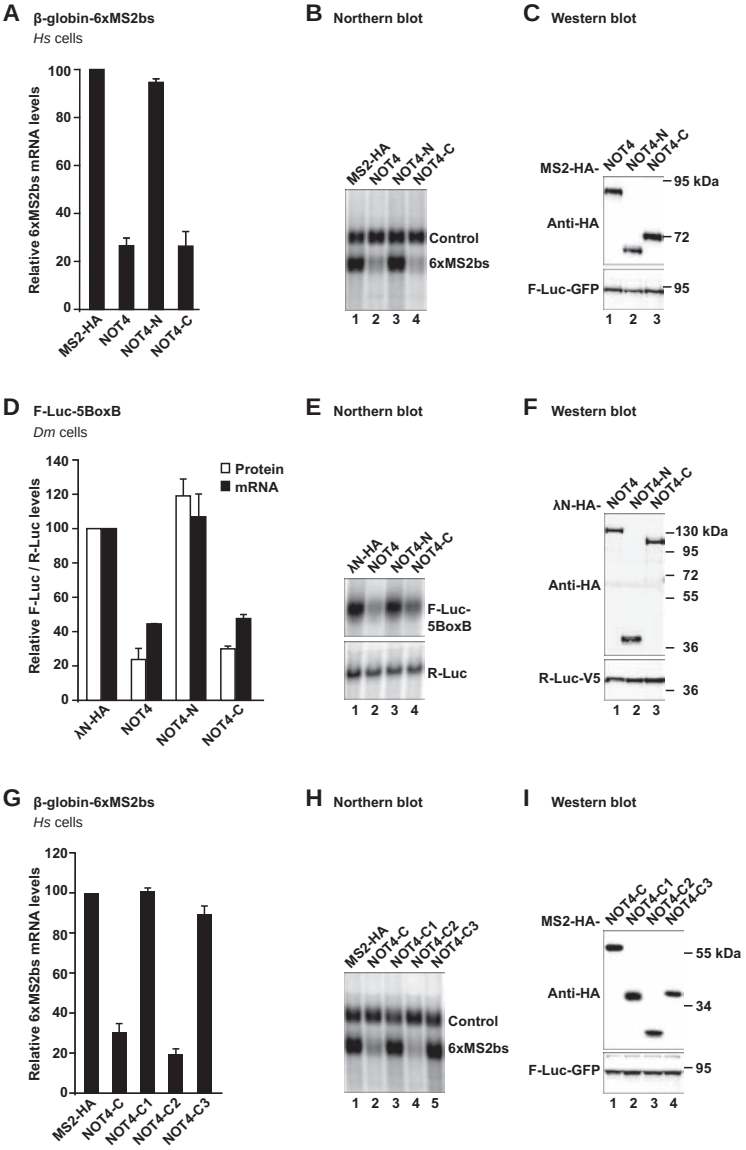


Figure 2. Metazoan NOT4 induces degradation of tethered mRNA reporters. (*A–C*) Tethering assay with *Hs* NOT4 and a β-globin mRNA reporter in HEK293T cells. *Hs* NOT4 or its fragments carried an N-terminal MS2-HA tag. β-Globin mRNA served as a reporter and contained six binding sites for the MS2 protein (6xMS2bs). β-Globin-GAPDH mRNA served as a reference and transfection control (control). (*A*) mRNA levels of the β-globin-6xMS2bs mRNA reporter normalized to the reference and plotted with respect to the values obtained from the expression of MS2-HA alone (set to 100). Error bars correspond to standard deviations. $n = 3$. (*B*) Northern blot of representative RNA samples. (*C*) Western blot demonstrating equal expression of MS2-HA-tagged proteins with F-Luc-GFP as a transfection control. (*D–F*) Tethering assay with *Dm* NOT4 and a luciferase reporter in *Dm* S2 cells. *Dm* NOT4 or its fragments carried an N-terminal λN-HA tag. Firefly luciferase mRNA served as a reporter and contained five BoxB binding sites for the λN peptide (F-Luc-5BoxB). Renilla luciferase mRNA served as a reference and transfection control (R-Luc). (*D*) F-Luc activity (white bars) and mRNA levels (black bars) normalized to the reference and plotted with respect to the values obtained from the expression of λN-HA alone (set to 100). Error bars correspond to standard deviations. $n = 3$. (*E*) Northern blot. (*F*) Western blot with R-Luc-V5 as a transfection control. (*G–I*) Tethering assay with fragments of *Hs* NOT4-C and the β-globin mRNA reporter. (*G*) Relative mRNA levels, with error bars corresponding to standard deviations. $n = 3$. For additional details, see *A*. (*H*) Northern blot. (*I*) Western blot.

5′-to-3′ mRNA decay by XRN1 (Loh et al. 2013; Bhandari et al. 2014; Chang et al. 2014; Kuzuoğlu-Öztürk et al. 2016; Sgromo et al. 2017). Indeed, we observed the accumulation of a shorter deadenylated decay intermediate of the β-globin-6xMS2bs reporter mRNA upon tethering *Hs* NOT4 or upon tethering the *Hs* Nanos2 mRNA-binding protein (Bhandari et al. 2014), which served as a positive control (Supplemental Fig. S3A–C). We therefore attributed reporter mRNA decay to the recruitment of the CCR4–NOT complex, although we could not formally exclude contributions from other deadenylases at this stage.

The capacity of NOT4-C to mediate tethered mRNA decay is conserved in metazoans

The C-terminal region of NOT4 is not conserved in sequence and length (Fig. 1B). We therefore investigated the functionality of *Dm* NOT4 as an example from an invertebrate species and to allow more general conclusions on NOT4 recruitment to the CCR4–NOT complex in metazoans. For *Dm* NOT4, we used a λN-based tethering assay in *Dm* S2 cells with an F-Luc-5BoxB reporter mRNA (Behm-Ansmant et al. 2006). Similar to *Hs* NOT4 and despite highly divergent sequences of NOT4-C, tethered *Dm* NOT4 and *Dm* NOT4-C efficiently mediated reporter mRNA decay (Fig. 2D,E). *Dm* NOT4 and its fragments were expressed at equal levels (Fig. 2F), and none of them affected the expression of an F-Luc reporter lacking the BoxB sites (Supplemental Fig. S2E,F).

Again, reporter mRNA decay was deadenylation-dependent. Deadenylated F-Luc-5BoxB reporter mRNA was stabilized in the presence of tethered *Dm* NOT4, when a GFP-tagged *Dm* DCP1 mutant (GFP-DCP1 mut; R70G, L71S, N72S, and T73G) known to prevent mRNA decapping in a dominant-negative manner (Chang et al. 2014; Kuzuoğlu-Öztürk et al. 2016) was overexpressed (Supplemental Fig. S3D–F). Tethered GW182 protein was used

as a positive control (Kuzuoğlu-Öztürk et al. 2016). Together, our results indicate that the interaction of NOT4-C with the CCR4–NOT complex is conserved between humans and flies.

Hs *NOT4 directly interacts with the NOT1 and CAF40 subunits of the CCR4–NOT complex*

To test whether the interaction between *Hs* NOT4 and the CCR4–NOT complex is direct, we expressed MBP-tagged *Hs* NOT4, *Hs* NOT4-N, and *Hs* NOT4-C in *Escherichia coli* and performed pull-down experiments with a reconstituted and purified subcomplex of human CCR4–NOT components that we had described previously (Sgromo et al. 2017) and that we here call the "pentameric" complex (Figs. 1A, 3A). This subcomplex comprises the C-terminal portion of NOT1 together with the CAF1 and CAF40 subunits and the C-terminal fragments of NOT2 and NOT3. Indeed, we observed a direct interaction with the pentameric complex. In contrast to the result from the SBP pull-down experiment in HEK293T cells (Fig. 1C, lane 6), the interaction occurred even with the full-length *Hs* NOT4 (Fig. 3A, lane 7). Furthermore, *Hs* NOT4-C interacted with the pentameric complex as efficiently as the recombinant full-length protein (Fig. 3A, lane 9), whereas recombinant *Hs* NOT4-N did not interact (Fig. 3A, lane 8).

To map the interactions to individual components of the CCR4–NOT complex, we also used previously described smaller subassemblies of the complex (Petit et al. 2012; Boland et al. 2013; Chen et al. 2014; Sgromo et al. 2017) in MBP pull-down experiments (Figs. 1A, 3B). These were the NOT1–NOT10–NOT11 module (comprising the N-terminal region of NOT1 and NOT10 and the C-terminal half of NOT11), CAF1 bound to the central α-helical domain in NOT1 (labeled "MIF4G"), CAF40 bound to the CAF40-binding domain of NOT1 (labeled "CN9BD"), the NOT1 connector domain (labeled "MIF4G-C"), and the NOT module (comprising the C-terminal regions of NOT1, NOT2, and NOT3). We detected interactions of *Hs* NOT4 with both the CAF40-containing subcomplex and the NOT module (Fig. 3B, lanes 13,17), pointing to multiple NOT4-binding sites within the CCR4–NOT complex. Finally, we narrowed down these interactions even further, primarily to the NOT1-SHD (Fig. 3C, lane 7) and CAF40 alone (Fig. 3C, lane 11), with only minor contributions from the NOT2–NOT3 heterodimer (Fig. 3C, lane 9).

The central C2 region of Hs *NOT4-C is sufficient to contact CCR4–NOT and trigger decay of tethered mRNA*

It is not uncommon for the long unstructured portions of regulatory mRNA-associated proteins (such as in GW182/

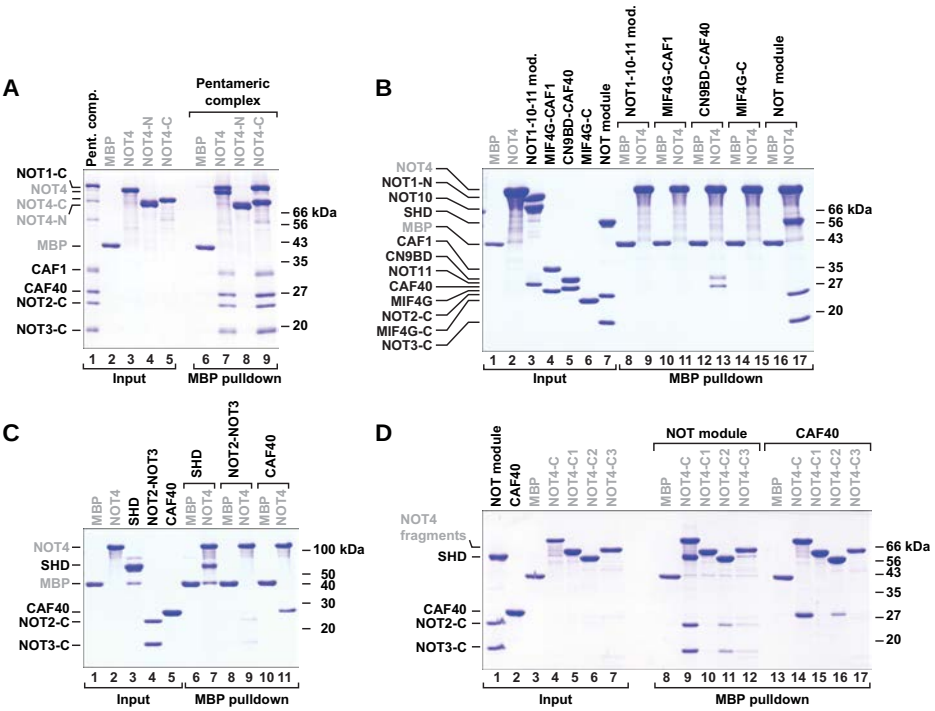


Figure 3. *Hs* NOT4 directly interacts with the NOT1-SHD and CAF40. (*A–D*) MBP pull-down assays with MBP-tagged *Hs* NOT4 and purified components of the human CCR4–NOT complex. MBP-tagged NOT4 or NOT4 fragments were used to pull down potential interaction partners. MBP alone served as a control. The respective starting materials ("input") and pull-down samples were analyzed by SDS-PAGE and Coomassie blue staining. Potential interaction partners included a pentameric assembly of recombinant human CCR4–NOT proteins (*A*), various subassemblies of the CCR4–NOT proteins (*B*), and individual proteins (*C*). (*D*) To further confine individual interactions, *Hs* NOT4-C was subdivided into the C1, C2, and C3 regions. See Figure 1B for details. MBP-tagged constructs are labeled in gray.

TNRC6, Nanos, or Roquin) to show multiple interactions with the CCR4–NOT complex, targeting several of its subunits (Jonas and Izaurralde 2015; Raisch et al. 2016; Sgromo et al. 2017). In order to map the interactions with the NOT module and CAF40 more precisely, we subdivided *Hs* NOT4-C into three regions (C1, C2, and C3) based on initial secondary structure and disorder prediction and tested them individually in MBP pull-down experiments with the NOT module or CAF40. These experiments identified the central C2 region (residues E377–Q428) (Fig. 1B) as a major interaction site for both the NOT module and CAF40 (Fig. 3D, lanes 11,16), although the interactions were weaker than with the entire NOT4-C region (Fig. 3D, lanes 9,14). The NOT module also interacted very weakly with the C3 region of *Hs* NOT4 (Fig. 3D, lane 12), whereas CAF40 showed no additional interactions in this context.

Considering the importance of the C2 region in the pull-down experiments, we also tested it in a tethering assay using the β-globin-6xMS2bs mRNA reporter. Strikingly, the C2 region was sufficient and highly efficient to trigger reporter mRNA decay. In contrast, the C1 or C3 region failed to elicit mRNA decay when tethered to the reporter (Fig. 2G,H). All tested fragments were expressed at comparable levels (Fig. 2I).

Alignments of NOT4 proteins reveal highly conserved sequence motifs in NOT4-C

To identify a potential sequence motif that could be responsible for the interaction of the C2 region with the NOT module and CAF40, we generated and analyzed separate alignments of NOT4 proteins from metazoans, plants, and yeasts. These alignments revealed locally conserved sequences at different positions in the C-terminal regions of NOT4 (Fig. 1B; Supplemental Figs. S4, S5A; Supplemental Alignment Files SF1–SF3).

Most striking is the conservation of a 23-amino-acid motif with α-helical propensity in the C2 region of *Hs* NOT4, which is present throughout all metazoans (Fig. 1B; Supplemental Fig. S4). A similar sequence also exists in plants but is not found in fungi. Conversely, the NOT1-binding motif of *Sc* Not4 (Bhaskar et al. 2015) is conserved in yeast but cannot be identified in plants and metazoans (Supplemental Fig. S5A; Supplemental Alignment Files SF1–SF3).

Beside the 23-amino-acid motif, the alignments also uncovered the presence of a proline-rich PPPGΦ motif (Φ = F, L, or I) at least once in each of the NOT4 sequences from metazoans, plants, and yeasts. The position of the PPPGΦ motif within NOT4-C varies with the phylogeny and can be found before or after the 23-amino-acid motif with the potential for misaligning distantly related sequences (Supplemental Figs. S4, S5A; Supplemental Alignment Files SF1–SF3). PPPGΦ motifs are known to be recognized by proteins containing proline-binding GYF domains (Kofler and Freund 2006).

The 23-amino-acid motif (residues E400–E422) occupies the second half of the C2 region (Supplemental Fig. S5B). We therefore again performed MBP pull-down experiments with the NOT module and CAF40, where either the first half (C2a) or the second half (C2b) of the C2 region or the entire C2 region was deleted from *Hs* NOT4-C. These deletions did not detectably affect the interaction with the NOT module, indicating that the C1 and C3 regions of *Hs* NOT4-C together are still sufficient to pull down the NOT module (Supplemental Fig. S5C, lanes 9–12). However, the interaction with CAF40 was clearly diminished by deleting the second half of the C2 region (Supplemental Fig. S5C, lanes 15,17). This observation suggests that the 23-amino-acid motif acts as a conserved CBM. This assumption was confirmed in the following by X-ray crystallography. The motif is hence called the NOT4 CBM.

Crystal structure of the CBM of Dm *NOT4 in complex with* Hs *CAF40*

To understand in molecular detail how NOT4 interacts with the CCR4–NOT complex, we used peptides corresponding to the putative CBM of *Hs* NOT4 or *Dm* NOT4 to set up cocrystallization trials with *Hs* CAF40. We obtained crystals—but only of a heterologous complex consisting of *Hs* CAF40 and the CBM of *Dm* NOT4. We obtained two distinct crystal forms, each with two crystallographically independent complexes per asymmetric unit and diffracting X-rays to a maximum resolution of 2.1 Å (Table 1). CBM binding is highly similar among the four CAF40–CBM complexes, and therefore only one of them (polypeptide chains A and B from space group $P2_12_12_1$) is described (Fig. 4A–C; Supplemental Figs. S5D, E, S6).

From *Dm* NOT4 L816 to E835 (see Fig. 4D for human sequence numbers), the *Dm* CBM adopts a common conformation in all of the available structures, with *Dm* NOT4 P820 to E835 folding into four turns of an amphipathic α helix that is bent between turns two and three toward the surface of *Hs* CAF40. Only the very N-terminal and very C-terminal residues of the crystallized peptide show differing orientations in the four available complexes, probably due to crystal-packing interactions and indicating structural flexibility (Supplemental Fig. S5E). *Hs* CAF40 adopts its rigid and previously described crescent-like shape, made from six armadillo repeats. Despite its highly negative overall charge (pI = 3.6), the *Dm* CBM interacts with *Hs* CAF40 primarily via hydrophobic contacts. It engages the concave surface of *Hs* CAF40, contacting residues from three parallel α helices (α5, α8, and α11) and burying a surface on *Hs* CAF40 of 842 $Å^2$ (Fig. 4A–C).

Importantly, the very same surface of CAF40 was described previously to be engaged also by the CBMs of Roquin and Bam (Fig. 4E,F; Sgromo et al. 2017, 2018). These CBMs also fold into amphipathic helices, covering surface areas of 841 $Å^2$ and 748 $Å^2$, respectively, and excluding any simultaneous associations of multiple CBMs with CAF40. Strikingly, however, whereas the CBMs of Roquin and Bam run in parallel to the α helices α5, α8, and α11, the CBM of NOT4 runs in an antiparallel fashion and hence is structurally and evolutionarily unrelated to the other two CBMs.

Table 1. *Data collection and refinement statistics for the* Hs *CAF40*–Dm *NOT4 CBM complex*

	Crystal form 1	Crystal form 2
Space group	P 2_1 2_1 2	I 2_1 2_1 2_1
Unit cell		
Dimensions *a*, *b*, *c*	83.9 Å, 109.6 Å, 69.7 Å	85.6 Å, 90.3 Å, 197.0 Å
Angles *α*, *β*, *γ*	90.0°, 90.0°, 90.0°	90.0°, 90.0°, 90.0°
Data Collection[a]		
Wavelength	1.0000 Å	1.0000 Å
Resolution range	50 Å–2.1 Å (2.14 Å–2.10 Å)	50 Å–2.2 Å (2.25 Å–2.20 Å)
R_{sym}	7.0% (91.4%)	6.5% (157.3%)
Completeness	99.6% (99.8%)	99.9% (99.8%)
Mean $I/\sigma(I)$	12.2 (1.5)	16.6 (1.6)
CC 1/2	99.0 (55.6)	100.0 (89.9)
Unique reflections	38,143 (2780)	39,129 (2831)
Multiplicity	4.0 (4.2)	8.9 (9.0)
Refinement		
R_{work}	19.1%	19.3%
R_{free}	22.2%	23.0%
Number of atoms		
All atoms	4971	4805
Protein	4760	4738
Ligands	34	3
Water	177	64
Average B factor		
All atoms	52.7 $Å^2$	81.0 $Å^2$
Protein	52.3 $Å^2$	81.1 $Å^2$
Ligands	95.5 $Å^2$	93.7 $Å^2$
Water	53.7 $Å^2$	66.9 $Å^2$
Ramachandran plot		
Favored regions	99.2%	98.3%
Disallowed regions	0.0%	0.0%
RMSD from ideal geometry		
Bond lengths	0.002 Å	0.010 Å
Bond angles	0.426°	1.040°

[a]Values in parentheses are for the highest-resolution shell.
(RMSD) Root mean square deviation.

Details of the interaction between the CBM of Dm *NOT4 and* Hs *CAF40*

The amphipathic α helix of the *Dm* NOT4 CBM is preceded by an "LGFDP" sequence motif that is invariant in our alignment of metazoan species (Fig. 4D). This sequence motif by itself forms a characteristic structure that helps to pin down the N-terminal half of the α helix (Fig. 4G, H). Probably due to the backbone flexibility provided by *Dm* NOT4 G817, the side chains of *Dm* NOT4 L816, F818, and P820 can join to form a small hydrophobic cluster, which is centrally contacted and completed by V181 from the α helix α11 of *Hs* CAF40. Furthermore, V181 is assisted by F184 to fix *Dm* NOT4 L816 and by L177 to fix *Dm* NOT4 F818, allowing *Dm* NOT4 L816 and F818 to intercalate between the side chains on the α helix α11 of *Hs* CAF40. *Dm* NOT4 P820 also initiates the α helix of NOT4 and is assisted by *Dm* NOT4 D819, which caps the helix and compensates for the positive charge of the helix dipole. Moreover, P820 is positioned straight over G141 in the α helix α8 of *Hs* CAF40 and would be spatially incompatible with any other residue at this position apart from a glycine.

Following P820, the hydrophobic surface of the NOT4 α helix probes the groove between α helices α8 and α5 of *Hs* CAF40 (Fig. 4D,H,I) using side chains of *Dm* NOT4 F821, T824, L828, and L831. These side chains are lined by residues from α helices α8 (*Hs* CAF40 T138, L137, Y134, and P131) and α5 (*Hs* CAF40 A95, N92, C91, N88, S87, H85, and A84), all of which are within van der Waals distance and frequently contribute to the interactions with the aliphatic portions of their side chains. At the C-terminal end of the CBM (Fig. 4D,H,I), the invariant *Dm* NOT4 E835 pins down the C-terminal half of the NOT4 α helix, using two hydrogen bonds to coordinate the free main chain nitrogens of *Hs* CAF40 A84 and H85 at the beginning of α helix α5 and compensating for the positive charge of the helix dipole. As a consequence of this interaction, the side chain of A84 gets completely surrounded by residues from NOT4 (*Dm* NOT4 L828, L831, M832, and E835), tolerating no side chains at this position that are larger than alanine. Additional specificity arises from an H bond between the side chains of *Dm* NOT4 T824 and *Hs* CAF40 T138 on α helix α8, which is deeply buried in the interface (Fig. 4D,H). Finally, it is important for *Hs* CAF40 Y134 on α helix α8 to rotate away from its preferred

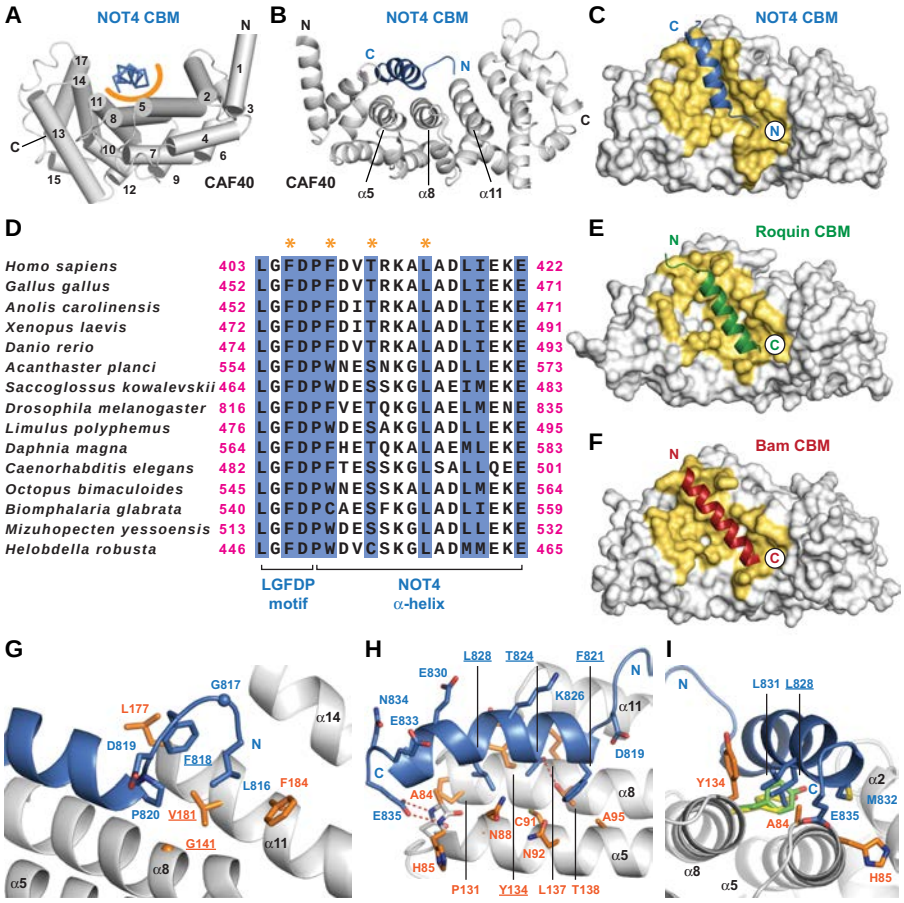


Figure 4. Crystal structure of the NOT4 CBM bound to CAF40. (*A*) Complex of the *Dm* NOT4 CBM peptide (blue, shown in ribbon representation) bound to *Hs* CAF40 (gray). The α helices of CAF40 are numbered and depicted as cylinders. The orange semicircle marks the predominantly hydrophobic interface between the CBM peptide and CAF40. The structurally variable flanks of the CBM peptide are excluded. (*B*) Rotated view of the CBM–CAF40 complex in cartoon representation marking the three central α helices of the concave CBM-binding surface. (*C*) Top view of CAF40 in surface representation, with the CBM of NOT4 as a cartoon. Interacting residues are colored in yellow. (*D*) Sequence alignment of metazoan NOT4 CBMs excluding the structurally variable flanks. The N-terminal portion of the CBM contains an extended LGFDP motif, and the C-terminal portion of the CBM consists of a bent α helix. Residues that directly contact CAF40 are shaded in blue, and residues that were mutated in this study are marked by orange asterisks. (*E*,*F*) Complexes of *Hs* CAF40 with the CBMs of *Dm* Roquin (*E*; Protein Data Bank [PDB] ID 5lsw) (Sgromo et al. 2017) and *Dm* Bam (*F*; PDB ID 5onb) (Sgromo et al. 2018), shown in the same style and orientation as in *C* and excluding structurally variable flanks of the CBMs. Note the inverted orientation of the CBMs. (*G*–*I*) Close-up views of the interface between *Dm* NOT4 and *Hs* CAF40. Selected side chains of NOT4 and CAF40 are shown as blue and orange sticks, respectively, with nitrogens in dark blue and oxygens in red. Hydrogen bonds are indicated by red dashed lines. Residues mutated in this study are underlined. (*I*) Rotamers of CAF40 Y134 as found in the complexes of CAF40 with the CBMs of *Dm* Roquin (yellow) and *Dm* Bam (lime).

rotamer position that is observed in free CAF40 (Garces et al. 2007; Chen et al. 2014; Mathys et al. 2014) and in the complexes with Roquin (Sgromo et al. 2017) and Bam (Sgromo et al. 2018). Hence, the orientation of *Hs* CAF40 Y134 could help to discriminate between the three binding partners, liberating access for T824 and L828 from the *Dm* NOT4 α helix to the groove between α helices α5 and α8 (Fig. 4D,H,I). As a result, the CBM and CAF40 interact via highly complementary shapes with a hydrophobic interface that excludes any water molecule and by exposing polar residues (*Dm* NOT4 K826, E830, E833, and N834) to the solvent on the hydrophilic side of the *Dm* NOT4 α helix (Fig. 4D,H).

Validation of the binding interface

To validate the specificity of the interface observed in the crystal structure, we generated mutations in the CBM of *Dm* NOT4 and in *Hs* CAF40 and tested them in MBP pull-down assays. First, we demonstrated that the CBM of *Dm* NOT4 indeed also interacts with *Hs* CAF40 in solution (Fig. 5A, lane 14), confirming that it is a bona fide CBM. We disrupted the interaction from the side of CAF40 using either a single point mutation targeted at the "LGFDP" motif (*Hs* CAF40 V181E) (Fig. 4G) or a double point mutation targeted at the α-helical part of the CBM (*Hs* CAF40 2x mut; Y134D and G141W) (Fig. 4G,I).

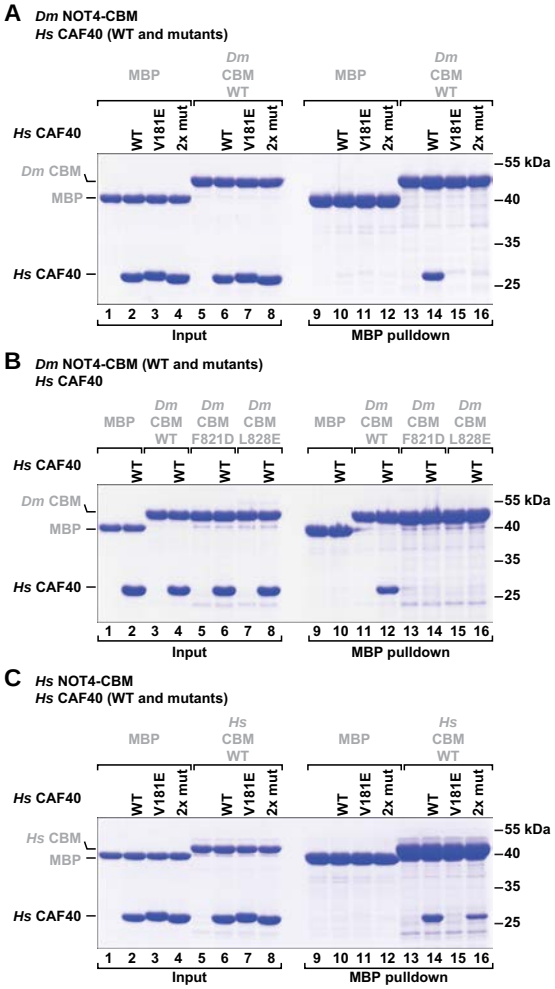


Figure 5. The interface between the NOT4 CBM and CAF40 is conserved in metazoans. MBP pull-down assays with MBP-tagged NOT4 CBMs and purified recombinant *Hs* CAF40 were done as described in Figure 3, A–D. (*A*) Mutants of *Hs* CAF40 in the presence of the *Dm* NOT4 CBM. Mutations target the interaction with the LGFDP motif (*Hs* CAF40 V181E) or the interaction with the CBM helix (*Hs* CAF40 2x mut; Y134D or G141W). (*B*) Mutants of the *Dm* NOT4 CBM in the presence of *Hs* CAF40. Mutations target the interaction of the CBM helix (*Dm* NOT4 F821D or L828E). (*C*) Mutants of *Hs* CAF40 in the presence of the *Hs* NOT4 CBM.

Both mutations had been used previously to disrupt the interactions of CAF40 with the CBMs of Roquin and Bam (Sgromo et al. 2017, 2018); in the present structural context, they abolished the interaction with the *Dm* NOT4 CBM (Fig. 5A, lanes 15,16). Conversely, single substitutions in the *Dm* NOT4 CBM (F821D or L828E) (Fig. 4D, H) were sufficient to abrogate the interaction with *Hs* CAF40 (Fig. 5B, lanes 14,16).

Furthermore, the His_6-NusA-tagged *Dm* NOT4 CBM and the His_6-NusA-tagged *Dm* Bam CBM competed with the MBP-tagged *Dm* NOT4 CBM for binding to *Hs* CAF40, confirming that they target overlapping binding surfaces on *Hs* CAF40 (Supplemental Fig. S5F). As in the case of the *Dm* Roquin CBM (Sgromo et al. 2018), we could not determine a precise dissociation constant for the *Dm* NOT4 CBM because it aggregated at concentrations needed to perform isothermal titration calorimetry and microscale thermophoresis experiments.

Finally, we also tested the interaction of the *Hs* NOT4 CBM with *Hs* CAF40. Similar to the results obtained with the *Dm* NOT4 CBM, the single *Hs* CAF40 V181E mutation (Fig. 4G) was sufficient to abolish the interaction with the *Hs* NOT4 CBM (Fig. 5C, lane 15 vs. 14). The double point mutation (*Hs* CAF40 2x mut) (Fig. 4G,I) also reduced the interaction but was not sufficient to abolish it, indicating species-specific adaptations in the molecular details of the coevolved interface of the human proteins (Fig. 5C, lane 16 vs. 14). Together, the results confirm that the interactions observed in the crystal structure also occur in solution.

The CBM is essential for the interaction of NOT4 with the human CCR4–NOT complex

To investigate the significance of the CBM–CAF40 interaction in cells and in the context of the entire CCR4–NOT complex, we repeated SBP pull-down assays with *Hs* NOT4-C where either the CBM (ΔC2b), the C2a region (ΔC2a), or the entire C2 region (ΔC2) were deleted from the construct (see Supplemental Fig. S5B for boundaries). We found that the presence of the CBM is required to pull down the CCR4–NOT complex (Fig. 6A, lane 7), as detected by the absence of NOT1, NOT2, NOT3, and CAF40 in the pull-down fraction when the C2b region was deleted (Fig. 6A, lane 10). Consistently, the deletion of the entire C2 region as well disrupted the interaction (Fig. 6A, lane 8). Surprisingly however, the presence of the C2a region was also required for an efficient pull-down of the CCR4–NOT complex (Fig. 6A, lane 9). The C2a region could simply act as a spacer to allow for a proper interaction of the CBM with CAF40. More likely, however, its requirement reflects and underlines the importance of auxiliary and possibly species-specific interactions of NOT4-C with other parts of the CCR4–NOT complex, such as the NOT module. In line with these results, the deletion of the entire C2 region strongly diminished the NOT4-mediated degradation of the β-globin-6xMS2bs reporter mRNA in HEK293T cells (Fig. 6B,C). The separate deletion of the C2a region (ΔC2a) or the CBM (ΔC2b) did so as well but to a lesser degree. All tethered proteins were expressed at a similar level (Fig. 6D).

To probe more directly for the interface observed in the crystal structure, we engineered a quadruple point mutant of the *Hs* NOT4 CBM (4x mut; F405D, F408A, T411E, and L415E) (Fig. 4D) affecting both the LGFDP motif and the hydrophobic surface of the α helix. This mutation strongly reduced the interaction with the CCR4–NOT complex in the SBP pull-down assay (NOT4-C 4x mut) (Fig. 6E), whereas individual point mutations were not as effective (Supplemental Fig. S7A). Also in the tethering assay, the quadruple point mutation (NOT4 4x mut) impaired the ability of NOT4 to degrade the β-globin-6xMS2bs reporter

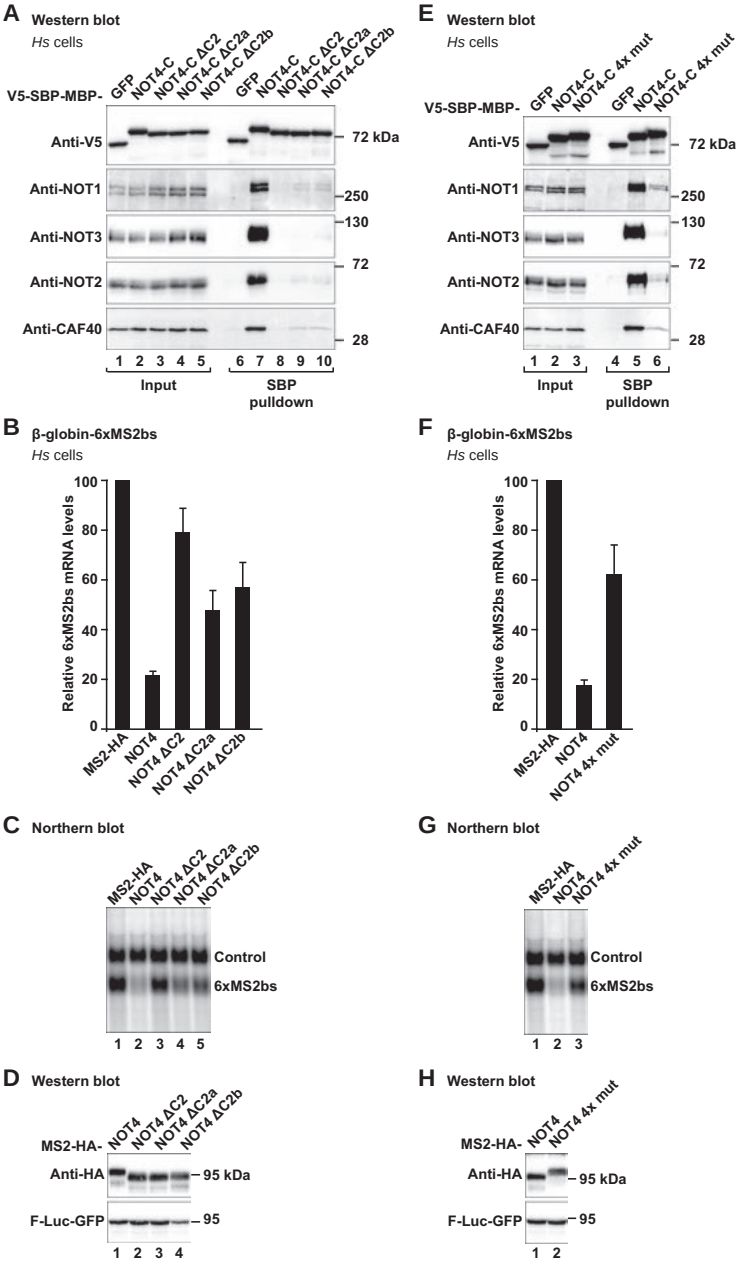


Figure 6. The NOT4 CBM plays a crucial role in the interaction with the CCR4–NOT complex in HEK293T cells. (*A*) SBP pull-down of endogenous human CCR4–NOT proteins with V5-SBP-MBP-tagged deletion variants of *Hs* NOT4-C. NOT4-C ΔC2 lacks residues E377–S424. NOT4-C ΔC2a lacks residues E377–D402. NOT4-C ΔC2b lacks residues E400–Q428, including the CBM. For additional details, see Figure 1C and Supplemental Figure S5B. (*B–D*) Tethering assay with the deletion variants of *Hs* NOT4 and the β-globin mRNA reporter. (*B*) Relative mRNA levels, with error bars corresponding to standard deviations. $n = 3$. For additional details, see Figure 2A. (*C*) Northern blot. (*D*) Western blot. (*E*) SBP pull-down of endogenous human CCR4–NOT proteins with a V5-SBP-MBP-tagged CBM mutation variant of *Hs* NOT4-C (4x mut; F405D, F408A, T411E, and L415E). For additional details, see Figure 1C. (*F–H*) Tethering assay with the CBM mutation variant of *Hs* NOT4 (4x mut) and the β-globin mRNA reporter. (*F*) Relative mRNA levels, with error bars corresponding to standard deviations. $n = 3$. For additional details, see Figure 2A. (*G*) Northern blot. (*H*) Western blot.

mRNA to a degree that is comparable with the deletion of the entire CBM (Fig. 6, F,G vs. B,C). All of the tethered proteins were equally expressed (Fig. 6H). These findings demonstrate the importance of the CBM for the function of NOT4 and indicate that this is a major site of interaction with the CCR4–NOT complex.

CAF40 plays a crucial and conserved role for the recruitment of NOT4 to the CCR4–NOT complex in metazoans

The present crystal structures show how the CBM interacts with CAF40, but our previous experiments did not formally exclude that the CBM makes similarly important contacts with other subunits of the CCR4–NOT complex. Therefore, we took advantage of HEK293T cells in which CAF40 had been knocked out by CRISPR–Cas9 genome editing (CAF40 knockout cells) (Fig. 7A–D; Supplemental Fig. S7B–D; Sgromo et al. 2018). In this cell line, the levels of endogenous NOT1, NOT2, and NOT3 proteins were not altered (Fig. 7B), but the NOT4-mediated decay of the tethered R-Luc-6xMS2bs mRNA reporter (Fig. 7A,D) or of the tethered β-globin-6xMS2bs mRNA reporter (Supplemental Fig. S7B,C) was impaired. This observation underlines the importance of CAF40 for the efficient recruitment of the CCR4–NOT complex.

To test whether the remaining mRNA repression in CAF40 knockout cells (Fig. 7A; Supplemental Fig. S7B,C)

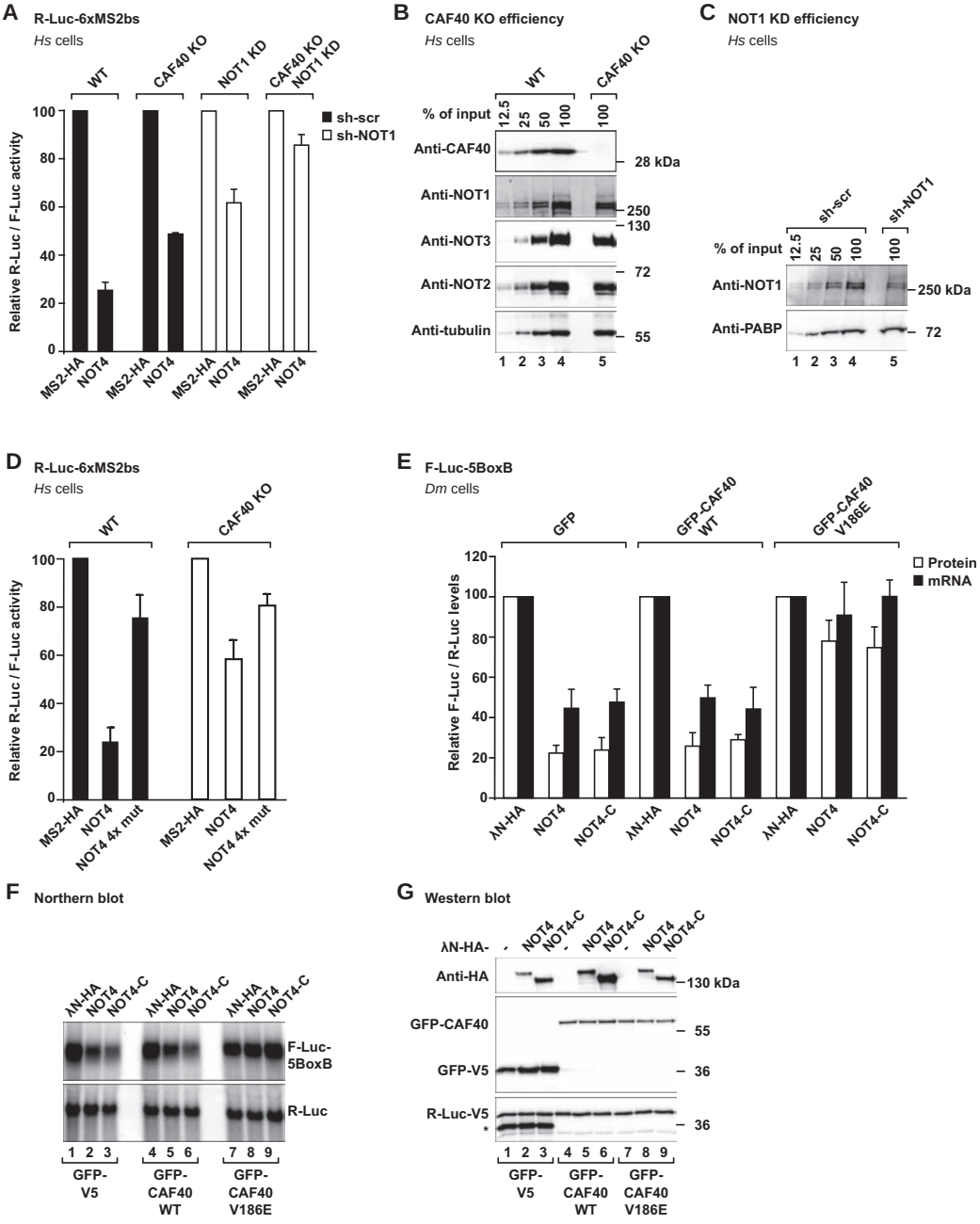


Figure 7. CAF40 is a central mediator for the interaction of NOT4 with the CCR4–NOT complex in HEK293T and *Dm* S2 cells. (*A*) Tethering assay with *Hs* NOT4 and a luciferase reporter in HEK293T cells lacking CAF40 (CAF40 knockout [KO]) (Sgromo et al. 2018). Renilla luciferase mRNA served as a reporter and contained six binding sites for the MS2 protein (R-Luc-6xMS2bs). Firefly luciferase mRNA served as a reference and transfection control (F-Luc). R-Luc activity was normalized to the reference and plotted with respect to the values obtained from the expression of MS2-HA alone (set to 100). The shRNA-mediated depletion of the CCR4–NOT complex is indicated by white bars (sh-NOT1 RNA and NOT1 knockdown [KD]) as compared with black bars (sh-scr RNA, control). Error bars correspond to standard deviations. $n = 3$. (*B*) Efficiency of CRISPR–Cas9-mediated gene editing of CAF40. The Western blot shows CAF40 knockout cells in comparison with a dilution series of wild-type HEK293T cells using tubulin as a loading control. (*C*) Efficiency of shRNA-mediated depletion of NOT1. The Western blot shows HEK293T cells expressing sh-NOT1 RNA in comparison with a dilution series of HEK293T cells expressing sh-scr RNA using PABP as a loading control. (*D*) Tethering assay with the CBM mutation variant of *Hs* NOT4 (4x mut) and the luciferase mRNA reporter in CAF40 knockout cells. Error bars correspond to standard deviations. $n = 3$. (*E–G*) Tethering assay with *Dm* NOT4 in *Dm* S2 cells overexpressing *Dm* CAF40 with a mutated CBM-binding surface. Experiments were done with λN-HA-tagged *Dm* NOT4 or *Dm* NOT4-C and analyzed as in Figure 2D, but cells were additionally overexpressing either wild-type *Dm* CAF40 or mutant *Dm* CAF40 (GFP-CAF40 V186E, corresponding to *Hs* CAF40 V181E). GFP-V5 was overexpressed as a negative control. (*E*) Relative protein and mRNA levels, with error bars corresponding to standard deviations. $n = 3$. (*F*) Northern blot. (*G*) Western blot. The asterisk denotes the additional detection of GFP-V5 on the anti-V5 blot.

was due to the CCR4–NOT complex, we additionally disrupted and depleted the remainder of the complex by a shRNA-mediated knockdown of NOT1 (Fig. 7C; Supplemental Fig. S7D; Boland et al. 2013). Under these conditions, tethering of NOT4 left reporter mRNA expression almost unaffected (CAF40 knockout + NOT1 knockdown) (Fig. 7A; Supplemental Fig. S7B,C). Consequently, the remaining subunits of the complex in CAF40 knockout cells still interact with tethered NOT4, which seems to act exclusively via the CCR4–NOT complex.

Finally, there was only a small difference between tethering the quadruple point mutation of NOT4 (NOT4 4x mut) and tethering wild-type NOT4 in CAF40 knockout cells, whereas this difference was considerable in wild-type cells (Fig. 7D). This result confirms CAF40 as the primary interaction partner of the NOT4 CBM in human cells.

The conservation of the CBM suggests that the interaction of NOT4 with CAF40 is preserved in metazoans, albeit modulated by additional contacts, such as with the NOT1-SHD (Fig. 3C). Likely due to such taxon-specific or species-specific modulation, it was therefore possible in *Dm* S2 cells to obtain a dominant-negative effect on CCR4–NOT recruitment by overexpressing a V186E mutant of GFP-tagged *Dm* CAF40 (Fig. 7E–G). This mutant (corresponding to *Hs* CAF40 V181E) was also shown previously to impair CCR4–NOT recruitment by the CBM of Bam (Sgromo et al. 2018). Apparently, in this case, the overexpression of GFP-tagged CAF40 can functionally replace the endogenous protein, and a single point mutation is then sufficient to disrupt the interaction with NOT4. Again, these observations demonstrate the central and conserved role of the CBM for the recruitment of NOT4 to the CCR4–NOT complex.

Discussion

The present work demonstrates that metazoan NOT4 contains a conserved CBM in its variable C-terminal tail and elucidates the molecular details of the CBM–CAF40 complex. The CBM is required in cells for an efficient recruitment of NOT4 to the CCR4–NOT complex or a recruitment of the complex to NOT4-mediated cellular processes. The interaction of the CBM is assisted by auxiliary flanking sequences in NOT4-C that vary between metazoan species. These sequences also contact other subunits of the CCR4–NOT complex, such as the SHD of NOT1. From an evolutionary point of view, the CBM therefore appears to represent the ancestral mode of coupling the NOT4-dependent E3 ubiquitin ligase activity with the CCR4- and CAF1-dependent deadenylation activity of the CCR4–NOT complex. In yeast, however, the CBM seems to have become dispensable, possibly because the contacts with NOT1 are sufficient to maintain the interaction (Bhaskar et al. 2015). The conservation of the CBM marks NOT4 as a ubiquitous but apparently facultative cofactor of the metazoan CCR4–NOT complex that likely has important functions in a subset of CCR4–NOT-dependent cellular processes.

Facultative and regulated interaction of the CCR4–NOT complex with metazoan NOT4

In *S. cerevisiae* and *S. pombe*, NOT4 copurifies with the CCR4–NOT complex (Chen et al. 2001; Stowell et al. 2016), suggesting that it is an integral component of the complex. In metazoan species, however, NOT4 is apparently not generally available to interact with the CCR4–NOT complex in a constitutive manner. This is indicated by the fact that endogenous *Hs* NOT4 or *Dm* NOT4 do not copurify with the core of the complex (Lau et al. 2009; Temme et al. 2010) and by our observation that full-length *Hs* NOT4 does not pull down the CCR4–NOT complex from HEK293T cell extracts, in contrast to *Hs* NOT4-C. However, full-length *Hs* NOT4 that was expressed in bacteria does interact with a reconstituted subassembly of the human CCR4–NOT complex.

We therefore speculate that metazoan NOT4-N somehow prevents NOT4-C from interacting with the CCR4–NOT complex, with possible assistance from posttranslational modifications or additional binding partners. For example, it is conceivable that a structural reorganization of NOT4 is required in eukaryotic cells to release the CBM for an interaction with CCR4–NOT and/or that the negatively charged CBM (pI = 4.0 in *Hs* NOT4) gets sequestered by the highly positively charged coiled-coil linker and RRM domains of NOT4-N (pI = 10.2 in *Hs* NOT4) when NOT4 is not bound to an mRNA. In this way, it would be possible to regulate the availability of the NOT4 ubiquitin ligase activity to only a subset of the CCR4–NOT-mediated cellular processes, but whether such regulation indeed exists and how it might be achieved in detail remains to be investigated.

Mutual corecruitment of NOT4 and the CCR4–NOT complex

The widespread conservation of the CBM in NOT4 proteins reveals NOT4 as an ancient cofactor of the CCR4–NOT complex. Furthermore, we show that NOT4 is able to cause CCR4–NOT-mediated mRNA decay if tethered to an mRNA target. However, it remains unclear from these experiments whether metazoan NOT4 is needed primarily as an upstream recruitment factor that directs the CCR4–NOT complex to selective mRNA targets or rather as a downstream effector that recruits additional proteins to the CCR4–NOT complex and/or ubiquitinates nearby protein targets; e.g., to mark them for proteasomal degradation. In contrast to selective mRNA-binding proteins such as TTP, Nanos, or Roquin (Newman et al. 2016), there are currently no known RNA targets for the coiled-coil linker, RRM, or ZNF domains of NOT4; i.e., for its putative RNA-binding domains. This argues against an RNA-specific recruitment function.

In the context of cotranslational mRNA quality control, however, NOT4 could act as both an upstream recruitment factor of the CCR4–NOT complex to mRNAs with stalled ribosomes and a downstream effector for the ubiquitination and degradation of protein targets (Panasenko

2014). Additional protein-binding partners may modulate or stabilize the interactions in this case.

Finally, it is worthwhile to follow up also on the PPPGΦ motifs that we found to be highly conserved in the NOT4 proteins from metazoans, plants, and yeasts. PPPGΦ motifs tend to interact with GYF domain proteins (Kofler and Freund 2006), such as the GIGYF1/2 translational repressors (Kryszke et al. 2016; Peter et al. 2017; Amaya Ramirez et al. 2018) that were described to bind CAF40 in human cancer cells (Ajiro et al. 2009). Quite likely, therefore, the CCR4–NOT complex and NOT4 frequently support each other in a mutual corecruitment that is difficult to disentangle experimentally.

Competition of mRNA-associated proteins for the CBM-binding site of CAF40

In most of the known cases where CCR4–NOT gets recruited to an mRNA target, the CBM-binding surface on CAF40 appears to remain available for a simultaneous recruitment of NOT4. This is, for example, the case for the TNRC6/GW182-mediated microRNA-dependent mRNA regulation, where tryptophans of TNRC6/GW182 bind to the convex side of CAF40 (Chen et al. 2014; Mathys et al. 2014). Similarly, mRNA-specific CCR4–NOT recruitment proteins such as TTP or Nanos apparently do not structurally interfere with NOT4 binding to the concave surface of CAF40 (Fabian et al. 2013; Bhandari et al. 2014; Bulbrook et al. 2018), allowing for a combinatorial mRNA regulation.

In contrast, the CBMs of Roquin and Bam were shown to target the exact same binding surface on CAF40 as the CBM of NOT4 (Sgromo et al. 2017, 2018), making their binding mutually exclusive. It is therefore possible that Roquin proteins have evolved to displace NOT4 in a context-specific manner, since they bring along their own E3 ubiquitin ligase domain. Conversely, in the case of Bam, the CBM might serve to prevent NOT4-mediated and ubiquitination-dependent processes downstream from CCR4–NOT recruitment in the germline of *D. melanogaster*. Future work will show whether such competition indeed occurs in vivo and whether there are additional CBM-containing mRNA-binding proteins in fungi, plants, or metazoans that operate in a similar manner.

Clearly, however, the present identification of a conserved CBM in the NOT4 protein underlines the role of CAF40 as a hub for peptide-mediated interactions and adds to the ever more complex regulation of mRNA expression in eukaryotic cells.

Materials and methods

DNA plasmid constructs

For bacterial expression of recombinant *Hs* NOT4 constructs in *E. coli*, cDNA sequences were inserted between the XhoI and NheI restriction sites of the pnEA-pM plasmid, resulting in fusion proteins carrying N-terminal MBP tags cleavable by the human rhinovirus 3C (HRV3C) protease and, in addition, C-terminal GB1 and hexahistidine tags. For bacterial expression of recombinant *Dm* NOT4 constructs, cDNA sequences were inserted between the XhoI and BamHI restriction sites of the pnYC-vM plasmid, resulting in tobacco etch virus (TEV) protease-cleavable MBP fusion proteins. For the expression of *Hs* NOT4 constructs in human (HEK293T) cells, cDNA sequences were inserted into the pCIneo-V5-SBP-MBP plasmid or the pcDNA3.1-MS2-HA plasmid using the XhoI and NotI restriction sites. For the expression of *Dm* NOT4 constructs in *Dm* S2 cells, cDNA sequences were inserted into the pAc5.1B-λN-HA plasmid using the NotI and BstBI restriction sites. All of the plasmid constructs generated in this study, including backbone references, are listed in Supplemental Table S1.

MBP pull-down assays with bacterially expressed proteins

For initial pull-down assays with full-length *Hs* NOT4 and its fragments, the proteins were expressed in *E. coli* BL21 (DE3) Star cells (Invitrogen) overnight in LB medium at 20°C. Cells were homogenized in lysis buffer (50 mM Na/HEPES at pH 7.5, 300 mM NaCl, 5 µg/mL DNaseI, 1 mg/mL lysozyme, Roche "Complete" EDTA-free protease inhibitors) supplemented with 20 mM imidazole and 2 mM β-mercaptoethanol. The proteins were immobilized and isolated from the lysate on Ni-NTA resin (Qiagen) followed by elution in lysis buffer supplemented with 500 mM imidazole. They were then immobilized on 50 µL of amylose resin and incubated with an excess of purified CCR4–NOT proteins for 1 h in 500 µL of binding buffer (50 mM Na/HEPES at pH 7.5, 300 mM NaCl, 2 mM β-mercaptoethanol). Finally, the amylose beads were washed five times with binding buffer, and the proteins were eluted in 50 µL of binding buffer supplemented with 25 mM D(+)-maltose.

For pull-down assays with *Hs* and *Dm* NOT4 CBM constructs, proteins were purified from cells homogenized in lysis buffer supplemented with 2 mM DTT. Proteins were immobilized and isolated from the lysate on an amylose resin (New England Biolabs) followed by anion exchange chromatography over a HiTrap Q column (GE Healthcare) and size exclusion chromatography over a Superdex 200 26/600 column (GE Healthcare) in a buffer containing 10 mM Na/HEPES (pH 7.5), 200 mM NaCl, and 2 mM DTT. Forty micrograms of purified MBP-tagged NOT4 fragments or 20 µg of MBP were then incubated with approximately equimolar amounts of the respective purified CCR4–NOT proteins and 50 µL of amylose resin in 500 µL of binding buffer (50 mM Na/HEPES at pH 7.5, 200 mM NaCl, 2 mM DTT). After the incubation and washing steps, the proteins were eluted in 200 µL of binding buffer supplemented with 25 mM D(+)-maltose and precipitated with trichloroacetic acid.

The purifications of other human CCR4–NOT proteins, including *Hs* CAF40 for crystallization, were described previously (Petit et al. 2012; Boland et al. 2013; Bhandari et al. 2014; Chen et al. 2014; Raisch et al. 2016; Sgromo et al. 2017, 2018). The protein samples were resolved and analyzed by SDS-PAGE.

Crystallization

Hs CAF40 (GPHMLE-R19–E285) (Chen et al. 2014) was mixed with a twofold molar excess of the *Dm* NOT4 CBM peptide (D813–Q838, chemically synthesized and purchased from EMC Microcollections) in 10 mM Na/HEPES (pH 7.5), 200 mM NaCl, and 2 mM DTT. Initial screens were carried out in sitting drops at 22°C by mixing 200 nL of sample solution (6 mg/mL CAF40, 1.2 mg/mL NOT4) with 200 nL of reservoir solution. Crystals appeared within 1 d in many conditions. Crystals of crystal form 1 appeared in the initial screen over a reservoir

containing 0.2 M ammonium sulfate, 0.1 M Bis-Tris/Cl (pH 5.5), and 25% (w/v) PEG 3350. The crystals were cryoprotected in reservoir solution supplemented with 25% glycerol and flash-cooled in liquid nitrogen. Optimized crystals of crystal form 2 grew at 18° C in hanging drops mixing 1 µL of sample solution and 1 µL of reservoir solution containing 0.9 M K_2HPO_4 and 0.3 M NaH_2PO_4. Crystals were cryoprotected in 4.0 M sodium formate and flash-cooled in liquid nitrogen.

Data collection and structure determination

X-ray diffraction data were collected at a wavelength of 1.0000 Å on a Pilatus 6M detector (Dectris) at the PXII beamline of the Swiss Light Source (Villigen) and processed using XDS and XSCALE (Kabsch 2010). Crystal form 1 (space group $P2_12_12$) diffracted X-rays to a resolution of 2.1 Å, whereas crystal form 2 (space group $I2_12_12_1$) diffracted X-rays to a comparable resolution of 2.2 Å but with an increased B_{Wilson} (56.7 $Å^2$ vs. 39.2 $Å^2$). For each crystal form, we identified two copies of *Hs* CAF40 per asymmetric unit by molecular replacement using PHASER (McCoy et al. 2007) from the CCP4 package (Winn et al. 2011) and using chain A of Protein Data Bank (PDB) ID 2fv2 (Garces et al. 2007) as a search model. Initial models were improved and completed by iterative cycles of model building in COOT (Emsley et al. 2010) and refinement using PHENIX (Afonine et al. 2012). The NOT4 CBM peptides were then built manually into the remaining electron density and improved by additional building and refinement cycles. For crystal form 1, final refinement rounds were done using PHENIX with one TLS group per polypeptide chain and including small molecule ligands (one molecule each of Tris and glycerol plus four sulfate ions) in addition to 177 water molecules. This resulted in an R_{work} of 19.0% and an R_{free} of 21.8%. For crystal form 2, final refinement rounds were done using BUSTER (https://www.globalphasing.com/buster), also with one TLS group per polypeptide chain but in addition to small molecule ligands (one sodium and two chloride ions) and 64 water molecules, also autorefining NCS restraints. This resulted in an R_{work} of 19.3% and an R_{free} of 23.0% (Table 1). Illustrations were prepared in PyMOL (http://www.pymol.org).

SBP pull-down assays from HEK293T cells

HEK293T cells were seeded in 10-cm dishes (4×10^6 cells per plate and experiment) and transfected with pCIneo-V5-SBP-MBP plasmids after 1 d using Turbofect (Thermo Scientific) according to the manufacturer's protocol. Two days after transfection, the cells were lysed on ice in 1 mL of NET lysis buffer (50 mM Tris/Cl at pH 7.5, 150 mM NaCl, 0.1% Triton X-100, 1 mM EDTA, 10% glycerol) supplemented with protease inhibitors (Roche). After 15 min, lysates were centrifuged at 20,000*g* for 15 min at 4°C. The cleared lysate was then treated with 200 µg/mL RNase A (Qiagen) for 30 min at 4°C and centrifuged again as before, resulting in the input fraction for the experiment (1 mL = 100%). The input fraction was then incubated for 1 h at 4°C with 50 µL of streptavidin sepharose resin (GE Healthcare). The beads were washed three times with NET buffer and finally resuspended in protein sample buffer, resulting in the pull-down fraction (100 µL = 100%). The samples were analyzed by Western blot (for antibodies, see Supplemental Table S2) using the ECL Western blotting detection system (GE Healthcare).

Tethering assays in HEK293T cells

For MS2-dependent tethering assays with the β-globin mRNA reporter (Lykke-Andersen et al. 2000), HEK293T cells were seeded in six-well plates (0.7×10^6 cells per well) and transfected on the following day using Lipofectamine 2000 (Invitrogen). The transfection mixtures contained 0.5 µg of the β-globin reporter plasmid encoding six MS2-binding sites (β-globin-6xMS2bs); 0.5 µg of the β-globin reference and transfection control plasmid lacking MS2-binding sites and containing a partial sequence of GAPDH (control; β-globin-GAPDH) (Lykke-Andersen et al. 2000); and variable amounts (0.05–0.75 µg) of pcDNA3.1-MS2-HA plasmids (Supplemental Table S1) to achieve equivalent expression of MS2-HA fusion proteins. The cells were harvested 2 d after transfection. The total RNA was isolated using the peqGOLD TriFast reagent (Peqlab) and analyzed by Northern blot as described previously (Behm-Ansmant et al. 2006). Equivalent expression of MS2-HA-tagged proteins was tested in parallel by Western blot, expressing F-Luc-GFP (Kuzuoğlu-Öztürk et al. 2016) as a transfection control.

For the experiment shown in Supplemental Figure S3, A–C, cells were additionally cotransfected with 0.5 µg of a plasmid expressing either wild-type *Hs* DCP2 (GFP-DCP2 wild-type) or the *Hs* DCP2 mutant (GFP-DCP2 mut; E148Q) (Loh et al. 2013). Equivalent expression of the GFP-tagged proteins was tested in parallel by Western blot, expressing V5-SBP-MBP as a transfection control.

For MS2-dependent tethering assays with the luciferase reporter system (Kuzuoğlu-Öztürk et al. 2016), the transfection mixtures contained 0.2 µg of reporter plasmid containing or lacking six MS2-binding sites (R-Luc-6xMS2bs or R-Luc), 0.2 µg of reference and transfection control plasmid lacking six MS2-binding sites (F-Luc-GFP), and variable amounts of pcDNA3.1-MS2-HA plasmids (0.1–1.5 µg) (Supplemental Table S1). The cells were harvested 2 d after transfection, mRNA levels were determined by Northern blot, and R-Luc and F-Luc activities were measured using a "dual-luciferase reporter assay" system (Promega).

Tethering assays in HEK293T cells with knockdown of NOT1

The shRNA-mediated depletion of NOT1 has been described previously (Boland et al. 2013) using shRNA (*Hs* NOT1 target: ATT CAACATTCCCTTATA) and control shRNA (scr, control target: ATTCTCCGAACGTGTCACG). For tethering assays in cells depleted of NOT1, wild-type HEK293T cells or HEK293T CAF40 knockout cells (Sgromo et al. 2018) were transfected twice. For the first transfection, cells were seeded in six-well plates (0.7×10^6 cells per well) and transfected on the following day with mixtures containing 4 µg of plasmid expressing the respective shRNA. After 1 d, cells were selected for 24 h in DMEM supplemented with 1.5 µg/mL puromycin and subsequently seeded in six-well plates in medium without puromycin (0.7×10^6 cells per well). The following day, cells were transfected again with mixtures containing 2 µg of the respective shRNA plasmids but also containing the reporter and reference/transfection control plasmids (0.2 µg of R-Luc-6xMS2bs and 0.2 µg of F-Luc-GFP) and 0.125–0.25 µg of pcDNA3.1-MS2-HA plasmids (Supplemental Table S1). After 1 d, cells were selected for 48 h in DMEM (supplemented with 1.5 µg/mL puromycin) and analyzed as before.

Tethering assays in Dm *S2 cells*

For the λN-dependent tethering assay (Behm-Ansmant et al. 2006) with the luciferase reporter system in *Dm* S2 cells, cells were seeded in six-well plates (2.5×10^6 cells per well) and transfected just thereafter using Effectene (Qiagen). The transfection mixtures contained 0.1 µg of the F-Luc-5BoxB reporter plasmid, 0.4 µg of an R-Luc reference and transfection control plasmid encoding a deadenylation-resistant mRNA lacking BoxB sequences

(R-Luc; R-Luc-A_{90}-HhR) (Raisch et al. 2016), and variable amounts (0.01–0.08 µg) of pAC5.1B-λN-HA plasmids (Supplemental Table S1) to achieve equivalent expression of λN-HA-fusion proteins. The cells were harvested 3 d after transfection and analyzed as described.

For the experiments in Supplemental Figure S3, D–F, *Dm* S2 cells were additionally cotransfected with 1 µg of a plasmid expressing either wild-type *Dm* DCP1 (GFP-DCP1 wild-type) or the *Dm* DCP1 mutant (GFP-DCP1 mut; R70G, L71S, N72S, and T73G) (Kuzuoğlu-Öztürk et al. 2016), and, for the experiments in Figure 7, E–G, cells were cotransfected with 1.5 µg of a plasmid expressing either wild-type *Dm* CAF40 (GFP-CAF40 wild-type) or the *Dm* CAF40 mutant (GFP-CAF40 V186E) (Sgromo et al. 2018).

Accession numbers

Atomic coordinates and structure factors for the reported crystal structures have been deposited with the PDB under accession number 6hom for space group $P2_12_12$ and 6hon for space group $I2_12_12_1$.

Acknowledgments

We dedicate this work to the memory of Elisa Izaurralde, who passed away while this manuscript was still in preparation. We deeply miss her as a cherished mentor and colleague. We are grateful to Heike Budde, Maria Fauser, Sigrun Helms, and Catrin Weiler for excellent technical support. We also thank the staff at the PX beamlines of the Swiss Light Source, Villigen, CH, for assistance with X-ray data collection. This work was supported by the Max Planck Society. Funding for the open access charge was provided by the Max Planck Society.

Author contributions: E.I. conceived the project. C.K., T.R., and A.S. performed experiments and analyzed data, together with C.I., D.B., O.W., and E.I. C.K., T.R., A.S., C.I., and D.B. wrote the initial manuscript, which was annotated by E.I. and finalized with help from O.W. All authors commented on and approved the manuscript.

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. doi:10.1107/S0907444912001308

Ajiro M, Katagiri T, Ueda K, Nakagawa H, Fukukawa C, Lin ML, Park JH, Nishidate T, Daigo Y, Nakamura Y. 2009. Involvement of RQCD1 overexpression, a novel cancer-testis antigen, in the Akt pathway in breast cancer cells. *Int J Oncol* **35:** 673–681.

Albert TK, Hanzawa H, Legtenberg YI, de Ruwe MJ, van den Heuvel FA, Collart MA, Boelens R, Timmers HT. 2002. Identification of a ubiquitin-protein ligase subunit within the CCR4–NOT transcription repressor complex. *EMBO J* **21:** 355–364. doi:10.1093/emboj/21.3.355

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

Bai Y, Salvadore C, Chiang YC, Collart MA, Liu HY, Denis CL. 1999. The CCR4 and CAF1 proteins of the CCR4–NOT complex are physically and functionally separated from NOT2, NOT4, and NOT5. *Mol Cell Biol* **19:** 6642–6651. doi:10.1128/MCB.19.10.6642

Basquin J, Roudko VV, Rode M, Basquin C, Séraphin B, Conti E. 2012. Architecture of the nuclease module of the yeast CCR4–NOT complex: the Not1–Caf1–Ccr4 interaction. *Mol Cell* **48:** 207–218. doi:10.1016/j.molcel.2012.08.014

Bawankar P, Loh B, Wohlbold L, Schmidt S, Izaurralde E. 2013. NOT10 and C2orf29/NOT11 form a conserved module of the CCR4–NOT complex that docks onto the NOT1 N-terminal domain. *RNA Biol* **10:** 228–244. doi:10.4161/rna.23018

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

Bhandari D, Raisch T, Weichenrieder O, Jonas S, Izaurralde E. 2014. Structural basis for the Nanos-mediated recruitment of the CCR4–NOT complex and translational repression. *Genes Dev* **28:** 888–901. doi:10.1101/gad.237289.113

Bhaskar V, Roudko V, Basquin J, Sharma K, Urlaub H, Séraphin B, Conti E. 2013. Structure and RNA-binding properties of the Not1–Not2–Not5 module of the yeast CCR4–NOT complex. *Nat Struct Mol Biol* **20:** 1281–1288. doi:10.1038/nsmb.2686

Bhaskar V, Basquin J, Conti E. 2015. Architecture of the ubiquitylation module of the yeast CCR4–NOT complex. *Structure* **23:** 921–928. doi:10.1016/j.str.2015.03.011

Boland A, Chen Y, Raisch T, Jonas S, Kuzuoğlu-Öztürk D, Wohlbold L, Weichenrieder O, Izaurralde E. 2013. Structure and assembly of the NOT module of the human CCR4–NOT complex. *Nat Struct Mol Biol* **20:** 1289–1297. doi:10.1038/nsmb.2681

Bulbrook D, Brazier H, Mahajan P, Kliszczak M, Fedorov O, Marchese FP, Aubareda A, Chalk R, Picaud S, Strain-Damerell C, et al. 2018. Tryptophan-mediated interactions between tristetraprolin and the CNOT9 subunit are required for CCR4–NOT deadenylase complex recruitment. *J Mol Biol* **430:** 722–736. doi:10.1016/j.jmb.2017.12.018

Chang CT, Bercovich N, Loh B, Jonas S, Izaurralde E. 2014. The activation of the decapping enzyme DCP2 by DCP1 occurs on the EDC4 scaffold and involves a conserved loop in DCP1. *Nucleic Acids Res* **42:** 5217–5233. doi:10.1093/nar/gku129

Chekulaeva M, Mathys H, Zipprich JT, Attig J, Colic M, Parker R, Filipowicz W. 2011. miRNA repression involves GW182-mediated recruitment of CCR4–NOT through conserved W-containing motifs. *Nat Struct Mol Biol* **18:** 1218–1226. doi:10.1038/nsmb.2166

Chen J, Rappsilber J, Chiang YC, Russell P, Mann M, Denis CL. 2001. Purification and characterization of the 1.0 MDa CCR4–NOT complex identifies two novel components of the complex. *J Mol Biol* **314:** 683–694. doi:10.1006/jmbi.2001.5162

Chen Y, Boland A, Kuzuoğlu-Öztürk D, Bawankar P, Loh B, Chang CT, 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

Chicoine J, Benoit P, Gamberi C, Paliouras M, Simonelig M, Lasko P. 2007. Bicaudal-C recruits CCR4–NOT deadenylase to target mRNAs and regulates oogenesis, cytoskeletal organization, and its own expression. *Dev Cell* **13:** 691–704. doi:10.1016/j.devcel.2007.10.002

Collart MA. 2016. The CCR4–NOT complex is a key regulator of eukaryotic gene expression. *Wiley Interdiscip Rev RNA* **7:** 438–454. doi:10.1002/wrna.1332

Cooke A, Prigge A, Wickens M. 2010. Translational repression by deadenylases. *J Biol Chem* **285:** 28506–28513. doi:10.1074/jbc.M110.150763

Cooper KF, Scarnati MS, Krasley E, Mallory MJ, Jin C, Law MJ, Strich R. 2012. Oxidative-stress-induced nuclear to cytoplasmic relocalization is required for Not4-dependent cyclin C destruction. *J Cell Sci* **125:** 1015–1026. doi:10.1242/jcs.096479

Dimitrova LN, Kuroha K, Tatematsu T, Inada T. 2009. Nascent peptide-dependent translation arrest leads to Not4p-mediated protein degradation by the proteasome. *J Biol Chem* **284:** 10343–10352. doi:10.1074/jbc.M808840200

Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of *Coot*. *Acta Crystallogr D Biol Crystallogr* **66:** 486–501. doi:10.1107/S0907444910007493

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

Garces RG, Gillon W, Pai EF. 2007. Atomic model of human Rcd-1 reveals an *armadillo*-like-repeat protein with in vitro nucleic acid binding properties. *Protein Sci* **16:** 176–188. doi:10.1110/ps.062600507

Gulshan K, Thommandru B, Moye-Rowley WS. 2012. Proteolytic degradation of the Yap1 transcription factor is regulated by subcellular localization and the E3 ubiquitin ligase Not4. *J Biol Chem* **287:** 26796–26805. doi:10.1074/jbc.M112.384719

Jonas S, Izaurralde E. 2015. Towards a molecular understanding of microRNA-mediated gene silencing. *Nat Rev Genet* **16:** 421–433. doi:10.1038/nrg3965

Kabsch W. 2010. XDS. *Acta Crystallogr D Biol Crystallogr* **66:** 125–132. doi:10.1107/S0907444909047337

Kofler MM, Freund C. 2006. The GYF domain. *FEBS J* **273:** 245–256. doi:10.1111/j.1742-4658.2005.05078.x

Kryszke MH, Adjeriou B, Liang F, Chen H, Dautry F. 2016. Post-transcriptional gene silencing activity of human GIGYF2. *Biochem Biophys Res Commun* **475:** 289–294. doi:10.1016/j.bbrc.2016.05.022

Kuzuoğlu-Öztürk 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. doi:10.15252/embj.201592901

Lau NC, Kolkman A, van Schaik FM, Mulder KW, Pijnappel WW, Heck AJ, Timmers HT. 2009. Human CCR4–NOT complexes contain variable deadenylase subunits. *Biochem J* **422:** 443–453. doi:10.1042/BJ20090500

Loh B, Jonas S, 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. doi:10.1101/gad.226951.113

Lykke-Andersen J, Shu MD, Steitz JA. 2000. Human Upf proteins target an mRNA for nonsense-mediated decay when bound downstream of a termination codon. *Cell* **103:** 1121–1131. doi:10.1016/S0092-8674(00)00214-2

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

Matsuda R, Ikeuchi K, Nomura S, Inada T. 2014. Protein quality control systems associated with no-go and nonstop mRNA surveillance in yeast. *Genes Cells* **19:** 1–12. doi:10.1111/gtc.12106

Mauxion F, Prève B, Séraphin B. 2013. C2ORF29/CNOT11 and CNOT10 form a new module of the CCR4–NOT complex. *RNA Biol* **10:** 267–276. doi:10.4161/rna.23065

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. 2007. Phaser crystallographic software. *J Appl Crystallogr* **40:** 658–674. doi:10.1107/S0021889807021206

Mersman DP, Du HN, Fingerman IM, South PF, Briggs SD. 2009. Polyubiquitination of the demethylase Jhd2 controls histone methylation and gene expression. *Genes Dev* **23:** 951–962. doi:10.1101/gad.1769209

Nasertorabi F, Batisse C, Diepholz M, Suck D, Böttcher B. 2011. Insights into the structure of the CCR4–NOT complex by electron microscopy. *FEBS Lett* **585:** 2182–2186. doi:10.1016/j.febslet.2011.05.071

Newman R, McHugh J, Turner M. 2016. RNA binding proteins as regulators of immune cell biology. *Clin Exp Immunol* **183:** 37–49. doi:10.1111/cei.12684

Panasenko OO. 2014. The role of the E3 ligase Not4 in cotranslational quality control. *Front Genet* **5:** 141. doi:10.3389/fgene.2014.00141

Panasenko OO, Collart MA. 2012. Presence of Not5 and ubiquitinated Rps7A in polysome fractions depends upon the Not4 E3 ligase. *Mol Microbiol* **83:** 640–653. doi:10.1111/j.1365-2958.2011.07957.x

Panasenko O, Landrieux E, Feuermann M, Finka A, Paquet N, Collart MA. 2006. The yeast CCR4–NOT complex controls ubiquitination of the nascent-associated polypeptide (NAC–EGD) complex. *J Biol Chem* **281:** 31389–31398. doi:10.1074/jbc.M604986200

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

Petit AP, Wohlbold L, Bawankar P, Huntzinger E, Schmidt S, Izaurralde E, Weichenrieder O. 2012. The structural basis for the interaction between the CAF1 nuclease and the NOT1 scaffold of the human CCR4–NOT deadenylase complex. *Nucleic Acids Res* **40:** 11058–11072. doi:10.1093/nar/gks883

Preissler S, Reuther J, Koch M, Scior A, Bruderek M, Frickey T, Deuerling E. 2015. Not4-dependent translational repression is important for cellular protein homeostasis in yeast. *EMBO J* **34:** 1905–1924. doi:10.15252/embj.201490194

Raisch T, Bhandari D, Sabath K, Helms S, Valkov E, Weichenrieder O, Izaurralde E. 2016. Distinct modes of recruitment of the CCR4–NOT complex by *Drosophila* and vertebrate Nanos. *EMBO J* **35:** 974–990. doi:10.15252/embj.201593634

Raisch T, Sandmeir F, Weichenrieder O, Valkov E, Izaurralde E. 2018. Structural and biochemical analysis of a NOT1 MIF4G-like domain of the CCR4–NOT complex. *J Struct Biol* **204:** 388–395. doi:10.1016/j.jsb.2018.10.009

Sgromo A, Raisch T, Bawankar P, Bhandari D, Chen Y, Kuzuoğlu-Öztürk D, Weichenrieder O, Izaurralde E. 2017. A CAF40-binding motif facilitates recruitment of the CCR4–NOT complex to mRNAs targeted by *Drosophila* Roquin. *Nat Commun* **8:** 14307. doi:10.1038/ncomms14307

Sgromo A, Raisch T, Backhaus C, Keskeny C, Alva V, Weichenrieder O, Izaurralde E. 2018. *Drosophila* Bag-of-marbles directly interacts with the CAF40 subunit of the CCR4–NOT complex to elicit repression of mRNA targets. *RNA* **24:** 381–395. doi:10.1261/rna.064584.117

Simonetti F, Candelli T, Leon S, Libri D, Rougemaille M. 2017. Ubiquitination-dependent control of sexual differentiation in fission yeast. *Elife* **6:** 28046. doi:10.7554/eLife.28046

Stowell JAW, Webster MW, Kögel A, Wolf J, Shelley KL, Passmore LA. 2016. Reconstitution of targeted deadenylation by the Ccr4–Not complex and the YTH domain protein Mmi1. *Cell Rep* **17:** 1978–1989. doi:10.1016/j.celrep.2016.10.066

Temme C, Zhang L, Kremmer E, Ihling C, Chartier A, Sinz A, Simonelig M, Wahle E. 2010. Subunits of the *Drosophila* CCR4–NOT complex and their roles in mRNA deadenylation. *RNA* **16:** 1356–1370. doi:10.1261/rna.2145110

Ukleja M, Cuellar J, Siwaszek A, Kasprzak JM, Czarnocki-Cieciura M, Bujnicki JM, Dziembowski A, Valpuesta JM. 2016. The architecture of the *Schizosaccharomyces pombe* CCR4–NOT complex. *Nat Commun* **7:** 10433. doi:10.1038/ncomms10433

The UniProt Consortium. 2018. UniProt: the universal protein knowledgebase. *Nucleic Acids Res* **46:** 2699. doi:10.1093/nar/gky092

Wahle E, Winkler GS. 2013. RNA decay machines: deadenylation by the CCR4–NOT and Pan2–Pan3 complexes. *Biochim Biophys Acta* **1829:** 561–570. doi:10.1016/j.bbagrm.2013.01.003

Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A, et al. 2011. Overview of the *CCP4* suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67:** 235–242. doi:10.1107/S0907444910045749

Wu Z, Wang Y, Lim J, Liu B, Li Y, Vartak R, Stankiewicz T, Montgomery S, Lu B. 2018. Ubiquitination of ABCE1 by NOT4 in response to mitochondrial damage links co-translational quality control to PINK1-directed mitophagy. *Cell Metab* **28:** 130–144.e7. doi:10.1016/j.cmet.2018.05.007

A conserved CAF40 binding motif in metazoan NOT4 mediates association with the CCR4-NOT complex

Csilla Keskeny, Tobias Raisch, Annamaria Sgromo, Cátia Igreja,

Dipankar Bhandari*, Oliver Weichenrieder* and Elisa Izaurralde†

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

* To whom correspondence should be addressed. Tel: +49-7071-601-1358, Fax: +49-7071-601-1353;
Email: oliver.weichenrieder@tuebingen.mpg.de or dipankar.bhandari@tuebingen.mpg.de

† deceased 30 April 2018

Present Address:
Tobias Raisch, Department of Structural Biochemistry, Max Planck Institute of Molecular Physiology, Otto-Hahn-Strasse 11, D-44227 Dortmund, Germany.
Annamaria Sgromo, IMBA - Institute of Molecular Biotechnology GmbH, Dr. Bohr-Gasse 3, 1030 Vienna, Austria.

- Supplemental Material -

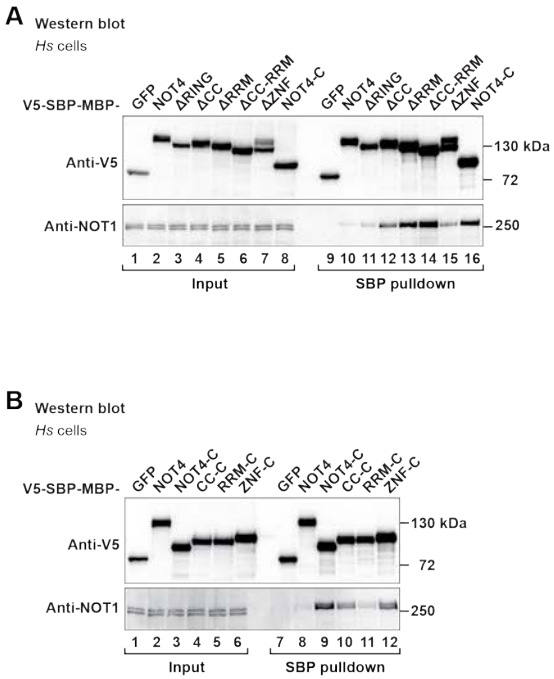
Supplemental figures S1–S7

Supplemental tables S1–S2

Supplemental alignment files

Supplemental references

Supplemental Figure S1. The presence of the positively charged CC linker and RRM domain prevents NOT4-C from interacting with the CCR4-NOT complex

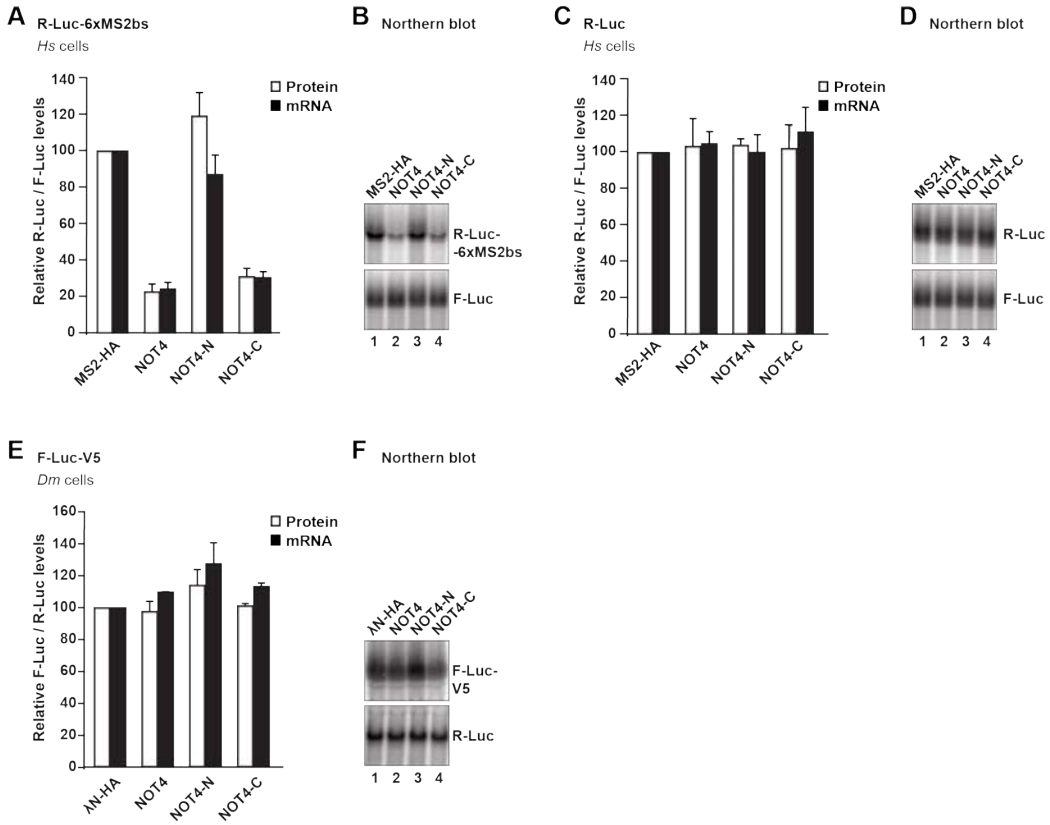


Supplemental Figure S1. The presence of the positively charged CC linker and RRM domain prevents NOT4-C from interacting with the CCR4-NOT complex

(A) SBP pulldown of endogenous human NOT1 from HEK293T cell lysates with V5-SBP-MBP-tagged deletion variants of *Hs* NOT4. V5-SBP-GFP-MBP served as negative control. Input samples correspond to 3 % of the total lysate for V5-tagged proteins and to 2 % for NOT1. Pulldown samples correspond to 7 % of the total pulldown for V5-tagged proteins and to 35 % for NOT1. For abbreviations and construct details see Figure 1B and Supplemental Table S1.

(B) SBP pulldown of endogenous human NOT1 from HEK293T cell lysates with V5-SBP-MBP-tagged fragments of *Hs* NOT4-N fused to *Hs* NOT4-C. CC-C: Putative coiled coil region fused to NOT4-C; RRM-C: RRM domain fused to NOT4-C; ZNF-C: zinc-finger domain fused to NOT4-C.

Supplemental Figure S2. Metazoan NOT4 induces degradation of tethered mRNA reporters



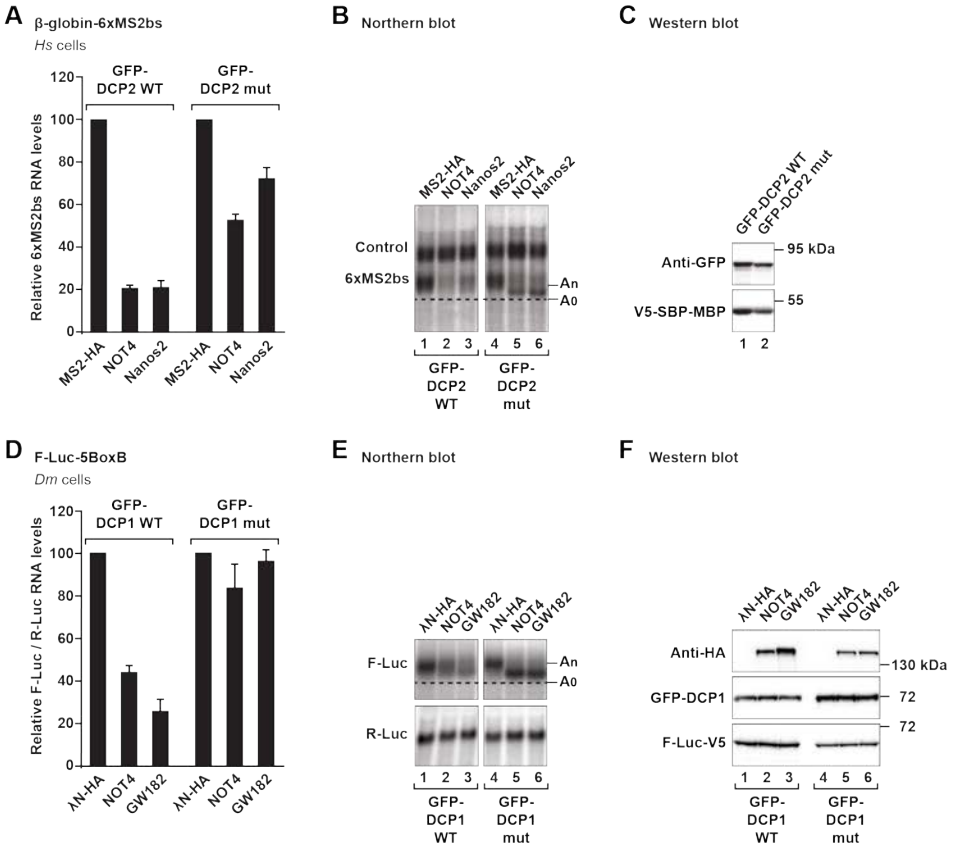
Supplemental Figure S2. Metazoan NOT4 induces degradation of tethered mRNA reporters

(A, B) Tethering assay with *Hs* NOT4 and a luciferase reporter in HEK293T cells. *Hs* NOT4 or its fragments carried an N-terminal MS2-HA tag. Renilla luciferase mRNA served as a reporter and contained six binding sites for the MS2 protein (R-Luc-6xMS2bs). Firefly luciferase mRNA served as a reference and transfection control (F-Luc). (A) R-Luc activity (white bars) and mRNA levels (black bars) normalized to the reference and plotted with respect to the values obtained from the expression of MS2-HA alone (set to 100). Error bars correspond to standard deviations (n=3). (B) Representative Northern blot.

(C, D) Control experiment for the tethering assay in panels (A, B). Renilla luciferase mRNA without binding sites for the MS2 protein (R-Luc) served as a tethering control. (C) Relative protein and mRNA levels analyzed as described in (A). (D) Northern blot.

(E, F) Control experiment for the tethering assay in Figure 2D, E with *Dm* NOT4 and a luciferase reporter in *Dm* S2 cells. Firefly luciferase mRNA without BoxB binding sites (F-Luc-V5) served as a tethering control. (E) Relative protein and mRNA levels analyzed as described in Figure 2D. (F) Northern blot.

Supplemental Figure S3. Tethered NOT4 causes reporter mRNA degradation via the 5'-to-3' decay pathway



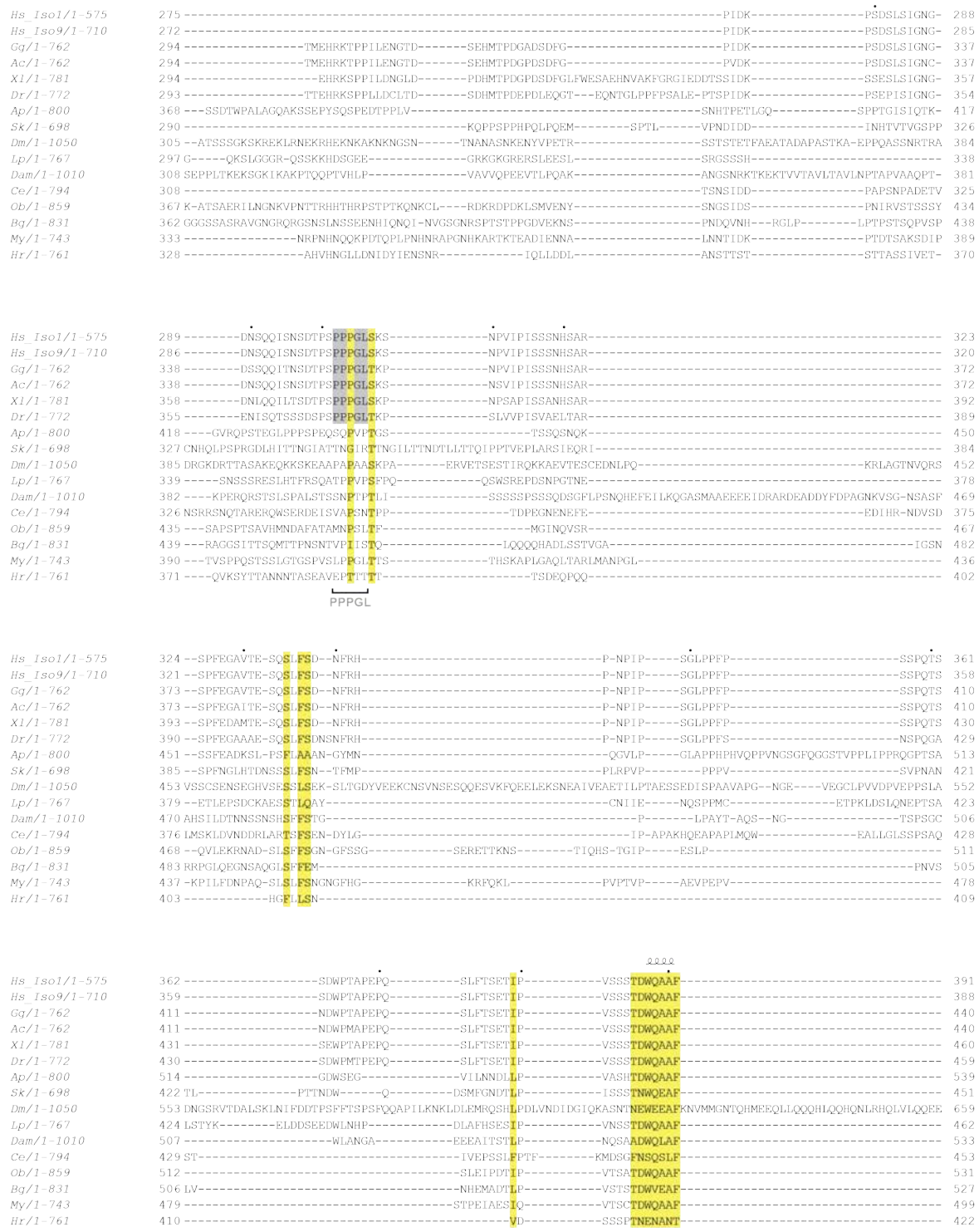
Supplemental Figure S3. Tethered NOT4 causes reporter mRNA degradation via the 5'-to-3' decay pathway

(A–C) Tethering assay with *Hs* NOT4 and the β-globin mRNA reporter in HEK293T cells overexpressing a DCP2 catalytically inactive mutant. Experiments were done and analyzed as described in Figure 2A, but cells were additionally overexpressing either wildtype *Hs* DCP2 (GFP-DCP2 WT) or the *Hs* DCP2 mutant (GFP-DCP2 mut; E148Q). MS2-HA-tagged *Hs* Nanos2 was overexpressed as a positive control for a protein eliciting 5'-to-3' mRNA decay (Bhandari et al. 2014). (A) Relative mRNA levels with error bars corresponding to standard deviations (n=3). (B) Northern blot demonstrating the accumulation of deadenylated mRNA degradation intermediates (A0) as compared to polyadenylated mRNA (An). (C) Western blot demonstrating equivalent expression of GFP-tagged DCP2 proteins with V5-SBP-MBP as a transfection control.

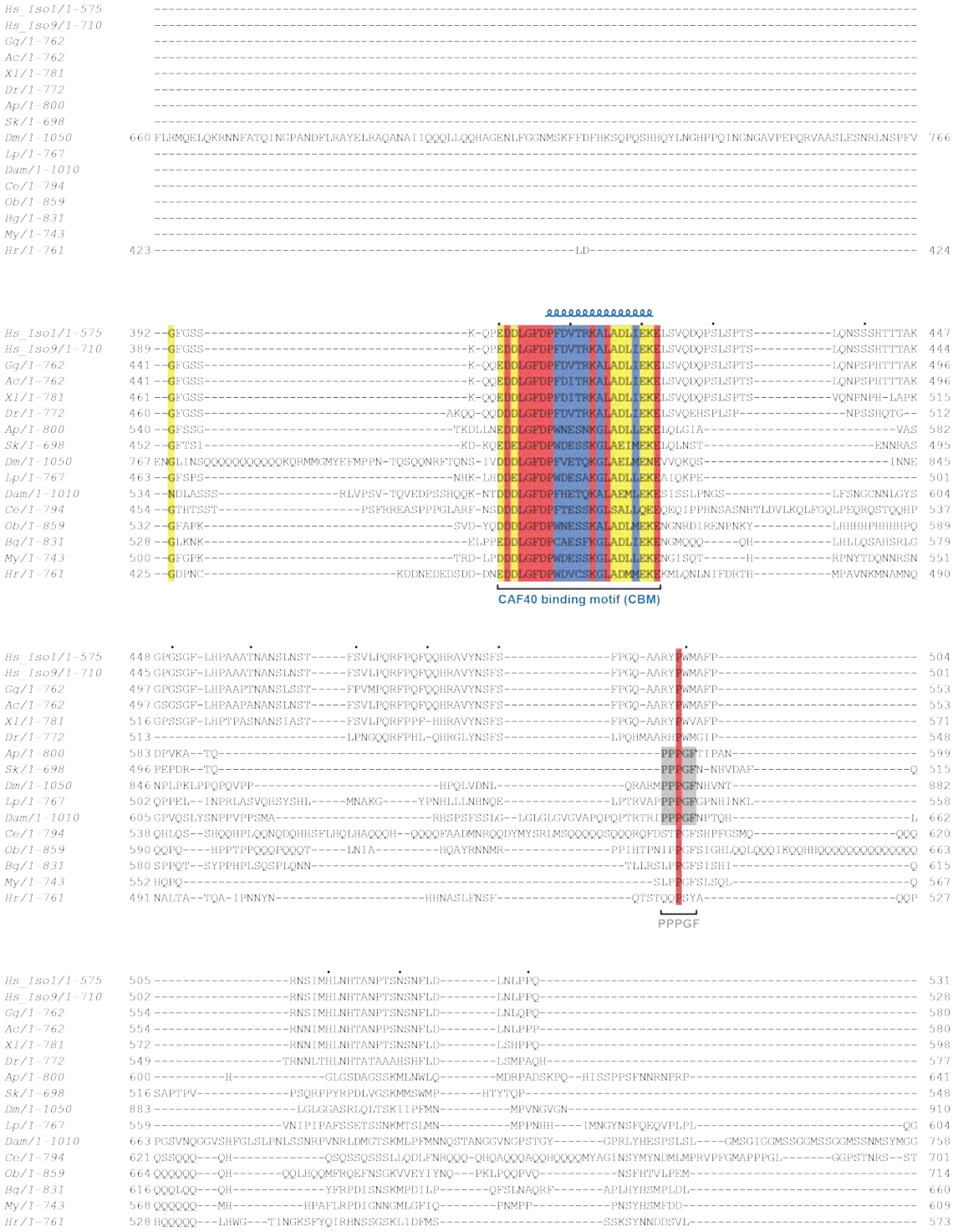
(D–F) Tethering assay with *Dm* NOT4 in *Dm* S2 cells overexpressing a DCP1 mutant. Experiments were done and analyzed as described in Figure 2D, but cells were additionally overexpressing either wildtype *Dm* DCP1 (GFP-DCP1 WT) or the *Dm* DCP1 mutant (GFP-DCP1 mut; R70G, L71S, N72S, T73G). λN-HA-tagged GW182 was overexpressed as a positive control for a protein eliciting 5'-to-3' mRNA decay (Kuzuoğlu-Öztürk et al. 2016). (D) Relative mRNA levels with error bars corresponding to standard deviations (n=3). (E) Northern blot demonstrating the accumulation of deadenylated mRNA degradation intermediates (A0) as compared to polyadenylated mRNA (An). (F) Western blot demonstrating equivalent expression of GFP-tagged DCP1 proteins and λN-HA-tagged proteins (NOT4 and GW182), with F-Luc-V5 as a transfection control.

Supplemental Figure S4. Sequence alignment of metazoan NOT4

Supplemental Figure S4, continued. Sequence alignment of metazoan NOT4



Supplemental Figure S4, continued. Sequence alignment of metazoan NOT4

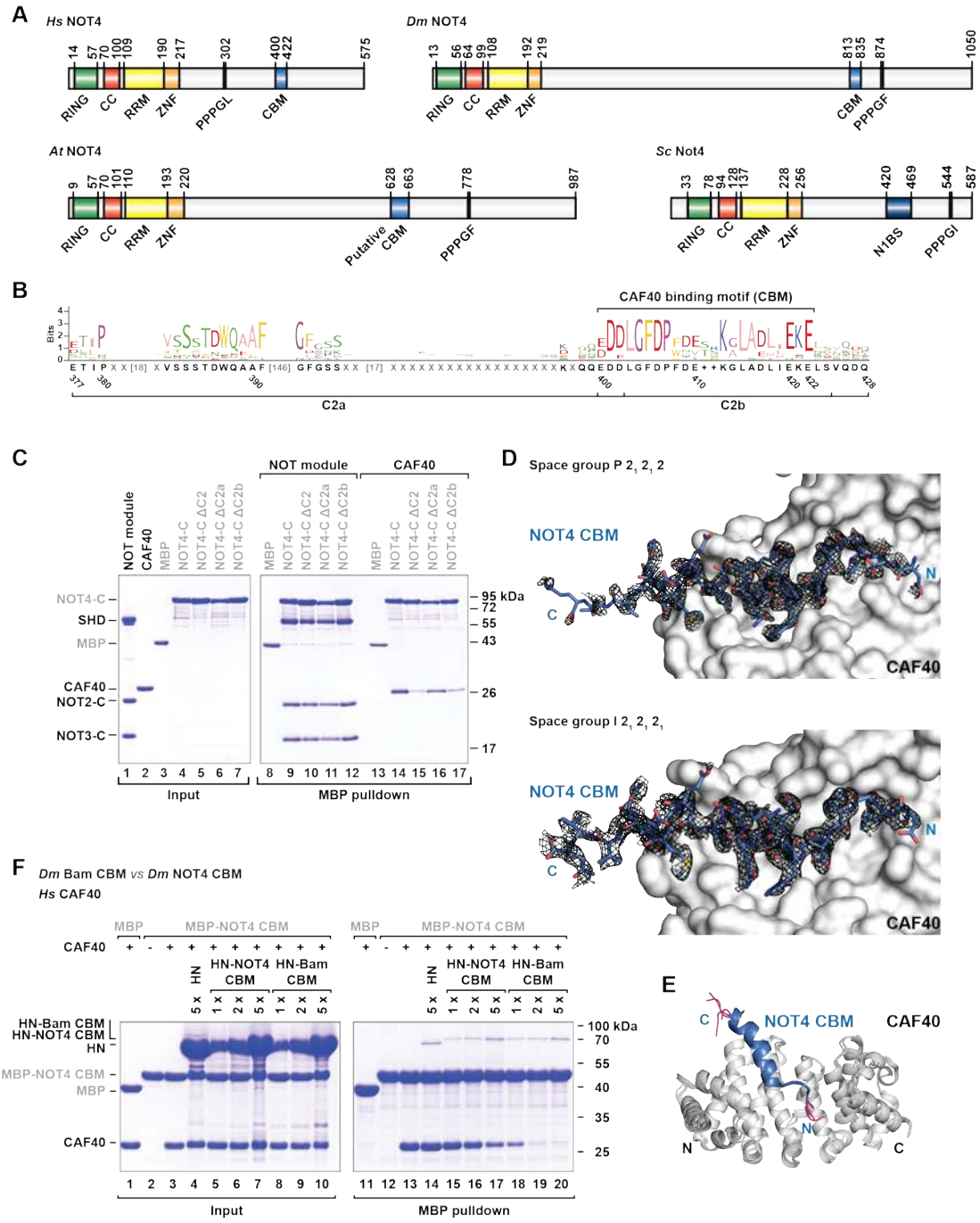


Supplemental Figure S4, continued. Sequence alignment of metazoan NOT4

Supplemental Figure S4. Sequence alignment of metazoan NOT4

Sequences cover the following phyla. Chordata: *Hs*, *Homo sapiens*; *Gg*, *Gallus gallus*; *Ac*, *Anolis carolinensis*; *Xl*, *Xenopus laevis*; *Dr*, *Danio rerio*. Echinodermata: *Ap*, *Acanthaster planci*. Hemichordata: *Sk*, *Saccoglossus kowalevskii*. Arthropoda: *Dm*, *Drosophila melanogaster*; *Lp*, *Limulus polyphemus*; *Dam*, *Daphnia magna*. Nematoda: *Ce*, *Caenorhabditis elegans*. Mollusca: *Ob*, *Octopus bimaculoides*; *Bg*, *Biomphalaria glabrata*; *My*, *Mizuhopecten yessoensis*. Annelida: *Hr*, *Helobdella robusta*. Uniprot or NCBI accession numbers are provided in the Supplemental alignment file SF1. Sequences were aligned using MAFFT (Katoh et al. 2002) as implemented in the MPI Bioinformatics Toolkit (Zimmermann et al. 2018). Structural domains in the conserved N-terminal region of NOT4 are marked below the alignment. Similarly, the presently identified CBM and PPPGΦ motifs are marked in the non-conserved C-terminal region of NOT4 and respectively shaded in blue and gray. Conserved residues are shaded in red, and residues with >70% similarity are shaded in yellow, with conservation scores calculated using the SCORECONS webserver (Valdar 2002). Secondary structure elements are indicated above the alignment and taken from the previously reported NMR structures of the RING and RRM domains (PDB-ID 1ur6, Dominguez et al. 2004, and PDB-ID 2cpi), from our presently determined crystal structure of the CBM, or from PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/) secondary structure prediction.

Supplemental Figure S5. Identification and analysis of the NOT4 CBM



Supplemental Figure S5. Identification and analysis of the NOT4 CBM

(A) Domain composition of selected NOT4 proteins from *Homo sapiens*, *Drosophila melanogaster*, *Arabidopsis thaliana* and *Saccharomyces cerevisiae*, indicating the relative positions of conserved sequence motifs in NOT4-C. The sequences corresponding to the 23 amino acid motif are labelled 'CBM' or 'CBM-like' in *Arabidopsis thaliana*, and the sequences corresponding to the PPPGΦ motifs are listed as PPPGL, PPPGF or PPPGI. For alignments see Supplemental alignment files SF1–SF3.

(B) Sequence logo (Waterhouse et al. 2009) of the NOT4 C2 region as derived from the alignment in Supplemental Figure S4. The consensus sequence is indicated below the logo, along with residue numbers corresponding to the human sequence. The separately deleted C2a and C2b regions are indicated as well as the 23 amino acid motif which is labelled as 'CBM'.

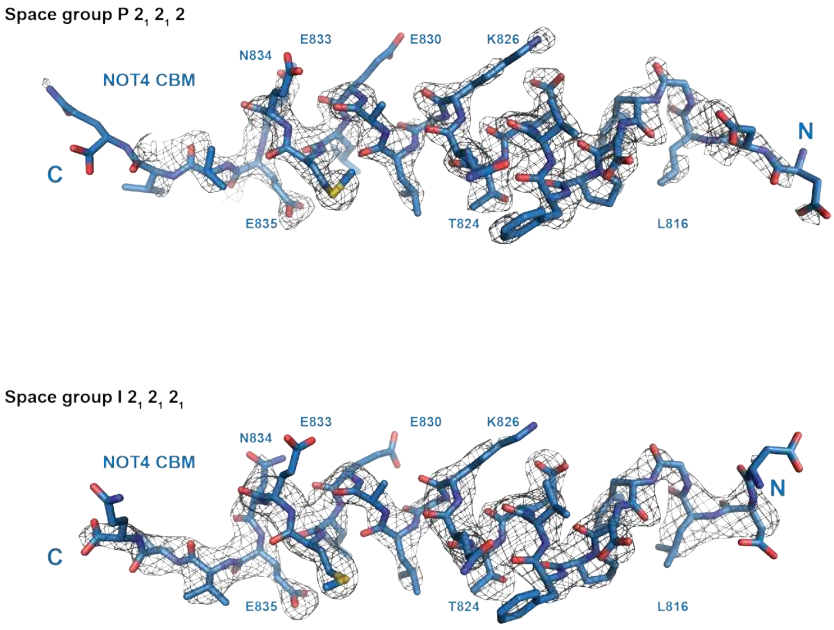
(C) MBP pulldown assay with MBP-tagged deletion variants of *Hs* NOT4-C and purified recombinant NOT module or CAF40. Experiments were done as described in Figure 3A–D, identifying the 23 amino acid motif as a CAF40 binding motif (CBM). NOT4-C ΔC2 lacks residues E377–S424. NOT4-C ΔC2a lacks residues E377–D402. NOT4-ΔC2b lacks residues E400–Q428 including the CBM (E400–E422).

(D) Composite omit maps for the NOT4 CBM peptide. $2F_O$-F_C type electron density surrounding the CBM peptide is contoured at 1.0 σ. The maps were generated with phenix.composite_omit_map (Afonine et al. 2012) using the final refined models, respectively. Top panel: space group $P2_12_12$. Bottom panel: space group $I2_12_12_1$.

(E) Superposition of the four crystallographically independent complexes of CAF40 with the NOT4 CBM. The structurally variable and non-conserved flanks of the CBM are shown in magenta. The C-terminal flanks of the *Dm* NOT4 CBM (V836–Q838) mediate crystal contacts and differ from the *Hs* NOT4 CBM (L423–V425), possibly explaining why the *Hs* NOT4 CBM did not crystallize.

(F) Competition assay. *Hs* CAF40 was incubated with equimolar amounts of MBP-tagged *Dm* NOT4 CBM and increasing amounts of His_6-NusA-tagged *Dm* NOT4 CBM (HN-NOT4 CBM) or His_6-NusA-tagged *Dm* Bam CBM (HN-Bam CBM). His_6-NusA (HN) was used as a negative control. The amount of *Hs* CAF40 pulled down with the MBP-tagged *Dm* NOT4 CBM was analyzed by SDS-PAGE and subsequent Coomassie staining. Molar equivalents (1x, 2x, 5x) are relative to the MBP-tagged *Dm* NOT4 CBM. MBP-tagged constructs are labelled in gray. See Sgromo et al. (2018) for additional experimental details.

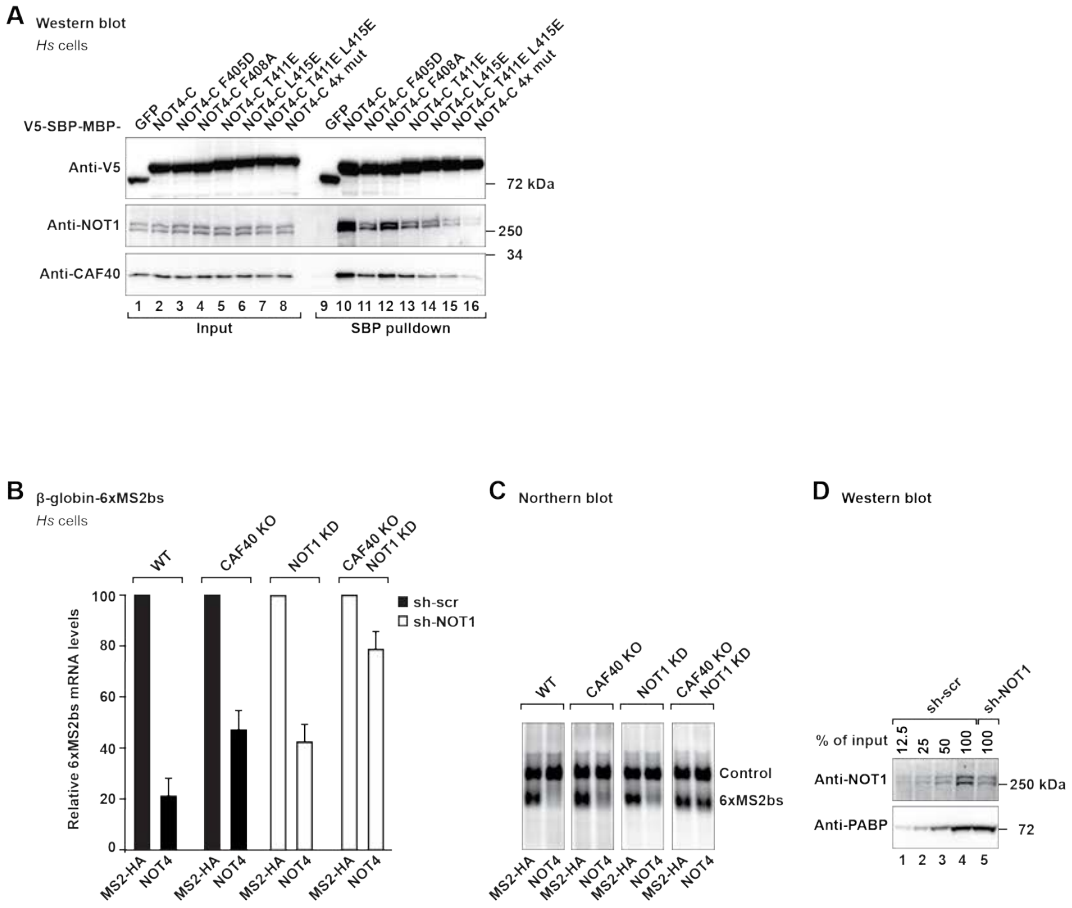
Supplemental Figure S6. Difference electron density for the CBM peptide



Supplemental Figure S6. Difference electron density for the CBM peptide

Classic F_O-F_C type difference electron density for the CBM peptide was calculated using the final refined models, respectively, but with the CBM peptide, ligands and water molecules omitted. The maps are contoured at 2.5 σ. The difference density is rather poor for the N-terminal and C-terminal residues of the modeled peptide and for the solvent-exposed side chains of K826, E830, E833 and N834. These residues were nevertheless modeled without truncating their side chains. The structurally conserved part of the CBM peptide as shown in Figure 4 ranges from L816 to E835. Top panel: space group $P2_12_12$. Bottom panel: space group $I2_12_12_1$.

Supplemental Figure S7. Importance of the NOT4 CBM and of the CCR4-NOT deadenylase for the degradation of tethered mRNA



Supplemental Figure S7. Importance of the NOT4 CBM and of the CCR4-NOT deadenylase for the degradation of tethered mRNA

(A) SBP pulldown of endogenous *Hs* NOT1 and *Hs* CAF40 proteins with V5-SBP-MBP-tagged CBM mutation variants of *Hs* NOT4-C. The CBM contained single or multiple point mutations. 4x mut; F405D, F408A, T411E, L415E. For additional details see Figure 6E.

(B–D) Tethering assay with *Hs* NOT4 and the β-globin mRNA reporter in HEK293T cells lacking CAF40 (CAF40 KO) (Sgromo et al. 2018). Efficiency of CRISPR-Cas9-mediated gene editing of CAF40 is shown in Figure 7B. Experiments were done and analyzed as described in Figure 2A. (B) Relative mRNA levels with error bars corresponding to standard deviations (n=3). The shRNA-mediated depletion of the CCR4-NOT complex is indicated by white bars (sh-NOT1 RNA, NOT1 KD) as compared to black bars (sh-scr RNA, control). WT: wildtype, KO: knockout, KD: knockdown. (C) Northern blot. **(**D**)** Efficiency of shRNA-mediated depletion of NOT1. The Western blot shows HEK293T cells expressing sh-NOT1 RNA in comparison to a dilution series of HEK293T cells expressing sh-scr RNA using PABP as a loading control.

Supplemental Table S1. Generated plasmid constructs

Plasmids expressing *Hs* NOT4 protein and *Hs* NOT4 fragments (Uniprot-ID O95628-1)

Name	Construct details	Plasmid backbone *
Hs NOT4	MS2-HA-*Hs* NOT4	pcDNA3.1-MS2-HA
	V5-SBP-MBP-*Hs* NOT4	pCIneo-V5-SBP-MBP
	MBP-*Hs* NOT4-GB1-6xHis	pnEA-pM, GB1
Hs NOT4-N	MS2-HA-*Hs* NOT4_1–274	pcDNA3.1-MS2-HA
	V5-SBP-MBP-*Hs* NOT4_1–274	pCIneo-V5-SBP-MBP
	MBP-*Hs* NOT4_1–274-GB1-6xHis	pnEA-pM, GB1
Hs NOT4-C	MS2-HA-*Hs* NOT4_275–575	pcDNA3.1-MS2-HA
	V5-SBP-MBP-*Hs* NOT4_275–575	pCIneo-V5-SBP-MBP
	MBP-*Hs* NOT4_275–575-GB1-6xHis	pnEA-pM, GB1
Hs NOT4 ΔRING	V5-SBP-MBP-*Hs* NOT4 (Δ1–63)	pCIneo-V5-SBP-MBP
Hs NOT4 ΔCC	V5-SBP-MBP-*Hs* NOT4 (Δ64–104)	pCIneo-V5-SBP-MBP
Hs NOT4 ΔRRM	V5-SBP-MBP-*Hs* NOT4 (Δ109–189)	pCIneo-V5-SBP-MBP
Hs NOT4 ΔCC-RRM	V5-SBP-MBP-*Hs* NOT4 (Δ64–189)	pCIneo-V5-SBP-MBP
Hs NOT4 ΔZNF	V5-SBP-MBP-*Hs* NOT4 (Δ190–274)	pCIneo-V5-SBP-MBP
Hs NOT4-CC-C	V5-SBP-MBP-*Hs* NOT4_64–104_GSSG_275–575	pCIneo-V5-SBP-MBP
Hs NOT4-RRM-C	V5-SBP-MBP-*Hs* NOT4_109–189_GSSG_275–575	pCIneo-V5-SBP-MBP
Hs NOT4-ZNF-C	V5-SBP-MBP-*Hs* NOT4_190–575	pCIneo-V5-SBP-MBP
Hs NOT4-C1	MS2-HA-*Hs* NOT4_275–376	pcDNA3.1-MS2-HA
	MBP-*Hs* NOT4_275–376-GB1-6xHis	pnEA-pM, GB1
Hs NOT4-C2	MS2-HA-*Hs* NOT4_377–428	pcDNA3.1-MS2-HA
	MBP-*Hs* NOT4_377–428-GB1-6xHis	pnEA-pM, GB1
Hs NOT4-C3	MS2-HA-*Hs* NOT4_429–575	pcDNA3.1-MS2-HA
	MBP-*Hs* NOT4_429–575-GB1-6xHis	pnEA-pM, GB1
Hs NOT4 ΔC2	MS2-HA-*Hs* NOT4 (Δ377–424)	pcDNA3.1-MS2-HA
Hs NOT4-C ΔC2	V5-SBP-MBP-*Hs* NOT4_275–575 (Δ377–424)	pCIneo-V5-SBP-MBP
	MBP-*Hs* NOT4_275–575-GB1-6xHis (Δ377–424)	pnEA-pM, GB1
Hs NOT4 ΔC2a	MS2-HA-*Hs* NOT4 (Δ377–402)	pcDNA3.1-MS2-HA
Hs NOT4-C ΔC2a	V5-SBP-MBP-*Hs* NOT4_275–575 (Δ377–402)	pCIneo-V5-SBP-MBP
	MBP-*Hs* NOT4_275–575-GB1-6xHis (Δ377–402)	pnEA-pM, GB1
Hs NOT4 ΔC2b	MS2-HA-*Hs* NOT4 (Δ400–428)	pcDNA3.1-MS2-HA
Hs NOT4-C ΔC2b	V5-SBP-MBP-*Hs* NOT4_275–575 (Δ400–428)	pCIneo-V5-SBP-MBP
	MBP-*Hs* NOT4_275–575-GB1-6xHis (Δ400–428)	pnEA-pM, GB1
Hs NOT4-CBM	MBP-*Hs* NOT4_400–427	pnEA-pM

* Plasmid backbones are described in Jonas et al. (2013) for pcDNA3.1-MS2-HA and pCIneo-V5-SBP-MBP, in Diebold et al. (2011) for pnEA-pM and in Cheng and Patel (2004) for GB1.

Supplemental Table S1, continued. Generated plasmid constructs

Plasmids expressing *Hs* NOT4 point mutations (Uniprot-ID O95628-1)

Name	Construct details	Plasmid backbone *
Hs NOT4 4x mut	MS2-HA-*Hs* NOT4 (F405D, F408A, T411E, L415E)	pcDNA3.1-MS2-HA
Hs NOT4-C 4x mut	V5-SBP-MBP-*Hs* NOT4_275–575 (F405D, F408A, T411E, L415E)	pCIneo-V5-SBP-MBP
Hs NOT4-C 2x mut	V5-SBP-MBP-*Hs* NOT4_275–575 (T411E, L415E)	pCIneo-V5-SBP-MBP
Hs NOT4-C F405D	V5-SBP-MBP-*Hs* NOT4_275–575 (F405D)	pCIneo-V5-SBP-MBP
Hs NOT4-C F408A	V5-SBP-MBP-*Hs* NOT4_275–575 (F408A)	pCIneo-V5-SBP-MBP
Hs NOT4-C T411E	V5-SBP-MBP-*Hs* NOT4_275–575 (T411E)	pCIneo-V5-SBP-MBP
Hs NOT4-C L415E	V5-SBP-MBP-*Hs* NOT4_275–575 (L415E)	pCIneo-V5-SBP-MBP

* Plasmid backbones are described in Jonas et al. (2013).

Plasmids expressing *Dm* NOT4 proteins (Uniprot-ID M9PCL9)

Name	Construct details	Plasmid backbone *
Dm NOT4	λN-HA-*Dm* NOT4	pAc5.1B-λN-HA
Dm NOT4-N	λN-HA-*Dm* NOT4_1–249	pAc5.1B-λN-HA
Dm NOT4-C	λN-HA-*Dm* NOT4_248–1050	pAc5.1B-λN-HA
Dm NOT4-CBM	MBP-*Dm* NOT4_813–836	pnYC-vM
	His_6-NusA-*Dm* NOT4_813–836	pnYC-vHN
Dm NOT4-CBM F821D	MBP-*Dm* NOT4_813–836 (F821D)	pnYC-vM
Dm NOT4-CBM L828E	MBP-*Dm* NOT4_813–836 (L828E)	pnYC-vM

* Plasmid backbones are described in Rehwinkel et al. (2005) for pAc5.1B-λN-HA and in Diebold et al. (2011) for pnYC-vM.

Supplemental Table S2. Antibodies

Antibody	Source	Catalog number	Dilution	Monoclonal/polyclonal
Anti-NOT1	In house	-	1:1000	Rabbit polyclonal
Anti-NOT2	Bethyl	A302-562A	1:2000	Rabbit polyclonal
Anti-NOT3	Abcam	ab55681	1:2000	Mouse monoclonal
Anti-CAF40	Proteintech	22503-1-AP	1:1000	Rabbit polyclonal
Anti-EDC4	Santa Cruz Biotech	SC-8418	1:1000	Mouse monoclonal
Anti-PABP	Abcam	ab21060	1:10000	Rabbit polyclonal
Anti-tubulin	Sigma	T6199	1:5000	Mouse monoclonal
Anti-V5	LSBio	LS-C57305	1:5000	Mouse monoclonal
Anti-GFP	Roche	11 814 460 001	1:2000	Mouse monoclonal
Anti-HA-HRP	Roche	12 013 819 001	1:5000	Rat monoclonal

Supplemental alignment files

Supplemental alignment files contain CLUSTAL-formatted text files with sub-alignments of NOT4 proteins and their Uniprot or NCBI accession codes.

SF1: SF1_metazoan_NOT4_alg.rtf

SF2: SF2_plant_NOT4_alg.rtf

SF3: SF3_yeast_NOT4_alg.rtf

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.

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

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.

Dominguez C, Bonvin AM, Winkler GS, van Schaik FM, Timmers HT, Boelens R. 2004. Structural model of the UbcH5B/CNOT4 complex revealed by combining NMR, mutagenesis, and docking approaches. *Structure* **12**: 633-644.

Jonas S, Weichenrieder O, 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.

Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* **30**: 3059-3066.

Kuzuoğlu-Öztürk 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.

Rehwinkel J, Behm-Ansmant I, Gatfield D, Izaurralde E. 2005. A crucial role for GW182 and the DCP1:DCP2 decapping complex in miRNA-mediated gene silencing. *RNA* **11**: 1640-1647.

Sgromo A, Raisch T, Backhaus C, Keskeny C, Alva V, Weichenrieder O, Izaurralde E. 2018. Drosophila Bag-of-marbles directly interacts with the CAF40 subunit of the CCR4-NOT complex to elicit repression of mRNA targets. *RNA* **24**: 381-395.

Valdar WS. 2002. Scoring residue conservation. *Proteins* **48**: 227-241.

Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. 2009. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**: 1189-1191.

Zimmermann L, Stephens A, Nam SZ, Rau D, Kübler J, Lozajic M, Gabler F, Söding J, Lupas AN, Alva V. 2018. A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. *J Mol Biol* **430**: 2237-2243.

Drosophila Bag-of-marbles directly interacts with the CAF40 subunit of the CCR4–NOT complex to elicit repression of mRNA targets

ANNAMARIA SGROMO,[1,3] TOBIAS RAISCH,[1,3] CHARLOTTE BACKHAUS,[1] CSILLA KESKENY,[1] VIKRAM ALVA,[2] OLIVER WEICHENRIEDER,[1] and ELISA IZAURRALDE[1]

[1]Department of Biochemistry, Max Planck Institute for Developmental Biology, Tübingen, D-72076, Germany
[2]Department of Protein Evolution, Max Planck Institute for Developmental Biology, Tübingen, D-72076, Germany

ABSTRACT

***Drosophila melanogaster* Bag-of-marbles (Bam) promotes germline stem cell (GSC) differentiation by repressing the expression of mRNAs encoding stem cell maintenance factors. Bam interacts with Benign gonial cell neoplasm (Bgcn) and the CCR4 deadenylase, a catalytic subunit of the CCR4–NOT complex. Bam has been proposed to bind CCR4 and displace it from the CCR4–NOT complex. Here, we investigated the interaction of Bam with the CCR4–NOT complex by using purified recombinant proteins. Unexpectedly, we found that Bam does not interact with CCR4 directly but instead binds to the CAF40 subunit of the complex in a manner mediated by a conserved N-terminal CAF40-binding motif (CBM). The crystal structure of the Bam CBM bound to CAF40 reveals that the CBM peptide adopts an α-helical conformation after binding to the concave surface of the crescent-shaped CAF40 protein. We further show that Bam-mediated mRNA decay and translational repression depend entirely on Bam's interaction with CAF40. Thus, Bam regulates the expression of its mRNA targets by recruiting the CCR4–NOT complex through interaction with CAF40.**

Keywords: deadenylation; mRNA decay; translational repression; germ cell differentiation

INTRODUCTION

The CCR4–NOT deadenylase complex is a major downstream effector complex in post-transcriptional mRNA regulation in eukaryotes (Wahle and Winkler 2013). Beyond its role in global mRNA degradation, the complex regulates the expression of a large number of specific mRNAs, to which it is recruited via interactions with RNA-associated proteins. These proteins include the GW182 family, which is involved in miRNA-mediated gene silencing in animals (Chen et al. 2014a; Mathys et al. 2014); tristetraprolin (TTP), a protein required for the degradation of mRNAs containing AU-rich elements (Fabian et al. 2013); the germline determinant Nanos (Suzuki et al. 2012; Bhandari et al. 2014; Raisch et al. 2016); and human (*Hs*) and *Drosophila melanogaster* (*Dm*) Roquin proteins (Leppek et al. 2013; Sgromo et al. 2017).

In metazoans, the CCR4–NOT complex comprises a core of seven proteins, which bind to independently folding α-helical domains in the central NOT1 scaffold subunit, forming four subcomplexes or modules: the catalytic module, the CAF40 module, the NOT module, and the NOT10-NOT11 module (Wahle and Winkler 2013). The catalytic module comprises two deadenylases, namely CAF1 or its paralog POP2 (also known as CNOT7 and CNOT8, respectively, in humans), and CCR4a or its paralog CCR4b (also known as CNOT6 and CNOT6L, respectively, in humans). CAF1 (or POP2) binds to a central domain of NOT1 that is structurally related to the middle portion of eIF4G (termed the NOT1 MIF4G domain) (Basquin et al. 2012; Petit et al. 2012). CAF1 or POP2 also bind to a leucine-rich repeat domain (LRR) in CCR4a/b, thus bridging the interaction of CCR4 paralogs with NOT1, and consequently with the assembled CCR4–NOT complex (Draper et al. 1994, 1995; Dupressoir et al. 2001; Basquin et al. 2012; Petit et al. 2012; Bawankar et al. 2013). The NOT1 MIF4G domain also serves as a binding platform for the DEAD-box protein DDX6 (also known as RCK), which functions as a translational repressor and decapping activator (Chen et al. 2014a; Mathys et al. 2014). The CAF40 module consists of the highly conserved CAF40 subunit (also known as NOT9) bound to the NOT1 CAF40/NOT9-binding domain (CN9BD), which is located

[3]**These authors contributed equally to this work.**

Corresponding author: elisa.izaurralde@tuebingen.mpg.de

Article is online at http://www.rnajournal.org/cgi/doi/10.1261/rna.064584.117. Freely available online through the RNA Open Access option.

C-terminal to the MIF4G domain (Chen et al. 2014a; Mathys et al. 2014). The NOT module consists of the NOT2–NOT3 heterodimer bound to the C-terminal NOT1 superfamily homology domain SHD (Bhaskar et al. 2013; Boland et al. 2013), whereas the NOT10 and NOT11 subunits bind to the N-terminal end of NOT1 (Bawankar et al. 2013; Mauxion et al. 2013).

Bag-of-marbles (Bam) is a key differentiation factor that determines the fate of germline stem cells (GSCs) (Cooley et al. 1988; McKearin and Spradling 1990; Carreira and Buszczak 2014). Loss of Bam results in uncontrolled stem cell proliferation, thus giving rise to germ cell tumors that characterize the mutant phenotype (McKearin and Ohlstein 1995). In contrast, ectopic Bam expression causes stem cell loss (Ohlstein and McKearin 1997). Bam is conserved in *Drosophila* and other dipteran species and contains several predicted α-helices (Fig. 1A), thus suggesting that it is mainly a folded protein. However, Bam does not display detectable similarity to other known proteins or domains.

Bam controls GSC differentiation by post-transcriptionally repressing the expression of *nanos* and *E-cadherin* mRNAs (Li et al. 2009). Bam function requires the assembly of a protein complex, which includes Benign gonial cell neoplasm (Bgcn), a putative DEXH RNA helicase protein, and additional proteins such as Tumorous testis (Tut) (Chen et al. 2014b), Sex-lethal (Sxl) (Chau et al. 2012) and Mei-P26 (Neumüller et al. 2008; Li et al. 2013). All of these proteins have been implicated in germ cell differentiation in *Dm*, but their individual contributions to mRNA binding and repression, as well as their interaction modes are not well understood. Bam has also been shown to interact with the translation initiation factor eIF4A and to antagonize its role in translation initiation (Shen et al. 2009).

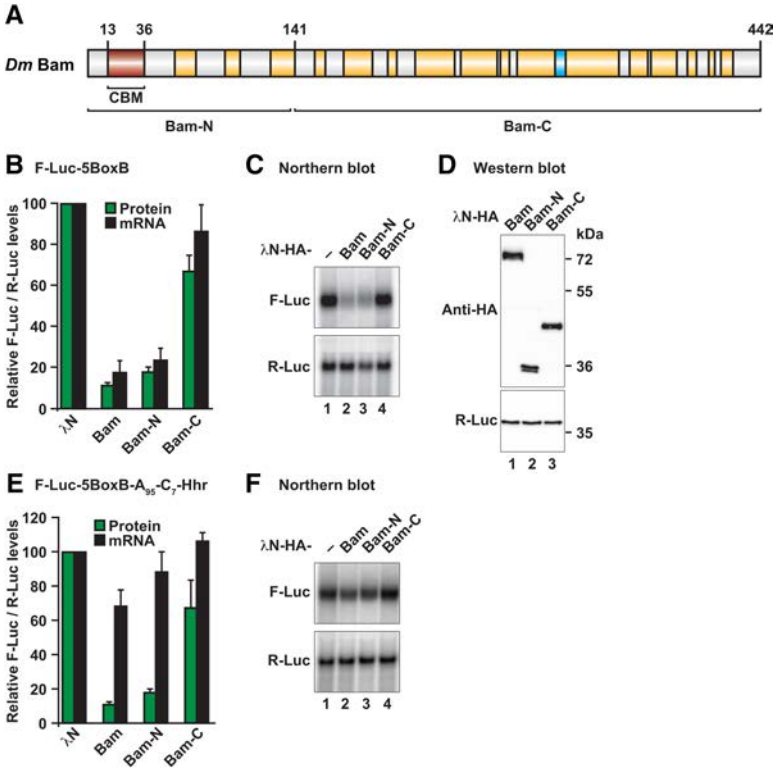


FIGURE 1. Bam induces degradation of bound mRNAs through its N-terminal region. (*A*) Bam consists of several predicted α-helices (shown in yellow) and a β-strand (shown in cyan). The position of the CAF40-binding motif (CBM, in red) as well as the boundaries of the Bam-N and Bam-C fragments are indicated. (*B*) Tethering assay using the F-Luc-5BoxB reporter and λN-HA-tagged Bam (full-length or the indicated fragments) in *Dm* S2 cells. A plasmid expressing R-Luc mRNA served as a transfection control. For each experiment, F-Luc activity and mRNA levels were normalized to those of the R-Luc transfection control and set to 100% in cells expressing the λN-HA peptide. (*C*) Northern blot of representative RNA samples shown in *B*. (*D*) Western blot analysis showing the equivalent expression of the λN-HA-tagged proteins used in the tethering assays shown in *B* and *C*. Protein size markers (kDa) are shown on the *right* of the panel. Full-length Bam and Bam-N display an aberrant mobility in SDS–PAGE, thus resulting in a higher apparent molecular weight. (*E*,*F*) Tethering assay using the F-Luc-5BoxB-A_{95}-C_7-HhR reporter and λN-HA-tagged Bam (full-length or the indicated fragments) in *Dm* S2 cells. The samples were analyzed as described in *B* and *C*. In *B* and *E*, bars represent mean values and error bars represent standard deviations from at least three independent experiments.

Although the mechanism through which Bam-containing complexes repress the expression of specific mRNA targets has not been fully elucidated, it apparently involves interaction with the CCR4 deadenylase subunit of the CCR4–NOT complex (Fu et al. 2015). Bam has been proposed to compete with CAF1/POP2 for direct binding to the CCR4 LRR domain, thereby displacing CCR4 from the CCR4–NOT complex. In this model, CCR4 participates in Bam-mediated repression as an isolated deadenylase and not as an integral component of the CCR4–NOT complex. The model was proposed on the basis of the observation that mutations in the CCR4 LRR domain disrupt binding to both Bam and CAF1/POP2 (Fu et al. 2015). However, the mutated residues are located in the hydrophobic core of the LRR domain (Basquin et al. 2012) and most probably destabilize the domain fold. Therefore, it remains unclear whether free CCR4 or the assembled CCR4–NOT complex is required for Bam-mediated repression.

In the present study, we investigated the role of the CCR4–NOT complex in Bam-mediated mRNA regulation. We found that Bam promotes translational repression and degradation of bound mRNAs and that these activities depend on the N-terminal region of Bam, which does not contain the previously identified Bgcn-binding region and putative CCR4-binding site (Supplemental Fig. S1; Pan et al. 2014; Fu et al. 2015). We further

show that this N-terminal region contains a CAF40-binding motif (CBM) that interacts directly with CAF40. A crystal structure of the Bam CBM peptide bound to CAF40 reveals a binding mode similar to that observed for the *Dm* Roquin CBM (Sgromo et al. 2017). However, in contrast to *Dm* Roquin, which recruits the CCR4–NOT complex through multiple redundant binding sites, Bam relies entirely on the interaction with CAF40. Disruption of the Bam–CAF40 interaction also disrupts the interaction with CCR4 and the additional subunits of the CCR4–NOT complex and abolishes Bam activity. Thus, Bam recruits the assembled CCR4–NOT deadenylase complex through a direct interaction with CAF40 and this interaction is essential for Bam to repress bound mRNAs.

RESULTS

The Bam N-terminal region mediates translational repression and degradation of mRNA targets

Bam promotes stem cell differentiation by repressing the expression of specific mRNA targets through a mechanism that involves interaction with the CCR4 deadenylase (Fu et al. 2015). To obtain detailed mechanistic insights into this repressive mechanism and more precisely define the Bam sequences responsible for its repressive activity, we used a λN-based tethering assay in *Drosophila melanogaster* S2 cells (Behm-Ansmant et al. 2006). On the basis of sequence alignments, we designed Bam N- and C-terminal fragments for tethering assays (Fig. 1A; Supplemental Fig. S1; Supplemental Table S1). Full-length Bam and the Bam-N and Bam-C fragments were expressed with λN-HA tags that bind to a coexpressed firefly luciferase mRNA reporter containing five λN-binding sites (BoxB hairpins) in the 3′ UTR (F-Luc-5BoxB mRNA). An mRNA encoding *Renilla* luciferase (R-Luc) served as a transfection control.

λN-HA-tagged Bam decreased the F-Luc expression level to 10% relative to the λN-HA fusion protein, which was used as a negative control (Fig. 1B). The decrease in F-Luc activity was predominantly explained by a corresponding decrease in the mRNA abundance (Fig. 1B,C) and a shortening of the mRNA half-life (Supplemental Fig. S2A,B), thus indicating that Bam induces mRNA degradation in S2 cells. Furthermore, the Bam-N fragment retained the activity of the full-length protein, whereas the activity of the Bam-C fragment was strongly impaired (Fig. 1B,C). All proteins were expressed at comparable levels (Fig. 1D), and none of the proteins affected the expression of an F-Luc reporter lacking the BoxB hairpins (Supplemental Fig. S2C,D), thus confirming that Bam must bind to the mRNA to induce degradation.

To determine whether Bam might repress translation independently of mRNA degradation, we used an mRNA reporter containing a 3′ end generated by a self-cleaving hammerhead ribozyme (HhR) that consequently lacks a poly(A) tail (F-Luc-5BoxB-$A_{95}C_7$-HhR) (Zekri et al. 2013). Additionally, the reporter contains an internal, DNA-encoded, oligo(A) stretch of 95 nucleotides and a 3′ oligo(C) stretch of seven nucleotides upstream of the ribozyme cleavage site. This reporter is resistant to deadenylation and subsequent degradation and is efficiently translated in S2 cells (Zekri et al. 2013). Full-length Bam and the Bam-N fragment repressed the expression of this reporter in S2 cells (Fig. 1E, F). This repression occurred mainly at the translational level, because mRNA levels were not decreased to a similar extent as with the polyadenylated reporter. Together, our results indicated that Bam promotes the degradation of polyadenylated mRNAs and also represses translation independently of mRNA degradation when deadenylation is blocked. Furthermore, the Bam activity resides primarily in the Bam-N fragment, which does not contain the putative CCR4-binding region (Supplemental Fig. S1; Fu et al. 2015).

Bam directs bound mRNAs to the 5′-to-3′ decay pathway

We next investigated whether Bam elicits mRNA degradation via the 5′-to-3′ decay pathway. In this pathway, deadenylation is followed by decapping and 5′-to-3′ exonucleolytic degradation of the mRNA body. We therefore performed tethering assays in S2 cells depleted of the decapping enzyme DCP2 and overexpressing a catalytically inactive DCP2 mutant (DCP2 E361Q), which inhibits decapping in a dominant negative manner (Chang et al. 2014). In these cells, degradation of the F-Luc-5BoxB mRNA by tethered Bam or the Bam-N fragment was impaired (Fig. 2A). The F-Luc-5BoxB mRNA accumulated as a fast-migrating form corresponding to a deadenylated decay intermediate (A_0; Fig. 2B, lanes 5 and 6). Despite the restoration of mRNA levels, F-Luc activity was not restored (Supplemental Fig. S2E), most likely because deadenylated transcripts are translated less efficiently. The expression of the tethered proteins was not affected by coexpression of the DCP2 mutant (Fig. 2C). Together, these results indicated that Bam directs mRNA targets to the 5′-to-3′ decay pathway.

Bam recruits the CCR4–NOT complex to induce mRNA degradation

Our results indicated that Bam promotes deadenylation-dependent decapping. To determine whether Bam-mediated deadenylation requires the assembled CCR4–NOT complex or, alternatively, whether only the CCR4 subunit acts in isolation, as suggested previously (Fu et al. 2015), we disrupted CCR4–NOT complex assembly by depleting NOT1, the scaffold subunit of the complex (Wahle and Winkler 2013). NOT1 depletion partially suppressed degradation of F-Luc-5BoxB mRNA mediated by Bam and Bam-N (Fig. 2D,E; Supplemental Fig. S2F), thus suggesting that the assembled CCR4–NOT complex is required for Bam's repressive activity. Furthermore, NOT1 depletion also suppressed Bam-

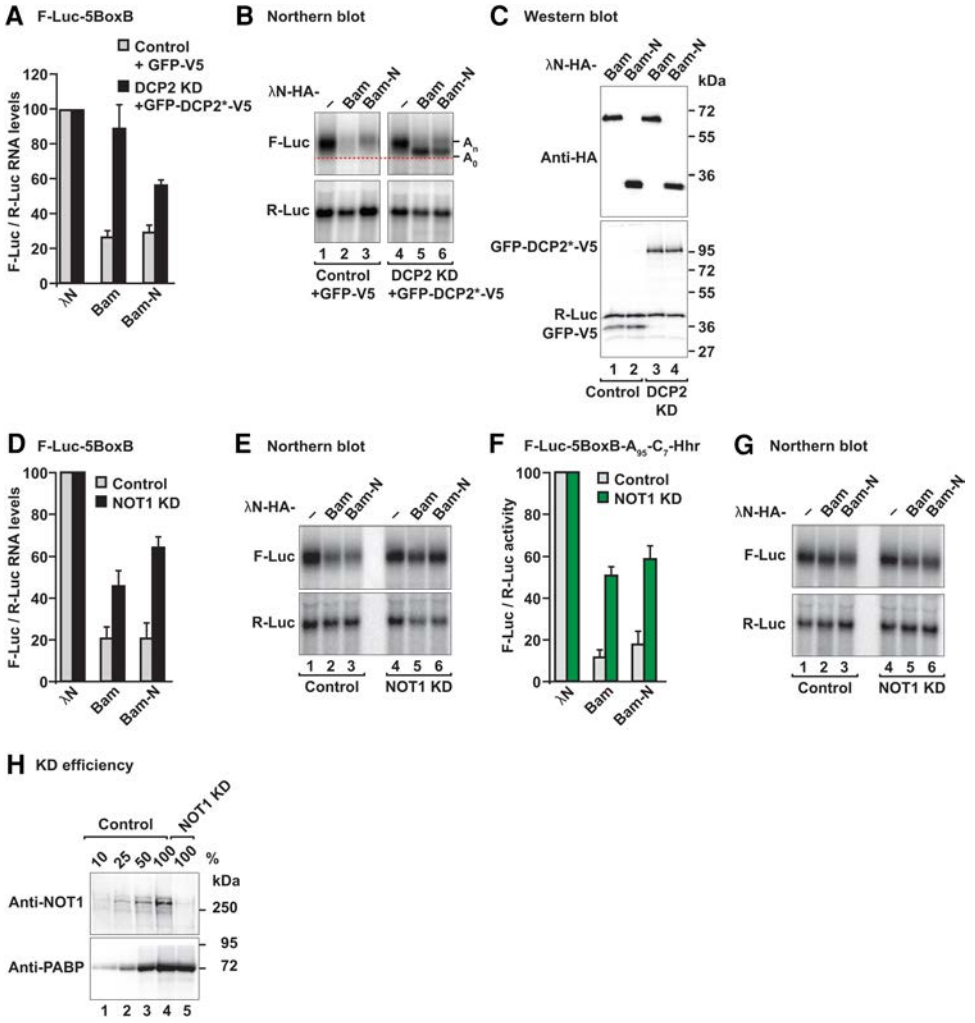


FIGURE 2. Bam degrades mRNAs through the 5′-to-3′ mRNA decay pathway. (*A*) Tethering assay using the F-Luc-5BoxB reporter was performed in control S2 cells or cells depleted of the decapping enzyme DCP2 (DCP2 KD). The transfection mixture included plasmids expressing either GFP-V5 or a GFP-tagged catalytically inactive DCP2 mutant (DCP2*; E361Q). The F-Luc-5BoxB mRNA levels were normalized to those of the R-Luc transfection control and set to 100% in control and knockdown cells expressing the λN-HA peptide. The gray bars represent the normalized F-Luc-5BoxB mRNA values in control cells expressing GFP-V5. The black bars represent the normalized F-Luc-5BoxB mRNA values in DCP2-depleted cells expressing GFP-DCP2*-V5. (*B*) Northern blot of representative RNA samples shown in *A*. The positions of the polyadenylated (A_n) and deadenylated (A_0, dashed red line) mRNA reporter are indicated on the *right* of the panel. (*C*) Western blot analysis showing equivalent expression of λN-HA-tagged proteins in the experiments shown in *A* and *B*. (KD) Knockdown. (*D*,*E*) Tethering assay using the F-Luc-5BoxB reporter in control S2 cells or in NOT1-depleted cells. Samples were analyzed as described in Figure 1B–D. (*F*,*G*) Tethering assay using the F-Luc-5BoxB-A_{95}-C_7-HhR reporter in control cells and in NOT1-depleted cells. Samples were analyzed as described in Figure 1B–D. In *A*, *D,* and *F*, bars represent mean values and error bars represent standard deviations from at least three independent experiments. (*H*) Western blot analysis showing the efficiency of NOT1 depletion in the experiments shown in *D*–*G*. Dilutions of control cell lysates were loaded in lanes *1*–*4*. PABP served as a loading control. Protein size markers (kDa) are shown on the *right* in each panel.

mediated translational repression of the reporter that was resistant to deadenylation and decay (F-Luc-5BoxB-$A_{95}C_7$-HhR; Fig. 2F,G; Supplemental Fig. S2G). Western blot analysis indicated that NOT1 levels were indeed decreased to <25% of the control levels in the knockdown cells (Supplemental Fig. S2H).

Because Bam activity depends on the integrity of the CCR4–NOT complex and it resides in the N-terminal fragment, which does not contain the putative CCR4-binding region (Fu et al. 2015), we re-examined Bam interactions with subunits of the CCR4–NOT complex. We expressed Bam with an HA tag in S2 cells and tested for interactions with GFP-tagged subunits of the CCR4–NOT complex in coimmunoprecipitation assays. Bam interacted with NOT1, NOT2, NOT3, CCR4, and CAF40 (Supplemental Fig. S3A–E), thus suggesting that it interacts with the assembled CCR4–NOT complex. All of these interactions were observed in the presence of RNaseA. Together, our results indicated that the CCR4–NOT complex is an important downstream effector of Bam-mediated mRNA regulation.

Bam interacts with the CAF40 subunit of the CCR4–NOT complex

To discriminate between direct and indirect interactions between Bam and the subunits of the CCR4–NOT complex, we performed pull-down assays in vitro, using purified recombinant proteins expressed in *Escherichia coli*. Because some *Dm* NOT1 domains are not expressed in a soluble form in bacteria, we first tested whether Bam could also interact with the human CCR4–NOT complex. To this end, we expressed Bam with a V5-SBP tag in human HEK293T cells and tested for interactions with endogenous subunits of the CCR4–NOT complex in pull-down assays. Bam pulled down all of the tested subunits of the endogenous CCR4–NOT complex (NOT1, NOT2, NOT3 and CAF40; Fig. 3A, lane 4) as well as HA-tagged CCR4 (Fig. 3B, lane 4) in the presence of RNaseA, thereby indicating that the Bam-binding surface on the CCR4–NOT complex is conserved.

This result allowed us to test for interactions with individual purified human CCR4–NOT subcomplexes in vitro, including the NOT1-10-11 module, the catalytic module comprising the NOT1 MIF4G domain bound to CAF1 and CCR4a, the NOT1 CN9BD domain bound to CAF40, a C-terminal connector domain of unknown function (CD), the NOT module comprising the NOT1 SHD and the C-terminal regions of NOT2 and NOT3, and an N-terminal coiled coil domain of NOT3 (Supplemental Fig. S3F). MBP-tagged Bam interacted only with the CN9BD–CAF40 module but not with any other subcomplex (Supplemental Fig. S3G, lane 25). The CN9BD–CAF40 module is highly conserved between *Hs* and *Dm* (CN9BD and CAF40 display 50% and 75% sequence identity, respectively). Accordingly, Bam also interacted with the *Dm* CN9BD–CAF40 module in pull-down assays (Fig. 3C, lane 12).

The CAF40-binding motif (CBM) is required for Bam repressive activity

To more precisely define the region of Bam that interacts with the CAF40

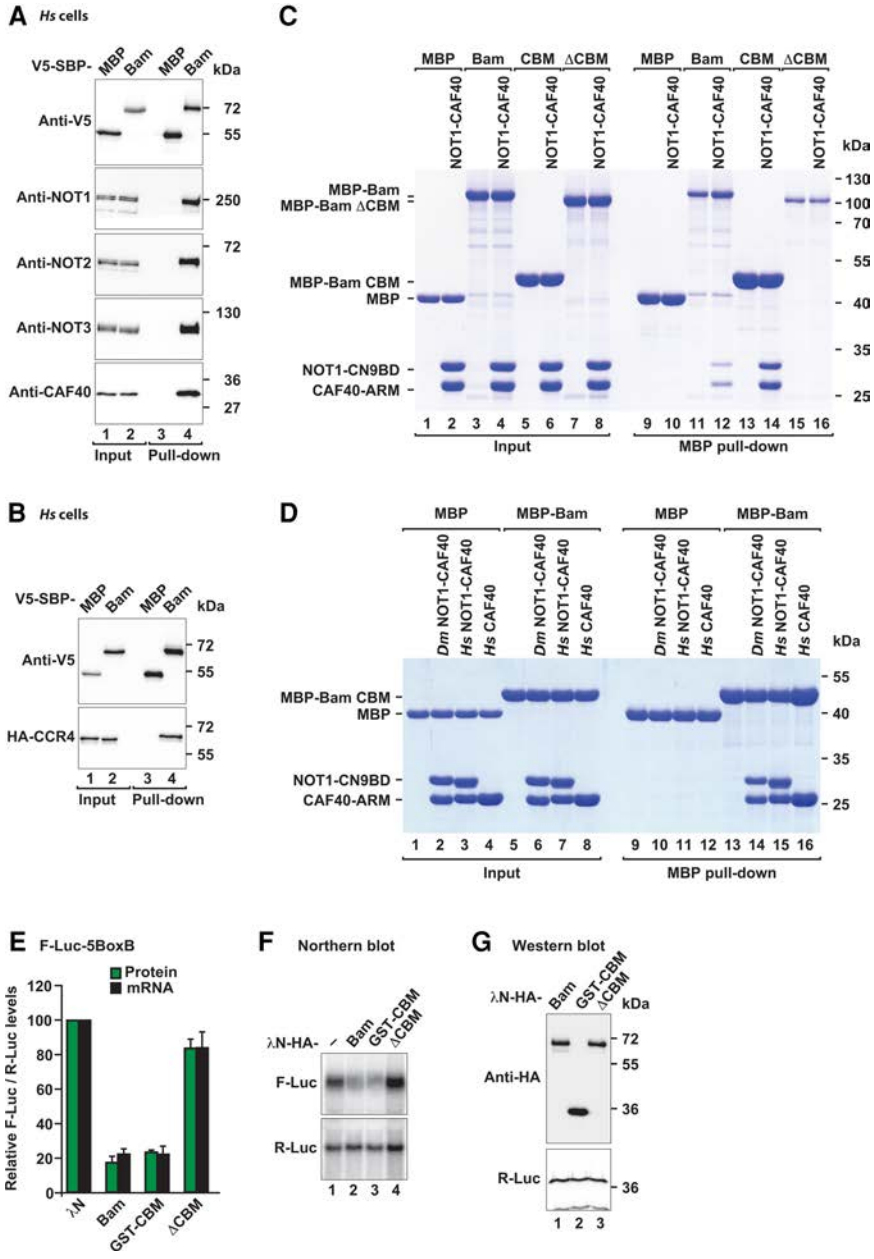


FIGURE 3. Bam binds directly to CAF40 by using an N-terminal CAF40-binding motif (CBM). (*A*) SBP pull-down assay in HEK293T cells expressing V5-SBP-tagged full-length Bam. V5-SBP-tagged MBP served as negative control. Input (1.5% for the V5-SBP tagged proteins and 1% for endogenous CCR4–NOT subunits) and bound fractions (10% for the V5-SBP tagged proteins and 30% for the CCR4–NOT subunits) were analyzed by western blotting using the indicated antibodies. (*B*) SBP pull-down assay in HEK293T cells expressing V5-SBP-tagged full-length Bam and HA-CCR4. V5-SBP-tagged MBP served as negative control. Samples were analyzed as described in *A*. (*C*) MBP pull-down assay testing the interaction of MBP-tagged full-length Bam, the CBM or Bam ΔCBM with the *Dm* CN9BD–CAF40 complex. MBP served as a negative control. The inputs (10%) and bound fractions (50%) were analyzed by SDS–PAGE and subsequent Coomassie staining. (*D*) MBP pull-down assay showing the interaction of MBP-tagged Bam CBM with the *Dm* and *Hs* CN9BD–CAF40 complex and *Hs* CAF40. Samples were analyzed as in *C*. (*E*) Tethering assay using the F-Luc-5BoxB reporter and λN-HA-tagged Bam (full-length or the indicated fragments) in S2 cells. The samples were analyzed as described in Figure 1B–D. The mean values ±SD from three independent experiments are shown. (*F*) Northern blot of representative RNA samples shown in *E*. (*G*) Western blot showing the equivalent expression of λN-HA-tagged proteins used in *E* and *F*. Protein size markers (kDa) are shown on the *right* in each panel.

module, we performed a series of pull-down assays using various Bam fragments, which led to the identification of a CAF40-binding motif (CBM, residues D13–E36) within the Bam N-terminal fragment. The Bam CBM was sufficient for binding to the *Dm* and human CN9BD–CAF40 modules in pull-down assays (Fig. 3C, lane 14 and Fig. 3D, lanes 14 and 15). Furthermore, the CBM interacted directly with the isolated *Hs* CAF40 subunit in the absence of the NOT1 CN9BD (Fig. 3D, lane 16). Importantly, deletion of the CBM in the context of full-length Bam abolished the interaction with the *Dm* CN9BD–CAF40 module in vitro (Fig. 3C, lane 16), thereby indicating that the CBM is the principal CAF40-binding site in Bam.

To determine the contribution of the CBM to Bam's repressive activity, we performed tethering assays in S2 cells. Remarkably, the CBM alone (fragment D13–E36 fused to GST) was sufficient to induce the repression and degradation of the F-Luc-5BoxB mRNA to a similar extent as full-length Bam (Fig. 3E,F). Furthermore, deletion of the CBM was sufficient to abolish the repressive activity of Bam in tethering assays (Fig. 3E,F). All proteins were expressed at comparable levels (Fig. 3G) and did not affect the expression of an F-Luc reporter lacking the BoxB hairpins (Supplemental Fig. S4A, B). We therefore concluded that the CBM is essential for Bam's repressive activity.

Crystal structure of the Bam CBM bound to CAF40

To elucidate the molecular principles underlying the interaction of Bam with the CAF40 module, we sought to determine the crystal structure of the CBM peptide (residues D13–E36) bound to the *Dm* and *Hs* CAF40 armadillo (ARM) domain (*Dm* CAF40 E25–G291 and *Hs* CAF40 R19–E285) as well as to the CAF40 modules containing the NOT1 CN9BD (residues *Dm* NOT1 Y1468–T1719 and *Hs* NOT1 V1351–L1588). However, only the complexes containing the human proteins yielded well-diffracting crystals. We solved the structures of the Bam CBM peptide bound to CAF40 and to the CN9BD–CAF40 complex and refined them to 3.0 Å and 2.7 Å resolution, respectively (Table 1; Fig. 4A–C).

The asymmetric unit of the CAF40-CBM crystal contained four complexes that were highly similar to each other

TABLE 1. Data collection and refinement statistics

	CAF40–Bam	NOT1–CAF40–Bam
Space group	P 2_1 2_1 2	P 3_2 2 1
Unit cell		
Dimensions *a*, *b*, *c* (Å)	105.6, 200.9, 59.6	106.6, 106.6, 263.4
Angles α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 120.0
Data collection[a]		
Wavelength (Å)	1.0396	1.0000
Resolution range (Å)	50.–3.0 (3.08–3.00)	50–2.7 (2.77–2.70)
R_{sym} (%)	9.5 (100.8)	11.4 (222.4)
Completeness (%)	99.5 (98.8)	99.9 (99.5)
Mean ($I/\sigma I$)	13.2 (1.7)	15.5 (1.2)
Unique reflections	26,082 (1852)	48,613 (3529)
Multiplicity	5.5 (5.7)	11.0 (10.7)
CC(1/2)	1.00 (0.65)	1.00 (0.70)
Refinement		
R_{work} (%)	21.6	20.9
R_{free} (%)	26.8	23.7
Number of atoms		
All atoms	9358	8481
Protein	9352	8424
Ordered solvent	6	57
Average *B* factor (Å^2)		
All atoms	100.8	97.9
Protein	100.7	97.5
Ordered solvent	103.7	149.3
Ramachandran plot		
Favored regions (%)	96.4	98.9
Disallowed regions (%)	0.2	0.0
RMSD from ideal geometry		
Bond lengths (Å)	0.010	0.002
Bond angles (°)	1.080	0.437

[a]Values in parentheses are for highest-resolution shell.

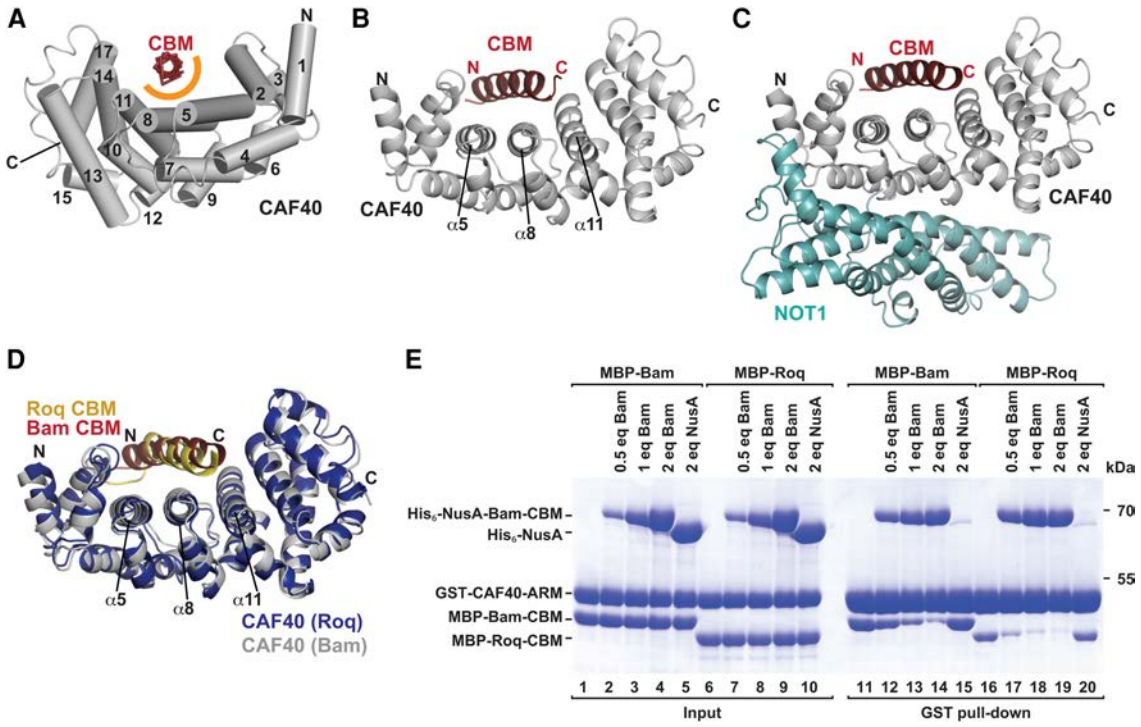


FIGURE 4. Structure of the Bam CBM bound to CAF40 and to the NOT1 CN9BD–CAF40 module. (*A*) The Bam CBM peptide (red, backbone shown in ribbon representation) bound to *Hs* CAF40 (gray). CAF40 helices are depicted as tubes and numbered in black. The orange semicircle marks the predominantly hydrophobic interface between the CBM peptide and CAF40. (*B*) Cartoon representation of the Bam CBM peptide bound to *Hs* CAF40. Selected CAF40 secondary structure elements are labeled in black. (*C*) Structure of the CBM peptide bound to the NOT1 CN9BD–CAF40 complex. (*D*) Superposition of the CAF40–Bam CBM structure with the structure of CAF40 bound to the Roq CBM (PDB 5LSW; Sgromo et al. 2017). The Roq CBM is shown in yellow and CAF40 from the Roq complex in blue. (*E*) In vitro competition assay. GST-tagged *Hs* CAF40 was incubated with equimolar amounts of MBP-tagged Bam or Roq CBMs and increasing amounts of His_6-NusA-tagged Bam CBM. His_6-NusA was used as a negative control. Proteins bound to GST-CAF40 were pulled down and analyzed by SDS-PAGE and subsequent Coomassie staining. Molar equivalents (eq) are relative to GST-CAF40.

(RMSDs between 0.31 and 0.75 Å; over 237–278 Cα atoms) and that were arranged as two pairs of dimers (Supplemental Fig. S5A,B). The dimer interface corresponds to the one previously observed in the structure of free CAF40 (Supplemental Fig. S5C; Garces et al. 2007). The asymmetric unit of the CN9BD–CAF40-CBM crystals contained two almost identical complexes (RMSD of 0.28 Å over 457 Cα atoms; Supplemental Fig. S5D,E). In all complexes, the interaction of the CBM peptide with the CAF40 concave surface was found to be almost identical (Fig. 4B,C; Supplemental Fig. S5F), and the CBM does not contact the NOT1 CN9BD (Fig. 4A–C), thus confirming that the CBM interacts exclusively with CAF40. Superposition of the CAF40 dimer bound to the CBM with the previously determined ligand-free CAF40 dimer (Supplemental Fig. S5C; RMSD of 0.90 Å over 509 Cα atoms; Garces et al. 2007) or with CAF40 bound to the NOT1 CN9BD (Supplemental Fig. S5G; RMSD of 0.94 Å over 416 Cα atoms; Chen et al. 2014a), indicated that binding of the CBM peptide does not induce any major conformational changes in the CAF40 ARM domain.

The CBM peptide folds into an amphipathic α-helix that is bound centrally across the concave surface of the crescent-shaped CAF40 ARM domain, which consists of 17 α-helices arranged into six armadillo (ARM) repeats (Fig. 4A–C; Supplemental Fig. S6A,B; Garces et al. 2007; Chen et al. 2014a; Mathys et al. 2014). The α-helix binds to a conserved hydrophobic patch close to the previously proposed nucleic acid-binding groove (Garces et al. 2007). In the structure of the Bam CBM bound to the CAF40 module, the NOT1 CN9BD binds to the convex surface of CAF40 and prevents CAF40 dimerization, as previously observed (Chen et al. 2014a; Mathys et al. 2014). Importantly however, the NOT1 CN9BD does not interfere with Bam CBM binding on the concave surface of CAF40, thus indicating that Bam can interact with CAF40 also in the context of the fully assembled CCR4–NOT complex (Fig. 4C).

The Bam CBM competes with the Roquin CBM for binding to CAF40

Remarkably, the Bam CBM occupies the same binding surface as the previously described CBM of *Dm* Roquin (Roq) (Sgromo et al. 2017) and binds CAF40 in a similar manner (Fig. 4D). The two CBM peptides fold into amphipathic helices that bind via their hydrophobic sides along a groove on the concave face of CAF40. Consequently, the two peptides

cannot bind CAF40 simultaneously and compete for binding to CAF40 when tested in vitro in a competition assay. In this assay, GST-tagged CAF40 was incubated with equimolar amounts of MBP-tagged Bam or Roq CBMs and increasing concentrations of His_6-NusA-tagged Bam CBM. The peptides bound to CAF40 were pulled down by using glutathione-agarose beads. Increasing concentrations of the His_6-NusA-Bam CBM competed with the two MBP-tagged CBMs for CAF40 binding (Fig. 4E). Interestingly, the Roquin CBM was competed out more efficiently than the Bam CBM (Fig. 4E, e.g., lane 12 versus 17), thus suggesting that Bam has a competitive advantage.

To obtain information on the affinities of the CBM peptides for CAF40, we performed isothermal titration calorimetry (ITC) experiments. The Bam CBM bound to the *Dm* CN9BD–CAF40 complex with a dissociation constant (K_D) in the nanomolar range (183 ± 44 nM; Supplemental Fig. S7A). In contrast, the affinity of the Roq CBM was too low to be measured by ITC (i.e., the necessary peptide concentrations for measurement could not be reached), thereby explaining why the Roq CBM competed rather poorly with the Bam CBM.

The Bam and Roquin CBMs use similar binding modes

The Bam CBM forms a single amphipathic α-helix extending through residues D13–E33 and the hydrophobic face of this helix binds in a groove formed by helices α5, α8 and α11 on the CAF40 concave surface (Fig. 5A,B). The interaction buries a total surface area of 1638 $Å^2$ and involves the side chains of Bam residues L17, F21, M24, L28, M31 and V32, which engage the hydrophobic CAF40-binding surface consisting of residue A84 in helix α5, residues R130, Y134, L137, T138, G141 and G144 in helix α8, and residues L177, T180, V181 and F184 in helix α11 on the CAF40 side (Fig. 5A,B). In addition, the N and C termini of the CBM helix contact the CAF40 surface through hydrogen bonds between Bam N20 and CAF40 N88, and Bam E33 and CAF40 K230, respectively (Fig. 5B). However, these interactions were not observed in all six complexes, thus indicating some degree of flexibility of the helix ends.

In the Roq CBM, the N-terminal portion (residues E790–M797) is no longer α-helical, owing to the insertion of a glycine (G796), which is conserved among Roq proteins from different *Drosophila* species (Fig. 5C,D). Instead, the residues form an extended "hook" that is stabilized by internal hydrogen bonds. In contrast, the Bam N-terminal residues (D13–N20) extend the amphipathic α-helix by another two turns. Despite this structural difference, Bam residue L17 engages the same binding pocket as Roq residue I793. Thus, critical contacts are preserved in both peptides despite the fact that Roq is no longer helical (Fig. 5C,D). Overall, the all α-helical conformation of the Bam CBM is likely to be more stable on its own than the more extended conformation of the Roq CBM, which probably does not form in the absence of CAF40. The resulting difference in the binding entropy could contribute to the higher affinity of the Bam CBM for CAF40 and to its competitive advantage over the Roquin CBM. Alternatively, differences in the hydrophobic interface residues may also potentially explain the observed differences in affinity and competition between the two CBMs, e.g., the side chain of residue F21 in the center of the Bam CBM establishes a more extensive network of hydrophobic interactions along the interface than the side chain of residue M798 at the same structural position in the Roq CBM.

The interaction of Bam with CAF40 is required for mRNA repression

To assess the importance of the interactions observed in the crystal structure, we substituted Bam residues L17 or M24 with glutamic acid. These substitutions abolished the interaction of the MBP-tagged Bam with the *Dm* CAF40 module in pull-down assays (Fig. 5E; Supplemental Table S1), thus indicating that the CBM is the only CAF40-binding site in Bam. We also analyzed the effects of amino acid substitutions in the CAF40 interface on complex formation. A single V186E substitution or the double Y139D, G146W substitution (2xMut) in *Dm* CAF40 (corresponding to *Hs* CAF40 residues V181, Y134 and G141) were sufficient to disrupt the interaction with Bam in vitro (Fig. 5F; Supplemental Table S1). The equivalent substitutions in *Hs* CAF40 were also sufficient to disrupt binding to the Roq CBM (Sgromo et al. 2017), thus further confirming the similarity in the CBM-binding modes.

Next, we assessed the relevance of the interface in S2 cells. The single amino acid substitution in *Dm* CAF40 (V186E) was sufficient to abolish binding to full-length Bam in cell lysates (Supplemental Fig. S3E, lane 6). Conversely, substitutions of CBM residues (4xMut, Supplemental Table S1) in the context of full-length Bam abolished binding to *Dm* CAF40 (Supplemental Fig. S7B).

To assess the functional relevance of the CAF40-Bam interaction in mRNA target repression, we performed tethering assays in S2 and human cells. Single amino acid substitutions in the Bam CBM abolished Bam activity in tethering assays in S2 cells (Fig. 6A,B) a result indicating that the CBM provides a major contribution to Bam's repressive activity. All mutants were expressed at comparable levels (Fig. 6C) and did not affect the expression of a F-Luc mRNA lacking the BoxB hairpins (Supplemental Fig. S7C,D).

In human cells, we tethered MS2-HA-tagged full-length Bam (wild-type or the 4xMut) to a β-globin reporter containing six binding sites for the MS2 protein in the 3′ UTR. As observed in *Dm* cells, wild-type Bam caused degradation of the β-globin-6xMS2bs reporter, whereas the Bam 4xMut was inactive (Supplemental Fig. S7E–G). Furthermore, the CBM alone fused to MBP was as active as full-length Bam (Supplemental Fig. S7E–G). Thus, Bam depends on the integrity of the CBM to repress mRNA expression both in human and S2 cells.

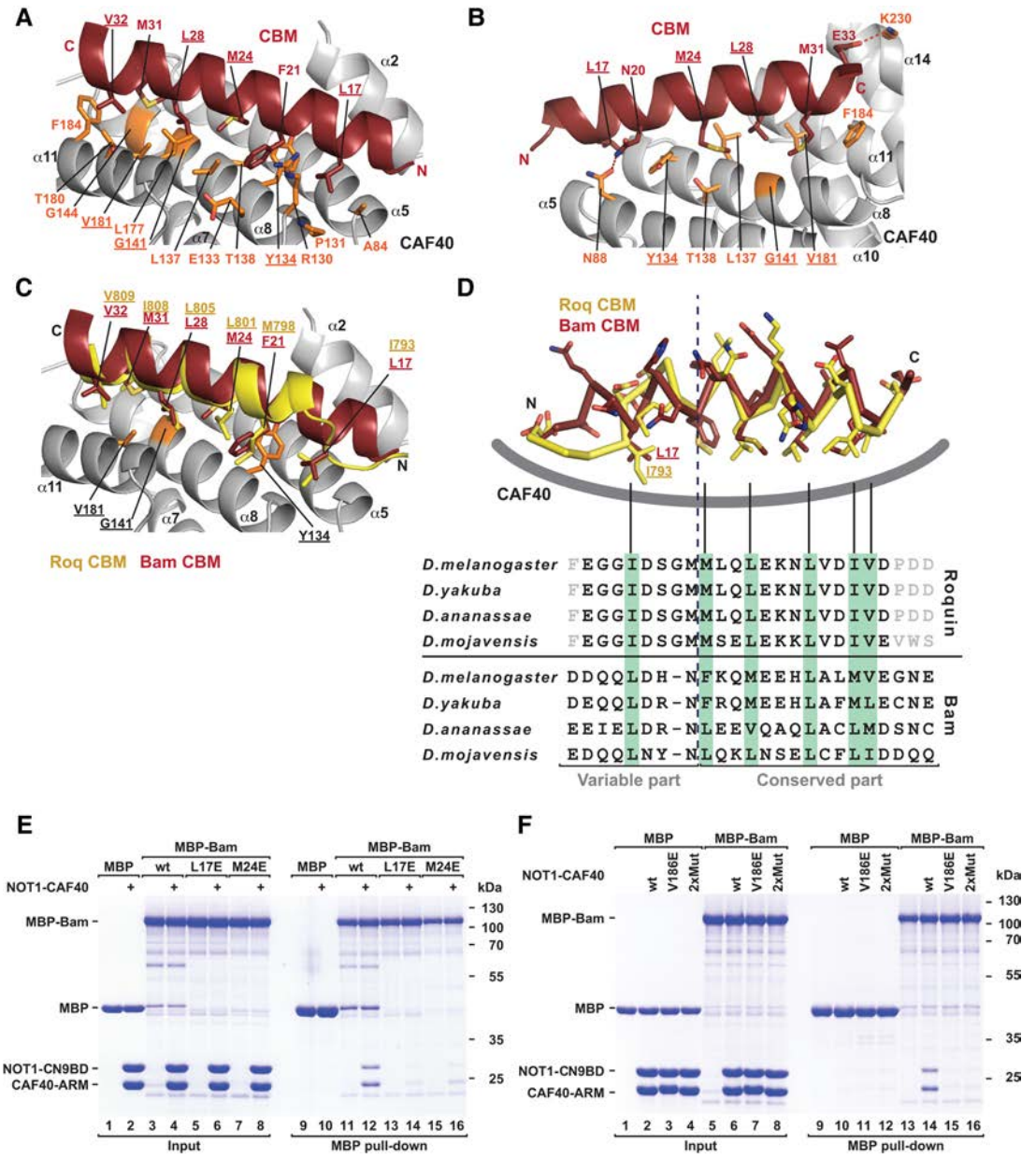


FIGURE 5. The Bam and Roq CBMs use a similar CAF40-binding mode. (*A*,*B*) Close-up views of the CAF40-Bam CBM-binding interface in two orientations. Selected residues of CAF40 and Bam are shown as orange and red sticks, respectively. Hydrogen bonds are indicated by red dashed lines. Residues mutated in this study are underlined. (*C*) Close-up view of the structural superposition of the CAF40-Bam CBM structure with the structure of the Roq CBM bound to CAF40. The Bam and Roq CBMs are shown in red and yellow, respectively. (*D*) (*Upper* panel) Superposition of the Bam and Roq CBM peptides bound to CAF40. The backbones are shown in ribbon representation, and side chains are shown as sticks. CAF40 is indicated as a thick gray line. (*Lower* panel) Sequence alignment of the Bam and Roq CBMs from the indicated *Drosophila* species. Hydrophobic residues interacting with CAF40 are highlighted by a light green background. Gray letters indicate residues that were not included in the crystallization setup. (*E*) MBP pull-down assay testing the interaction of MBP-tagged Bam (wild-type or mutants L17E and M24E) with the *Dm* NOT1-CN9BD–CAF40 complex. MBP served as a negative control. (*F*) MBP pull-down assay testing the interaction of MBP-tagged Bam with *Dm* NOT1-CN9BD–CAF40 complex (containing CAF40 wild-type or the indicated mutants). MBP served as a negative control.

Bam interaction with CCR4 is indirect and mediated by CAF40

In the pull-down assays using recombinant proteins, we did not observe a direct interaction between Bam and the catalytic module (containing *Hs* CCR4a, which is 57% identical to *Dm* CCR4; Supplemental Fig. S3G). Furthermore, Bam did not competitively displace the CAF1-NOT1 subcomplex from CCR4a (Supplemental Fig. S3G, lane 24), as has previously been suggested (Fu et al. 2015). To determine whether Bam interaction with CCR4 was direct or mediated by CAF40, we used CRISPR-Cas9 gene editing to generate a

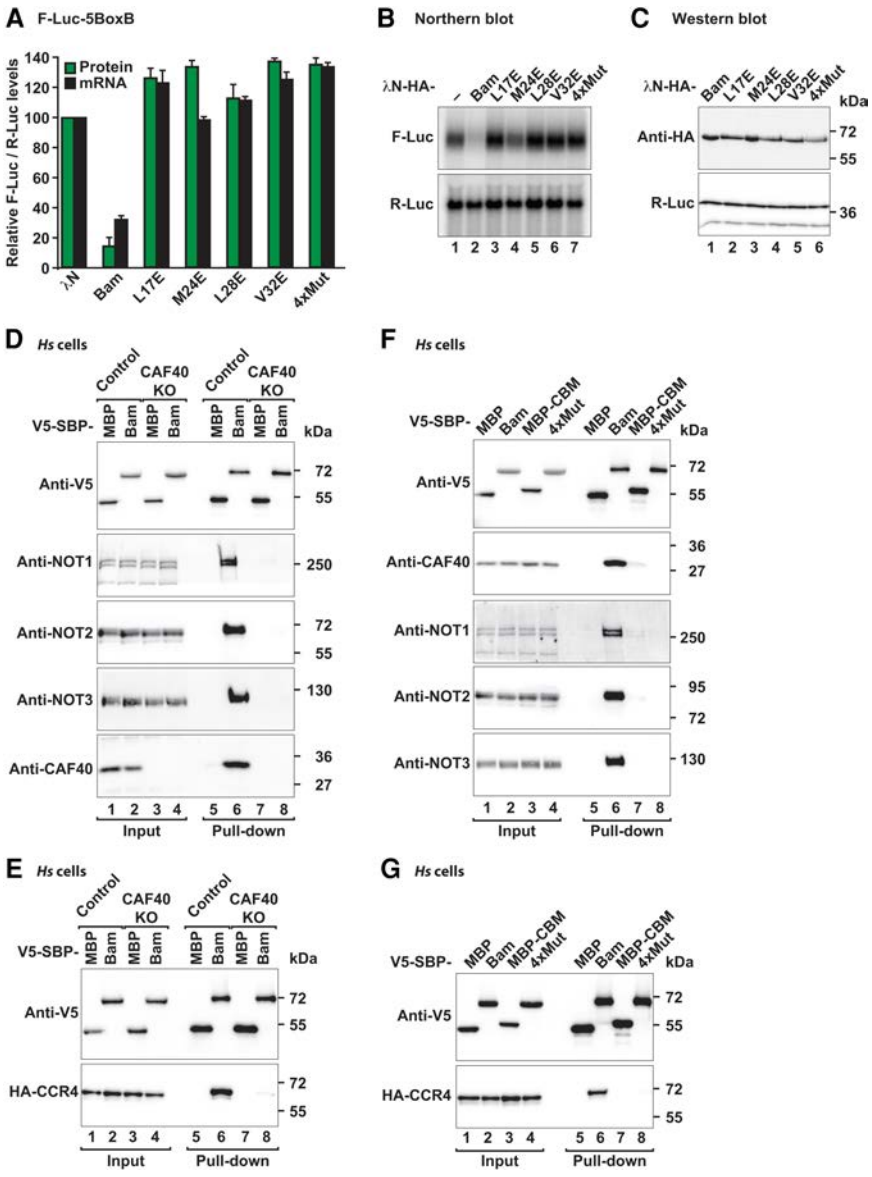


FIGURE 6. The CBM is necessary for Bam-mediated mRNA repression. (*A*) Tethering assay using the F-Luc-5BoxB reporter and λN-HA-tagged Bam (wild-type or the indicated mutants) in S2 cells. The samples were analyzed as described in Figure 1B. (*B*) Northern blot of representative RNA samples shown in *A*. (*C*) Western blot showing the equivalent expression of λN-HA-tagged proteins used in *A* and *B*. (*D*) SBP pull-down assay in control and CAF40-null HEK293T cells expressing V5-SBP-tagged full-length Bam. V5-SBP-tagged MBP served as a negative control. Input (1.5% for the V5-SBP tagged proteins and 1% for endogenous CCR4–NOT subunits) and bound fractions (10% for the V5-SBP tagged proteins and 30% for the CCR4–NOT subunits) were analyzed by western blotting using the indicated antibodies. (KO) Knockout. (*E*) SBP pull-down assay in control and CAF40-null HEK293T cells expressing V5-SBP-tagged full-length Bam and HA-CCR4. Samples were analyzed as in *D*. (*F*) SBP pull-down assay in HEK293T cells expressing V5-SBP-tagged full-length Bam or the 4xMut. V5-SBP-tagged MBP served as a negative control. Input (1.5% for the V5-SBP tagged proteins and 1% for CCR4–NOT subunits) and the bound fractions (10% for the V5-SBP tagged proteins and 30% for CCR4–NOT subunits) were analyzed by western blotting using the indicated antibodies. (*G*) SBP pull-down assay in HEK293T cells expressing V5-SBP-tagged full-length Bam or the 4xMut and HA-tagged CCR4.

CAF40-null HEK293T human cell line in which the CAF40 levels were decreased below detectable levels (Fig. 6D, lanes 3 and 4 versus 1 and 2 and Supplemental Fig. S7H), whereas the expression of the additional subunits of the CCR4–NOT complex was not affected (Fig. 6D, lane 3 and 4 versus 1 and 2). In this cell line, Bam did not interact with endogenous NOT1, NOT2 and NOT3 (Fig. 6D, lane 8 versus 6) or with HA-tagged CCR4 (Fig. 6E, lane 8 versus 6), thus indicating that the interaction of Bam with the subunits of the CCR4–NOT complex is indeed mediated by CAF40. Similarly, the combined quadruple substitutions in the Bam CBM (4xMut) abrogated the interaction with the endogenous subunits of the CCR4–NOT complex in human cells (Fig. 6F, lane 8 versus 6) as well as the interaction with HA-tagged CCR4 (Fig. 6G, lane 8 versus 6). Similarly, the Bam 4xMut did not interact with CCR4 or NOT2 in S2 cells (Supplemental Fig. S7I,J). Together, our results indicated that the previously reported interaction of Bam with CCR4 (Fu et al. 2015), is most likely indirect and mediated by CAF40 in the context of the fully assembled CCR4–NOT complex.

CAF40 is the only Bam-binding site within the CCR4–NOT complex

To further validate the relevance of Bam interaction with CAF40 for the recruitment of the CCR4–NOT complex, we performed tethering assays in *Dm* S2 cells overexpressing CAF40 wild-type or the CAF40 V186E mutant, which does not interact with Bam and was thus expected to suppress Bam activity in a dominant negative manner. Accordingly, Bam activity in tethering assays was suppressed in cells overexpressing the CAF40 V186E mutant but not when CAF40 wild-type was overexpressed (Fig. 7A,B). For a control, we tethered *Dm* Roq, which in addition to the CBM contains additional binding sites for the CCR4–NOT complex (Sgromo et al. 2017). Consequently, Roq activity was only slightly affected in cells overexpressing the CAF40 mutant (Fig. 7A,B, lane 9). Overexpression of CAF40 did not affect the Bam and Roq expression levels (Fig. 7C). The differential effect of the CAF40 mutant on Bam and Roq activities further confirmed that Bam, in contrast to Roq, depends entirely on its interaction with CAF40 for

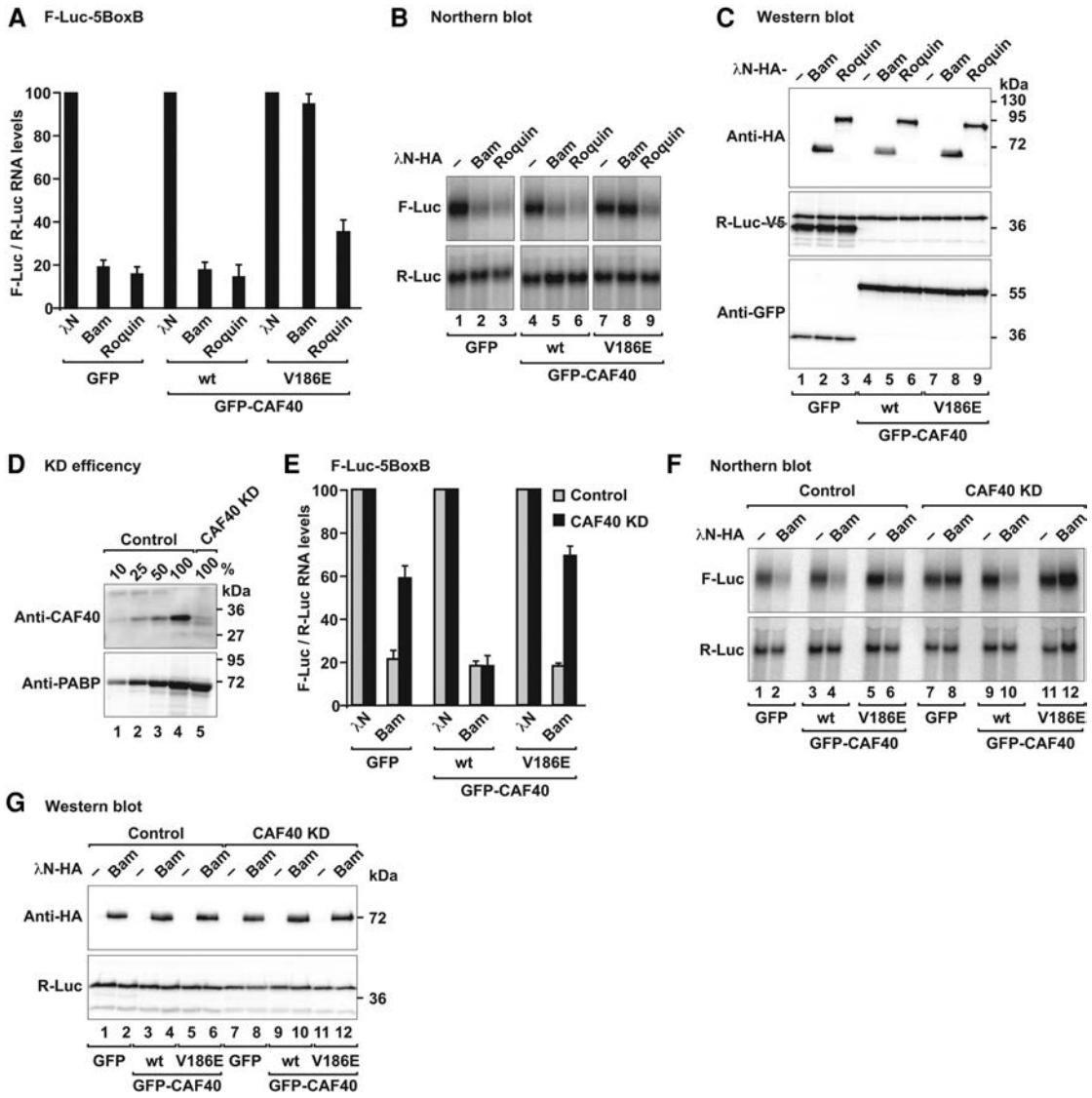


FIGURE 7. Bam depends on CCR4–NOT complex recruitment to induce mRNA decay. (*A*) Tethering assay using the F-Luc-5BoxB reporter and λN-HA-tagged Bam and Roq in S2 cells. The transfection mixtures also contained plasmids for expression of GFP (control) or GFP-CAF40 (wild-type or the V186E mutant) as indicated. The samples were analyzed as described in Figure 1B. (*B*) Northern blot of representative RNA samples shown in *A*. (*C*) Western blot showing the equivalent expression of the λN-HA-tagged proteins in cells expressing GFP or GFP-CAF40 (either wild-type or the V186E mutant) used in *A* and *B*. (*D*) Western blot showing the efficiency of the CAF40 depletion in *Dm* S2 cells. Dilutions of control cell lysates were loaded in lanes *1–4* to estimate the efficacy of the depletion. PABP served as a loading control. (KD) Knockdown. (*E*) Complementation assay using the F-Luc-5BoxB reporter and λN-HA-tagged Bam in S2 cells depleted of CAF40 (CAF40 KD) or in control cells (control). Samples were analyzed as described in Figure 1B. (*F*) Northern blot of representative RNA samples shown in *E*. (*G*) Western blot showing the equivalent expression of the λN-HA-tagged Bam constructs used in *E* and *F*.

repression, whereas Roq can recruit the CCR4–NOT complex through additional binding sites.

In an independent approach, we tethered Bam in S2 cells depleted of CAF40 in which CAF40 levels were decreased to ~10% of the control levels (Fig. 7D). CAF40 depletion partially suppressed Bam activity in tethering assays in S2 cells (Fig. 7E,F). The Bam-mediated repression was restored by transient expression of wild-type CAF40 but not by expression of the CAF40 V186E mutant, which does not interact with Bam (Fig. 7E,F), despite comparable expression levels (Fig. 7G). Thus, Bam requires interactions with CAF40 for full repressive activity.

DISCUSSION

In this study, we showed that Bam represses the translation and promotes the degradation of bound mRNAs by directly recruiting the CCR4–NOT deadenylase complex through an interaction with CAF40. This interaction is mediated by a short CAF40-binding motif (CBM) that is necessary and

sufficient for Bam's repressive activity. We further elucidated the structural basis of the interaction of the Bam CBM with CAF40 and identified the concave surface of CAF40 as a binding site for amphipathic helices in RNA-associated proteins that recruit the CCR4–NOT complex.

CCR4–NOT complex recruitment is required for Bam repressive activity

The recruitment of the CCR4–NOT complex via the CAF40-CBM interaction is required for Bam to repress the translation of mRNA targets. Unlike other proteins, such as GW182, TTP, Roq and *Dm* Nanos, that use multiple redundant motifs to recruit the CCR4–NOT complex (Fabian et al. 2013; Chen et al. 2014a; Mathys et al. 2014; Raisch et al. 2016; Sgromo et al. 2017), Bam depends entirely on the interaction between the CBM and CAF40. Indeed, single point mutations in the CBM that abolished the interaction with CAF40 also disrupted the interaction with CCR4 and additional subunits of the CCR4–NOT complex and abrogated Bam's repressive activity. Similar results were obtained in cells depleted of CAF40, thus indicating that the previously reported interaction between Bam and CCR4 is indirect and is mediated by CAF40 in the context of the assembled CCR4–NOT complex. These results also indicated that the CCR4–NOT complex is the main downstream effector complex in Bam-mediated mRNA regulation.

CAF40 serves as a binding platform of the CCR4–NOT complex

Bam adds to the growing number of examples of RNA-associated proteins that directly recruit the CCR4–NOT complex via short linear motifs to down-regulate mRNA targets. To date, the motifs that have been characterized have been found to bind non-overlapping surfaces on the CCR4–NOT complex. For example, vertebrate and *Dm* Nanos and *Saccharomyces cerevisiae* (*Sc*) NOT4 bind to non-overlapping surfaces on the NOT module (Bhandari et al. 2014; Bhaskar et al. 2015; Raisch et al. 2016). The CAF40 subunit also provides interaction sites for RNA-associated proteins within the CCR4–NOT complex. The convex surface of the CAF40 armadillo-repeat domain features two tryptophan-binding sites that are used by proteins of the GW182 protein family, which recruit the CCR4–NOT complex to miRNA targets (Chen et al. 2014a; Mathys et al. 2014). The CAF40 concave surface provides binding sites for the CBM in the *Dm* Roq protein (Sgromo et al. 2017). Here, we found that this surface also binds to the Bam CBM, thus indicating that Bam and Roq binding to CAF40 is mutually exclusive. However, the Bam and Roq proteins share no apparent sequence similarity and thus their CBMs may have evolved independently to bind to the same surface of CAF40. The functional relevance of this competitive binding remains unclear, because it is not known whether the two proteins are expressed in the same cell type under the same physiological conditions.

The high conservation of CAF40 (75% sequence identity between the *Hs* and *Dm* proteins, and 57% identity between *Hs* and *Sc*), particularly of the CBM-binding surface, suggests the existence of additional CBM-containing proteins in eukaryotes. Through an in silico search using a consensus pattern derived from the Bam and *Dm* Roq CBM sequences, we could indeed identify several potential CBMs in other proteins (Supplemental Fig. S8). However, none of the tested sequences interacted with *Hs* CAF40 in vitro in pull-down assays (data not shown), thus indicating that the tested fragments are not bona fide CBMs and that the rules guiding CAF40 binding are still incompletely understood. From what we know, it is possible and quite likely indeed that if CBMs exist in other proteins, they do not share an evolutionary origin with Bam and Roq and therefore also have no phylogenetic sequence conservation. Indeed, sequence searches conducted with either Bam or Roq did not identify the respective other protein as a CBM-containing protein.

CCR4–NOT complex recruitment is a recurring mechanism for targeted repression of gene expression

With the expanding repertoire of RNA-binding proteins that are known to interact with the CCR4–NOT complex, some underlying principles of recruitment are emerging. First, many RNA-associated proteins use extended peptide motifs embedded in unstructured regions for binding to CCR4–NOT. Interactions of such short linear motifs (SLiMs) are generally characterized by high specificity and at the same time relatively low individual affinity (Tompa 2012; Van Roey et al. 2014). This aspect is important in regulatory complexes such as the CCR4–NOT complex, because the complex must be recruited in a highly specific manner and need to be released again after exerting its specific function. Additionally, these motifs usually show high evolutionary plasticity (Tompa 2012; Van Roey et al. 2014) and are not conserved in orthologous proteins accross species.

Another common theme is that RNA binding is often spatially separated from CCR4–NOT complex recruitment. In many cases including Nanos and Roq, RNA binding is mediated by highly conserved RNA-binding domains, whereas CCR4–NOT complex recruitment is mediated by SLiMs in long unstructured regions of up to several hundred amino acids in length. In other cases, RNA binding and CCR4–NOT recruitment are associated with different polypeptides. For example, in the miRNA-induced silencing complexes (miRISCs), RNA binding is achieved by Argonaute proteins (AGOs), whereas CCR4–NOT complex recruitment is mediated by the GW182 proteins that act as adaptor molecules downstream from AGOs (Jonas and Izaurralde 2015). In the case of Bam, it is unknown whether the RNA-binding activity resides in the Bam protein itself or whether additional factors mediate mRNA binding.

Finally, it is interesting to note that some proteins such as Bam and vertebrate Nanos (this study; Bhandari et al. 2014), use a single motif with relatively high affinity to interact with the CCR4–NOT complex, whereas others such as *Dm* Roq and the GW182 proteins, use avidity effects in a distributive binding mode involving multiple lower-affinity motifs in disordered protein regions for recruitment (Chen et al. 2014a; Mathys et al. 2014; Sgromo et al. 2017). The highly diverse sequence motifs bind to several structured surfaces on the complex. Nevertheless, independently of the mode of interaction, the recruitment of the CCR4–NOT complex by diverse RNA-binding proteins results in a common functional outcome: the repression of the mRNA target through deadenylation-dependent and independent mechanisms and, in cellular contexts in which deadenylation is coupled to decapping, the degradation of the mRNA through the 5′-to-3′ mRNA decay pathway. Thus, the CCR4–NOT complex, through its ability to provide binding sites for diverse sequence motifs, is a major downstream effector complex in post-transcriptional mRNA regulation.

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 mutants used in this study were generated by site-directed mutagenesis using a QuikChange mutagenesis kit (Stratagene). All constructs and mutations were confirmed by sequencing.

Coimmunoprecipitation and SBP pull-down assays

All coimmunoprecipitation and SBP pull-down assays in S2 and HEK293T cell lysates were performed in the presence of RNaseA as previously described (Sgromo et al. 2017). All western blots were developed using an ECL western blotting detection system (GE Healthcare). The antibodies used in this study are listed in Supplemental Table S2. A detailed description of these assays is included in the Supplemental Material.

Tethering and complementation assays

Knockdown of DCP2, NOT1 and CAF40 in S2 cells using dsRNA was performed as previously described (Behm-Ansmant et al. 2006). For the λN-tethering assays in *Dm* S2 cells, 2.5×10^6 cells per well were seeded in six-well plates and transfected using Effectene (Qiagen). The transfection mixtures contained the following plasmids: 0.1 µg of Firefly luciferase reporters (F-Luc-5BoxB, F-Luc-V5 or F-Luc-5BoxB-$A_{95}C_7$-HhR), 0.4 µg of the R-Luc transfection control and various amounts of plasmids expressing λN-HA-tagged full-length Bam or Bam fragments (0.05 µg for wild-type or mutant full-length Bam, 0.02 µg for Bam-N, 0.1 µg for Bam-C, 0.1 µg GST-CBM and 0.05 µg of Bam ΔCBM). Cells were harvested 3 days after transfection.

In the experiment shown in Figure 2A, control and DCP2 knockdown cells were additionally transfected with plasmids expressing GFP-V5 (0.08 µg) and GFP-DCP2*-V5 mutant (E361Q; 1 µg), respectively. In the experiment shown in Figure 7A and B, cells were also transfected with plasmids expressing GFP (0.05 µg) or GFP-tagged CAF40 (1.5 µg) either wild-type or mutant. In the complementation assay shown in Figure 7E,F, knockdown cells were also transfected with plasmids expressing GFP (0.002 µg) or GFP-tagged CAF40 (0.005 µg) either wild-type or mutant (V186E). To measure the mRNA half-life, S2 cells were treated with actinomycin D (5 µg/ml final concentration) 3 d after transfection and collected at the indicated time points. RNA samples were analyzed by northern blotting.

A detailed description of the procedure to generate the HEK293T CAF40-null cell line is included in the Supplemental Material. For the Bam tethering assays in human cells, HEK293T cells (0.7×10^6 cells per well) were seeded in six-well plates and transfected using Lipofectamine 2000 (Thermo Fisher Scientific). The transfection mixtures contained 0.5 µg of the β-globin reporter containing six MS2-binding sites (β-globin-6xMS2bs), 0.5 µg of the control plasmid containing the β-globin gene fused to a fragment of the GAPDH gene inserted in the 3′ UTR but lacking MS2-binding sites (control: β-globin-GAP), and various amounts of pT7-MS2-HA plasmids for the expression of MS2-HA-fusion proteins [full-length Bam (1 µg), MBP-Bam CBM (0.2 µg) and Bam 4xMut (0.5 µg)].

Firefly and *Renilla* luciferase activities were measured 3 d (S2 cells) or 2 d (HEK293T cells) after transfection by using a Dual-Luciferase Reporter Assay System (Promega). The total RNA was isolated using a Trifast Reagent (Peqlab) and analyzed by northern blotting, as previously described (Braun et al. 2011).

Protein expression and purification

All recombinant proteins were expressed in *E. coli* BL21 (DE3) Star cells (Thermo Fisher Scientific) grown in LB medium overnight at 20°C. The cells were lysed with an EmulsiFlex-C3 homogenizer (AVESTIN) in the indicated lysis buffer supplemented with DNase I (5 µg/mL), lysozyme (1 mg/mL) and complete EDTA-free protease inhibitor cocktail (Roche). Bam constructs were expressed as fusions with an N-terminal, TEV-cleavable MBP tag. The cells were lysed in a buffer containing 50 mM HEPES (pH 7.5), 300 mM NaCl and 2 mM DTT. The proteins were purified from cleared cell lysates by using amylose resin (New England Biolabs), and this was followed by anion chromatography using a HiTrapQ column (GE Healthcare). The Bam constructs were further purified on a Superdex 200 26/600 column (GE Healthcare) in a buffer containing 10 mM HEPES (pH 7.5), 200 mM NaCl and 2 mM DTT.

The purification of *Hs* CAF40 (ARM domain, residues R19–E285) was as previously described (Sgromo et al. 2017). Briefly, the protein was expressed with an N-terminal His_6 tag cleavable by the HRV3C protease. Cells were lysed in a buffer containing 50 mM potassium phosphate (pH 7.5), 500 mM NaCl, 10% (v/v) glycerol, 20 mM imidazole and 2 mM β-mercaptoethanol. The protein was purified from cleared cell lysates with a HiTrap IMAC column (GE Healthcare). The His_6 tag was removed by overnight cleavage using HRV3C protease during dialysis in a buffer containing 50 mM Tris–HCl (pH 7.5), 150 mM NaCl and 1 mM DTT. After

cleavage of the tag, CAF40 was further purified using a HiTrap Heparin column (GE Healthcare) followed by gel filtration on a Superdex 200 26/600 column (GE Healthcare) in a buffer containing 20 mM Tris–HCl (pH 7.5), 150 mM NaCl and 1 mM DTT.

The purification of the *Hs* NOT1 CN9BD–CAF40 complex has been previously described (Chen et al. 2014a). The complex was obtained by co-expression of MBP-tagged NOT1-CN9BD (residues V1351–L1588) with His_6-tagged CAF40 (R19–E285). The cells were lysed in a buffer containing 50 mM potassium phosphate (pH 7.5), 300 mM NaCl and 2 mM β-mercaptoethanol. The complex was purified from the cleared lysate by using amylose resin, and this was followed by removal of the His_6 and MBP tags by cleavage with HRV3C protease overnight at 4°C during dialysis in a buffer containing 50 mM HEPES (pH 7.5), 150 mM NaCl, 10% (v/v) glycerol and 2 mM DTT. The complex was separated from the tags by binding to a HiTrap Heparin column (GE Healthcare), and this was followed by elution with a linear gradient to 1 M NaCl. Finally, size exclusion chromatography was performed using a Superdex 200 26/600 column in a buffer containing 10 mM HEPES (pH 7.5), 150 mM NaCl, 10% (v/v) glycerol and 2 mM DTT.

A detailed description of the purification of the additional modules of the human and *Drosophila* CCR4–NOT complex can be found in the Supplemental Material.

Crystallization, data collection, and structure determination

A detailed description of the crystallization conditions and the structure determination process are included in the Supplemental Material. Diffraction data sets of the CN9BD–CAF40–Bam CBM complex were recorded on a PILATUS 6M detector at the PXII beamline of the Swiss Light Source at a temperature of 100 K. The best data set of the CAF40–Bam CBM complex was recorded on a PILATUS 6M fast detector (DECTRIS) at the DESY beamline P11. The diffraction data and refinement statistics are summarized in Table 1.

In vitro MBP pull-down assays

Purified MBP (20 μg) or MBP-tagged full-length Bam or fragments (40 μg) were incubated with equimolar amounts of purified CCR4–NOT complex modules or subunits and amylose resin (New England Biolabs) in pull-down buffer containing 50 mM HEPES (pH 7.5), 200 mM NaCl and 2 mM DTT. After 1 h incubation, the beads were washed five times with pull-down buffer and the proteins were eluted with pull-down buffer supplemented with 25 mM D-(+)-Maltose. The eluted proteins were precipitated with trichloroacetic acid and analyzed by SDS-PAGE and subsequent Coomassie staining.

In vitro competition assays

Purified GST-tagged CAF40 (ARM domain, 50 μg) was incubated with equimolar amounts of either MBP-tagged Bam CBM or MBP-Roquin CBM, increasing amounts of His_6-NusA-tagged Bam CBM as a competitor, and 50 μL 50% slurry of Protino glutathione agarose 4B (Macherey-Nagel). Purified His_6-NusA served as a negative control. The experiment was performed in pull-down buffer. After 1 h of incubation, the beads were pelleted and washed three times with pull-down buffer. The proteins bound to the beads were eluted by boiling the beads in 2× protein sample buffer. The eluted proteins were analyzed by SDS-PAGE and subsequent Coomassie staining.

Isothermal titration calorimetry (ITC) and bioinformatics analysis

The ITC measurements were performed as previously described (Igreja et al. 2014). A detailed description of the ITC conditions and the bioinformatic analysis can be found in the Supplemental Material.

DATA DEPOSITION

The coordinates for the structure of the Bam CBM peptide bound to CAF40 and to the CAF40 module were deposited in the Protein Data Bank (PDB) under ID code 5ONB and 5ONA, respectively.

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

We are grateful to Heike Budde and Catrin Weiler for excellent technical support. We thank Lara Wohlbold for cloning the *Dm* Bam construct, Daniel Peter and Stefan Grüner for assistance with the ITC experiments, and Eugene Valkov for comments on the manuscript. We thank the staff at the PX beamlines of the Swiss Light Source, Villigen, and the staff at the P11 beamline of the DESY, Hamburg, for assistance with X-ray data collection.

Received October 20, 2017; accepted December 2, 2017.

REFERENCES

Basquin J, Roudko VV, Rode M, Basquin C, Séraphin B, Conti E. 2012. Architecture of the nuclease module of the yeast Ccr4-not complex: the Not1-Caf1-Ccr4 interaction. *Mol Cell* **48:** 207–218.

Bawankar P, Loh B, Wohlbold L, Schmidt S, Izaurralde E. 2013. NOT10 and C2orf29/NOT11 form a conserved module of the CCR4-NOT complex that docks onto the NOT1 N-terminal domain. *RNA Biol* **10:** 228–244.

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.

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

Bhaskar V, Basquin J, Conti E. 2015. Architecture of the ubiquitylation module of the yeast Ccr4-Not complex. *Structure* **23:** 921–928.

Bhaskar V, Roudko V, Basquin J, Sharma K, Urlaub H, Séraphin B, Conti E. 2013. Structure and RNA-binding properties of the Not1-Not2-Not5 module of the yeast Ccr4-Not complex. *Nat Struct Mol Biol* **20:** 1281–1288.

Boland A, Chen Y, Raisch T, Jonas S, Kuzuoğlu-Öztürk D, Wohlbold L, Weichenrieder O, Izaurralde E. 2013. Structure and assembly of the

NOT module of the human CCR4-NOT complex. *Nat Struct Mol Biol* **20:** 1289–1297.

Braun JE, Huntzinger E, Fauser M, Izaurralde E. 2011. GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets. *Mol Cell* **44:** 120–133.

Carreira-Rosario A, Buszczak M. 2014. A competitive cell fate switch. *Dev Cell* **31:** 261–262.

Chang CT, Bercovich N, Loh B, Jonas S, Izaurralde E. 2014. The activation of the decapping enzyme DCP2 by DCP1 occurs on the EDC4 scaffold and involves a conserved loop in DCP1. *Nucleic Acids Res* **42:** 5217–5233.

Chau J, Kulnane LS, Salz HK. 2012. Sex-lethal enables germline stem cell differentiation by down-regulating Nanos protein levels during *Drosophila* oogenesis. *Proc Natl Acad Sci* **109:** 9465–9470.

Chen Y, Boland A, Kuzuoğlu-Öztürk D, Bawankar P, Loh B, Chang CT, Weichenrieder O, Izaurralde E. 2014a. A DDX6-CNOT1 complex and W-binding pockets in CNOT9 reveal direct links between miRNA target recognition and silencing. *Mol Cell* **54:** 737–750.

Chen D, Wu C, Zhao S, Geng Q, Gao Y, Li X, Zhang Y, Wang Z. 2014b. Three RNA binding proteins form a complex to promote differentiation of germline stem cell lineage in *Drosophila*. *PLoS Genet* **10:** e1004797.

Cooley L, Kelley R, Spradling A. 1988. Insertional mutagenesis of the *Drosophila* genome with single P elements. *Science* **239:** 1121–1128.

Draper MP, Liu HY, Nelsbach AH, Mosley SP, Denis CL. 1994. CCR4 is a glucose-regulated transcription factor whose leucine-rich repeat binds several proteins important for placing CCR4 in its proper promoter context. *Mol Cell Biol* **14:** 4522–4531.

Draper MP, Salvadore C, Denis CL. 1995. Identification of a mouse protein whose homolog in *Saccharomyces cerevisiae* is a component of the CCR4 transcriptional regulatory complex. *Mol Cell Biol* **15:** 3487–3495.

Dupressoir A, Morel AP, Barbot W, Loireau MP, Corbo L, Heidmann T. 2001. Identification of four families of yCCR4- and Mg^{2+}-dependent endonuclease-related proteins in higher eukaryotes, and characterization of orthologs of yCCR4 with a conserved leucine-rich repeat essential for hCAF1/hPOP2 binding. *BMC Genomics* **2:** 9.

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.

Fu Z, Geng C, Wang H, Yang Z, Wenig C, Li H, Deng L, Liu L, Liu N, Ni J, et al. 2015. Twin promotes the maintenance and differentiation of germline stem cell lineage through modulation of multiple pathways. *Cell Rep* **13:** 1366–1379.

Garces RG, Gillon W, Pai EF. 2007. Atomic model of human Rcd-1 reveals an armadillo-like-repeat protein with in vitro nucleic acid binding properties. *Protein Sci* **16:** 176–188.

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.

Jonas S, Izaurralde E. 2015. Towards a molecular understanding of microRNA-mediated gene silencing. *Nat Rev Genet* **16:** 421–433.

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

Li Y, Minor NT, Park JK, McKearin DM, Maines JZ. 2009. Bam and Bgcn antagonize *Nanos*-dependent germ-line stem cell maintenance. *Proc Natl Acad Sci* **106:** 9304–9309.

Li Y, Zhang Q, Carreira-Rosario A, Maines JZ, McKearin DM, Buszczak M. 2013. Mei-P26 cooperates with Bam, Bgcn and Sxl to promote early germline development in the *Drosophila* ovary. *PLoS One* **8:** e58301.

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.

Mauxion F, Prève B, Séraphin B. 2013. C2ORF29/CNOT11 and CNOT10 form a new module of the CCR4-NOT complex. *RNA Biol* **10:** 267–276.

McKearin D, Ohlstein B. 1995. A role for the *Drosophila* bag-of-marbles protein in the differentiation of cystoblasts from germline stem cells. *Development* **121:** 2937–2947.

McKearin DM, Spradling AC. 1990. bag-of-marbles: a *Drosophila* gene required to initiate both male and female gametogenesis. *Genes Dev* **4:** 2242–2251.

Neumüller RA, Betschinger J, Fischer A, Bushati N, Poernbacher I, Mechtler K, Cohen SM, Knoblich JA. 2008. Mei-P26 regulates microRNAs and cell growth in the *Drosophila* ovarian stem cell lineage. *Nature* **454:** 241–245

Ohlstein B, McKearin D. 1997. Ectopic expression of the *Drosophila* Bam protein eliminates oogenic germline stem cells. *Development* **124:** 3651–3662.

Pan L, Wang S, Lu T, Weng C, Song X, Park JK, Sun J, Yang ZH, Yu J, Tang H, et al. 2014. Protein competition switches the function of COP9 from self-renewal to differentiation. *Nature* **514:** 233–236.

Petit AP, Wohlbold L, Bawankar P, Huntzinger E, Schmidt S, Izaurralde E, Weichenrieder O. 2012. The structural basis for the interaction between the CAF1 nuclease and the NOT1 scaffold of the human CCR4-NOT deadenylase complex. *Nucleic Acids Res* **40:** 11058–11072.

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

Sgromo A, Raisch T, Bawankar P, Bhandari D, Chen Y, Kuzuoğlu-Öztürk D, Weichenrieder O, Izaurralde E. 2017. A CAF40-binding motif facilitates recruitment of the CCR4-NOT complex to mRNAs targeted by *Drosophila* Roquin. *Nat Commun* **8:** 14307.

Shen R, Weng C, Yu J, Xie T. 2009. eIF4A controls germline stem cell self-renewal by directly inhibiting BAM function in the *Drosophila* ovary. *Proc Natl Acad Sci* **106:** 11623–11628.

Suzuki A, Saba R, Miyoshi K, Morita Y, Saga Y. 2012. Interaction between NANOS2 and the CCR4-NOT deadenylation complex is essential for male germ cell development in mouse. *PLoS One* **7:** e33558.

Tompa P. 2012. Intrinsically disordered proteins: a 10-year recap. *Trends Biochem Sci* **37:** 509–516.

Van Roey K, Uyar B, Weatheritt RJ, Dinkel H, Seiler M, Budd A, Gibson TJ, Davey NE. 2014. Short linear motifs: ubiquitous and functionally diverse protein interaction modules directing cell regulation. *Chem Rev* **114:** 6733–6778.

Wahle E, Winkler GS. 2013. RNA decay machines: deadenylation by the Ccr4-not and Pan2-Pan3 complexes. *Biochim Biophys Acta* **1829:** 561–570.

Zekri L, Kuzuoğlu-Öztürk D, Izaurralde E. 2013. GW182 proteins cause PABP dissociation from silenced miRNA targets in the absence of deadenylation. *EMBO J* **32:** 1052–1065.

SUPPLEMENTAL MATERIAL

Drosophila Bag-of-marbles directly interacts with the CAF40 subunit of the CCR4-NOT complex to elicit repression of mRNA targets

Annamaria Sgromo, Tobias Raisch, Charlotte Backhaus, Csilla Keskeny, Vikram Alva, Oliver Weichenrieder and Elisa Izaurralde

Supplemental Table S1. Constructs and mutants used in this study.

Name	**Bag of marbles** (Uniprot P22745)	**Comment**
Bam	λN-HA-Bam 1–442	
	SBP-Bam 1–442	
	MS2-HA-Bam 1–442	
	MBP-Bam 1–442	
Bam-N	λN-HA-Bam 1–140	
	GFP-Bam 1–140	
	MS2-HA-Bam 1–140	
	MBP-Bam 1–140	
Bam-C	λN-HA-Bam 141–442	
	GFP-Bam 141–442	
	MBP-Bam 141–442	
CBM	λN-HA-GST-Bam 13–36	
	SBP-MBP-Bam 13–36	
	MS2-HA-MBP-Bam 13–36	
	MBP-Bam 13–36	
	His_6-NusA-Bam 13–36	
ΔCBM	λN-HA-Bam Δ13–36	Disrupts CAF40 binding
	MBP-Bam Δ13–36	Disrupts CAF40 binding
L17E	λN-HA-Bam L17E	Disrupts CAF40 binding
	MBP-Bam 1–442 L17E	Disrupts CAF40 binding
M24E	λN-HA-Bam M24E	Disrupts CAF40 binding
	MBP-Bam 1–442 M24E	Disrupts CAF40 binding
L28E	λN-HA-Bam L28E	Disrupts CAF40 binding
V32E	λN-HA-Bam V32E	Disrupts CAF40 binding
4xMut	λN-HA-Bam L17E, M24E, L28E, V32E	Disrupts CAF40 binding
	SBP-Bam L17E, M24E, L28E, V32E	Disrupts CAF40 binding
	MS2-HA-Bam L17E, M24E, L28E, V32E	Disrupts CAF40 binding

Name	***Hs* NOT1** (Uniprot A5YKK6)	**Comment**
NOT1 N	*Hs* NOT1 1–1000	
NOT1 MIF4G	His_6-*Hs* NOT1 1093–1317	MIF4G-like domain
NOT1 CN9BD	MBP-*Hs* NOT1 1351–1588	CNOT9-binding domain
NOT1 MIF4G-C	MBP-*Hs* NOT1 1607–1815	Predicted MIF4G-like domain
NOT1 SHD	MBP-*Hs* NOT1 1833–2361	Superfamily homology domain

Name	***Hs* NOT2** (Uniprot Q9NZN8)	**Comment**
NOT2-C	MBP-*Hs* NOT2 350–540	

Name	***Hs* NOT3** (Uniprot O75175)	**Comment**
NOT3-N	MBP-*Hs* NOT3 2–212	
NOT3-C	His_6-*Hs* NOT3 607–748	

Name	***Hs* CCR4a** (Uniprot Q9ULM6)	**Comment**
CCR4a full-length	MBP-*Hs* CCR4a	

Name	***Hs* CAF1** (Uniprot Q9UIV1)	**Comment**
CAF1 full-length	MBP-*Hs* CAF1	

Name	***Hs* CAF40** (Uniprot Q92600)	**Comment**
CAF40-ARM wt	His_6-*Hs* CAF40 19–285	
	GST-*Hs* CAF40 19–285	
CAF40 wt	SBP-MBP-CAF40 1–299	
V181E	SBP-MBP-CAF40 1–299 V181E	

Name	***Hs* NOT10** (Uniprot Q9H9A5)	**Comment**
NOT10 TPR	*Hs* NOT10 25–707	

Name	***Hs* NOT11** (Uniprot Q9UKZ1)	**Comment**
NOT11-C	*Hs* NOT11 257–498-His_6	

Name	***Dm* CAF40 (1–304)** (Uniprot Q7JVP2)	**Comment**
CAF40 wt	λN-HA-CAF40 1–304	
	GFP-CAF40 1–304	dsRNA resistant
CAF40-ARM wt	His_6-CAF40 25–291	
V186E	GFP-CAF40 V186E	Disrupts CBM binding; dsRNA resistant
	His_6-CAF40 25–291 V186E	Disrupts CBM binding
2xMut	His_6-CAF40 25–291 Y139E, G146E	Disrupts CBM binding

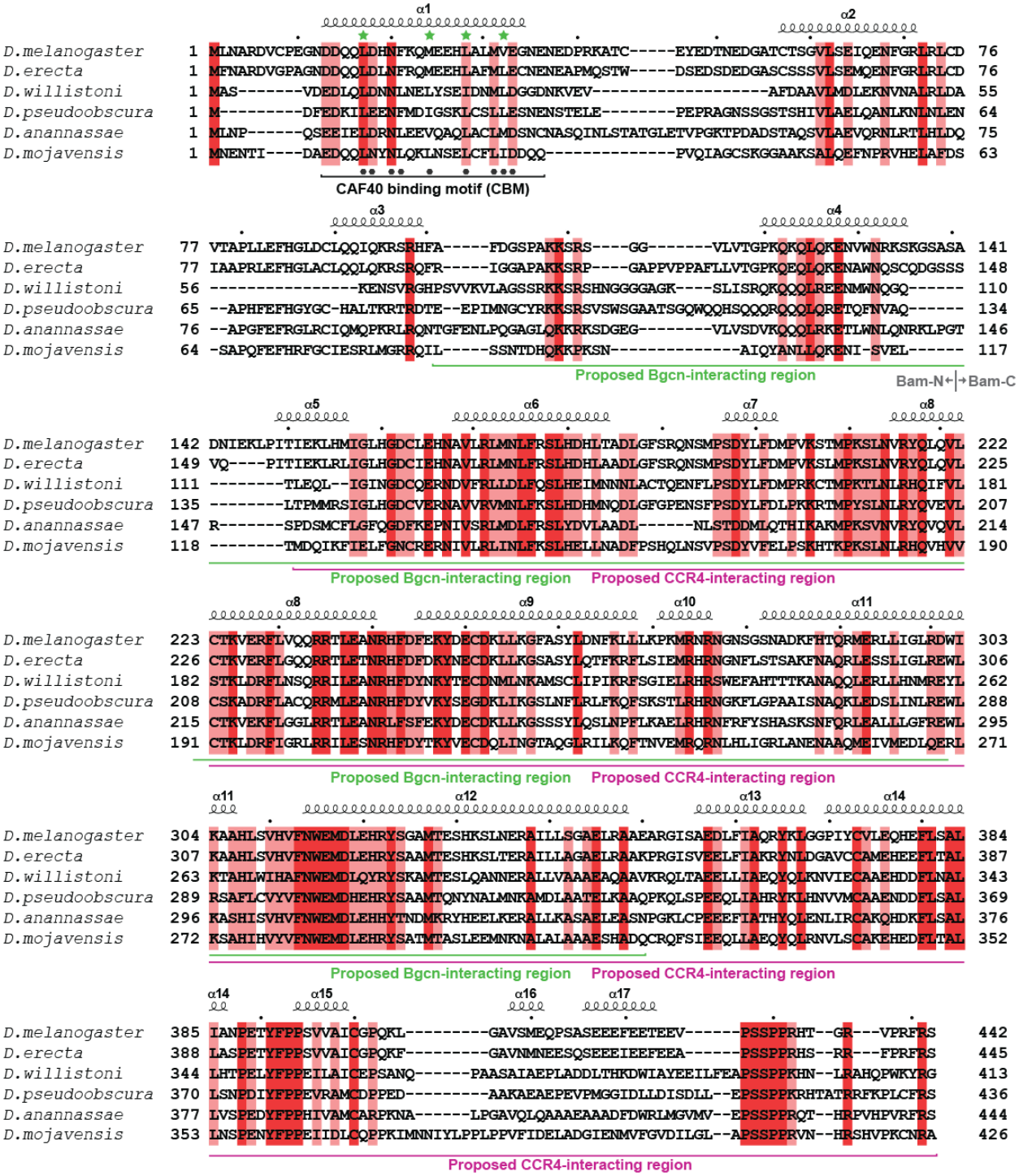
Name	***Dm* NOT1** (Uniprot A8DY81)	**Comment**
NOT1	λN-HA-NOT1	
NOT1-CN9BD	λN-HA-NOT1 1467–1719	CAF40-binding domain
	MBP-NOT1 1468–1719	CAF40-binding domain

Supplemental Table S2. Antibodies used in this study.

Antibody	**Source**	**Catalog Number**	**Dilution**	**Monoclonal/ Polyclonal**
Anti-HA-HRP	Roche	12 013 819 001	1:5,000	Monoclonal
Anti-GFP (for western blotting)	Roche	11 814 460 001	1:2,000	Mouse Monoclonal
Anti-GFP (for immunoprecipitation)	In house			Rabbit Polyclonal
Anti-*Dm* NOT1	Kind gift from E. Wahle	T6199	1:1,000	Rabbit Polyclonal
Anti-*Dm* CAF40	In house		1:1,000	Rabbit Polyclonal
Anti-*Hs* NOT1	In house		1:2,000	Rabbit Polyclonal
Anti-*Hs* NOT2	Bethyl	A302-562A	1:2,000	Rabbit Polyclonal
Anti-*Hs* NOT3	Abcam	Ab55681	1:2,000	Monoclonal
Anti-*Hs* CAF40 (RQCD1)	Proteintech	22503-1-AP	1:1,000	Rabbit Polyclonal
Anti-tubulin	Sigma-Aldrich	T6199	1:10,000	Monoclonal
Anti-*Dm* PABP	In house		1:10,000	Rabbit Polyclonal
Anti-V5	AbD Serotec	MCA1360GA	1:5,000	Monoclonal
Anti-mouse-HRP	GE Healthcare	NA931V	1:10,000	Monoclonal

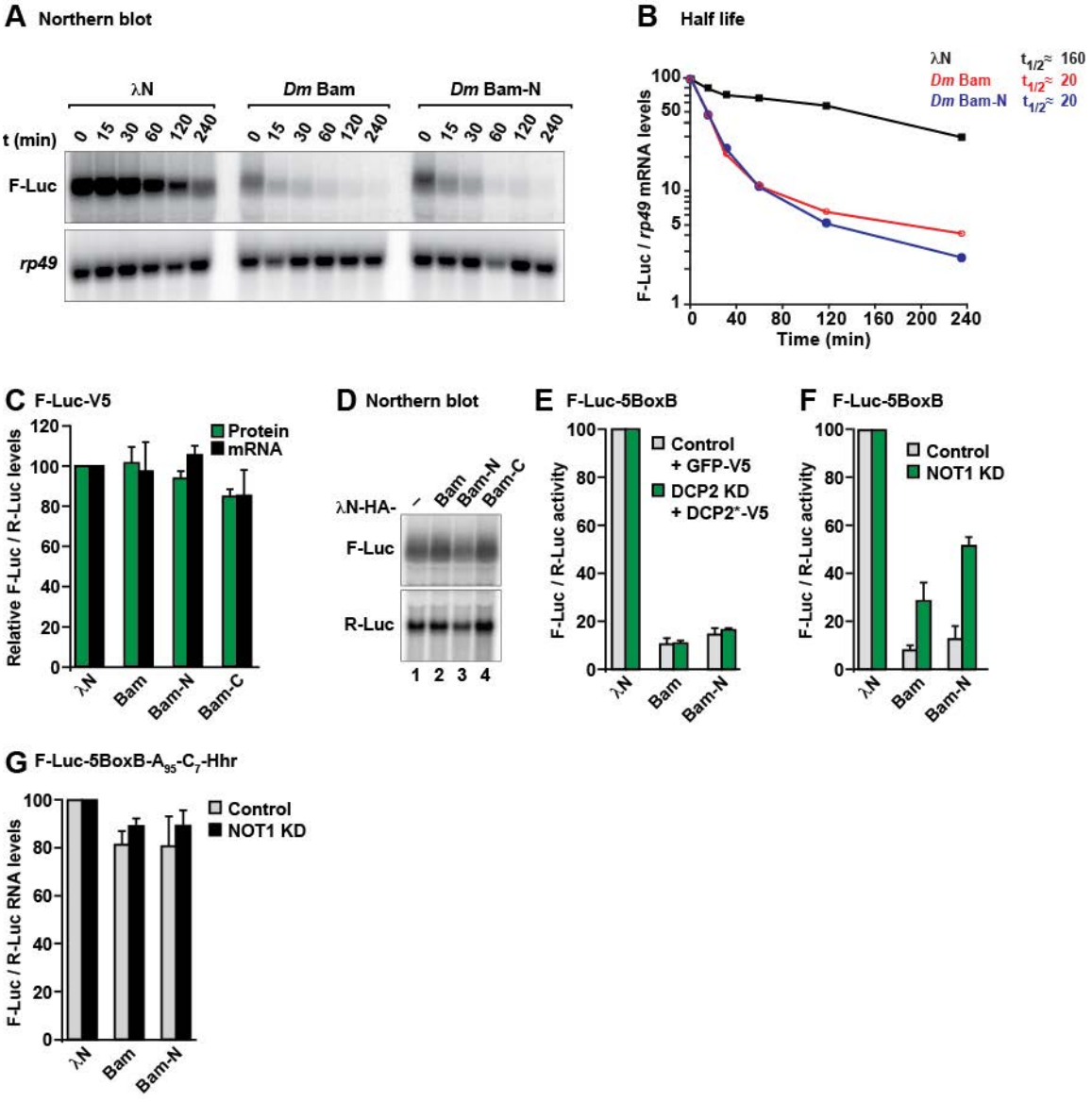
SUPPLEMENTAL FIGURES

Supplemental Figure S1



Supplemental Figure S1. Sequence alignment of *Drosophila* Bam. The secondary structure elements, as predicted by PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/), are indicated in black. Residues conserved in all aligned sequences are shown with a dark red background, and residues with >70% similarity are highlighted in light red; conservation scores were calculated using the SCORECONS webserver (Valdar 2002). The CAF40-binding motif (CBM) is indicated. Black dots indicate residues in the CBM that directly contact CAF40. Green asterisks indicate residues mutated in this study.

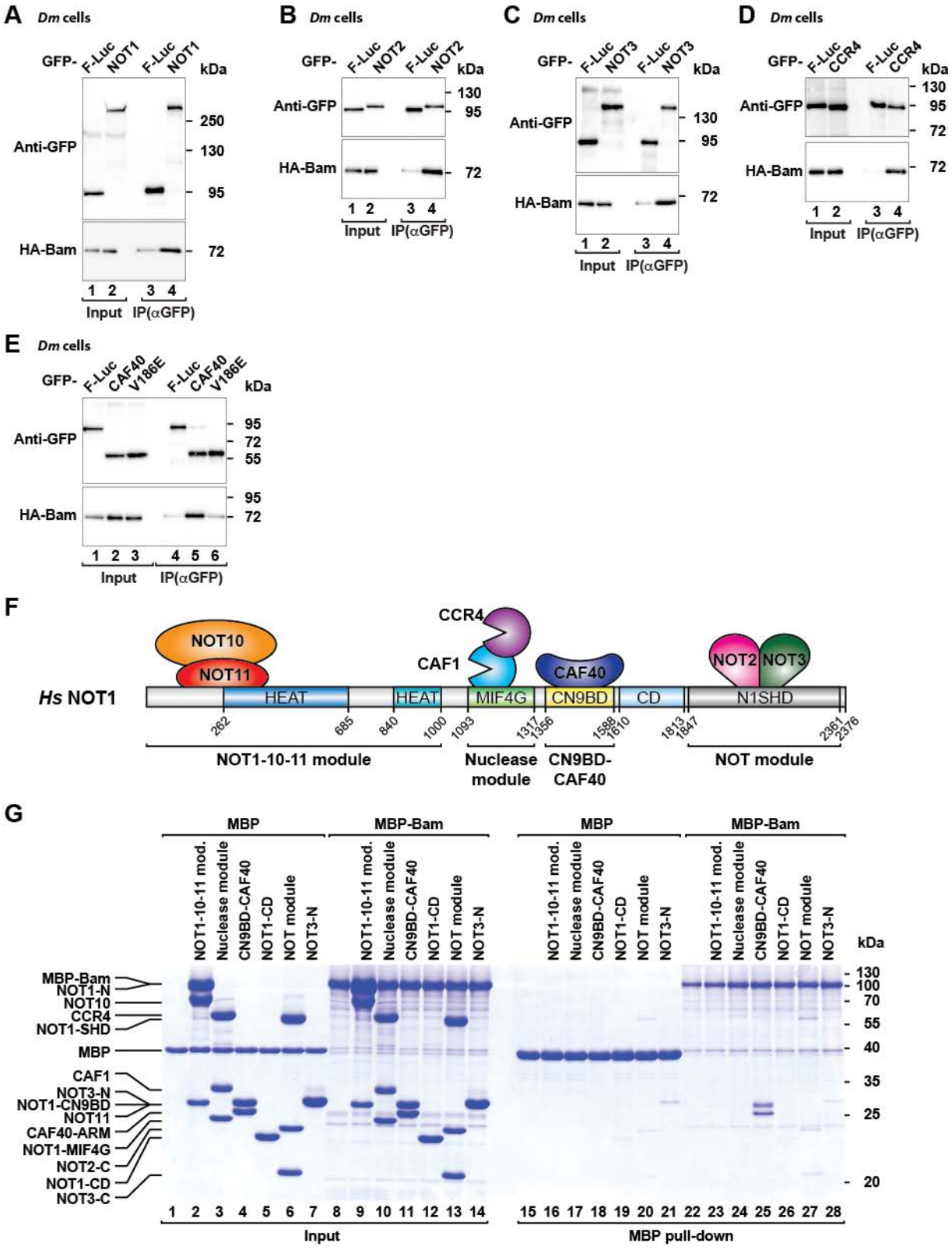
Supplemental Figure S2



Supplemental Figure S2. Bam promotes mRNA degradation. (*A*) Representative northern blot showing the decay of the F-Luc-5BoxB mRNA in S2 cells expressing λN-HA or λN-HA-tagged Bam or the Bam-N fragment after inhibition of transcription by actinomycin D. (*B*) F-Luc mRNA levels were normalized to those of the *rp49* mRNA and plotted against time. (*C,D*) Tethering assay using the F-Luc reporter lacking BoxB sites and λN-HA-tagged Bam (full-length or the indicated fragments) in S2 cells. The samples were analyzed as described in Figure 1B–D.

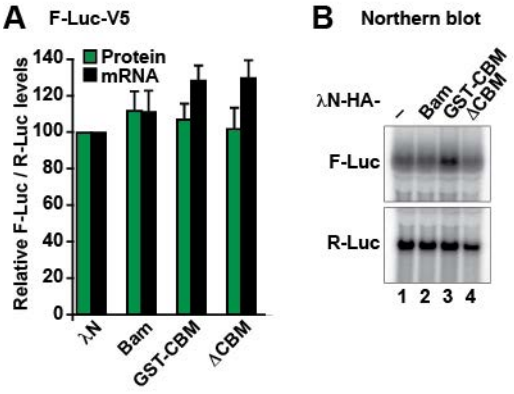
The corresponding experiment with the F-Luc-5BoxB reporter is shown in Figure 1B. (**E**) Normalized F-Luc activity values corresponding to the experiment described in Figure 2A and B. The F-Luc-5BoxB activity was normalized to that of the R-Luc transfection control and set to 100% in cells expressing the λN-HA peptide. The grey and green bars represent the normalized F-Luc-5BoxB activity in control cells expressing GFP-V5 and in DCP2-depleted cells expressing GFP-DCP2*-V5, respectively. (*F*) Normalized F-Luc-5BoxB activity values corresponding to the experiment described in Figure 2D and E. (G) Normalized F-Luc-5BoxB-A_{95}-C_7-HhR reporter mRNA levels corresponding to the experiment described in Figure 2F and G.

Supplemental Figure S3



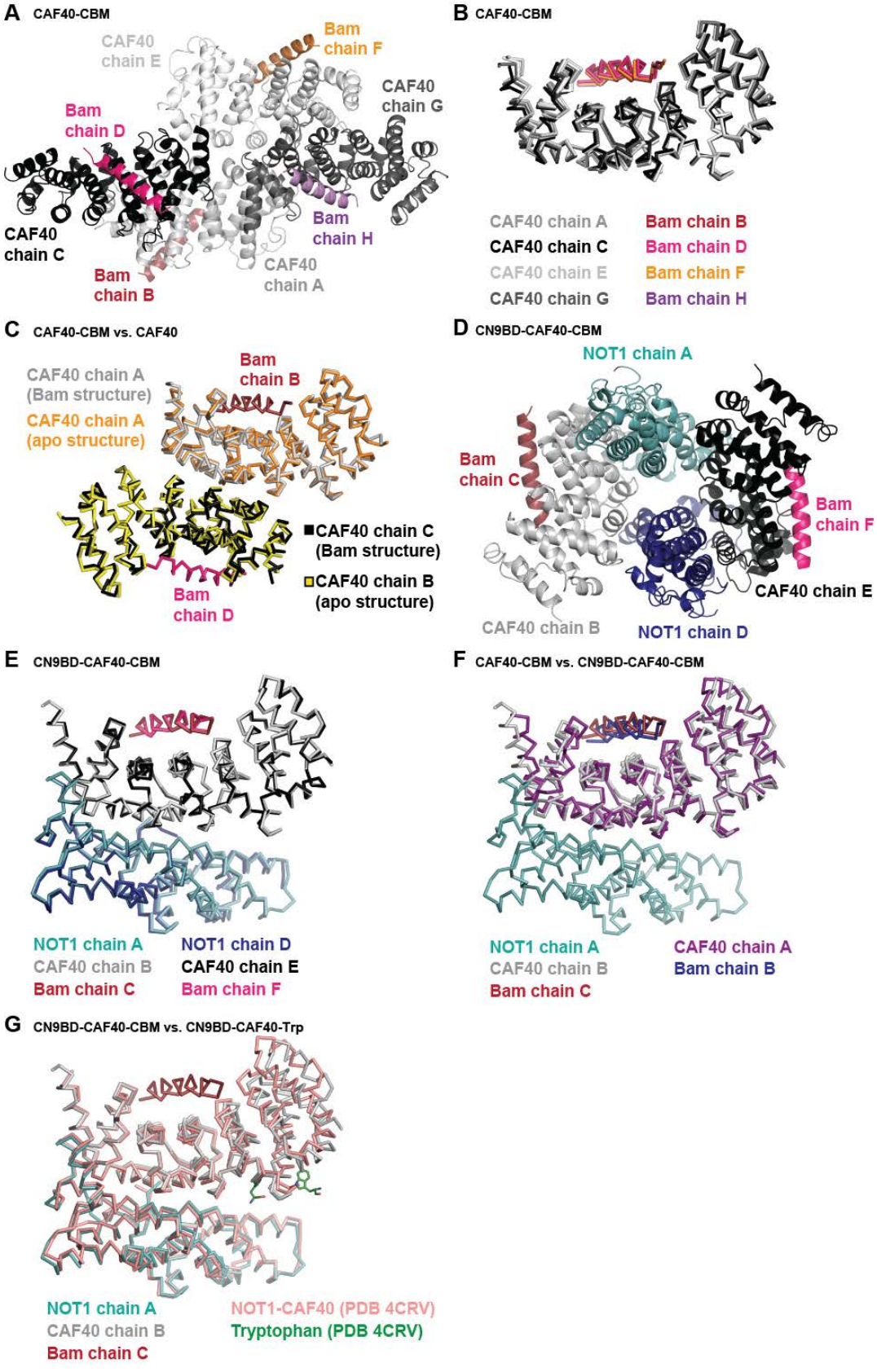
Supplemental Figure S3. Bam interacts with the CCR4-NOT complex. (*A*-*E*) Coimmunoprecipitation assays showing the interaction of HA-tagged Bam with the indicated GFP-tagged CCR4-NOT subunits in S2 cell lysates treated with RNase A. In all panels, GFP-F-Luc served as a negative control. Inputs (1% for the HA-tagged proteins and 3% for the GFP-tagged proteins) and immunoprecipitates (30% for the HA-tagged proteins and 10% for the GFP-tagged proteins) were analyzed by western blotting. Protein size markers are shown on the right in each panel. (*F*) Schematic representation of the *Hs* CCR4-NOT complex. NOT1 contains two HEAT repeat domains (shown in blue and petrol), a MIF4G domain composed of HEAT repeats (green), a three-helix bundle domain (CN9BD, yellow), a connector domain (CD, light blue) and a NOT1 superfamily homology domain (SHD, gray), which also consists of HEAT repeats. The additional subunits of the complex are shown at their binding positions on NOT1. (*G*) *In vitro* MBP pull-down assay testing the interaction of MBP-tagged full-length Bam with the indicated *Hs* CCR4-NOT subcomplexes. MBP served as a negative control.

Supplemental Figure S4



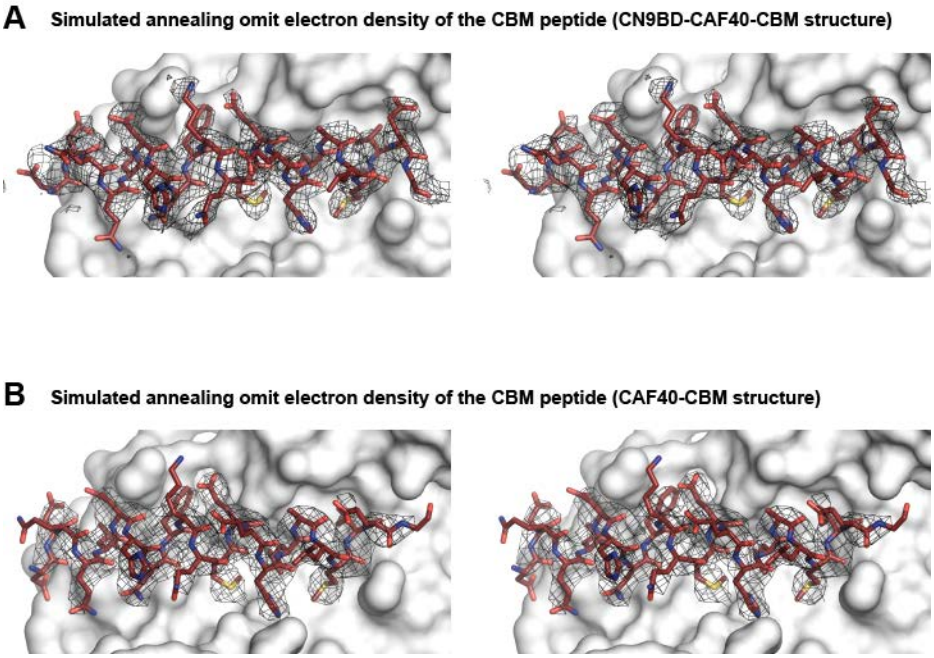
Supplemental Figure S4. Bam requires binding to the mRNA target to induce degradation. (*A,B*) Tethering assay using the F-Luc reporter lacking BoxB sites and λN-HA-tagged Bam (full-length or the indicated fragments) in S2 cells. The samples were analyzed as described in Figure 1B–D. The corresponding experiment with the F-Luc-5BoxB reporter is shown in Figure 3E and F.

Supplemental Figure S5



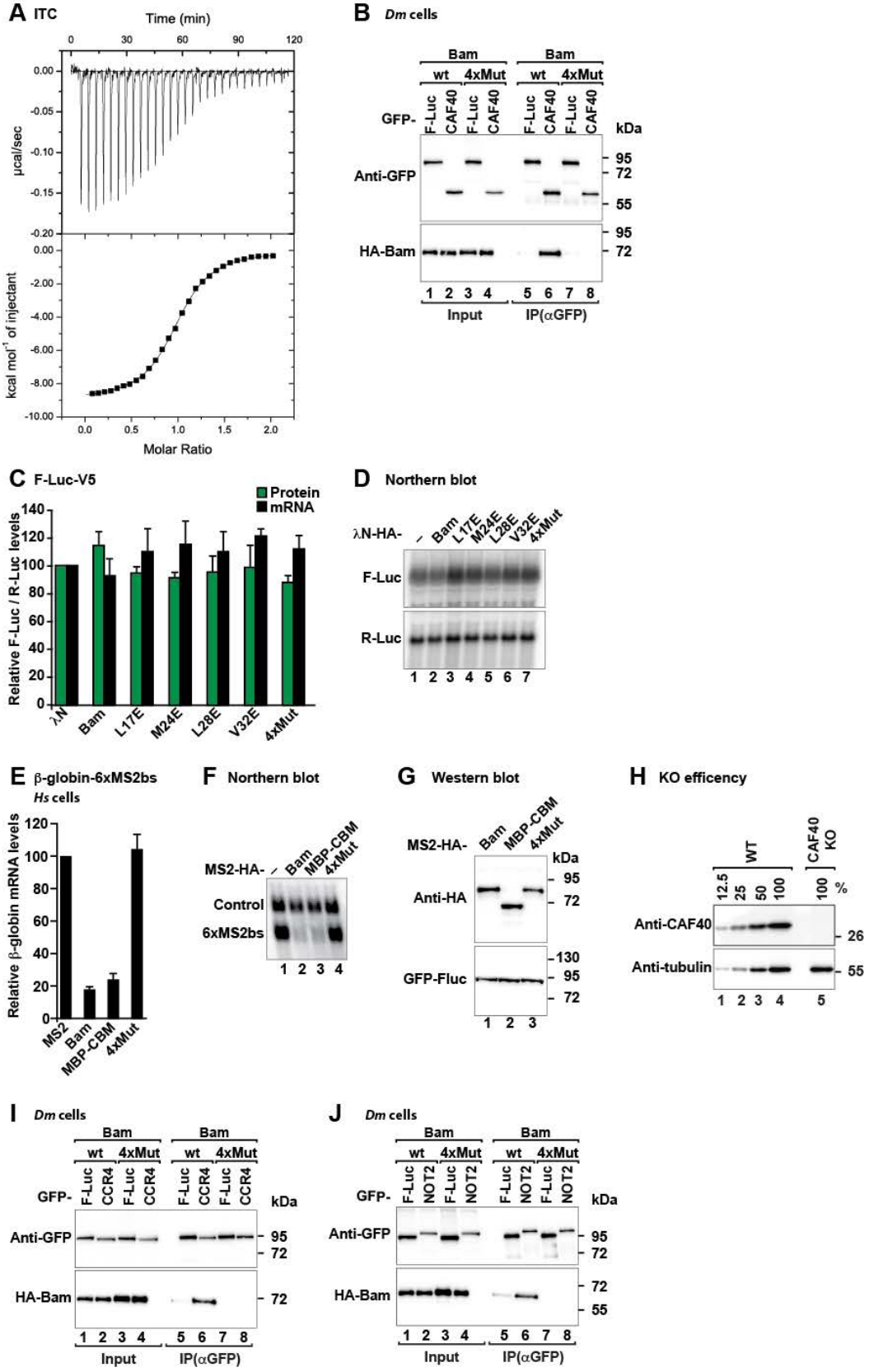
Supplemental Figure S5. Crystal structure of the Bam CBM bound to CAF40 and the CN9BD–CAF40 module. (*A*) Crystal packing of the CAF40–Bam CBM complex. The four copies of CAF40 (chains A, C, E and G) are shown in different shades of gray, the four Bam CBM peptides (chains B, D, F and H) in different colors. (*B*) Superposition of the four CAF40-Bam CBM complexes in the asymmetric unit in ribbon representation. Colors are as in (A). (*C*) Superposition of a CAF40 homodimer (orange and yellow, PDB 2FV2; Garces et al. 2007), with a CAF40 homodimer bound to the Bam CBM (chains A–D, colors are as in (A)). (*D*) Crystal packing of the NOT1 CN9BD-CAF40-Bam CBM complex in cartoon representation. The NOT1 CN9BD is shown in cyan and blue (chains A and D, respectively), CAF40 in gray and black (chains B and E), and Bam in red and pink (chains C and F). (*E*) Superposition of the two NOT1 CN9BD-CAF40-Bam complexes in the asymmetric unit. Colors are as in (D). (*F*) Superposition of the CAF40-Bam complex and NOT1 CN9BD-CAF40-Bam complex structures. (*G*) Superposition of the NOT1 CN9BD-CAF40-Bam CBM complex with the NOT1 CN9BD-CAF40 complex with bound tryptophan (PDB 4CRV) (Chen et al. 2014).

Supplemental Figure S6



Supplemental Figure S6. Simulated annealing electron density of the Bam CBM peptide. (*A*) Stereo view showing the $2F_O$-F_C simulated annealing composite omit map surrounding the CN9BD-CAF40-bound CBM peptide contoured at 1.0 σ. This map was generated with Phenix.Composite_omit_map (Afonine et al. 2012) using the final refined CN9BD-CAF40-Bam model. (*B*) Stereo view showing the $2F_O$-F_C simulated annealing composite omit map surrounding the CAF40-bound CBM peptide contoured at 1.0 σ. This map was generated with Phenix.Composite_omit_map (Afonine et al. 2012) using the final refined CAF40-Bam model.

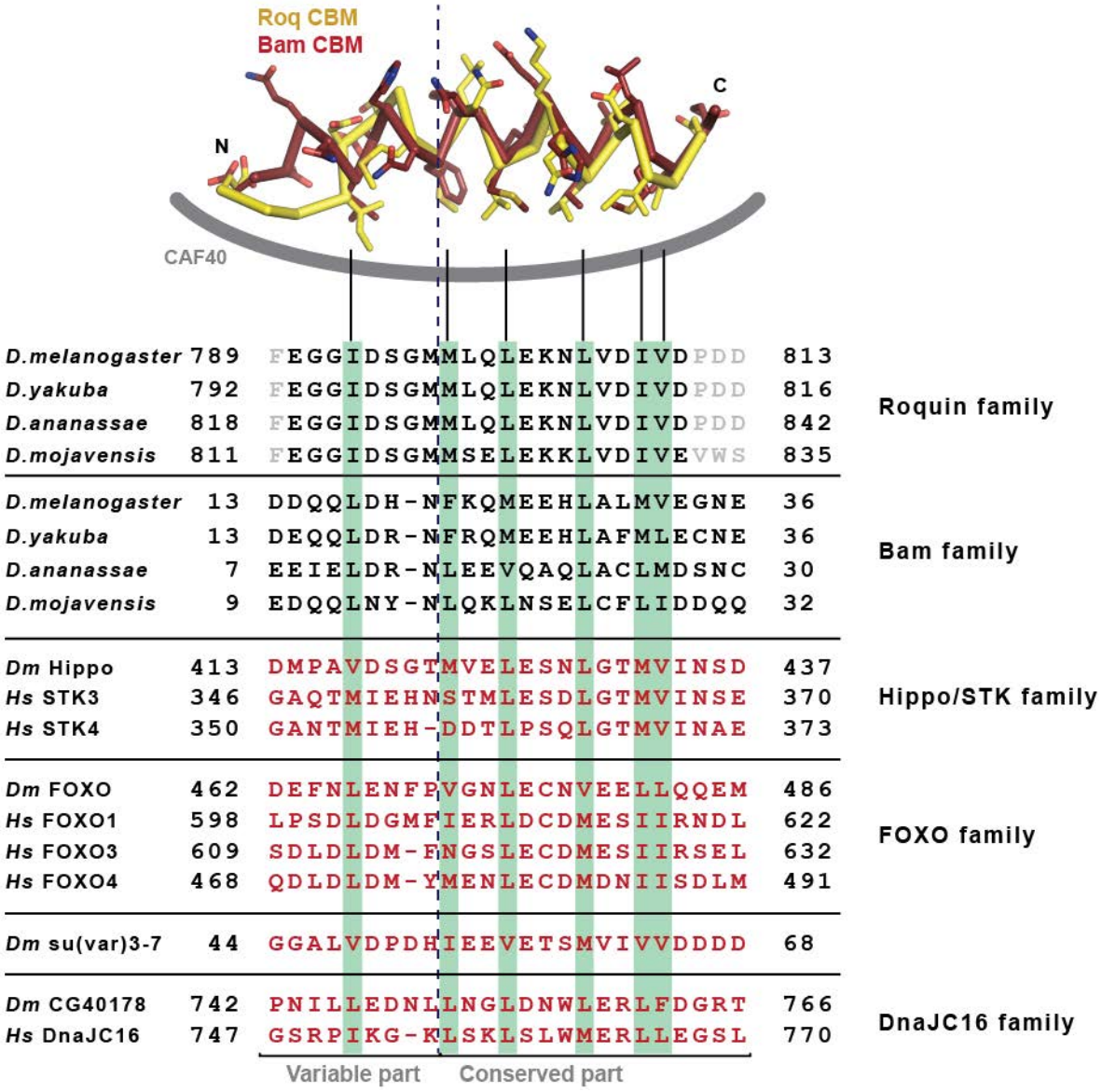
Supplemental Figure S7



Supplemental Figure S7. The CBM is required for Bam activity. (*A*) Representative isothermal titration calorimetry thermogram showing the interaction of the MBP-tagged Bam CBM with the NOT1 CN9BD-CAF40 complex. The upper panel shows raw data, and the lower panel shows the integration of heat changes associated with each injection. Data were fitted using a one-site binding model. (*B*) Coimmunoprecipitation assay showing the interaction of HA-tagged Bam with GFP-tagged CAF40 in S2 cell lysates treated with RNase A. GFP-F-Luc served as a negative control. Inputs (1% for the HA-tagged proteins and 3% for the GFP-tagged proteins) and immunoprecipitates (30% for the HA-tagged proteins and 10% for the GFP-tagged proteins) were analyzed by western blotting. (*C,D*) Tethering assay using the F-Luc reporter lacking BoxB sites and λN-HA-tagged Bam (wild-type or the indicated mutants) in S2 cells. The samples were analyzed as described in Figure 1B–D. The corresponding experiment with the F-Luc-5BoxB reporter is shown in Figure 6A and B. (*E*) Tethering assays using the β-globin-6xMS2bs reporter and MS2-HA-tagged Bam (full-length, MBP-Bam CBM or the 4xMut) in human HEK293T cells. A plasmid expressing a β-globin mRNA reporter lacking MS2-binding sites (Control) served as a transfection control. The β-globin-6xMS2bs mRNA level was normalized to that of the control mRNA and set to 100% in cells expressing MS2-HA. The mean values ± s.d. from three independent experiments are shown in (E). (*F*) Representative northern blot of samples shown in (E). (**G**) Western blot showing the equivalent expression of the MS2-HA-tagged Bam constructs used in (E) and (F). (*H*) Western blot showing the efficiency of the CAF40 depletion in HEK293T cells corresponding to the experiment shown in Figure 6D and E. Dilutions of control cell lysates were loaded in lanes (1–4) to estimate the efficacy of the depletion. Tubulin served as a loading control. KO: knockout. Protein size markers are shown on the right in each panel. (*I,J*) Coimmunoprecipitation assays showing the interaction of HA-tagged Bam with GFP-

tagged CCR4 (I) and NOT2 (J) in S2 cell lysates treated with RNaseA. GFP-F-Luc served as a negative control. Inputs (1% for the HA-tagged proteins and 3% for the GFP-tagged proteins) and immunoprecipitates (30% for the HA-tagged proteins and 10% for the GFP-tagged proteins) were analyzed by western blotting. Protein size markers (kDa) are shown on the right in each panel.

Supplemental Figure S8



Supplemental Figure S8. Profile-based sequence alignment. Profile-based sequence alignment of CBMs from Roquin and Bam of the indicated *Drosophila* species, as well as putative CBMs of proteins from *Hs* and *Dm* shown in red. Residues known or expected to interact with CAF40 are highlighted by a light green background. Gray letters indicate residues that were not included in the crystallization setup. Numbers on both sides of the alignment indicate the residue numbers of the respective fragment boundaries.

SUPPLEMENTAL MATERIAL AND METHODS

DNA constructs

The plasmids used for the expression of subunits of the human and *Dm* CCR4-NOT complex and *Dm* Roq in cells have been previously described (Brau et al. 2011; Bawankar et al. 2013; Sgromo et al. 2017). The plasmids for the expression of *Hs* NOT2-C, NOT3-C, CAF40 ARM domain and the NOT1 MIF4G, CN9BD, CD and SHD domains in *Escherichia coli* have been previously described (Petit et al. 2012; Boland et al. 2013; Chen et al. 2014a; Sgromo et al. 2017). The plasmids for expression of the β-globin-6xMS2bs and the control β-globin-GAP mRNA in human cells were kindly provided by Dr. Lykke-Andersen and have been previously described (Lykke-Andersen et al. 2000). The plasmids for tethering assays in S2 cells (F-Luc-5BoxB, F-Luc-V5, F-Luc-5BoxB-$A_{95}C_7$-HhR, and R-Luc) have been previously described (Behm-Ansmant et al. 2006; Zekri et al. 2013).

For expression of Bam (full-length and fragments) in *Dm* S2 cells, the corresponding cDNA was amplified from total *Dm* oocyte cDNA and cloned between the XhoI and ApaI restriction sites of the pAc5.1-λN-HA and pAc5.1-GFP vectors (Rehwinkel et al. 2005; Tritschler et al. 2007). For expression in HEK293T cells, the cDNA encoding Bam was inserted between the BglII and BamHI restriction sites of the pT7-V5-SBP-C1 and pT7-MS2-HA vectors (Jonas et al. 2013). The plasmids for expression of Bam (full-length, Bam CBM and Bam fragments) in *Escherichia coli* were obtained by inserting the corresponding Bam cDNA fragments between the XhoI and AvrII restriction of the pnYC-vM plasmid (Diebold et al. 2011), thus yielding fusion proteins carrying N-terminal MBP tags cleavable by the TEV protease. For expression of the *Hs* NOT1-10-11 complex, two plasmids were generated. A cDNA fragment encoding the *Hs* NOT1 N-terminus (residues M1–D1000) was inserted into the AvrII restriction site of the pnYC vector,

which does not encode a solubility tag. cDNA fragments encoding *Hs* NOT10 (residues D25–Q707) and *Hs* NOT11 (residues D257–D498) were cloned in a bicistronic plasmid based on the pnEA backbone, thus resulting in the expression of untagged NOT10 and NOT11 with a C-terminal, TEV-cleavable His_6 tag. For expression of the human catalytic module, the His_6-tagged human NOT1 MIF4G domain (residues E1093-S1317) was coexpressed with a bicistronic plasmid expressing untagged CAF1 and CCR4a with an N-terminal MBP-tag cleavable by the HRV3C protease. *Hs* NOT1-CD cDNA was cloned in the pnYC-pM plasmid (Diebold et al. 2011), thereby generating a fusion protein containing an N-terminal MBP tag that is cleavable by the HRV3C protease. The cDNA encoding the NOT3-N fragment (residues A2–D212) was inserted between the XhoI and BamH1 restriction sites of the pnEA-pM vector, thus resulting in an N-terminally MBP-tagged protein.

Coimmunoprecipitation and SBP-pull-down assays

For coimmunoprecipitation assays in S2 cells, 2.5×10^6 cells were seeded per well in 6-well plates and transfected using Effectene transfection reagent (Qiagen). The transfection mixtures contained plasmids expressing GFP-tagged CCR4-NOT of subunits (2 µg) or HA-tagged Bam (1 µg). Cells were harvested 3 days after transfection, and coimmunoprecipitation assays were performed using RIPA buffer [20 mM HEPES (pH 7.6), 150 mM NaCl, 2.5 mM $MgCl_2$, 1% NP-40, 1% sodium deoxycholate supplemented with protease inhibitors (Complete protease inhibitor mix, Roche)] as previously described (Tritschler et al. 2008). For SBP pull-down assays in human cells, HEK293T cells (ATCC, wild-type or CAF40-null cells) were grown in 10-cm dishes (4×10^6 / 10-cm dish) and transfected using TurboFect transfection reagent (Thermo Fisher Scientific). The transfection mixtures contained 20 µg, 5 µg and 25 µg of plasmids expressing

Bam, MBP-CBM and Bam 4xMut, respectively. For the pull-down assays in Figure 6E and G, cells were also co-transfected with 8 µg of a plasmid expressing HA-tagged CCR4. The cells were harvested 2 days after transfection, and pull-down assays were performed as previously described (Bhandari et al. 2014).

Generation of the CAF40-null cell line

An sgRNA (sequence: 5' CCCATGCTGTGGCATTCATT 3') targeting the second exon of the *Hs* CAF40 gene was designed using CHOPCHOP (http://chopchop.cbu.uib.no) and inserted 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 pSp-CAF40-sgRNA-Cas9(BB)-2A-Puro plasmid and selected with puromycin (3 µg/ml) to obtain stable CAF40 knockout cells. To obtain clonal cell lines, single cells were distributed in 96-well plates using serial dilutions. Expansion of single-cell clones was performed under non-selective conditions. CAF40-null clones were identified by western blotting using anti-CAF40 antibodies (Supplemental Table S2). Genomic DNA from single clones was isolated using a Wizard SV Genomic DNA Purification System (Promega) and the targeted CAF40 locus was amplified by PCR and sequenced to confirm gene editing. We observed a deletion of 22nt in one allele and an insertion of one nucleotide in the second exon of CAF40 in the other allele, both of which cause a frameshift.

Protein expression, purification and competition assays

To purify the *Dm* NOT1 CN9BD-CAF40 complex, MBP-tagged NOT1-CN9BD (residues Y1468-T1719) was co-expressed with His_6-tagged CAF40 (ARM domain, residues E25–G291). The cells were lysed in a buffer containing 50 mM HEPES (pH 7.5), 300 mM NaCl, 20 mM

imidazole and 2 mM β-mercaptoethanol. The complex was purified from cleared cell lysates by Nickel affinity chromatography using a HiTrap IMAC column and eluted by a linear gradient to 500 mM imidazole. The complex was further purified on a HiTrapQ column (GE Healthcare), and this was followed by removal of the His_6 and MBP tags by cleavage with HRV3C protease overnight at 4°C. The complex was separated from the tags by size exclusion chromatography using a Superdex 200 26/600 column in a buffer containing 10 mM HEPES (pH 7.5), 200 mM NaCl and 2 mM DTT.

For competition assays, the CAF40 ARM domain was expressed with an N-terminal GST tag. The cells were lysed in a buffer containing 50 mM HEPES (pH 7.5), 300 mM NaCl and 2 mM DTT. The protein was purified from cleared cell lysates by using Protino glutathione agarose 4B (Macherey-Nagel) followed by a HiTrapQ column and further purified by size exclusion chromatography on a Superdex 200 26/600 column in a buffer containing 10 mM HEPES (pH 7.5), 200 mM NaCl and 2 mM DTT. The Roquin CBM fused to MBP was purified as previously described (Sgromo et al. 2017). Cells expressing either His_6-NusA-tagged Bam CBM or His_6-NusA were lysed in a buffer containing 50 mM potassium phosphate (pH 7.5), 300 mM NaCl, 20 mM imidazole and 2 mM β-mercaptoethanol. The proteins were isolated from the crude cell lysate by Nickel affinity chromatography using a HiTrap IMAC column and eluted by a linear gradient to 500 mM imidazole. The eluted proteins were directly applied to size exclusion chromatography on a Superdex 200 16/600 column in a buffer containing 10 mM HEPES (pH 7.5), 200 mM NaCl and 2 mM DTT.

The assembled *Hs* NOT1-10-11 trimer was obtained by co-expression of C-terminally His_6-tagged NOT11 (residues D257–D498) and untagged NOT1 (residues M1–D1000) and NOT10 (residues D25–Q707). The cells were lysed in 50 mM potassium phosphate (pH 7.6), 300 mM

NaCl, 20 mM imidazole and 2 mM β-mercaptoethanol. The complex was purified from cleared cell lysates by using a HiTrap IMAC column and eluted by a linear gradient to 500 mM imidazole. The complex was dialyzed in a buffer containing 50 mM HEPES (pH 7.5), 200 mM NaCl and 2 mM DTT, and was further purified over a HiTrap Heparin column (GE Healthcare), then subjected to size exclusion chromatography over a Superdex 200 26/600 column in a buffer containing 10 mM HEPES (pH 7.5), 200 mM NaCl and 2 mM DTT.

To purify the assembled catalytic module, His_6-tagged NOT1 MIF4G domain (residues E1093–S1317), untagged CAF1 and MBP-tagged CCR4a were co-expressed. Cells were lysed in a buffer containing 50 mM potassium phosphate (pH 7.5), 300 mM NaCl and 2 mM β-mercaptoethanol. The complex was purified from cleared cell lysates using amylose resin and eluted with in a buffer containing 50 mM potassium phosphate (pH 7.5), 300 mM NaCl, 20 mM imidazole, 25 mM D(+)-maltose and 2 mM β-mercaptoethanol. The complex was further purified using a HiTrap IMAC column (GE Healthcare) and eluted by a linear gradient to 500 mM imidazole. The His_6 and MBP tags were removed by cleavage with the HRV3C protease overnight at 4°C. The catalytic module was further purified over a Superdex200 (26/600 column; GE Healthcare) in a buffer containing 10 mM HEPES (pH 7.5), 200 mM NaCl and 2 mM DTT.

The *Hs* NOT3 N-terminus (residues A2–D212) was expressed with an N-terminal MBP tag. Cells were lysed in a buffer containing 50 mM HEPES (pH 7.5), 300 mM NaCl and 2 mM DTT. The protein was purified from cleared cell lysates with amylose resin, then with a HiTrapQ column. The MBP tag was removed by cleavage using the HRV3C protease. After cleavage of the tag, the protein was further purified on a Superdex 75 26/600 column (GE Healthcare) using a buffer containing 10 mM HEPES (pH 7.5), 200 mM NaCl and 2 mM DTT.

The purification procedures for the human NOT1-CD (residues D1607–S1815) and of the NOT and CAF40 modules have been previously described (Chen et al. 2014; Raisch et al. 2016; Sgromo et al. 2017). The NOT module comprises the NOT1-SHD (residues H1833–M2361), NOT2-C (residues M350–F540) and NOT3-C (residues L607–E748). The *Hs* CAF40 module comprises NOT1-CN9BD (residues V1351–L1588) and the CAF40 ARM domain (residues R19-E285). The *Dm* Bam CBM peptide (residues D13–E36) used for crystallization was obtained from EMC microcollections and solubilized in a buffer containing 10 mM HEPES (pH 7.5), 200 mM NaCl and 2 mM DTT.

Crystallization

Crystals of *Hs* CAF40 (ARM domain) bound to Bam CBM peptide (residues D13–E36) were obtained at 22°C using the hanging-drop vapor diffusion method after the protein solution (6 mg/ml CAF40 and 1.1 mg/ml Bam CBM peptide; 200 nl) was mixed with the crystallization reservoir solution (200 nl). Crystals appeared within one day in many conditions. Optimized crystals grew at 18°C in hanging drops consisting of 1 µl protein solution (6 mg/ml CAF40 and 1.1 mg/ml Bam CBM peptide) and 1 µl crystallization reservoir solution containing 100 mM HEPES (pH 7.0), 200 mM $CaCl_2$ and 15% PEG 6,000. Crystals were soaked in reservoir solution supplemented with 15% ethylene glycol for cryoprotection before being flash-frozen in liquid nitrogen.

Crystals of the *Hs* NOT1 CN9BD–CAF40 complex bound to the Bam CBM peptide were obtained at 22°C by using the hanging-drop vapor diffusion method after mixing the protein solution (7.5 mg/ml NOT1 CN9BD–CAF40 and 0.8 mg/ml CBM peptide; 200 nl) with the crystallization reservoir solution (200 nl). Crystals appeared within one day in several conditions.

Optimized crystals grew in drops of 200 nl protein solution (5 mg/ml NOT1 CN9BD–CAF40 complex and 0.5 mg/ml CBM peptide) mixed with 200 nl crystallization reservoir solution comprising 1.5 M ammonium sulfate, 20 mM MES (pH 6.0) and 80 mM MES (pH 6.5). Crystals were soaked in reservoir solution supplemented with 25% glycerol for cryoprotection before flash-freezing in liquid nitrogen.

Data collection and structure determination

X-ray diffraction data for the *Hs* NOT1 CN9BD–CAF40 bound to the Bam CBM were collected at a wavelength of 1.0000 Å on a PILATUS 6M detector (Dectris) at the PXII beamline of the Swiss Light Source (SLS) and processed in space group $P3_221$ by using XDS and XSCALE (Kabsch 2010) to a resolution of 2.7 Å, aiming at a CC(1/2) value (Karplus and Diederichs 2012) of ~70 % as a high resolution cutoff. Initial phases were determined by molecular replacement, with two copies of the NOT1 CN9BD–CAF40 complex (PDB 4CRU) used as a search model in PHASER (McCoy et al. 2007) from the CCP4 package (Winn et al. 2011). The initial model was improved and completed by iterative cycles of building in COOT (Emsley et al. 2010) and refinement in PHENIX (Afonine et al. 2012), also optimizing TLS parameters (one TLS group per macromolecular chain). Finally, two copies of the Bam CBM peptide were manually built into the density (Supplemental Fig. S6A) and improved by further refinement cycles.

The best crystal of the CAF40 (ARM domain) bound to the Bam CBM peptide was recorded at a wavelength of 1.0396 Å on a PILATUS 6M fast detector (DECTRIS) at the DESY beamline P11. The dataset was processed in XDS and XSCALE in space group $P2_12_12$ to a resolution of 3.0 Å, aiming at a CC(1/2) value (Karplus and Diederichs 2012) of ~70 % as a high resolution cutoff. Four copies of the CAF40 ARM domain (PDB 2FV2, chain A) were found in the

asymmetric unit by molecular replacement using PHASER from the CCP4 package. This initial model was improved and completed by iterative cycles of building in COOT (Emsley et al. 2010) and refinement using PHENIX (Afonine et al. 2012) and BUSTER (Bricogne et al. 2011) using NCS restraints and TLS parameters (one TLS group per macromolecular chain). Finally, four copies of the Bam CBM peptide were manually built into the density (Supplemental Fig. S6B) and improved through further refinement cycles.

The stereochemical properties for all of the structures were verified with MOLPROBITY (Chen et al. 2010), and illustrations were prepared using PyMOL (http://www.pymol.org). The diffraction data and refinement statistics are summarized in Table 1.

Isothermal titration calorimetry (ITC)

The ITC experiments were performed on a VP-ITC microcalorimeter (Microcal) at 20°C. A solution containing the *Dm* NOT1 CN9BD bound to CAF40 (ARM domain) (6.0 μM in the experiments with the Bam CBM and up to 10 μM in the experiments with the Roq CBM) in a calorimetric cell was titrated with a solution of MBP-tagged Bam CBM (60 μM) or MBP-tagged Roquin CBM (up to 100 μM). All proteins were dissolved in a buffer containing 10 mM HEPES (pH 7.5), 200 mM NaCl and 0.5 mM TCEP. The titration experiments consisted of an initial injection of 2 μl followed by 28 injections of 10 μl at 240 s intervals. The binding experiment was repeated three times. The thermodynamic parameters were calculated using a one-site binding model (ORIGIN version 7.0; Microcal). Correction for dilution heating and mixing was achieved by subtracting the final baseline, which consisted of small peaks of similar size. The first injection was removed from the analysis (Mizoue and Tellinghuisen 2004).

Bioinformatic analysis

To identify proteins featuring potential CBMs in *Dm* and *Hs*, we followed a two-step approach. In the first step, we searched for homologs of *Dm* Bam and Roquin in the nonredundant (nr) protein sequence database using PSI-BLAST (Boratyn et al. 2013), as implemented in the MPI Bioinformatics toolkit (Alva et al. 2016), and extracted the CBMs from the obtained homologs originating from different *Drosophila* species. These motifs were then aligned, and a consensus pattern was derived by manual inspection (x-x-x-[LI]-[DENQ]-x(2,3)-[FLM]-x-x-[ILM]-x-x-x-[IL]-x-x-[ILM]-[LIV]-x-x-x-x). In the second step, the aforementioned consensus pattern was submitted to the PatternSearch tool of the MPI Bioinformatics Toolkit to identify proteins in *Dm* and *Hs* with potential CBMs. This search yielded a total of 1,200 candidate proteins. We next analyzed this set further to discard all proteins in which the detected motifs showed no helical propensity or were embedded within a domain (as opposed to being embedded in an intrinsically disordered region). We also excluded all proteins with obvious functional irrelevance (e.g. membrane proteins) from further consideration. Finally, we chose the *Hs* and *Dm* homologs of four protein families, on the basis of the presence of known or predicted RNA-binding domains in the proteins and on the percentage similarity of the putative CBMs to the CBMs of Bam and Roquin. These candidate CBMs were then expressed as MBP fusions and tested for CAF40 binding in *in vitro* MBP pull-down assays.

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.

Alva V, Nam SZ, Soding J, Lupas AN. 2016. The MPI bioinformatics Toolkit as an integrative platform for advanced protein sequence and structure analysis. *Nucleic Acids Res* **44:** W410-415.

Bawankar P, Loh B, Wohlbold L, Schmidt S, Izaurralde E. 2013. NOT10 and C2orf29/NOT11 form a conserved module of the CCR4-NOT complex that docks onto the NOT1 N-terminal domain. *RNA Biol* **10:** 228-244.

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.

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

Boland A, Chen Y, Raisch T, Jonas S, Kuzuoğlu-Öztürk D, Wohlbold L, Weichenrieder O, Izaurralde E. 2013. Structure and assembly of the NOT module of the human CCR4-NOT complex. *Nat Struct Mol Biol* **20:** 1289-1297.

Boratyn GM, Camacho C, Cooper PS, Coulouris G, Fong A, Ma N, Madden TL, Matten WT, McGinnis SD, Merezhuk Y. et al. 2013. BLAST: a more efficient report with usability improvements. *Nucleic Acids Res* **41:** W29-33.

Braun JE, Huntzinger E, Fauser M, Izaurralde E. 2011. GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets. *Mol Cell* **44:** 120-33.

Bricogne G, Blanc E, Brandl M, Flensburg C, Keller P, Paciorek W, Roversi P, Sharff A, Smart O, Vonrhein C. et al. 2011. BUSTER. *Cambridge, United Kingdom: Global Phasing Ltd.*

Chen VB, Arendall WB3rd, 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.

Chen Y, Boland A, Kuzuoğlu-Öztürk D, Bawankar P, Loh B, Chang CT, Weichenrieder O, Izaurralde E. 2014a. A DDX6-CNOT1 complex and W-binding pockets in CNOT9 reveal direct links between miRNA target recognition and silencing. *Mol Cell* **54:** 737-750.

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.

Garces RG, Gillon W, Pai EF. 2007. Atomic model of human Rcd-1 reveals an armadillo-like-repeat protein with in vitro nucleic acid binding properties. *Protein Sci* **16:** 176-188.

Jonas S, Weichenrieder O, 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.

Kabsch W. 2010. Xds. *Acta Crystallogr D Biol Crystallogr* **66:** 125-132.

Karplus PA, Diederichs K. 2012. Linking crystallographic model and data quality. *Science* **336:** 1030-1033.

Lykke-Andersen J, Shu MD, Steitz JA. 2000. Human Upf proteins target an mRNA for nonsense-mediated decay when bound downstream of a termination codon. *Cell* **103:** 1121-1131.

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.

Petit AP, Wohlbold L, Bawankar P, Huntzinger E, Schmidt S, Izaurralde E, Weichenrieder O. 2012. The structural basis for the interaction between the CAF1 nuclease and the NOT1 scaffold of the human CCR4-NOT deadenylase complex. *Nucleic Acids Res* **40:** 11058-11072.

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

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.

Rehwinkel J, Behm-Ansmant I, Gatfield D, Izaurralde E. 2005. A crucial role for GW182 and the DCP1:DCP2 decapping complex in miRNA-mediated gene silencing. *RNA* **11:** 1640-1647.

Sgromo A, Raisch T, Bawankar P, Bhandari D, Chen Y, Kuzuoğlu-Öztürk D, Weichenrieder O, Izaurralde E. 2017. A CAF40-binding motif facilitates recruitment of the CCR4-NOT complex to mRNAs targeted by Drosophila Roquin. *Nat Commun* **8:** 14307.

Tritschler F, Eulalio A, Truffault V, Hartmann MD, Helms S, Schmidt S, Cole ,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.

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.

Valdar WS. 2002. Scoring residue conservation. *Proteins* **48:** 227-241.

Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A. et al. 2011. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67:** 235-242.

Zekri L, Kuzuoğlu-Öztürk D, Izaurralde E. 2013. GW182 proteins cause PABP dissociation from silenced miRNA targets in the absence of deadenylation. *EMBO J* **32:** 1052-1065.