Drosophila melanogaster Roquin and Bam share a CAF40 binding motif to recruit the CCR4-NOT deadenylase complex and repress target mRNAs

DISSERTATION

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von Annamaria Sgromo

aus Lamezia Terme, Italien

Tübingen

2018

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

Tag der mündlichen Qualifikation:	25.07.2018
Dekan:	Professor Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Professor Dr. Ralf-Peter Jansen
2. Berichterstatter:	Professor Dr. Thilo Stehle

This thesis describes work conducted in the laboratory of Prof. Dr. Elisa Izaurralde in the Department of Biochemistry at the Max Planck Institute for Developmental Biology, Tübingen, Germany, from February 2013 to April 2018. The work was supervised by Prof. Dr. Ralf-Peter Jansen at the Eberhard Karls University Tübingen, Germany, and was supported by a fellowship from the International PhD program of the MPI for Developmental Biology. 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 carried out by my colleagues was also indicated accordingly.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I want to thank:

Prof. Dr. Elisa Izaurralde for her mentoring, support and guidance during my PhD.

Prof. Dr. Ralf-Peter Jansen at the Eberhard Karls University Tübingen for agreeing to be my supervisor.

Prof. Dr. Thilo Stehle at the Eberhard Karls University Tübingen for agreeing to evaluate my thesis.

Associate Prof. Dr. Fulvia Bono and Prof. Dr. Gáspár Jékely at the University of Exeter as members of my thesis advisory committee.

Dr. Oliver Weichenrieder for the interesting discussions and collaborations within the years. Within these years I also had the chance to join a trip to the synchrotron with him – it was a unique experience! I also want to thank him for commenting on the manuscript of this thesis.

Dr. Dagmar Sigurdardottir as PhD program coordinator for her support throughout my PhD.

Dr. Heike Budde for her constant help and technical support within the years also concerning student supervision.

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

Sigrun Helms and Maria Fauser for introducing me to the lab and for the great time we had in the lab! Special thanks to Suzi for the nice stories I heard during these years and for the "donation" of her whisker that helped me crystalize the Roquin CBM.

Dr. Duygu Kuzuoglu-Öztörk, Dr. Catia Igreja and Akin Öztürk for their precious friendship that got me through the good and rough times I had during these years.

Very special thanks to Aoife Hanet, I know I can always count on her! I will always have very nice memories of the holidays and festivals we went to. I will miss you but I am sure we will continue travelling together!

Dr. Tobias Raisch not only for the great collaborations with the Roquin, Bam and NOT4 projects but also and especially for his friendship. We really had a good time! I want to thank him also for helping me with Figure 20 containing the CBM structures in this thesis.

Many thanks to Charlotte Backhaus, Kevin Sabath, Andreas Blaha, and Matthias Becker for being great students, I hope you had a good time in the lab and enjoyed working at the bench.

Dr. Praveen Bawankar and Dr. Vikram Alva for the very nice collaboration with the Roquin and the Bam projects, respectively.

The former and current members of Department II for the nice working atmosphere: Dr. Valentina Ahl, Stefanie Becker, Dr. Dipankar Bhandari, Dr. Andreas Boland, Regina Butter, Dr. Chung-Te Chang, Dr. Ying Chen, Dr. Mary Christie, Min-Yi Chung, Maria Fernandez, Stefan Grüner, Dr. Stefanie Jonas, Csilla Keskeny, Dr. Elena Khazina, Simone Larivera, Yevgen Levdansky, Dr. Zita Liutkeviciute, Dr. Belinda Loh, Sowndarya Muthukumar, Michelle Noble, Felix Räsch, Vincenzo

Ruscica, Dr. Steffen Schmidt, Dr. Anna Schneider, Ayshwarya Seenivasan, Alina Stein, Dr. Eugene Valkov, Gabriele Wagner, Ramona Weber, Catrin Weiler, Dr. Lara Wohlbold and Dr. Latifa Zekri.

Ida Axtmann, Marie Palaj, Marion Weber, Nadine Weiss, Max Widmann for their technical help in the lab.

Dr. Antonio Strillacci for guidance during my studies.

David Mörsdorf for his precious support and understanding. I cannot thank him enough for everything he did and he is doing for me!

Above all, I want to thank all my friends and my big family for their constant and endless support. Especially, I want to express my deepest gratitude to my mother and my grandparents without whom none of this work would have been possible. GRAZIE!

The memories I have of Tübingen and of the MPI for Developmental Biology as well as the people I met will remain precious to me!

1.1 SUMMARY

Messenger RNA (mRNA) encodes the information for cellular protein biosynthesis. To maintain correct protein levels, which is essential for cell function, strict control over mRNA levels is crucial. The CCR4-NOT complex is a key regulator of gene expression and uses transcriptional as well as posttranscriptional mechanisms to fine-tune the expression of mRNAs in diverse biological contexts. Its best characterized function is to trigger cytoplasmic mRNA decay: The complex deadenylates bulk and targeted mRNAs, which ultimately leads to their degradation.

Over the past years, research in the RNA decay field has focused on understanding how the multisubunit CCR4-NOT complex assembles, and on identifying regulatory proteins that specifically modulate the repressive activity of the complex towards individual transcripts. RNA-associated proteins that recruit the CCR4-NOT complex to mRNA targets provide a unique opportunity to specifically control mRNA translation and decay. However, no general principle underlying this recruitment has been found. In fact, it appears that every RNA-associated protein uses an individual mode to interact with the different subunits of the complex. Thus, to understand the function of these regulatory proteins, it is crucial to reveal the molecular mechanisms they use to recruit CCR4-NOT.

During my doctoral studies, I aimed to shed light on the recruitment of the CCR4-NOT complex by RNA-associated proteins to target mRNAs. I characterized at the molecular level how the *Drosophila melanogaster (Dm)* proteins Roquin and Bag-of-marbles (Bam) interact with the CCR4-NOT complex to induce repression of their targets. Roquin carries at least two distinct motifs, which both contribute to a direct interaction with the CCR4-NOT complex. These are a NOT module binding motif (NBM) and a CAF40 binding motif (CBM_R; where the subscript R refers to Roquin). In contrast to Roquin, Bam harbors only one motif, a CBM_B, that directly recruits the fully assembled CCR4-NOT complex. Subsequently, my work on Roquin and Bam led to the identification of a CBM_N in *Dm* NOT4, a protein that is a stable subunit of CCR4-NOT in yeast.

In conclusion, my studies identified a previously unknown peptide-binding surface on the CAF40 subunit of the CCR4-NOT complex, which is bound by at least three distinct proteins via their CBMs, suggesting mutually exclusive binding. Surprisingly, Roquin, Bam and NOT4 CBMs show a similar way of interacting with CAF40 despite the lack of sequence conservation. This implies convergent evolution of the CBMs and suggests that CAF40 provides a binding platform within the CCR4-NOT complex for additional, yet unidentified proteins. The CBM-containing proteins presented in this study add to a growing body of peptide-mediated associations in the highly complex network of CCR4-NOT-mediated mRNA regulation.

1.2 ZUSAMMENFASSUNG

Messenger RNA (mRNA) enthält die kodierte Information für die zelluläre Proteinbiosynthese. Für die korrekte Ausführung vieler zellulärer Aufgaben ist es unabdingbar, korrekte Proteinkonzentrationen zu gewährleisten, was wiederum eine strenge Kontrolle der mRNA-Konzentrationen voraussetzt. Eine Schlüsselrolle bei der Kontrolle von mRNA Konzentrationen hat der CCR4-NOT-Komplex, einer der wichtigsten Regulatoren der Genexpression auf der transkriptionellen sowie der post-transkriptionellen Ebene. Die am besten charakterisierte Funktion des CCR4-NOT-Komplexes ist die allgemeine oder gezielt ausgelöste Deadenylierung von cytosolischen mRNAs, die letztendlich zu deren komplettem Abbau führt.

In den vergangenen Jahren hat sich die Erforschung des mRNA-Abbaus verstärkt damit befasst, zu verstehen wie sich der CCR4-NOT-Multiproteinkomplex aufbaut, und regulatorische Proteine zu identifizieren, welche den Komplex dahingehend modulieren, spezifische Transkripte zu reprimieren. RNA-assoziierte Proteine, die den CCR4-NOT-Komplex zu spezifischen mRNAs rekrutieren, stellen eine einzigartige Möglichkeit dar, um die Stabilität von einzelnen mRNAs gezielt zu kontrollieren. Ein genereller Mechanismus, der dieser Rekrutierung des CCR4-NOT-Komplexes unterliegt, ist allerdings nicht bekannt. Vielmehr scheint jedes RNA-assoziierte Protein die Untereinheiten des Komplexes über individuelle Interaktionen zu kontaktieren. Um die exakte Funktion dieser regulatorischen Proteine zu verstehen ist es somit nötig den molekularen Mechanismus der CCR4-NOT Rekrutierung zu charakterisieren.

Das Ziel meiner Doktorarbeit war es, zu verstehen, wie der CCR4-NOT-Komplex durch RNA-assoziierte Proteine zu Ziel-mRNAs gebracht wird. Ich habe auf molekularer Ebene charakterisiert, wie die *Drosophila melanogaster* (*Dm*) Proteine Roquin und Bag-of-marbles (Bam) mit dem CCR4-NOT-Komplex interagieren, um ihre Ziel-mRNAs zu reprimieren. Roquin enthält mindestens zwei Peptidmotive die beide zu einer direkten Proteininteraktion mit dem CCR4-NOT-Komplex beitragen. Diese sind ein NOT-Modul-bindendes Motiv (NBM) und ein CAF40-bindendes Motiv (CBM_R; wobei das tiefgestellte R für Roquin steht). Im Gegensatz zu Roquin enthält Bam nur ein Motiv (CBM_B) das direkt an CAF40 bindet um den kompletten CCR4-NOT Proteinkomplex zu rekrutieren. Meine Arbeit an Roquin und Bam führte auch zur Identifikation eines CBM_N in *Dm* NOT4, einem Protein das in der Hefe eine stabile Untereinheit des CCR4-NOT-Komplexes ist.

Schlussendlich identifizierte meine Arbeit eine bisher unbekannte Peptidbindestelle auf der CAF40 Untereinheit des CCR4-NOT-Komplexes, die von den CBMs von mindestens drei verschiedenen Proteinen genutzt wird und wobei sich die einzelnen Wechselwirkungen wohl gegenseitig

ausschließen. Überraschenderweise ähneln sich die Interaktionen der CBMs von Roquin, Bam und NOT4 mit CAF40 trotz fehlender Sequenzkonservierung. Dies lässt auf eine konvergente Evolution der CBMs schließen und darauf, dass die Bindetasche in CAF40 auch von weiteren, bislang noch nicht identifizierten Proteinen genutzt werden könnte. Die in dieser Arbeit präsentierten CBM-enthaltenden Proteine gesellen sich zu einer wachsenden Anzahl peptidvermittelter Assoziationen im hochkomplexen Netzwerk der von CCR4-NOT gesteuerten mRNA Regulation.

2. INTRODUCTION

Gene expression is a tightly controlled process that determines when, where and how much of a gene product is produced in a cell. Often gene products are proteins, however, genes can also encode for functional RNAs, including ribosomal RNAs (rRNAs), transfer RNAs (tRNAs) or small nuclear RNAs (snRNAs). The central dogma of molecular biology explains how genetic information, encoded in DNA, is transcribed into RNA by a process called transcription. The RNA can then be decoded into proteins in a process called translation. In eukaryotes, DNA is first transcribed into a precursor messenger RNA (pre-mRNA) by RNA polymerase II in the nucleus. The pre-mRNA is then processed into a mature messenger RNA (mRNA).

RNA processing includes co-transcriptional capping (addition of a 7-methyl guanosine cap structure at the 5'-end), removal of introns by splicing, and 3'-end cleavage and polyadenylation. The added 5'-cap structure and the poly(A) tail at the 3'-end protect the mRNA from nuclease digestion (Shatkin & Manley 2000). Subsequently, the mRNA is transported through the nuclear pore complex to the cytoplasm where it is translated into a functional protein by ribosomes, and eventually degraded. mRNA can also be localized to specific subcellular locations, a particularly important feature in neurons and embryogenesis, where spatially controlled protein synthesis is crucial (Parton et al. 2014). During their entire cellular lifespan, mRNAs are coated with a variety of proteins, forming messenger ribonucleoprotein particles (mRNPs) (Rissland 2017). mRNP composition dictates every phase of the mRNA life (Mitchell & Parker 2014).

The amount of proteins within a cell is controlled at both transcriptional and posttranscriptional levels to allow cells to rapidly adapt to various stimuli and environmental changes. mRNA degradation ensures a rapid regulation of transcript levels and is interconnected with translation. Both processes, together with transcription, have to be balanced to maintain cell function and survival (Moore 2005). An overview of gene regulation mechanisms is shown in Figure 1.

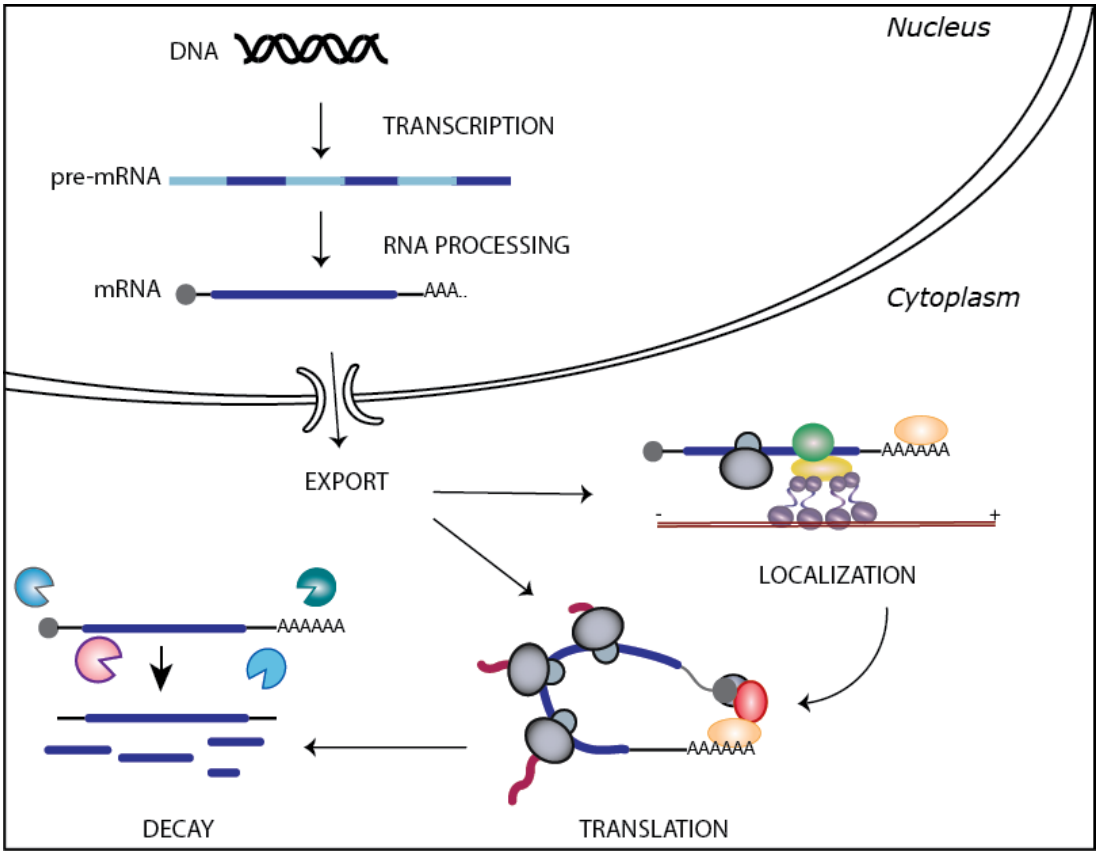


Figure 1: Schematic overview of eukaryotic gene expression. DNA is first transcribed into a pre-mRNA that is processed in the nucleus and then exported to the cytoplasm. Ribosomes associate with the mRNA to translate it into a functional protein. Prior to translation, the mRNA can be localized within the cell. Subsequently mRNA stability is controlled by degradation mechanisms.

2.1 Eukaryotic mRNA translation

Translation is the process by which an mRNA is translated into proteins by the ribosomes and it involves three steps: initiation, elongation and termination. The presence of the cap structure and the poly(A) tail stimulates the initiation of protein synthesis and consequently the amount of protein that is produced from the mRNA.

During cap-dependent translation, the cap structure is recognized by the eukaryotic translation initiation factor 4F multisubunit complex (eIF4F; containing the scaffold protein eIF4G, the cap-binding protein eIF4E and the RNA helicase eIF4A) (Marcotrigiano et al. 1997, Sonenberg & Dever 2003). The eIF4F complex (via eIF4G) in turn interacts with the cytoplasmic poly(A) binding protein (PABP) that binds to the poly(A) tail of the mRNA (Tarun & Sachs 1996, Imataka et al. 1998). Thus, the interaction between eIF4G and PABP induces the closed loop mRNA conformation, promoting mRNA stability and translation initiation (Wells et al. 1998, Amrani et al. 2008). Conversely, removal of the poly(A) tail and the 5'-cap structure destabilizes mRNAs leading

to their degradation (Coller & Parker 2004, Meyer et al. 2004). Notably, poly(A) tail length generally does not correlate linearly with translation efficiency and stability of mRNAs. In fact, recently it was shown that eukaryotic mRNAs with a short, "pruned" poly(A) tail (less than 60 nt and thus able to accommodate 1 to 2 PABP molecules) are more abundant and translated efficiently (Lima et al. 2017).

2.2 Eukaryotic mRNA turnover

Multiple mRNA decay pathways ensure a tight surveillance of gene expression by eliminating the templates for protein production in a cell (Houseley & Tollervey 2009). mRNA decay is initiated by the shortening of the poly(A) tail at the 3'-end - a process known as deadenylation - which results in the release of PABP proteins and disruption of the closed loop conformation (Garneau et al. 2007, Parker & Sheth 2007, Wahle & Winkler 2013). In the canonical 5'-to-3' decay pathway, deadenylation is followed by decapping, where the 5'-cap structure is removed from the mRNA by the DCP1/DCP2 decapping complex. Several accessory factors are required for efficient decapping, such as the (L)Sm (Sm and Like-Sm) protein EDC3 (enhancer of decapping 3), the DEAD-box RNA helicase DDX6 (also known as RCK/Me31B/Dhh1) and PATL1 (Coller et al. 2001, Harigaya et al. 2010, Marnef & Standart 2010). Decapping is an irreversible step that exposes bulk mRNA to complete degradation by the 5'-to-3' exoribonuclease XRN1 (Chang et al. 2011, Jones et al. 2012, Arribas-Layton et al. 2013). Alternatively, deadenylated mRNAs can be degraded by the cytoplasmic RNA exosome in the 3'-to-5' direction (Schmid & Jensen 2008). An illustration of eukaryotic mRNA turnover is shown in Figure 2.

Deadenylation is often the rate-limiting step for mRNA decay and translational repression (Parker & Song 2004). There are three main deadenylases: the poly(A)-specific ribonuclease PARN (which is not present in budding yeast and *Drosophila*) (Harnisch et al. 2012, Virtanen et al. 2013), the PAN2-PAN3 complex (Wahle & Winkler 2013, Wolf & Passmore 2014) and the CCR4-NOT complex (Goldstrohm & Wickens 2008). PARN belongs to the DEDD (Asp-Glu-Asp-Asp) family of nucleases and is not involved in bulk mRNA deadenylation (Yoda et al. 2013) but rather acts on specific substrates and upon stress (Godwin et al. 2013). The PAN2-PAN3 complex, with PAN2 as a catalytic subunit, is stimulated by PABP proteins and is involved in general mRNA decay. It acts at the earliest stages of deadenylation in a distributive manner. Ultimately, the CCR4-NOT complex removes the remaining poly(A) tail in a processive manner (Yamashita et al. 2005, Chen et al. 2017). Importantly, the CCR4-NOT complex deadenylates mRNAs even in the absence of the PAN2-PAN3 complex (Tucker et al. 2001).

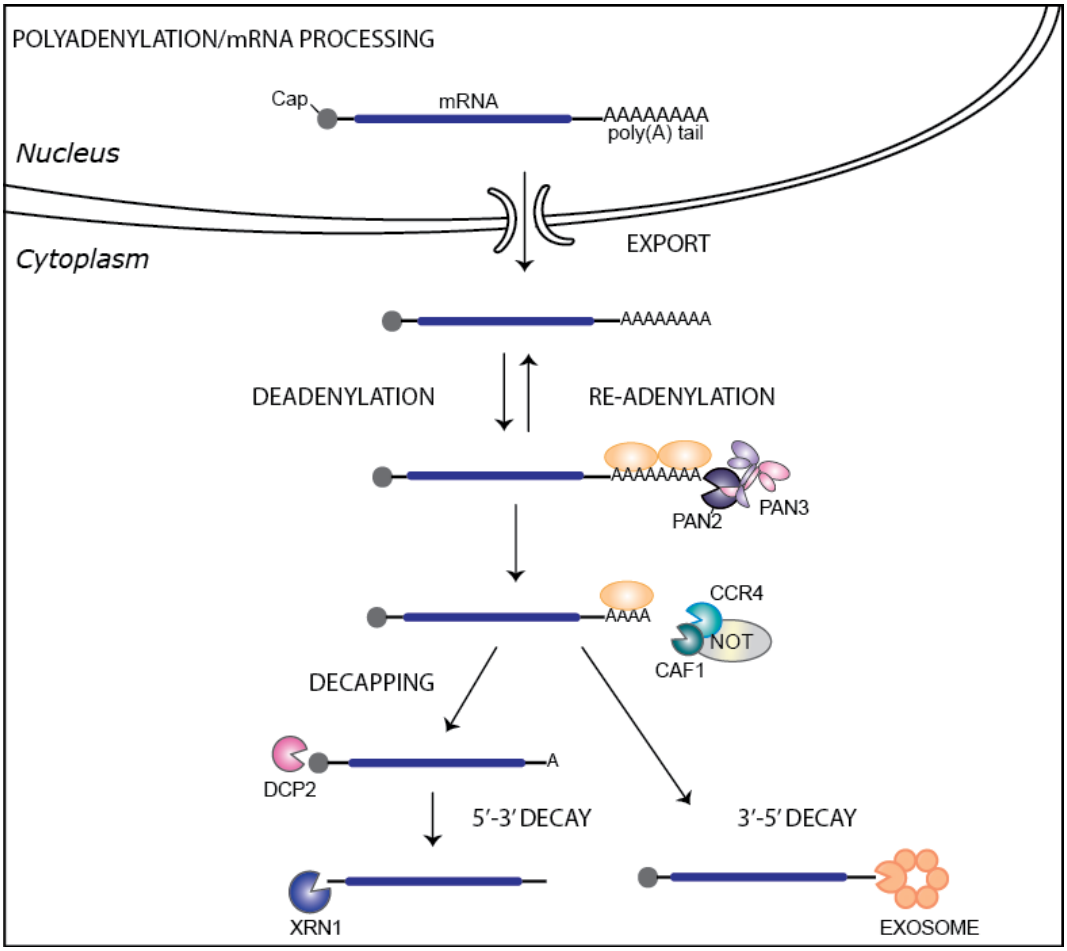


Figure 2: Mechanisms of eukaryotic mRNA turnover. mRNA decay is initiated by the removal of the poly(A) tail by the two major cytoplasmic deadenylases: the PAN2-PAN3 and CCR4-NOT complexes. Deadenylated mRNA can be decapped and degraded via the exonucleolityc activity of XRN1 from the 5'-to-3' end. Alternatively, mRNA can be degraded 3'-to-5' by the exosome. Additionally, derepression of mRNAs can occur through re-adenylation.

2.3 The CCR4-NOT deadenylase complex

2.3.1 CCR4-NOT architecture

The CCR4-NOT complex was first discovered in budding yeast as a negative transcriptional regulator for TATA-less promoters, hence the name NOT – negative on TATA-less (Collart & Struhl 1994, Oberholzer & Collart 1998, Bai et al. 1999). However, further studies highlighted its central role as cytoplasmic deadenylase (Tucker et al. 2002). In yeast, the complex consists of nine subunits, the NOT1-5 proteins, CCR4 (carbon catabolite repressor 4), and three CCR4-associated factors, CAF1, CAF40, and CAF130. In metazoans, all of the nine subunits have orthologues apart from CAF130. Metazoans encode NOT3 as a functional homolog of both NOT5 and NOT3, and they have two orthologues for CAF1 known as CNOT7 and CNOT8 (also known as POP2) and two

orthologues for CCR4 known as CNOT6 (or CCR4a) and CNOT6L (or CCR4b) that bind to the complex in a mutually exclusive manner. Additionally, metazoan NOT10 and NOT11 are absent in yeast (Albert et al. 2000, Bawankar et al. 2013, Mauxion et al. 2013). Finally, NOT4 does not appear to be stably associated with CCR4-NOT in metazoans, although it is an integral component of the complex in yeast (Albert et al. 2000, Wahle & Winkler 2013, Bhaskar et al. 2015).

Within the CCR4-NOT deadenylase complex NOT1 is the largest protein and acts as a scaffold. Its depletion in cells induces cell death and impairs deadenylation activity, revealing the crucial role of NOT1 for the assembly and function of the complex (Ito et al. 2011). Multiple biochemical and structural studies have provided snapshots of subcomplexes of the CCR4-NOT complex (Basquin et al. 2012, Collart & Panasenko 2012, Petit et al. 2012, Bawankar et al. 2013, Boland et al. 2013, Bhandari et al. 2014, Mathys et al. 2014, Xu et al. 2014). The CCR4-NOT complex exhibits four distinct structural and functional modules that bind to the α-helical NOT1 scaffold subunit, described below and shown in Figure 3.

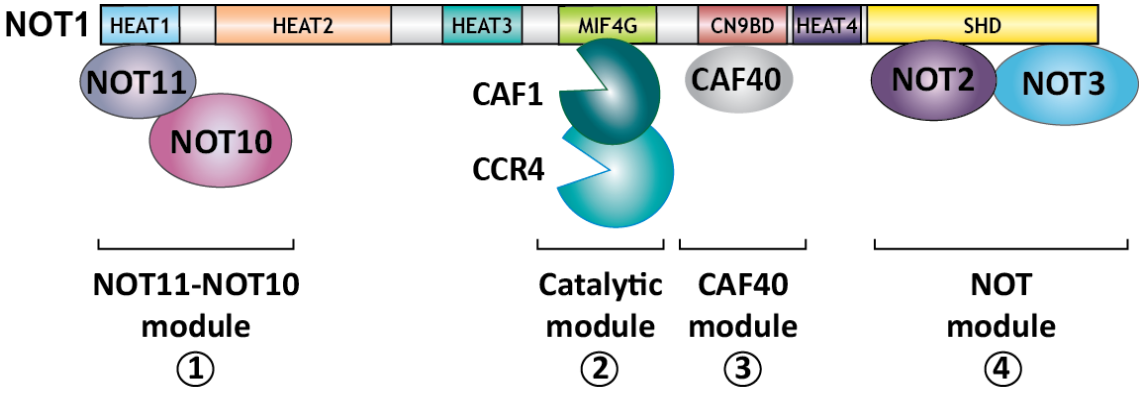


Figure 3: Architecture of the CCR4-NOT deadenylase complex in metazoans. Four distinct structural and functional modules assemble around the NOT1 scaffold protein and characterize the CCR4-NOT deadenylase complex. Adapted from Jonas & Izaurralde (2015).

① First, the NOT11-NOT10 module comprises the NOT11 and NOT10 subunits that bind to the HEAT1 repeat of NOT1 (Bawankar et al. 2013, Mauxion et al. 2013). Structurally, NOT11 has a domain of unknown function DUF2363 in its C-terminal region and NOT10 has a TPR structure. This module is not conserved in all species and is not required for mRNA deadenylation (Bawankar et al. 2013). Currently, the function of the NOT11-NOT10 module remains unknown. ② Second, the catalytic module includes the MIF4G domain (structurally similar to the middle domain of eIF4G) of NOT1 and the two deadenylases CCR4 and CAF1. The exonuclease-endonuclease-phosphatase (EEP) CCR4 uses its N-terminal LRR domain (leucine-rich repeat) to contact CAF1 (a member of the DEDD family of nucleases). CAF1, in turn, binds to the MIF4G domain of NOT1

(Dupressoir et al. 2001, Basquin et al. 2012, Petit et al. 2012). As the name suggests, this module is the catalytic region of the CCR4-NOT complex and its activity is essential for mRNA deadenylation (Tucker et al. 2002). However, it is unclear why the presence of two exonucleases is required. ③ Third, the CAF40 module consists of the armadillo repeat (ARM) of CAF40 (also known as Rcd1/RQCD1/CNOT9) bound to the three helix bundle CN9BD (CNOT9 binding domain) of NOT1 (Chen et al. 2014b, Mathys et al. 2014). The CAF40 concave surface is rich in positively charged residues that bind to nucleic acids *in vitro* (Garces et al. 2007). ④ Fourth, the NOT module is composed of the NOT2-NOT3 subunits that bind to the C-terminal superfamily homology domain (SHD) of NOT1 (Bhaskar et al. 2013, Boland et al. 2013). The C-terminal regions of NOT2 and NOT3 share conserved NOT-box domains that are involved in the interaction with the NOT1-SHD (Zwartjes et al. 2004). A schematic domain organization of the CCR4-NOT subunits is illustrated in Figure 4 and their binding sites on the NOT1 scaffold protein as well as their conservation between species are listed in Table 1.

Importantly, some of the modules discussed above provide a versatile binding platform for the interaction with RNA-binding proteins that have been shown to bring the deadenylase complex in contact with their target mRNAs. An overview of the recruitment of the CCR4-NOT complex mediated by RNA-binding proteins will follow in section 2.4.

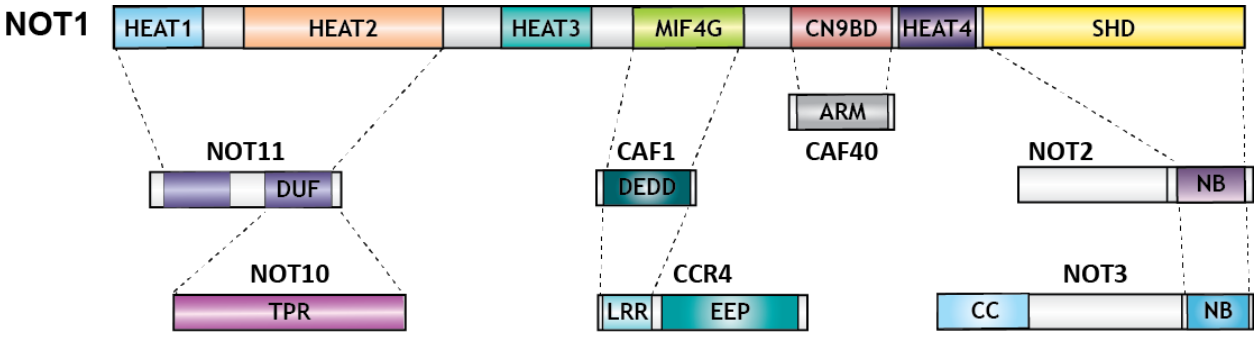


Figure 4: Domain organization of the basic subunits of the CCR4-NOT deadenylase complex. Shown are the subunits of the CCR4-NOT complex as in Figure 3 including their domain organization.

Table 1: CCR4-NOT complex subunits. Note that NOT4 has been identified as a stable subunit of the CCR4-NOT complex only in yeast. Adapted from Collart & Panasenko (2017).

S. cerevisiae	***D. melanogaster***	***H. sapiens***	**Binding regions on NOT1**
NOT1	NOT1	CNOT1	
NOT2	NOT2	CNOT2	SHD
NOT3	NOT3/5	CNOT3	SHD
NOT4	(NOT4)	(CNOT4)	
NOT5		-	SHD
CCR4	CCR4/TWIN	CNOT6/CCR4a CNOT6L/CCR4b	
CAF1/POP2	CAF1	CNOT7/CAF1a CNOT8/POP2/CAF1b	MIF4G
CAF40	CAF40	CNOT9/Rcd1/RQCD1	CN9BD
CAF130	-	-	unknown
-	NOT10	CNOT10	N-terminal region of NOT1
-	NOT11	CNOT11 (C2orf29)	N-terminal region of NOT1

2.3.2 CCR4-NOT functions

Numerous studies have highlighted the role of CCR4-NOT as a multifunctional complex that controls gene expression at multiple levels within the cell.

NOT proteins were originally identified as transcriptional repressors in yeast (Collart & Struhl 1994) and the CCR4-NOT complex has, since then, been shown to act at different levels of transcription. For example, several components of CCR4-NOT bind transcription factors and repress transcription initiation (Haas et al. 2004, Winkler et al. 2006). Additionally, the CCR4-NOT complex interacts with RNA polymerase II and was suggested to function in transcription elongation (Kruk et al. 2011). Nuclear functions of the yeast CCR4-NOT complex also include mRNA export and nuclear RNA quality control: Multiple subunits of the CCR4-NOT complex bind to nuclear pore complex proteins (Kerr et al. 2011) as well as to the nuclear exosome and the TRAMP complex (Azzouz et al. 2009). However, many of the above-mentioned functions were

described only in yeast and are still controversial. Thus, their relevance in metazoans remains elusive.

The most prominent and conserved function of the CCR4-NOT complex is deadenylation, the first step of mRNA decay that is performed by the CCR4 and CAF1 enzymes. The CCR4-NOT complex is widely conserved as a deadenylase for bulk mRNA turnover and it has also been shown to deadenylate aberrant transcripts that undergo nonsense-mediated decay (NMD) (Loh et al. 2013). Furthermore, targeted deadenyation by the CCR4-NOT complex is used to repress a plethora of target mRNAs by triggering their degradation in many physiological contexts such as immune response and germline stem cell (GSC) maintenance and differentiation. Targeting can occur via specific micro RNAs (miRNAs) or mRNA-binding proteins (see section 2.4). Indeed, the CCR4-NOT complex is a central component of the miRNA-mediated mRNA decay pathway (Braun et al. 2011, Chekulaeva et al. 2011, Zekri et al. 2013) and is recruited to miRNA by proteins of the TNRC6 family (GW182 in *Dm*). Importantly, the CCR4-NOT complex, independently of deadenylation, can also repress translation via the interaction with the translational repressor and decapping activator DDX6 (Kuzuoglu-Ozturk et al. 2016). However, the mechanisms underlying this process are yet unknown.

In addition to its role in deadenylation and translational repression, the yeast CCR4-NOT complex can also ubiquitinate via the E3 ubiquitin ligase NOT4 (Albert et al. 2000). Ubiquitination by NOT4 was extensively studied in yeast, where NOT4 was shown to interact with the E2 enzyme UbcH5B (Albert et al. 2002, Winkler et al. 2004). Yeast NOT4 ubiquitinates a wide range of substrates, like the small ribosomal protein Rps7A (Collart & Panasenko 2012), the nascent polypeptide associated complex (NAC) (Panasenko et al. 2006) and the transcription factor YAP1 (Gulshan et al. 2012). Even though the exact function of NOT4 is currently unknown, it is involved in the assembly of the proteasome and in co-translational quality control (Panasenko & Collart 2011, Collart & Panasenko 2012, Panasenko 2014). These findings imply functions of the yeast CCR4-NOT complex also in protein degradation.

Function of the CCR4-NOT complex in immune response

A tight control of mRNA expression and thus protein levels in the immune system is fundamental for efficient immune and inflammatory responses. Via the interaction with RNA-binding proteins (RBPs) the CCR4-NOT deadenylase complex is a crucial effector in regulating inflammatory mRNAs, such as *tumor necrosis factor alpha* (*TNF-α*), *Interleukin-1* (*IL-1*), *IL-6, IL-8, IL-10, cyclooxygenase-2,* and *interferon-γ (IFN-γ)* (Sanduja et al. 2012, Leppek et al. 2013, Carpenter et al. 2014, Chapat & Corbo 2014). RBPs recognize multiple elements within the 3’ untranslated region (UTR) of inflammatory mRNAs to ensure a fast and specific regulation of

the immune response. For example, the cytokine *TNF-α* mRNA contains both ARE (AU-rich element) and CDE (constitutive decay element) sequences in its 3'UTR. The posttranscriptional regulator tristetraproline (TTP, also known as TIS11 or ZFP36) binds ARE sequences and recruits the CCR4-NOT deadenylase complex (see section 2.4). This mechanism is fine-tuned via phosphorylation of TTP by the MAP-kinase pathway (Clement et al. 2011). Additionally, in macrophages, the CDE sequence is recognized by Roquin1, which can also recruit the CCR4-NOT complex (Leppek et al. 2013). Thus, the recruitment of the CCR4-NOT complex by multiple RBPs (here TTP and Roquin) to the same mRNA (e.g. *TNF-α*) provides a robust regulation of inflammatory mRNAs through its deadenylase activity.

Function of the CCR4-NOT complex in GSC differentiation

Deadenylation by the CCR4-NOT complex is also crucial for posttranscriptional regulation during oogenesis in *Drosophila melanogaster* (*Dm*) (Morris et al. 2005). NOT1 knockdown in germ cells, with consequently impaired deadenylation, results in a loss of GSCs (Fu et al. 2015). Among the subunits of the CCR4-NOT complex, several studies revealed that CCR4 is required in *Dm* germ cells and early *Dm* embryos (Joly et al. 2013, Fu et al. 2015). Concurrently, CCR4, which is encoded by the gene *twin* in *Dm*, has a strong phenotype: *twin/CCR4* mutants show GSC loss. This is due to the role of CCR4 in regulating the poly(A) tail length of *cyclin A* and *B* in GSCs (Morris et al. 2005). Regulation of specific mRNAs in the context of GSCs is achieved by the recruitment of the CCR4-NOT complex via RBPs. To date, Nanos and Pumilio proteins have been reported to regulate specific targets such as *cyclin B* and *mei-P26* in *Dm* germ cells through the CCR4-NOT complex (Joly et al. 2013). Additionally, Fu et al. (2015) showed that the CCR4 deadenylase controls *Dm* GSC self-renewal by maintaining E-cadherin accumulation in the niche cells and by interacting with the key differentiation factor Bag-of-marbles.

2.4 Recruitment of the CCR4-NOT complex by mRNA-associated proteins

In addition to its central role in bulk mRNA degradation, the CCR4-NOT complex controls the expression of specific mRNAs to which it is recruited via interactions with RBPs. RBPs directly or indirectly - through additional regulatory proteins - interact with the complex to determine mRNA fate in multiple biological processes. Transcript recognition by RBPs is determined by mRNA sequences and/or structures (Rissland 2017).

DND1 (dead end homolog 1), TTP, GW182 (also known as TNRC6A-C in human), Smaug, Bicaudal-C, Pumilio, Nanos and Roquin1 are examples of RNA-associated proteins that have been

shown to bind to CCR4-NOT and to induce repression of target mRNAs. Their biological functions and interacting partners within the CCR4-NOT complex are described in Table 2.

Table 2: Overview of the RNA-associated proteins that were known to interact with the CCR4-NOT complex (or its subunits) at the beginning of my PhD. *Dm*, *Drosophila melanogaster*; *Hs, Homo sapiens*; *Mm, Mus musculus; Sc, Saccharomyces cerevisiae*. The asterisk (*) highlights proteins that I investigated during my PhD. Adapted and extended from Chapat & Corbo (2014).

RBPs/effectors	CCR4-NOT interacting partner	Species	RNA recognition element	Biological functions
DND1	CNOT1	*Mm*	URR	Germ cell development and maintenance
TTP	CNOT1	*Hs, Mm*	ARE	Inflammation, immunity
GW182	CNOT1/CAF40	*Hs, Dm*	miRNA-binding site	miRNA-mediated gene silencing
Smaug		*Dm*	SRE	Embryonic development
Bicaudal-C	NOT3	*Dm*		Oogenesis Embryonic development
Pumilio	CAF1	*Dm*	PRE	Abdominal patterning Germline stem cell self-renewal
Nanos	CNOT1 CNOT1-CNOT3	*Hs, Mm, Dm*		Embryonic patterning Germ cell development
Roquin1*		*Hs*	CDE	Inflammation, immunity Development
NOT4*	NOT1	*Sc*		
Bam*	CCR4	*Dm*		Germline stem cell differentiation

Recently Yamaji et al. revealed the RNA-binding protein DND1 as a putative binding partner for the N-terminal region of NOT1 in the mouse germ line. During embryogenesis in mice, DND1 recruits the CCR4-NOT complex to mRNA targets thus regulating inflammation, apoptosis and stem cell pluripotency. Specifically, DND1 interacts with U-rich regions (URRs) in the 3'UTR of target mRNAs (Yamaji et al. 2017).

Many proteins have also been characterized structurally in their binding to the CCR4-NOT complex. For example, TTP directly interacts with NOT1 and induces degradation of mRNA targets. TTP uses a tandem CCHC zinc finger (ZnF) domain to bind and destabilize mRNAs containing ARE sequences within their 3'UTR, such as *TNF-α*, *cyclooxygenase-2* and *IL-1*, playing a crucial role in inflammation, as mentioned before (see section 2.3.2) (Sandler et al. 2011, Brooks & Blackshear 2013, Fabian et al. 2013).

The CCR4-NOT complex can also be recruited by the RNA-induced silencing complex to control translation and stability of miRNA targets (Inada & Makino 2014). The interaction of the miRNA machinery with the CCR4-NOT complex is mediated by GW182 proteins. GW182 proteins directly interact with the PAN3 subunit of the PAN2-PAN3 deadenylase complex as well as with NOT1 and CAF40 of the CCR4-NOT complex to induce mRNA degradation and translational repression of miRNA-targets (Chen et al. 2014b, Mathys et al. 2014).

Moreover, proteins functioning during *Dm* embryonic development have been shown to recruit the CCR4-NOT complex to specific mRNAs. Smaug, for example, regulates the decay of maternal mRNAs during maternal-to-zygotic transition by recruiting the CCR4-NOT complex to mRNAs containing Smaug recognition elements (SREs) (Semotok et al. 2005, Zaessinger et al. 2006).

During *Dm* oogenesis, Bicaudal-C binds to the NOT3 subunit of the CCR4-NOT complex to negatively regulate mRNAs involved in cytoskeletal regulation and oogenesis, including its own *bicaudal-C* mRNA (Chicoine et al. 2007).

The RNA binding protein Pumilio binds to the Pumilio response element (PRE) of mRNA targets. The Pumilio/Nanos complex specifically represses *hunchback* mRNA to control *Dm* embryonic development (Wreden et al. 1997). This complex then recruits the CCR4-NOT machinery by binding to CAF1 via Pumilio, and NOT1-2-3 via Nanos (Van Etten et al. 2012, Raisch et al. 2016).

In vertebrates and invertebrates, Nanos proteins contain a C-terminal CCHC-type ZnF that binds nucleic acids, and an unstructured, poorly conserved, N-terminal region. Human and *Dm* Nanos use different short motifs at their N-terminal regions to recruit the CCR4-NOT complex to mRNA targets. Human Nanos binds to the C-terminal end of NOT1, while *Dm* Nanos uses a

bipartite binding mode contacting the N-terminal part of NOT1 and NOT3 (Bhandari et al. 2014, Raisch et al. 2016).

Recently, Roquin proteins have been shown to interact with conserved RNA stem-loop sequences, which are highly conserved and are found in more than 50 vertebrate mRNAs and code for regulators of inflammation, development and immunity. Additionally, Roquin1 was shown to recruit the CCR4-NOT complex, although the molecular details of the recruitment remain poorly characterized (Leppek et al. 2013).

Furthermore, Bhaskar and colleagues revealed the *Saccharomyces cerevisiae* (*Sc*) NOT4 to bind to the NOT1 SHD via a short motif, which is not conserved in other species (Bhaskar et al. 2015). Although human NOT4 was shown to interact with the C-terminal region of NOT1 in yeast-two-hybrid assays (Albert et al. 2000), detailed molecular insights into this association with the CCR4-NOT complex are missing. It is also unclear whether NOT4 uses its RRM domains to recruit the CCR4-NOT complex to mRNA targets (Bhaskar et al. 2015).

The fly-specific differentiation factor Bag-of-marbles (Bam) was recently shown to bind the CCR4-NOT deadenylase CCR4. However, Bam was proposed to do so independently of the CCR4-NOT complex (Fu et al. 2015) and it its currently unknown if Bam can bind mRNAs in the context of a posttranscriptional repressor complex to regulate stem cell differentiation (Li et al. 2009, Shen et al. 2009, Insco et al. 2012).

All of these examples indicate that the CCR4-NOT complex acts as a major regulator of mRNA stability during diverse cellular processes such as differentiation, immune response and proliferation. Below, I provide detailed information about two RNA-associated, regulatory proteins: the E3 ubiquitin ligase Roquin and the stem cell differentiation factor Bam, which are the major focus of my studies.

2.5 RNA-binding E3 ubiquitin ligases

Posttranslational modifications (PTMs) influence protein functions, and ubiquitination is a PTM that typically causes protein degradation (Varshavsky 2017). In addition, ubiquitination of RBPs can alter interactions, affinities and localization and thus influences mRNA translation and stability (Lee 2012). Ubiquitin is usually covalently attached to a lysine residue of the substrate proteins via an isopeptide bond. Ubiquitination requires three enzymatic reactions that include the ubiquitin-activating enzymes (E1), the ubiquitin conjugating enzymes (E2) and the ubiquitin ligases (E3) (Komander & Rape 2012).

The human genome encodes more than 600 potential E3 ligases and fifteen of these contain an additional RNA binding domain, implying another layer of complexity in posttranscriptional regulation (Deshaies & Joazeiro 2009, Cano et al. 2010, Hildebrandt et al. 2017). Little is known about the impact of ubiquitination on mRNA turnover. However, in the past few years several pieces of evidence suggested a link, as is the case for the MEX-family proteins. The human RNA-binding proteins MEX-3A to -D are conserved phosphoproteins that contain two K homology (KH) domains and a C-terminal RING domain. Among the MEX-family proteins, MEX-3C has been associated with immune disease and cancer (Pereira et al. 2013, Cano et al. 2015). MEX-3C binds a consensus RNA motif in the 3'UTR of *HLA-A2* (*human leukocyte antigen A2*) mRNA through its KH domains and with its RING domain interacts with and ubiquitinates the CAF1 subunit of the CCR4-NOT complex. Ubiquitination of CAF1 does not cause its proteolysis, but promotes its deadenylation activity, resulting in posttranscriptional regulation of specific transcripts (Cano et al. 2015, Yang et al. 2017). This example implies a role for ubiquitin ligase activity in mRNA decay. However, the molecular mechanism of mRNA decay and translational repression mediated by MEX-3C lacks detailed characterization.

Roquin and NOT4 are also RNA-binding E3 ubiquitin ligases that may connect ubiquitination to mRNA stability. Moreover, both proteins were shown to interact with the CCR4-NOT deadenylase complex (Leppek et al. 2013, Bhaskar et al. 2015), thus, they might regulate the activity of the complex by ubiquitination, similar to MEX-3C.

2.6 Roquin proteins

Roquin proteins are E3 ubiquitin ligases of the RING-type family with an important role in immune response. Roquin is an essential factor for preventing T-cell mediated autoimmune disease. It was identified in a mutagenesis screen for autoimmune regulators in mice where a point mutation in the *Rc3h1* gene, encoding for Roquin1, caused autoimmunity in the *sanroque* mice (Vinuesa et al. 2005). The *sanroque* mice showed aberrant follicular helper T cell expansion, production of autoantibodies and resembled the human systemic lupus erythematosus (SLE) disease (Vinuesa et al. 2005).

Roquin proteins are highly conserved in metazoans with only one homolog found in Dm and *Caenorhabditis elegans* (*Ce*) (Li et al. 2007). In vertebrates, there are two paralogs, Roquin1 and Roquin2 also called MNAB (membrane-associated nucleic-acid binding protein; Siess et al. 2000). The Ce and Dm homologs are RLE-1 (regulation of longevity by E3) and Roquin, respectively. The

N-terminal domain organization of all Roquin proteins is highly similar, whereas the C-terminal regions are unstructured and strongly divergent (see Figure 5 for Roquin domain organization).

The presence of a RING domain suggests that Roquin may target proteins for ubiquitin-dependent degradation, but until now, little is known about this activity. Previous studies have shown that Roquin2 ubiquitinates apoptosis signal-regulating kinase 1 (ASK1) for proteasomal degradation (Maruyama et al. 2014). Additionally, the Ce RLE-1 protein regulates the lifespan of the worm by ubiquitinating the transcriptional activator DAF-16 (homolog of mammalian Foxo3a) that regulates aging in Ce (Li et al. 2007). The ROQ domain is a highly conserved domain among Roquin proteins and mediates RNA binding. It recognizes a conserved stem-loop motif (CDE) in the 3'UTR of mRNA targets, such as the inducible costimulator Icos, the costimulatory receptor Ox40, neuropilin-1, IFN-γ, TNF-α and Roquin1 and 2 (Yu et al. 2007, Leppek et al. 2013, Vogel et al. 2013, Schlundt et al. 2014, Tan et al. 2014, Sakurai et al. 2015, Janowski et al. 2016).

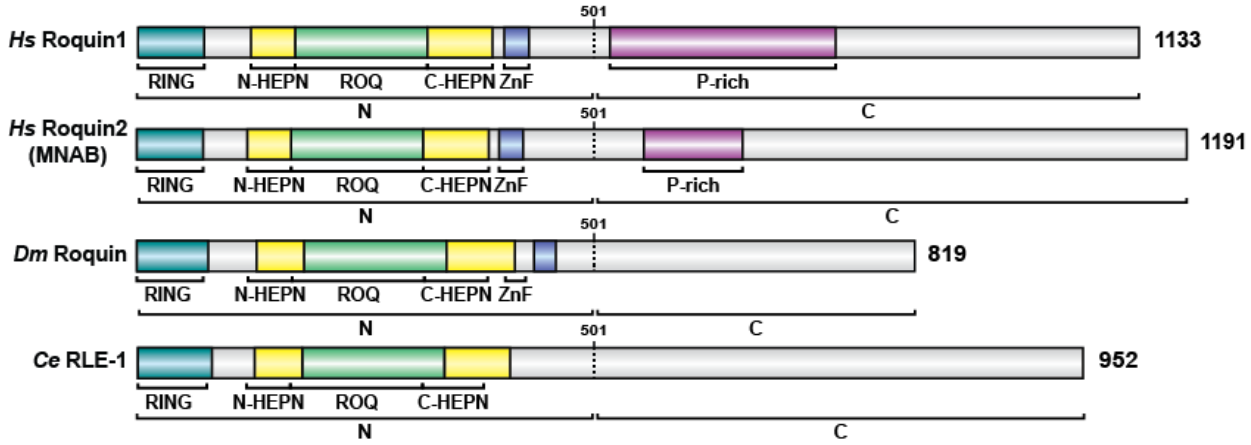


Figure 5: Domain organization of Roquin proteins. Roquin proteins share a conserved N-terminal region containing a RING-finger E3 ubiquitin ligase domain, a ROQ RNA-binding domain flanked by helical HEPN (higher eukaryotes and prokaryotes nucleotide-binding) domains and a CCHC-type ZnF domain. The C-terminal half of Roquin proteins is not conserved and is predicted to be an intrinsically disordered region (shown in grey). Adapted from Sgromo et al. (2017).

Importantly, human Roquin1 was shown to recruit the CCR4-NOT complex via its C-terminal region, leading to a complete degradation of bound transcripts (Leppek et al. 2013). However, the molecular mechanism of this recruitment was still unknown when I started my PhD project.

2.7 Bag-of-marbles protein

Dm Bag-of-marbles (Bam) is a master differentiation factor that determines the fate of germline stem cells. Loss of Bam results in uncontrolled stem cell proliferation and tumor formation, causing the tumorous *bam* mutant phenotype, where the germ cells resemble marbles in a bag (McKearin & Ohlstein 1995, Ohlstein & McKearin 1997). Bam lacks homology to other known proteins and it is only present in flies. Secondary structure prediction suggests Bam to be mainly α-helical (Figure 6).

Figure 6: Schematic view of a secondary structure prediction of Bam. The Bam protein is predicted to be mainly α-helical (yellow) and to have a single β-strand (blue). Alignments indicate a possible separation into N- and C- fragments. Adapted from Sgromo et al. (2018).

Dm gametogenesis is a common model to study stem cell behavior (Lin 2002). Bam is transcriptionally silent in GSCs and upregulated in cystoblasts (Chen & McKearin 2003). Via E-cadherin, GSCs are anchored to neighboring niche cells that produce bone morphogenetic proteins (BMPs). In GSCs, the BMP ligands Decapentaplegic (Dpp) and Glass bottom boat (Gbb) activate the BMP receptor complex, consisting of type I serine/threonine kinase receptors Saxophone (Sax) and Thickveins (Tkv), as well as the type II receptor Punt (Xie & Spradling 1998). BMP signaling prevents differentiation by activating downstream transcriptional effectors (Mad, Medea and Schnurri) that silence *bam* transcription. Transcriptional repression of the *bam* gene is the major requirement to maintain GSC fate. Many additional factors have been identified as crucial regulators of GSC self-renewal. As an example, the repressor complex Pumilio/Nanos induces GSC fate within these cells via the CCR4-NOT complex (Kadyrova et al. 2007). In addition, microRNAs inhibit differentiation genes in GSCs (Neumuller et al. 2008).

Following GSC division, one daughter cell migrates away from the niche to become a cystoblast. The absence of BMP signaling in cystoblasts leads to expression of Bam, which initiates the differentiation program (Spradling et al. 2011). Bam controls GSC differentiation by forming a protein complex with Benign gonial cell neoplasm (Bgcn; a putative DEXH RNA-helicase-like protein) (Lavoie et al. 1999, Li et al. 2009). This complex is thought to promote cystoblast formation by posttranscriptionally repressing the expression of stem cell maintaining factors like Nanos, Pumilio and E-cadherin (Li et al. 2009, Shen et al. 2009, Insco et al. 2012). Bam has also been found to cooperate with the proteins Tumorous testis (Tut; containing a tandem RRM) (Chen

et al. 2014a), Sex-lethal (Sxl) (Chau et al. 2012, Li et al. 2013) and Mei-P26 (Neumuller et al. 2008, Li et al. 2013) leading to differentiation. The process of germline stem cell differentiation in *Drosophila* is summarized in Figure 7.

Despite the many protein-protein interactions identified for Bam, the mechanism by which it regulates mRNAs remains poorly understood. Currently it is unknown whether Bam can directly bind mRNAs or requires additional adaptor proteins.

Recently, the CCR4 deadenylase subunit of the CCR4-NOT complex was shown to interact with Bam by immunoprecipitation in *Dm* S2 cells and germ cells (Fu et al. 2015). Fu et al. suggested a functional connection between Bam and CCR4 independent of the assembled CCR4-NOT complex. However, it remained unclear which subunit of the CCR4-NOT complex is the direct interactor of Bam and if Bam requires an intact CCR4-NOT complex or isolated CCR4 to perform its repressive function.

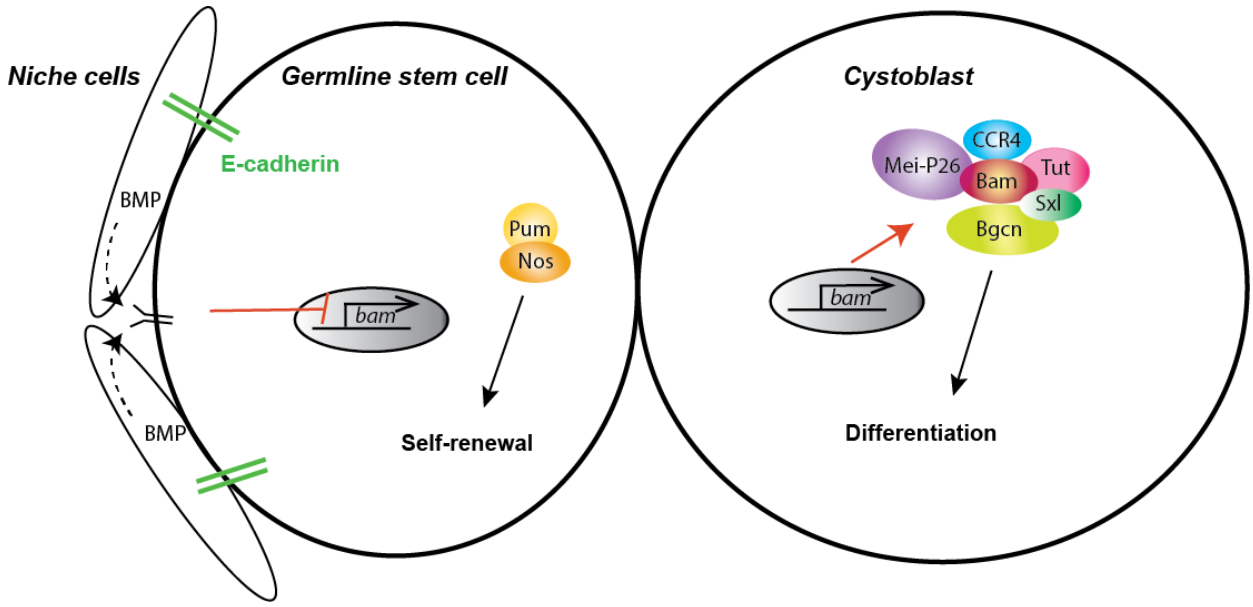


Figure 7: Schematic representation of *Drosophila* gametogenesis. Niche cells secrete BMP signals to maintain the nearby germline stem cell (GSC) identity by repressing *bam* expression. In the GSCs the repressor complex Pumilio/Nanos induces a self-renewal fate. After division, cells migrate away from the niche and the self-renewal signals. In the absence of BMP signaling, those cells differentiate into cystoblasts due to upregulation of *bam* expression. Here Bam, together with additional proteins such as Benign gonial cell neoplasm (Bgcn), Sex-lethal (Sxl), Mei-P26, Tumorous testis (Tut) and CCR4, regulates differentiation.

3. AIMS & OBJECTIVES

To influence mRNA fate and ultimately protein production, a wide variety of biological processes rely on posttranscriptional regulation of mRNA (Vogel & Marcotte 2012). A versatile and widely acknowledged effector of mRNA regulation is the CCR4-NOT complex. The multisubunit complex deadenylates bulk mRNAs and specific transcripts, leading to their degradation. Recruitment of the CCR4-NOT complex to specific mRNA targets can be achieved by RBPs. Multiple RBPs and RNA-associated proteins are known to contact subunits or modules of the large CCR4-NOT complex (see section 2.4). Regulatory proteins that associate with RNA and recruit the CCR4-NOT complex to regulate mRNA expression are the central element of my PhD studies.

In my work, I combined structural and biochemical approaches in order to elucidate how two *Drosophila* proteins, Roquin and Bam, regulate mRNA stability through their interaction with the CCR4-NOT complex, unveiling the molecular mechanisms of posttranscriptional regulation by Roquin and Bam.

In the first part of my doctoral research, I investigated the regulatory mechanism underlying Roquin-mediated mRNA degradation. The human RBP Roquin1 has previously been shown to recruit the CCR4-NOT complex resulting in targeted mRNA decay (Leppek et al. 2013). Although human Roquin proteins have been studied for their important role in the immune system and inflammation, *Dm* Roquin was not yet characterized when I started my PhD. Therefore, it remained unclear whether and how *Dm* Roquin recruits the CCR4-NOT complex to degrade target mRNAs. My work aimed to investigate whether Roquin-mediated CCR4-NOT recruitment is conserved among species, given the poor conservation of the C-terminal region that binds to the CCR4-NOT complex in human Roquin1 (Leppek et al. 2013).

In the second part of my doctoral research, I focused on the function of the fly-specific GSC differentiation factor Bam in mRNA decay. Fu et al. (2015) identified an interaction between Bam and CCR4, and proposed that Bam acts independently of the CCR4-NOT complex. In my studies, I wanted to test this model and fully characterize the binding of Bam to the CCR4-NOT complex to address how Bam posttranscriptionally regulates mRNA targets. Furthermore, I investigated the physical connection between Bam, Bgcn and Tut, the proposed repressor complex involved in GSC differentiation. Understanding how individual components of Bam's network connect was expected to provide insights into the posttranscriptional regulation mechanisms behind GSC differentiation and the role of Bam in this process.

Finally, I also contributed to a collaborative project on the characterization of NOT4, which is an E3 ubiquitin ligase like Roquin. The presence of an E3 ubiquitin ligase domain together with an RNA-binding domain is a striking feature shared by other proteins associated with the CCR4-NOT complex, potentially linking protein turnover to posttranscriptional control. NOT4 carries an RRM domain, which suggests RNA binding and is a stable component of the CCR4-NOT complex in yeast. However, it remained unclear whether, in metazoans, NOT4 interacts with the CCR4-NOT complex.

Deciphering the intricate binding network of the CCR4-NOT complex and associated RNA-binding proteins in molecular detail is crucial for a thorough understanding of how cytoplasmic control of gene expression is achieved in a cell.

4. RESULTS

4.1 Roquin-mediated mRNA degradation is conserved across species

The work described in this chapter has been published in Sgromo et al. 2017. Detailed experimental data and methods are available in the attached publication (see chapter 9).

4.1.1 Roquin proteins induce mRNA decay via their unstructured C-terminal regions

The Roquin RNA-binding proteins are evolutionarily conserved (see Figure 5). Previous work identified the CCR4-NOT deadenylase complex as a major downstream factor for the repressive activity of *Hs* Roquin1 (Leppek et al. 2013) but the molecular details of this interaction were not investigated thoroughly. In addition, it remained unclear how Roquin proteins induce mRNA degradation and if they have maintained this activity throughout evolution.

Hs Roquin1 was shown to recruit the CCR4-NOT complex via its unstructured C-terminal region (Leppek et al. 2013). The divergent nature of the C-terminal tails of Roquin proteins raised the question whether all Roquin proteins share a similar way to recruit the CCR4-NOT complex. In order to understand how Roquin proteins induce mRNA decay, I first cloned and expressed the *Hs* Roquin1 and Roquin2 proteins in the Human Embryonic Kidney 293T cell line (HEK293T) and the *Dm* Roquin protein in the *Dm* S2 cell line. To test the activity of Roquin proteins in mRNA decay and translational repression I performed tethering assays. The tethering assay is a powerful technique that allows to analyze the functional activity of a tethered protein on an artificial reporter mRNA, without knowing the natural mRNA target of the tethered protein. Different reporter systems have been developed, such as the λN/BoxB and MS2 systems. The protein of interest to be tethered is expressed with a tag of either λN or MS2 coat protein. This tag specifically binds to an RNA sequence, for example within the 3'UTR of the reporter, thereby recruiting the protein of interest to the reporter mRNA (Coller & Wickens 2007). The effect of the tethered protein on reporter levels can then be studied using several experimental techniques such as northern blots (revealing the effect on RNA levels) or luciferase assays (revealing the effect on protein levels).

Using the tethering assay, I first tested the effect of *Hs* Roquin1 and 2 on a β-globin reporter mRNA containing 6xMS2 binding hairpins in the 3'UTR. The full-length proteins as well as the N- or C-terminal fragments (defined based on sequence alignments and analysis) were used with an N-terminal MS2-tag. The N-terminal fragments include all conserved, structured domains (see Figure 5), whereas the C-terminal fragments comprise the unstructured, non-conserved tails. Tethering of both full-length *Hs* Roquin1 and *Hs* Roquin2 induced reporter mRNA decay. Tethering of the C-terminal fragment of *Hs* Roquin1 and *Hs* Roquin2 resulted in a comparable degradation of reporter

mRNA, whereas the N-terminal regions were inactive (Figure 8 A-D). These results confirmed previous experiments that linked the C-terminal region of *Hs* Roquin1 to mRNA decay (Leppek et al. 2013), and additionally showed that *Hs* Roquin2 can trigger mRNA degradation using the C-terminal region which shares no sequence similarity with *Hs* Roquin1. Given that *Hs* Roquin1 and 2 both induce mRNA decay through their C-terminal regions, I wanted to address whether this function is conserved, despite the C-terminal sequence divergence, also in *Dm* Roquin.

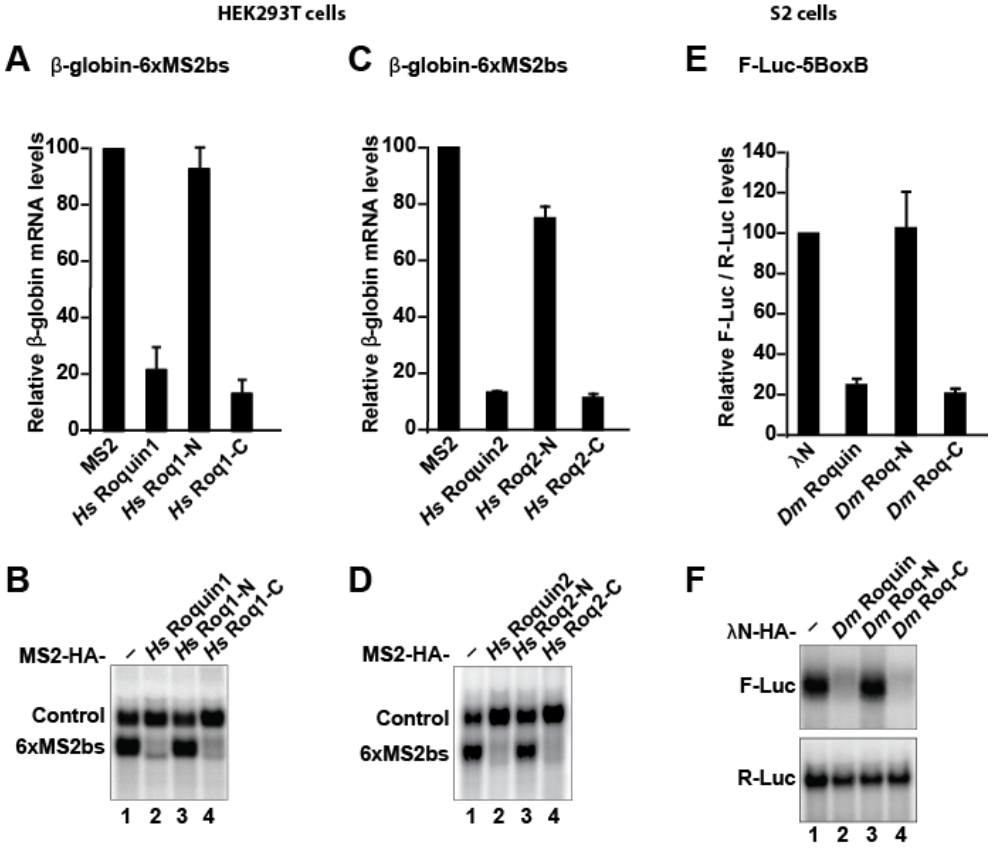


Figure 8: Roquin proteins trigger mRNA decay via their C-terminal tails. Tethering assays using *Hs* Roquin1 and *Hs* Roquin2 in HEK293T cells and *Dm* Roquin in S2 cells. Full-length constructs or fragments (N- and C- terminal regions) of Roquin proteins were tethered. Panels A, C and E show the relative mRNA levels in three independent experiments. mRNA levels were normalized to those of the control (control or R-Luc) and set to 100% in cells expressing the tag peptide alone. Panels B, D and F show representative northern blots for each tethering assay. Figure modified from Sgromo et al. (2017).

Therefore, I performed similar experiments in S2 cells, using the λN/BoxB tethering system. N-terminally λN-tagged *Dm* Roquin fusion proteins (full-length, N- or C-terminal constructs) were tethered to the reporter encoding the firefly luciferase and carrying 5BoxB elements in the 3'UTR. Interestingly, also *Dm* Roquin induced degradation of the mRNA reporter via its C-terminal region (Figure 8 E,F), revealing that this feature is maintained across species. These results show that all

tested Roquin proteins trigger decay of reporter mRNA via their C-terminal regions that are unstructured and are not conserved.

4.1.2 *Dm* Roquin interacts directly with the CCR4-NOT complex to elicit mRNA decay

The 5'-to-3' mRNA decay pathway promotes deadenylation followed by decapping and rapid degradation (see chapter 2). To address whether Roquin proteins use this pathway to elicit mRNA decay, I overexpressed an inactive mutant of the catalytic subunit DCP2 of the decapping complex, which out-competes the endogenous protein, resulting in a dominant negative effect with inhibition of decapping activity and consequently inhibition of mRNA decay. In this condition, tethering of Roquin proteins to a reporter mRNA resulted in an accumulation of deadenylated mRNA, implying that Roquin proteins require 5'-to-3' mRNA decay.

Since the CCR4-NOT complex is the major deadenylase of the 5'-to-3' decay pathway, it was of interest to study how it contributes to Roquin-mediated mRNA degradation. I therefore performed RNA interference (RNAi)-mediated knockdown experiments (Clemens et al. 2000, Zamore 2001), using double-stranded (dsRNA) RNA against the open reading frame (ORF) of NOT1 in S2 cells to deplete NOT1, the scaffold subunit of the CCR4-NOT complex. Tethering of *Dm* Roquin in cells depleted of NOT1 reduced Roquin-mediated mRNA decay activity, in comparison to a control knockdown, revealing that Roquin function requires the CCR4-NOT complex.

To gain a mechanistic insight into the specificity of Roquin proteins in mRNA decay, I decided to biochemically characterize the interactions between Roquin proteins and the CCR4-NOT deadenylase complex first via co-immunoprecipitation (co-IP) assays in both *Hs* and *Dm* cells. The co-IP studies showed that Roquin proteins interact with the core subunits of the CCR4-NOT complex *in vivo*. In order to identify direct interactions, I then performed *in vitro* pulldown assays with purified proteins. I therefore expressed the *Dm* Roquin C-terminal region, which induced mRNA decay in the tethering assays, with an N-terminal MBP tag in *Escherichia coli* (*E. coli*). I performed pulldown assays of the MBP-tagged *Dm* Roquin C-terminal fragment with a purified subcomplex of the CCR4-NOT complex (consisting of the C-terminal fragments of NOT1, NOT2 and NOT3, as well as the full-length NOT7 and the structured core of CAF40) and its purified distinct modules (the catalytic module, the CAF40 module and the NOT module) (see Figure 3). My results showed that *Dm* Roquin directly binds to the CAF40 subunit and to the NOT module. To map the minimal interacting regions in *Dm* Roquin, required to bind the CAF40 subunit and the NOT module, I performed *in vitro* pulldown assays with different overlapping fragments of the *Dm* Roquin C-terminal region. Using this approach, I identified two minimal binding peptides. One is

responsible for the binding to CAF40 (the CAF40 binding motif or CBM_R, where the subscript R refers to Roquin) and one is responsible for the binding to the NOT module (the NOT module binding motif or NBM).

To unravel the molecular details of the CBM_R-CAF40 interaction, a crystal structure of the *Dm* Roquin CBM_R peptide bound to the *Hs* CAF40 protein was solved in collaboration with a former PhD student in the lab, Tobias Raisch (note that *Hs* and *Dm* CAF40 share 75% of sequence identity). In the crystal structure, the CAF40 ARM domain folds into six armadillo repeats with a concave surface where the CBM_R is bound. The CBM_R uses an amphipathic helix to mediate binding to a conserved hydrophobic surface of CAF40 (see section 4.4.1). To validate the binding interface I tested several mutations, which were designed to disrupt the binding interface observed in the crystal structure, both in the context of *Dm* Roquin and in the context of CAF40. All of the mutations disrupted the interaction *in vitro* as well as *in vivo* (Figure 9 A,B).

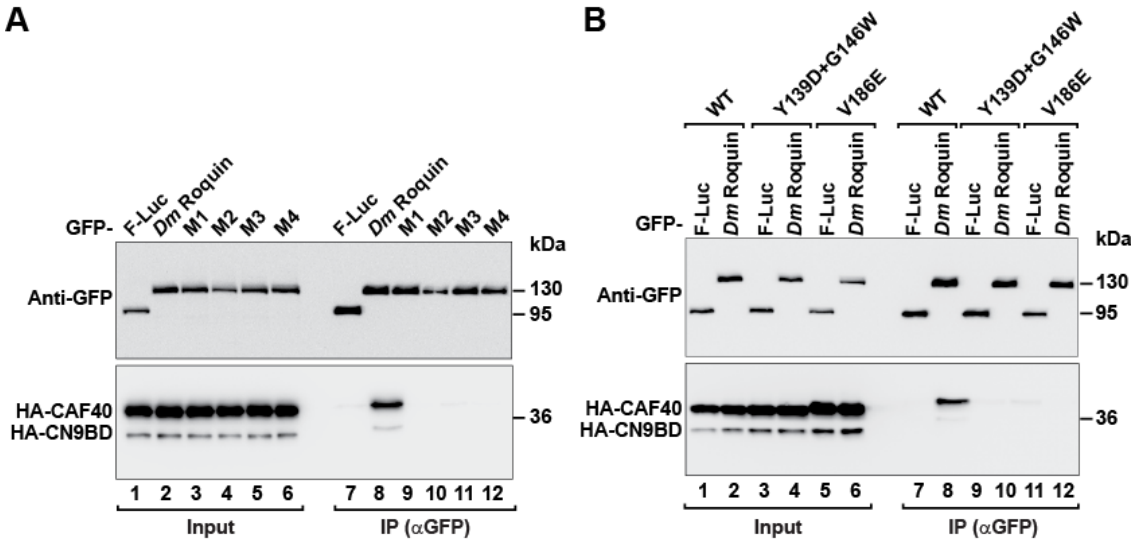


Figure 9: Validation of the CBM_R-CAF40 interface. Anti-GFP i*n vivo* co-IPs of GFP-tagged *Dm* Roquin proteins with the HA-tagged CAF40 module in the presence of RNaseA to exclude RNA-mediated interactions. Panel A shows that wild type (wt) *Dm* Roquin immunoprecipitates the CAF40 module, whereas *Dm* Roquin mutants (M1, M2, M3 and M4; design based on the crystal structure) are not able to bind the CAF40 module anymore. Panel B shows that mutations in the concave CAF40 interface (Y139D + G146W and V186E; mutations based on the structure) disrupt the interaction with wt *Dm* Roquin. Protein size markers (kDa) are shown on the right in each panel. Figure modified from Sgromo et al. (2017).

Interestingly, even though the *Hs* Roquin1 C-terminal region also directly contacts CAF40, the CAF40 mutations (Y134D + G141W and V181E) that disrupt the interaction with *Dm* Roquin did not disrupt the interaction with *Hs* Roquin1 (Figure 10). This result indicates that *Hs* Roquin1

has a more extended binding interface with CAF40 and/or that the interactions of *Dm* Roquin and *Hs* Roquin1 require different residues on CAF40.

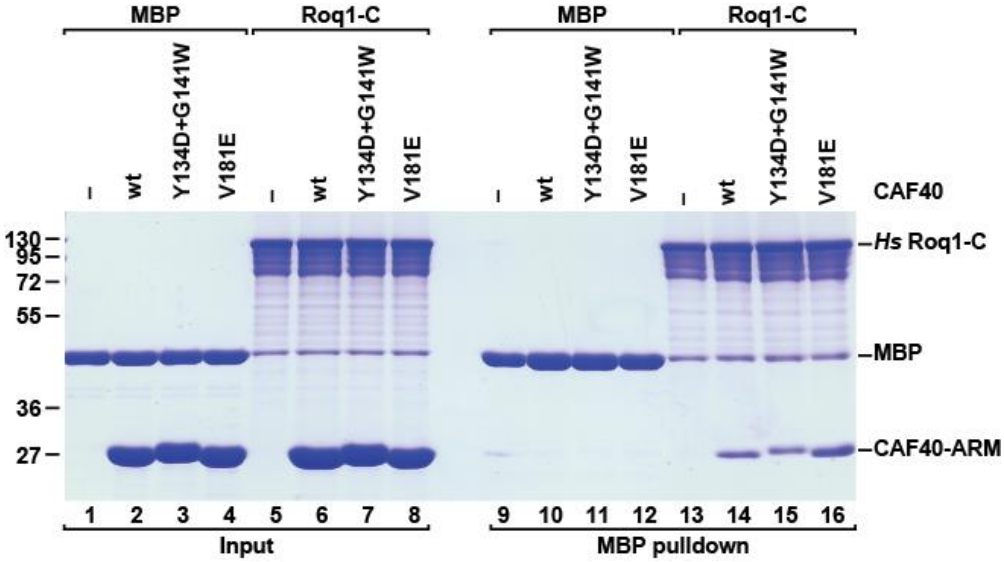


Figure 10: Interaction of the *Hs* Roquin1 C-terminal region with CAF40. MBP pulldown assay of the MBP-tagged *Hs* Roquin1 C-terminal fragment and purified *Hs* CAF40 wt (lane 14) or mutants (Y134D + G141W and V181E, lanes 15 and 16, respectively). MBP alone served as a negative control (lanes 9 to 12). Protein size markers are shown on the left. Figure adapted from Sgromo et al. (2017).

Regarding the interaction of *Hs* Roquin1 and *Hs* Roquin2 with the CCR4-NOT complex, I could not define a single fragment within their unstructured C-terminal regions, sufficient to mediate the interaction, but instead my results suggested the presence of multiple low-affinity binding sites that all contribute in an additive manner to bind CCR4-NOT.

4.1.3 NBM and CBM_R sequences in *Dm* Roquin act redundantly to mediate mRNA decay

To investigate the role and the contribution of the CBM_R and the NBM peptides in *Dm* Roquin to the interaction with the CCR4-NOT complex, I tested them in tethering assays. Single deletions of each of those peptides or combined deletions in *Dm* Roquin did not abolish its activity in tethering assays, suggesting the presence of other redundant and uncharacterized binding sites within the *Dm* Roquin C-terminal region involved in the regulation of mRNA targets. Therefore, both identified motifs contribute to the activity of *Dm* Roquin in S2 cells but they are not sufficient for its full activity.

To test the impact of the CBM_R on an endogenous Roquin target, I used *Dm* Roquin in a hybrid experiment in HEK293T cells. In this experiment, I modified the MS2-reporter system of the tethering assay by replacing the 6xMS2 binding hairpins in the 3'UTR of the β-globin reporter with

the CDE of TNF-α. The CDE of TNF-α was previously shown to be directly bound by *Hs* Roquin1 (Leppek et al. 2013; see Figure 11 A). Strikingly, wild type (wt) *Dm* Roquin promoted CDE-mediated mRNA decay of this reporter, whereas the *Dm* Roquin with a mutated CBM_R (M4), which is not able to bind CAF40, was strongly impaired in its mRNA decay activity in human cells (Figure 11 B,C). Thus, the CBM_R provides a strong contribution to the interaction of *Dm* Roquin with the CCR4-NOT complex in human cells. Furthermore, the experiment implies that *Dm* Roquin has the capability to directly recognize the stem loop structure of the TNF-α CDE by the ROQ domain, which is conserved across species. However, endogenous *Dm* Roquin mRNA targets are still unknown.

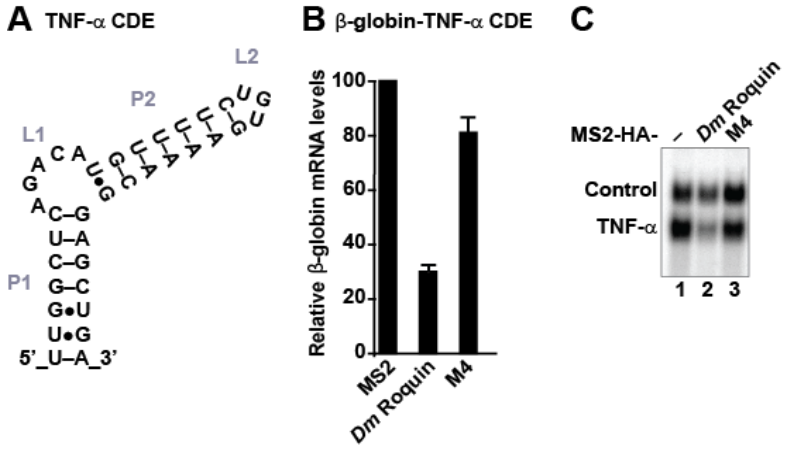


Figure 11: *Dm* Roquin induces mRNA decay of a reporter containing the TNF-α CDE element in the 3'UTR. Panel A shows the secondary structure model of the human CDE of TNF-α (adapted from Leppek et al. 2013) that I cloned into the 3'UTR of the β-globin reporter. Panel B shows the relative mRNA levels of three independent experiments. mRNA levels were normalized to those of the control and set to 100% in cells expressing the MS2 tag. *Dm* Roquin induces degradation of an mRNA containing the CDE of TNF-α in its 3'UTR, in HEK293T cells. In contrast, a mutant version of *Dm* Roquin (M4), which is not able to interact with the CAF40 subunit of the CCR4-NOT complex, retains the ability to induce degradation of the same reporter mRNA. Panel C shows a representative northern blot for the experiment shown in panel B. Figure modified from Sgromo et al. (2017).

Overall, this study shows that all tested Roquin proteins elicit mRNA degradation by recruiting the CCR4-NOT complex. My results also shed light onto the evolution of Roquin proteins. It is evident that Roquin proteins have maintained the ability to recognize secondary structure elements in the 3'UTR of target mRNAs via their conserved ROQ domain. Moreover, the recruitment of the CCR4-NOT complex is conserved during evolution, although *Hs* Roquin proteins and *Dm* Roquin seem to rely on different motifs to mediate the interaction.

4.2 Bam requires the assembled CCR4-NOT complex to induce decay of bound mRNA

The work described in this chapter has been published in Sgromo et al. 2018. Detailed experimental data and methods are available in the attached publication (see chapter 9).

4.2.1 Bam carries an N-terminal CBM_B to recruit the CCR4-NOT complex

This section describes my studies of *Dm* Bam, which is a crucial factor in germ line stem cell maintenance and has been proposed to interact with the CCR4 subunit of the CCR4-NOT complex (Fu et al. 2015). Fu et al. proposed a model in which Bam-mediated mRNA repression required the deadenylase CCR4 isolated from the CCR4-NOT complex. This model was based on the observation that mutations in the LRR domain of CCR4 disrupted both the binding to Bam and CAF1 (the other deadenylase of the CCR4-NOT complex that bridges CCR4 and NOT1, see Figure 3). However, the mutations introduced by Fu et al. most likely destabilized the fold of the LRR domain of CCR4, as suggested by the available CCR4 structural information (Basquin et al. 2012). A misfolded CCR4 would explain why the mutant binds neither Bam nor CAF1 and cannot be incorporated into the CCR4-NOT complex anymore. Hence, it remained unclear whether Bam activity required isolated CCR4 or the assembled CCR4-NOT complex.

To address this issue, I worked together with a master student in the lab, Charlotte Backhaus, to investigate the role of Bam in mRNA decay and translational repression. To assess whether Bam elicits mRNA degradation, I first performed tethering assays in the *Dm* S2 cell line. Different fragments of the Bam protein were tested (N- and C-terminal regions, as well as full-length protein; see Figure 6). Bam triggered reporter mRNA degradation via its N-terminal region, which does not include the previously proposed binding site for CCR4. Furthermore, overexpression of an inactive mutant of the catalytic subunit DCP2 blocked Bam-mediated mRNA decay, demonstrating that Bam acts via the 5'-to-3' mRNA decay pathway, as I previously showed for Roquin (see section 4.1). Moreover, depletion of NOT1 in S2 cells impaired Bam-mediated mRNA degradation, suggesting that the decay of a Bam-bound reporter mRNA requires the fully-assembled CCR4-NOT complex.

To identify components of the CCR4-NOT complex that mediate Bam activity I performed co-IPs in S2 cells. HA-tagged Bam co-immunoprecipitated with multiple components of the CCR4-NOT deadenylase complex, including NOT1, NOT2, NOT3, CCR4 (as previously shown) and CAF40 (Figure 12 A-E). Since these experiments were done in cell lysates, the observed interactions could be indirect and mediated by other proteins. To identify which of the interactions are direct, I performed *in vitro* pulldown assays using MBP-tagged full-length Bam and all of the purified modules of the CCR4-NOT complex, as previously shown for the experiments with *Dm*

Roquin (see section 4.1). Surprisingly, MBP-tagged Bam directly bound to the CAF40 module but not to the catalytic module, which contains CCR4. Further *in vitro* mapping of MBP-tagged, overlapping Bam fragments revealed the presence of a short motif (CBM_B, where the subscript B refers to Bam) in the N-terminal region, responsible for the direct interaction with the CAF40 module. This CBM_B was necessary and sufficient to mediate the interaction with the CCR4-NOT complex both *in vitro* and *in vivo*.

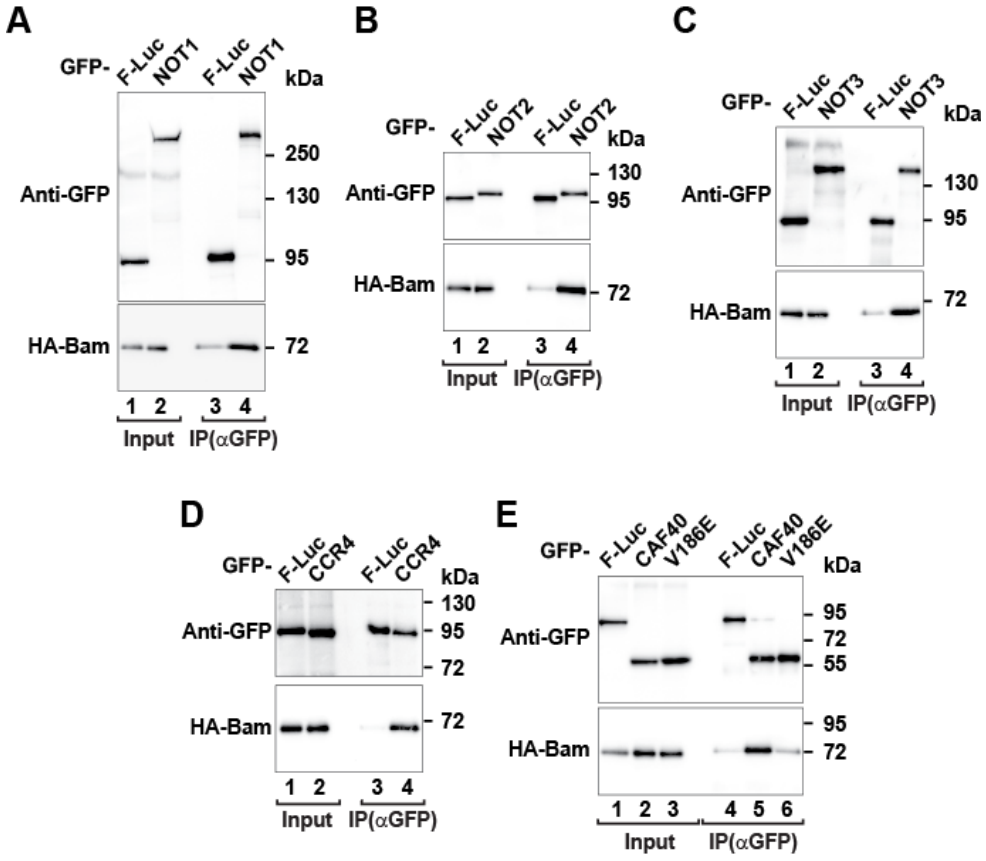


Figure 12: *Dm* Bam co-immunoprecipitates with several subunits of the CCR4-NOT complex. Western blots showing anti-GFP co-IPs of GFP-tagged CCR4-NOT subunits with HA-tagged Bam in *Dm* S2 cells, in presence of RNaseA. Bam interacts with NOT1 (A), NOT2 (B), NOT3 (C), CCR4 (D) and CAF40 (E, lane 5). Protein size markers (kDa) are shown on the right in each panel. Figure modified from Sgromo et al. (2018).

In collaboration with Tobias Raisch, two crystal structures were obtained for a detailed characterization of the Bam-CAF40 interaction: the *Dm* Bam peptide (CBM_B) bound to the *Hs* CAF40 protein (see section 4.4.1) and the *Dm* Bam CBM_B bound to the *Hs* CAF40-NOT1 (CN9BD) complex. The latter clearly showed that Bam interacts with CAF40 via a different surface than NOT1; therefore, both interactions can occur simultaneously and do not interfere with each other. Strikingly, the structures revealed that the CBM_B peptide binds to CAF40 in a similar manner as the CBM_R peptide that I previously identified (see section 4.1). To validate the functional significance of the Bam-CAF40 binding interface, I tested single point mutations in both the *Dm*

Bam full-length protein and on the CAF40 surface. All of the mutations were designed based on the structural information and they disrupted the interaction *in vitro* as well as *in vivo* (see Figure 12 E, lane 6 for the V186E). Additionally, tethering of these Bam mutants also impaired Bam-mediated mRNA decay. This confirmed that the CBM_B is necessary and sufficient to promote degradation of bound mRNAs (Figure 13 A,B).

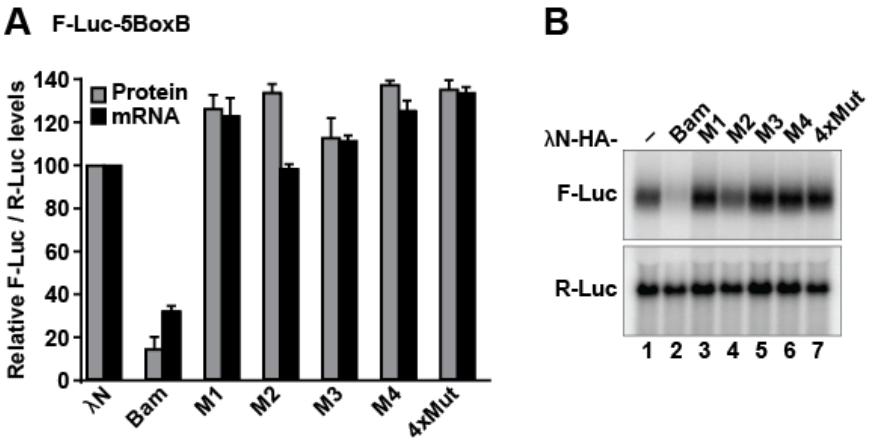


Figure 13: Mutations in the Bam CBM_B impair the repression of a reporter. Tethering assay using λN-HA-tagged Bam (full-length or mutant versions) and the F-Luc-5BoxB reporter in *Dm* S2 cells. Bam induces decay of the reporter mRNA, whereas each of the tested Bam mutants with disrupted CAF40 binding fails to do so. An mRNA reporter expressing *Renilla* luciferase (R-Luc) served as transfection control. Panel A shows F-Luc activity (grey bars) and mRNA levels (black bars) normalized to those of the R-Luc control and set to 100% in cells expressing the λN-HA peptide. The bars represent mean values and error bars represent standard deviations from three independent experiments. Panel B shows a northern blot of representative RNA samples. Figure modified from Sgromo et al. (2018).

4.2.2 The CBM_B-CAF40 interaction is required for Bam-mediated mRNA decay

My results clearly showed that Bam requires the interaction with CAF40 for the recruitment of the assembled CCR4-NOT complex and induces target mRNA degradation. To exclude that other factors facilitated Bam repressive activity, I depleted S2 cells of the CAF40 subunit of the CCR4-NOT complex using dsRNA against *CAF40* mRNA. Tethering of Bam in cells depleted of CAF40 strongly impaired reporter repression. This effect could be rescued by overexpressing a dsRNA resistant version of wt CAF40 (where silent mutations impair dsRNA-mediated recognition without affecting the protein product) in a complementation assay shown in Figure 14. To support the idea that Bam-mediated repression depends entirely on CAF40, a dsRNA resistant version of the CAF40 mutant (V186E), that is unable to bind Bam or Roquin, was tested. Accordingly, the CAF40 mutant was not able to rescue Bam activity (Figure 14 A,B).

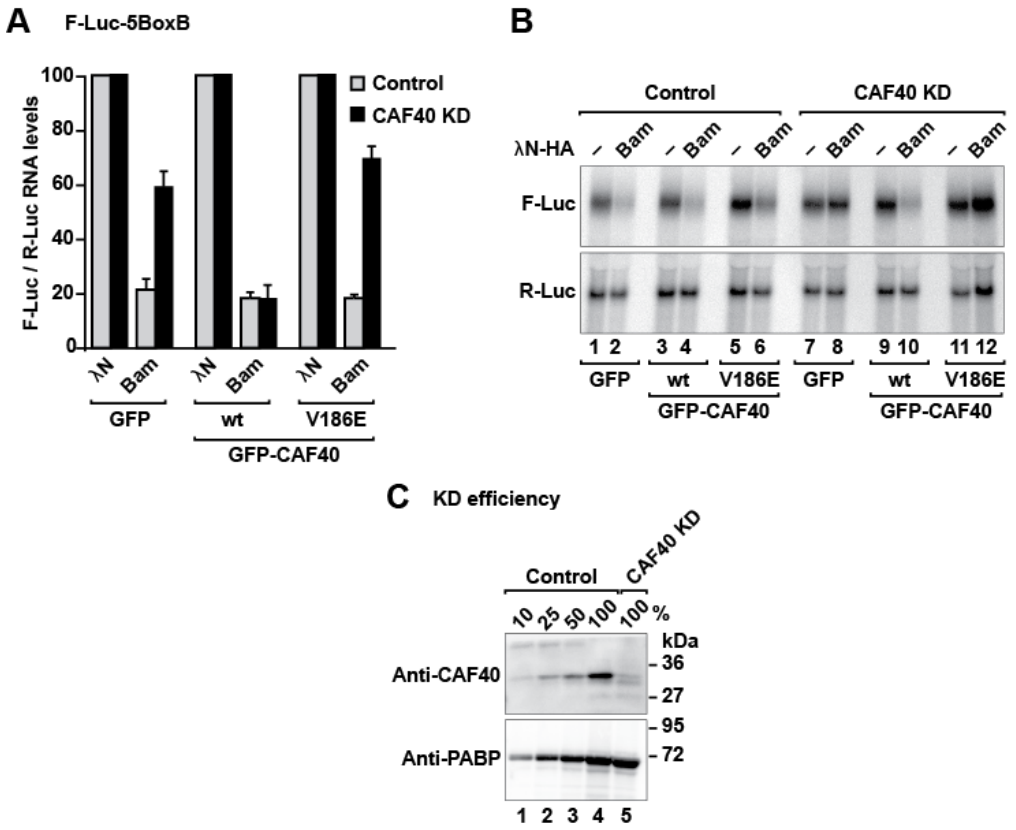


Figure 14: Complementation assay in *Dm* S2 cells. Complementation assay using the F-Luc-5BoxB reporter and λN-HA-tagged Bam in knockdown control S2 cells (control, gray bars) or in cells depleted of CAF40 (CAF40 KD, black bars). Panel A shows the relative mRNA levels from three independent experiments and panel B shows a representative northern blot. Panel C shows the efficiency of the CAF40 knockdown. Protein size markers (kDa) are shown on the right. Figure modified from Sgromo et al. (2018).

This new insight into the role of Bam reveals that Bam requires CAF40 and the assembled CCR4-NOT complex to elicit its activity in mRNA degradation. GSC differentiation requires a Bam repressor complex to regulate key transcripts and my work suggests how such repression could occur: Bam could posttranscriptionally repress specific GSC maintenance factors through the direct recruitment of the CCR4-NOT complex. However, it remains to be elucidated how the Bam repressor complex recognizes and targets the respective mRNA transcripts.

4.3 Bam forms a complex together with Bgcn and Tut

The work described in this section is not published yet. Additional information regarding experimental procedures and methods are provided in section 8.3.

4.3.1 Bgcn induces mRNA decay via its N-terminal region

Bam was proposed to be part of a large repressor complex that regulates GSC differentiation in flies (Neumuller et al. 2008, Li et al. 2013, Chen et al. 2014a, Shan et al. 2017), however, the molecular details of Bam function remained poorly defined. To fully investigate the role of Bam in GSC differentiation, it is important to characterize Bam interactions with proposed co-factors (Shen et al. 2009, Chen et al. 2014a, Malik et al. 2017). Therefore, I started investigating the putative Bam binding partners Bgcn and Tut and their interaction with Bam. Bgcn was already shown to be an intimate binding partner of Bam (Ohlstein et al. 2000, Li et al. 2009) and to interact with the C-terminal region of Bam (Pan et al. 2014). Tut is an interesting binding partner since it is an RBP and may represent the missing link between Bam and target mRNAs. Interestingly, *tut* and *bgcn* mutant flies show a similar phenotype as *bam* mutants, suggesting genetic interaction. Indeed, Bgcn and Bam were suggested to regulate germ cell differentiation (Li et al. 2009). To examine whether Bam and Bgcn share similar functions and repress target mRNAs, I performed a λN-based tethering assay in S2 cells together with a master student, Charlotte Backhaus.

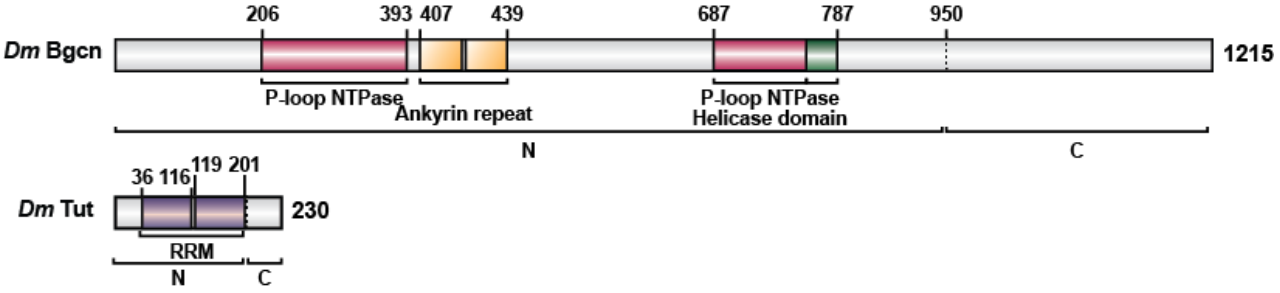


Figure 15: Domain organization of *Dm* Bgcn and *Dm* Tut proteins. Bgcn is a DExH-box RNA helicase but it lacks ATP-binding and RNA unwinding motifs (Ohlstein et al., 2000). It contains two ankyrin repeats (yellow). Tut is a fly-specific RNA-binding protein that contains two RRM domains (purple). Protein domain organizations and N- and C- fragments are indicated.

I expressed either full-length Bgcn or N- or C-terminal regions (Figure 15) with an N-terminal λN-HA tag that binds to a co-transfected firefly luciferase mRNA reporter, containing 5BoxB hairpins (which are λN-binding sites) in the 3'UTR (for detailed experimental procedures see Supplemental information in section 8.3). Tethering of Bgcn to the F-Luc-5BoxB reporter mRNA decreased the F-Luc reporter levels to approximately 20% relative to the negative control (tethering of λN-HA fusion protein). Importantly, Bgcn activity resided in the Bgcn-N fragment,

while the Bgcn-C fragment was not active (Figure 16 A). The decrease in F-Luc activity corresponds to a decrease in the mRNA levels (Figure 16 B), thereby showing that similarly to Bam, Bgcn (in particular the N-terminal region; Bgcn-N) also induces mRNA decay in S2 cells. This seems to be independent of the presence of Bam, since endogenous Bam levels are very low in S2 cells (Shen et al. 2009).

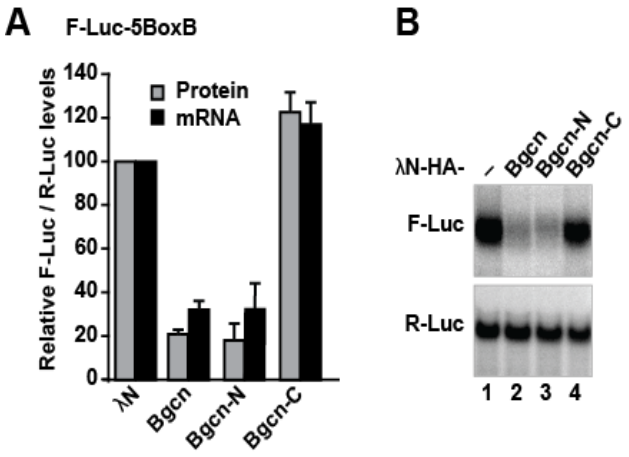


Figure 16: The Bgcn N-terminal region induces degradation of the F-Luc-5BoxB mRNA reporter. Tethering assay using λN-HA-tagged Bgcn (full-length, N-terminal and C-terminal regions) to the F-Luc-5BoxB reporter in *Dm* S2 cells. Bgcn induces decay of the reporter mRNA via its N-terminal region. An mRNA reporter expressing *Renilla* luciferase (R-Luc) served as transfection control. Panel A shows F-Luc activity (grey bars) and mRNA levels (black bars) normalized to those of the R-Luc control and set to 100% in cells expressing the λN-HA peptide. The bars represent mean values and error bars represent standard deviations from three independent experiments. Panel B shows a representative northern blot of the experiment in panel A.

Bgcn-mediated mRNA decay could depend on the interaction with the CCR4-NOT deadenylase complex. I therefore used co-IPs to identify several interacting subunits within the CCR4-NOT complex (NOT1, NOT2, NOT11 and CCR4) that bind to Bgcn *in vivo* (Figure 17 A-D). However, a precise mapping of the interaction *in vitro* is currently missing due to difficulties in expressing and purifying Bgcn protein fragments from *E.coli*.

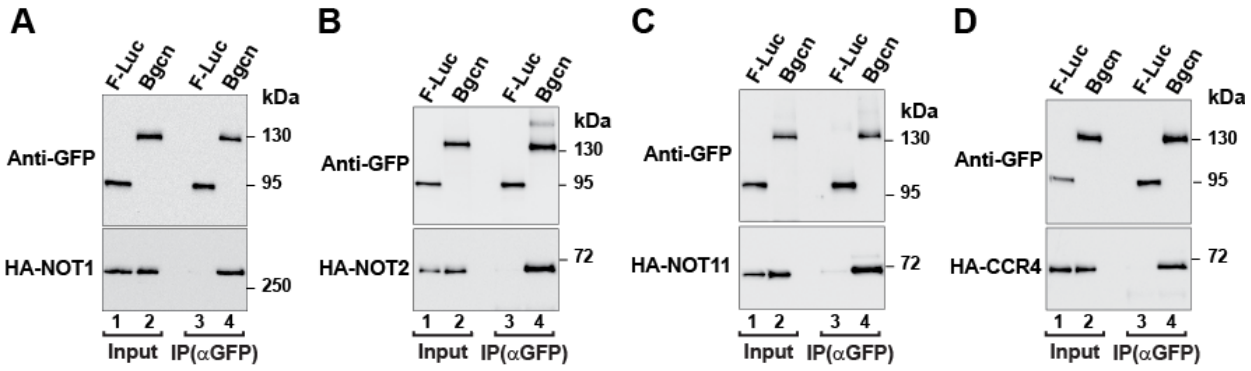


Figure 17: *Dm* Bgcn co-immunoprecipitates several subunits of the CCR4-NOT complex. Western blots showing anti-GFP co-IPs of GFP-tagged Bgcn with HA-tagged subunits of the CCR4-NOT complex in *Dm* S2 cells, in presence of RNaseA. Bgcn interacts with NOT1 (A), NOT2 (B), NOT11 (C) and CCR4 (D). GFP-F-Luc served as 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.

4.3.2 Tut recruits the CCR4-NOT complex via Bam

Recently it was shown that Tut interacts with the N-terminal region of Bam in yeast two-hybrid assays (Chen et al. 2014a). A detailed mapping of this interaction was however still missing. In collaboration with two former students in the lab, Kevin Sabath and Andreas Blaha, I determined the minimal regions in Bam and Tut required for their direct interaction, combining *in vivo* co-IPs and *in vitro* pulldown assays. Bam uses a TBM (Tut binding motif; residues 79-103) in its N-terminal region to interact directly with the Tut N-terminal region (which includes the two RRMs) (result obtained by Kevin Sabath). This peptide is distinct from the CBM_B (residues 17-36) used to bind to CAF40. These data, together with the already known interaction between Bam and Bgcn, suggested the formation of a trimeric Bam-Bgcn-Tut complex.

To gain molecular insights into the repressive activity of the Bam-Tut complex I performed luciferase assays using a λN-based tethering approach in S2 cells. I tethered λN-HA-tagged Tut to the F-Luc-5BoxB mRNA reporter, while co-expressing GFP-tagged Bam constructs. I used either full-length Bam, N-terminal or C-terminal Bam fragments, the Bam mutant (4xMut) that does not bind to CAF40 as shown previously (Sgromo et al. 2018) or GFP alone (as a control). Under control conditions, λN-HA-tagged Tut did not cause mRNA decay but rather increased reporter levels (Figure 18 A,B; lanes 1,2). In contrast, tethering of λN-HA-tagged Tut decreased the F-Luc expression levels to 30% relative to the negative control (λN-HA fusion protein) when co-transfected with GFP-Bam full-length or the GFP-Bam N-terminal fragment that both carry the CBM_B and the TBM (Figure 18 A,B; lanes 4 and 6). Notably, tethered Tut requires Bam to repress bound mRNA, most likely via Bam-mediated recruitment of the CCR4-NOT complex. Thus, Tut

triggered mRNA degradation only when co-expressed with functional Bam in S2 cells. Importantly, tethering of λN-HA-Tut did not decrease the reporter mRNA levels when co-expressed with the GFP-Bam C-terminal fragment or the GFP-Bam mutant (4xMut) (Figure 18 A,B; lanes 8 and 10). Thus, Tut and Bam were functional only when Bam was able to directly recruit the CCR4-NOT complex via its N-terminal region where the CBM_B is located (Figure 18 A,B). These experiments revealed that Tut is a direct binding partner of Bam and requires Bam to degrade bound mRNAs.

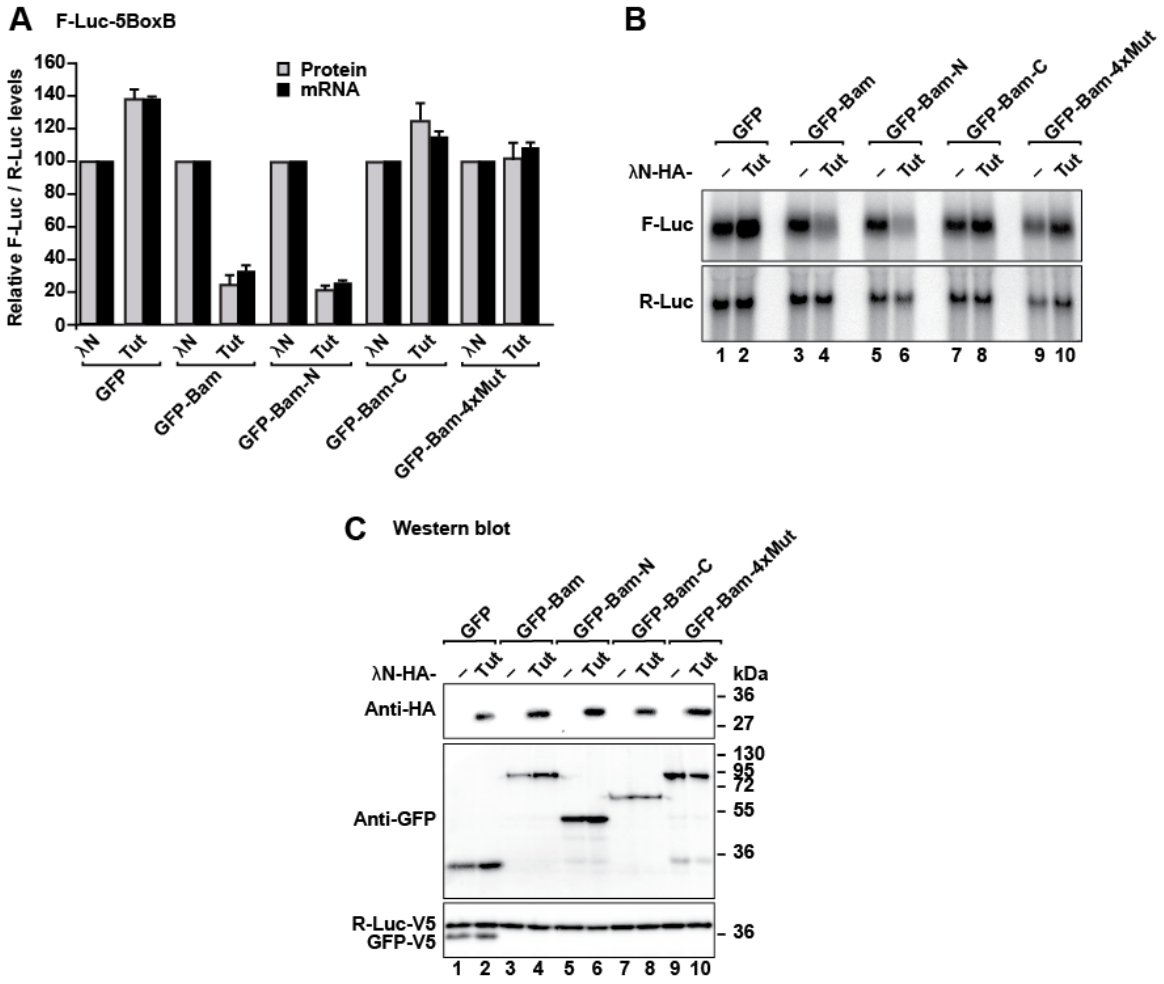


Figure 18: Tut degrades bound mRNAs when expressed together with Bam. Tethering assay of λN-HA-Tut onto an F-Luc-5BoxB reporter in the presence of Bam. Panel A shows F-Luc activity and mRNA levels normalized to those of the R-Luc control and set to 100% in cells expressing the λN-HA peptide. The bars represent mean values ± standard deviations from three independent experiments. Panel B shows a northern blot of representative RNA samples. Panel C shows the equivalent expression of proteins used in the experiments shown in panels A and B. Protein size markers (kDa) are shown on the right.

Moreover, Tut may provide RNA-binding to the Bam-Bgcn-Tut repressor complex, since it contains a tandem RRM domain. To validate this hypothesis, I performed a similar experiment to the one shown in Figure 18, using a modified reporter system, where I replaced the 5BoxB elements

with the 3'UTR of *mei-P26* (F-Luc-mei-P26), a proposed endogenous mRNA target of the Bam-Bgcn-Tut complex (Chen et al. 2014a). This reporter assay does not rely on the artificial tethering of Tut, but rather on its intrinsic RNA binding ability. Unexpectedly, Tut did not affect the expression of this reporter. Figure 19 A shows protein levels of the firefly luciferase reporter carrying the *mei-P26* 3'UTR when Bam (either wt protein, the N- or C- terminal regions or the mutant version) and Bgcn were overexpressed. This finding is in conflict with previous data (Chen et al. 2014a) and suggests that the complex might require additional co-factors to modulate and/or bind the RNA.

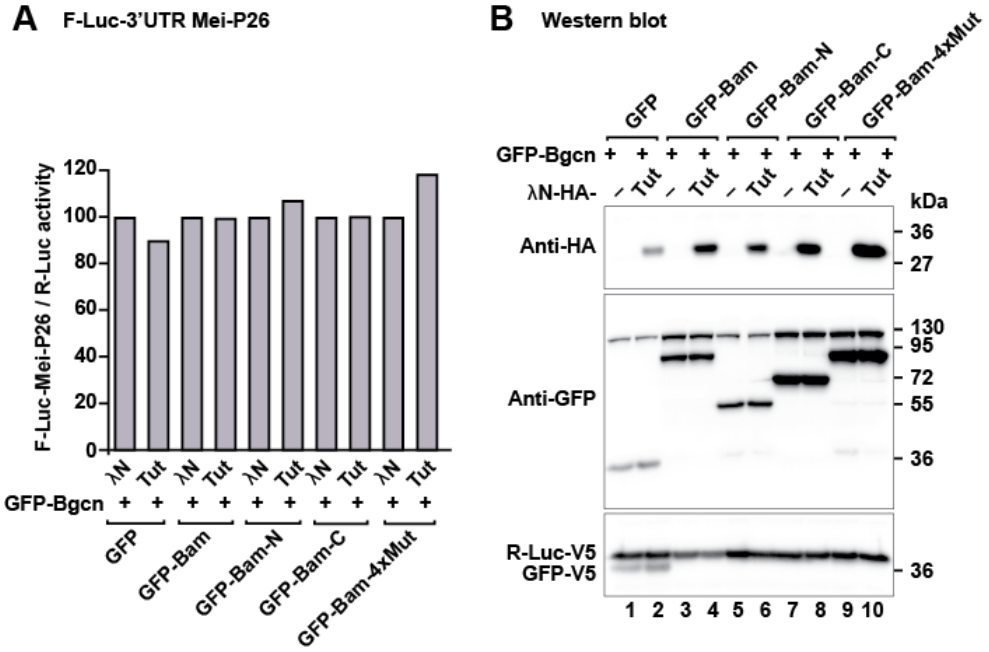


Figure 19: Bam, Bgcn and Tut do not affect the expression of a reporter containing the *mei-P26* 3'UTR, when co-expressed. Panel A shows F-Luc activity normalized to the R-Luc control and set to 100% in cells expressing the λN-HA peptide alone. Panel B shows the equivalent expression of proteins used in the experiment shown in panel A. Protein size markers (kDa) are shown on the right.

Overall, these results identified for the first time the minimal regions required for a direct interaction between Bam and Tut. The data suggest the formation of a trimeric Bam-Bgcn-Tut complex with Tut acting as an RNA-binding protein, possibly for the direct recognition of specific mRNAs via its tandem RRM domains. Unexpectedly, I could not confirm previous findings that Tut binds to the *mei-P26* 3'UTR (Chen et al. 2014a). Thus, it would be of great interest to further characterize at the molecular and structural level the architecture of this trimeric complex as well as its recruitment to target mRNAs. This will help to elucidate the complete mechanism by which Bam, Bgcn and Tut act through the CCR4-NOT complex to regulate GSC differentiation.

4.4 CAF40 is an important hub for the recruitment of the CCR4-NOT complex

My two publications (Sgromo et al. 2017 and Sgromo et al. 2018) contain parts of the work summarized in this section. The papers including detailed experimental data and methods are attached. The work involving NOT4 is described in the attached manuscript in preparation Keskeny et al. (see chapter 9).

4.4.1 *Dm* NOT4 also carries a CBM that directly binds the CCR4-NOT complex

The identification of two distinct CBMs in Roquin and Bam targeting the same surface on the CAF40 subunit of the CCR4-NOT complex raised the question whether there are yet other RNA-binding proteins containing a CBM. Indeed, in collaboration with the PhD students Csilla Keskeny and Tobias Raisch, we identified, by *in vitro* pulldown assays, metazoan NOT4 as another protein that carries a CBM (CBM_N, where the subscript N refers to NOT4).

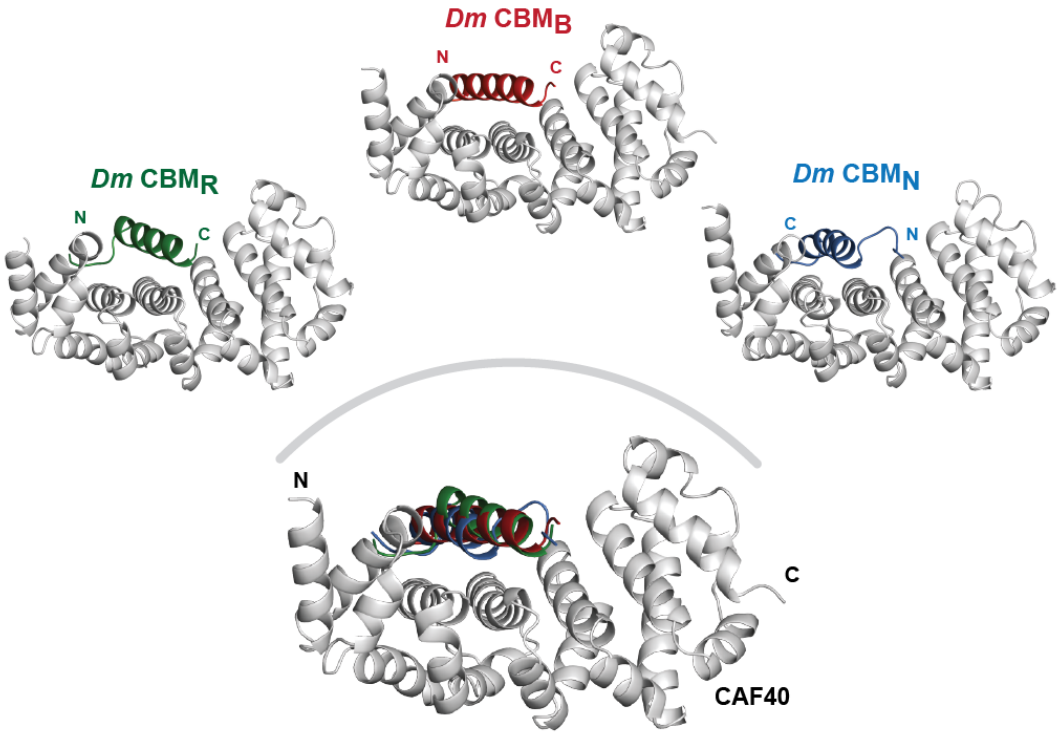


Figure 20: Crystal structures of the CBMs bound to CAF40. *Dm* Roquin (green), *Dm* Bam (red) and *Dm* NOT4 (blue) CBMs bound to *Hs* CAF40 are illustrated. *Hs* CAF40 consists of six ARM repeats (gray). All CBMs fold into an α-helix and bind to the concave surface of CAF40. The bottom of the figure shows a superimposition of all CBM structures. Figure adapted from Sgromo et al. (2017), Sgromo et al. (2018) and the Keskeny et al. NOT4 manuscript.

The CBM_N sequence resides in the C-terminal region of *Dm* NOT4 and binds to the same concave surface of CAF40 as the two previously identified CBMs, CBM_R and CBM_B. In contrast to the *Dm* CBM_R, the *Dm* CBM_N is conserved in metazoan NOT4 proteins and functional in *Hs* NOT4, which

additionally contains a NOT module binding motif (NBM) that shares no sequence homology with the *Dm* Roquin NBM.

Interestingly, the crystal structure of the *Dm* CBM_N in complex with CAF40 determined by Tobias Raisch, showed that the *Dm* NOT4 CBM_N peptide also forms an α-helix, but that it binds to CAF40 in an N-to-C inverted orientation, when compared to the CBMs of Roquin and Bam. Even though all CBMs lack sequence similarity, they all fold into an amphipathic helix to interact with the same hydrophobic surface on CAF40 (Figure 20). These findings imply that the same binding surface within the CCR4-NOT complex (the concave CAF40 interface) can be bound by multiple motifs (CBMs) that evolved independently of each other by convergent evolution to bind CAF40 and recruit the CCR4-NOT complex.

To assess the role of the three CAF40-CBM interactions, I first validated the structures both *in vivo* and *in vitro* (the *in vitro* validation of the CAF40-CBM_N interaction was performed by Tobias Raisch). Similar to Roquin and Bam (see sections 4.1 and 4.2) NOT4 can induce decay of bound mRNAs via the direct interaction with the CCR4-NOT complex, through a CMB_N and an NBM.

4.4.2 CBMs show common properties but differences in binding CAF40

Even though the CBMs of Roquin, Bam and NOT4 bind to the CAF40 subunit of the CCR4-NOT complex in a very similar way, they contribute differently to the recruitment of the CCR4-NOT complex and to the resulting mRNA decay activity. While the *Dm* CBM_B is necessary and sufficient to trigger Bam-mediated mRNA degradation (section 4.2), the *Dm* CBM_R and *Dm* CBM_N alone only trigger partial target mRNA decay in the tethering assay (section 4.1 and Keskeny et al. manuscript in preparation). This reflects the fact that NOT4 and Roquin rely on additional interactions outside of their CBMs for the recruitment of the CCR4-NOT deadenylase complex.

The requirement for additional interactions in the recruitment of the CCR4-NOT complex could be explained by a low-affinity binding of the CBM_R and CBM_N to CAF40. It was thus interesting to determine the binding affinities of the different CBMs for the purified CAF40 module. To measure the respective dissociation constants, I performed isothermal titration calorimetry (ITC) measurements together with Tobias Raisch. We were able to determine the binding affinity between the CBM_B peptide and the CAF40 module (dissociation constant of ~180 nM). However, measuring the Roquin and NOT4 CBM binding affinities was not possible because the elevated protein concentrations required for the measurements led to protein precipitation. The need for high CBM_R and CBM_N peptide concentrations implies a rather low-affinity interaction of these peptides with the CAF40 module in comparison to the CBM_B peptide. Assuming that Roquin,

Bam and NOT4 all recruit the CCR4-NOT complex with comparable efficiency, a low-affinity binding of the CBM_R and CBM_N peptides would explain why additional binding sites (leading to avidity effects) are needed within Roquin and NOT4 to recruit the CCR4-NOT complex.

The presence of a CBM in Roquin, Bam and NOT4 raises the question whether these three proteins compete *in vivo* for the binding to CAF40. Currently it is unknown if the three identified CBM-containing proteins are present at the same time in a cell which is a prerequisite for a competitive binding scenario. It is thus possible that Roquin, Bam and NOT4 compete for binding to CAF40 under specific cellular conditions.

In conclusion, my results reveal for the first time that the highly conserved CAF40 subunit of the CCR4-NOT complex is an important hub since it is not only bound by TNRC6 proteins (Chen et al. 2014b, Mathys et al. 2014), which contact the CAF40 convex surface, but also by Roquin, Bam and NOT4, which contact the CAF40 concave surface via their CBMs. However, the identification of new CBMs in other proteins remains challenging, since all three CBMs identified here lack sequence conservation.

5. DISCUSSION

The CCR4-NOT complex regulates mRNA metabolism from transcription to decay. In addition to its role in bulk mRNA turnover, it is required for deadenylation and translational repression of a plethora of mRNAs to which it is recruited via its interaction with RNA-associated proteins. My doctoral work, together with other studies, highlights the CCR4-NOT complex as a major effector for mRNA repression mediated by specific binding partners, such as Roquin and Bam. My findings not only reveal the structural and molecular basis for the recruitment of the CCR4-NOT deadenylase complex by these two proteins, but they also show how Roquin and Bam trigger mRNA decay. Importantly, I uncover a previously unknown binding surface in CAF40 as a contact site within the CCR4-NOT complex that is shared by these two proteins, and by NOT4. This is the first time that functionally distinct proteins are found to contact an overlapping interface within the CCR4-NOT complex, using a short α-helical motif. The discovery of a common CBM-binding surface on CAF40 identifies CAF40 as a common hub for mRNA-associated proteins to recruit the CCR4-NOT complex. A cartoon summary of the recruitment of the CCR4-NOT complex by the mRNA-associated proteins characterized in my studies is shown in Figure 21.

5.1 Roquin-mediated mRNA decay is conserved through evolution

In my PhD work, I first investigated the role of *Dm* Roquin, a previously uncharacterized protein, in the recruitment of the CCR4-NOT complex. My results showed that Roquin proteins induce mRNA degradation via the direct interaction with the CCR4-NOT deadenylase complex in both human and fly. In *Drosophila*, Roquin interacts with the CCR4-NOT complex through two distinct motifs in its unstructured C-terminal region that contact the CAF40 module (CBM_R) and the NOT module (NBM) to mediate deadenylation and to target mRNAs for decay.

Interestingly, both motifs (CBM_R and NBM) are not conserved between *Dm* and *Hs* Roquin proteins, suggesting that the proteins have maintained the ability to recruit the CCR4-NOT complex to mRNA targets even though the precise mode of binding has changed during evolution. Previous studies on *Hs* Roquin1 showed that it recruits the CCR4-NOT complex via its unstructured C-terminal tail (Leppek et al. 2013). I could confirm this interaction and additionally show that also *Hs* Roquin2 recruits the CCR4-NOT complex via its C-terminal tail (see section 4.1). Since the C-terminal tails of Roquin proteins are divergent (also between *Hs* Roquin1 and 2), this supports the concept that the recruitment of the CCR4-NOT complex is functionally conserved but differs at the molecular level.

Another feature that appears to be maintained through evolution is the recognition of specific RNA elements by the ROQ domain. This characteristic RNA binding domain is highly conserved between all Roquin proteins. My results showed that *Dm* Roquin is capable of using its ROQ domain to recognize the CDE of TNF-α, a known target of the *Hs* Roquin1 ROQ domain, but it is unknown whether the *Dm* Roquin ROQ domain has the same mRNA targets in the fly. Even though the *Drosophila* immune system differs largely from the vertebrate immune system (Govind 2008), *Drosophila* has a TNF-α homolog called Eiger. Eiger induces cell-death and is involved in the immune response (Perez-Garijo et al. 2013, Igaki & Miura 2014). However, the *Dm eiger* 3'UTR lacks evident CDEs, suggesting that *Dm* Roquin could have other target mRNAs than its human counterpart. To date, the identity of endogenous *Dm* Roquin targets remains unknown, raising the question whether they also regulate inflammation and immunity. Since *Dm* Roquin binds the CDE of *Hs* TNF-α, it would be interesting to screen the *Drosophila* transcriptome for consensus CDEs, as has been previously done for the mouse transcriptome (Leppek et al. 2013). This would reveal endogenous *Dm* CDE-containing mRNAs as potential *Dm* Roquin targets. To identify mRNAs that are bound by Roquin in *Drosophila*, endogenous Roquin immunoprecipitation followed by RNA sequencing could be used in future experiments. Alternatively, RNA sequencing experiments in combination with *Dm* Roquin knockdown will provide a general picture of mRNAs that are affected by *Dm* Roquin. It will be interesting to discover the biological context in which *Dm* Roquin recruits the CCR4-NOT complex, considering the existing evidence that *roquin* is essential for fly development (Smibert et al., abstract for the 48th Annual Drosophila Research Conference, 2007).

5.2 Bam and NOT4 use a CBM to recruit the CCR4-NOT complex

From simple sequence analysis of Bam, the second protein that I investigated as a putative CCR4-NOT interactor, the presence of a CBM was not obvious. Thus, it was surprising to discover that Bam carries a CBM_B and that the CBM_B-CAF40 interaction is highly similar to the CBM_R-CAF40 interaction identified previously. Bam is a GSC differentiation factor in *Dm* and the identified CBM_B is crucial for the direct interaction with CAF40 within the CCR4-NOT deadenylase complex. Importantly, my results shed light on the function of *Dm* Bam and show that Bam requires the integrity of the CCR4-NOT complex for mRNA repression, in contrast to the previous work of Fu et al. (2015). The model proposed by Fu et al. was based on an interaction between Bam and CCR4. They proposed a mutually exclusive binding of either Bam or CAF1 to CCR4. The authors based their observation on mutations in the LRR domain of CCR4 that disrupted both the binding to Bam and CAF1. However, the mutations introduced most likely

destabilized the fold of the LRR domain of CCR4, as suggested by the available CCR4 structural information (Basquin et al. 2012), thus impairing its incorporation into the CCR4-NOT complex. Fu et al. showed that the CCR4 LRR mutants cannot rescue *twin* mutant flies, which is consistent with their model where the LRR mutations disrupt a common binding interface for CAF1 and Bam (Fu et al. 2015). However, Fu et al. did not prove the functional integrity of these CCR4 LRR mutants, raising the possibility that the mutants are misfolded and bind neither Bam nor CAF1. Thus, it remained unclear whether Bam acts by recruiting the CCR4-NOT complex or by binding to CCR4 independently of the complex. My results confirmed the binding between Bam and CCR4, but also demonstrated that this interaction is actually indirect. Moreover, my studies revealed that Bam directly interacts with CAF40 to recruit the whole CCR4-NOT complex to target mRNAs (see section 4.2). Bam requires the CCR4-NOT complex for its activity since depletion of the scaffold subunit NOT1, which disrupts the assembly of the complex, abolishes Bam-mediated mRNA degradation.

While Bam is a known GSC differentiation factor that represses GSC maintenance (McKearin & Ohlstein 1995), a function of CAF40 in GSC differentiation has not been described yet. A role for CAF40 in this process seems likely, considering that the Bam-CAF40 interaction is necessary for the direct recruitment of the CCR4-NOT complex, a major deadenylase. Thus, it is possible that Bam triggers deadenylation, translational repression and degradation of yet unknown GSC-maintaining target mRNAs, via CAF40. Bam might act as an adaptor between specific transcripts and the mRNA decay machinery in differentiating germ cells. In this light, it will be crucial to study the physiological importance of the observed CAF40-CBM_B interaction, which is sufficient to recruit the CCR4-NOT complex in S2 cells. Future experiments could address this question by assessing the impact of Bam point mutations in fly GSCs.

Unexpectedly, a CBM_N that directly binds CAF40 within the CCR4-NOT complex was also identified in the *Hs* and *Dm* NOT4 proteins, in a collaborative project with my colleagues Csilla Keskeny and Tobias Raisch. In contrast to the CBM_R, the *Dm* CBM_N sequence is conserved in metazoans. Interestingly, *Hs* and *Dm* NOT4 both carry additional motifs to interact with the CCR4-NOT complex. The presence of additional binding sites within NOT4 indicates a possible redundancy in the recruitment of the CCR4-NOT complex, similar to Roquin proteins. The fact that also yeast NOT4 interacts with the CCR4-NOT complex suggests that NOT4 proteins have evolved to maintain the interaction with the CCR4-NOT complex. There are however, two striking differences between *Hs*/*Dm* NOT4 and the yeast NOT4. First, yeast NOT4 is a stable component of the CCR4-NOT complex, whereas the interaction in *Hs* and *Dm* appears to be transient or regulated (Lau et al. 2009, Temme et al. 2010). Second, yeast NOT4 recruits the CCR4-NOT complex via the

NOT module (Bhaskar et al. 2015), whereas *Hs* and *Dm* NOT4 additionally bind to CAF40. It is clear that the interaction of NOT4 proteins with the CCR4-NOT complex is important, however it is currently unknown if *Hs* and *Dm* NOT4 have additional functions other than binding the CCR4-NOT complex. Furthermore, it remains elusive, whether NOT4 uses its RRM domain to recruit the CCR4-NOT complex to specific mRNA targets or whether the CCR4-NOT complex recruits the E3 ubiquitin ligase NOT4 for independent functions. Finally, it is conceivable that the main function of *Hs* and *Dm* NOT4 is independent of the CCR4-NOT complex and requires a transient interaction with the CCR4-NOT machinery.

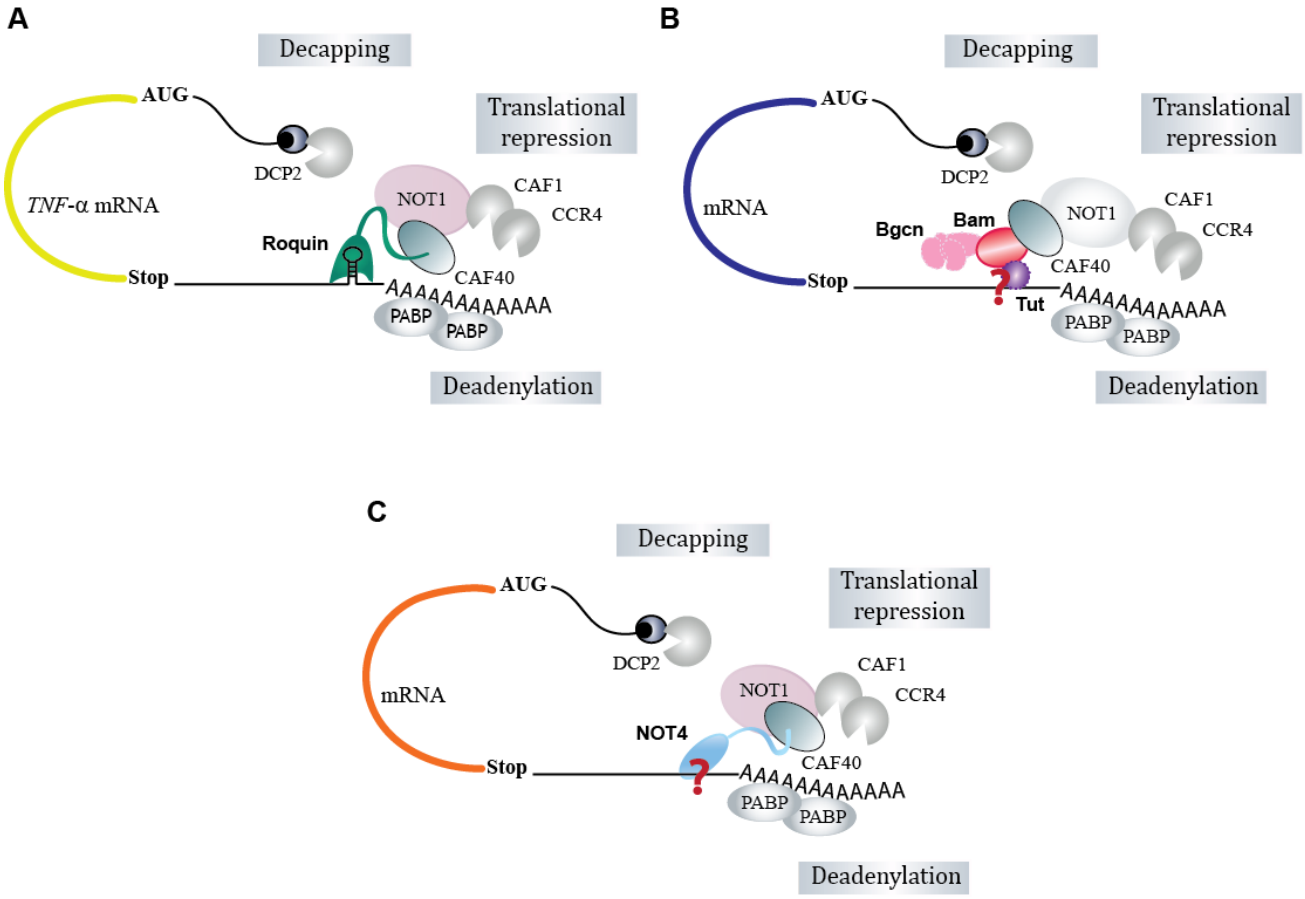


Figure 21: Models for the recruitment of the CCR4-NOT deadenylase complex by *Dm* Roquin, *Dm* Bam and *Dm* NOT4 proteins. A) *Dm* Roquin with its ROQ domain directly recognizes stem loop-structured elements in the 3'UTR of target mRNAs and with the CBM_R and NBM directly recruits the CCR4-NOT complex. B) *Dm* Bam directly interacts with the CAF40 subunit of the CCR4-NOT complex via a CBM_B. Additionally, Bam forms a complex together with Bgcn and Tut proteins. It is still unknown if Bam or Tut directly contact target RNAs or require binding to additional RBPs. C) *Dm* NOT4 might contact target mRNAs with its N-terminal RRM domain and it directly binds to CAF40 (via a CBM_N) and carries additional yet unidentified binding sites to contact the CCR4-NOT complex. In all of the proposed models, the recruitment of the CCR4-NOT complex leads to deadenylation followed by decapping and degradation, as well as translational repression of target mRNAs.

5.3 The Bam-Bgcn-Tut repressor complex

The identification of the Bam CBM_B indicates that the recruitment of the CCR4-NOT complex is important for Bam's function in GSC differentiation. However, it is not clear whether Bam requires other binding partners such as Bgcn or Tut to elicit its function in posttranscriptional regulation. My preliminary results indicate that Bam and Bgcn can both induce mRNA degradation independently of each other, via the CCR4-NOT effector complex. However, it remains unclear whether Bgcn can directly recruit the CCR4-NOT complex or requires additional proteins. In the context of a trimeric Bam-Bgcn-Tut repressor complex (Chen et al. 2014a), the finding that both Bam and Bgcn execute a similar function is surprising. Further studies are required to understand whether Bam and Bgcn act redundantly or cooperate to recruit the CCR4-NOT complex via multiple contacts. Additionally, it remains unknown whether Bam can directly bind mRNA by itself, or requires other RBPs such as Tut.

Bam, Bgcn and Tut form a trimeric complex that represses GSC maintenance (Chen et al. 2014a), however, a detailed characterization of the interactions within the complex is currently missing. Tut is an especially interesting protein within this complex, because it contains two RRMs (section 4.3) which might explain how Bam associates with RNAs *in vivo*. It can be speculated that within the Bam-Bgcn-Tut complex functions such as RNA binding and CCR4-NOT recruitment are distributed among the different proteins, rather than coming from a single polypeptide chain as it is the case for Roquin.

Work done by Kevin Sabath demonstrated that Bam directly contacts Tut and identified the minimal regions for this interaction. Moreover, my results showed that Tut requires the interaction with Bam to elicit repression of an mRNA reporter via the CCR4-NOT complex in tethering assays. Thus, Tut may confer RNA-binding activity to the Bam-Bgcn-Tut complex, which in turn recruits the CCR4-NOT complex via Bam. However, Tut does not mediate the degradation of an mRNA reporter carrying the *mei-P26* 3'UTR, even in the presence of Bam and Bgcn (see section 4.3). Previous work showed by RNA immunoprecipitations that Tut can bind a long and a short isoform of the *mei-P26* 3'UTR (Chen et al. 2014a). Interestingly, binding seems to be weaker for the short, wild type isoform of the *mei-P26* 3'UTR (Chen et al. 2014a) which I used in my reporter construct. This could explain my result, where Tut did not reduce the levels of this reporter mRNA. Thus, it would be interesting to identify additional endogenous Tut target mRNAs to test their regulation by the Bam-Bgcn-Tut complex in reporter binding assays. Photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation (PAR-CLIP) experiments of Tut (in the presence or absence of Bam/Bgcn) could be used to identify direct targets of the complex.

The CCR4-NOT complex has critical roles in the germline, where it regulates Nanos mRNA targets in GSCs. My finding that the GSC master differentiation factor Bam directly recruits CAF40 indicates that the CCR4-NOT complex also regulates stem cell differentiation. However, the functional relevance of the proposed Bam-Bgcn-Tut repressor complex and its connection to the CCR4-NOT complex *in vivo* remain to be elucidated.

5.4 CAF40 as a common binding platform for RNA-associated proteins

An important and unexpected common feature of Roquin, Bam and NOT4 is the interaction with a previously unknown binding surface on CAF40. Binding interfaces on the CCR4-NOT complex are known for other RNA-associated proteins, such as for GW182 (which binds CAF40 on its convex surface; Chen et al. 2014b, Mathys et al. 2014), as well as for *Hs* and *Dm* Nanos (which contact the NOT module on distinct surfaces; Bhandari et al. 2014, Raisch et al. 2016). In general, these proteins use distinct interaction modes to bind the CCR4-NOT complex. It is thus surprising that the sequence-wise unrelated CBMs of Roquin, Bam and NOT4 all fold into an amphipathic α-helix to contact the same site on the concave surface of CAF40. Importantly, the concave CAF40 surface is highly conserved, even in species where CBMs where not identified so far, for example in yeast. Thus, it is likely that CBM-containing proteins are also present in these species, but remain to be discovered.

Notably, even though all of the CBM-containing proteins that I identified in my studies fold into an amphipathic α-helix to contact CAF40, they share no sequence conservation. This strongly suggests that the CBMs arose from convergent evolution, a characteristic feature of short linear motifs (SLiMs) (Tompa 2012, Van Roey et al. 2014). Moreover, the presence of a CBM in Roquin, Bam and NOT4 indicates that their binding to CAF40 is mutually exclusive. Thus, it is possible that the CBMs compete for the same binding site under certain cellular conditions. Competition for CAF40 binding might provide a stringent posttranscriptional mRNA regulation to coordinate different biological processes that depend on CCR4-NOT activity. However, to date, there is no information regarding the simultaneous expression of these proteins within a cell. Since Bam is exclusively expressed in the germline (Brown et al. 2014), competition is more likely to occur between the two E3 ubiquitin ligases Roquin and NOT4 that are broadly expressed in *Dm*.

Several non-catalytic modules of the CCR4-NOT complex have regulatory functions and are now known to be bound by mRNA-associated proteins. For example, proteins like *Hs* and *Dm* Nanos and yeast NOT4 have been shown to interact with the NOT module on distinct surfaces (Bhaskar et al. 2015). With the present work however, the CAF40 module emerges as a major hub

within the CCR4-NOT complex that is competitively targeted by several mRNA-associated proteins, whereas proteins of the TNRC6/GW182 family can probably still contact CAF40 on the convex surface and recruit the CCR4-NOT complex to miRNA targets (Chen et al. 2014b, Mathys et al. 2014). The fact that the CBM-containing proteins bind CAF40 implies that CAF40 plays a general role beyond miRNA-mediated gene silencing and represents a central contact site in the CCR4-NOT complex for mRNA-associated proteins.

5.5 An evolutionary perspective on CBMs in the recruitment of the CCR4-NOT complex

Multiple motifs carry out CCR4-NOT recruitment in proteins like TNRC6/GW182 (Chen et al. 2014b, Mathys et al. 2014), TTP (Fabian et al. 2013), Roquin (see section 4.1), Nanos (Bhandari et al. 2014, Raisch et al. 2016), Bam (see section 4.2) and NOT4 (Bhaskar et al. 2015 and unpublished work by Keskeny et al.). Typically, such proteins use SLiMs (Tompa 2012, Van Roey et al. 2014) located in intrinsically disordered regions to bind directly and specifically to the CCR4-NOT complex, consequently triggering translational repression and mRNA degradation. These motifs are linear, meaning that their secondary structure is often induced upon binding. SLiMs mediate transient low-affinity interactions that generally act cooperatively to promote a more stable binding. They are subject to high evolutionary plasticity, meaning that they evolve rapidly and typically show convergent evolution, which hampers their identification. A clear example of this plasticity is found in Roquin and NOT4 proteins. The Roquin CBM and NBM are not conserved in orthologous proteins across species and represent an example of how multiple low affinity motifs can result in high avidity interactions. At least two redundant motifs are also found in the *Hs* NOT4 protein, however, the NOT4 CBM is conserved between *Hs* and *Dm*.

In contrast to Roquin and NOT4, Bam relies entirely on the interaction between the CBM_B and CAF40 to recruit the CCR4-NOT complex and elicit its repressive activity. This might explain the relatively high binding affinity of Bam for CAF40 measured by ITC. The crucial role of Bam in GSC differentiation might require a single, defined interaction with the CCR4-NOT deadenylase complex rather than a modulated interaction via multiple redundant binding motifs to strictly regulate target mRNAs.

In general, SLiMs appear to be a common feature of the above-described mRNA-associated proteins that recruit the CCR4-NOT complex. SLiMs have been proposed to mediate binding interactions of dynamic processes (Davey et al. 2015). This is possibly also the case for the CBM-containing proteins, where in all of the cases, the recruitment of the CCR4-NOT complex leads to a common repressive mechanism. The removal of the poly(A) tail from target mRNAs, their

translational repression and, in a cell context-dependent manner, their full degradation are all highly regulated processes that certainly benefit from a dynamic interaction with the CCR4-NOT complex.

5.6 A general view on posttranscriptional mRNA regulation

To date, multiple interactions of RNA-associated proteins with the CCR4-NOT complex have been characterized. Based on these findings it is possible to draw a general picture of the recruitment of the CCR4-NOT complex to mRNA targets (Figure 22). Typically, RNA-associated proteins contain an RNA-binding domain recognizing specific mRNA targets, for example the ROQ domain in Roquin (Leppek et al. 2013) or the ZnF in Nanos proteins (Curtis et al. 1997, Hashimoto et al. 2010). This RNA-binding domain is phyiscally distinct from the motifs required to bind the CCR4-NOT complex. RNA-associated proteins have evolved versatile binding modes to recruit the CCR4-NOT complex: i) by using distinct motifs, which usually are not conserved between species and ii) by contacting different modules within the complex.

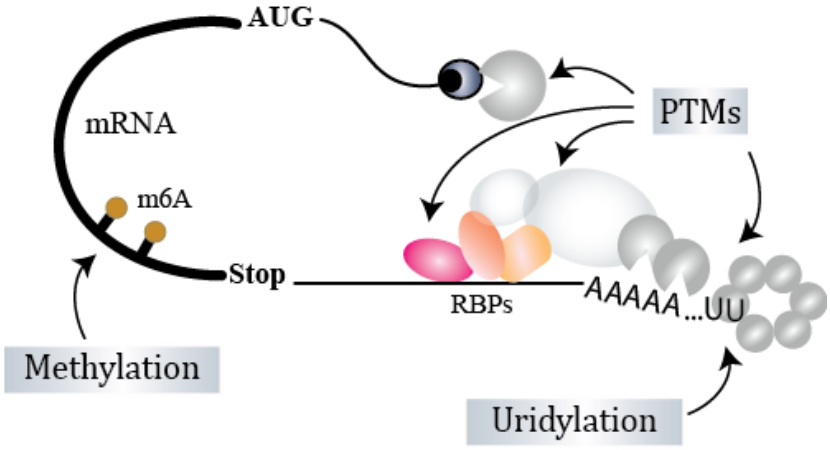


Figure 22: Overview of the complexity of posttranscriptional mRNA regulation. General determinants of mRNA stability are the poly(A) tail and the 5'-cap structure. Specific mRNAs are often recognized by RBPs (red and orange) that recruit effector complexes (grey) to regulate mRNA stability. The complexity of this network is increased by PTMs of RBPs and effector complexes. Furthermore, RNA modifications (here methylation and uridylation are shown as examples) have an impact on mRNA stability.

The context-specific expression of mRNA-associated proteins allows a fine-tuned regulation of numerous mRNA targets. Moreover, PTMs of either the RNA-associated proteins (as is the case for TTP phosphorylation; Clement et al. 2011) or of subunits of the CCR4-NOT complex (as is the case for CAF1 ubiquitination by MEX-3C; Cano et al. 2015) play a key role in posttranscriptional regulation. mRNAs can also be modified, for example by methylation or uridylation (Lee et al.

2014), thus affecting their stability, activity and localization. This overwhelming complexity is additionally increased by the cooperativity of multiple regulatory proteins binding to a single mRNA. It will be interesting, but also challenging considering the level of complexity, to fully understand the molecular mechanisms behind the function and specificity of RNA-associated proteins in posttranscriptional mRNA regulation.

My studies on Roquin proteins and Bam highlight the plasticity in the recruitment of the CCR4-NOT complex to posttranscriptionally regulate mRNAs. Here CAF40 acts as regulatory hub that is contacted by multiple RNA-associated proteins. This emphasizes the importance of mRNA regulation in diverse biological contexts ranging from regulation of autoimmunity to GSC differentiation.

6. REFERENCES

Albert TK, Lemaire M, van Berkum NL, Gentz R, Collart MA, Timmers HT (2000) Isolation and characterization of human orthologs of yeast CCR4-NOT complex subunits. Nucleic Acids Res 28: 809-17.

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: 355-64.

Amrani N, Ghosh S, Mangus DA, Jacobson A (2008) Translation factors promote the formation of two states of the closed-loop mRNP. Nature 453: 1276-80.

Arribas-Layton M, Wu D, Lykke-Andersen J, Song H (2013) Structural and functional control of the eukaryotic mRNA decapping machinery. Biochim Biophys Acta 1829: 580-9.

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.

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: 6642-51.

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: 207-18.

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: 228-44.

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, Roudko V, Basquin J, Sharma K, Urlaub H, Seraphin 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: 1281-8.

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

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: 1289-97.

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

Brooks SA, Blackshear PJ (2013) Tristetraprolin (TTP): interactions with mRNA and proteins, and current thoughts on mechanisms of action. Biochim Biophys Acta 1829: 666-79.

Brown JB, Boley N, Eisman R, May GE, Stoiber MH, Duff MO, Booth BW, Wen J, Park S, Suzuki AM, Wan KH, Yu C, Zhang D, Carlson JW, Cherbas L, Eads BD, Miller D, Mockaitis K, Roberts J, Davis CA et al. (2014) Diversity and dynamics of the *Drosophila* transcriptome. Nature 512: 393-9.

Cano F, Miranda-Saavedra D, Lehner PJ (2010) RNA-binding E3 ubiquitin ligases: novel players in nucleic acid regulation. Biochem Soc Trans 38: 1621-6.

Cano F, Rapiteanu R, Sebastiaan Winkler G, Lehner PJ (2015) A non-proteolytic role for ubiquitin in deadenylation of *MHC-I* mRNA by the RNA-binding E3-ligase MEX-3C. Nat Commun 6: 8670.

Carpenter S, Ricci EP, Mercier BC, Moore MJ, Fitzgerald KA (2014) Post-transcriptional regulation of gene expression in innate immunity. Nat Rev Immunol 14: 361-76.

Chang JH, Xiang S, Xiang K, Manley JL, Tong L (2011) Structural and biochemical studies of the 5' → 3' exoribonuclease Xrn1. Nat Struct Mol Biol 18: 270-6.

Chapat C, Corbo L (2014) Novel roles of the CCR4-NOT complex. Wiley Interdiscip Rev RNA 5: 883-901.

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 U S A 109: 9465-70.

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: 1218-26.

Chen CA, Zhang Y, Xiang Y, Han L, Shyu AB (2017) Antagonistic actions of two human Pan3 isoforms on global mRNA turnover. RNA 23: 1404-1418.

Chen D, McKearin D (2003) Dpp signaling silences bam transcription directly to establish asymmetric divisions of germline stem cells. Curr Biol 13: 1786-91.

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 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: 737-50.

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.

Clemens JC, Worby CA, Simonson-Leff N, Muda M, Maehama T, Hemmings BA, Dixon JE (2000) Use of double-stranded RNA interference in *Drosophila* cell lines to dissect signal transduction pathways. Proc Natl Acad Sci U S A 97: 6499-503.

Clement SL, Scheckel C, Stoecklin G, Lykke-Andersen J (2011) Phosphorylation of tristetraprolin by MK2 impairs AU-rich element mRNA decay by preventing deadenylase recruitment. Mol Cell Biol 31: 256-66.

Collart MA, Struhl K (1994) NOT1(CDC39), NOT2(CDC36), NOT3, and NOT4 encode a global-negative regulator of transcription that differentially affects TATA-element utilization. Genes Dev 8: 525-37.

Collart MA, Panasenko OO (2012) The Ccr4-not complex. Gene 492: 42-53.

Collart MA, Panasenko OO (2017) The Ccr4-Not Complex: Architecture and Structural Insights. Subcell Biochem 83: 349-379.

Coller J, Parker R (2004) Eukaryotic mRNA decapping. Annu Rev Biochem 73: 861-90.

Coller J, Wickens M (2007) Tethered function assays: an adaptable approach to study RNA regulatory proteins. Methods Enzymol 429: 299-321.

Coller JM, Tucker M, Sheth U, Valencia-Sanchez MA, Parker R (2001) The DEAD box helicase, Dhh1p, functions in mRNA decapping and interacts with both the decapping and deadenylase complexes. RNA 7: 1717-27.

Curtis D, Treiber DK, Tao F, Zamore PD, Williamson JR, Lehmann R (1997) A CCHC metal-binding domain in Nanos is essential for translational regulation. EMBO J 16: 834-43.

Davey NE, Cyert MS, Moses AM (2015) Short linear motifs - ex nihilo evolution of protein regulation. Cell Commun Signal 13: 43.

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

Dupressoir A, Morel AP, Barbot W, Loireau MP, Corbo L, Heidmann T (2001) Identification of four families of yCCR4- and Mg2+-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: 735-9.

Fu Z, Geng C, Wang H, Yang Z, Weng C, Li H, Deng L, Liu L, Liu N, Ni J, Xie T (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: 176-88.

Garneau NL, Wilusz J, Wilusz CJ (2007) The highways and byways of mRNA decay. Nat Rev Mol Cell Biol 8: 113-26.

Godwin AR, Kojima S, Green CB, Wilusz J (2013) Kiss your tail goodbye: the role of PARN, Nocturnin, and Angel deadenylases in mRNA biology. Biochim Biophys Acta 1829: 571-9.

Goldstrohm AC, Wickens M (2008) Multifunctional deadenylase complexes diversify mRNA control. Nat Rev Mol Cell Biol 9: 337-44.

Govind S (2008) Innate immunity in *Drosophila*: Pathogens and pathways. Insect Sci 15: 29-43.

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: 26796-805.

Haas M, Siegert M, Schurmann A, Sodeik B, Wolfes H (2004) c-Myb protein interacts with Rcd-1, a component of the CCR4 transcription mediator complex. Biochemistry 43: 8152-9.

Harigaya Y, Jones BN, Muhlrad D, Gross JD, Parker R (2010) Identification and analysis of the interaction between Edc3 and Dcp2 in Saccharomyces cerevisiae. Mol Cell Biol 30: 1446-56.

Harnisch C, Moritz B, Rammelt C, Temme C, Wahle E (2012) Activity and Function of Deadenylases. Enzymes 31: 181-211.

Hashimoto H, Hara K, Hishiki A, Kawaguchi S, Shichijo N, Nakamura K, Unzai S, Tamaru Y, Shimizu T, Sato M (2010) Crystal structure of zinc-finger domain of Nanos and its functional implications. EMBO Rep 11: 848-53.

Hildebrandt A, Alanis-Lobato G, Voigt A, Zarnack K, Andrade-Navarro MA, Beli P, Konig J (2017) Interaction profiling of RNA-binding ubiquitin ligases reveals a link between posttranscriptional regulation and the ubiquitin system. Sci Rep 7: 16582.

Houseley J, Tollervey D (2009) The many pathways of RNA degradation. Cell 136: 763-76.

Igaki T, Miura M (2014) The *Drosophila* TNF ortholog Eiger: emerging physiological roles and evolution of the TNF system. Semin Immunol 26: 267-74.

Imataka H, Gradi A, Sonenberg N (1998) A newly identified N-terminal amino acid sequence of human eIF4G binds poly(A)-binding protein and functions in poly(A)-dependent translation. EMBO J 17: 7480-9.

Inada T, Makino S (2014) Novel roles of the multi-functional CCR4-NOT complex in post-transcriptional regulation. Front Genet 5: 135.

Insco ML, Bailey AS, Kim J, Olivares GH, Wapinski OL, Tam CH, Fuller MT (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.

Ito K, Takahashi A, Morita M, Suzuki T, Yamamoto T (2011) The role of the CNOT1 subunit of the CCR4-NOT complex in mRNA deadenylation and cell viability. Protein Cell 2: 755-63.

Janowski R, Heinz GA, Schlundt A, Wommelsdorf N, Brenner S, Gruber AR, Blank M, Buch T, Buhmann R, Zavolan M, Niessing D, Heissmeyer V, Sattler M (2016) Roquin recognizes a non-canonical hexaloop structure in the 3'-UTR of Ox40. Nat Commun 7: 11032.

Joly W, Chartier A, Rojas-Rios P, Busseau I, Simonelig M (2013) The CCR4 deadenylase acts with Nanos and Pumilio in the fine-tuning of Mei-P26 expression to promote germline stem cell self-renewal. Stem Cell Reports 1: 411-24.

Jonas S, Izaurralde E (2015) Towards a molecular understanding of microRNA-mediated gene silencing. Nat Rev Genet 16: 421-33.

Jones CI, Zabolotskaya MV, Newbury SF (2012) The 5' → 3' exoribonuclease XRN1/Pacman and its functions in cellular processes and development. Wiley Interdiscip Rev RNA 3: 455-68.

Kadyrova LY, Habara Y, Lee TH, Wharton RP (2007) Translational control of maternal Cyclin B mRNA by Nanos in the *Drosophila* germline. Development 134: 1519-27.

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.

Komander D, Rape M (2012) The ubiquitin code. Annu Rev Biochem 81: 203-29.

Kruk JA, Dutta A, Fu J, Gilmour DS, Reese JC (2011) The multifunctional Ccr4-Not complex directly promotes transcription elongation. Genes Dev 25: 581-93.

Kuzuoglu-Ozturk D, Bhandari D, Huntzinger E, Fauser M, Helms S, Izaurralde E (2016) miRISC and the CCR4-NOT complex silence mRNA targets independently of 43S ribosomal scanning. EMBO J 35: 1186-203.

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: 443-53.

Lavoie CA, Ohlstein B, McKearin DM (1999) Localization and function of Bam protein require the benign gonial cell neoplasm gene product. Dev Biol 212: 405-13.

Lee EK (2012) Post-translational modifications of RNA-binding proteins and their roles in RNA granules. Curr Protein Pept Sci 13: 331-6.

Lee M, Kim B, Kim VN (2014) Emerging roles of RNA modification: m(6)A and U-tail. Cell 158: 980-987.

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: 869-81.

Li W, Gao B, Lee SM, Bennett K, Fang D (2007) RLE-1, an E3 ubiquitin ligase, regulates *C. elegans* aging by catalyzing DAF-16 polyubiquitination. Dev Cell 12: 235-46.

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: 9304-9.

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.

Lin H (2002) The stem-cell niche theory: lessons from flies. Nat Rev Genet 3: 931-40.

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: 2125-38.

Malik S, Jang W, Kim C (2017) Protein Interaction Mapping of Translational Regulators Affecting Expression of the Critical Stem Cell Factor Nos. Dev Reprod 21: 449-456.

Marcotrigiano J, Gingras AC, Sonenberg N, Burley SK (1997) Cocrystal structure of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. Cell 89: 951-61.

Marnef A, Standart N (2010) Pat1 proteins: a life in translation, translation repression and mRNA decay. Biochem Soc Trans 38: 1602-7.

Maruyama T, Araki T, Kawarazaki Y, Naguro I, Heynen S, Aza-Blanc P, Ronai Z, Matsuzawa A, Ichijo H (2014) Roquin-2 promotes ubiquitin-mediated degradation of ASK1 to regulate stress responses. Sci Signal 7: ra8.

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: 751-65.

Mauxion F, Preve B, Seraphin B (2013) C2ORF29/CNOT11 and CNOT10 form a new module of the CCR4-NOT complex. RNA Biol 10: 267-76.

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: 2937-47.

Meyer S, Temme C, Wahle E (2004) Messenger RNA turnover in eukaryotes: pathways and enzymes. Crit Rev Biochem Mol Biol 39: 197-216.

Mitchell SF, Parker R (2014) Principles and properties of eukaryotic mRNPs. Mol Cell 54: 547-58.

Moore MJ (2005) From birth to death: the complex lives of eukaryotic mRNAs. Science 309: 1514-8.

Morris JZ, Hong A, Lilly MA, Lehmann R (2005) twin, a CCR4 homolog, regulates cyclin poly(A) tail length to permit *Drosophila* oogenesis. Development 132: 1165-74.

Neumuller 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: 241-5.

Oberholzer U, Collart MA (1998) Characterization of NOT5 that encodes a new component of the Not protein complex. Gene 207: 61-9.

Ohlstein B, McKearin D (1997) Ectopic expression of the *Drosophila* Bam protein eliminates oogenic germline stem cells. Development 124: 3651-62.

Ohlstein B, Lavoie CA, Vef O, Gateff E, McKearin DM (2000) The *Drosophila* cystoblast differentiation factor, benign gonial cell neoplasm, is related to DExH-box proteins and interacts genetically with bag-of-marbles. Genetics 155: 1809-19.

Pan L, Wang S, Lu T, Weng C, Song X, Park JK, Sun J, Yang ZH, Yu J, Tang H, McKearin DM, Chamovitz DA, Ni J, Xie T (2014) Protein competition switches the function of COP9 from self-renewal to differentiation. Nature 514: 233-6.

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: 31389-98.

Panasenko OO, Collart MA (2011) Not4 E3 ligase contributes to proteasome assembly and functional integrity in part through Ecm29. Mol Cell Biol 31: 1610-23.

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

Parker R, Song H (2004) The enzymes and control of eukaryotic mRNA turnover. Nat Struct Mol Biol 11: 121-7.

Parker R, Sheth U (2007) P bodies and the control of mRNA translation and degradation. Mol Cell 25: 635-46.

Parton RM, Davidson A, Davis I, Weil TT (2014) Subcellular mRNA localisation at a glance. J Cell Sci 127: 2127-33.

Pereira B, Le Borgne M, Chartier NT, Billaud M, Almeida R (2013) MEX-3 proteins: recent insights on novel post-transcriptional regulators. Trends Biochem Sci 38: 477-9.

Perez-Garijo A, Fuchs Y, Steller H (2013) Apoptotic cells can induce non-autonomous apoptosis through the TNF pathway. Elife 2: e01004.

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: 11058-72.

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: 974-90.

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: 1640-7.

Rissland OS (2017) The organization and regulation of mRNA-protein complexes. Wiley Interdiscip Rev RNA 8

Sakurai S, Ohto U, Shimizu T (2015) Structure of human Roquin-2 and its complex with constitutive-decay element RNA. Acta Crystallogr F Struct Biol Commun 71: 1048-54.

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: 4373-86.

Sanduja S, Blanco FF, Young LE, Kaza V, Dixon DA (2012) The role of tristetraprolin in cancer and inflammation. Front Biosci (Landmark Ed) 17: 174-88.

Schlundt A, Heinz GA, Janowski R, Geerlof A, Stehle R, Heissmeyer V, Niessing D, Sattler M (2014) Structural basis for RNA recognition in roquin-mediated post-transcriptional gene regulation. Nat Struct Mol Biol 21: 671-8.

Schmid M, Jensen TH (2008) The exosome: a multipurpose RNA-decay machine. Trends Biochem Sci 33: 501-10.

Semotok JL, Cooperstock RL, Pinder BD, Vari HK, Lipshitz HD, Smibert CA (2005) Smaug recruits the CCR4/POP2/NOT deadenylase complex to trigger maternal transcript localization in the early *Drosophila* embryo. Curr Biol 15: 284-94.

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. Nature Communications 8: 14307.

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.

Shan L, Wu C, Chen D, Hou L, Li X, Wang L, Chu X, Hou Y, Wang Z (2017) Regulators of alternative polyadenylation operate at the transition from mitosis to meiosis. J Genet Genomics 44: 95-106.

Shatkin AJ, Manley JL (2000) The ends of the affair: capping and polyadenylation. Nat Struct Biol 7: 838-42.

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 U S A 106: 11623-8.

Siess DC, Vedder CT, Merkens LS, Tanaka T, Freed AC, McCoy SL, Heinrich MC, Deffebach ME, Bennett RM, Hefeneider SH (2000) A human gene coding for a membrane-associated nucleic acid-binding protein. J Biol Chem 275: 33655-62.

Sonenberg N, Dever TE (2003) Eukaryotic translation initiation factors and regulators. Curr Opin Struct Biol 13: 56-63.

Spradling A, Fuller MT, Braun RE, Yoshida S (2011) Germline stem cells. Cold Spring Harb Perspect Biol 3: a002642.

Tan D, Zhou M, Kiledjian M, Tong L (2014) The ROQ domain of Roquin recognizes mRNA constitutive-decay element and double-stranded RNA. Nat Struct Mol Biol 21: 679-85.

Tarun SZ, Jr., Sachs AB (1996) Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G. EMBO J 15: 7168-77.

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: 1356-70.

Tompa P (2012) Intrinsically disordered proteins: a 10-year recap. Trends Biochem Sci 37: 509-16.

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: 6695-708.

Tucker M, Valencia-Sanchez MA, Staples RR, Chen J, Denis CL, Parker R (2001) The transcription factor associated Ccr4 and Caf1 proteins are components of the major cytoplasmic mRNA deadenylase in Saccharomyces cerevisiae. Cell 104: 377-86.

Tucker M, Staples RR, Valencia-Sanchez MA, Muhlrad D, Parker R (2002) Ccr4p is the catalytic subunit of a Ccr4p/Pop2p/Notp mRNA deadenylase complex in *Saccharomyces cerevisiae*. EMBO J 21: 1427-36.

Van Etten J, Schagat TL, Hrit J, Weidmann CA, Brumbaugh J, Coon JJ, Goldstrohm AC (2012) Human Pumilio proteins recruit multiple deadenylases to efficiently repress messenger RNAs. J Biol Chem 287: 36370-83.

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: 6733-78.

Varshavsky A (2017) The Ubiquitin System, Autophagy, and Regulated Protein Degradation. Annu Rev Biochem 86: 123-128.

Vinuesa CG, Cook MC, Angelucci C, Athanasopoulos V, Rui L, Hill KM, Yu D, Domaschenz H, Whittle B, Lambe T, Roberts IS, Copley RR, Bell JI, Cornall RJ, Goodnow CC (2005) A RING-

type ubiquitin ligase family member required to repress follicular helper T cells and autoimmunity. Nature 435: 452-8.

Virtanen A, Henriksson N, Nilsson P, Nissbeck M (2013) Poly(A)-specific ribonuclease (PARN): an allosterically regulated, processive and mRNA cap-interacting deadenylase. Crit Rev Biochem Mol Biol 48: 192-209.

Vogel C, Marcotte EM (2012) Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. Nat Rev Genet 13: 227-32.

Vogel KU, Edelmann SL, Jeltsch KM, Bertossi A, Heger K, Heinz GA, Zoller J, Warth SC, Hoefig KP, Lohs C, Neff F, Kremmer E, Schick J, Repsilber D, Geerlof A, Blum H, Wurst W, Heikenwalder M, Schmidt-Supprian M, Heissmeyer V (2013) Roquin paralogs 1 and 2 redundantly repress the Icos and Ox40 costimulator mRNAs and control follicular helper T cell differentiation. Immunity 38: 655-68.

Wahle E, Winkler GS (2013) RNA decay machines: deadenylation by the Ccr4-not and Pan2-Pan3 complexes. Biochim Biophys Acta 1829: 561-70.

Wells SE, Hillner PE, Vale RD, Sachs AB (1998) Circularization of mRNA by eukaryotic translation initiation factors. Mol Cell 2: 135-40.

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: 157-65.

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: 3089-99.

Wolf J, Passmore LA (2014) mRNA deadenylation by Pan2-Pan3. Biochem Soc Trans 42: 184-7.

Wreden C, Verrotti AC, Schisa JA, Lieberfarb ME, Strickland S (1997) Nanos and pumilio establish embryonic polarity in *Drosophila* by promoting posterior deadenylation of hunchback mRNA. Development 124: 3015-23.

Xie T, Spradling AC (1998) decapentaplegic is essential for the maintenance and division of germline stem cells in the *Drosophila* ovary. Cell 94: 251-60.

Xu K, Bai Y, Zhang A, Zhang Q, Bartlam MG (2014) Insights into the structure and architecture of the CCR4-NOT complex. Front Genet 5: 137.

Yamaji M, Jishage M, Meyer C, Suryawanshi H, Der E, Yamaji M, Garzia A, Morozov P, Manickavel S, McFarland HL, Roeder RG, Hafner M, Tuschl T (2017) DND1 maintains germline stem cells via recruitment of the CCR4-NOT complex to target mRNAs. Nature 543: 568-572.

Yamashita A, Chang TC, Yamashita Y, Zhu W, Zhong Z, Chen CY, Shyu AB (2005) Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. Nat Struct Mol Biol 12: 1054-63.

Yang L, Wang C, Li F, Zhang J, Nayab A, Wu J, Shi Y, Gong Q (2017) The human RNA-binding protein and E3 ligase MEX-3C binds the MEX-3-recognition element (MRE) motif with high affinity. J Biol Chem 292: 16221-16234.

Yoda M, Cifuentes D, Izumi N, Sakaguchi Y, Suzuki T, Giraldez AJ, Tomari Y (2013) Poly(A)-specific ribonuclease mediates 3'-end trimming of Argonaute2-cleaved precursor microRNAs. Cell Rep 5: 715-26.

Yu D, Tan AH, Hu X, Athanasopoulos V, Simpson N, Silva DG, Hutloff A, Giles KM, Leedman PJ, Lam KP, Goodnow CC, Vinuesa CG (2007) Roquin represses autoimmunity by limiting inducible T-cell co-stimulator messenger RNA. Nature 450: 299-303.

Zaessinger S, Busseau I, Simonelig M (2006) Oskar allows nanos mRNA translation in *Drosophila* embryos by preventing its deadenylation by Smaug/CCR4. Development 133: 4573-83.

Zamore PD (2001) RNA interference: listening to the sound of silence. Nat Struct Biol 8: 746-50.

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: 1052-65.

Zwartjes CG, Jayne S, van den Berg DL, Timmers HT (2004) Repression of promoter activity by CNOT2, a subunit of the transcription regulatory Ccr4-not complex. J Biol Chem 279: 10848-54.

7. ABBREVIATIONS

ARE	Au-rich element
ARM	Armadillo repeat
ASK1	Apoptosis signal-regulating kinase 1
Bam	Bag-of-marbles
Bgcn	Benign gonial cell neoplasm
BMP	Bone morphogenetic protein
CAF	CCR4-associated factor
CBM	CAF40-binding motif
CC	Coiled-coil
CCR4	Carbone catabolite repressor 4
CDE	Constitutive decay element
Ce	*Caenorhabditis elegans*
CN9BD	CAF40/CNOT9 binding domain
co-IP	Co-immunoprecipitation
DAF-16	Abnormal dauer formation 16
DCP1/DCP2	Decapping enzyme subunit 1/2
DDX6	DEAD box protein 6
DEDD	Asp-Glu-Asp-Asp
DExH	Asp-Glu-X-His
Dpp	Decapentaplegic
Dm	*Drosophila melanogaster*
DNA	Deoxyribonucleic acid
DND1	Dead end homolog 1
ds	Double stranded
DUF	Domain of unknown function
EDC3	Enhancer of decapping 3
E.coli	*Escherichia coli*
EEP	Endonuclease-exonuclease-phosphatase domain
eIF	Eukaryotic translation initiation factor
F-Luc	Firefly luciferase
Foxo	Forkhead box protein O
Gbb	Glass bottom boat

GFP	Green fluorescent protein
GSC	Germline stem cell
GW182	Glycine-tryptophan repeat containing protein of 182 kDa size, Gawky
HEK293T	Human embryonic kidney 293T cell line
HEPN	Higher eukaryotes and prokaryotes nucleotide-binding
HLA-A2	Human leukocyte antigen 2
Hs	*Homo sapiens*
Icos	Inducible costimulatory
IFN- γ	Interferon γ
IL-1	Interleukin-1
ITC	Isothermal titration calorimetry
KD	Knockdown
KH	K homology
LRR	Leucine rich repeat
MAPK	Mitogen-activated protein kinases
MBP	Maltose binding protein
MIF4G	Middle domain of eukaryotic initiation factor 4G
miRNA	MicroRNA
Mm	*Mus musculus*
MNAB	Membrane-associated nucleic-acid binding protein
mRNA	Messenger RNA
mRNP	Messenger ribonucleoprotein
NAC	Nascent polypeptide associated complex
NBM	NOT module-binding motif
NMD	Nonsense-mediated decay
NOT	Negative on TATA less
ORF	Open reading frame
Ox40	Ox40 receptor
PABP	Poly(A) binding protein
PAN2-PAN3	Poly(A) nuclease 2-3
PAR-CLIP	Photo-activable ribonucleoside-enhanced crosslinking and immunoprecipitation
PARN	Poly(A) ribonuclease
PATL1	Protein PAT1 homolog 1
Poly(A)	Poly adenosine

POP2	PGK promoter directed over production 2
PRE	Pumilio response element
pre-mRNA	Precursor messenger RNA
PTM	Posttranslational modification
RBP	RNA-binding protein
Rcd-1	Required for cell differentiation-1
RING	Really interesting new gene
RLE-1	Regulation of longevity by E3
R-Luc	*Renilla* luciferase
RNA	Ribonucleic acid
RNAi	RNA interference
Rps7A	40S ribosomal protein s7-A
RQCD1	Required for cell differentiation 1
RRM	RNA recognition motif
rRNA	Ribosomal RNA
S2	Schneider 2 cell line
Sax	Saxophone
Sc	*Saccharomyces cerevisiae*
SHD	NOT1 superfamily homology domain
SLE	systemic lupus erythematosus
SLiMs	Short linear motifs
snRNA	Small nuclear RNA
SRE	Smaug recognition element
Sxl	Sex-lethal
TBM	Tut binding motif
TNF-α	Tumor necrosis factor alpha
TNRC6	Trinucleotide repeat containing gene protein 6
Tkv	Thickveins
TRAMP	Trf4/Air2/Mtr4p polyadenylation
tRNA	Transfer RNA
TTP	Tristetraproline
TPR	Tetratricopeptide repeat
Tut	Tumorous testis
URR	U-rich repeat

UTR	Untranslated region
wt	wild type
XRN1	5’-3’ exoribonuclease 1
YAP1	Yes-associated protein 1
ZnF	Zinc-finger

8. APPENDIX

8.1 List of publications

In this section, I describe my contribution as author for the publications discussed in this thesis.

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

My contribution: I designed and cloned all constructs for expression in *Dm* S2 cells and *E.coli*. Together with Praveen Bawankar, I designed and cloned all constructs for expression in human cells. I performed all of the tethering assays and knockdown experiments both in human and *Dm* S2 cells and co-IP assays in *Dm* S2 cells. I identified both CBM and NBM motifs involved in the direct interaction between Roquin and the CCR4-NOT complex via *in vitro* pulldown assays. Together with Tobias Raisch, I crystalized the Roquin CBM_R bound to CAF40. I contributed to writing the manuscript and prepared all of the figures.

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
*Equal contributions

My contribution: I designed and cloned all constructs for expression in human and *Dm* S2 cells as well as *E.coli*. Together with Charlotte Backhaus, I performed tethering assays and co-IP assays in *Dm* S2 cells. I performed all overexpression experiments and knockdown experiments in *Dm* S2 cells. I performed all of the tethering assays and pulldown experiments in human cells. Together with Charlotte Backhaus and Tobias Raisch, I identified the Bam CBM involved in the direct interaction with CAF40 via *in vitro* pulldown assays. I contributed to writing the manuscript and prepared all of the figures.

8.2 Manuscript in preparation

This section describes my contribution as author for the manuscript discussed in this thesis.

Keskeny C, Raisch T, **Sgromo A**, , Bhandari D, Igreja C, Weichenrieder O, Izaurralde E (in preparation) Metazoan NOT4 associates with the CCR4-NOT complex via interactions with CAF40 and NOT1.

My contribution: I designed and cloned all constructs for expression in *Dm* S2 cells. I performed all of the tethering assays and western blots in *Dm* S2 cells. Together with Tobias Raisch, I identified the NOT4 CBM bound to CAF40. I contributed to writing the manuscript.

8.3 Supplemental information

Table 3. Constructs used in section 4.3.

Name	**Bag-of-marbles** (Uniprot P22745)	**Comment**
Bam	GFP-Bam 1–442	
Bam-N	GFP-Bam 1–140	Contains the CBM and the TBM
Bam-C	GFP-Bam 141–442	Interacts with Bgcn
Bam-4xMut	GFP-Bam L17E, M24E, L28E, V32E	Disrupts CAF40 binding

Name	**Benign gonial cell neoplasm** (Uniprot Q9W1I2)	**Comment**
Bgcn	HA- Bgcn 1-1215	
	GFP- Bgcn 1-1215	
Bgcn-N	HA- Bgcn 1-950	
Bgcn-C	HA- Bgcn 951-1215	

Name	**Tumorous testis** (Uniprot Q8IQA2)	**Comment**
Tut	HA-Tut 1-230	

Table 4. Antibodies used in section 4.3.

Antibody	Source	Catalog number	Dilution	Monoclonal/ Polyclonal
Anti-HA-HRP	Roche	12 013 819 001	1:5,000	Monoclonal
Anti-GFP (for immunoprecipitation)	In house			Rabbit Polyclonal
Anti-V5	AbD Serotec	MCA1360GA	1:5,000	Monoclonal
Anti-mouse-HRP	GE Healthcare	NA931V	1:10,000	Monoclonal
Anti-GFP (for western blotting)	Roche	11 814 460 001	1:2,000	Mouse Monoclonal

Material and Methods of section 4.3

DNA constructs

The DNA constructs used are listed in Table 3 and were confirmed by sequencing. For expression of Bgcn (full-length and fragments) in *Dm* S2 cells, the corresponding cDNA was amplified from total *Dm* oocyte cDNA and cloned between the HindIII and XhoI restriction sites of the pAc5.1-λN-HA and pAc5.1-GFP vectors (Rehwinkel et al. 2005, Tritschler et al. 2008). For expression of Tut in *Dm* S2 cells, the corresponding cDNA was amplified from total *Dm* oocyte cDNA and cloned between the KpnI and XhoI restriction sites of the pAc5.1-λN-HA vector. The plasmids used for the expression of *Dm* Bam constructs and for the expression of the subunits of the CCR4-NOT complex in S2 cells have been previously described (Sgromo et al. 2017, Sgromo et al. 2018).

Co-immunoprecipitation assays

All co-immunoprecipitation assays in *Dm* S2 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 are listed in Table 4.

2.5×10^6 of S2 cells were seeded per well in six-well plates and transfected using Effectene transfection reagent (Qiagen). The transfection mixtures contained plasmids expressing GFP-tagged Bgcn (2 µg) and HA-tagged subunits of the CCR4-NOT complex (1.5 µg for NOT1, 1 µg for NOT2, NOT11 and CCR4). GFP-F-Luc (0.015 µg) served as a negative control. Cells were harvested 3 days after transfection, and co-immunoprecipitation assays were performed using NET buffer (50 mM Tris [pH 7.4], 150 mM NaCl, 1 mM EDTA, 0.1% Triton X-100 supplemented with protease inhibitors [Complete protease inhibitor mix, Roche]) as previously described (Braun et al. 2011).

Tethering assays

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 transfection reagent (Qiagen). For the experiment shown in Figure 16 the transfection mixtures contained the following plasmids: 0.1 µg of firefly luciferase reporter F-Luc-5BoxB, 0.4 µg of the R-Luc transfection control and various amounts of plasmids expressing λN-HA-tagged full-length Bgcn or Bgcn fragments (0.005 µg for full-length, 0.025 µg for Bgcn-N, 0.01 µg for Bgcn-C). In the experiment shown in Figure 18, cells were transfected with 0.01 µg of plasmid expressing λN-HA-tagged Tut. Cells were also transfected with plasmids expressing GFP (0.05 µg) or GFP-tagged Bam either wild type (0.1 µg) or fragments or the mutant (0.05 µg for Bam-N, 0.2 µg for Bam-C and 0.15 µg for Bam-4xMut). For the experiment shown in Figure 19 the transfection mixtures contained the following plasmids: 0.1 µg of firefly luciferase reporter F-Luc-mei-P26, 0.4 µg of the R-Luc transfection control and 0.01 µg of plasmids expressing λN-HA-tagged Tut. Cells were also co-transfected with 0.2 µg of GFP-tagged full-length Bgcn.

To generate F-Luc-mei-P26 the 3'UTR of *mei-P26* was amplified from wt ovary cDNA using the oligos AGCCgctagcTTGCAAATCAAAGCGCGCAAC and CGAActcgagAGTAGTAGCGCTAATTG and cloned into the 3'UTR of the F-Luc reporter using the NheI and XhoI restriction sites.

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

9. ORIGINAL MANUSCRIPTS DISCUSSED IN THIS THESIS

In this section, the originals of my publications and manuscripts are attached.

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

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
*Equal contributions

Keskeny C, Raisch T, **Sgromo A**, Bhandari D, Igreja C, Weichenrieder O, Izaurralde E (in preparation) Metazoan NOT4 associates with the CCR4-NOT complex via interactions with CAF40 and NOT1.

nature COMMUNICATIONS

ARTICLE

Received 26 Jul 2016 | Accepted 15 Dec 2016 | Published 6 Feb 2017

DOI: 10.1038/ncomms14307 OPEN

A CAF40-binding motif facilitates recruitment of the CCR4-NOT complex to mRNAs targeted by *Drosophila* Roquin

Annamaria Sgromo[1], Tobias Raisch[1], Praveen Bawankar[1], Dipankar Bhandari[1], Ying Chen[1], Duygu Kuzuoğlu-Öztürk[1], Oliver Weichenrieder[1] & Elisa Izaurralde[1]

Human (*Hs*) Roquin1 and Roquin2 are RNA-binding proteins that promote mRNA target degradation through the recruitment of the CCR4-NOT deadenylase complex and are implicated in the prevention of autoimmunity. Roquin1 recruits CCR4-NOT via a C-terminal region that is not conserved in Roquin2 or in invertebrate Roquin. Here we show that Roquin2 and *Drosophila melanogaster* (*Dm*) Roquin also interact with the CCR4-NOT complex through their C-terminal regions. The C-terminal region of *Dm* Roquin contains multiple motifs that mediate CCR4-NOT binding. One motif binds to the CAF40 subunit of the CCR4-NOT complex. The crystal structure of the *Dm* Roquin CAF40-binding motif (CBM) bound to CAF40 reveals that the CBM adopts an α-helical conformation upon binding to a conserved surface of CAF40. Thus, despite the lack of sequence conservation, the C-terminal regions of Roquin proteins act as an effector domain that represses the expression of mRNA targets via recruitment of the CCR4-NOT complex.

[1] Department of Biochemistry, Max Planck Institute for Developmental Biology, Spemannstrasse 35, Tübingen 72076, Germany. Correspondence and requests for materials should be addressed to E.I. (email: elisa.izaurralde@tuebingen.mpg.de).

The CCR4-NOT deadenylase complex plays a central role in bulk mRNA degradation by catalysing the removal of mRNA poly(A) tails, which is the first step in general mRNA decay[1]. In addition to its role in global mRNA degradation, the CCR4-NOT complex regulates the expression of a large number of specific mRNAs, to which it is recruited via interactions with RNA-associated proteins. Consequently, CCR4-NOT functions as a major downstream effector complex in posttranscriptional mRNA regulation in eukaryotes.

The CCR4-NOT complex consists of several structurally and functionally distinct modules, which assemble around the NOT1 scaffold subunit[1]. NOT1 contains several α-helical domains that provide binding surfaces for the individual modules. A central domain of NOT1 that is structurally related to the middle portion of eIF4G (termed the NOT1 MIF4G domain) provides a binding site for the catalytic module, which comprises two deadenylases, namely CAF1 or its paralogue POP2 (also known as CNOT7 and CNOT8, respectively, in humans), and CCR4a or its paralogue CCR4b (also known as CNOT6 and CNOT6L, respectively, in humans). 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[2,3]. C-terminal to the MIF4G domain, NOT1, contains a CAF40/NOT9-binding domain, CN9BD, that binds to the highly conserved CAF40 subunit, which is also known as CNOT9 (refs 2,3), followed by a NOT1 superfamily homology domain SHD, which interacts with NOT2-NOT3 heterodimers to form the NOT module[4,5].

The CAF40 and NOT modules have no catalytic activity and have been implicated in mediating interactions with RNA-associated proteins that recruit the CCR4-NOT complex to their targets[2–10]. These proteins include the GW182 family, which is involved in miRNA-mediated gene silencing in animals[2,3], tristetraprolin (TTP), a protein required for the degradation of mRNAs containing AU-rich elements[6], the germline determinant Nanos[7–9] and the human Roquin1 and Roquin2 proteins[10].

The vertebrate Roquin proteins are negative regulators of T follicular helper cell differentiation and autoimmunity in vertebrates[11,12]. There are two partially redundant paralogues, Roquin1 and Roquin2 (initially named membrane-associated nucleic acid-binding protein), in vertebrates and only one family member in invertebrate species[11,12]. The proteins feature an N-terminal folded region followed by a C-terminal extension of variable length and low sequence complexity that is predicted to be predominantly unstructured[13,14] (Fig. 1a). The N-terminal region contains a RING finger E3 ubiquitin ligase domain, a ROQ RNA-binding domain flanked by a bilobed HEPN domain and a CCCH-type zinc finger domain, all of which are highly conserved in metazoans and define the protein family[14–20]. The RING domains of *Hs* Roquin2 and the *Caenorhabditis elegans* homologue of Roquin1, RLE-1, (regulation of longevity by E3) exhibit E3 ubiquitin ligase activity[21,22].

The ROQ domain of *Hs* Roquin1 and Roquin2 recognizes specific stem-loop structures in the 3′-untranslated region (UTR) of target mRNAs. These targets include mRNAs encoding regulators of inflammation such as the inducible T-cell costimulator, the costimulatory receptor Ox40, neuropilin-1, interleukin-6, interferon γ (IFN-γ) and the tumor necrosis factor-α (TNF-α)[10,14,16,23–31]. *Hs* Roquin1 and Roquin2 downregulate these mRNA targets through interactions with the CCR4-NOT deadenylase complex and decapping factors[10,32].

For *Hs* Roquin1, it has been shown that the interaction with the CCR4-NOT complex is mediated by the C-terminal region of the protein that is conserved only among vertebrate Roquin1 orthologues[10]. However, it is not known how Roquin2 recruits the CCR4-NOT complex, because its C-terminal region shows no similarity with that of Roquin1. In addition, the C-terminal regions of the invertebrate Roquin proteins are highly divergent[18,20–22], and it is unclear whether Roquin proteins recruit the CCR4-NOT complex in invertebrates. However, the conservation of the ROQ domain indicates that Roquin proteins also bind RNA in invertebrates, although their specific RNA targets are currently unknown.

Despite extensive information regarding the mode of RNA recognition by Roquin proteins[15–19], a detailed molecular understanding of how the proteins interact with the CCR4-NOT complex is lacking, and it is not even known whether the interactions are direct. Here we investigate the molecular details of how Roquin proteins recruit the CCR4-NOT complex. First, we show that *Hs* Roquin2 and *Dm* Roquin interact with the CCR4-NOT complex and promote target mRNA degradation via their C-terminal regions, suggesting conserved functional principles among all Roquin proteins. Furthermore, we find that the *Dm* Roquin C-terminal region contains multiple binding sites for the CCR4-NOT complex and that these sites act redundantly to promote mRNA degradation. Among these sites, we identify a short linear motif (SLiM) that is necessary and sufficient to mediate direct binding to the CAF40 subunit of the CCR4-NOT complex. This motif is termed the CAF40-binding motif (CBM), and we determine its crystal structure bound to CAF40. Structure-based mutations of the CAF40-CBM interface prevent binding of *Dm* Roquin to CAF40 and reduce the ability of the protein to degrade mRNA targets, indicating that CAF40 is an important mediator of the recruitment of the CCR4-NOT complex. Together with previous studies[10], our results reveal a common role of the Roquin C-terminal region as an effector domain that regulates mRNA target expression through the recruitment of the CCR4-NOT complex despite the lack of sequence conservation.

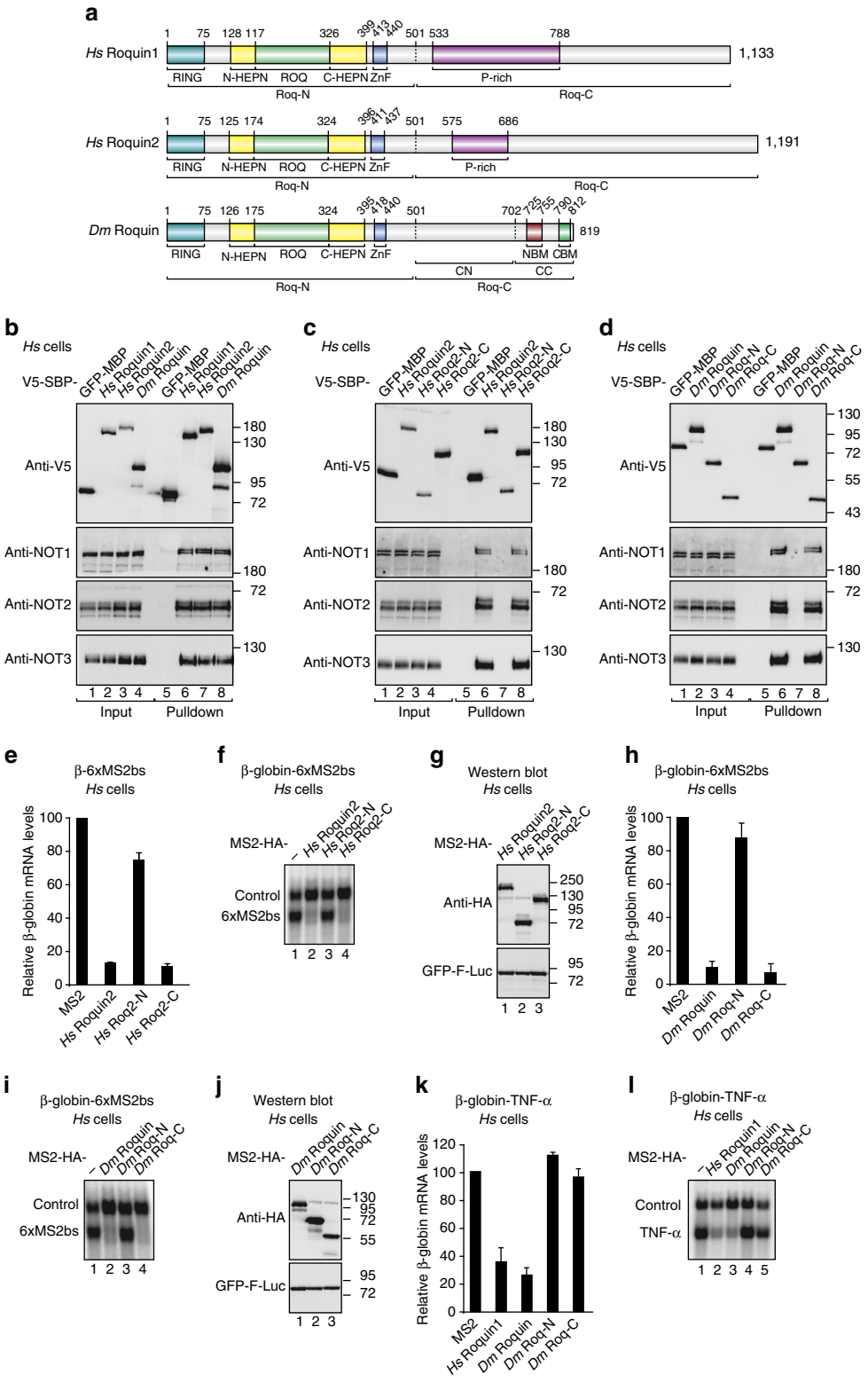
Results

Roquin C-terminal regions recruit the CCR4-NOT complex. *Hs* Roquin1 interacts with the CCR4-NOT deadenylase complex through a C-terminal region that shows very low sequence similarity to the C-terminal regions of the corresponding vertebrate Roquin2 paralogues and invertebrate Roquin[10] (Fig. 1a). We therefore asked whether the C-terminal regions of *Hs* Roquin2 and *Dm* Roquin have the ability to interact with the CCR4-NOT complex. *Hs* Roquin2 and *Dm* Roquin expressed with a tag consisting of the V5 epitope followed by the streptavidin-binding peptide (V5-SBP) pulled down the endogenous CCR4-NOT complex in human HEK293T cells to a similar extent as *Hs* Roquin1 (Fig. 1b). Moreover, the C-terminal regions of *Hs* Roquin2 and *Dm* Roquin were necessary and sufficient for the interaction, as observed for *Hs* Roquin1 (Fig. 1c,d; Supplementary Fig. 1a)[10]. The observation that *Dm* Roquin interacts with the CCR4-NOT complex in human cells further suggests that the protein recognizes surfaces on the CCR4-NOT complex that are conserved across species.

Roquin C-terminal regions mediate mRNA degradation. We next investigated whether the C-terminal regions of *Hs* Roquin2 and *Dm* Roquin elicit the degradation of mRNA targets. To this end, we used an MS2-based tethering assay in human HEK293T cells. The full-length proteins and the corresponding N- and C-terminal fragments (*Hs* Roq2-N, *Hs* Roq2-C, *Dm* Roq-N and *Dm* Roq-C, respectively; Fig. 1a, Supplementary Table 1) were expressed with an MS2-HA tag that mediates binding to a β-globin reporter mRNA containing six MS2-binding sites in its 3′-UTR (β-globin-6xMS2bs)[33]. Tethered *Hs* Roquin2 and *Dm*

Roquin reduced the level of the β-globin-6xMS2bs mRNA relative to the MS2-HA fusion protein, which was used as a negative control (Fig. 1e–j). Furthermore, the C-terminal fragments retained the mRNA degradation activity of the full-length proteins, whereas the N-terminal fragments were inactive (Fig. 1e–j). Similar results were obtained for *Hs* Roquin1 (Supplementary Fig. 1b–d). The N- and C-terminal fragments were expressed at levels comparable to those of the full-length proteins (Fig. 1g,j), and none of the proteins affected the expression of the control β-globin mRNA lacking MS2-binding sites (Fig. 1f,i; control). Thus we conclude that, despite the lack of sequence conservation, the C-terminal regions of Roquin proteins interact with the CCR4-NOT complex and promote the degradation of bound mRNAs.

As shown for *Hs* Roquin1 and Roquin2, *Dm* Roquin regulates the expression of a β-globin mRNA reporter containing the



constitutive decay element (CDE) from the TNF-α mRNA in its 3′-UTR (CDE-37; ref. 10) in human HEK293T cells (Fig. 1k,l), consistent with the conservation of the ROQ domain[15,16,23]. Regulation of the β-globin-TNF-α reporter by *Dm* Roquin was abolished by mutations in the CDE that disrupt the binding of the ROQ domain (Supplementary Fig. 1e,f; mutant MUT16; ref. 10). Furthermore, in contrast to the results obtained in the tethering assay, degradation of the β-globin-TNF-α reporter required both the N- and the C-terminal regions of *Dm* Roquin (Fig. 1k,l). Indeed, the C-terminal region alone was not sufficient to cause degradation of the β-globin-TNF-α reporter (Fig. 1k,l), most likely because it does not bind RNA.

Roquin proteins direct mRNAs to the 5′-to-3′ decay pathway. Given that Roquin proteins interact with the CCR4-NOT complex, we next investigated whether the proteins elicit degradation of mRNA targets via the 5′-to-3′ decay pathway, in which deadenylation is followed by decapping and 5′-to-3′ exonucleolytic degradation of the mRNA body. To this end, we performed tethering assays in HEK293T cells overexpressing a catalytically inactive DCP2 mutant (DCP2 E148Q), which inhibits decapping in a dominant-negative manner[34]. We observed that degradation of the β-globin-6xMS2bs mRNA by tethered *Hs* Roquin1, *Hs* Roquin2 and *Dm* Roquin was impaired in cells expressing the DCP2 mutant; in these cells, the reporter accumulated in a shorter deadenylated form (Fig. 2a,b, lanes 6–8). The expression of the tethered proteins was not affected by the expression of DCP2 mutant (Fig. 2c). Our results indicate that the three Roquin proteins direct mRNA targets to the 5′-to-3′ decay pathway.

***Dm* Roquin degrades bound mRNAs in *Drosophila* cells.** To investigate whether the *Dm* Roquin protein has the ability to repress and degrade bound mRNAs in *Dm* cells, we used a λN-based tethering assay in *Drosophila melanogaster* Schneider S2 cells[35]. Tethered *Dm* Roquin caused 10-fold repression of a firefly luciferase (F-Luc) reporter containing five binding sites for the λN-tag (BoxB hairpins) in its 3′-UTR (Fig. 3a–c). The reduction in F-Luc activity was accompanied by a corresponding decrease in mRNA abundance (Fig. 3a–c) and a decrease in the half-life of the mRNA (Fig. 3d,e), indicating that *Dm* Roquin induces mRNA degradation in S2 cells. *Dm* Roquin did not affect the expression of an F-Luc reporter that lacked the BoxB hairpins (Supplementary Fig. 1g,h).

As observed in human cells, the Roq-C fragment retained the repressive activity of the full-length protein in the tethering assay and accelerated the degradation of the mRNA reporter (Fig. 3a–e), whereas the Roq-N fragment was inactive (Fig. 3a,b). Furthermore, the full-length *Dm* Roquin and the Roq-C fragment repressed the translation of a F-Luc mRNA reporter with a 3′-end generated by a self-cleaving hammerhead ribozyme (F-Luc-5BoxB-$A_{95}C_7$-HhR; Fig. 3f,g). This reporter is resistant to deadenylation and is therefore not degraded in S2 cells[36]. Similar results have been reported for other RNA-associated proteins that recruit the CCR4-NOT complex[8,9,36], although the involvement of other factors cannot be excluded.

To confirm that mRNA degradation caused by *Dm* Roquin is dependent on the CCR4-NOT complex, we depleted NOT1 in S2 cells. The ability of *Dm* Roquin to elicit the degradation of the F-Luc-5BoxB mRNA was partially suppressed in NOT1-depleted cells (Fig. 3h,i). Western blotting analysis indicated that the NOT1 levels were reduced to $<25\%$ of their control levels in depleted cells (Supplementary Fig. 1i). Thus *Dm* Roquin promotes mRNA degradation by recruiting the CCR4-NOT complex in *Drosophila* cells.

***Dm* Roquin interacts directly with CAF40 and the NOT module.** To identify the subunits of the CCR4-NOT complex that interact with *Dm* Roquin, we expressed the GFP-tagged protein in S2 cells and determined whether it interacts with HA-tagged subunits of the CCR4-NOT complex using co-immunoprecipitation assays. We also tested for interactions of the protein with the PAN2-PAN3 deadenylase complex and with decapping factors. *Dm* Roquin interacted with NOT1, NOT2, NOT3, CAF40, NOT10, CAF1 and PAN3 (Supplementary Fig. 2a–j). These interactions were observed in the presence of RNase A, suggesting that they are not mediated by RNA. *Dm* Roquin also interacted with the decapping factor HPat in an RNA-independent manner but not with other decapping factors (Supplementary Fig. 2k–o). In particular, and in contrast to *Hs* Roquin1 (ref. 32), we observe no interaction of *Dm* Roquin with *Dm* EDC4 or Me31B, which is the *Dm* orthologue of DDX6/RCK (Supplementary Fig. 2n,o).

To discriminate between direct and indirect interactions with subunits of the CCR4-NOT complex, we performed pulldown assays *in vitro* using purified recombinant proteins expressed in *Escherichia coli*. Because *Dm* NOT1 is not expressed in a soluble form in bacteria, we expressed the human proteins, which interact with *Dm* Roquin (as shown in Fig. 1b) and tested them for interaction with *Dm* Roquin *in vitro*.

Initially, we used a purified human pentameric complex consisting of a NOT1 fragment comprising residues 1093–2376, CAF1, CAF40 and the C-terminal domains of NOT2 and NOT3 (Fig. 4a). The *Dm* Roq-C fragment carrying an N-terminal maltose-binding protein (MBP) tag pulled down the purified pentameric complex (Fig. 4b, lane 20), thus demonstrating its direct interaction with the complex.

Figure 1 | The C-terminal regions of Roquin proteins interact with the CCR4-NOT complex and induce degradation of bound mRNA. (**a**) Roquin proteins consist of a conserved N-terminal region containing a RING-finger E3 ubiquitin ligase domain, a ROQ RNA-binding domain flanked by a bilobed HEPN domain and a CCCH-type zinc-finger (ZnF) domain. The N-terminal region is followed by a variable C-terminal extension (shown in grey) that often contains proline-rich sequences (P-rich). The positions of the CBM and the NBM in *Dm* Roquin are indicated. The numbers above the protein outline indicate the residues at domain/motif boundaries. (**b**–**d**) SBP pulldown assays in HEK293T cells expressing V5-SBP-tagged *Hs* Roquin1, *Hs* Roquin2 and *Dm* Roquin (full-length or N- and C-terminal fragments). A V5-SBP-tagged GFP-MBP fusion served as a negative control. The presence of endogenous NOT1, NOT2 and NOT3 in the bound fractions was analysed by western blotting using specific antibodies. The inputs (1.5% for the V5-SBP tagged proteins and 1% for NOT1, NOT2 and NOT3) and bound fractions (10% for the V5-SBP tagged proteins and 30% for NOT1, NOT2 and NOT3) were analysed by western blotting. (**e**–**j**) Tethering assays using the β-globin-6xMS2bs reporter and MS2-HA-tagged *Hs* Roquin2 and *Dm* Roquin (full-length or the indicated fragments) 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**,**h**. (**f**,**i**) show representative northern blottings. (**g**,**j**) show the equivalent expression of the MS2-HA-tagged proteins used in the corresponding tethering assays. (**k**,**l**) Effect of *Dm* Roquin on the expression of the β-globin-TNF-α mRNA reporter analysed as described in **e**–**j**. Protein size markers (kDa) are shown on the right of the western blotting panels. Error bars represent s.d. from three independent experiments. Full images of western and northern blottings are shown in Supplementary Fig. 10.

To map the binding site more precisely, we tested interactions with individual CCR4-NOT subcomplexes, including the NOT1 MIF4G domain bound to CAF1, the NOT1 CN9BD bound to CAF40, a C-terminal connector domain of unknown function (CD) and the NOT module (comprising the NOT1 SHD and the C-terminal regions of NOT2 and NOT3). MBP-tagged *Dm* Roq-C pulled down the CN9BD-CAF40 complex as well as the NOT module but not the NOT1 MIF4G-CAF1 complex or the CD (Fig. 4b, lanes 21–24).

Further analysis indicated that *Dm* Roq-C interacted directly with both *Hs* and *Dm* CAF40 in the absence of the NOT1 CN9BD (Fig. 4c, lanes 11 and 12). We also investigated whether the interaction of *Dm* Roq-C with the NOT module was mediated by the NOT1 C-terminal SHD domain or by the NOT2-NOT3 dimer. However, splitting the NOT module resulted in severely reduced binding to both NOT1 and the NOT2-NOT3 dimer, demonstrating that only the assembled module is recognized efficiently (Supplementary Fig. 3a). In summary, the *Dm* Roq-C fragment contains at least two distinct binding sites for the CCR4-NOT complex, one site that contacts CAF40 and a second site that contacts the NOT module.

Redundancy of CCR4-NOT-binding motifs in *Dm* Roq-C. To define more precisely how *Dm* Roq-C interacts with CAF40 and the NOT module, we sought to identify conserved motifs within the primary sequence. Using only sequences from *Drosophila* species, it was possible to align the *Dm* Roq-C sequences across their entire length (Supplementary Fig. 4). The alignment revealed clusters of conserved residues dispersed throughout the sequence with a higher level of conservation evident at the C-terminal end of Roq-C (Supplementary Fig. 4). We therefore generated two fragments, which we termed Roq-CN and Roq-CC (Fig. 1a, Supplementary Fig. 4). Remarkably, each of these fragments in isolation exhibited repressive activity in tethering assays, indicating functional redundancy (Supplementary Fig. 3b–h). However, only the Roq-CC fragment, comprising residues 702–819, bound to the purified CAF40 protein *in vitro* as efficiently as the entire Roq-C fragment, whereas binding of the Roq-CN fragment (residues 501–702) was strongly impaired (Fig. 4d, lanes 23 and 20 versus lane 17). In contrast, both Roq-CN and Roq-CC retained binding activity for the NOT module, although their binding was reduced compared with that of the full Roq-C fragment (Fig. 4d, lanes 21 and 24 versus lane 18). These results indicate the presence of multiple binding sites for the NOT module within Roq-C.

Through a deletion analysis combined with binding, we then identified a motif comprising residues 790–812 within Roq-CC that is necessary and sufficient for binding to the CAF40 armadillo repeat (ARM) domain (Figs 1a and 4e). Indeed, deletion of the CBM in Roq-CC abolished its binding to CAF40 (Figs 4e, lane 14). Conversely, the CBM is sufficient for binding to CAF40 (Fig. 4e, lane 16 versus lane 12). We also identified

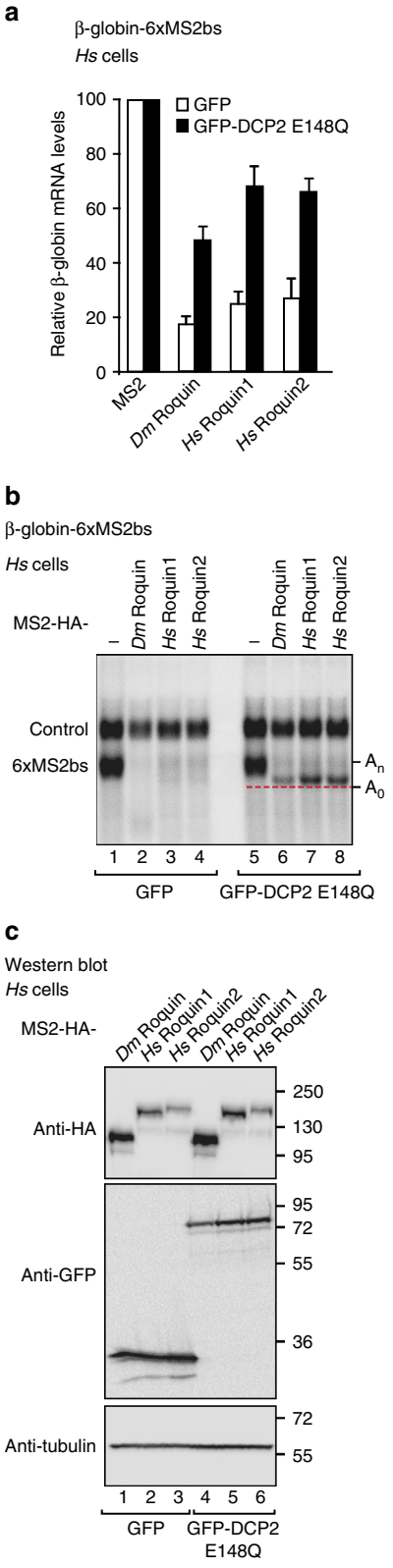


Figure 2 | Roquin proteins degrade mRNAs through the 5′-to-3′ mRNA decay pathway. (**a**,**b**) A tethering assay was performed in HEK293T cells using the β-globin-6xMS2bs reporter as described in Fig. 1e–j, except that different amounts of plasmid were transfected (see Methods section). The transfection mixture included plasmids expressing either GFP or a GFP-tagged catalytically inactive DCP2 mutant (E148Q). The β-globin-6xMS2bs mRNA levels were normalized to those of the control mRNA and set to 100% in the presence of MS2-HA for each condition. The mean values ± s.d. from three independent experiments are shown in **a**. The white and black bars represent the normalized β-globin-6xMS2bs mRNA values in cells expressing GFP and the GFP-DCP2 mutant, respectively. (**b**) shows a representative northern blotting. The positions of the polyadenylated (A_n) and deadenylated (A_0, dashed red line) mRNA reporter are indicated on the right. (**c**) Western blotting analysis showing equivalent expression of MS2-HA-tagged proteins in cells expressing GFP or the DCP2 mutant. Error bars represent s.d. from three independent experiments. Full images of western and northern blottings are shown in Supplementary Fig. 11.

residues 725–755 as being required for binding to the NOT module in the context of Roq-CC (this region was termed the NOT-module binding motif (NBM); Fig. 1a, Supplementary Fig. 5a, lane 14); however, in isolation the NBM was not sufficient for binding, indicating a more complicated binding mode (Supplementary Fig. 5a, lane 16).

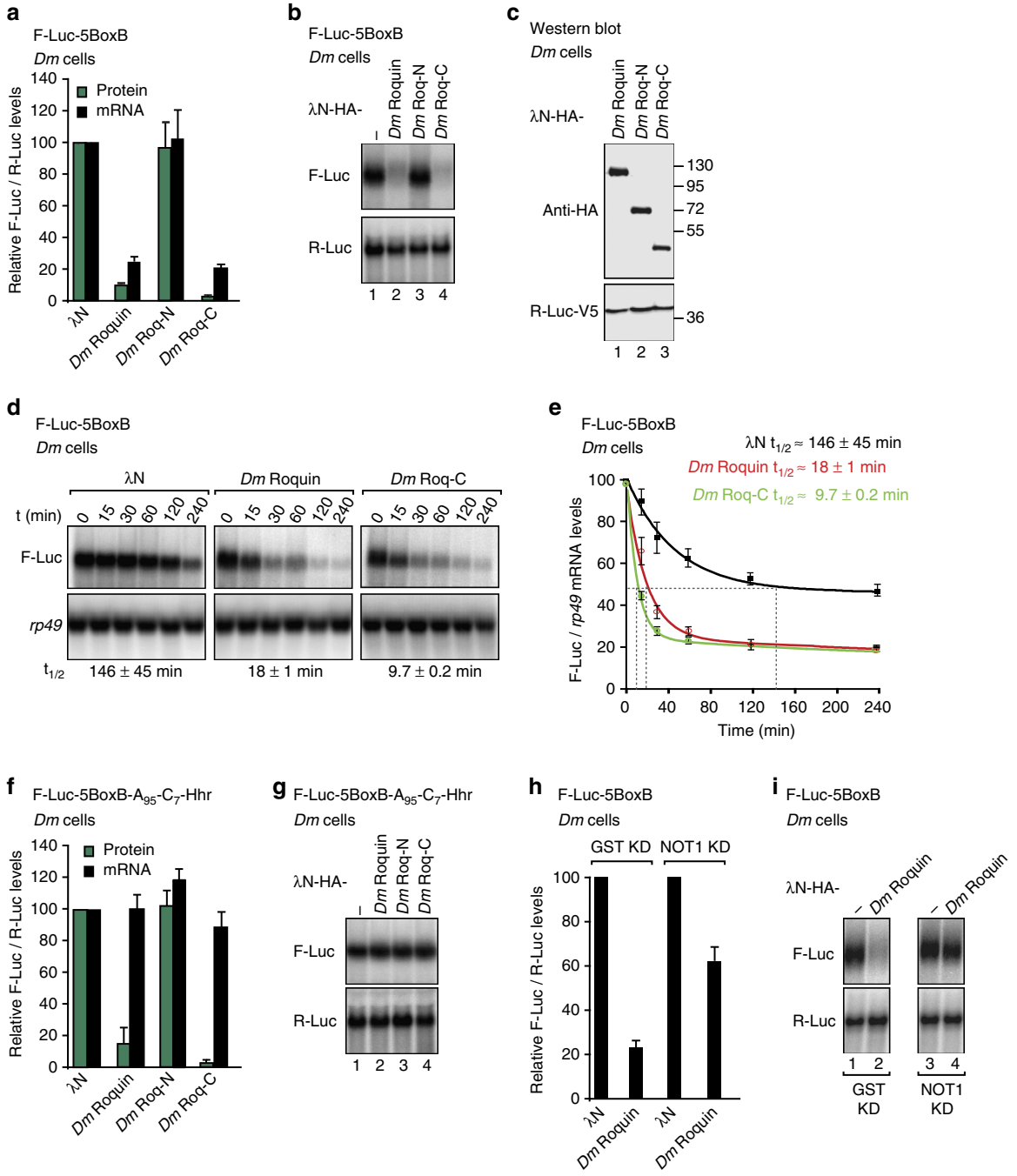


Figure 3 | *Dm* Roquin degrades bound mRNAs. (**a**,**b**) Results of tethering assays using the F-Luc-5BoxB reporter and λN-HA-tagged *Dm* Roquin (full-length or the indicated fragments) in *Dm* S2 cells. A plasmid expressing R-Luc served as a transfection control. F-Luc activity and mRNA levels were normalized to those of the R-Luc transfection control and set to 100% in cells expressing the λN-HA peptide. The mean values ± s.d. from three independent experiments are shown in **a**. (**b**) shows a representative northern blotting. The corresponding control experiment with a F-Luc reporter lacking the BoxB sites is shown in Supplementary Fig. 1g,h. (**c**) Western blotting showing the equivalent expression of the λN-HA-tagged proteins used in **a**,**b**. (**d**,**e**) Representative northern blotting showing the decay of the F-Luc-5BoxB mRNA in S2 cells expressing λN-HA or λN-HA-tagged *Dm* Roquin or the Roq-C fragment after inhibition of transcription by actinomycin D. F-Luc mRNA levels were normalized to those of the *rp49* mRNA and plotted against time. The mRNA half-life ($t_{1/2}$) ± s.d. was calculated from the decay curve shown in **e**. (**f**,**g**) Results of the tethering assays using the F-Luc-5BoxB-A_{95}-C_7-HhR reporter and λN-HA-tagged *Dm* Roquin (full-length or the indicated fragments) in *Dm* S2 cells. The samples were analysed as described in **a**,**b**. (**h**,**i**) Tethering assay using the F-Luc-5BoxB reporter and λN-HA-tagged *Dm* Roquin in *Dm* S2 cells depleted of NOT1 or in control cells (treated with a dsRNA targeting GST). The samples were analysed as described in **a**,**b**. The efficacy of the depletion is shown in Supplementary Fig. 1i. Error bars represent s.d. from three independent experiments. Full images of western and northern blottings are shown in Supplementary Fig. 12.

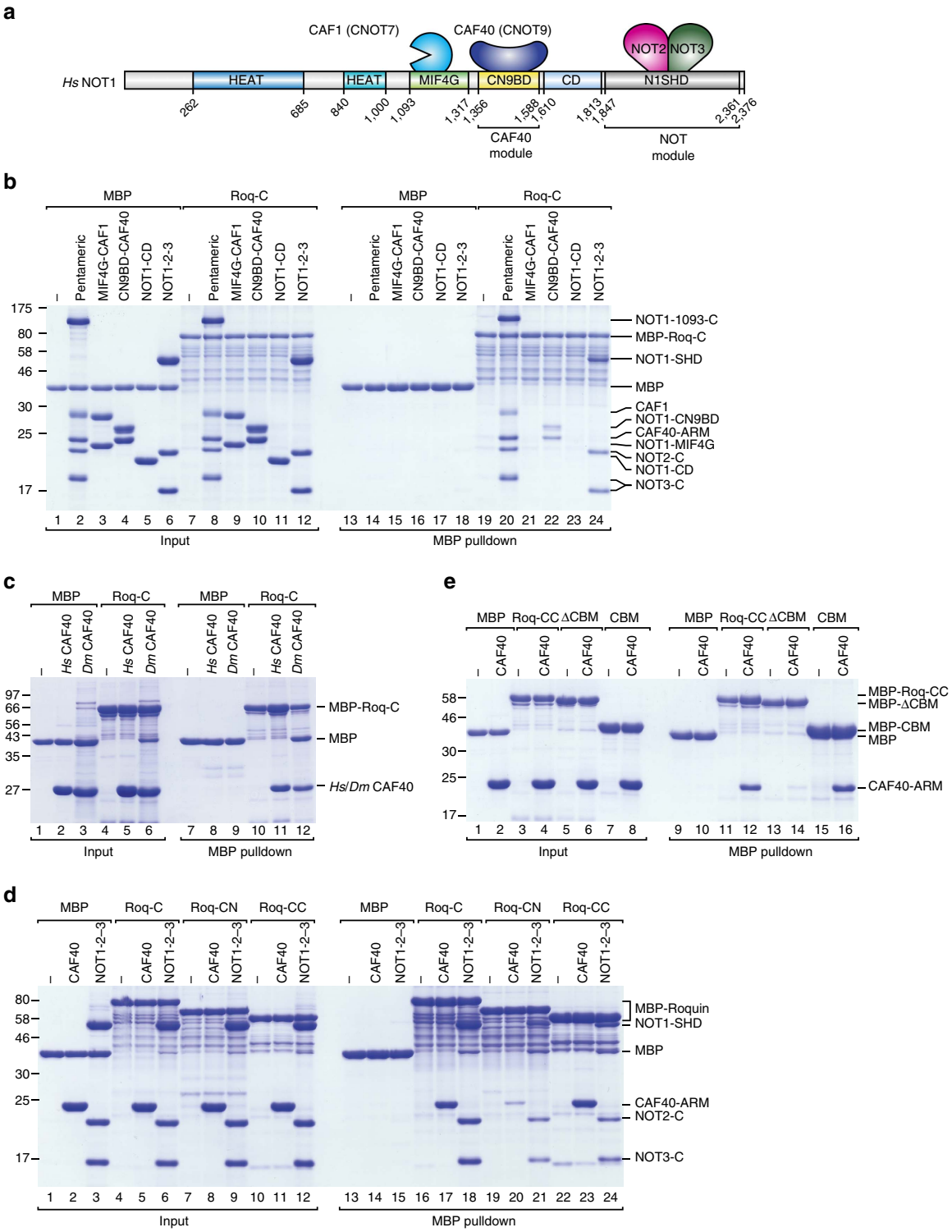
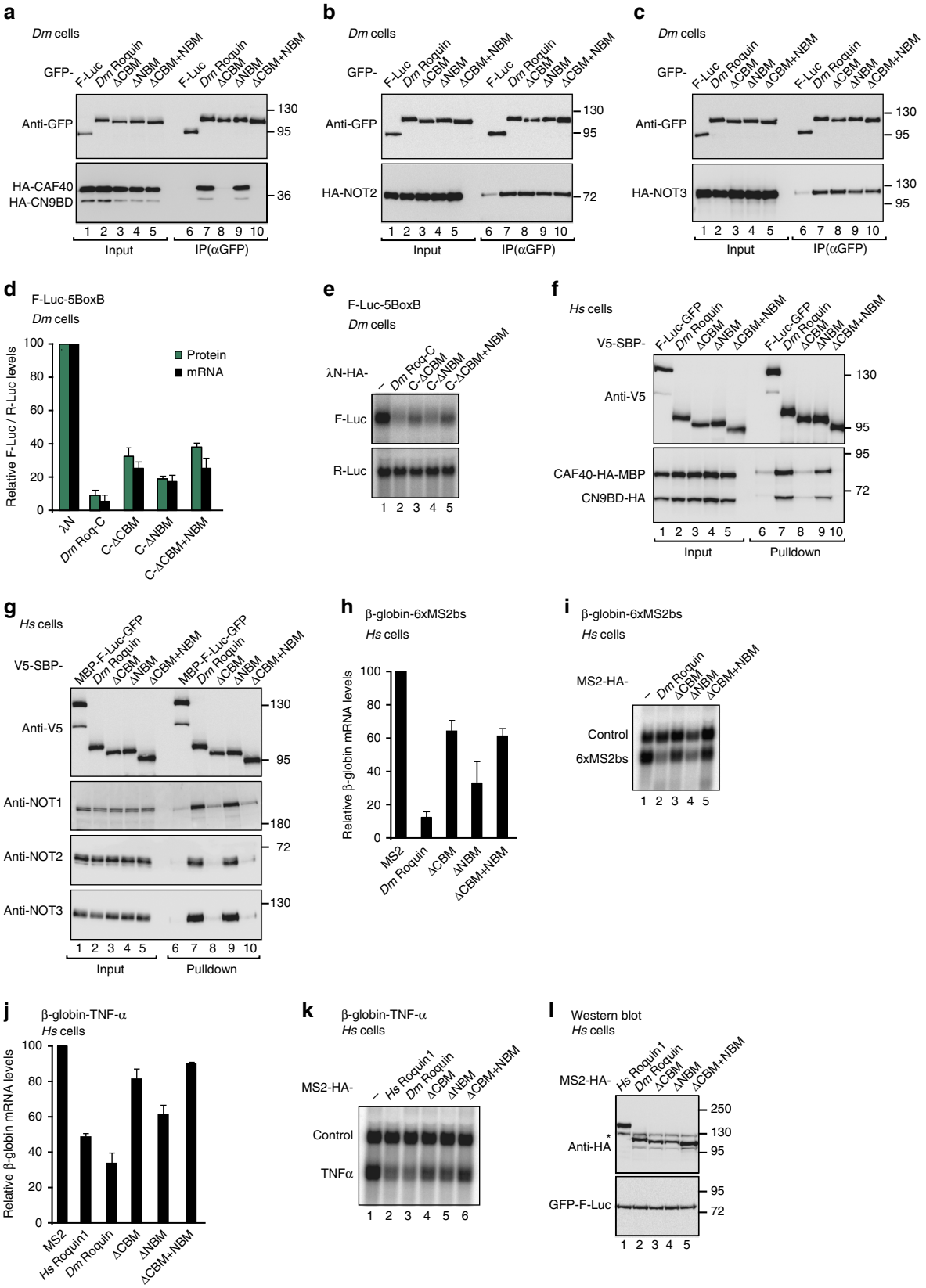


Figure 4 | *Dm* Roq-C interacts directly with CAF40 and the NOT module. (**a**) Schematic representation of the *Hs* CCR4-NOT complex with the subunits used in this study. NOT1 contains two HEAT repeat domains (shown in blue and petrol), a MIF4G domain composed of HEAT repeats (shown in green), a three-helix bundle domain (CN9BD, yellow), a connector domain (CD, light blue) and a NOT1 superfamily homology domain (SHD, grey), which also consists of HEAT repeats. The positions of the other subunits indicate their binding sites on NOT1. (**b**) MBP pulldown assay showing the interaction of MBP-tagged *Dm* Roq-C with purified pentameric NOT1-2-3-7-9 complex, the indicated NOT1 domains and CCR4-NOT subcomplexes. MBP served as a negative control. The difference in the migration of the NOT3-C protein in different samples is due to the presence of a non-cleavable 6xHis-tag in the case of the pentameric NOT1-2-3-7-9 complex, whereas in the case of the NOT1-2-3 complex the HRV3C-cleavable 6xHis tag was removed during purification. (**c**) MBP pulldown assay showing the interaction of MBP-tagged *Dm* Roq-C with purified *Hs* and *Dm* CAF40 proteins. (**d**) MBP pulldown assay showing the interaction of MBP-tagged *Dm* Roquin fragments (Roq-C, Roq-CN and Roq-CC) with the *Hs* CAF40 proteins and the recombinant *Hs* NOT module (NOT1-2-3) containing the NOT1 SHD and the NOT2 and NOT3 C-terminal fragments. (**e**) MBP pulldown assay showing the interaction of MBP-tagged *Dm* Roquin fragments (Roq-CC, Roq-CC-ΔCBM or CBM alone) with the *Hs* CAF40-ARM domain. Protein size markers (kDa) are shown on the left in each panel. Full images of protein gels are shown in Supplementary Fig. 13.

We next tested whether the interaction with CAF40 is also observed for the human Roquin proteins. We observed that *Hs* CAF40 interacts directly with the C-terminal region of *Hs* Roquin1 (Supplementary Fig. 5b); however, more detailed mapping to identify a single CBM was unsuccessful, thus suggesting the presence of multiple binding sites.



The CBM contributes to the activity of *Dm* Roquin. To assess the relative contribution of the CBM and the NBM to the binding of *Dm* Roquin to the CCR4-NOT complex and to the repressive activity of *Dm* Roquin, we deleted these motifs from the protein and performed co-immunoprecipitations and tethering assays in *Dm* S2 cells and human HEK293T cells. Deletion of the CBM abolished the interaction of full-length *Dm* Roquin with the *Dm* CAF40-CN9BD module in S2 cells (Fig. 5a, lane 8). However, this deletion did not affect the binding of *Dm* Roquin to NOT2 or NOT3 (Fig. 5b,c, lane 8). Deletion of the NBM had no effect on CAF40, NOT2 or NOT3 binding (Fig. 5a–c, lane 9), consistent with the observation that *Dm* Roquin harbours multiple binding sites for the NOT module. Accordingly, deletion of the CBM and NBM in the context of Roq-C reduced but did not abolish the activity of this fragment in tethering assays (Fig. 5d,e, Supplementary Fig. 5c–e).

Similarly, in human cells, deletion of the CBM abolished the interaction of full-length *Dm* Roquin with the CAF40-CN9BD module (Fig. 5f, lane 8), indicating that the CBM represents the only binding site for CAF40 in *Dm* Roquin. Importantly, deletion of the CBM also abolished the interaction of full-length *Dm* Roquin with the endogenous CCR4-NOT complex in human cells (Fig. 5g, lane 8). In agreement with these results, deletion of the CBM reduced the ability of *Dm* Roquin to degrade β-globin-6xMS2bs and β-globin-TNF-α mRNAs (Fig. 5h–l). Furthermore, depletion of CAF40 partially suppressed the activity of *Hs* Roquin1, *Hs* Roquin2 and *Dm* Roquin in tethering assays (Supplementary Fig. 6a–c), indicating that CAF40 is indeed an important recruitment factor, but other, redundant interactions compensate for the lack of CAF40 in *Dm* and human cells.

In summary, the interaction of the CBM with CAF40 contributes to the recruitment of the CCR4-NOT complex by *Dm* Roquin. However, the observation that deletion of the NBM and CBM have little effect on the activity of full-length *Dm* Roquin in S2 cells probably indicates that *Dm* Roquin establishes additional uncharacterized interactions with the CCR4-NOT complex or additional binding partners to regulate mRNA targets.

Table 1 | Data collection and refinement statistics.

	CAF40-*Dm*CBM
Data collection	
Space group	$P2_1$
No. of reflections	34,975
Cell dimensions	
a, *b*, *c* (Å)	56.9, 103.7, 60.8
α, *β*, *γ* (°)	90.0, 113.0, 90.0
Wavelength (Å)	1.00001
Resolution (Å)	46.8–2.15 (2.20–2.15)*
R_{sym}	0.048 (0.52)
$I/\sigma I$	13.6 (2.2)
Completeness (%)	98.7 (98.8)
Redundancy	2.9 (2.7)
Refinement	
Resolution (Å)	46.7–2.15
No. of reflections	34,964
R_{work}/R_{free}	18.4%/22.6%
No. of atoms	4,905
Protein	4,678
Water	191
Other solvent molecules	36
B-factors (Å^2)	60.0
Protein	60.0
Water	52.4
Other solvent molecules	102.5
Ramachandran plot	
Favoured regions (%)	98.8
Disallowed regions (%)	0.0
r.m.s. deviations	
Bond lengths (Å)	0.004
Bond angles (°)	0.597

*Values in parentheses are for the highest-resolution shell.

Crystal structure of the *Dm* Roquin CBM bound to CAF40. To elucidate the molecular principles underlying the recruitment of CAF40 by *Dm* Roquin, we sought to determine the crystal structure of the CBM peptide bound to the *Dm* and *Hs* CAF40 ARM domains, which exhibit 81% identity (Supplementary Fig. 7). However, only the complex containing *Hs* CAF40 yielded well-diffracting crystals, from which we obtained a structure at 2.15 Å resolution (Table 1). Two copies of the complex, which are structurally highly similar, are present in the crystal asymmetric unit (Supplementary Fig. 8a,b; root-mean-square deviation (r.m.s.d.) 0.24 Å over 254 C_α atoms).

As previously described, the CAF40 ARM domain consists of 17 α-helices arranged into six armadillo (ARM) repeats. These repeats adopt the typical crescent-like shape of ARM domains (Fig. 6a–c)[2,3,37], with the concave surface that accommodates the CBM peptide (residues 790–810). The CBM peptide folds into an amphipathic helix (residues 795–810) that runs centrally across the concave surface of CAF40 and binds to a conserved hydrophobic patch close to the previously postulated nucleic acid-binding groove (Fig. 6d–f)[37].

Figure 5 | The CBM contributes to the mRNA degradation activity of *Dm* Roquin. (**a**–**c**) Immunoprecipitation assays showing the interaction of GFP-tagged *Dm* Roquin (wild-type or the indicated deletion mutants) with HA-tagged CAF40, NOT2 and NOT3 in *Dm* S2 cells. In (**a**) the interaction was tested in the presence of HA-tagged CN9BD. GFP-tagged firefly luciferase (F-Luc) served as a negative control. Input and immunoprecipitates were analysed using anti-GFP and anti-HA antibodies. For the GFP-tagged proteins, 3% of the input and 10% of the immunoprecipitates were loaded. For the HA-tagged proteins, 1% of the input and 30% of the immunoprecipitates were analysed. In all panels, the cell lysates were treated with RNase A prior to immunoprecipitation. (**d**,**e**) Tethering assay using the F-Luc-5BoxB reporter and λN-HA-tagged *Dm* Roq-C or the indicated deletion mutants in *Dm* S2 cells. The samples were analysed as described in Fig. 3a,b. The corresponding control experiment with a F-Luc reporter lacking the BoxB sites and a western blotting showing the equivalent expression of the tethered proteins are shown in Supplementary Fig. 5c–e. (**f**,**g**) Interaction of V5-SBP-tagged *Dm* Roquin (full-length or the indicated deletion mutants) with HA-tagged CAF40 (in the presence of the HA-tagged CN9BD) and with endogenous NOT1, NOT2 and NOT3 in HEK293T cells. A V5-SBP-tagged MBP-F-Luc-GFP fusion served as a negative control. The inputs (0.75% for V5-SBP-tagged proteins and 1% for NOT1, 2, 3) and bound fractions (5% for SBP-V5-tagged proteins and 30% for NOT1, NOT2 and NOT3) were analysed by western blotting. (**h**,**i**) Tethering assay using the β-globin-6xMS2bs reporter and the indicated MS2-HA-tagged proteins in HEK293T cells. The samples were analysed as described in Fig. 1e–j. (**j**,**k**) The effect of full-length *Dm* Roquin or the indicated deletion mutants on the expression of the β-globin-TNF-α mRNA reporter was analysed as described in Fig. 1k,l. (**l**) Western blotting analysis showing comparable expression of the MS2-HA-tagged proteins used in **h**–**k**. The asterisk indicates cross-reactivity with the anti-HA antibody. Error bars represent s.d. from three independent experiments. Full images of western and northern blottings are shown in Supplementary Fig. 14.

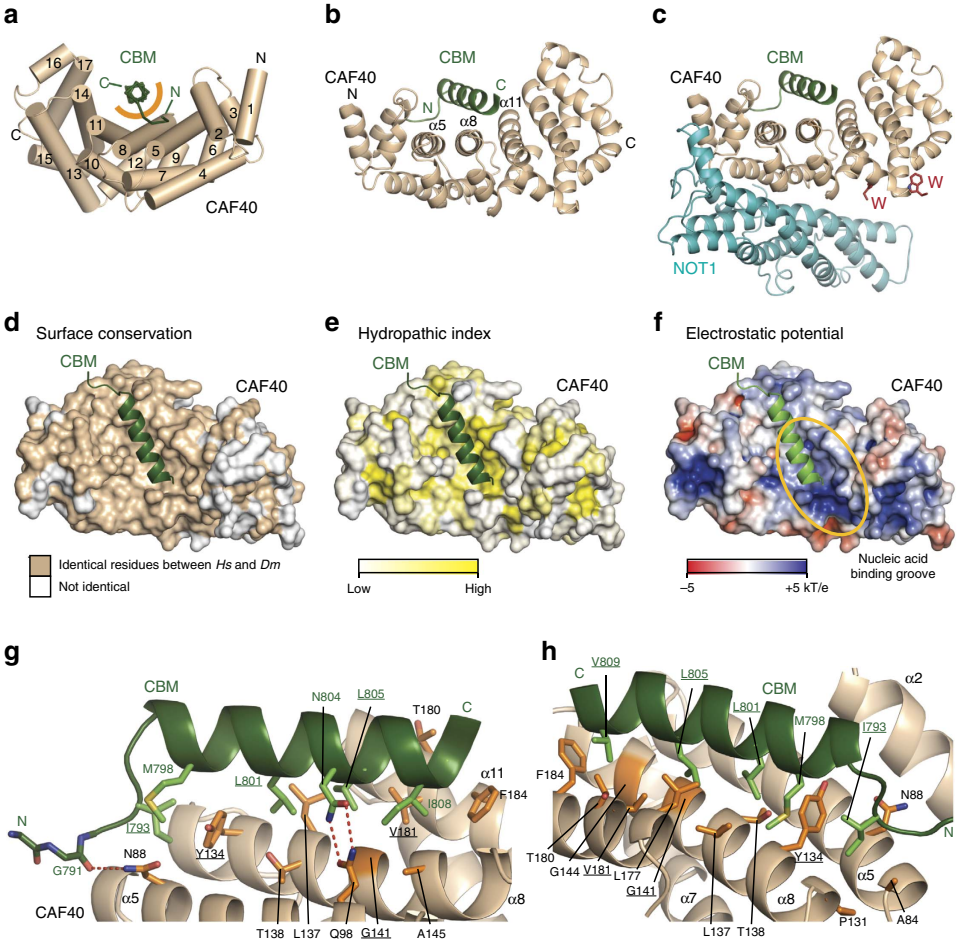


Figure 6 | Structure of the *Dm* Roquin CBM bound to CAF40. (**a**) The *Dm* CBM peptide (green; backbone shown in ribbon representation) bound to *Hs* CAF40 (light brown). The helices of CAF40 are depicted as tubes and are numbered in black. The orange semicircle marks the predominantly hydrophobic interface between the CBM peptide and CAF40. (**b**) Cartoon representation of the *Dm* CBM peptide bound to *Hs* CAF40. Secondary structure elements of CAF40 are labelled in black. (**c**) Structural model of the CBM peptide bound to the CAF40 module (consisting of CAF40 bound to the NOT1 CN9BD domain). The model was obtained by superimposing the structure shown in (**b**) with the structure of CAF40 bound to the NOT1 CN9BD domain and free tryptophan (PDB 4CRU)[2]. (**d**) Conservation of the CBM-binding surface on CAF40. CAF40 is shown in surface representation. Surface residues that are identical between *Hs* and *Dm* are shown in light brown; all other residues are shown in white. (**e**) Surface representation of CAF40 with residues coloured in a gradient from white to yellow with increasing hydophobicity[57]. The CBM is shown in green. (**f**) Surface representation of CAF40 with residues coloured in a gradient from red over white to blue according to the electrostatic potential from -5 to $+5\,kTe^{-1}$. The electrostatic potential was calculated using the APBS tools plugin within PyMOL (http://www.pymol.org). The proposed nucleic acid-binding groove[37] is indicated by an orange circle. (**g**,**h**) Close-up views of the CAF40–CBM-binding interface in two orientations. Selected residues of CAF40 and Roquin are shown as orange and green sticks, respectively. Hydrogen bonds are indicated by red dashed lines. The residues that were mutated in this study are underlined.

Superposition of the CAF40-CBM complex with the structure of the CAF40 dimer[37] (r.m.s.d. of 0.87 Å over 260 Cα atoms; PDB 2FV2) or with that of CAF40 bound to the NOT1 CN9BD (r.m.s.d. of 0.59 Å over 254 Cα atoms, PDB 4CRV)[2] shows that binding of the CBM peptide does not induce major conformational changes in the CAF40 ARM domain (Supplementary Fig. 8c,d). Importantly, binding of the CBM does not interfere with NOT1 binding (Fig. 6c), suggesting that *Dm* Roquin can interact with CAF40 in the context of the CCR4-NOT complex. Finally, binding of the CBM does not block access to the tryptophan-binding pockets on the convex surface of CAF40 that serve as binding sites for the GW182/TNRC6 proteins involved in miRNA-mediated gene silencing (Fig. 6c)[2,3].

The amphipathic helix of the CBM peptide lies almost parallel to helices α5, α8 and α11 and it uses residues M798, L801, L805, I808 and V809 on its hydrophobic side to interact with CAF40 residues Y134, L137, G141, G144 and A145 (helix α8) and L177, T180, V181 and F184 (helix α11; Fig. 6g,h and Supplementary Fig. 8e,f). Furthermore, I793 anchors the N-terminal extension of the CBM helix between CAF40 residues A84, R130, P131 and Y134, resulting in a total buried surface of 1903 Å^2 that does not include any water molecules (Fig. 6h). Finally, the CBM peptide is fixed by two hydrogen bonds between N804 and the CAF40 residue Q98 (helix α5) and by a hydrogen bond from CAF40 N88 (helix α5) to the carbonyl oxygen of G791 (Fig. 6g). The CAF40 residues R130 and K148 may have additional roles in anchoring the CBM peptide, but they have distinct orientations in the two copies of the complex.

CCR4-NOT is recruited via the concave surface of CAF40. To validate the interfaces determined from the crystal structure, we introduced mutations in CBM and CAF40 and tested them in MBP pulldown assays *in vitro*. Substitution of *Dm* Roquin interface residues L805 and V809 by glutamic acid (mutant M2) and the further introduction of I793E and L801E substitutions to create a quadruple mutant (mutant M4) abolished the interaction

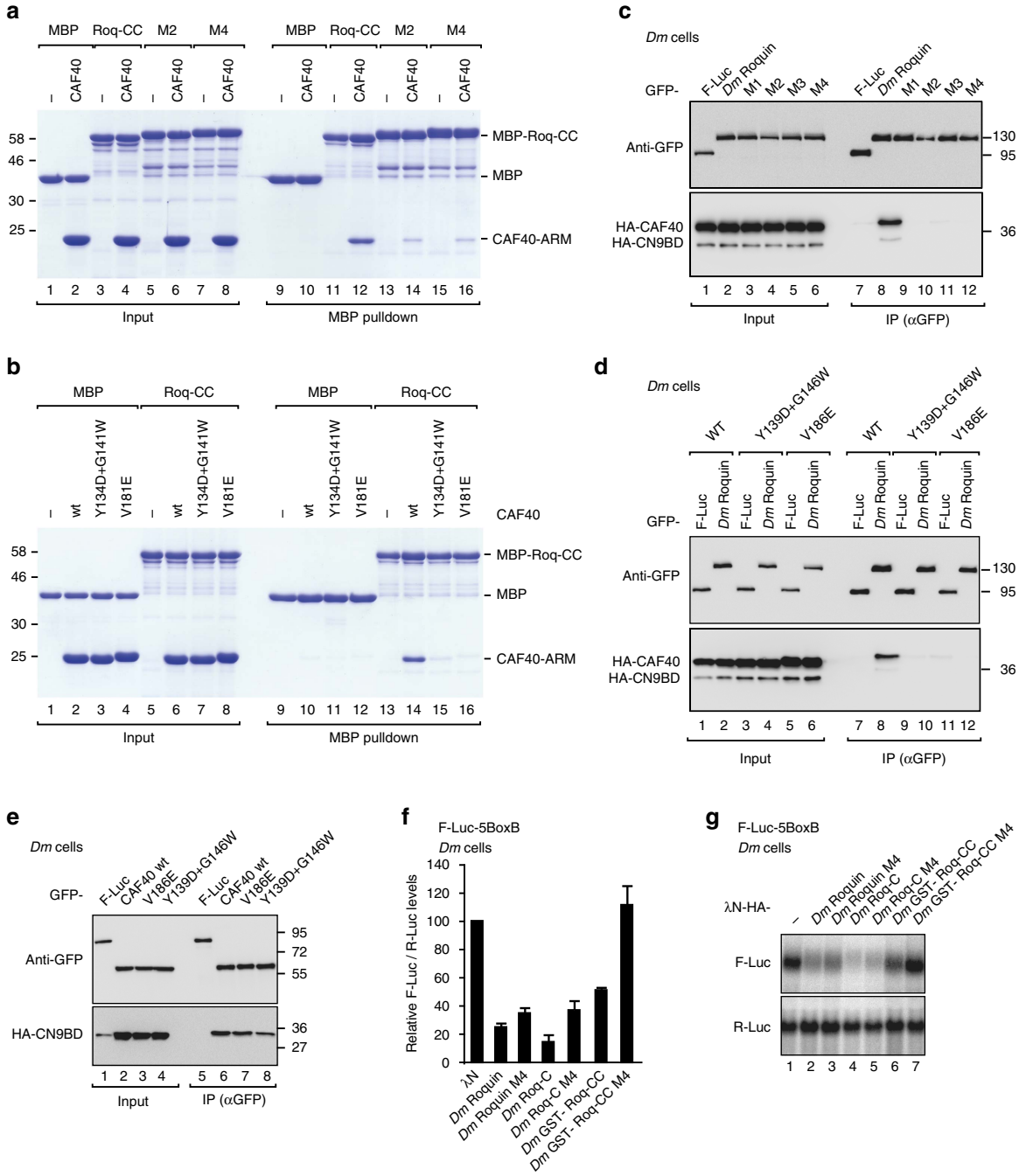


Figure 7 | The CBM is the only CAF40-binding site in *Dm* Roquin. (**a**) *In vitro* MBP pulldown assay showing the interaction of MBP-tagged *Dm* Roq-CC and the indicated mutants (M2 and M4; see Supplementary Table 1) with the purified *Hs* CAF40-ARM domain. MBP served as a negative control. (**b**) *In vitro* MBP pulldown assay showing the interaction of MBP-tagged *Dm* Roq-CC with the *Hs* CAF40-ARM domain (wild-type or the indicated mutants). (**c**) Interaction of GFP-tagged *Dm* Roquin (wild-type or the M1, M2, M3 and M4 mutants; see Supplementary Table 1) with HA-tagged CAF40 in the presence of the HA-tagged CN9BD in *Dm* S2 cells. F-Luc-GFP served as a negative control. (**d**) Interaction of GFP-tagged *Dm* Roquin wild-type with HA-tagged CAF40 (wild-type or the indicated mutants) in the presence of the HA-tagged CN9BD in *Dm* S2 cells. (**e**) Interaction of GFP-tagged *Dm* CAF40 (wild-type or mutants) with HA-tagged CN9BD in *Dm* S2 cells. (**f**,**g**) A tethering assay using the F-Luc-5BoxB reporter and λN-HA-tagged *Dm* Roquin, Roq-C and GST-Roq-CC (full-length or mutant M4) was performed in *Dm* S2 cells as described in Fig. 3a,b. (**f**) shows mean values ± s.d. for normalized F-Luc mRNA levels from three independent experiments. (**g**) shows a northern blotting of representative RNA samples corresponding to the experiment shown in **f**. Error bars represent s.d. from three independent experiments. Full images of western and northern blottings are shown in Supplementary Fig. 15.

of the MBP-tagged Roq-CC with purified *Hs* CAF40 *in vitro* (Fig. 7a, Supplementary Table 1). Conversely, a single V181E substitution in *Hs* CAF40 or substitution of residues Y134 and G141 by aspartic acid and tryptophan, respectively, abolished the interaction of *Hs* CAF40 with MBP-Roq-CC (Fig. 7b, Supplementary Table 1). In contrast, mutation of the concave surface of CAF40 was not sufficient to disrupt the interaction with *Hs* Roq1-C (Supplementary Fig. 5b). Thus, although *Hs* Roq1-C binds to CAF40 directly, it must contact additional and/or alternative CAF40-binding surfaces.

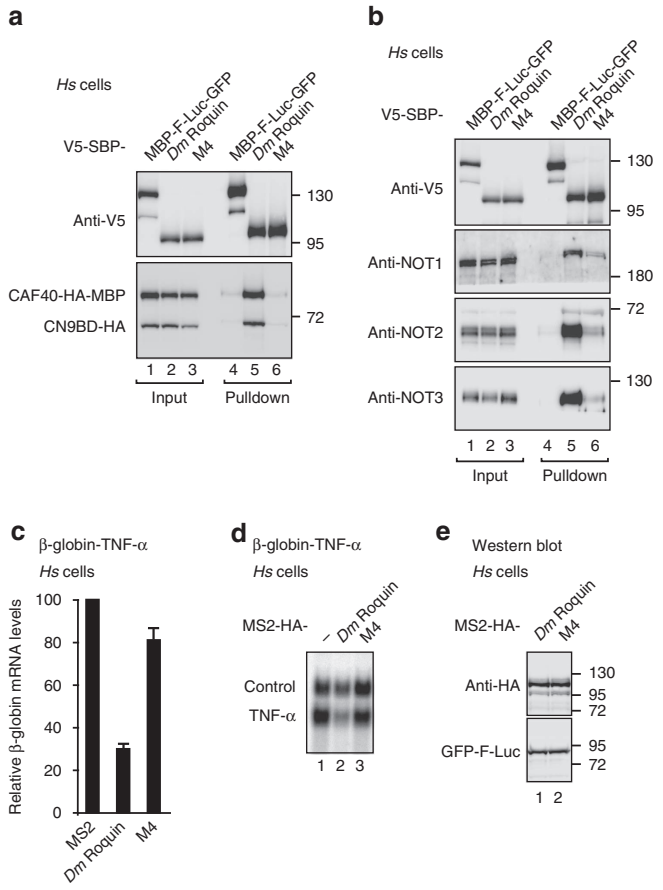


Figure 8 | The CBM mediates the activity of *Dm* Roquin in human cells. (**a**) SBP pulldown assays in HEK293T cell lysates showing the interaction of V5-SBP-tagged *Dm* Roquin (full-length or the M4 mutant) with HA-tagged CAF40-MBP in the presence of CN9BD-HA. A V5-SBP-tagged MBP-F-Luc-GFP fusion served as a negative control. The presence of the HA-tagged proteins in the bound fractions was tested by western blotting using anti-HA antibodies. The samples were analysed by western blotting as described in Fig. 5f. (**b**) SBP pulldown assays in HEK293T cell lysates showing the interaction of V5-SBP-tagged *Dm* Roquin (full-length or the M4 mutant) with endogenous NOT1, NOT2 and NOT3. A V5-SBP-tagged MBP-F-Luc-GFP fusion served as a negative control. The samples were analysed by western blotting as described in Fig. 1b–d. (**c**,**d**) The effect of *Dm* Roquin full-length or the M4 mutant on the expression of the β-globin-TNF-α mRNA reporter was analysed in HEK293T cells as described in Fig. 1k,l. Error bars represent s.d. from three independent experiments. (**e**) Western blotting analysis showing the equivalent expression of the proteins used in the tethering assays shown in **c**,**d**. Full images of western and northern blottings are shown in Supplementary Fig. 16.

The residues in CAF40 that interact with the *Dm* Roquin CBM peptide are conserved in *Hs* and *Dm* (Fig. 6d, Supplementary Fig. 7), suggesting that the CBM interacts with *Dm* CAF40 via a similar binding mode. Nevertheless, it was important to test whether *Dm* Roquin binds to the concave surface of *Dm* CAF40. All tested mutations in the CBM (M1, M2, M3 and M4; Supplementary Table 1) were sufficient to disrupt the interaction of full-length *Dm* Roquin with *Dm* CAF40 in co-immunoprecipitation assays in *Dm* S2 cells (Fig. 7c, lanes 9–12), confirming that the CBM is the only motif in *Dm* Roquin that mediates binding to CAF40. Conversely, substitutions in *Dm* CAF40 corresponding to the mutations of *Hs* CAF40 shown in Fig. 7b abolished the interaction of full-length *Dm* CAF40 with full-length *Dm* Roquin (Fig. 7d). The mutations in CAF40 did not abolish its binding to NOT1 CN9BD, indicating that they do not disrupt the CAF40 fold (Fig. 7e).

Notably, mutations in the CBM did not disrupt the interaction of full-length *Dm* Roquin with NOT2 or NOT3 in *Dm* S2 cells (Supplementary Fig. 9a,b), in agreement with the observation that *Dm* Roquin contains multiple sequences that mediate binding to the CCR4-NOT complex in a redundant manner. Consistent with this redundancy, mutations in the CBM abolished the activity of the Roq-CC fragment in tethering assays in *Dm* S2 cells but impaired the activity of full-length *Dm* Roquin only slightly (Fig. 7f,g, Supplementary Fig. 9c–e). In contrast, when it was tested in human cells, mutations in the *Dm* Roquin CBM not only abolished interaction with CAF40 but strongly reduced the interaction with the endogenous CCR4-NOT complex (Fig. 8a,b) indicating that the CBM provides a major contribution to the interaction of *Dm* Roquin with the CCR4-NOT complex in human cells. Accordingly, the CBM mutants strongly reduced the ability of *Dm* Roquin to degrade the β-globin-TNF-α reporter containing the CDE-37 element in the 3′-UTR in human HEK293T cells (Fig. 8c–e).

Discussion

In this study, we show that Roquin proteins (*Hs* Roquin1, *Hs* Roquin2, *Dm* Roquin) use their C-terminal extensions to directly recruit the CCR4-NOT complex to mRNA targets, promoting the degradation of these mRNAs. We show that this function is conserved among Roquin proteins despite the fact that the sequences of the unstructured C-terminal regions of these proteins are highly divergent and do not share similar motifs.

In *Dm* Roquin, the interaction with the CCR4-NOT complex is mediated by multiple and partially redundant motifs that include a CBM. We further elucidated the structural basis of the interaction of the CBM with CAF40 and identified the concave surface of CAF40 as a target for amphipathic helices to recruit the CCR4-NOT complex.

The finding that Roquin proteins use a combination of SLiMs (or eukaryotic linear motifs)[38] to recruit the CCR4-NOT complex has important functional implications. First, although SLiMs mediate relatively low-affinity interactions, these interactions can be highly specific, and stable binding can be achieved through avidity effects generated by contributions from the flanking disordered regions that extend the interaction interface[38,39]. In the case of *Dm* Roquin, the sequences flanking the CBM provide binding sites for the NOT module, indicating that the C-terminal region of *Dm* Roquin contacts multiple subunits within the CCR4-NOT complex.

A second consequence of the nature of SLiMs is their evolutionary plasticity[38,39]. Owing to their short length and lack of sequence constraints in the absence of a protein fold, even single point mutations can render an existing motif nonfunctional or generate a new motif in another location of the same protein. In particular, the CBM is present in *Drosophila* species, but sequence analysis of Roquin proteins from other insects, worms and vertebrates does not reveal a detectable CBM. Nevertheless, *Hs* Roquin1 and Roquin2 also interact with the CCR4-NOT complex via their unstructured C-terminal regions (ref. 10 and this study), indicating that the overall principle of CCR4-NOT complex recruitment and mRNA target repression is evolutionarily conserved even though the interaction details have diverged across species.

In addition to the previously identified tryptophan-binding pockets on the CAF40 convex surface[2,3], our crystal structure demonstrates that the concave surface of CAF40 is also used to

recruit the CCR4-NOT complex by RNA-associated proteins. The hydrophobic properties of the concave surface are highly conserved, making it an ideal partner for amphipathic helices or disordered, short hydrophobic peptides that are frequently present in RNA-binding proteins.

Therefore, the C-terminal regions of *Hs* Roquin1 and Roquin2 may also target this surface of CAF40, although no clear CBM motifs are detectable by sequence analysis. However, the proteins may contain 'cryptic' CBMs that are likely discontinuous and probably also target neighbouring surfaces on CAF40 because mutations affecting the concave surface of *Hs* CAF40 did not disrupt its interaction with the *Hs* Roquin1 C-terminal fragment.

In addition to the Roquin proteins, a large number of RNA-associated proteins have been shown to recruit the CCR4-NOT complex to their mRNA targets, thereby repressing translation and/or promoting mRNA degradation. These include GW182 proteins, TTP, Nanos and the *Dm* proteins CUP, Bicaudal C, Smaug and Pumilio[2,3,6,8,9,40–44]. For most of these proteins, it has been shown that interaction with the CCR4-NOT complex is mediated by SLiMs embedded in peptide regions of predicted disorder. However, detailed characterization of the binding mode on the structural level is only available for TTP, GW182, vertebrate and *Dm* Nanos and *Dm* Roquin (refs 2,3,6,8,9 and this study). Similar to Roquin, *Dm* Nanos, GW182 proteins and TTP all contain multiple binding sites for different subunits of the CCR4-NOT complex that act redundantly to recruit the complex to mRNA targets. This modular recruitment mode likely enhances the binding affinity and confers redundancy and robustness to the repression mechanism.

As observed for Roquin, the motifs in TTP and vertebrate and *Dm* Nanos, which have been structurally characterized, adopt an α-helical conformation that possibly forms only upon binding (refs 6,8,9 and this study). In contrast, GW182 peptides likely bind to the CCR4-NOT complex in an extended conformation and insert tryptophan residues into tandem hydrophobic pockets exposed at the convex surface of CAF40 and into additional pockets in NOT1 that remain to be identified[2,3].

Because *Dm* Roquin and GW182 proteins can bind CAF40 simultaneously (Fig. 6c), it is possible that the proteins cooperate to recruit the CCR4-NOT complex to mRNAs. The presence of multiple CCR4-NOT-binding partners on an mRNA likely enhances the efficiency of recruitment and the extent of the regulation.

In summary, together with previous studies[2,3,6,8,9], our results indicate that SLiMs in unstructured and poorly conserved regions of RNA-associated proteins represent a common and widespread mode of recruitment of the CCR4-NOT complex to mRNA targets, resulting in a common downstream repressive mechanism that has a major role in posttranscriptional mRNA regulation in eukaryotic cells.

Methods

DNA constructs. Luciferase reporters and plasmids for the expression of GFP- and HA-tagged subunits of the CCR4-NOT and PAN2-PAN3 deadenylase complexes and decapping factors in S2 cells were previously described[35,45,46]. *Dm* Roquin cDNA was purchased from the *Drosophila* Genomics Resource Center, amplified by PCR and inserted into the HindIII and NotI restriction sites of the pAc5.1-λN-HA or pAc5.1-GFP vectors. For expression in HEK293T cells, the cDNA encoding *Dm* Roquin was inserted into the NotI and ApaI restriction sites of the pcDNA3.1-MS2-HA vector[33] and between the HindIII and KpnI sites of the pT7-V5-SBP-C1 vector[47].

A cDNA sequence encoding *Hs* Roquin1 open-reading frame was amplified by PCR from human HEK293T cell total cDNA and inserted between the EcoRI and SacII sites of the pT7-V5-SBP-C1 vector and between the NotI and ApaI sites of the pcDNA3.1-MS2-HA vector. A cDNA encoding *Hs* Roquin2 open-reading frame was amplified by PCR from human HeLa cell total cDNA and inserted between the BamHI and NotI restriction sites of the pcDNA3.1-MS2-HA vector and between the SalI and BamHI restriction sites of the pT7-V5-SBP-C1 vector.

β-Globin reporters containing the wild-type or mutant (MUT16) TNF-α CDE were obtained by replacing the 6xMS2-binding sites in vector β-globin-6xMS2bs[33] with the CDE-37 (TTGGCTCAGACATGTTTTCCGTGAAAACGGAGCTGAA) or CDE-37-MUT16 (TTGGCTCAGACATGTTTTCCGTGAAATGGGAGCT GAA) sequences[10].

For expression of recombinant proteins in *E.coli*, cDNAs encoding *Hs* Roq1-C and *Dm* Roquin fragments were inserted between the AflII and AvrII and the AflII and XbaI restriction sites of the pnYC-pM plasmid[48], respectively, resulting in Roquin fusion proteins carrying N-terminal MBP tags that are cleavable by HRV3C protease. The Roq-C and Roq-CC cDNAs and all of the constructs derived from them also contain a C-terminal GB1 tag[49] followed by a MGSS linker and a 6xHis tag.

A cDNA encoding *Dm* CAF40 was inserted between the NdeI and XbaI restriction sites of the pnEA-pM plasmid[48], generating a fusion protein containing an N-terminal MBP tag cleavable by the HRV3C protease.

For purification of the pentameric NOT1-2-3-7-9 complex, a cDNA encoding human NOT1 (residues 1093–2371) was inserted between the XhoI and BamHI restriction sites of the pnYC-pM vector, resulting in a fusion protein containing an N-terminal MBP-tag. A multicistronic plasmid was generated by inserting an expression cassette containing 6xHis-NOT3-C, MBP-NOT2-C, 6xHis-CAF40-ARM and GST-NOT7 (all tags except the CNOT3 6xHis tag are cleavable by HRV3C protease) into the pnEA vector.

A cDNA encoding NOT1-CD (residues 1607–1815) was inserted between the XhoI and BamHI restriction sites of the pnYC-pM plasmid[48], generating a fusion protein containing an N-terminal MBP tag cleavable by HRV3C protease.

Plasmids for the expression of NOT1-MIF4G, NOT1-CN9BD, NOT1-SHD, NOT2-C, NOT3-C, NOT7 and CAF40-ARM have been previously described[2,5,50]. The DNA constructs used in this study are listed in Supplementary Table 1.

Co-immunoprecipitation and SBP-pulldown assays. For co-immunoprecipitation assays in S2 cells (ATCC), 2.5×10^6 cells were seeded per well in six-well plates and transfected using Effectene transfection reagent (Qiagen). The transfection mixtures contained 1 μg of plasmid expressing HA-tagged deadenylase or decapping factors and 1.5 μg of GFP-tagged Roquin (either full length or fragments). S2 cells were harvested 3 days after transfection and co-immunoprecipitation assays were performed as previously described[51].

For SBP pulldown assays in human cells, HEK293T cells (ATCC) were grown in 10-cm dishes and transfected using TurboFect transfection reagent (Thermo Fisher Scientific). The transfection mixtures contained 6, 30 and 2 μg of plasmids expressing *Hs* Roquin1, *Hs* Roquin2 and *Dm* Roquin, respectively. In the experiment shown in Fig. 1c, the transfection mixture contained 25 μg of plasmids expressing *Hs* Roquin2 and Roq2-N and 12 μg of a plasmid expressing Roq2-C. In the experiment shown in Fig. 1d, the transfection mixtures contained 15, 20 and 25 μg of plasmids expressing *Dm* Roquin, Roq-N and Roq-C, respectively. In the experiments shown in Fig. 5f,g, 10 μg of plasmids expressing *Dm* Roquin, ΔCBM and ΔNBM and 15 μg of plasmid expressing *Dm* Roquin ΔCBM + NBM were included. In the experiments shown in Fig. 8a,b, 10 μg of the *Dm* Roquin and *Dm* Roquin M4 plasmids was transfected. In the experiments shown in Figs 5f and 8a, the transfection mixtures contained 7.5 μg of a plasmid expressing CAF40-HA-MBP and 5 μg of a plasmid expressing CN9BD-HA. In the experiments shown in Supplementary Fig. 1a, the transfection mixtures contained 1 μg of plasmids expressing *Hs* Roquin1 and Roq1-C and 10 μg of a plasmid expressing Roq1-N. Human cells were harvested 2 days after transfection, and co-immunoprecipitation assays were performed as previously described[8]. Western blottings were developed using the ECL Western Blotting Detection System (GE Healthcare) according to the manufacturer's recommendations. The antibodies used in this study are listed in Supplementary Table 2.

Tethering assays in human and S2 cells and RNA interference. Tethering assays in human HEK293T cells using the β-globin reporter containing six MS2-binding sites (6xMS2bs)[33] were performed as previously described[8]. Briefly, cells were seeded in six-well plates (0.8×10^6 cells per well) and transfected using Lipofectamine 2000 (Thermo Fisher Scientific).

The transfection mixtures contained 0.5 μg of the control plasmid (containing the β-globin gene fused to a fragment of the GAPDH gene but lacking MS2-binding sites)[33], 0.5 μg of the β-globin-6xMS2bs reporter and varying amounts of pcDNA3.1-MS2-HA plasmids expressing MS2-HA-tagged proteins. The plasmid amounts were as follows: for *Hs* Roquin1, 0.6 μg full length and 0.5 μg each Roq1-N and Roq1-C; for *Hs* Roquin2, 1.5 μg full length, 0.6 μg Roq2-N and 0.3 μg Roq2-C; for *Dm* Roquin, 0.3 μg full length, 0.1 μg each Roq-N and Roq-C, 0.2 μg each Roq-ΔCBM and ΔNBM and 0.3 μg Roq-ΔCBM + NBM. In the experiment shown in Fig. 2a,b, 0.175 μg of *Dm* Roquin, 0.4 μg of *Hs* Roquin1 and 1 μg of *Hs* Roquin2 were transfected. The transfection mixtures also contained plasmids expressing GFP (0.2 μg) or a GFP-tagged catalytically inactive DCP2 mutant (1.5 μg), as indicated.

When the β-globin-TNF-α reporter was used, cells were co-transfected with 0.5 μg β-globin-TNF-α reporter (CDE37 wild-type or MUT16; ref. 10) and plasmids expressing the HA-MS2 tagged proteins (*Hs* Roquin1: 0.5 μg full-length, 0.4 μg each Roq1-N and Roq1-C; *Dm* Roquin: 0.2 μg full-length, 0.1 μg each Roq-N, Roq-C, ΔCBM and ΔNBM and 0.2 μg ΔCBM + ΔNBM). In the

experiments shown in Fig. 8c–e, the cells were transfected with 0.05 μg of *Dm* Roquin and *Dm* Roquin M4. Cells were harvested 2 days after transfection.

For the λN-tethering assay 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 0.1 μg firefly luciferase reporter (F-Luc-5BoxB, F-Luc-V5 or F-Luc-5BoxB-$A_{95}C_7$-HhR), 0.4 μg *Renilla* luciferase transfection control and various amounts of plasmids expressing λN-HA-tagged *Dm* Roquin constructs (0.01 μg full-length, 0.003 μg Roq-N, 0.025 μg Roq-C, 0.05 μg GST-Roq-CN, and 0.02 μg GST-Roq-CC). In the experiment described in Fig. 7f,g and Supplementary Fig. 9c–e, half of these amounts were transfected. In the experiment described in Supplementary Fig. 3g,h, the GST tags were replaced by GFP tags and 0.02 μg GFP-Roq-CN, and 0.005 μg GFP-Roq-CC were transfected. Cells were harvested 3 days after transfection.

NOT1 knockdowns using dsRNA were performed as previously described[35]. To measure the mRNA half-life, cells were treated with actinomycin D (5 μg ml^{-1} final concentration) 3 days after transfection and collected at the indicated time points. RNA samples were analysed by northern blotting. The level of reporter mRNA was normalized to the levels of endogenous *rp49* mRNA in three independent experiments, averaged and plotted against time. The data were fitted to a double exponential decay function prior to averaging. The reported half-lives ($t_{1/2}$) correspond to 50% decay with respect to the initial amount of reporter RNA. Half-life errors are calculated from the standard fitting error.

Knockdowns in HeLa cells (provided by O. Mühlemann) were performed as described previously[2]. The 19 nt target sequences are as follows: control 5′-ATT CTCCGAACGTGTCACG-3′, CAF40 5′-GATCTATCAGTGGATCAAT-3′. Cells were transfected in six-well plates using Lipofectamine 2000 according to the manufacturer's protocol. Transfection mixtures contained 0.2 μg of *Dm* Roq-C; 0.3 μg of *Hs* Roq1-C and 0.4 μg of *Hs* Roq2-C.

Firefly and *Renilla* luciferase activities were measured using a Dual-Luciferase Reporter Assay system (Promega). Northern blotting was performed as previously described[35].

Protein expression and purification. All recombinant proteins were expressed in *E. coli* BL21 (DE3) Star cells (Invitrogen) in LB medium at 20 °C overnight. *Dm* Roquin fragments were expressed as fusion proteins containing N-terminal MBP tags cleavable by HRV3C protease[48]. In addition, the *Dm* Roq-C, -CN and -CC fragments carried an HRV3C-cleavable C-terminal GB1-6xHis tag[49]. The cells were lysed using an Avestin Emulsiflex-C3 homogenizer in lysis buffer (50 mM HEPES-NaOH (pH 7.5), 200 mM NaCl, 20 mM imidazole and 2 mM β-mercaptoethanol) supplemented with complete EDTA-free protease inhibitors (Roche), 5 μg ml^{-1} DNaseI and 1 mg ml^{-1} lysozyme. The proteins were separated from the crude lysate using amylose resin (New England Biolabs) and subsequently eluted from the resin in lysis buffer containing 25 mM D-(+)-maltose. The proteins containing a GB1-6xHis tag were further purified by nickel affinity chromatography using a HiTrap IMAC column (GE Healthcare). Proteins without GB1-6xHis tags were further purified by anion exchange chromatography using a HiTrapQ column (GE Healthcare). The final purification step for all proteins was size exclusion chromatography using a Superdex 200 16/600 column (GE Healthcare) in a buffer containing 10 mM HEPES-NaOH (pH 7.5), 200 mM NaCl and 2 mM DTT.

Hs CAF40 (residues 19–285) was expressed with a 6xHis tag cleavable by the HRV3C protease. Lysis was carried out in lysis buffer containing 50 mM potassium phosphate (pH 7.5), 500 mM NaCl, 10% glycerol, 20 mM imidazole and 2 mM β-mercaptoethanol supplemented with complete EDTA-free protease inhibitors, DNaseI and lysozyme. The *Hs* 6xHis-CAF40 ARM domain was isolated from the lysate using a HiTrap IMAC column (GE Healthcare). The 6xHis tag was removed by cleavage using HRV3C protease during dialysis in low salt buffer containing 50 mM Tris-HCl (pH 7.5), 150 mM NaCl and 1 mM DTT. Subsequently, *Hs* 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 gel filtration buffer containing 20 mM Tris-HCl (pH 7.5), 150 mM NaCl and 1 mM DTT.

Dm CAF40 (residues 25–291) was expressed with an N-terminal MBP tag. The cells were lysed in a buffer containing 50 mM HEPES-NaOH (pH 7.5), 500 mM NaCl and 2 mM DTT supplemented with complete EDTA-free protease inhibitors, DNaseI and lysozyme. The protein was isolated from the crude lysate using amylose resin followed by anion exchange chromatography using a HiTrapQ column (GE Healthcare). The MBP tag was removed by cleavage with HRV3C protease during dialysis in a buffer containing 10 mM potassium phosphate (pH 8.0), 400 mM NaCl, 10% glycerol, 50 mM ammonium sulfate and 2 mM DTT. The final purification step was size exclusion chromatography using a Superdex 200 26/600 column equilibrated with the same buffer.

The pentameric human NOT1-2-3-7-9 complex was obtained by co-expression of MBP-NOT1 (residues 1093–2371), MBP-NOT2-C, 6xHis-NOT3-C, GST-NOT7 and 6xHis-CAF40-ARM. The cells were lysed in buffer containing 50 mM potassium phosphate (pH 7.5), 300 mM NaCl and 2 mM DTT supplemented with complete EDTA-free protease inhibitors, DNaseI and lysozyme. The complex was purified using amylose resin, and the tags were removed by cleavage with HRV3C protease during dialysis in buffer containing 20 mM HEPES-NaOH (pH 7.5), 100 mM NaCl, 5% glycerol and 2 mM DTT. The complex was further purified using a HiTrap Heparin column (GE Healthcare) followed by size exclusion chromatography on a Superdex 200 26/600 column equilibrated in 20 mM HEPES-NaOH (pH 7.5), 300 mM NaCl, 5% glycerol and 2 mM DTT.

To purify the complex containing 6xHis-NOT1-MIF4G (residues 1093–1317) bound to MBP-CAF1, cells coexpressing the proteins were lysed in a buffer containing 50 mM HEPES-NaOH (pH 7.5), 300 mM NaCl, 20 mM imidazole and 2 mM β-mercaptoethanol supplemented with complete EDTA-free protease inhibitors, DNaseI and lysozyme. The complex was isolated from the crude lysate using amylose resin followed by chromatography on a HiTrap IMAC column. The affinity tags were removed by cleavage using the HRV3C protease, followed by final size exclusion chromatography on a Superdex 200 26/600 column in a buffer containing 10 mM HEPES-NaOH (pH 7.5), 200 mM NaCl and 2 mM DTT.

The NOT1 CD (residues 1607–1815) was expressed with an N-terminal MBP tag. The cells were lysed in a buffer containing 50 mM HEPES (pH 7.5), 300 mM NaCl and 2 mM DTT supplemented with complete EDTA-free protease inhibitors, DNaseI and lysozyme. The protein was isolated from the lysate using amylose resin followed by a HiTrapQ column. The MBP tag was removed by cleavage with HRV3C protease, followed by final size exclusion chromatography on a Superdex 200 26/600 column using a buffer containing 10 mM HEPES-NaOH (pH 7.5), 200 mM NaCl and 2 mM DTT.

Purification of the *Hs* NOT module comprising NOT1-SHD (residues 1833–2361), NOT2-C (residues 350–540) and NOT3-C (residues 607–748) and the complex comprising NOT1-CN9BD (residues 1356–1588) bound to the CAF40-ARM has been previously described[2,9].

The *Dm* Roquin CBM peptide (residues 790–810) used for crystallization was synthesized by EMC Microcollections. The peptide was dissolved in 20 mM Tris-HCl (pH 7.5), 150 mM NaCl and 1 mM DTT.

Crystallization. *Dm* Roquin CBM was mixed with the purified *Hs* CAF40 ARM domain in an equimolar ratio. Initial crystallization screens were carried out using the sitting drop vapor diffusion method at 22 °C by mixing 200 nl of the CAF40-Roquin complex solution (at 6.75 mg ml^{-1}) with 200 nl of reservoir solution. Crystals appeared within 1 day under several conditions. Optimized crystals grew at 18 °C after 1 day using hanging drops after mixing 0.8 μl of the CAF40-Roquin complex solution (at 2.25 mg ml^{-1}) with 0.8 μl of reservoir solution containing 0.1 M NaOAC (pH 5.0), 17.5% (w v^{-1}) PEG 4000 and 0.1 M $AmSO_4$. Crystals were cryoprotected using reservoir solution supplemented with 15% glycerol and flash-frozen in liquid nitrogen.

Data collection and structure determination. Diffraction data were collected on a PILATUS 6M detector (Dectris) at the PXII beamline of the Swiss Light Source. The best data set extending to a resolution of 2.15 Å was recorded at a wavelength of 1.000 Å and processed in space group $P2_1$ using XDS and XSCALE[52]. Two copies of the CAF40 ARM domain (PDB-ID 2FV2, chain A) were found in the asymmetric unit by molecular replacement using PHASER[53] from the CCP4 package[54]. This initial model was improved by iterative cycles of model building in COOT[55] and refinement using PHENIX[56]. As the final step, two Roquin CBM peptides were modeled into the density and improved by further refinement cycles.

The final model was refined with excellent stereochemistry to $R_{work} = 18.4\%$ and $R_{free} = 22.6\%$ and includes all residues of the two CAF40 molecules (residues 19–285 plus six residues from the HRV3C cleavage site and linker sequences; chains A and C) and all residues (790–810) of the Roquin peptides for both chains (B and D). For the C-terminal two pairs of α-helices in CAF40, the final atomic B-factors are clearly above average, pointing to an elevated mobility and/or statistical disorder in this part of the molecule, which is not involved in binding the CBM.

***In vitro* MBP-pulldown assays.** Purified MBP (20 μg) or MBP-tagged Roquin fragments (40 μg) were incubated with equimolar amounts of purified CCR4-NOT subcomplexes and 50 μl of the amylose resin slurry (New England Biolabs) in 1 ml of pulldown buffer (50 mM HEPES-NaOH (pH 7.5), 200 mM NaCl, 2 mM DTT). After a 1-h incubation, the beads were washed five times with pulldown buffer and the proteins were eluted with pulldown buffer supplemented with 25 mM D-(+)-maltose. The eluted proteins were precipitated with trichloroacetic acid and analysed by SDS-PAGE followed by Coomassie blue staining.

Data availability. The coordinates for the structure of the *Dm* Roquin CBM peptide bound to CAF40 were deposited in the Protein Data Bank (PDB) under ID code 5LSW. The authors declare that the data supporting the findings of this study and relevant source data are available within the article and its Supplementary Information file. Other data and materials are available from the authors upon request.

References

1. Wahle, E. & Winkler, G. S. RNA decay machines: deadenylation by the Ccr4–Not and Pan2–Pan3 complexes. *Biochim. Biophys. Acta* **1829,** 561–570 (2013).
2. Chen, Y. *et al.* A DDX6-CNOT1 complex and W-binding pockets in CNOT9 reveal direct links between miRNA target recognition and silencing. *Mol. Cell* **54,** 737–750 (2014).

3. Mathys, H. *et al.* Structural and biochemical insights to the role of the CCR4-NOT complex and DDX6 ATPase in MicroRNA repression. *Mol. Cell* **54,** 751–765 (2014).
4. Bhaskar, V. *et al.* Structure and RNA-binding properties of the Not1–Not2–Not5 module of the yeast Ccr4–Not complex. *Nat. Struct. Mol. Biol.* **20,** 1281–1288 (2013).
5. Boland, A. *et al.* Structure and assembly of the NOT module of the human CCR4–NOT complex. *Nat. Struct. Mol. Biol.* **20,** 1289–1297 (2013).
6. Fabian, M. R. *et al.* Structural basis for the recruitment of the human CCR4–NOT deadenylase complex by Tristetraprolin. *Nat. Struct. Mol. Biol.* **20,** 735–739 (2013).
7. Suzuki, A., Saba, R., Miyoshi, K., Morita, Y. & Saga, Y. Interaction between NANOS2 and the CCR4-NOT deadenylation complex is essential for male germ cell development in mouse. *PLoS ONE* **7,** e33558 (2012).
8. Bhandari, D., Raisch, T., Weichenrieder, O., Jonas, S. & Izaurralde, E. Structural basis for the Nanos-mediated recruitment of the CCR4–NOT complex and translational repression. *Genes Dev.* **28,** 888–901 (2014).
9. Raisch, T. *et al.* Distinct modes of recruitment of the CCR4-NOT complex by Drosophila and vertebrate Nanos. *EMBO J.* **35,** 974–990 (2016).
10. Leppek, K. *et al.* Roquin promotes constitutive mRNA decay via a conserved class of stem-loop recognition motifs. *Cell* **153,** 869–881 (2013).
11. Heissmeyer, V. & Vogel, K. U. Molecular control of Tfh-cell differentiation by Roquin family proteins. *Immunol. Rev.* **253,** 273–289 (2013).
12. Schaefer, J. S. & Klein, J. R. Roquin – a multifunctional regulator of immune homeostasis. *Genes Immun.* **17,** 79–84 (2016).
13. Newman, J. R. S. & Keating, A. E. Comprehensive identification of human bZIP interactions with coiled-coil arrays. *Science* **300,** 2097–2101 (2003).
14. Vinuesa, C. G. *et al.* A RING-type ubiquitin ligase family member required to repress follicular helper T cells and autoimmunity. *Nature* **435,** 452–458 (2005).
15. Schlundt, A. *et al.* Structural basis for RNA recognition in roquin-mediated post-transcriptional gene regulation. *Nat. Struct. Mol. Biol.* **21,** 671–678 (2014).
16. Tan, D., Zhou, M., Kiledjian, M. & Tong, L. The ROQ domain of Roquin recognizes mRNA constitutive decay element and double-stranded RNA. *Nat. Struct. Mol. Biol.* **21,** 679–685 (2014).
17. Murakawa, Y., Rosenbaum, E., Landthaler, M. & Heinemann, U. Roquin binding to target mRNAs involves a winged helix-turn-helix motif. *Nat. Commun.* **5,** 5701 (2014).
18. Zhang, Q. *et al.* New insights into the RNA-binding and E3 ubiquitin ligase activities of roquins. *Sci. Rep.* **5,** 15660 (2015).
19. Sakurai, S., Ohto, U. & Shimizu, T. Structure of human Roquin-2 and its complex with constitutive-decay element RNA. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **71,** 1048–1054 (2015).
20. Schlundt, A., Niessing, D., Heissmeyer, V. & Sattler, M. RNA recognition by Roquin in posttranscriptional gene regulation. *Wiley Interdiscip. Rev. RNA* **7,** 455–469 (2016).
21. Li, W., Gao, B., Lee, S.-M., Bennett, K. & Fang, D. RLE-1, an E3 ubiquitin ligase, regulates *C. elegans* aging by catalyzing DAF-16 polyubiquitination. *Dev. Cell* **12,** 235–246 (2007).
22. Maruyama, T. *et al.* Roquin-2 promotes ubiquitin-mediated degradation of ASK1 to regulate stress responses. *Sci. Signal.* **7,** ra8–ra8 (2014).
23. Yu, D. *et al.* Roquin represses autoimmunity by limiting inducible T-cell co-stimulator messenger RNA. *Nature* **450,** 299–303 (2007).
24. Linterman, M. A. *et al.* Roquin differentiates the specialized functions of duplicated T cell costimulatory receptor genes Cd28 and Icos. *Immunity* **30,** 228–241 (2009).
25. Vogel, K. U. *et al.* Roquin paralogs 1 and 2 redundantly repress the Icos and Ox40 costimulator mRNAs and control follicular helper T cell differentiation. *Immunity* **38,** 655–668 (2013).
26. Schuetz, A., Murakawa, Y., Rosenbaum, E., Landthaler, M. & Heinemann, U. Roquin binding to target mRNAs involves a winged helix-turn-helix motif. *Nat. Commun.* **5,** 5701 (2014).
27. Schaefer, J. S., Montufar-Solis, D., Nakra, N., Vigneswaran, N. & Klein, J. R. Small intestine inflammation in Roquin-mutant and Roquin-deficient mice. *PLoS ONE* **8,** e56436 (2013).
28. Murakawa, Y. *et al.* RC3H1 post-transcriptionally regulates A20 mRNA and modulates the activity of the IKK/NF-κB pathway. *Nat. Commun.* **6,** 7367 (2015).
29. Mino, T. *et al.* Regnase-1 and Roquin regulate a common element in inflammatory mRNAs by spatiotemporally distinct mechanisms. *Cell* **161,** 1058–1073 (2015).
30. Janowski, R. *et al.* Roquin recognizes a non-canonical hexaloop structure in the 3′-UTR of Ox40. *Nat. Commun.* **7,** 11032 (2016).
31. Jeltsch, K. M. *et al.* Cleavage of roquin and regnase-1 by the paracaspase MALT1 releases their cooperatively repressed targets to promote TH17 differentiation. *Nat. Immunol.* **15,** 1079–1089 (2014).
32. Glasmacher, E. *et al.* Roquin binds inducible costimulator mRNA and effectors of mRNA decay to induce microRNA-independent post-transcriptional repression. *Nat. Immunol.* **11,** 725–733 (2010).
33. Lykke-Andersen, J., Shu, M.-D. & Steitz, J. A. Human Upf proteins target an mRNA for nonsense-mediated decay when bound downstream of a termination codon. *Cell* **103,** 1121–1131 (2000).
34. Chang, C.-T., Bercovich, N., Loh, B., Jonas, S. & Izaurralde, E. 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 (2014).
35. Behm-Ansmant, I. *et al.* mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. *Genes Dev.* **20,** 1885–1898 (2006).
36. Zekri, L., Kuzuoğlu-Öztürk, D. & Izaurralde, E. GW182 proteins cause PABP dissociation from silenced miRNA targets in the absence of deadenylation. *EMBO J.* **32,** 1052–1065 (2013).
37. Garces, R. G., Gillon, W. & Pai, E. F. Atomic model of human Rcd-1 reveals an armadillo-like-repeat protein with *in vitro* nucleic acid binding properties. *Protein Sci. Publ. Protein Soc.* **16,** 176–188 (2007).
38. Davey, N. E. *et al.* SLiMPrints: conservation-based discovery of functional motif fingerprints in intrinsically disordered protein regions. *Nucleic Acids Res.* **40,** 10628–10641 (2012).
39. Tompa, P. Intrinsically disordered proteins: a 10-year recap. *Trends Biochem. Sci.* **37,** 509–516 (2012).
40. Semotok, J. L. *et al.* Smaug recruits the CCR4/POP2/NOT deadenylase complex to trigger maternal transcript localization in the early drosophila embryo. *Curr. Biol.* **15,** 284–294 (2005).
41. Goldstrohm, A. C., Hook, B. A., Seay, D. J. & Wickens, M. PUF proteins bind Pop2p to regulate messenger RNAs. *Nat. Struct. Mol. Biol.* **13,** 533–539 (2006).
42. Chicoine, J. *et al.* Bicaudal-C recruits CCR4-NOT deadenylase to target mRNAs and regulates oogenesis, cytoskeletal organization, and its own expression. *Dev. Cell* **13,** 691–704 (2007).
43. Igreja, C. & Izaurralde, E. CUP promotes deadenylation and inhibits decapping of mRNA targets. *Genes Dev.* **25,** 1955–1967 (2011).
44. Van Etten, J. *et al.* Human Pumilio proteins recruit multiple deadenylases to efficiently repress messenger RNAs. *J. Biol. Chem.* **287,** 36370–36383 (2012).
45. Tritschler, F. *et al.* 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 (2008).
46. Haas, G. *et al.* HPat provides a link between deadenylation and decapping in metazoa. *J. Cell Biol.* **189,** 289–302 (2010).
47. Jonas, S., Weichenrieder, O. & Izaurralde, E. 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 (2013).
48. Diebold, M.-L., Fribourg, S., Koch, M., Metzger, T. & Romier, C. 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 (2011).
49. Cheng, Y. & Patel, D. J. An efficient system for small protein expression and refolding. *Biochem. Biophys. Res. Commun.* **317,** 401–405 (2004).
50. Petit, A.-P. *et al.* 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 (2012).
51. Braun, J. E., Huntzinger, E., Fauser, M. & Izaurralde, E. GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets. *Mol. Cell* **44,** 120–133 (2011).
52. Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 125–132 (2010).
53. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674 (2007).
54. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67,** 235–242 (2011).
55. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 486–501 (2010).
56. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68,** 352–367 (2012).
57. Kyte, J. & Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157,** 105–132 (1982).

Acknowledgements

We thank Dr Oliver Mühlemann for the kind gift of the HeLa cell line used in knockdown experiments and Dr Jens Lykke-Andersen for providing the plasmids for performing MS2 tethering assays in human cells. We are grateful to Heike Budde, Sigrun Helms, Maria Fauser and Catrin Weiler for excellent technical support. We thank the staff at the PX beamlines of the Swiss Light Source for assistance with data collection. This work was supported by the Max Planck Society.

Author contributions

A.S. and P.B. cloned proteins for expression in human and S2 cells. A.S. conducted tethering assays in human and *Drosophila* cells and co-immunoprecipitation assays in *Drosophila* cells with help from D.K.-O. and D.B. P.B. conducted pulldown assays in

human cells. T.R. and A.S. expressed and purified proteins from *E. coli* for crystallization and pulldown assays, crystallized the CAF40-CBM complex, conducted *in vitro* pulldown assays and analysed the data. T.R. solved and refined the crystal structure of the CAF40-CBM complex. Y.C. cloned and purified the pentameric complex. O.W. supervised the structural part of the study. E.I. was the principal investigator and conceived and supervised the project. A.S., T.R., O.W. and E.I. wrote the manuscript with contributions from P.B. and D.K.-O.

Additional information

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

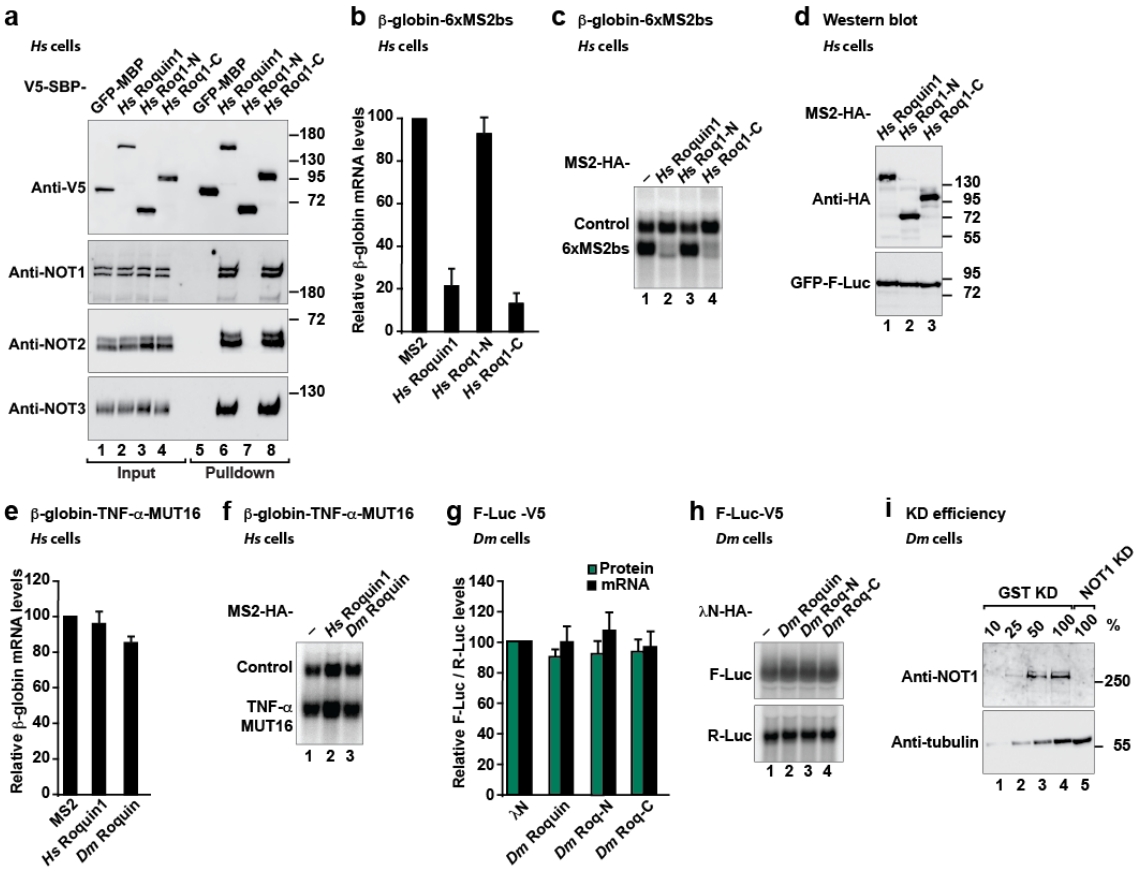
Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Sgromo, A. *et al.* 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 (2017).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

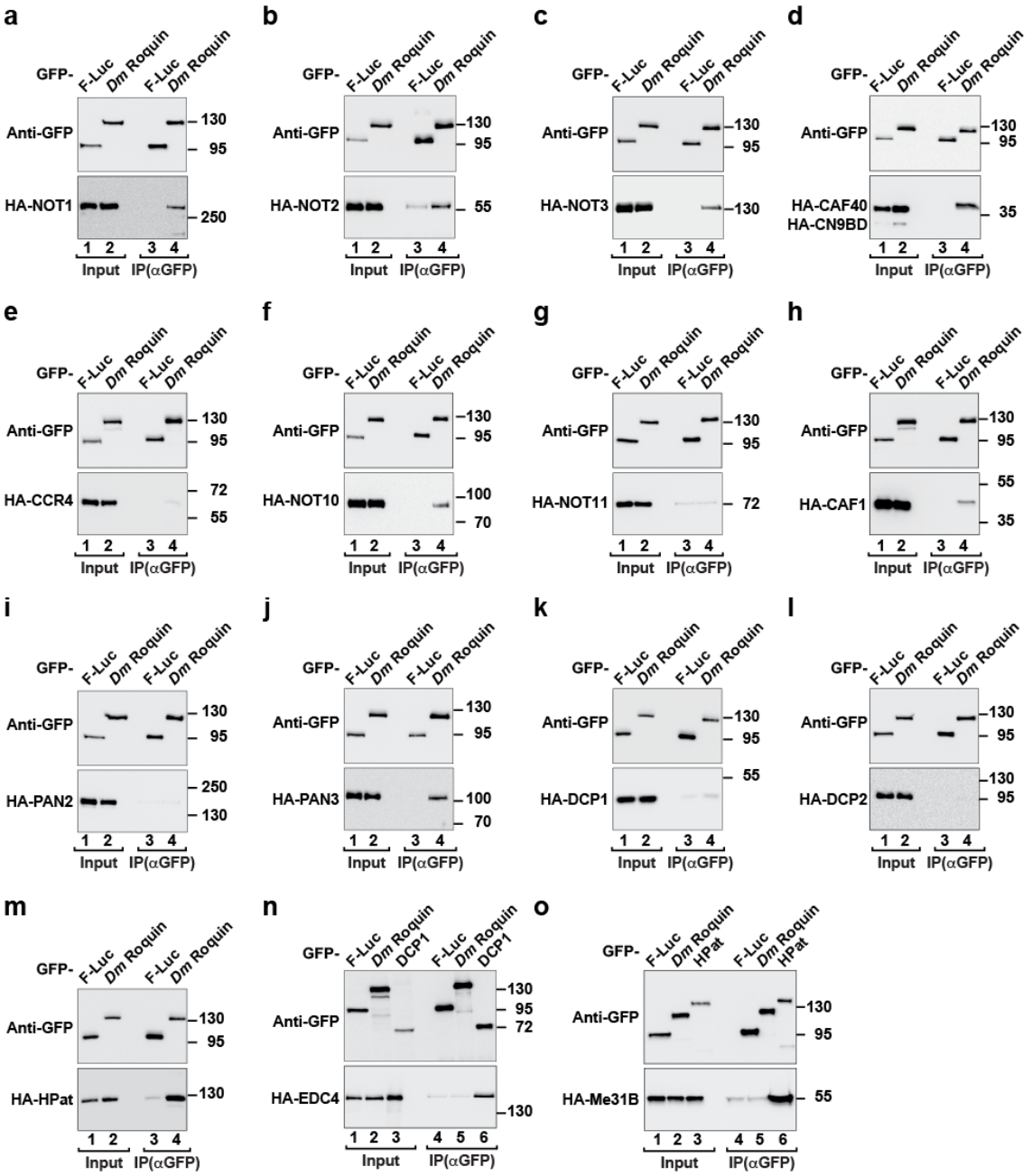
Supplementary Figure 1



Supplementary Figure 1 | The C-terminal regions of Roquin proteins recruit the CCR4-NOT complex and trigger degradation of bound mRNA. (**a**) SBP pulldown assays in HEK293T cell lysates showing the interaction of V5-SBP-tagged *Hs* Roquin1 (full-length or the indicated fragments) with endogenous NOT1, NOT2 and NOT3. Input and immunoprecipitates were analyzed as described in Fig. 1b–d. (**b,c**) Tethering assays using the β-globin-6xMS2bs reporter and MS2-HA-tagged *Hs* Roquin1 (full-length or the indicated fragments) 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 levels were normalized to those of the control mRNA and set to 100% in the presence of MS2-HA. The mean values ± s.d. from three independent experiments are shown in panel (b). Panel (c) shows a representative

northern blot. (**d**) Western blot showing the equivalent expression of the MS2-HA-tagged proteins used in the tethering assay shown in panels (b,c). (**e,f**) Effect of *Hs* Roquin1 and *Dm* Roquin on the expression of the β-globin-TNFα-MUT16 mRNA reporter carrying a mutated CDE (MUT16; ref. 10) analyzed as described in Fig. 1k,l. (**g,h**) A tethering assay using the F-Luc-V5 (lacking the 5BoxB sites) reporter in *Dm* S2 cells co-expressing λN-HA-*Dm* Roquin (full-length or the indicated fragments) was performed as described in Fig. 3a,b. (g) Mean values ± s.d. from three independent experiments. (h) Northern blot of representative RNA samples. (**i**) Western blot analysis of S2 cells depleted of NOT1 corresponding to the experiment shown in Fig. 3h,i. Dilutions of control cell lysates were loaded in lanes (1–4) to estimate the efficacy of the depletion. Tubulin served as a loading control. KD: knockdown.

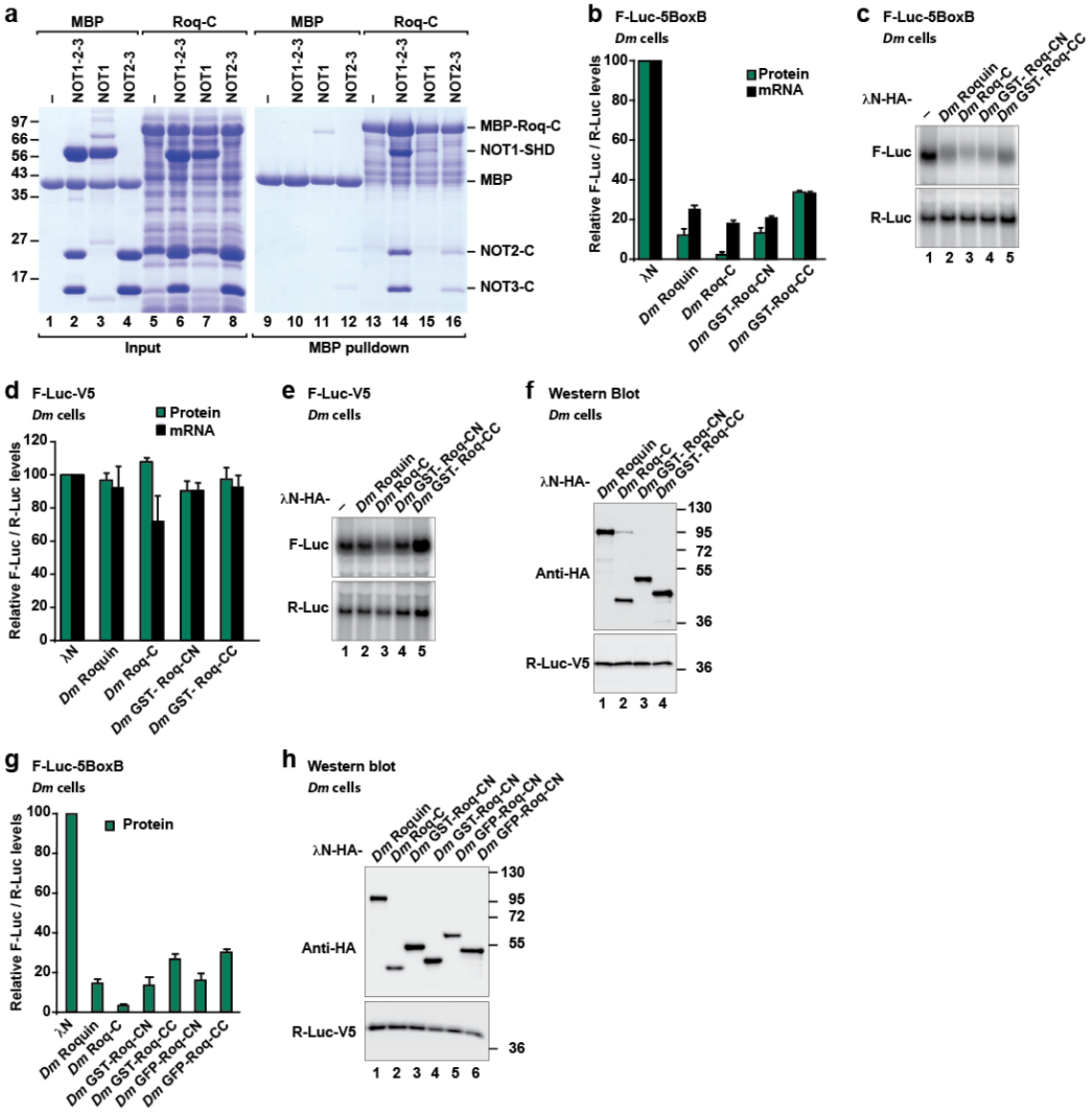
Supplementary Figure 2



Supplementary Figure 2 | *Dm* Roquin interacts with subunits of the deadenylase and decapping complexes. (**a–o**) Co-immunoprecipitation assays showing the interaction of GFP-tagged *Dm* Roquin with the indicated HA-tagged proteins in *Dm* S2 cell lysates treated with RNase A. In panels (n) and (o), GFP-tagged DCP1 and HPat,

respectively, served as positive controls. In all panels, F-Luc-GFP served as a negative control. Protein size markers (kDa) are shown on the right in each panel.

Supplementary Figure 3

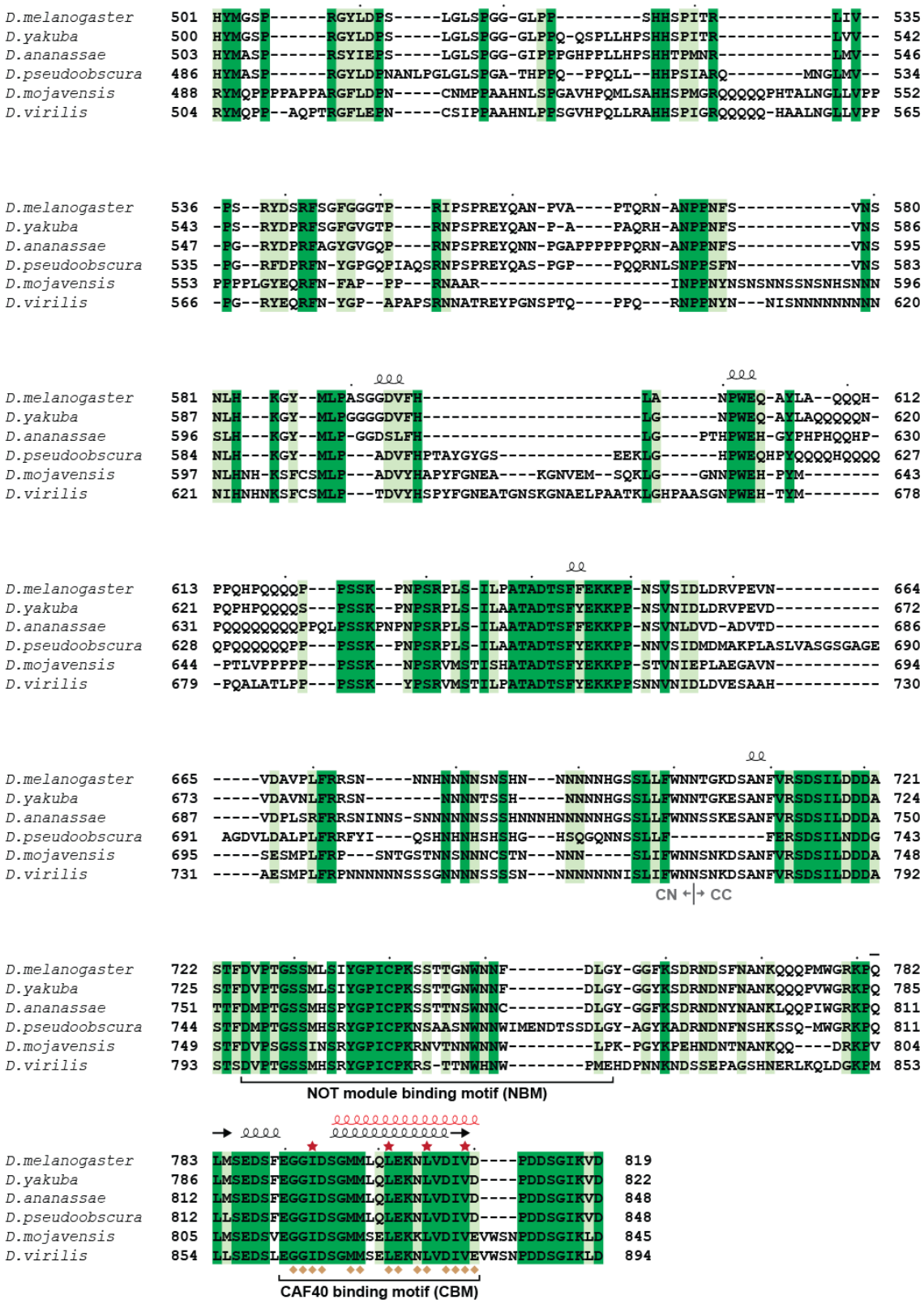


Supplementary Figure 3 | The *Dm* Roquin C-terminal region binds to CAF40 and to the NOT-module. (**a**) *In vitro* MBP pulldown assay showing the interaction of the MBP-tagged *Dm* Roq-C with the assembled NOT module (NOT1-2-3), the isolated NOT1-SHD (NOT1) and the NOT2-NOT3 dimers (NOT2-3). MBP served as a negative

control. (**b–e**) Tethering assays using the F-Luc-5BoxB or the F-Luc-V5 (lacking the 5BoxB sites) reporters and λN-HA-tagged *Dm* Roquin (full-length or the indicated fragments) were performed in *Dm* S2 cells as described in Fig. 3a,b. A plasmid expressing R-Luc served as a transfection control. 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. Panels (b,d) show mean values ± s.d. of the normalized F-Luc activities and mRNA levels from three independent experiments. Panels (c,e) show northern blots of representative RNA samples. (**f**) Western blot analysis showing the equivalent expression of the λN-HA-tagged proteins used in panels (b–e). (**g**) A tethering assay using the F-Luc-5BoxB reporter and λN-HA-tagged *Dm* Roquin (full-length or the indicated fragments) was performed in *Dm* S2 cells as described in Fig. 3a,b. The *Dm* Roq-CC and Roq-CN contained in addition either a GST or a GFP tag, with the purpose of visualization. A plasmid expressing R-Luc served as the transfection control. F-Luc activity was normalized to that of the R-Luc transfection control and set to 100% in cells expressing the λN-HA peptide. The panel shows mean values ± s.d. of the normalized F-Luc activities from three independent experiments. The GST or GFP tags do not interfere with the activity of the fragments. (**h**) Western blot analysis showing the equivalent expression of the λN-HA-tagged proteins used in the tethering assays shown in panel (g).

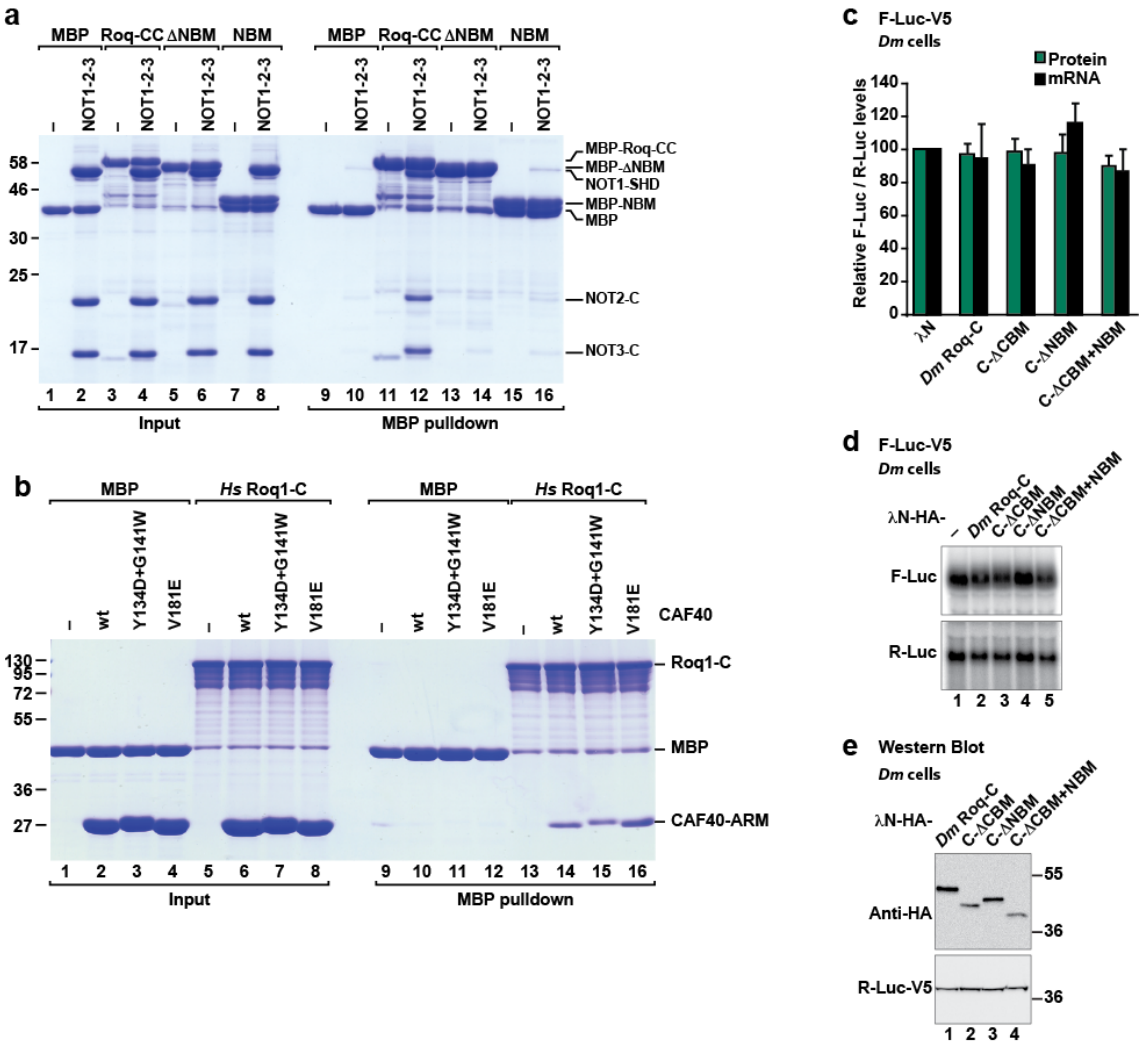
Error bars represent s. d. from three independent experiments.

Supplementary Figure 4



Supplementary Figure 4 | Sequence alignment of the *Drosophila* Roquin C-terminal regions. The CBM helix as determined from the crystal structure is indicated in red above the sequences. The secondary structure elements as predicted by PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/) are indicated in black. Residues conserved in all of the aligned sequences are shown with a dark green background, and residues with >70% similarity are highlighted in light green; conservation scores were calculated using the SCORECONS webserver[1]. The sequences required for binding to the NOT module (NBM) and the CAF40-binding motif (CBM) are indicated. Residues in the CBM that directly contact CAF40 are marked by brown diamonds. Residues that were mutated in this study are indicated by red asterisks. The boundary between the CN and CC fragments is indicated below the alignment.

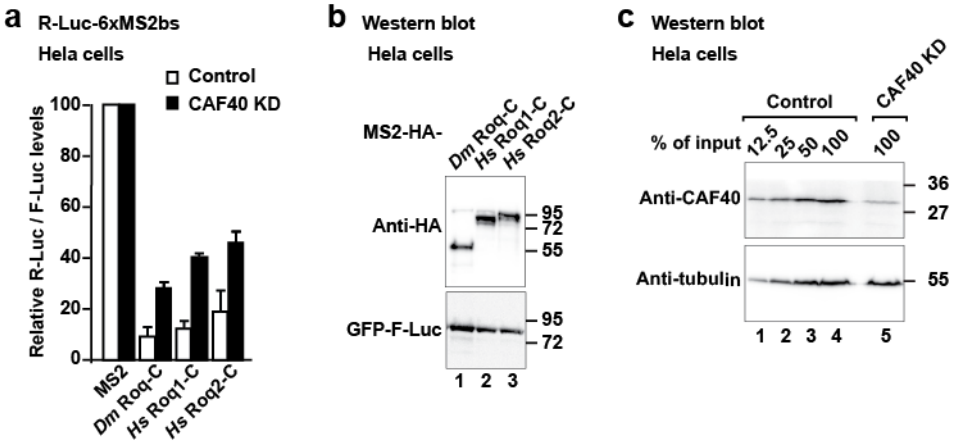
Supplementary Figure 5



Supplementary Figure 5 | The CBM contributes to the activity of *Dm* Roquin. (**a**) *In vitro* MBP pulldown assay showing the interaction of MBP-tagged *Dm* Roq-CC, Roq-CC-ΔNBM or the NBM alone with the assembled NOT module (NOT1-2-3). MBP served as a negative control. (**b**) MPB pulldown assay showing the interaction of MPB-tagged *Hs* Roq1-C with the purified *Hs* CAF40-ARM domain (wild-type and the indicated mutants). MBP served as a negative control. (**c,d**) A tethering assay using the F-Luc-V5 reporter lacking λN binding sites and the indicated λN-HA-tagged *Dm* Roquin fragments was performed in *Dm* S2 cells as described in Fig. 3a,b. The corresponding experiment using the F-Luc-5BoxB reporter is shown in Fig. 5d,e. Error

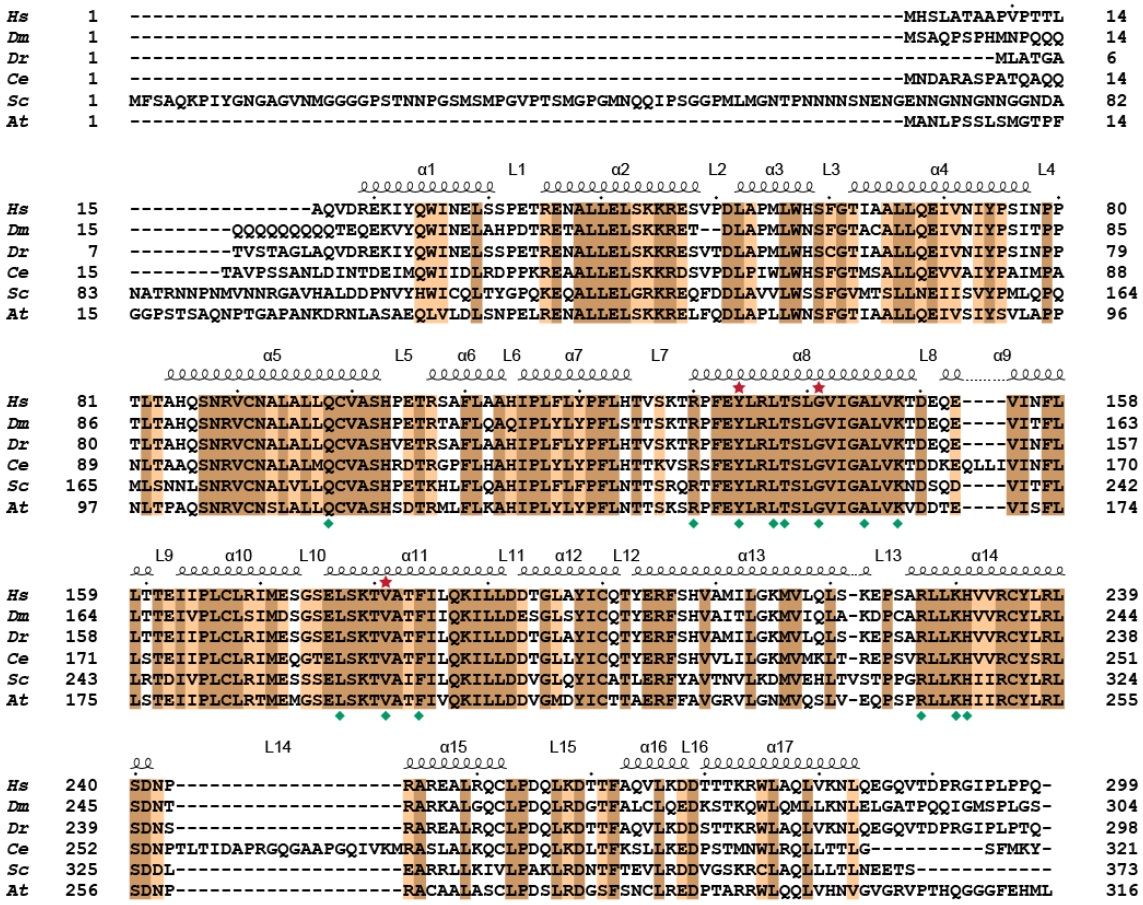
bars represent s.d. from three independent experiments. (**e**) Western blot showing the equivalent expression of the λN-HA-tagged proteins used in panels (c,d) and in Fig. 5d,e. Protein size markers (kDa) are shown on the right of the panels.

Supplementary Figure 6



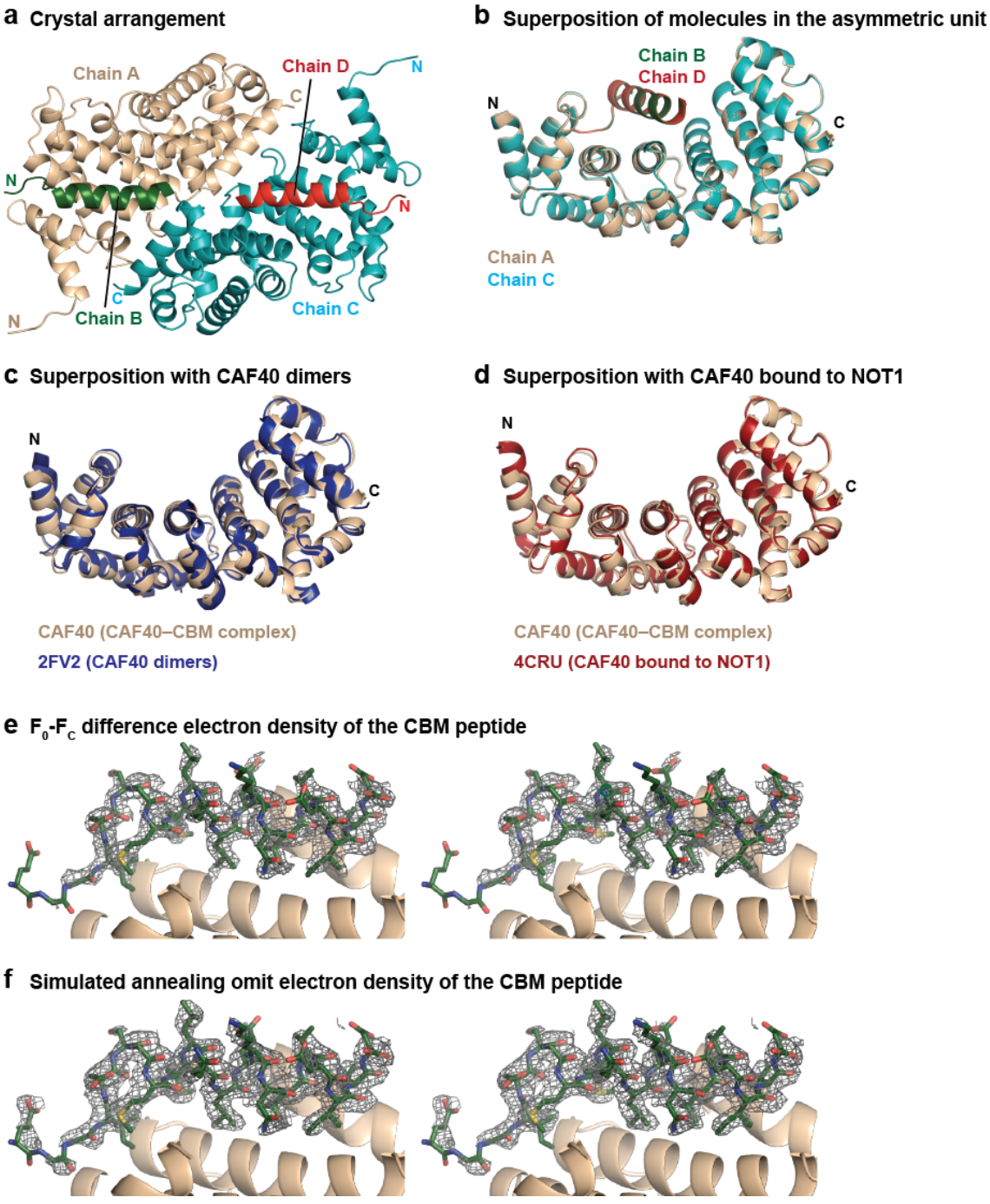
Supplementary Figure 6 | Depletion of CAF40 partially suppresses Roquin activity in HeLa cells. A tethering assay using the R-Luc-6xMS2bs was performed in HeLa cells depleted of CAF40 or treated with a control shRNA (Control). A plasmid expressing F-Luc served as a transfection control. For each condition, *Renilla* luciferase activity was measured, normalized to that of the F-Luc transfection control and set to 100% in cells expressing MS2-HA. (**a**) Mean values ± s.d. of the normalized R-Luc activities from three independent experiments. (**b**) Western blot analysis showing the equivalent expression of the MS2-HA-tagged proteins used in the tethering assay shown in panel (a). (**c**) Western blot showing the efficiency of the CAF40 knockdown. Dilutions of control cell lysates were loaded in lanes 1–4 to estimate the efficacy of the depletion. Tubulin served as a loading control.

Supplementary Figure 7



Supplementary Figure 7 | Structure-based sequence alignment of CAF40. Secondary structure elements as determined by the *Hs* CAF40-*Dm* Roquin structure are shown above the alignment. Residues conserved in all aligned species are highlighted by a brown background, and residues that show conservation of at least 70% are shown with an orange background. Conservation scores were calculated using the SCORECONS webserver[1]. Residues that directly contact the *Dm* Roquin CBM peptide are marked by green diamonds. The residues that were mutated in this study are indicated by red asterisks. Species abbreviations are as follows: *Hs*, *Homo sapiens*; *Dm*, *Drosophila melanogaster*; *Dr*, *Danio rerio*; *Ce*, *Caenorhabditis elegans*; *Sc*, *Saccharomyces cerevisiae*; *At*, *Arabidopsis thaliana*.

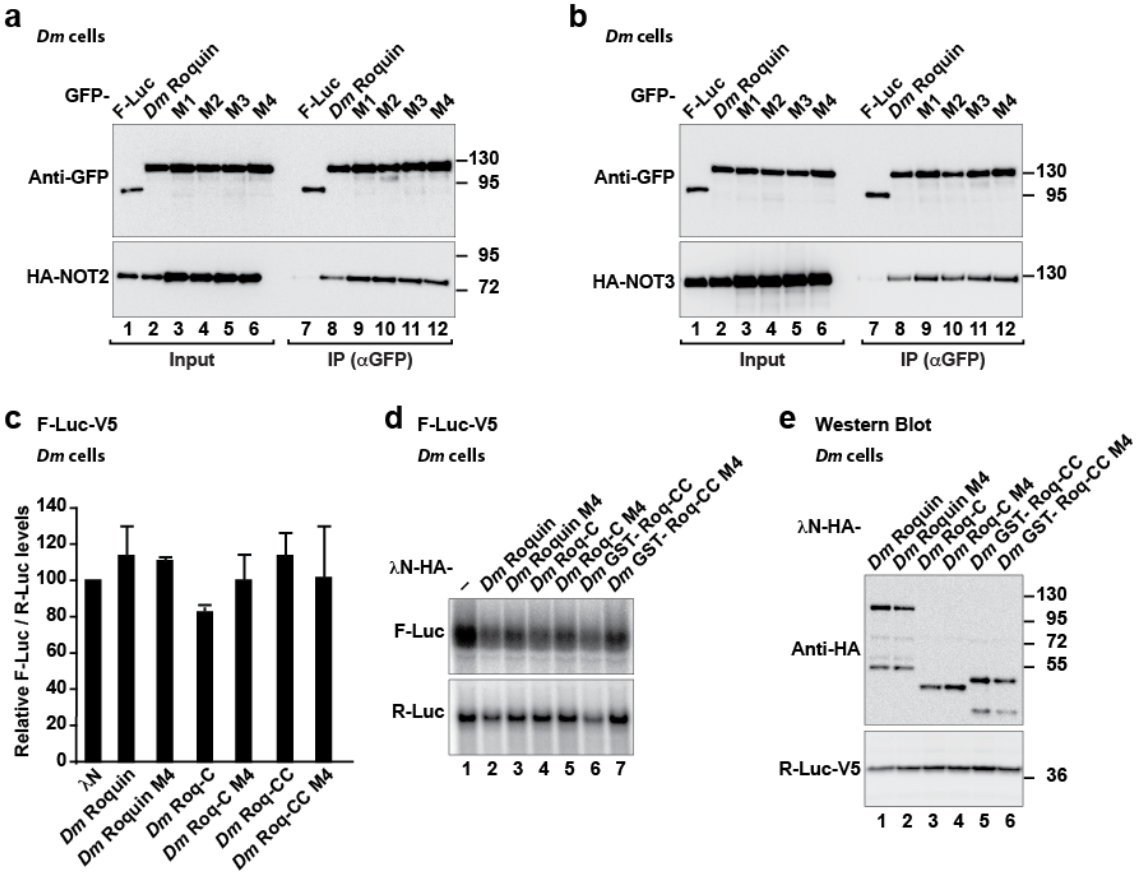
Supplementary Figure 8



Supplementary Figure 8 | Structure of CAF40 bound to the *Dm* Roquin CBM peptide. (**a**) Crystal packing of the CAF40-Roquin CBM complex. CAF40 (chain A) and *Dm* Roquin CBM (chain B) from complex 1 are colored in light brown and green, respectively, whereas the molecules from complex 2 are colored in cyan (CAF40, chain

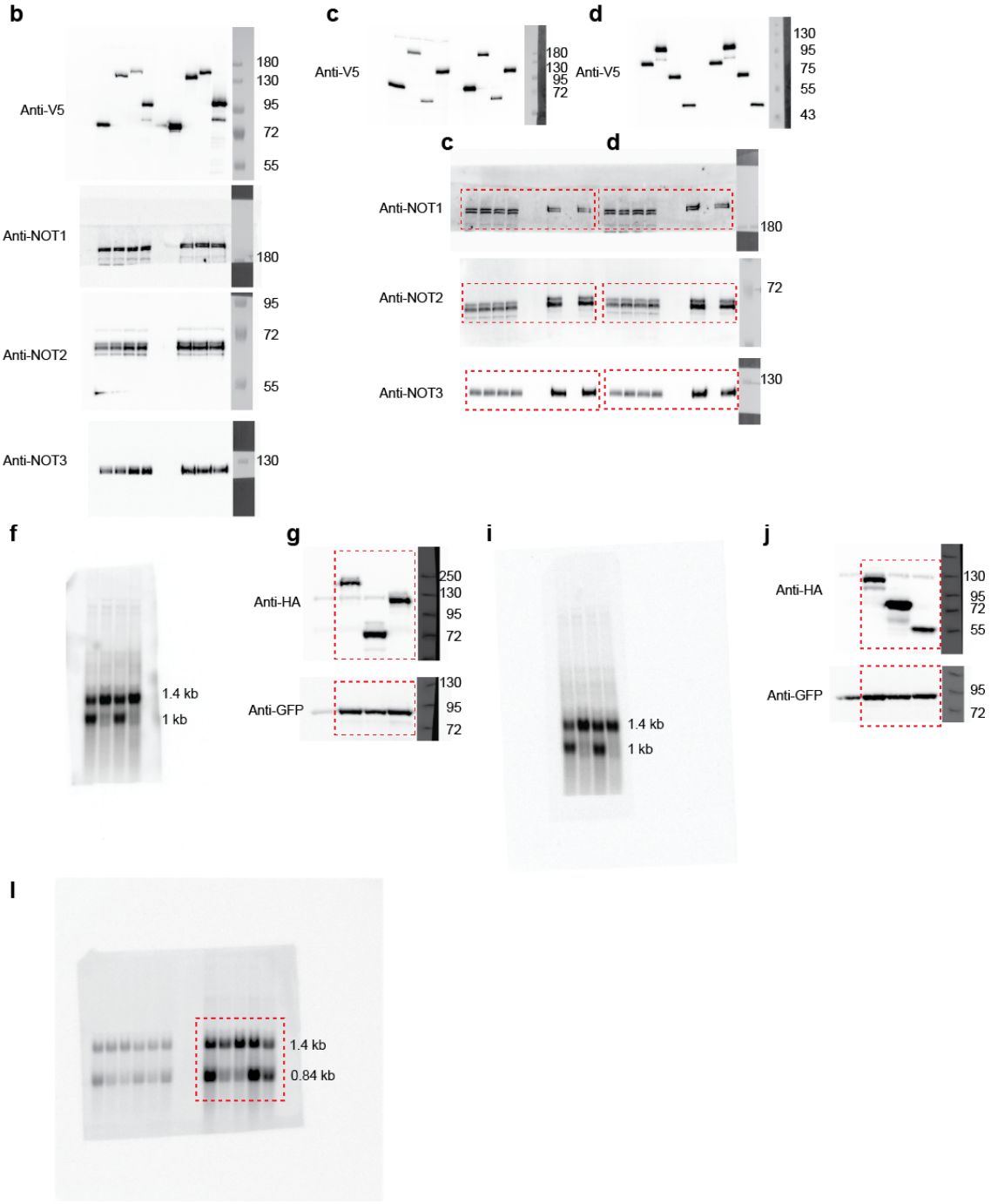
C) and red (CBM, chain D). (**b**) Superposition of the two CAF40-Roquin complexes from the asymmetric unit of the crystal. The colors are as described in (a). The structures superpose with an RMSD of 0.24 Å over 254 Cα atoms. (**c**) Superposition of the structure of the CAF40 ARM domain bound to the CBM (light brown, this study) with the structure of the isolated CAF40 ARM domain (blue, PDB 2FV2; ref. 37). (**d**) Superposition of the structure of the CAF40 ARM domain bound to the CBM (light brown, this study) with that of the CAF40 ARM domain bound to the NOT1 CN9BD (red, PDB 4CRU; ref. 2). (**e**) Stereo view showing the F_o-F_c difference electron density for the CBM peptide. The Roquin CBM peptide is shown in stick representation bound to CAF40 in cartoon representation. The difference electron density of the CBM peptide is shown as a grey mesh. The F_o-F_c-type map, which is contoured at 2.0 σ, was calculated using a refined CAF40 model before the CBM peptide was modelled. (**f**) Stereo view showing the $2F_o$-F_c simulated annealing composite omit map surrounding the CBM peptide. This $2F_o$-F_c map (which is largely devoid of model bias[2]), is shown around the CBM and contoured at 1.0 σ. It was generated with Phenix.Autobuild[3] using the final refined CAF40-Roquin model.

Supplementary Figure 9



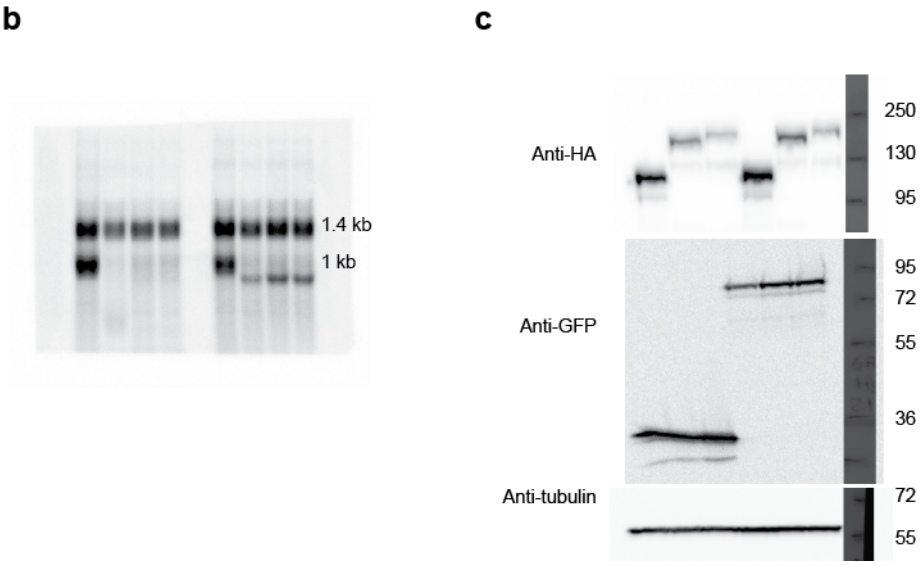
Supplementary Figure 9 | The CBM of *Dm* Roquin contributes to the recruitment of the CCR4-NOT complex. (**a,b**) Interaction of GFP-tagged *Dm* Roquin (wild-type and the M1, M2, M3 and M4 mutants; see Supplementary Table 1) with HA-tagged NOT2 and NOT3 in *Dm* S2 cells. F-Luc-GFP served as negative control. (**c,d**) A tethering assay using the F-Luc-V5 reporter lacking λN binding sites and the indicated λN-HA-tagged *Dm* Roquin mutants was performed in *Dm* S2 cells as described in Fig. 3a,b. The corresponding experiment using the F-Luc-5BoxB reporter is shown in Fig. 7f,g. Error bars represent s.d. from three independent experiments. (**e**) Western blot showing the equivalent expression of the λN-HA-tagged proteins used in panels (c,d) and Fig. 7f,g.

Supplementary Figure 10



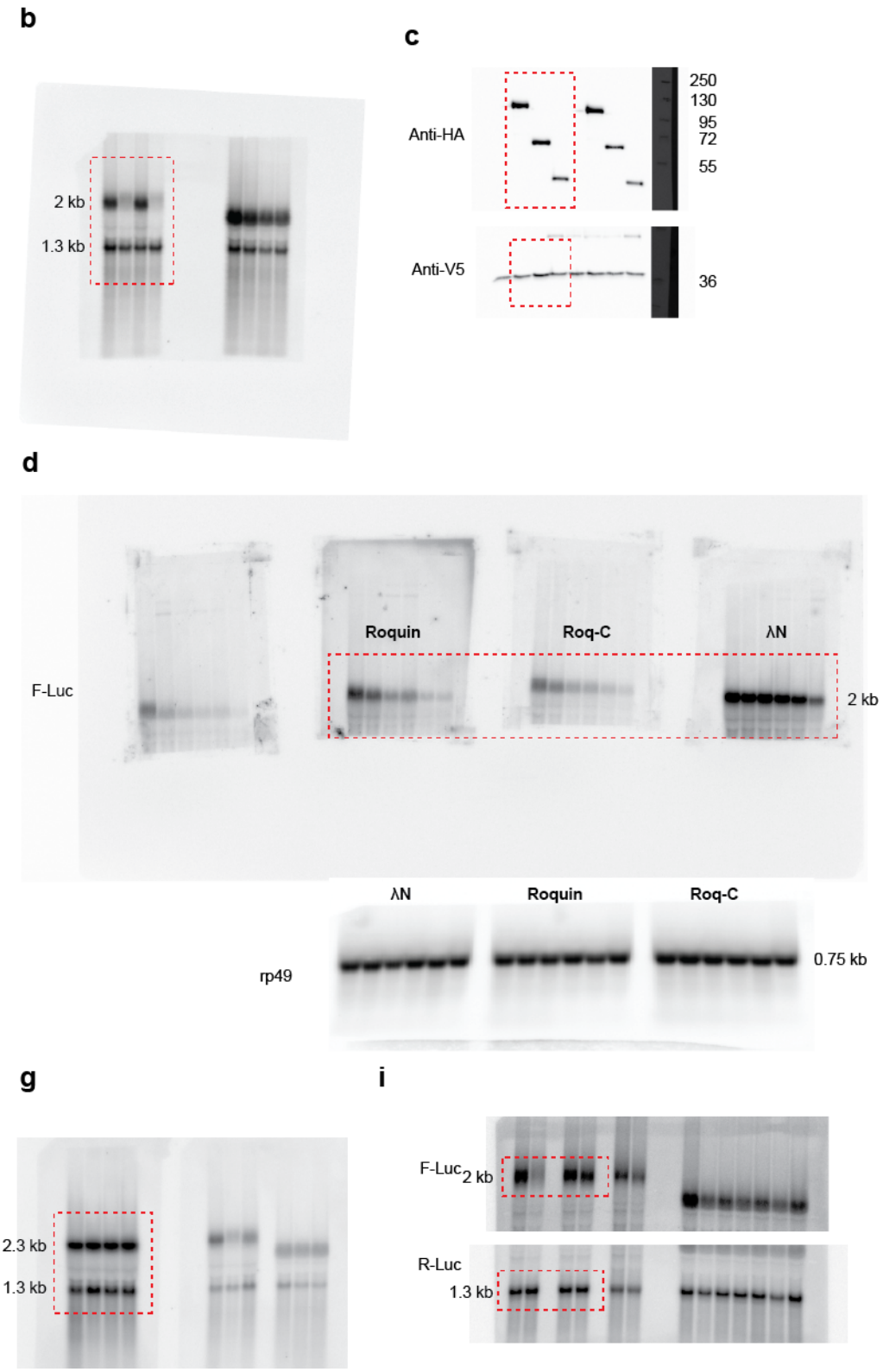
Supplementary Figure 10 | Original images of western and northern blots used in the corresponding panels in Fig. 1. The estimated sizes of the mRNA reporters without poly(A) and protein size markers (kDa) are shown on the right of the Northern and Western blots, respectively.

Supplementary Figure 11



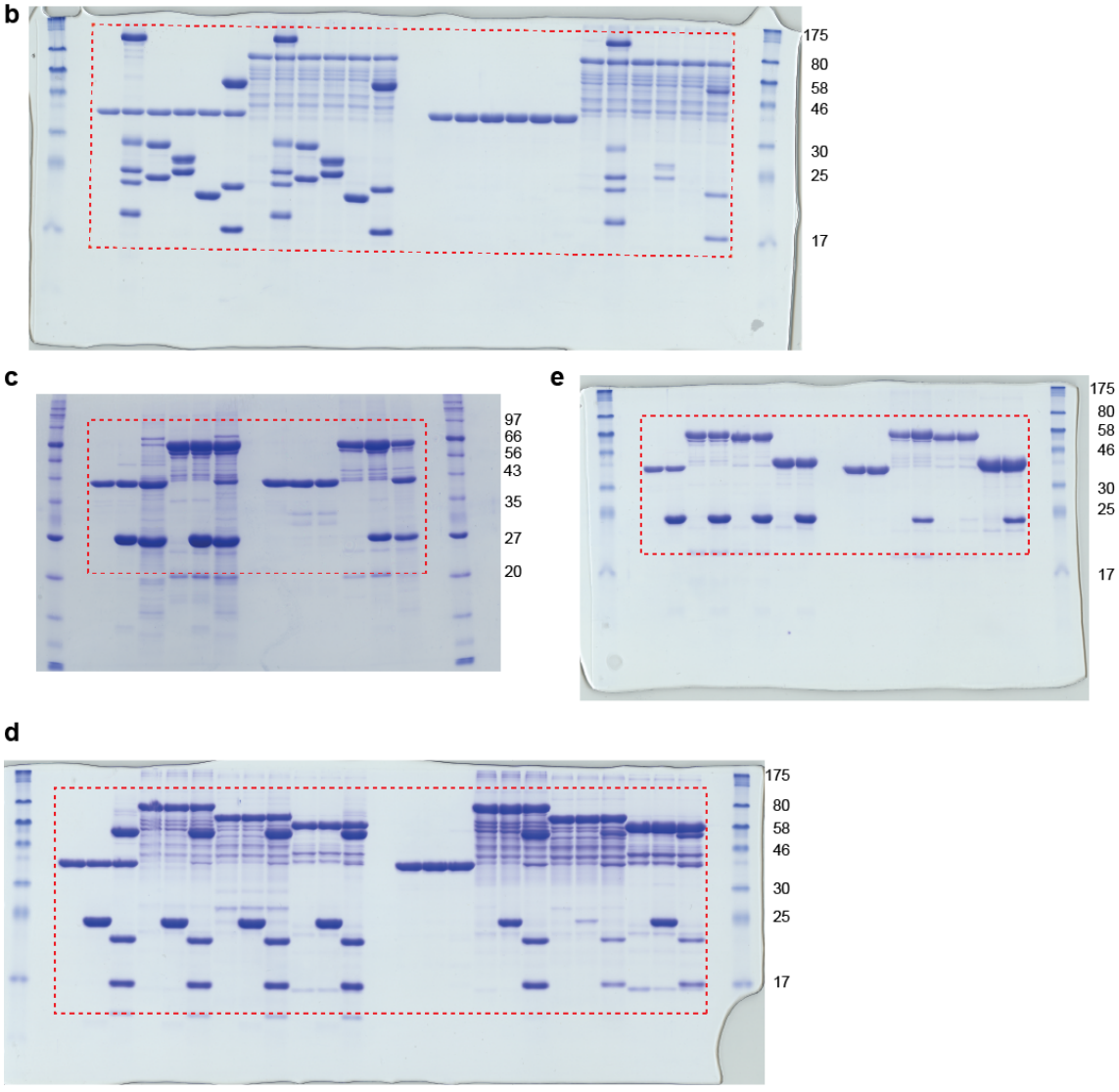
Supplementary Figure 11 | Original images of western and northern blots used in the corresponding panels in Fig. 2. The estimated sizes of the mRNA reporters without poly(A) and protein size markers (kDa) are shown on the right of the Northern and Western blots, respectively.

Supplementary Figure 12



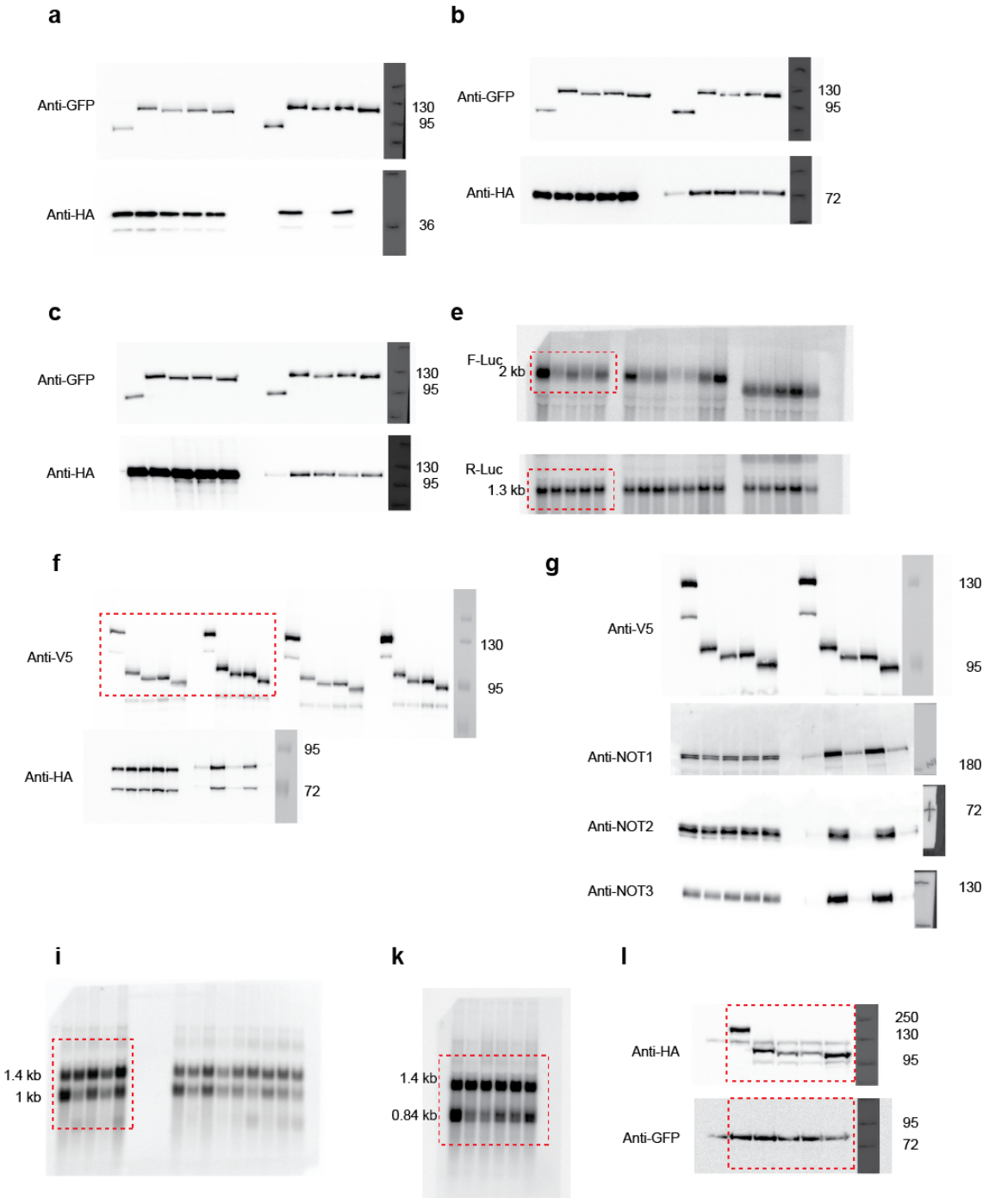
Supplementary Figure 12 | Original images of western and northern blots used in the corresponding panels in Fig. 3. The estimated sizes of the mRNA reporters without poly(A) and protein size markers (kDa) are indicated.

Supplementary Figure 13



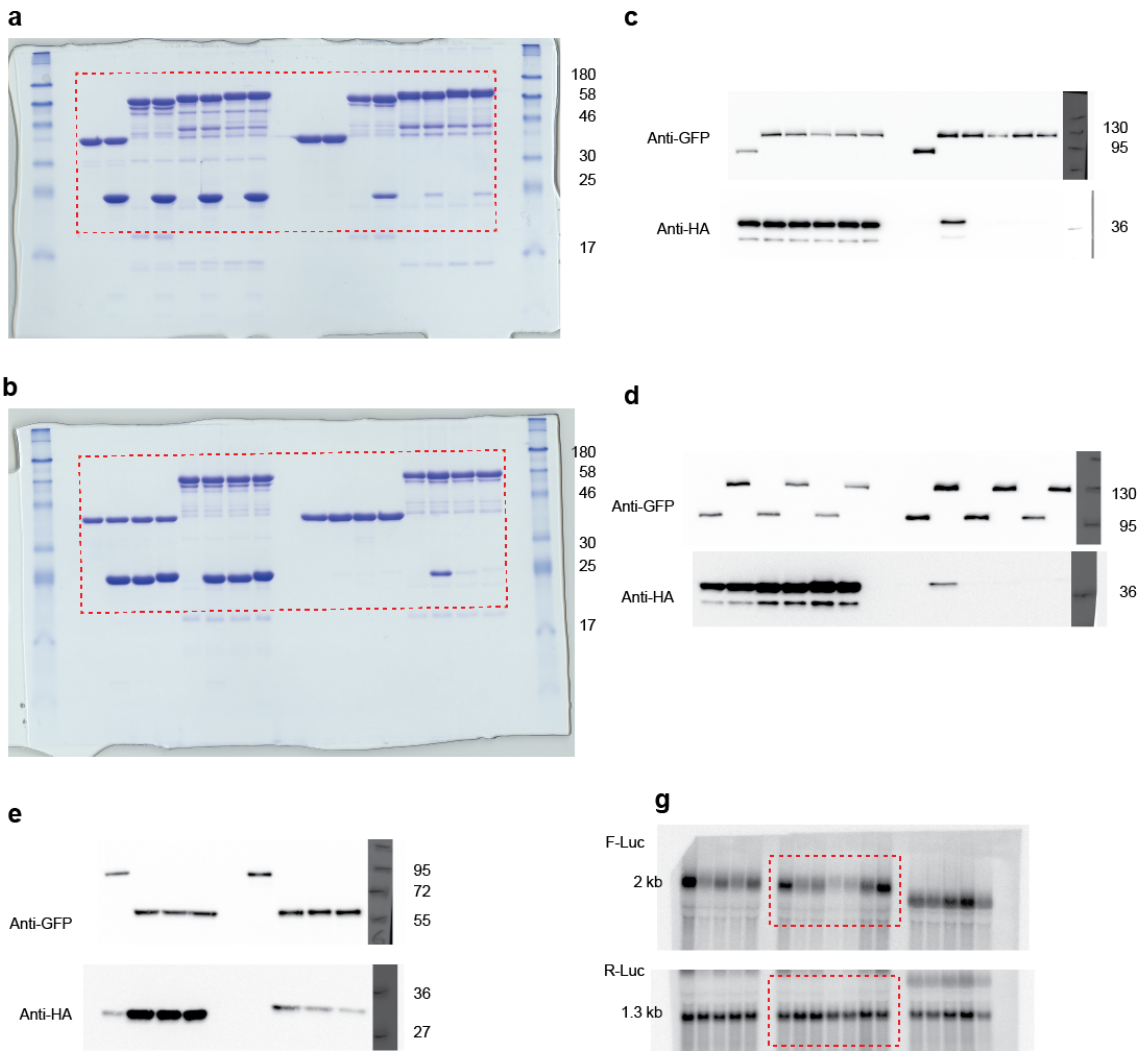
Supplementary Figure 13 | Original images of protein gels used in the corresponding panels in Fig. 4. Protein size markers (kDa) are shown on the right of the panels.

Supplementary Figure 14



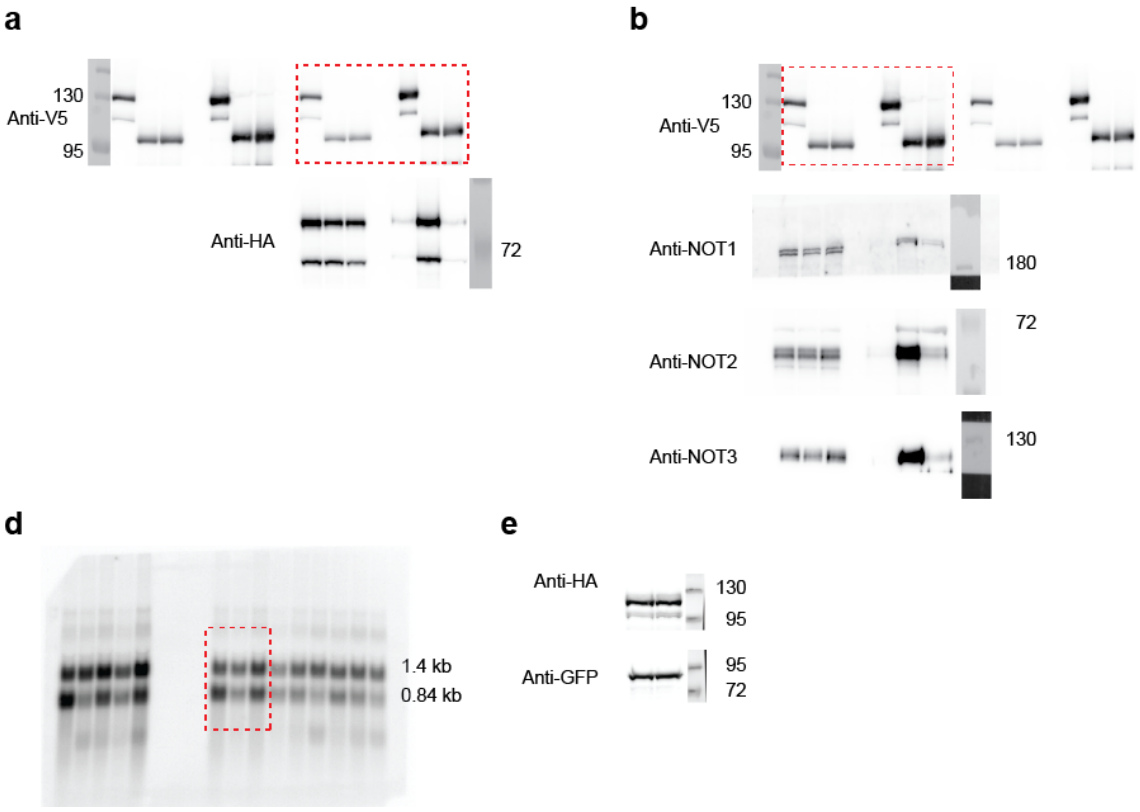
Supplementary Figure 14 | Original images of western and northern blots used in the corresponding panels in Fig. 5. The estimated sizes of the mRNA reporters without poly(A) and protein size markers (kDa) are indicated.

Supplementary Figure 15



Supplementary Figure 15 | Original images of protein gels, western and northern blots used in the corresponding panels in Fig. 7. The estimated sizes of the mRNA reporters without poly(A) and protein size markers (kDa) are indicated

Supplementary Figure 16



Supplementary Figure 16 | Original images of western and northern blots used in the corresponding panels in Fig. 8. The estimated sizes of the mRNA reporters without poly(A) and protein size markers (kDa) are indicated.

Supplementary Table 1. Constructs and mutants used in this study.

***Dm* Roquin (1-819)** (Uniprot Q9VV48)		**Comment**
Roquin	λN-HA-Roquin 1-819	
	GFP-Roquin 1-819	
	MS2-HA-Roquin 1-819	
	V5-SBP-Roquin 1-819	
Roq-N	λN-HA-Roquin 1-500	
	MS2-HA Roquin 1-500	
	V5-SBP-Roquin 1-500	
Roq-C	λN-HA-Roquin 501-819	
	MS2-HA-Roquin 501-819	
	V5-SBP-Roquin 501-819	
	MBP-Roquin 501-819-GB1-6xHis	
Roq-CN	λN-HA-GST-Roquin 501-702	
	λN-HA-GFP-Roquin 501-702	
	MBP-Roquin 501-702-GB1-6xHis	
Roq-CC	λN-HA-GST-Roquin 702-819	
	λN-HA-GFP-Roquin 702-819	
	MBP-Roquin 702-819-GB1-6xHis	
NBM	MBP-Roquin 725-755	NOT module-binding region
CBM	MBP-Roquin 790-812	CAF40-binding region
ΔNBM	MS2-HA-Roquin Δ725-755	NOT module-binding region
	GFP-Roquin Δ725-755	NOT module-binding region
	V5-SBP-Roquin Δ725-755	NOT module-binding region
Roq-C-ΔNBM	λN-HA-Roquin 501-819 Δ725-755	NOT module-binding region
Roq-CC-ΔNBM	MBP-Roquin 702-819 Δ725-755-GB1-6xHis	NOT module-binding region
ΔCBM	MS2-HA-Roquin Δ790-812	CAF40-binding region
	GFP-Roquin Δ790-812	CAF40-binding region
	V5-SBP-Roquin Δ790-812	CAF40-binding region
Roq-C-ΔCBM	λN-HA-Roquin 501-819 Δ790-812	CAF40-binding region
Roq-CC-ΔCBM	MBP-Roquin 702-819 Δ790-812-GB1-6xHis	CAF40-binding region
ΔCBM+NBM	MS2-HA-Roquin Δ725-755-Δ790-812	CAF40-binding region, NOT module-binding region
	GFP-Roquin Δ725-755-Δ790-812	CAF40-binding region, NOT module-binding region
	V5-SBP-Roquin Δ725-755-Δ790-812	CAF40-binding region, NOT module-binding region
Roq-C-ΔCBM+NBM	λN-HA-Roquin 702-819 Δ725-755-Δ790-812	CAF40-binding region, NOT module-binding region
Roq-CC ΔCBM+NBM	MBP-Roquin 702-819 Δ725-755-Δ790-812-GB1-6xHis	CAF40-binding region, NOT module-binding region
M1	GFP-Roquin I793E-L801E	CAF40-binding region
M2	GFP-Roquin L805E-V809E	CAF40-binding region
Roq-CC M2	MBP-Roquin 702-819 L805E-V809E-GB1-6xHis	CAF40-binding region
M3	GFP-Roquin L801E-L805E	CAF40-binding region

M4	GFP-Roquin I793E-L801E-L805E-V809E	CAF40-binding region
	λN-HA-Roquin I793E-L801E-L805E-V809E	CAF40-binding region
	MS2-HA-Roquin I793E-L801E-L805E-V809E	CAF40-binding region
	V5-SBP-Roquin I793E-L801E-L805E-V809E	CAF40-binding region
Roq-C M4	λN-HA-Roquin 501-819 I793E-L801E-L805E-V809E	CAF40-binding region
Roq-CC M4	λN-HA-GST-Roquin 702-819 I793E-L801E-L805E-V809E	CAF40-binding region
	MBP-Roquin 702-819 I793E-L801E-L805E-V809E-GB1-6xHis	CAF40-binding region
***Hs* Roquin1 (1-1133)** (Uniprot Q5TC82)		
Roquin1	MS2-HA-Roquin1 1-1133	
	V5-SBP-Roquin1 1-1133	
Roq1-N	MS2-HA-Roquin1 1-500	
	V5-SBP-Roquin1 1-500	
Roq1-C	MS2-HA-Roquin1 501-1133	
	V5-SBP-Roquin1 501-1133	
	MBP-Roq1 501-1133	
***Hs* Roquin2 (1-1191)** (Uniprot Q9HBD1)		
Roquin2	MS2-HA-Roquin2 1-1191	
	V5-SBP-Roquin2 1-1191	
Roq2-N	MS2-HA-Roquin2 1-500	
	V5-SBP-Roquin2 1-500	
Roq2-C	MS2-HA-Roquin2 501-1191	
	V5-SBP-Roquin2 501-1191	
***Hs* CNOT1** (Uniprot A5YKK6)		
CNOT1-MID+C	MBP-Hs CNOT1 1093-2371	Includes MIF4G, CN9BD, CD and SHD
CNOT1-MIF4G	6xHis-Hs CNOT1 1093-1317	MIF4G-like domain
CNOT1-CN9BD	MBP-Hs CNOT1 1351-1588	CNOT9-binding domain
CNOT1-CD	MBP-Hs CNOT1 1607-1815	Connector domain
CNOT1-SHD	MBP-Hs CNOT1 1833-2361	Superfamily homology domain
***Hs* CNOT2** (Uniprot Q9NZN8)		
CNOT2-C	MBP-Hs CNOT2 350-540	
***Hs* CNOT3** (Uniprot O75175)		
CNOT3-C	6xHis-Hs CNOT3 607-748	
***Hs* CNOT7** (Uniprot Q9UIV1)		
CNOT7	MBP-Hs CNOT7	

***Hs* CNOT9** (Uniprot Q92600)		
CNOT9-ARM	6xHis-HsCNOT9 19-285	
CNOT9-ARM V181E	6xHis-HsCNOT9 19-285-V181E	Disrupts Roquin binding
CNOT9-ARM 2xMut	6xHis-HsCNOT9 19-285-Y134D-G141W	Disrupts Roquin binding
***Dm* NOT9 (1-304)** (Uniprot Q7JVP2)		
CAF40	λN-HA-CAF40 1-304	
	GFP-CAF40 1-304	
CAF40 V186E	GFP-CAF40 V186E	Roquin-binding region
CAF40 2xMut	GFP-CAF40 Y139D-G146E	Roquin-binding region
CAF40-ARM	MBP-CAF40 25-291	Roquin-binding region
***Dm* NOT1** (Uniprot A8DY81)		
NOT1	λN-HA-NOT1	
CN9BD	λN-HA-CN9BD 1468-1717	CAF40-binding domain

Supplementary Table 2. 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	Monoclonal
Anti-GFP (for immunoprecipitations)	In house			Rabbit Polyclonal
Anti-*Dm* NOT1	Kind gift from E. Wahle	T6199	1:1,000	Rabbit Polyclonal
Anti-tubulin	Sigma-Aldrich	T6199	1:10,000	Monoclonal
Anti-V5	AbD Serotec	MCA1360GA	1:5,000	Monoclonal
Anti-mouse-HRP	GE Healthcare	NA931V	1:10,000	Monoclonal
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

Supplementary References

1. Valdar, W. S. J. Scoring residue conservation. *Proteins Struct. Funct. Bioinforma.* **48,** 227–241 (2002).
2. Terwilliger T. C. *et al.* Iterative-build OMIT maps: map improvement by iterative model building and refinement without model bias. *Acta Crystallogr. D Biol. Crystallogr.* **64,** 515–524 (2008).
3. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68,** 352–367 (2012).

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).

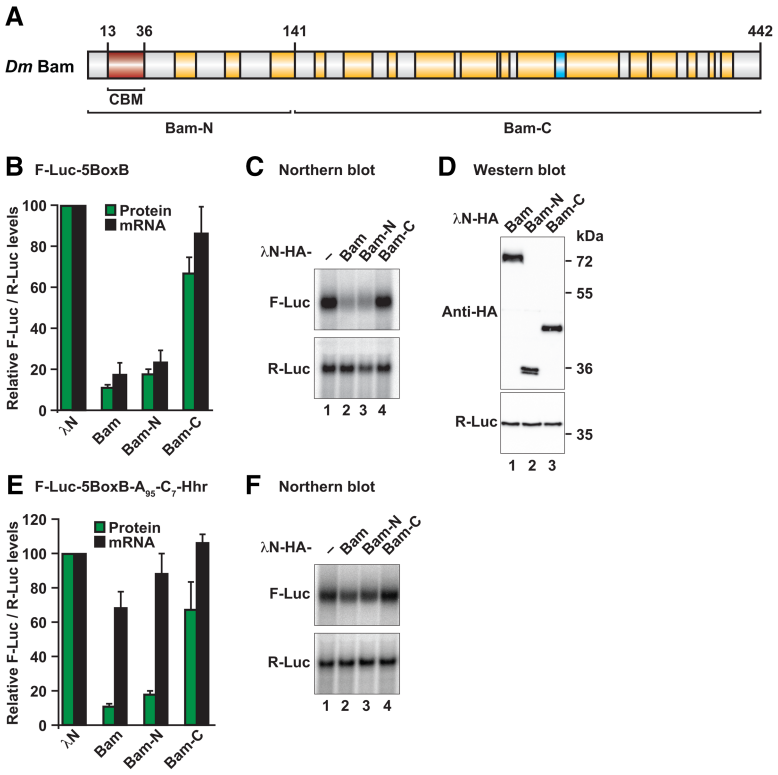


FIGURE 1. Bam induces degradation of bound mRNAs through its N-terminal region. (*A*) Bam consists of several predicted α-helices (shown in yellow) and a β-strand (shown in cyan). The position of the CAF40-binding motif (CBM, in red) as well as the boundaries of the Bam-N and Bam-C fragments are indicated. (*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-

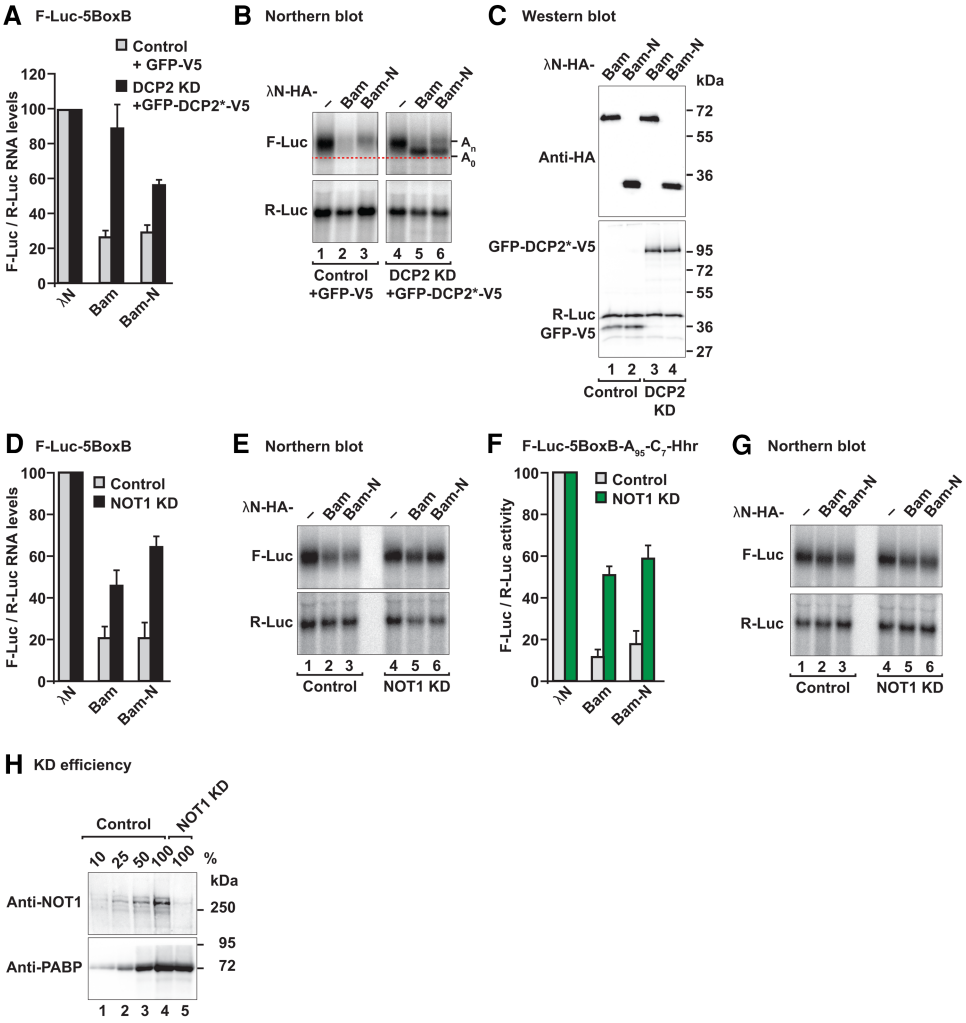


FIGURE 2. Bam degrades mRNAs through the 5′-to-3′ mRNA decay pathway. (*A*) Tethering assay using the F-Luc-5BoxB reporter was performed in control S2 cells or cells depleted of the decapping enzyme DCP2 (DCP2 KD). The transfection mixture included plasmids expressing either GFP-V5 or a GFP-tagged catalytically 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

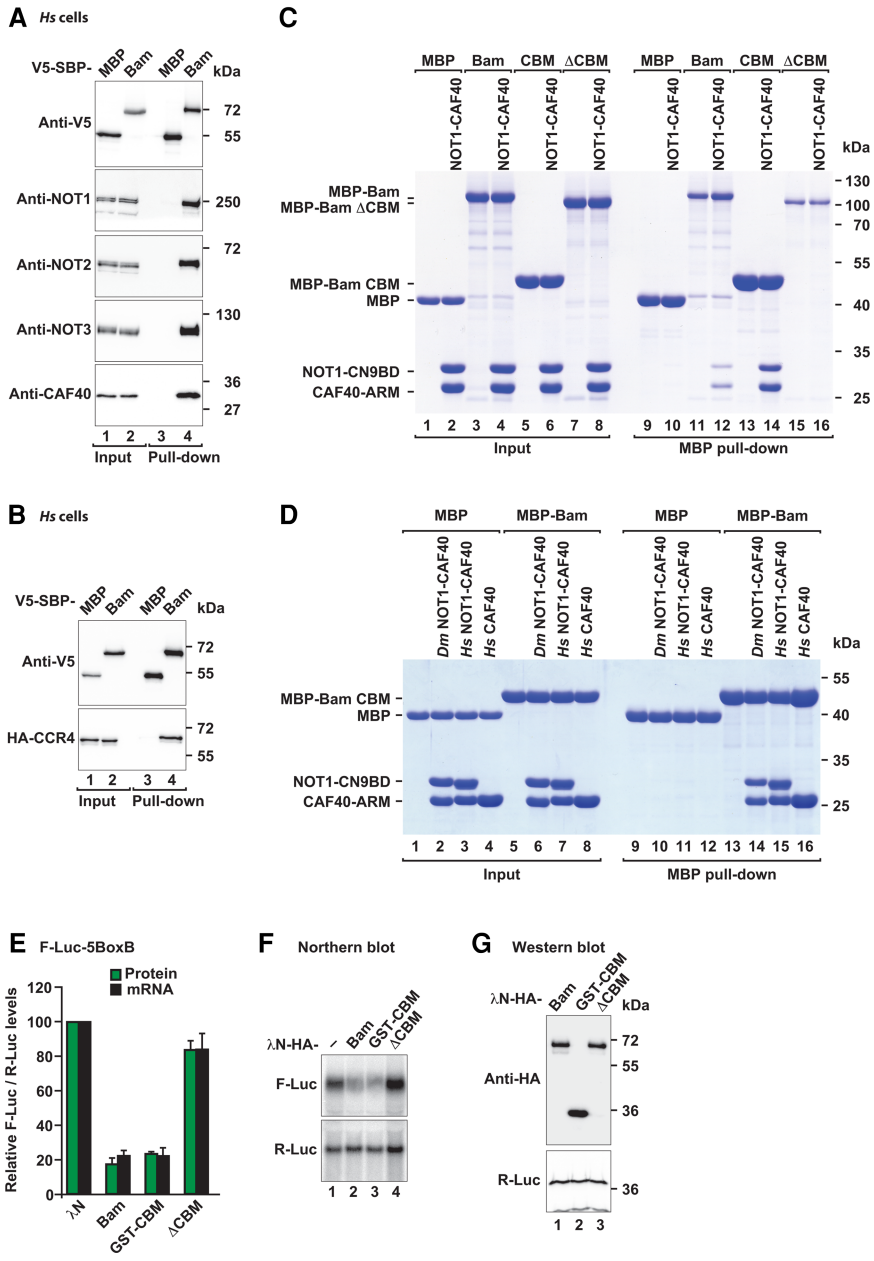


FIGURE 3. Bam binds directly to CAF40 by using an N-terminal CAF40-binding motif (CBM). (*A*) SBP pull-down assay in HEK293T cells expressing V5-SBP-tagged full-length Bam. V5-SBP-tagged MBP served as negative control. Input (1.5% for the V5-SBP tagged proteins and 1% for endogenous CCR4–NOT subunits) 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.

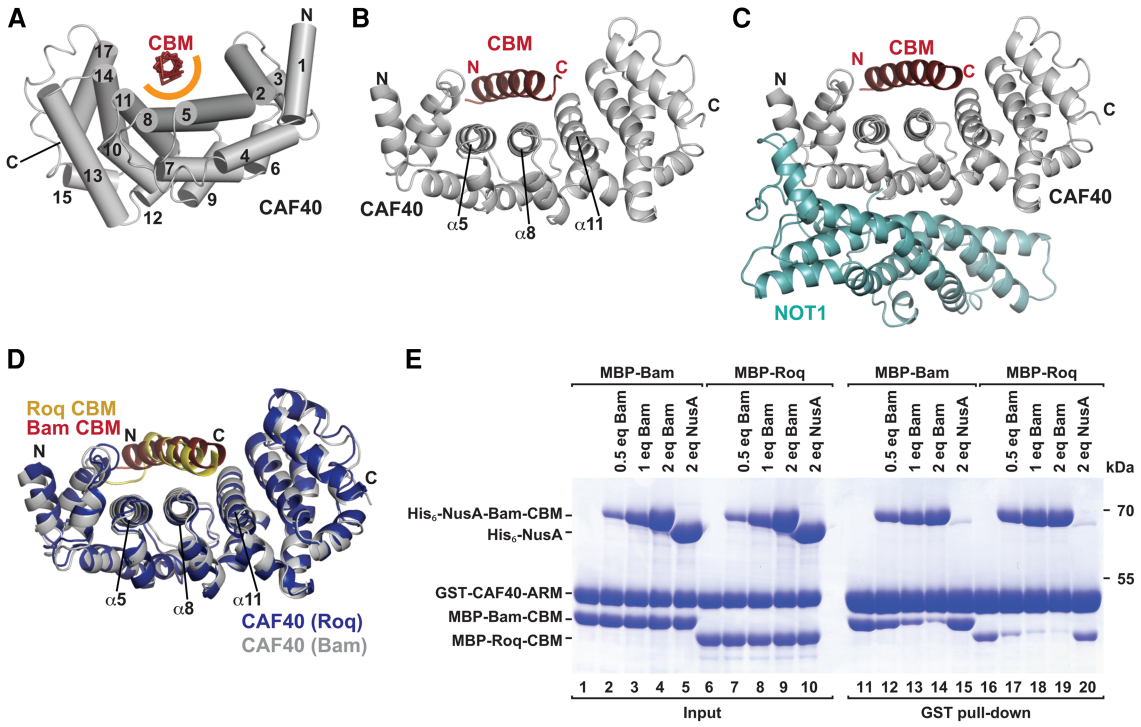


FIGURE 4. Structure of the Bam CBM bound to CAF40 and to the NOT1 CN9BD–CAF40 module. (*A*) The Bam CBM peptide (red, backbone shown in ribbon representation) bound to *Hs* CAF40 (gray). CAF40 helices are depicted as tubes and numbered in black. The orange semicircle marks the predominantly hydrophobic 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.

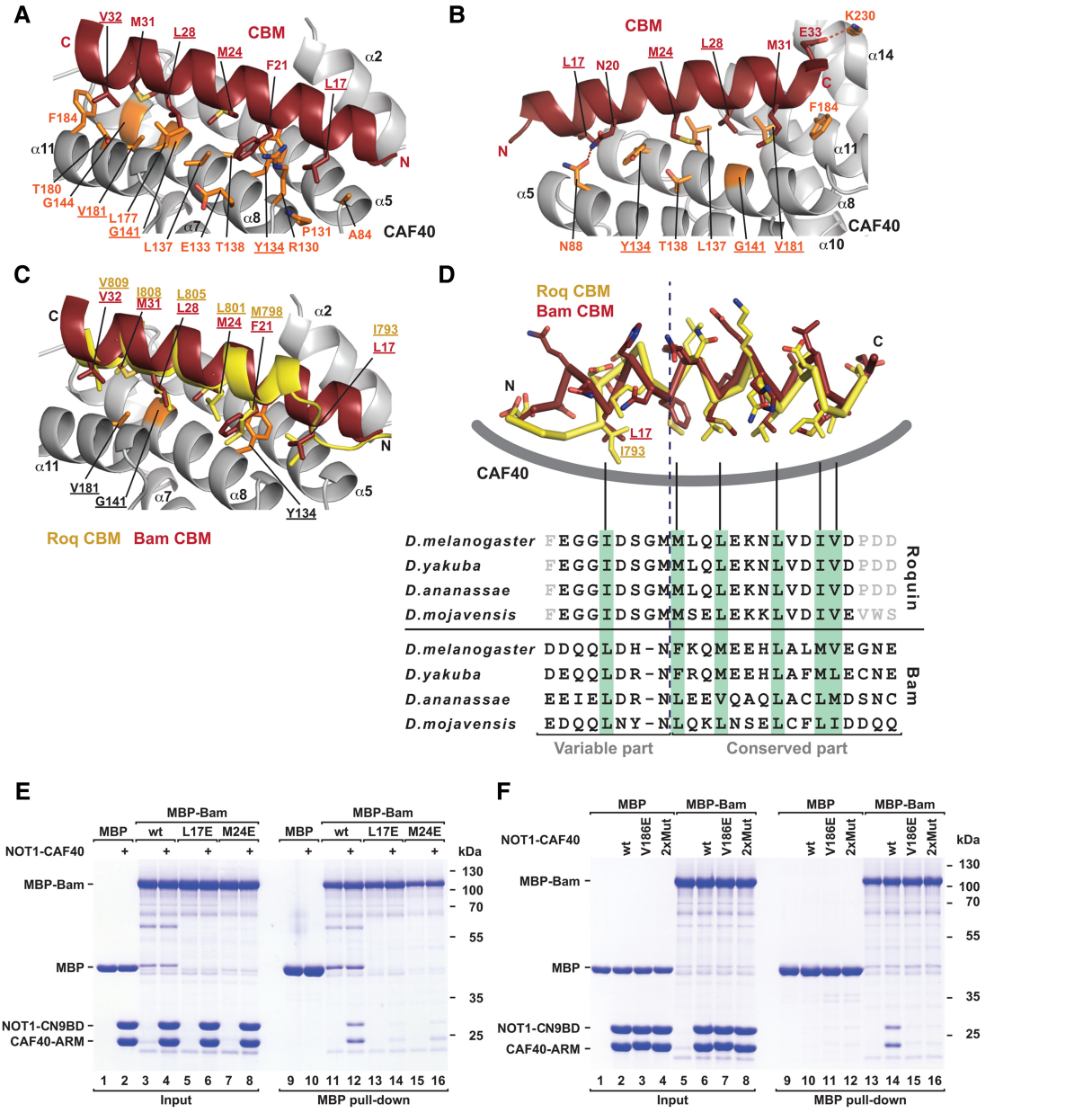


FIGURE 5. The Bam and Roq CBMs use a similar CAF40-binding mode. (*A*,*B*) Close-up views of the CAF40-Bam CBM-binding interface in two orientations. Selected residues of CAF40 and Bam are shown as orange and red sticks, respectively. Hydrogen bonds are indicated by red dashed lines. Residues mutated 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

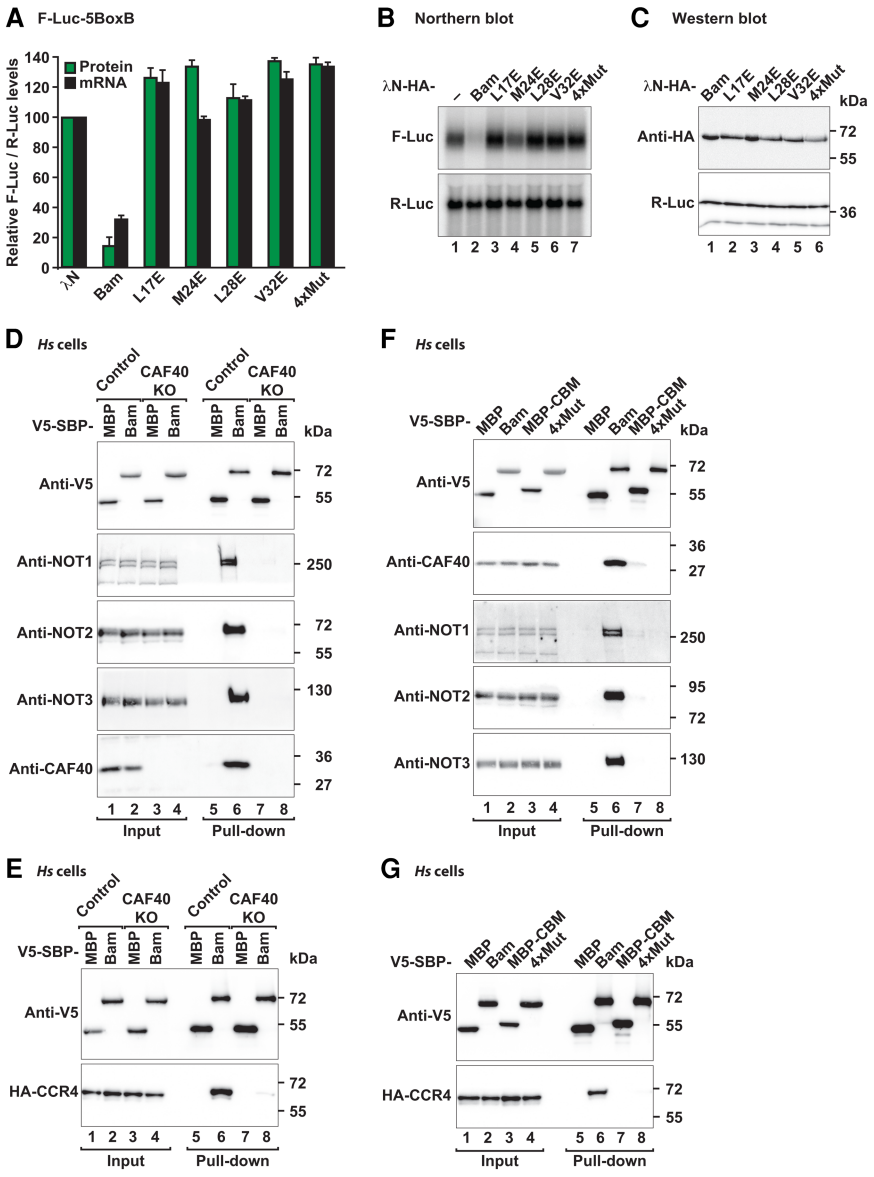


FIGURE 6. The CBM is necessary for Bam-mediated mRNA repression. (*A*) Tethering assay using the F-Luc-5BoxB reporter and λN-HA-tagged Bam (wild-type or the indicated mutants) in S2 cells. The samples were analyzed as described in Figure 1B. (*B*) Northern blot of representative RNA samples shown in *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

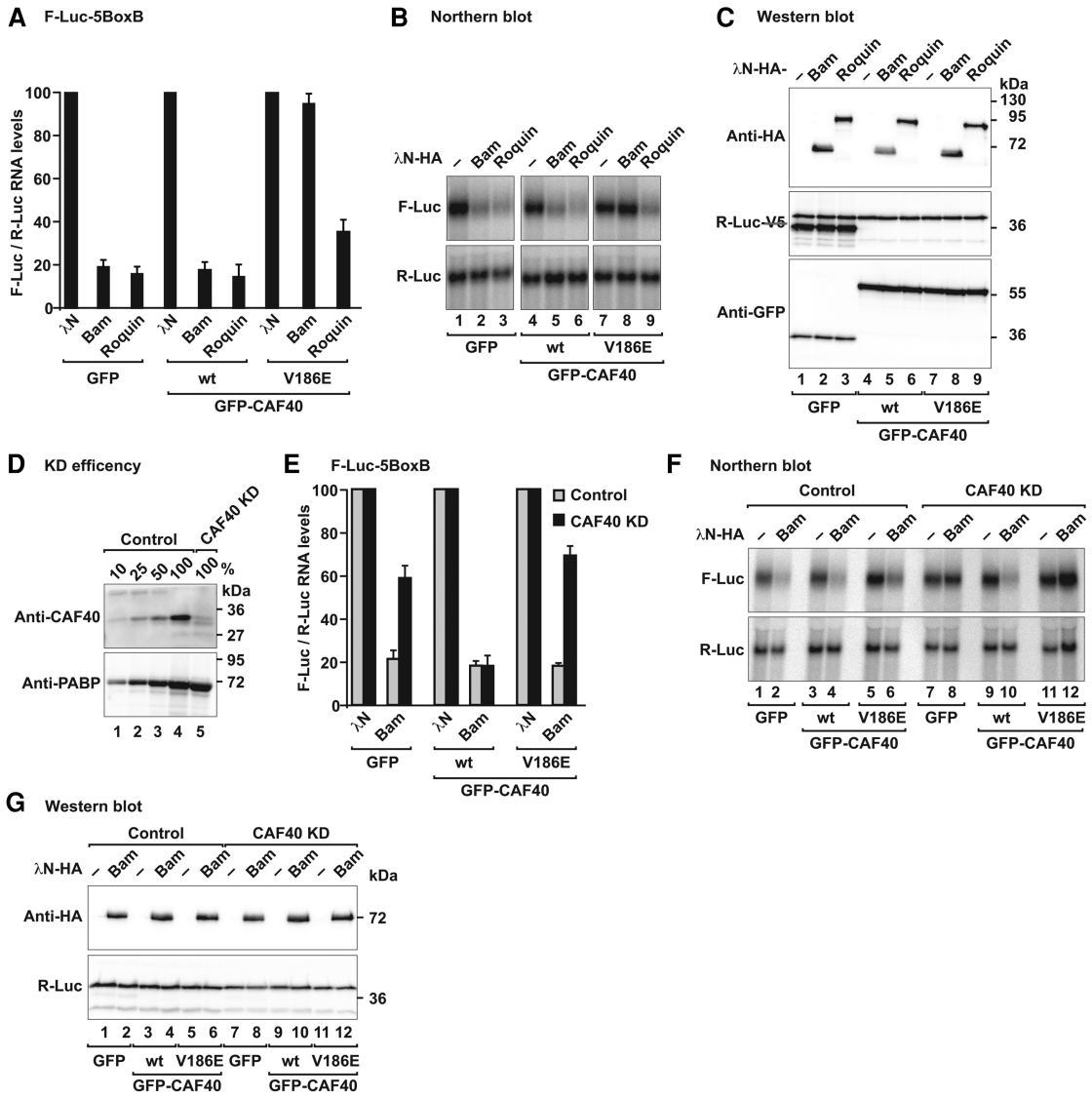


FIGURE 7. Bam depends on CCR4–NOT complex recruitment to induce mRNA decay. (*A*) Tethering assay using the F-Luc-5BoxB reporter and λN-HA-tagged Bam and Roq in S2 cells. The transfection mixtures also contained plasmids for expression of GFP (control) or GFP-CAF40 (wild-type or the V186E mutant) as 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.

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, et al.

RNA 2018 24: 381-395 originally published online December 18, 2017
Access the most recent version at doi:10.1261/rna.064584.117

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

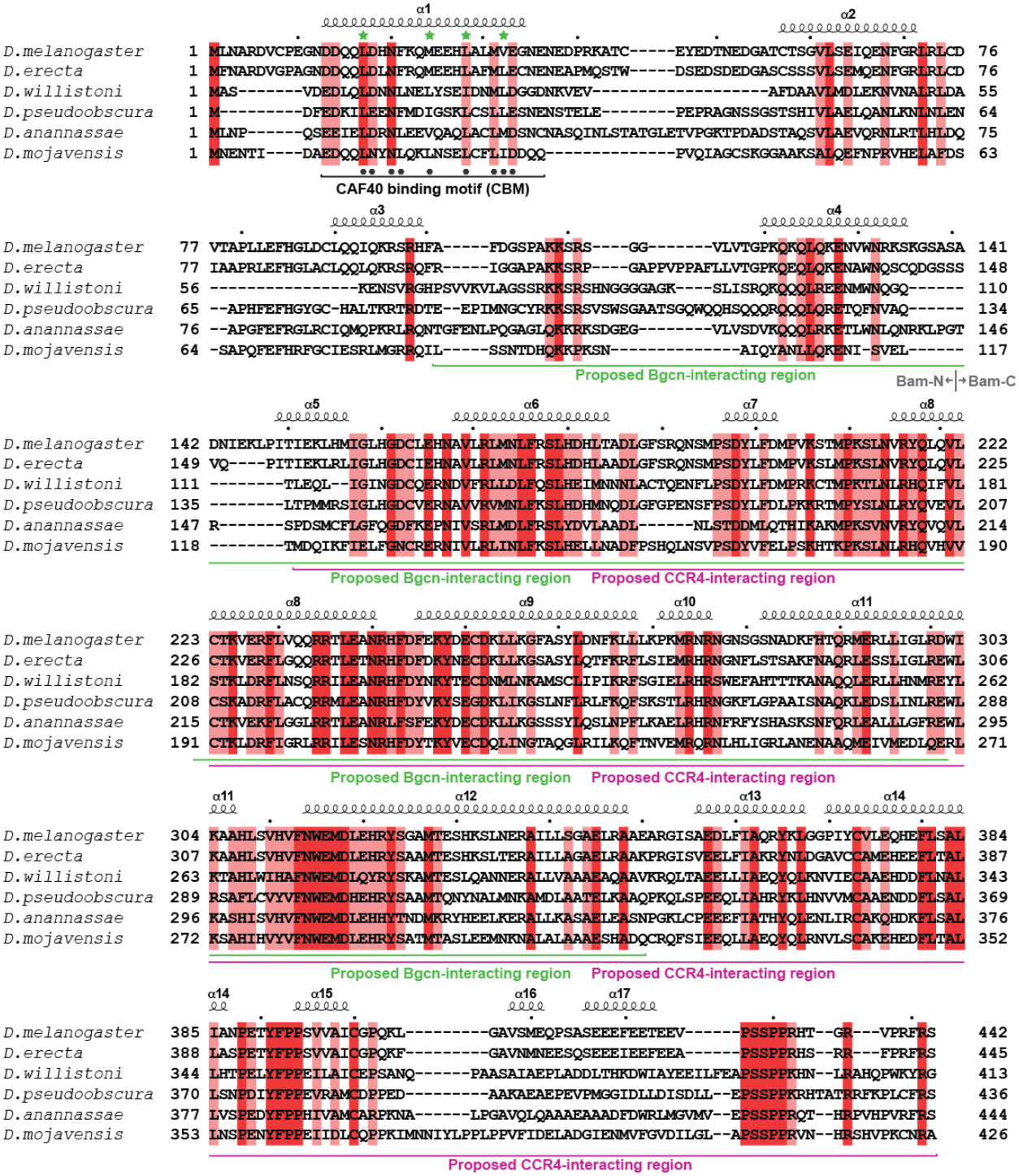
Name	***Dm* NOT1** (Uniprot A8DY81)	**Comment**
NOT1	λN-HA-NOT1	
NOT1-CN9BD	λN-HA-NOT1 1467–1719	CAF40-binding domain
	MBP-NOT1 1468–1719	CAF40-binding domain

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

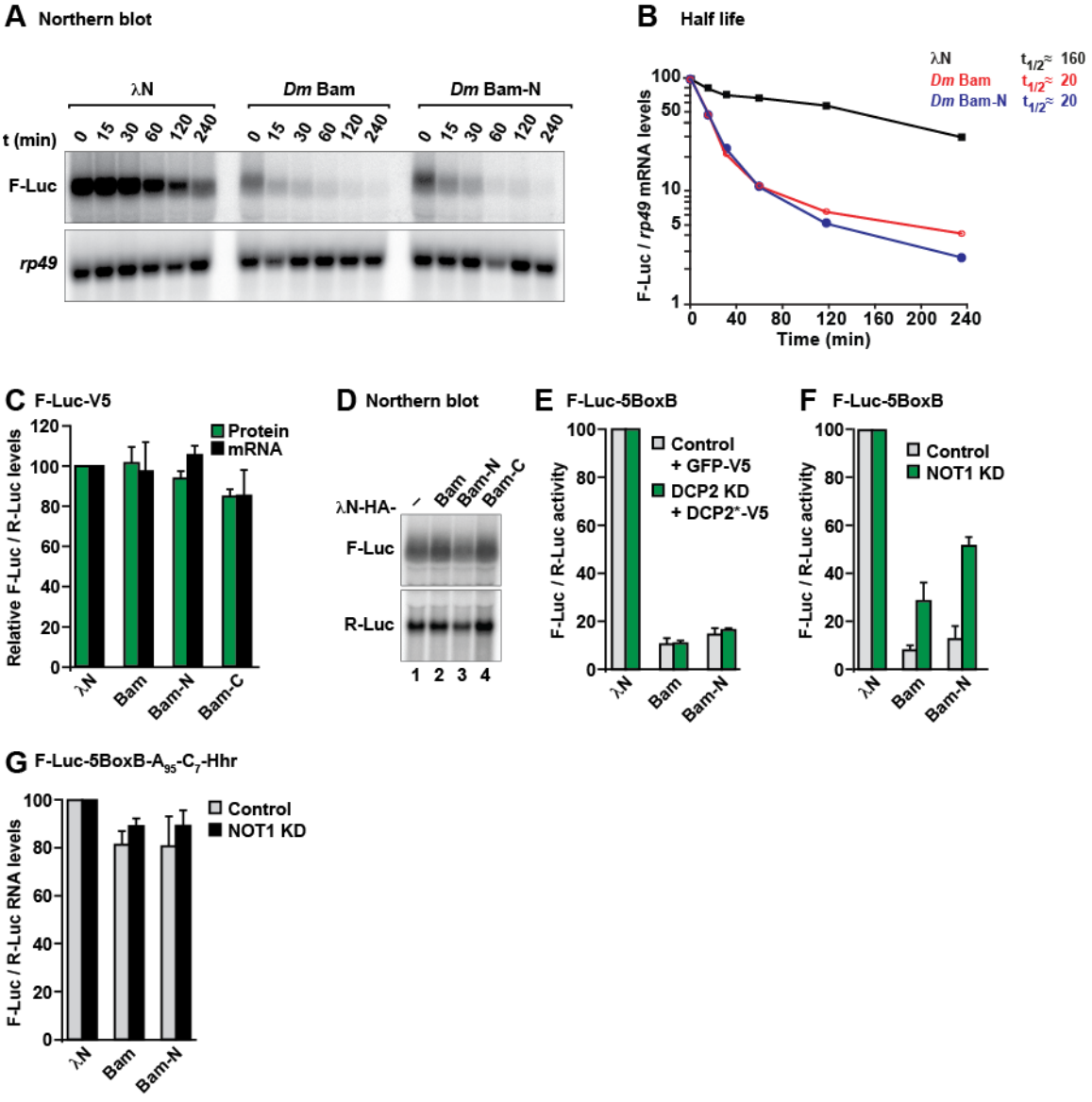
SUPPLEMENTAL FIGURES

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.

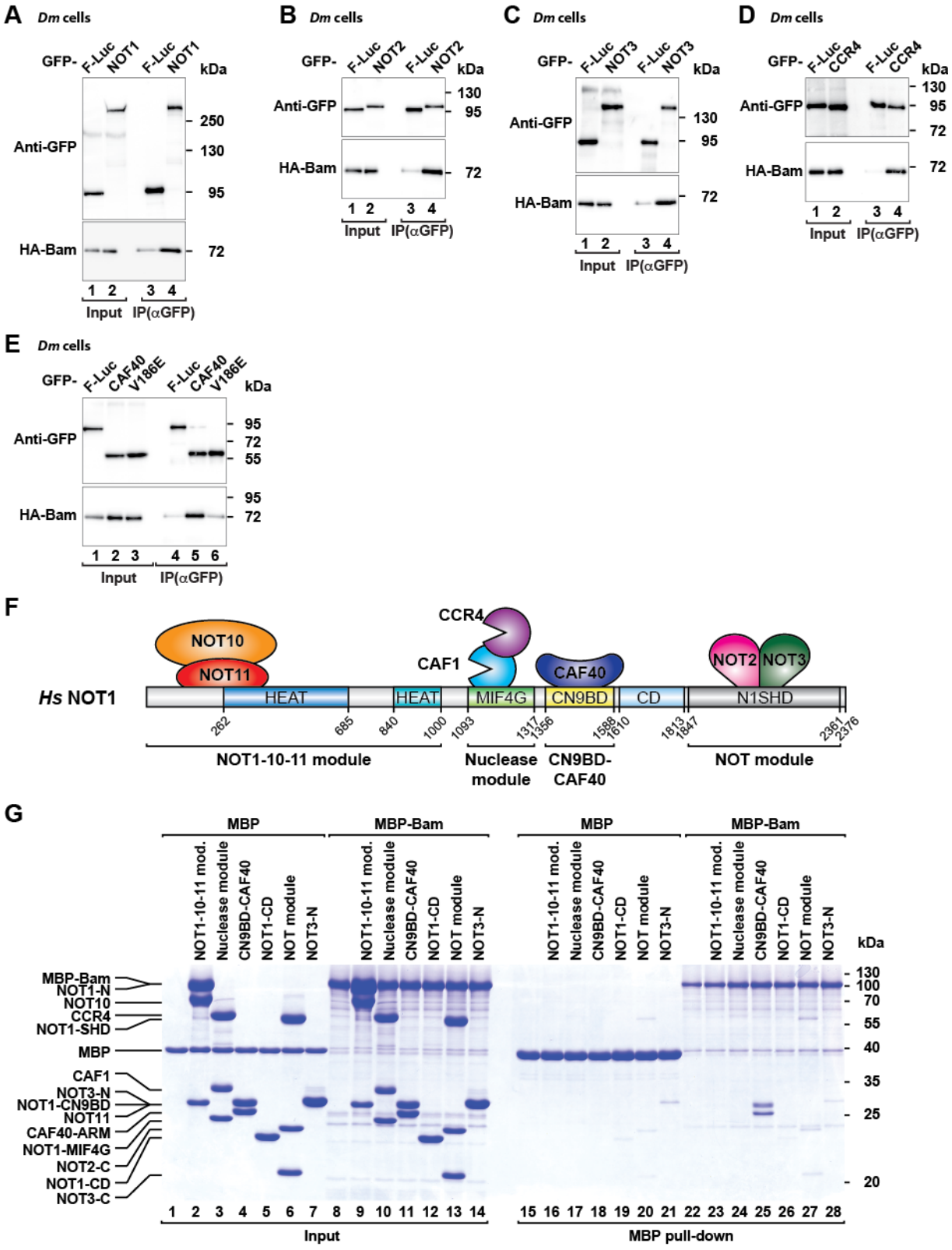
Supplemental Figure S2



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.

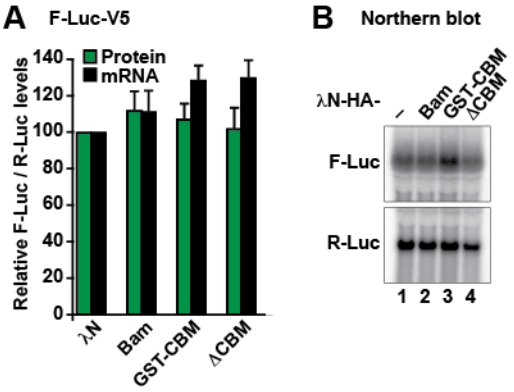
The corresponding experiment with the F-Luc-5BoxB reporter is shown in Figure 1B. (**E**) Normalized F-Luc activity values corresponding to the experiment described in Figure 2A and B. The F-Luc-5BoxB activity was normalized to that of the R-Luc transfection control and set to 100% in cells expressing 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



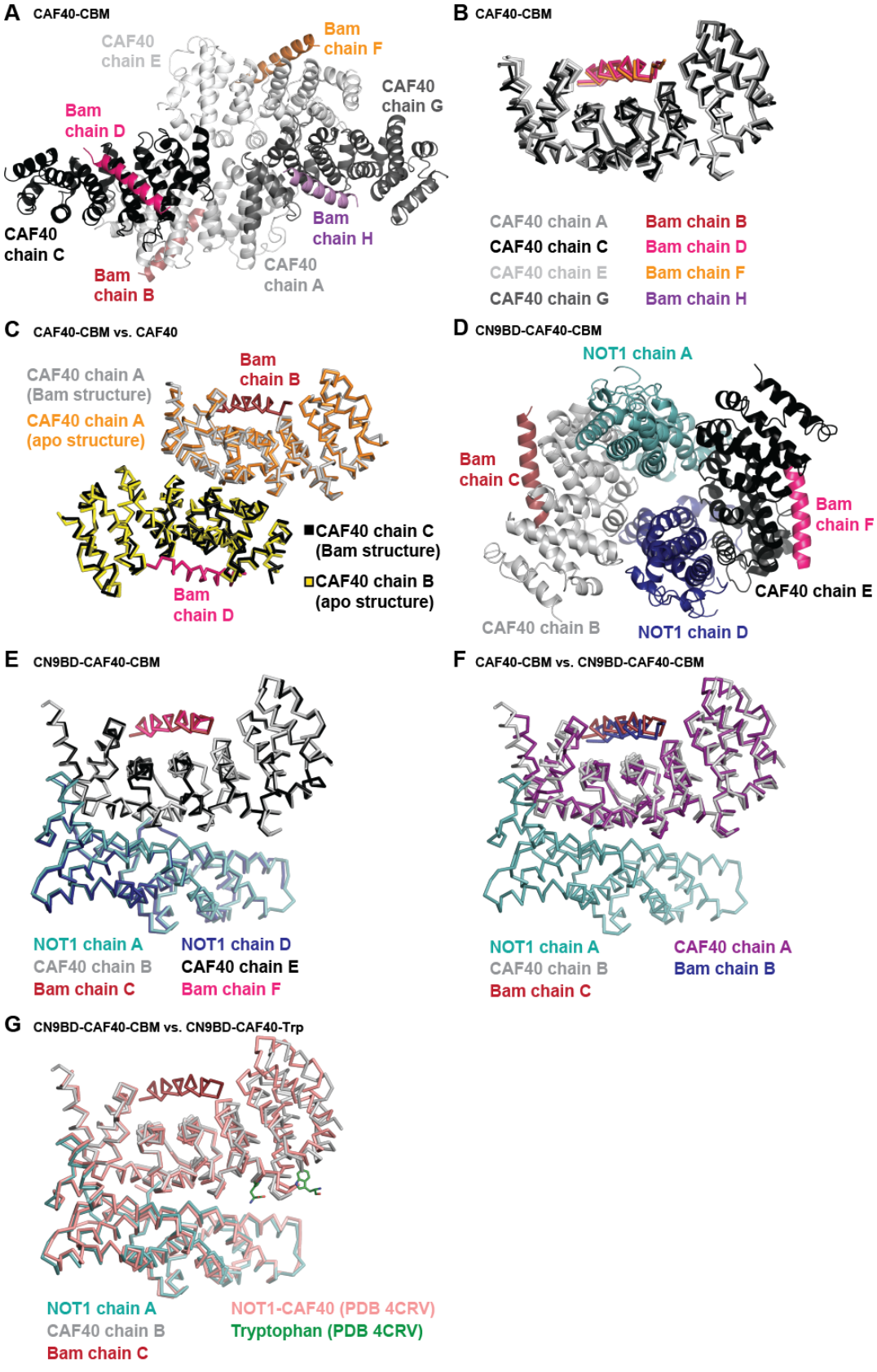
Supplemental Figure S3. Bam interacts with the CCR4-NOT complex. (*A*-*E*) Coimmunoprecipitation assays showing the interaction of HA-tagged Bam with the indicated GFP-tagged CCR4-NOT subunits in S2 cell lysates treated with RNase A. In all panels, GFP-F-Luc served as a negative control. Inputs (1% for 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



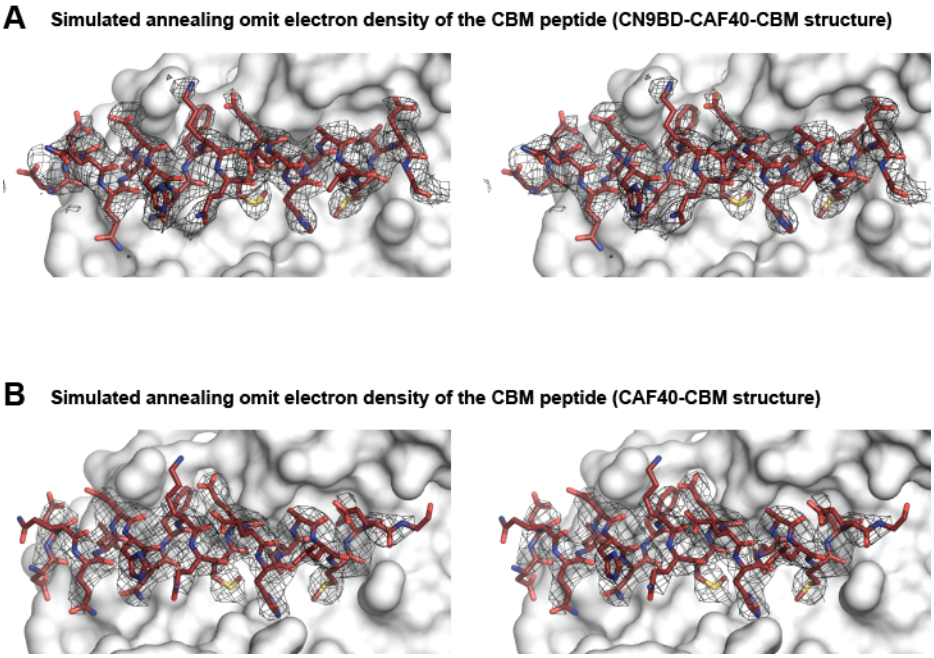
Supplemental Figure S4. Bam requires binding to the mRNA target to induce degradation. (*A,B*) Tethering assay using the F-Luc reporter lacking BoxB sites and λN-HA-tagged Bam (full-length or the indicated fragments) in S2 cells. The samples were analyzed as described in Figure 1B–D. The corresponding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

Supplemental Figure S5



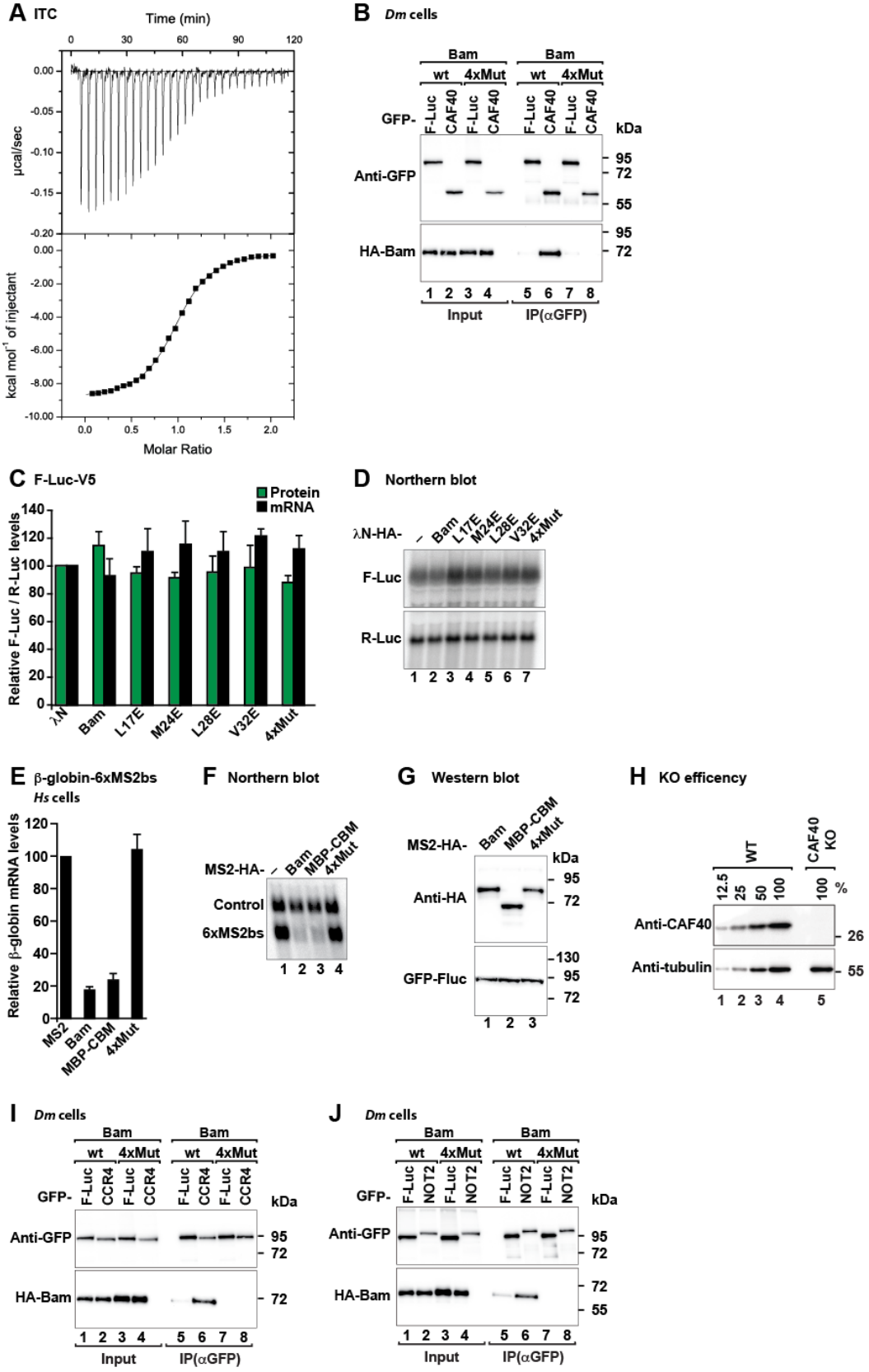
Supplemental Figure S5. Crystal structure of the Bam CBM bound to CAF40 and the CN9BD–CAF40 module. (*A*) Crystal packing of the CAF40–Bam CBM complex. The four copies of CAF40 (chains A, C, E and G) are shown in different shades of gray, the four Bam CBM peptides (chains B, D, F and H) in different 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.

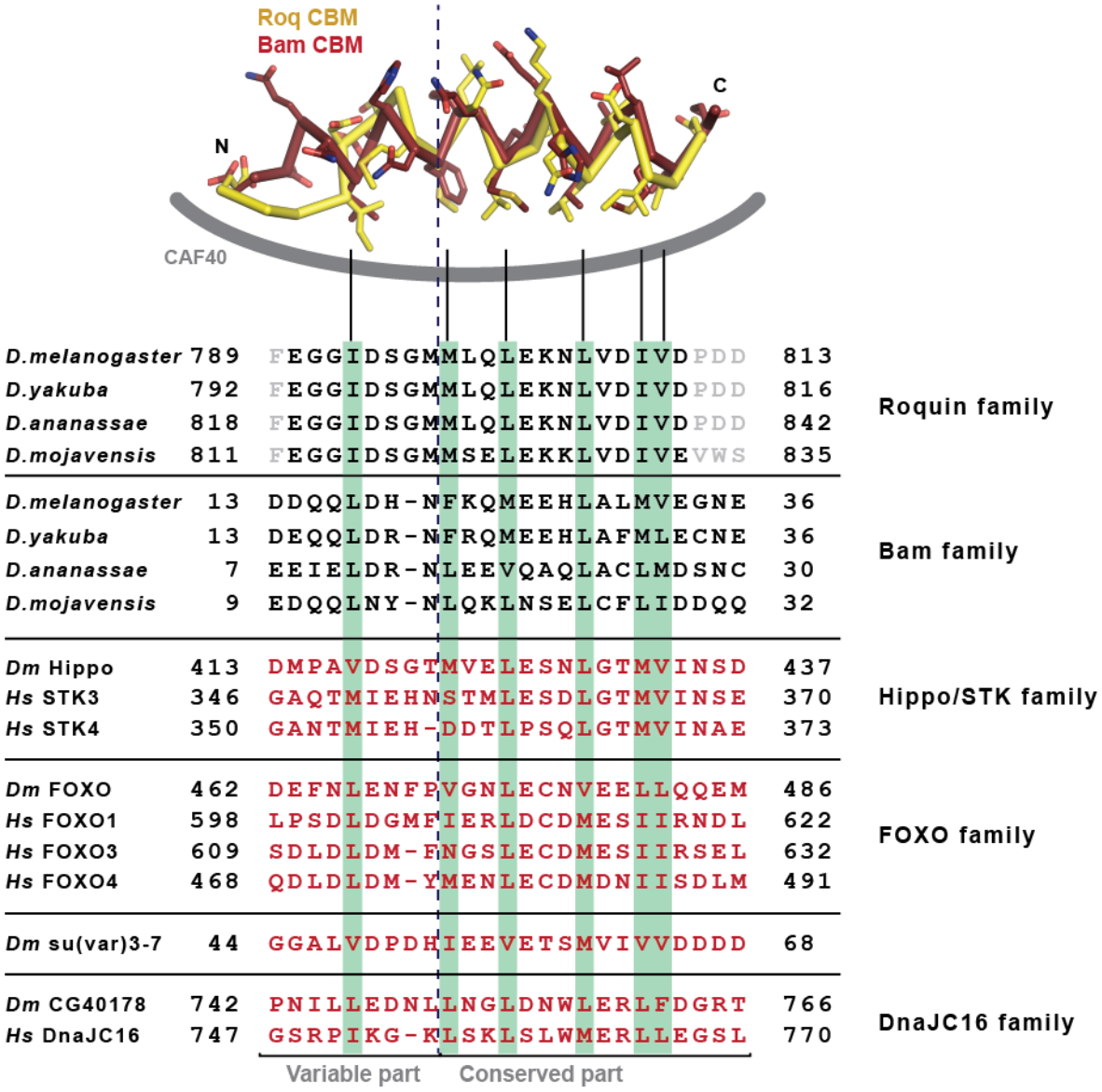
Supplemental Figure S7



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.

Metazoan NOT4 associates with the CCR4-NOT complex via interactions with CAF40 and NOT1

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

Department of Biochemistry, Max Planck Institute for Developmental Biology, Spemannstrasse 35, 72076 Tübingen, Germany

*Correspondence should be addressed to:

elisa.izaurralde@tuebingen.mpg.de

Running title: CCR4-NOT recruitment by metazoan NOT4

Keywords: Deadenylation / decapping / mRNA decay / translational repression / ubiquitination

Abstract

The CCR4-NOT deadenylase complex plays a crucial role in post-transcriptional regulation of gene expression in eukaryotes. The complex assembles around the NOT1 scaffold subunit which contains α-helical domains that provide binding sites for the other subunits. The NOT4 E3 ubiquitin ligase is one of the core subunits of the complex in yeast. However, it is not stably associated with the human (*Hs*) and *Drosophila melanogaster* (*Dm*) CCR4-NOT complexes. Here, we present data showing that both *Hs* and *Dm* NOT4 contain motifs in their C-terminal regions that directly interact with the CAF40 subunit of the complex. These CAF40 binding motifs (CBM) are necessary for NOT4 to induce degradation of bound mRNAs. Co-crystal structures of the *Dm* NOT4 CBM bound to CAF40 reveal striking similarities to the interaction of *Dm* Roquin with CAF40. Disruption of the NOT4-CAF40 interaction using structure-based mutations or by depleting CAF40 impairs the ability of NOT4 to elicit mRNA degradation. Our results explain the molecular basis for the association of metazoan NOT4 proteins with the core of the CCR4-NOT complex.

Introduction

The CCR4-NOT complex plays a central role in post-transcriptional regulation of gene expression by catalyzing the removal of mRNA poly(A) tails, thereby repressing translation and promoting mRNA degradation (Collart, 2016; Wahle and Winkler, 2013). 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 contains a catalytically inactive core of four to six subunits, depending on the organism (Chen et al., 2001; Lau et al., 2009; Temme et al., 2010). These subunits interact with the central NOT1 protein consisting of a series of α-helical domains which serve as binding sites for the other subunits (Bawankar et al., 2013). The N-terminal region of NOT1 contains helical repeat domains that interact with the NOT10 and NOT11 subunits of the complex in many species including metazoans (Basquin et al., 2012; Bawankar et al., 2013; Mauxion et al., 2013). A central MIF4G (middle domain of eIF4G) domain of NOT1 provides a binding site for the catalytic module, which comprises two deadenylases, namely CAF1 (or its paralog POP2) and CCR4a (or its paralog CCR4b). The NOT1 MIF4G domain is followed by a helical bundle domain, CN9BD (CAF40/CNOT9 binding domain), which interacts with the CAF40 subunit. CAF40 in turn interacts with the GW182/TNRC6 and Roquin proteins (Chen et al., 2014; Mathys et al., 2014; Sgromo et al., 2017). The C-terminal region of NOT1 contains the NOT1 superfamily homology domain (SHD), which interacts with the NOT2-NOT3 heterodimer to form the NOT module. The NOT module provides binding sites for the translational regulators Nanos, Roquin and Bicaudal-C, which recruit the CCR4-NOT complex to their mRNA targets (Sgromo et al., 2017; Raisch et al., 2016; Bhandari et al., 2014; Chicoine et al., 2007). In *Saccharomyces cerevisiae* (*Sc*), the NOT1 SHD also interacts with NOT4, an integral component of the yeast CCR4-NOT complex (Bai et al., 1999, Bhaskar et al., 2015).

NOT4 is an evolutionarily conserved E3 ubiquitin ligase that contains a RING domain, a coiled-coil region, an RRM domain and a C3H1 type zinc-finger domain at its N-terminus (NOT4-N). The C-terminal region (NOT4-C) is not well conserved and is predicted to be unstructured. The ubiquitin ligase activity of NOT4 depends on its specific interaction with the ubiquitin conjugating E2 enzyme UBCH5B, which has two orthologs in yeast: UBC4 and UBC5 (Albert et al., 2002; Bhaskar et al., 2015; Mulder et al., 2007). NOT4 ubiquitinates a large number of proteins including the small ribosomal protein RPS7A (Panasenko and Collart, 2012), the nascent polypeptide associated complex (NAC; Panasenko et al., 2006), the histone demethylase JHD2 (Mersman et al., 2009), Cyclin C (Cooper et al., 2012) and the transcription factor YAP1 (Gulshan et al., 2012). While the exact impact of the ubiquitination of these targets by NOT4 is not yet clear, it has been proposed that NOT4 plays a role in co-translational quality control in the context of stalled ribosomes (e.g. in the No-go decay; Dimitrova et al., 2009; Matsuda et al., 2014; Panasenko, 2014). In line with this, it has also been reported that NOT4 is required for global translational repression under nutritional stress and particularly for repression of mRNAs that cause transient ribosome stalling in yeast (Preissler et al., 2015).

Recently, it was shown that *Sc* NOT4 directly interacts with the C-terminal SHD of NOT1 (Bhaskar et al., 2015). A crystal structure illustrated how *Sc* NOT4 uses an elongated polypeptide in its unstructured C-terminal region to bind the NOT1 SHD via predominantly hydrophobic interactions. However, this region of the protein is not well conserved in species other than yeast (Bhaskar et al., 2015), an observation that suggested a different mode of interaction between metazoan NOT4 proteins and the core of the CCR4-NOT complex. Accordingly, although NOT4 is stably associated with the CCR4-NOT complex in yeast (Bai et al., 1999) and can be co-purified with the endogenous and overexpressed complexes (Nasertorabi et al., 2011; Stowell et al., 2016; Ukleja et al., 2016), the interaction seems to be transient *in vivo* in metazoans as NOT4 was not detected in mass spectrometric analyses of

the native *Hs* and *Drosophila melanogaster* (*Dm*) CCR4-NOT complexes (Lau et al., 2009; Temme et al., 2010). However, the human (*Hs*) NOT4 was shown to interact with the C-terminal half of NOT1 in yeast-two-hybrid assays (Albert et al., 2002), but the molecular basis and relevance of this association with the core complex has remained uncharacterized.

To elucidate the interaction between *Hs* and *Dm* NOT4 and the core CCR4-NOT complex in molecular detail, we initially confirm that the unstructured C-terminal region of *Hs* NOT4 directly interacts with the CCR4-NOT complex *in vitro* by multiple contacts, and the C-terminal regions of *Hs* and *Dm* NOT4 promote degradation of bound mRNAs in reporter assays. Furthermore, we identify a CAF40 binding motif (CBM) within the C-terminal regions of *Hs* and *Dm* NOT4 that is necessary and sufficient to mediate direct binding to the CAF40 subunit of the CCR4-NOT complex. Co-crystal structures of the *Dm* NOT4 CBM bound to *Hs* CAF40 confirm an overall similar binding mode as in the case of the *Dm* Roquin CBM, but also reveal striking differences to Roquin (Sgromo et al., 2017). Structure-based mutations in the CBMs abrogate the ability of *Hs* and *Dm* NOT4 to induce mRNA degradation. Moreover, depletion of CAF40 in HEK293T cells abolishes the interaction of *Hs* NOT4 with the CCR4-NOT complex and impairs NOT4-mediated mRNA degradation. Our results demonstrate that CAF40 is an important mediator for the association of NOT4 with the core CCR4-NOT complex in metazoa.

Results

Hs NOT4 directly interacts with the CCR4-NOT complex via its C-terminal region

Previous studies have shown that *Sc* NOT4 can directly interact with the C-terminal NOT1 SHD using an extended peptide motif in its C-terminal region (Fig. 1A) which is not well conserved outside yeast (Bhaskar et al., 2015; Panasenko and Collart, 2011). *Hs* NOT4 (the comparably short isoform 5) has also been shown to interact with the C-terminal half of NOT1 in yeast-two-hybrid assays (Albert et al., 2002), but the molecular basis of this interaction remained unknown.

In order to better characterize the interaction of *Hs* NOT4 with the CCR4-NOT complex, we tested the interaction between full-length NOT4 (the longer isoform 1; Fig. 1A and S1) and a pentameric complex comprising the C-terminal half of NOT1 plus CAF1 and CAF40, and C-terminal fragments of NOT2 and NOT3 (Sgromo et al., 2017) in *in vitro* MBP pulldown assays. Consistent with the previous yeast-two-hybrid experiment (Albert et al., 2002), we observed a direct interaction of NOT4 with this pentameric complex (Fig. 1B, lane 7). Furthermore, based on sequence alignments we generated two NOT4 fragments, one comprising the structured and conserved N-terminus (Fig. 1A; NOT4-N; residues M1-T274), and another corresponding to the unstructured C-terminal region (Fig. 1A; NOT4-C; residues P275-A575). While NOT4-N does not interact with the pentameric CCR4-NOT subcomplex (Fig. 1B, lane 8), NOT4-C binds as efficiently as the full-length protein (Fig. 1B, lane 9). This shows that the site of interaction is indeed located in the unstructured C-terminus of NOT4, as already proposed earlier (Albert et al., 2002). This direct interaction raises the question whether NOT4 can recruit the CCR4-NOT complex to bound mRNAs in order to induce their degradation.

The C-terminal regions of Hs and Dm NOT4 induce degradation of bound mRNAs

To investigate whether NOT4 can induce degradation of mRNAs, we used an MS2-based tethering assay in *Hs* HEK293T cells. Full-length NOT4, NOT4-N and NOT4-C were expressed with an MS2-HA tag that binds to a β-globin reporter containing six MS2 binding sites in the 3' UTR (β-globin-6xMS2bs; Lykke-Andersen et al., 2000). Tethered *Hs* NOT4 caused a strong reduction of the β-globin-6xMS2bs mRNA level compared to the negative control MS2-HA (Fig. 1C,D, lane 2). Similarly, tethered NOT4-C had the same effect whereas NOT4-N was inactive (Fig. 1C,D). All NOT4 fragments were expressed at comparable levels (Fig. 1E) and none of them affected the expression of the control β-globin mRNA lacking MS2 binding sites (Fig. 1D; control).

We also tested the effect of NOT4 tethering on a reporter that contains the Renilla luciferase open reading frame and six MS2 binding sites in the 3' UTR (R-Luc-6xMS2bs). Tethered NOT4 and NOT4-C caused a significant reduction of the R-Luc activity compared to the MS2-HA control (Fig. S3A). The reduction is predominantly explained by a corresponding decrease in mRNA abundance of the reporter (Fig. S3A,B). An R-Luc reporter lacking the MS2 binding sites was not affected by the expression of MS2-HA-tagged full-length NOT4 or fragments (Fig. S3C,D). This suggests that NOT4 can indeed reduce the levels of mRNAs by inducing degradation, and that this activity resides exclusively in the C-terminal region of the protein.

As NOT4-C is not well conserved between vertebrates and invertebrates (Fig. S1 and S2), we tested whether *Dm* NOT4 and its C-terminal region can actively induce mRNA degradation as well. To this end, we used a λN-based tethering assay in *Dm* S2 cells (Behm-Ansmant et al., 2006). Similar to the situation in *Hs*, *Dm* NOT4 and NOT4-C efficiently reduced the levels of the bound F-Luc-5BoxB reporter mRNA (Fig. 1F,G). These effects were matched by a similar reduction of F-Luc-5BoxB mRNA levels upon tethering NOT4 and NOT4-C (Fig. 1F). All three NOT4 fragments were expressed at comparable levels (Fig. 1H) and none of them affected the expression of an F-Luc reporter lacking the BoxB sites (Fig S3E,F). This

indicates that the repressive activity of the NOT4 C-terminal region is conserved between *Hs* and *Dm*.

The CCR4-NOT complex is known to induce translational repression independently of deadenylation (Cooke et al., 2010; Zekri et al., 2013). To assess whether NOT4 can repress the expression of mRNAs in the absence of mRNA decay, we used an R-Luc-6xMS2bs reporter mRNA that contains a DNA-encoded stretch of 95 adenosine residues followed by the 3'-end of the long noncoding RNA MALAT1 which is processed by RNase P to form the non-adenylated 3'end of the RNA (R-Luc-6MS2bs-A95-MALAT; Kuzuoglu-Ozturk et al.; 2016; Wilusz et al., 2012). NOT4 and NOT4-C efficiently repressed the translation of this reporter in HEK293T cells (Fig. S3G) without causing reporter mRNA decay (Fig. S3G,H). We thus conclude that the repressive activity of NOT4 resides exclusively in the C-terminal region.

NOT4 causes mRNA decay via the 5'-to-3' decay pathway

The interaction of *Hs* NOT4 with the CCR4-NOT complex (Fig. 1B) suggests that the repressive activity of the protein may depend on the recruitment of CCR4-NOT to bound mRNAs. To investigate this question, we inhibited decapping by overexpressing a catalytically inactive mutant of the GFP-tagged decapping enzyme DCP2 (E148Q) in HEK293T cells (Fig. S3I,J). In these cells, NOT4 caused the accumulation of a shorter, deadenylated decay intermediate of the β-globin-6xMS2bs reporter mRNA (Fig. S3I,J, lane 5 vs. 2). *Hs* Nanos2, which is known to induce mRNA decay via the 5'-to-3' decay pathway (Bhandari et al., 2014), was used as a positive control (Fig. S3I,J lane 6 vs. 3). These results suggest that NOT4 funnels bound mRNAs into the general 5'-to-3'-mRNA decay pathway where deadenylation is followed by decapping and exonucleolytic degradation.

Similar results were obtained in *Dm* S2 cells when overexpressing a GFP-tagged DCP1 mutant (R70G-L71S-N72S-T73G) that inhibits decapping in a dominant negative manner

(Chang et al., 2014; Kuzuoglu-Ozturk et al., 2016). As observed in *Hs* HEK293T cells, tethered NOT4 caused a stabilization of a deadenylated species of the F-Luc-5BoxB reporter mRNA similar to the positive control, the GW182 protein (Fig. S3K,L). The equivalent expression level of the overexpressed proteins was confirmed by western blot (Fig. S3M).

The accumulation of deadenylated mRNA reporters upon NOT4 tethering in decapping-compromised cells suggests that NOT4 mediates deadenylation-dependent decapping.

Hs NOT4 possesses binding sites for different modules of the CCR4-NOT complex

To determine which subunits of the CCR4-NOT complex interact with NOT4, we performed MBP pulldown assays with different purified modules, namely the NOT10-NOT11 module (comprising the N-terminus of NOT1, the C-terminal half of NOT11 and NOT10), the NOT1 MIF4G bound to CAF1, CAF40 in complex with the NOT1 CN9BD, the NOT1 connector domain (CD) and the NOT module consisting of the C-terminal regions of NOT1, NOT2 and NOT3 (Fig. 2A; Boland et al., 2013; Chen et al., 2014; Petit et al., 2012; Sgromo et al., 2017). Strikingly, both the CN9BD-CAF40 dimer as well as the NOT module, but none of the other subcomplexes, can be co-purified with MBP-NOT4 (Fig. 2B, lanes 13 and 17), indicating the existence of multiple binding sites for NOT4 in the CCR4-NOT complex. This binding pattern is similar to *Dm* Roquin which also interacts with the CN9BD-CAF40 dimer and the NOT module (Sgromo et al., 2017). Additionally, these results show that despite the lack of sequence conservation between *Sc* and *Hs* NOT4-C, both use the NOT module for docking to the CCR4-NOT complex.

Similarly to *Sc* NOT4 (Bhaskar et al., 2015), *Hs* NOT4 binds strongly to the NOT1 SHD (Fig. 2C, lane 7). On the other hand, we observed only a very weak interaction with the NOT2-3 dimer (Fig. 2C, lane 9), indicating that the main binding site resides in the NOT1 component

of the NOT module. Furthermore, NOT4 interacts directly with CAF40 in the absence of the NOT1 CN9BD (Fig.2C, lane 11).

A minimal NOT1 and CAF40 binding region (NCBR) within NOT4-C is sufficient to mediate the contact with the CCR4-NOT complex.

To determine more precisely how NOT4 interacts with NOT1 and CAF40, we generated several fragments of NOT4-C and tested them for binding to the NOT module and CAF40. We identified a short central region in NOT4-C (residues E377-D428) that is sufficient to bind to both complexes (Fig. 3A, lanes 11 and 16), albeit with reduced affinity compared to NOT4-C (Fig. 3A, lanes 9 and 14). We therefore termed this region the NOT1 and CAF40 binding region (NCBR). Neither of the two other fragments flanking the NCBR was able to interact with CAF40 (Fig. 3A, lanes 15 and 17), and the C-terminal fragment (residues Q429-A575) shows only very weak binding to the NOT module while the N-terminal fragment (residues P275-S376) does not bind at all (Fig. 3A, lanes 10 and 12).

To separate the regions in the NCBR responsible for binding to the NOT module and CAF40, respectively, we divided the NCBR in two fragments, and tested their interactions in MBP pulldowns. The interaction with the NOT module was reduced by deleting the N-terminal half of the NCBR (residues E377-D402), similar to the deletion of the full NCBR (Fig. S3N, lanes 10 and 11). Deletion of the second half (residues E400-D428) had no effect (Fig. S3N, lane 12). This suggests that the N-terminal half of the NCBR contributes to NOT module binding, but that in addition other auxiliary sequences must exist outside the NCBR. In contrast to that, deletion of the NCBR as well as the deletion of its C-terminal half strongly impaired the interaction of NOT4-C with CAF40 (Fig. S3N, lanes 15 and 17), indicating that the main binding site for CAF40 resides inside residues E400-D428. This fragment which includes a predicted α-helix was therefore termed a CAF40-binding motif (CBM). Deletion of the N-terminal half of the NCBR, on the other hand, had almost no effect on the interaction between

NOT4-C and CAF40 (Fig. S3N, lane 16).

The NCBR is necessary and sufficient for NOT4-mediated mRNA decay

To assess the importance of the NCBR interactions with the CCR4-NOT complex, we tested all NOT4-C fragments in tethering assays as described in Fig. 1C,D. In agreement with the MBP pulldown assay (Fig. 3A), the NCBR elicited degradation of the target mRNA, whereas the other two NOT4-C fragments were inactive (Fig. 3B,C). This result clearly indicates that the NCBR is important for NOT4-mediated mRNA decay (Fig. 3B,C). Furthermore, regions upstream and downstream of the NCBR may facilitate the interaction with NOT1 or CAF40, but are not sufficient for strong individual binding to the CCR4-NOT complex (Fig. 3A) and induction of mRNA decay (Fig. 3B,C). All tested fragments were expressed at comparable levels (Fig. 3D).

To determine the contribution of the NCBR to the interaction of NOT4 with the endogenous CCR4-NOT complex, we expressed V5-SBP-tagged NOT4-C and a deletion of the NCBR in HEK293T cells and performed SBP pulldown assays. Consistent with the MBP pulldown assays and the previous yeast-two-hybrid assays (Albert et al., 2002), NOT4-C co-purified with endogenous NOT1, NOT2, NOT3 and CAF40 (Fig. 3E, lane 7). Deletion of the NCBR from NOT4-C resulted in complete loss of binding to all four CCR4-NOT subunits (Fig. 3E, lane 8). Deletion of either of the two NCBR motifs individually was also sufficient to disrupt the interactions, indicating that both motifs are necessary to establish a strong interaction with the complex (Fig. 3E, lanes 9 and 10).

In line with these results, deletion of the complete NCBR or either of the two motifs abolished the NOT4-C-mediated degradation of the β-globin-6xMS2bs reporter mRNA in HEK293T cells (Fig 3F,G). These results underline the importance of the NCBR, and therefore most likely the recruitment of the CCR4-NOT complex, for the degradative activity of NOT4. All

tethered proteins were expressed at a similar level (Fig. 3H).

Crystal structures of the Dm NOT4 CBM bound to CAF40

To determine the molecular basis of the CAF40-NOT4 interaction, we set out to solve high-resolution crystal structures of the *Hs* and *Dm* NOT4 CBMs (residues E400-D428 and D813-Q838, respectively) with either *Hs* or *Dm* CAF40 (ARM domains). However, the only combination that yielded well-diffracting crystals in multiple conditions was the complex of the *Dm* NOT4 CBM and the *Hs* CAF40 ARM domain. We solved structures of the complex in two different crystal forms at 2.1 Å (crystal form 1, space group $P2_12_12$; Fig. S4A) and 2.2 Å (crystal form 2, space group $I2_12_12_1$; Fig. S4B), respectively. In both crystal forms, the asymmetric unit contains two CAF40-NOT4 heterodimers in highly similar arrangements (Fig S4C; rmsd 1.06 Å over 544 Cα atoms).

In each of the CAF40-NOT4 heterodimers, the CAF40 ARM domain adopts the well-known fold consisting of six armadillo repeats (Chen et al., 2014; Garces et al., 2007; Mathys et al., 2014; Sgromo et al., 2017). The *Dm* NOT4 CBM peptide (residues 813-838) folds partially into an α-helix and binds into the concave surface of CAF40 (Fig. 4A-C). This arrangement is similar between the two heterodimers within each crystal form (Crystal form 1: rmsd 0.67 Å over 267 Cα atoms; crystal form 2: rmsd 0.41 Å over 242 Cα atoms; Fig. S4D,E) and also between the dimers of the two crystal forms (rmsd 0.37 Å over 256 Cα atoms; Fig. S4F). Indeed, the most pronounced differences between the four heterodimers is observed in the N- and C-terminal ends of the NOT4 peptide, indicating structural flexibility of these termini (Fig. S4D-F).

Comparison with Roquin reveals significant differences in binding mode

Interestingly, the NOT4 CBM peptide binds to the same spot on the concave surface of CAF40 as the previously identified CBM of *Dm* Roquin (Fig. 4D and S4G; PDB 5LSW; Sgromo et al., 2017).

Both peptides fold into amphipathic α-helices that are recognized by the same conserved hydrophobic surface of CAF40 (Fig. 4C,D), despite the lack of sequence similarity between the CBMs. Nevertheless, an evolutionary relation of the two motifs is unlikely as the relative orientation of the peptides on CAF40 is very different, i.e. the NOT4 CBM is flipped by almost 180° with respect to the Roquin CBM (compare Fig.4C and D).

Consequently, the α-helix of the NOT4 CBM lies in an almost antiparallel fashion in the groove formed by the CAF40 helices α5, α8 and α11. The core of the predominantly hydrophobic interface is formed by residues F821, T824, L828, L831 and M832 of NOT4 that are inserted into the pocket lined by CAF40 residues R130, Y134, L137, T138 and L177, but also residues L816, F818 and P820 of the extended N-terminus of NOT4 strongly contribute as they extend the interface by interacting with CAF40 residues G141, G144, T180 and V181 (Fig. 4E). In particular, NOT4 P820 might have a key role as it marks the border between the α-helix and the N-terminal extension. Finally, there are additional hydrophilic interactions that stabilize the interaction, including a hydrogen bond between CAF40 R46 and NOT4 D819, and NOT4 E835 that stabilizes the C-terminus of the CBM by forming two hydrogen bonds to the backbone at the entrance of CAF40 helix α5 (Fig. 4E).

Compared to Roquin, the overall characteristics of the interaction seem to be very similar as also Roquin uses an amphipathic helix to mediate binding. On the other hand, there are pronounced differences apart from the obviously flipped orientation of the two peptides. In particular, the contribution of the N-terminal extension of the NOT4 CBM to the interaction is not observed in the Roquin CBM (Sgromo et al., 2017). Also, a comparison of the conformation of CAF40 in both structures shows a small but striking difference: Whereas the overall fold is highly similar (Fig. S4G; rmsd 0.65 Å over 239 Cα atoms), and also most side

chains involved in the interactions adopt similar conformations, Y134 adopts completely different orientations in both structures (Fig. S4H). This small difference in the center of the hydrophobic pocket is sufficient to allow the two peptides to bind in opposite directions.

Validation of the binding interface

To verify the interfaces obtained from the crystal structure, we generated mutations in the NOT4-CBM and CAF40 and tested them in MBP pulldown assays. First, we showed that similarly to the *Hs* NOT4 CBM, the *Dm* NOT4 CBM interacts with *Hs* CAF40 under these conditions (Fig. 4F, lane 14), confirming that it is a *bona fide* CBM in solution. To disrupt the interaction from the CAF40 side, we used either a single point mutation (V181E) or a double mutation (2x mut; Y134D+G141W), that were both previously shown to disrupt the interaction with the Roquin CBM (Sgromo et al., 2017). In both cases, these mutations abolished the interaction of CAF40 with the *Dm* NOT4-CBM (Fig 4F, lanes 15 and 16). Conversely, single F821D and L828E substitutions in the *Dm* NOT4-CBM were sufficient to abrogate the interaction with CAF40 (Fig 4G, lanes 14 and 16).

Similar to the results obtained with *Dm* NOT4, a single V181E mutation in CAF40 was sufficient to abolish the interaction with the *Hs* NOT4-CBM (Fig 5A, lanes 14 and 15). However, the CAF40 2x mut reduced but did not abolish the interaction with the *Hs* NOT4-CBM, indicating the existence of species-specific differences in the molecular details of the interaction in *Dm* and *Hs* (Fig 5A, lanes 14 and 16).

CAF40 mediates the recruitment of the CCR4-NOT complex by NOT4

To assess the role of the CAF40-CBM interface in a cellular context, we introduced either individual point mutations (F405D, F408A, T411E, L415E) or combinations into V5-SBP-MBP tagged *Hs* NOT4-C and tested them for the ability to interact with endogenous CCR4-NOT subunits in HEK293T cell lysates. All single point mutations modestly affected the

binding to NOT1 and CAF40 (Fig. S5A, lanes 11-14). However, a T411E-L415E double mutant was sufficient to strongly reduce the interaction (Fig. S5A, lane 15), and the combination of all four mutations (4x mut) abolished the interaction with all tested endogenous CCR4-NOT complex subunits (Fig. 5B, lane 6, and S5A, lane 16).

When tested in a tethering assay in HEK293T cells, the 4x mut impaired the ability of both NOT4-C and NCBR to degrade the β-globin-6xMS2bs reporter mRNA (Fig. 5C,D). All the tethered proteins were expressed at comparable levels (Fig. 5E). These findings underline the importance of the CBM for the function of NOT4 and indicate that the CBM is the major site for interaction with the CCR4-NOT complex.

The CCR4-NOT complex is required for Hs NOT4-induced mRNA decay

To assess the role of CAF40 in the recruitment of the CCR4-NOT complex by *Hs* NOT4, we depleted CAF40 from HEK293T cells by CRISPR-Cas9 genome editing. The depletion was confirmed by western blot (Fig. 6A, lane 5). CAF40 depletion did not alter the levels of endogenous NOT1, NOT2 and NOT3 proteins in the cell (Fig. 6A, lane 5). In this cell line, V5-SBP-MBP tagged NOT4-C was unable to pull down components of the endogenous CCR4-NOT complex (Fig. 6B, lane 12 vs. 10), confirming that CAF40 plays a major role in the interaction of NOT4 with the CCR4-NOT complex. Addition of purified recombinant *Hs* CAF40 (ARM domain) to the cell lysates prior to the pulldown rescued the interaction of NOT4-C with endogenous NOT1 (Fig. 6B, lanes 14 and 16), further confirming that CAF40 is required for association of NOT4 with the CCR4-NOT complex.

In the CAF40-depleted cell line, the repressive effect on the R-Luc-6xMS2bs reporter mRNA upon tethering MS2-HA-tagged NOT4 and NOT4-C was partially impaired in comparison to the control cells (Fig. 6C). The finding that CAF40 depletion alone was not sufficient to completely abolish the repressive activity of NOT4 suggests that other factors also contribute to the NOT4-mediated repression (Fig. 6C).

To test whether the remaining activity of NOT4 in the CAF40 knockout cell line is due to the recruitment of the CCR4-NOT complex via the NOT module, we knocked down NOT1 in control cells or in the CAF40 knockout cell line. In the control cell line, the NOT1 knockdown moderately reduced the repression by tethered NOT4, similar to the effect of the CAF40 knockout (Fig. 6C). However, combination of the CAF40 knockout with the NOT1 knockdown drastically reduced the activity of tethered NOT4 (Fig. 6C). This indicates that the major contribution to NOT4-mediated repression stems from the CCR4-NOT complex. This result suggests that binding of NOT4 to both CAF40 and the NOT module is required to synergistically maintain the full affinity of the overall interaction, but on the other hand, binding to either of the sites is sufficient for a weak association with the CCR4-NOT complex. Western blot showed that shRNA-mediated NOT1 knockdown reduced the expression of endogenous NOT1 to around 20 % of the control level (Fig. 6D). The equivalent expression of NOT4 in all samples was confirmed by western blot (Fig. 6E).

The CAF40-CBM interaction is essential for Dm NOT4-mediated mRNA decay

To validate the relevance of the CAF40-CBM interaction for the repressive function of *Dm* NOT4, we overexpressed GFP-tagged CAF40 (either wild-type or a V186E mutant; this mutant is equivalent to V181E in *Hs* CAF40 that is unable to interact with NOT4). In this context, we performed tethering assays using NOT4 and NOT4-C as described in Fig 1F. In this assay, decay of the F-Luc-5BoxB reporter mRNA was not affected by overexpressing wild-type CAF40 compared to the control situation where only GFP was overexpressed (Fig. S5B,C, lanes 1-6). However, overexpression of the CAF40 V186E mutant resulted in a complete inhibition of the NOT4-induced target mRNA degradation (Fig. S5B,C, lane 8 and 9). This result shows, firstly, that the V186E mutant can displace endogenous CAF40 when overexpressed, and, secondly, that this displacement of wild-type CAF40 by the non-

interacting mutant has a dominant negative effect on the activity of NOT4. This underlines the important function of the interaction with CAF40 for the activity of NOT4. The equivalent expression of NOT4 and NOT4-C in all samples was confirmed by western blot (Fig. S5D).

Discussion

In this study, we show that *Hs* and *Dm* NOT4 directly interact with the CCR4-NOT complex through their unstructured C-terminal regions. *Hs* NOT4 uses at least two binding sites for interaction with CCR4-NOT subunits, one site contacting the NOT module, and another motif that contacts CAF40. This CAF40-binding motif is conserved in *Dm*, and it is functionally relevant for the recruitment of the CCR4-NOT complex by NOT4. The crystal structures of the *Dm* CBM bound to CAF40 elucidate the molecular basis of the interaction and show differences in binding mode compared with the CBM of *Dm* Roquin. Specific disruption of the CAF40-CBM interaction by single residue substitutions or CAF40 depletion impairs the ability of NOT4 to induce degradation of bound mRNAs.

Binding of NOT4 to CCR4-NOT is conserved

Our data show that both *Hs* and *Dm* NOT4 interact directly with the CCR4-NOT complex and that these interactions are strong enough to be observed by different techniques (i.e., *in vitro* pulldown assays, pulldown assays in *Hs* cell lysates, and tethering assays in *Hs* and *Dm* cells). In addition, a previous study showed the association of *Hs* NOT4 with NOT1 in yeast-two-hybrid assays (Albert et al., 2002). In contrast to that, several studies that investigated the overall composition of the *Hs* and *Dm* CCR4-NOT complexes showed that NOT4 does not co-purify with the complex, indicating that its association with the complex may be transient or regulated (Lau et al., 2009; Temme et al., 2010).

The situation is different in yeast where NOT4 appears to be an integral component of the CCR4-NOT complex (Bai et al., 1999; Bhaskar et al., 2015). This difference in the association behavior of NOT4 and the rest of the CCR4-NOT complex may reflect the different binding modes in yeast and metazoans. Both *Hs* and *Dm* NOT4 share the conserved CBM which, as our tethering assays suggest, mediates the major contact with the CCR4-NOT

complex. In addition to that, *Hs* NOT4 also contacts the NOT module with one motif inside the NCBR and auxiliary sequences outside the NCBR. It is unclear whether *Dm* NOT4 also contains additional CCR4-NOT binding sites outside the CBM, but since the disruption of the CBM-CAF40 interaction is sufficient to completely suppress the activity in tethering assays, we propose that the CBM provides the major contact. In *Sc*, however, the situation is different as the peptide that mediates the contact to the NOT1 SHD is sufficient to tether NOT4 to the rest of the complex (Bhaskar et al., 2015). Analysis of the sequence of *Sc* NOT4 shows that the protein does not contain any conserved CBM, although it cannot be excluded that additional motifs exist that mediate additional contacts. In evolutionary terms, the NOT4 proteins are similar to other interactors of the CCR4-NOT complex like Nanos and Roquin which have also retained the overall interaction with the CCR4-NOT complex, but have diversified the modes of interaction during evolution (Bhandari et al., 2014; Raisch et al., 2016; Sgromo et al., 2017).

These different binding modes may result in a higher overall affinity of *Sc* NOT4 for the rest of the CCR4-NOT complex compared to the metazoan counterparts which would explain the obviously tighter incorporation of NOT4 into the yeast complex. On the other hand, we also cannot exclude the possibility that the association of NOT4 with the CCR4-NOT complex might be regulated, for example by post-translational modifications, in the different organisms.

Comparison with Roquin suggests the existence of other CBM-containing proteins

Comparing NOT4 with other proteins that interact with the core CCR4-NOT complex reveals striking similarities in the nature of these interactions. When comparing NOT4 with *Dm* Roquin, these similarities are quite obvious since also Roquin binds to the CCR4-NOT complex via a CBM and additional sequences contacting the NOT module (Sgromo et al., 2017), leading to an overall similar situation as in the case of NOT4. On the other hand, the

different orientations of the CBM peptides and the lack of sequence similarity between the regions that are responsible for NOT module binding are strong indicators for the absence of an evolutionary link between the two proteins. On the other hand, it is likely that both CBMs arose by convergent evolution. This raises the question why both proteins evolved to bind to the same binding site on CAF40 via chemically relatively similar binding modes. One possibility is that recruiting the CCR4-NOT complex via CAF40 to bound mRNA targets is generally beneficial, as it may for example position the nuclease domains of CAF1 and CCR4 optimally with respect to the 3'-end of the poly(A)-tail in order to assure high enzymatic activity. Another possibility is that NOT4 and Roquin might compete with each other for binding to the CCR4-NOT complex, which could be an efficient means of differentially regulating the translational repression of target mRNAs, and it might also ensure that only one E3 ubiquitin ligase activity, either from NOT4 or Roquin, is associated with the CCR4-NOT complex at any time.

The high conservation of the CBM-binding surface of CAF40, also in species like yeast where no CBM has been identified so far, suggests that more CBM-containing proteins may exist. However, the lacking evolutionary link and the differences in binding mode between NOT4 and Roquin suggest that these hypothetical CBMs may again have no sequence similarity to the known CBMs, which would make their identification by *in silico* sequence search difficult.

Materials and methods

DNA constructs

For recombinant expression in *E.coli*, *Hs* NOT4 cDNA constructs were inserted in the pnEA-pM plasmid (Diebold et al., 2011) between the XhoI and NheI restriction sites, resulting in fusion proteins carrying N-terminal MBP tags cleavable by the HRV3C protease, and in addition C-terminal GB1- and 6xHis-tags. For expression in human HEK293T cells, *Hs* NOT4 cDNA constructs were inserted in the pCIneo-V5-SBP-MBP plasmid between the XhoI and NotI restriction sites, or into the pcDNA3.1-MS2-HA plasmid using the same restriction sites.

Dm NOT4 cDNA constructs were inserted between the XhoI and BamHI restriction sites of the pnYC-vM plasmid, resulting in TEV-cleavable MBP fusion proteins. For expression in *Drosophila* S2 cells, *Dm* NOT4 cDNA constructs were inserted between the NotI and BstBI restriction sites of the pAc5.1B plasmids with N-terminal GFP or λN-HA tags, except for a full-length *Dm* NOT4 construct, where the GFP tag is C-terminal.

For expression of the *Hs* NOT1-10-11 complex, two plasmids were created. The N-terminus of NOT1 (residues M1-D1000) was cloned in a pnYC plasmid that did not contain any tag sequence. NOT10 (residues D25-Q707) and NOT11 (residues D257-D498) were cloned in a bicistronic plasmid based on the pnEA backbone, resulting in the expression of untagged NOT10 and NOT11 carrying a C-terminal, TEV-cleavable 6xHis tag.

Plasmids for the expression of the other subcomplexes of the *Hs* CCR4-NOT complex have been described before (Bhandari et al., 2014; Boland et al., 2013; Chen et al., 2014; Petit et al., 2012; Sgromo et al., 2017). The DNA constructs used in this study are listed in Table S1.

Protein expression and purification

All recombinant proteins were expressed in *E. coli* BL21 (DE3) Star cells (Invitrogen) in LB medium overnight at 20 °C. *Hs* and *Dm* NOT4 CBM constructs were expressed with N-terminal MBP tags. The cells were lysed in a buffer containing 50 mM HEPES pH 7.5, 300 mM NaCl and 2 mM DTT supplemented with complete EDTA-free protease inhibitors, DNaseI and Lysozyme. The proteins were isolated from the lysate using amylose resin (New England Biolabs) followed by anion chromatography using a HiTrapQ column (GE Healthcare). Finally, the NOT4 constructs were applied to size exclusion chromatography on a Superdex 200 26/600 column (GE Healthcare) using a buffer containing 10 mM HEPES pH 7.5, 200 mM NaCl and 2 mM DTT.

The purification of *Hs* CAF40 (residues R19-E285) has been described before (Sgromo et al., 2017). Briefly, the protein was expressed fused to a 6xHis tag cleavable by the HRV3C protease. Lysis was carried out in lysis buffer containing 50 mM potassium phosphate pH 7.5, 500 mM NaCl, 10 % glycerol, 20 mM imidazole and 2 mM β-mercaptoethanol supplemented with complete EDTA-free protease inhibitors, DNaseI and Lysozyme. 6xHis-CAF40 ARM was isolated using a HiTrap IMAC column. Then, the 6xHis 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. Subsequently, CAF40 was further purified using a HiTrap Heparin (GE Healthcare) column followed by gel filtration on a Superdex 200 26/600 column in a buffer containing 20 mM Tris-HCl pH 7.5, 150 mM NaCl and 1 mM DTT.

The assembled *Hs* NOT1-10-11 trimer was obtained by co-expression of C-terminally 6xHis-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 supplemented with complete EDTA-free protease inhibitors, DNaseI and Lysozyme. The complex was isolated from crude lysate by Nickel affinity chromatography using a HiTrap IMAC column and eluted by a linear gradient to 500 mM imidazole. The complex was dialyzed in Heparin binding buffer

containing 50 mM HEPES pH 7.5, 200 mM NaCl and 2 mM DTT, and further purified over a HiTrap Heparin column, followed by size exclusion chromatography over a Superdex 200 26/600 column in gel filtration buffer containing 10 mM HEPES pH 7.5, 200 mM NaCl and 2 mM DTT.

The purifications of the other *Hs* CCR4-NOT complex components have been described previously (Chen et al., 2014; Petit et al., 2012; Raisch et al., 2016; Sgromo et al., 2017).

The *Dm* NOT4 CBM peptide (residues D813-Q838) for crystallization was obtained chemically synthesized from EMC microcollections and solubilized in buffer containing 10 mM HEPES pH 7.5, 200 mM NaCl and 2 mM DTT.

In vitro MBP pulldown assays

For *in vitro* pulldown assays in Fig. 1B, Fig. 2B-C and Fig. 3A, full-length *Hs* NOT4 and fragments thereof were expressed as fusion proteins with N-terminal MBP and C-terminal GB1-6xHis tags. The cells were lysed in binding buffer containing 50 mM HEPES pH 7.5, 300 mM NaCl, 20 mM imidazole and 2 mM β-mercaptoethanol, supplemented with complete EDTA-free protease inhibitors, Lysozyme and DNaseI. The proteins were isolated from the lysate using Ni-NTA resin (Quiagen), followed by elution with binding buffer supplemented with 500 mM imidazole. Then, the NOT4 fragments were immobilized on amylose resin and subsequently incubated with purified interactors for 1h. Finally, the amylose beads were washed five times with binding buffer, and the proteins were eluted with binding buffer supplemented with 25 mM D-(+)-maltose. The eluted proteins were analyzed by SDS-PAGE.

For pulldowns with *Hs* and *Dm* NOT4 CBM constructs, purified MBP (20 µg) or MBP-tagged NOT4 fragments (40 µg) were incubated with equimolar amounts of the respective purified interactors and amylose resin in a 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 the same

buffer and the proteins were eluted with the same buffer supplemented with 25 mM D-(+)-Maltose. The eluted proteins were precipitated with trichloroacetic acid, and analyzed by SDS-PAGE.

Tethering assays in human HEK293T cells

For tethering assays, $0.7x10^6$ HEK293T cells were seeded in 6-well plates and transfected on the following day using Lipofectamine 2000 (Invitrogen). The transfection mixtures contained 0.5 μg of the control plasmid encoding the β-globin gene fused to the GAPDH sequences but lacking MS2 binding sites (Control; β-globin-GAP), 0.5 μg of the β-globin reporter containing six MS2-binding sites (β-globin-6xMS2bs) and various amounts of pT7-MS2-HA plasmids (MS2-HA: 0.5 μg; MS2-HA-NOT4-N: 0.1 μg; MS2-HA-NOT4-C: 1.5 μg) for the expression of MS2-HA-fusion proteins. The cells were harvested two days after transfection. The total RNA was isolated using the Trifast Reagent (Peqlab) and analyzed by Northern blot.

For tethering assays with Luciferase reporters, the transfection mixtures contained 0.2 μg of the control plasmid (F-Luc-GFP) and 0.2 μg of the reporter plasmid containing or lacking six MS2-binding sites (R-Luc-6xMS2bs, R-Luc or R-Luc-6xMS2bs-A95-MALAT) and various amounts of pT7-MS2-HA plasmids as described before for the expression of MS2-HA-NOT4 fusion proteins. The cells were harvested 2 days after transfection. Firefly and Renilla luciferase activities were measured using a Dual-Luciferase Reporter Assay system (Promega). Northern blotting was performed as previously described (Behm-Ansmant et al., 2006). For the experiment shown in Fig. S3G,H, cells were co-transfected with 0.5 μg of a plasmid expressing either wild-type GFP-DCP2 or the GFP-DCP2 E148Q mutant.

Knockdown in human HEK293T cells

For knockdown of NOT1 in *Hs* HEK293T cells using short-hairpin RNAs (shRNAs), target sequences (control: ATTCTCCGAACGTGTCACG; NOT1: ATTCAACATTCCCTTATA) were cloned in the pSUPER plasmid (a kind gift from O. Mühlemann) containing a puromycin selection marker. Cells were transfected in 12-well plates using Lipofectamine 2000. The transfection mixtures contained 2 μg of plasmids expressing the respective shRNAs. After 1 day, cells were selected for 24 hours in DMEM supplemented with 1.5 μg/ml puromycin and subsequently seeded in 12-well plates in medium without puromycin. The following day, the cells were transfected with the reporter plamsids and the respective shRNA plasmids. The transfection mixture contained 0.1 μg of R-Luc-6xMS2bs, 0.1 μg of pEGFP-N3-F-Luc transfection control and 1 μg of the respective shRNA-expressing plasmids. 24 hours after transfection, cells were selected with puromycin (1.5μg/ml). Cells were harvested and luciferase activities were measured 72 hours after the second transfection.

Tethering assays in Drosophila S2 cells

For the λN-tethering assay in *Dm* S2 cells, 2.5×10^6 S2 cells were seeded in 6-well plates and transfected with Effectene (Qiagen). The transfection mixtures contained the following plasmids: 0.1 μg of the F-Luc-5BoxB reporter plasmid, 0.4 μg of R-Luc-A_{90}-HhR and different amounts of either a plasmid expressing λN-HA or plasmids expressing λN-HA-*Dm*NOT4 proteins (λN-HA: 80ng; NOT4-fl: 10 ng; NOT4-N: 80 ng; NOT4-C: 10 ng). Three days after transfection, cells were harvested and luciferase activities were measured and Northern blotting was performed as described in the previous section.

For the experiment shown in Fig. S3I,J cells were co-transfected with 1 μg of a plasmid expressing either wild-type DCP1 or the R70G-L71S-N72S-T73G mutant.

SBP pulldown assays

HEK293T cells were seeded in 10 cm plates ($4x10^6$ cells/plate), and transfected after 1 day using Turbofect (Thermo Scientific). Two days after transfection, the cells were lysed on ice for 15 minutes in 1 ml NET lysis buffer (50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.1 % Triton X-100, 1 mM EDTA and 10 % (v/v) glycerol) supplemented with protease inhibitors (Roche), then centrifuged at 20,000 x g at 4°C for 15 min. The cleared lysate was treated with 100 µg/ml RNase A (Qiagen) for 30 min then centrifuged again at 20,000 x g at 4°C for 15 min. The lysates were incubated for 1 hour at 4°C with 50 µl Streptavidin sepharose resin (GE Healthcare). The beads were washed with NET buffer 3 times. The samples were analyzed by Western blot, using antibodies listed in Supplemental table S2. The Western blots were developed using the ECL Western blotting detection system (GE Healthcare).

For the experiment shown in Fig. 6B, 5 µg purified GST or the indicated amounts of purified recombinant *Hs* CAF40 ARM was added to the cell lysate and incubated for 1 hour prior to the incubation with the Streptavidin sepharose resin.

Crystallization

The *Hs* CAF40 ARM domain was mixed with a twofold molar excess of the *Dm* NOT4 CBM peptide. Initial screens were carried out in sitting drops at 22 °C by mixing 200 nl protein solution (6 mg/ml CAF40 and 1.2 mg/ml NOT4) with 200 nl crystallization condition solution. Crystals appeared within one day in many conditions. Optimized crystals of crystal form 2 grew at 18 °C in hanging drops consisting of 1 µl protein solution 6 mg/ml CAF40 and 1.2 mg/ml NOT4) and 1 µl crystallization condition containing 0.9 M K_2HPO_4 and 0.3 M NaH_2PO_4. Crystals were cryoprotected with 4 M sodium formate and flash-cooled in liquid nitrogen.

Crystals of crystal form 1 appeared in the initial screen in a condition containing 0.2 M Ammonium sulfate, 0.1 M Bis-Tris pH 5.5 and 25 % (w/v) PEG 3,350. The crystals were

cryoprotected with crystallization condition supplemented with 25 % glycerol and flash-cooled in liquid nitrogen.

Data collection and structure determination

X-ray diffraction data of the *Dm* NOT4-*Hs* CAF40 complex were collected on a PILATUS 6M detector (Dectris) at the PXII beamline of the Swiss Light Source (SLS). The best datasets were recorded at a wavelength of 0.99994 Å and processed XDS and XSCALE (Kabsch, 2010). The best dataset of crystal form 1 (space group $P2_12_12$) extended to 2.1 Å. Four copies of the CAF40 ARM domain (PDB 2FV2, chain A) were found in the asymmetric unit by molecular replacement with PHASER (McCoy et al., 2007) from the CCP4 package (Winn et al., 2011). 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). Four copies of the NOT4 CBM peptide were manually modelled into the density and improved by further refinement cycles. The final model was refined to R_{work} = 19.0 % and R_{free} = 21.8 %.

The best dataset of crystal form 2 (space group $I2_12_12_1$) extended to 2.2 Å. Initial phase information was obtained by molecular replacement using the CAF40 ARM domain (PDB 2FV2, chain A) as a search model. This initial model comprising four copies of CAF40 was refined using PHENIX and BUSTER (Bricogne et al., 2011); four copies of NOT4 were manually modelled and improved by further refinement cycles. The final model was refined to R_{work} = 19.3 % and R_{free} = 23.0 %.

Knockout of CAF40 in HEK293T cells using CRISPR-Cas9

An sgRNA (Sequence: 5’ CACCGCCCATGCTGTGGCATTCATT 3’) to target the second exon of *Hs* CAF40 was designed using CHOPCHOP (http://chopchop.cbu.uib.no) and cloned into the pSpCas9(BB)-2A-Puro (PX459) vector (Addgene plasmid 48139). 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. A single colony of knockout cells was obtained by serial dilution and expansion of single-cell clones under non-selective conditions. Positive clones were validated by Western blot analysis and DNA sequencing of the targeted genomic region. Sequencing the genomic region of the positive clones showed a 22-nt deletion in the targeted exon of CAF40 that causes a frameshift.

Accession numbers

The NOT4-CAF40 structures were deposited in the Protein Data Bank under accession codes XXXX (crystal form 1) and XXXX (crystal form 2).

Acknowledgments

We are grateful to Heike Budde and Catrin Weiler for excellent technical support. We thank the staff at the PX beamlines of the Swiss Light Source, Villigen for assistance with x-ray data collection. This work was supported by the Max Planck Society.

References

Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., and Adams, P.D. (2012). Towards automated crystallographic structure refinement with phenix.refine. Acta crystallographica Section D, Biological crystallography *68*, 352-367.

Albert, T.K., Hanzawa, H., Legtenberg, Y.I., de Ruwe, M.J., van den Heuvel, F.A., Collart, M.A., Boelens, R., and Timmers, H.T. (2002). Identification of a ubiquitin-protein ligase subunit within the CCR4-NOT transcription repressor complex. The EMBO journal *21*, 355-364.

Bai, Y., Salvadore, C., Chiang, Y.C., Collart, M.A., Liu, H.Y., and Denis, C.L. (1999). The CCR4 and CAF1 proteins of the CCR4-NOT complex are physically and functionally separated from NOT2, NOT4, and NOT5. Molecular and cellular biology *19*, 6642-6651.

Bawankar, P., Loh, B., Wohlbold, L., Schmidt, S., and 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 biology *10*, 228-244.

Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P., and Izaurralde, E. (2006). mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. Genes & development *20*, 1885-1898.

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

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

Boland, A., Chen, Y., Raisch, T., Jonas, S., Kuzuoglu-Ozturk, D., Wohlbold, L., Weichenrieder, O., and Izaurralde, E. (2013). Structure and assembly of the NOT module of the human CCR4-NOT complex. Nature structural & molecular biology *20*, 1289-1297.

Chang, C.T., Bercovich, N., Loh, B., Jonas, S., and 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 research *42*, 5217-5233.

Chen, J., Rappsilber, J., Chiang, Y.C., Russell, P., Mann, M., and Denis, C.L. (2001). Purification and characterization of the 1.0 MDa CCR4-NOT complex identifies two novel components of the complex. Journal of molecular biology *314*, 683-694.

Chen, Y., Boland, A., Kuzuoglu-Ozturk, D., Bawankar, P., Loh, B., Chang, C.T., Weichenrieder, O., and Izaurralde, E. (2014). A DDX6-CNOT1 complex and W-binding pockets in CNOT9 reveal direct links between miRNA target recognition and silencing. Molecular cell *54*, 737-750.

Cooke, A., Prigge, A., and Wickens, M. (2010). Translational repression by deadenylases. The Journal of biological chemistry *285*, 28506-28513.

Cooper, K.F., Scarnati, M.S., Krasley, E., Mallory, M.J., Jin, C., Law, M.J., and Strich, R. (2012). Oxidative-stress-induced nuclear to cytoplasmic relocalization is required for Not4-dependent cyclin C destruction. Journal of cell science *125*, 1015-1026.

Diebold, M.L., Fribourg, S., Koch, M., Metzger, T., and Romier, C. (2011). Deciphering correct strategies for multiprotein complex assembly by co-expression: application to complexes as large as the histone octamer. Journal of structural biology *175*, 178-188.

Dimitrova, L.N., Kuroha, K., Tatematsu, T., and Inada, T. (2009). Nascent peptide-dependent translation arrest leads to Not4p-mediated protein degradation by the proteasome. The Journal of biological chemistry *284*, 10343-10352.

Garces, R.G., Gillon, W., and Pai, E.F. (2007). Atomic model of human Rcd-1 reveals an armadillo-like-repeat protein with in vitro nucleic acid binding properties. Protein science : a publication of the Protein Society *16*, 176-188.

Gulshan, K., Thommandru, B., and Moye-Rowley, W.S. (2012). Proteolytic degradation of the Yap1 transcription factor is regulated by subcellular localization and the E3 ubiquitin ligase Not4. The Journal of biological chemistry *287*, 26796-26805.

Kabsch, W. (2010). Xds. Acta crystallographica Section D, Biological crystallography *66*, 125-132.

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

Lau, N.C., Kolkman, A., van Schaik, F.M., Mulder, K.W., Pijnappel, W.W., Heck, A.J., and Timmers, H.T. (2009). Human Ccr4-Not complexes contain variable deadenylase subunits. The Biochemical journal *422*, 443-453.

Lykke-Andersen, J., Shu, M.D., and Steitz, J.A. (2000). Human Upf proteins target an mRNA for nonsense-mediated decay when bound downstream of a termination codon. Cell *103*, 1121-1131.

Mathys, H., Basquin, J., Ozgur, S., Czarnocki-Cieciura, M., Bonneau, F., Aartse, A., Dziembowski, A., Nowotny, M., Conti, E., and Filipowicz, W. (2014). Structural and biochemical insights to the role of the CCR4-NOT complex and DDX6 ATPase in microRNA repression. Molecular cell *54*, 751-765.

Matsuda, R., Ikeuchi, K., Nomura, S., and Inada, T. (2014). Protein quality control systems associated with no-go and nonstop mRNA surveillance in yeast. Genes to cells : devoted to molecular & cellular mechanisms *19*, 1-12.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. Journal of applied crystallography *40*, 658-674.

Mersman, D.P., Du, H.N., Fingerman, I.M., South, P.F., and Briggs, S.D. (2009). Polyubiquitination of the demethylase Jhd2 controls histone methylation and gene expression. Genes & development *23*, 951-962.

Mulder, K.W., Inagaki, A., Cameroni, E., Mousson, F., Winkler, G.S., De Virgilio, C., Collart, M.A., and Timmers, H.T. (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., and Bottcher, B. (2011). Insights into the structure of the CCR4-NOT complex by electron microscopy. FEBS letters *585*, 2182-2186.

Panasenko, O., Landrieux, E., Feuermann, M., Finka, A., Paquet, N., and Collart, M.A. (2006). The yeast Ccr4-Not complex controls ubiquitination of the nascent-associated polypeptide (NAC-EGD) complex. The Journal of biological chemistry *281*, 31389-31398.

Panasenko, O.O. (2014). The role of the E3 ligase Not4 in cotranslational quality control. Frontiers in genetics *5*, 141.

Panasenko, O.O., and Collart, M.A. (2011). Not4 E3 ligase contributes to proteasome assembly and functional integrity in part through Ecm29. Molecular and cellular biology *31*, 1610-1623.

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

Petit, A.P., Wohlbold, L., Bawankar, P., Huntzinger, E., Schmidt, S., Izaurralde, E., and 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 research *40*, 11058-11072.

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

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

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

Stowell, J.A., Webster, M.W., Kogel, A., Wolf, J., Shelley, K.L., and Passmore, L.A. (2016). Reconstitution of Targeted Deadenylation by the Ccr4-Not Complex and the YTH Domain Protein Mmi1. Cell reports *17*, 1978-1989.

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

Ukleja, M., Cuellar, J., Siwaszek, A., Kasprzak, J.M., Czarnocki-Cieciura, M., Bujnicki, J.M., Dziembowski, A., and Valpuesta, J.M. (2016). The architecture of the Schizosaccharomyces pombe CCR4-NOT complex. Nature communications *7*, 10433.

Valdar, W.S. (2002). Scoring residue conservation. Proteins *48*, 227-241.

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

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A., *et al.* (2011). Overview of the CCP4 suite and

current developments. Acta crystallographica Section D, Biological crystallography *67*, 235-242.

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

Table 1. Data Collection and Refinement Statistics

	CNOT9-DmNot4	
	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* (Å)	83.9, 109.6, 69.7	85.6, 90.3, 197.0
Angles α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Data Collection*		
Wavelength (Å)	0.99994	0.99994
Resolution range (Å)	50-2.1 (2.14-2.10)	50-2.2 (2.25–2.20)
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)
Unique Reflections	38,143 (2,780)	39,129 (2,831)
Multiplicity	4.0 (4.2)	8.9 (9.0)
Refinement		
R_{work} (%)	19.0	19.3
R_{free} (%)	21.8	23.0
Number of atoms		
All atoms	4971	4805
Protein	4760	4738
Ligands	34	3
Water	177	64
Average B factor (Å^2)		
All atoms	52.7	81.0
Protein	52.3	81.1
Ligands	95.5	93.7
Water	53.7	66.9
Ramachandran plot		
Favored regions, %	99.2	98.3
Disallowed Regions, %	0.0	0.0
RMSD from ideal geometry		
Bond lengths (Å)	0.002	0.010
Bond angles (°)	0.426	1.040

*Values in parentheses are for highest-resolution shell.

Supplemental table S1. Constructs used in this study

Name	**NOT4**
NOT4	MS2-HA-*Hs* NOT4iso1 (Uniprot O95628-1**)**
	V5-SBP-MBP-*Hs* NOT4iso1
	MBP-*Hs* NOT4iso1-GB1-6xHis
	λN-HA-*Dm* NOT4isoL (Uniprot M9PCL9)
	Dm NOT4isoL-GFP
NOT4-N	MS2-HA-*Hs* NOT4iso1_1-274
	V5-SBP-MBP-*Hs* NOT4iso1_1-274
	MBP-*Hs* NOT4iso1_1-274-GB1-6xHis
	λN-HA-*Dm* NOT4isoL_1-249
	GFP-*Dm* NOT4isoL_1-249
NOT4-C	MS2-HA-*Hs* Not4iso1_275-575
	V5-SBP-MBP-*Hs* Not4iso1_275-575
	MBP-*Hs* NOT4iso1_275-575-GB1-6xHis
	λN-HA-*Dm* NOT4isoL_248-1050
	GFP-*Dm* NOT4isoL_248-1050
275-376	MS2-HA-*Hs* NOT4iso1_275-376
	MBP-*Hs* NOT4iso1_375-376-GB1-6xHis
NCBR	MS2-HA-*Hs* NOT4iso1_377-428
	MBP-*Hs* NOT4iso1_377-428-GB1-6xHis
429-575	MS2-HA-*Hs* NOT4iso1_429-575
	MBP-*Hs* NOT4iso1_429-575-GB1-6xHis
ΔNCBR	MS2-HA-*Hs* NOT4iso1_275-575 Δ424
	V5-SBP-MBP-*Hs* Not4iso1_275-575 Δ424
	MBP-*Hs* NOT4iso1_275-575 Δ424-GB1-6xHis
Δ377-402	MS2-HA-*Hs* NOT4iso1_275-575 Δ377-402
	V5-SBP-MBP-*Hs* Not4iso1_275-575 Δ377-402
	MBP-*Hs* NOT4iso1_275-575 Δ377-402-GB1-6xHis
ΔCBM	MS2-HA-*Hs* NOT4iso1_275-575 Δ400-428
	V5-SBP-MBP-*Hs* Not4iso1_275-575 Δ400-428
	MBP-*Hs* NOT4iso1_275-575 Δ400-428-GB1-6xHis
CBM	MBP-*Hs* NOT4iso1_400-427
	MBP-*Dm* NOT4isoL_813-838
F824D	MBP-*Dm* NOT4isoL_813-838 F824D
L832E	MBP-*Dm* NOT4isoL_813-838 L832E
F405D	V5-SBP-MBP-*Hs* NOT4iso1_275-575 F405D
F408A	V5-SBP-MBP-*Hs* NOT4iso1_275-575 F408A
T411E	V5-SBP-MBP-*Hs* NOT4iso1_275-575 T411E
L415E	V5-SBP-MBP-*Hs* NOT4iso1_275-575 L415E
2x mut	V5-SBP-MBP-*Hs* NOT4iso1_275-575 T411E L415E
NOT4-C 4x mut	V5-SBP-MBP-*Hs* NOT4iso1_275-575 F405D F408A T411E L415E
	MS2-HA-*Hs* NOT4iso1_275-575 F405D F408A T411E L415E
NCBR 4x mut	MS2-HA-*Hs* NOT4iso1_377-428 F405D F408A T411E L415E
Name	**CAF40**
CAF40-ARM	*Hs* CAF40_19-285 (Uniprot Q92600)
CAF40	V5-SBP-MBP-*Hs* CAF40 shRNAres2

HA-CAF40	HA-*Dm* CAF40 (Uniprot Q7JVP2)
GFP-CAF40	GFP-*Dm* CAF40 dsRNAres
CAF40 4x mut	V5-SBP-MBP-*Hs* CAF40 H58AF60AA64YV71Y-shRNAres2
V181E	V5-SBP-MBP-*Hs* CAF40 shRNAres2 V181E
	Hs CAF40_19-285 V181E
V186E	GFP-*Dm* CAF40 dsRNAres V186E
2x mut	*Hs* CAF40_19-285 Y134D+G141W
Name	***Dm* NOT1** (Uniprot A5YKK6)
HA-CN9BD	HA-*Dm* NOT1_1467-1717
HA-NOT1	HA-*Dm* NOT1 FL

Supplemental table S2. Antibodies used in this study

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-NOT4	Abcam	ab72049	1:1000	Rabbit polyclonal
Anti-CAF40	Proteintech	22503-1-AP	1:1000	Rabbit polyclonal
Anti-V5	LSBio	LS-C57305	1:5000	Mouse monoclonal
Anti-HA-HRP	Roche	12 013 819 001	1:5000	Monoclonal
Anti-GFP	Roche	11 814 460 001	1:2000	Mouse monoclonal
Anti-tubulin	Sigma	T6199	1:5000	Mouse monoclonal
Anti-PABP	Abcam	ab21060	1:10000	Rabbit polyclonal

Figure legends

Figure 1. Metazoan NOT4 directly interacts with the CCR4-NOT complex and induces degradation of bound mRNA.

(A) Schematic representations of NOT4 proteins. NOT4 comprises a RING-type E3 ubiquitin ligase domain, a predicted coiled coil region, an RNA-recognition motif (RRM) domain and a C3H1-type Zinc-finger domain (ZnF) in its conserved N-terminal region. The divergent C-terminal regions are predicted to be unstructured. *Hs* NOT4-C contains a NOT-module-and-CAF40-binding-region (NCBR) that includes a CAF40-binding motif (CBM). *Dm* NOT4 carries a CBM as well. *Sc* NOT4 harbors a binding site for NOT1 in its C-terminal region.

(B) MBP pulldown assay showing the interaction of MBP-tagged *Hs* NOT4 and NOT4-C with a pentameric complex including the C-terminal half of NOT1, the NOT2 and NOT3 C-terminal regions, the armadillo domain of CAF40 and full-length CAF1.

(C-E) Tethering assay in human HEK293T cells using a β-globin reporter containing 6 binding sites for the MS2 protein (β-globin-6xMS2bs) and MS2-HA-tagged *Hs* NOT4 or *Hs* NOT4 fragments. A β-globin-GAPDH reporter lacking MS2 binding sites was used as transfection control. The β-globin-6xMS2bs reporter levels were normalized to this control and set to 100 in the presence of MS2-HA. Mean values ± standard deviations from three independent experiments are shown in (**C**). **(D)** Northern blot of representative RNA samples. **(E)** Western blot showing the equivalent expression of MS2-HA-tagged proteins used in the tethering assay.

(F-H) Tethering assay in *Dm* S2 cells using an F-Luc-5BoxB reporter and the indicated λN-HA-tagged proteins. A plasmid expressing R-Luc served as a transfection control. The F-Luc-5BoxB reporter activity (white bars) and mRNA levels (black bars) were normalized to the R-Luc control and set to 100 in the presence of λN-HA. Mean values ± standard deviations from three independent experiments are shown in **(F)**. **(G)** Northern blot of representative RNA

samples. **(H)** Western blot showing the equivalent expression of λN-HA-tagged proteins used in the tethering assay.

Figure 2. *Hs* NOT4 directly interacts with NOT1-SHD and CAF40.

(A) Schematic representation of *Hs* NOT1 and the other CCR4-NOT subunits used in this study. NOT1 contains N-terminal HEAT repeat domains (shown in yellow) that interact with NOT10 and NOT11, a central MIF4G domain (light blue) that binds CAF1, a CN9BD (dark blue) that recruits CAF0, a connector domain (CD, green) and a NOT1 superfamily homology domain (N1SHD, yellow) that forms the NOT module together with NOT2 and NOT3.

(B) MBP pulldown assay testing the interaction of MBP-tagged *Hs* NOT4 with the following *Hs* CCR4-NOT subcomplexes: The NOT1 N-terminal fragment with NOT10 and NOT11, the NOT1 MIF4G domain with CAF1, the NOT1 CN9BD with CAF40, the NOT1 connector domain (CD), and the assembled NOT module (containing the NOT1-SHD and the NOT2 and NOT3 C-terminal fragments). SM: Starting material.

(C) MBP pulldown assay testing the interaction of MBP-tagged *Hs* NOT4 and the *Hs* NOT1-SHD, the NOT2–NOT3 dimer (NOT2–3) or CAF40. SM: Starting material.

Figure 3. The NCBR directly interacts with NOT1 and CAF40 and promotes target mRNA degradation.

(A) MBP pulldown assay testing the interaction of MBP-tagged *Hs* NOT4-C or fragments (residues 275-376, NCBR, 429-575) with the NOT module or CAF40. SM: Starting material.

(B-D) Tethering assay in human HEK293T cells using the β-globin-6xMS2bs reporter and MS2-HA-tagged *Hs* NOT4-C or fragments (residues 275-376, NCBR, 429-575). The β-globin-6xMS2bs reporter levels were normalized to those of the control mRNA and set to 100 in the presence of MS2-HA as described in Fig 1C. Mean values ± standard deviations from three independent experiments are shown in (**B)**. **(C)** Northern blot of representative RNA

samples. **(D)** Western blot showing the equivalent expression of MS2-HA-tagged proteins used in the tethering assay.

(E) SBP pulldown assay testing the interaction of V5-SBP-tagged *Hs* NOT4-C or deletion constructs (ΔNCBR, Δ377-402 or ΔCBM) with endogenous NOT1, NOT3, NOT2 and CAF40 in HEK293T cell lysates.

(F-H) Tethering assay using MS2-HA-tagged *Hs* NOT4-C or deletion constructs in human HEK293T cells as mentioned in (B). Mean values ± standard deviations from three independent experiments are shown in (F). **(G)** Northern blot of representative RNA samples. **(H)** Western blot showing the equivalent expression of MS2-HA-tagged proteins used in the tethering assay.

Figure 4. Structure of the *Dm* NOT4 CBM bound to CAF40.

(A) The *Dm* NOT4 CBM peptide (blue, backbone shown in ribbon representation) bound to *Hs* CAF40 (grey). The helices of CAF40 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 *Dm* NOT4 CBM peptide bound to *Hs* CAF40. Selected secondary structure elements are labeled in black.

(C-D) The CBM of *Dm* NOT4 (C) and Roquin (D) bound to CAF40 in the same orientation. The CBMs are displayed in cartoon view, CAF40 as surface. The CAF40 residues interacting with the respective CBMs are indicated in yellow.

(E) Close-up view of the CAF40-NOT4 binding interface. Selected residues of CAF40 and NOT4 are shown as orange and blue sticks, respectively. Hydrogen bonds are indicated by red dashed lines. Residues mutated in this study are underlined.

(F) MBP pulldown assay testing the interaction of MBP-tagged *Dm* NOT4-CBM with *Hs* CAF40 (wild-type or the indicated point mutants).

(G) MBP pulldown assay testing the interaction of MBP-tagged *Dm* NOT4-CBM (either wild-type or the indicated point mutants) with *Hs* CAF40.

Figure 5. The CBM is necessary for the mRNA degradative activity of *Hs* NOT4.

(A) MBP pulldown assay testing the interaction of MBP-tagged *Hs* NOT4-CBM with *Hs* CAF40 (either wild-type or the indicated point mutants).

(B) SBP pulldown assay testing the interaction of SBP-tagged *Hs* NOT4-C (either wild-type or a quadruple mutant, 4x mut) with endogenous *Hs* NOT1, NOT2, NOT3 and CAF40 in HEK293T cell lysates.

(C-E) Tethering assay using MS2-HA-tagged *Hs* NOT4-C or the NCBR (either wild-type or the 4x mut) in human HEK293T cells. The β-globin-6xMS2bs reporter levels were normalized to those of the control mRNA and set to 100 in the presence of MS2-HA. Mean values ± standard deviations from three independent experiments are shown in **(C)**. **(D)** Northern blot of representative RNA samples. **(E)** Western blot showing the equivalent expression of MS2-HA-tagged proteins used in the tethering assay.

Figure 6. CAF40 is required for NOT4-mediated mRNA-decay.

(A) Western blot showing the endogenous levels of CAF40, NOT1, NOT3 and NOT2 in the CAF0 knockout HEK293T cell line. Dilutions of control cell lysates were loaded in lanes (1–4) to estimate the efficacy of the depletion. Tubulin was used as a loading control.

(B) SBP pulldown assay testing the interaction of SBP-tagged *Hs* NOT4-C with endogenous *Hs* NOT1, NOT2 and NOT3 in wild-type and CAF40 knockout HEK293T cell lysates, in the presence of either 5 µg GST or the indicated amounts of purified *Hs* CAF40 ARM.

(C) Tethering assay using an R-Luc-6xMS2bs reporter and MS-HA tagged NOT4, either in wild-type HEK293T cells, in CAF40 knockout cells, NOT1 knockdown cells, or cells combining the CAF40 knockout and the NOT1 knockdown. R-Luc activity was normalized to

that of an F-Luc transfection control and set to 100 in cells expressing MS2-HA. Mean values ± standard deviations from three independent experiments are shown.

(D) Western blot showing the efficiency of the NOT1 knockdown in *Hs* HEK293T cells used in the experiment in (C). Dilutions of control cell lysates were loaded in lanes (1–4) to estimate the efficacy of the depletion. PABP served as a loading control.

(E) Western blot showing the expression of NOT4 and NOT4-C in the knockout and knockdown samples of the experiment shown in panel (C). GFP-F-Luc served as a transfection control.

Figure S1. Sequence alignment of vertebrate NOT4.

The secondary structure elements are indicated above the alignment. The CBM helix (as determined by the *Dm* NOT4-CAF40 structures) is indicated in red; the secondary structure elements of the RING and RRM domains according to the previously reported NMR structures (PDB 1UR6 and 2CPI, respectively) are shown in black; secondary structure elements of the structurally not characterized parts of NOT4 are indicated in black as predicted by PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/). Residues conserved in all aligned sequences are shown with a purple background, and residues with >70% similarity are highlighted in light pink; conservation scores were calculated using the SCORECONS webserver (Valdar, 2002). Residues that were mutated in this study are indicated by orange asterisks. Species abbreviations are as follows: *Hs*: *Homo sapiens*, human; *Ec*: *Equus caballus*, horse; *Gg*: *Gallus gallus*, chicken; *Xt*: *Xenopus tropicalis*, western clawed frog; *Dr*: *Danio rerio*, zebrafish.

Figure S2. Sequence alignment of *Drosophila* NOT4-C.

The secondary structure elements are indicated above the alignment. The CBM helix (as determined by the CAF40-NOT4 structures) is indicated in red; secondary structure elements

of the structurally not characterized parts of NOT4 are indicated in black as predicted by PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/). Residues conserved in all aligned sequences are shown with a dark blue background, and residues with >70% similarity are highlighted in light blue; conservation scores were calculated using the SCORECONS webserver (Valdar, 2002). Residues that were mutated in this study are indicated by orange asterisks. Residues in the CBM that directly contact CAF40 are marked by green diamonds.

Figure S3. NOT4 mediates mRNA degradation via the 5'-3' decay pathway as well as degradation-independent translational repression

(A,B) Tethering assay in *Hs* HEK293T cells using an R-Luc-6xMS2bs reporter, and MS2-HA-tagged *Hs* NOT4 and N- and C-terminal fragments. An F-Luc reporter without MS2 binding sites was used as transfection control. The R-Luc-6xMS2bs reporter levels were normalized to this control and set to 100 in the presence of MS2-HA. Luciferase activity values ('Protein') were also normalized to the F-Luc transfection control. Mean values ± standard deviations from three independent experiments are shown in **(A)**. **(B)** shows a representative northern blot.

(C,D) Tethering assay in *Hs* HEK293T cells using an R-Luc mRNA lacking MS2 binding sites and MS2-HA-tagged *Hs* NOT4 and N- and C-terminal fragments. Samples were processed as described in (A,B). The corresponding experiment with the R-Luc-6xMS2bs reporter is shown in (A,B).

(E,F) Tethering assay in *Dm* S2 cells using an F-Luc reporter lacking BoxB sites and the indicated λN-HA-tagged proteins. A plasmid expressing R-Luc served as a transfection control. The F-Luc reporter activity (white bars) and mRNA levels (black bars) were normalized to the R-Luc control and set to 100 in the presence of λN-HA. Mean values ± standard deviations from three independent experiments are shown in **(E)**. **(F)** Northern blot of representative RNA samples. The corresponding experiment with the F-Luc-5BoxB

reporter is shown in Fig. 1F-H.

(G,H) Tethering assay in *Hs* HEK293T cells using an R-Luc-6xMS2bs-A95-MALAT reporter which is resistant to deadenylation, and MS2-HA-tagged *Hs* NOT4 and N- and C-terminal fragments. An F-Luc reporter without MS2 binding sites was used as transfection control. The R-Luc-6xMS2bs reporter levels were normalized to this control and set to 100 in the presence of MS2-HA. Luciferase activity values ('Protein') were also normalized to the F-Luc transfection control. Mean values ± standard deviations from three independent experiments are shown in **(G)**. **(H)** shows a representative northern blot.

(I,J) Tethering assay in *Hs* HEK293T cells using a β-globin-6xMS2bs reporter and MS2-HA-tagged *Hs* NOT4 and Nanos2. In addition, either wild-type DCP2 or a catalytically inactive mutant (E148Q) was overexpressed. Samples were processed as described in (A,B).

(K-M) Tethering assay in *Dm* S2 cells using an F-Luc-5BoxB reporter and λN-HA-tagged *Dm* NOT4 and GW182. In addition, either wild-type DCP1 or a R70G-L71S-N72S-T73G mutant was overexpressed. An R-Luc reporter without BoxB elements was used as transfection control. The F-Luc-5BoxB reporter levels were normalized to this control and set to 100 in the presence of λN-HA. Mean values ± standard deviations from three independent experiments are shown in **(K)**. **(L)** shows a representative Northern blot. **(M)** Western blot showing the expression of the proteins used in the experiment in (I,J).

(N) MBP pulldown assay testing the interaction of MBP-tagged *Hs* NOT4-C or the indicated deletion constructs thereof with the NOT module or CAF40.

Figure S4. Molecular arrangement in the NOT4-CAF40 crystals and comparison with Roquin

(A) CAF40-NOT4 tetramer found in the asymmetric unit of the crystal in space group $P2_12_12$ (crystal form 1). The two CAF40 molecules are shown in grey (chain A) and black (chain C), respectively, and the NOT4 CBMs in blue (chain B) and pink (chain D).

(B) CAF40-NOT4 tetramer found in the asymmetric unit of the crystal in space group $I2_12_12_1$ (crystal form 2). Colors are chosen as described in panel (A).

(C) Superposition of CAF40-NOT4 heterotetramers of both crystal forms in ribbon representation. CAF40 from crystal form 1 is colored in light (chain A) and dark grey (chain C), NOT4 from the same crystal form in blue (chain B) and pink (chain D); the tetramer from crystal form 2 is shown in orange.

(D) Superposition of the two CAF40-NOT4 heterodimers in the asymmetric unit of crystal form 1 in ribbon representation. The CAF40 molecules are shown in grey (chain A) and yellow (chain C) while the NOT4 molecules are shown in blue (chain B) and orange (chain D), respectively.

(E) Superposition of the two CAF40-NOT4 heterodimers in the asymmetric unit of crystal form 2 in ribbon representation. The colors are chosen as described in panel (A).

(F) Superposition of CAF40-NOT4 heterodimers of both crystal forms in ribbon representation. CAF40 and NOT4 from crystal form 1 are colored in grey and blue, respectively, and CAF40 and NOT4 from crystal form 2 in black and cyan.

(G) Superposition of the CAF40-NOT4 heterodimer (crystal form 1) with the CAF40-Roquin structure (PDB 5LSW, Sgromo et al., 2017) in ribbon representation. The CAF40-NOT4 complex is colored as before. CAF40 from the Roquin structure is shown in purple, the Roquin CBM in green.

(H) Close-up view from the top onto the CBM-CAF40 interaction with CAF40 from the NOT4 structure (crystal form 1) shown as cartoon with helices depicted as cylinders and the CBM backbones of NOT4 (blue) and Roquin (green) shown as ribbons. Selected side chains of CAF40 and the CBMs are shown as sticks, highlighting the different orientation of CAF40 Y134 in both structures.

Figure S5. Disruption of the interaction with CAF40 impairs the degradative activity of

metazoan NOT4.

(A) SBP pulldown assay testing the interaction of V5-SBP-tagged *Hs* NOT4-C (wild-type or the indicated mutants) with endogenous NOT1 and CAF40 in HEK293T cell lysates.

(B,C) Tethering assay in *Drosophila* S2 cells using the F-Luc-5BoxB reporter and λN-HA-tagged *Dm* NOT4 and *Dm* NOT4-C. GFP or GFP-tagged *Dm* CAF40 (wild-type or the V186E mutant) was overexpressed. An R-Luc reporter without BoxB elements was used as transfection control. The F-Luc-5BoxB reporter levels were normalized to this control and set to 100 in the presence of λN-HA. Mean values ± standard deviations from three independent experiments are shown in (B). (C) shows a representative Northern blot.

(D) Western blot showing the expression of GFP, GFP-tagged *Dm* CAF40 (wild-type or the V186E mutant) and λN-HA-NOT4 (full-length or the C-terminus). The additional bands in lanes 1-3 marked by the asterisk correspond to GFP which also carries a V5 epitope tag.

Figure 1. Keskeny et al.

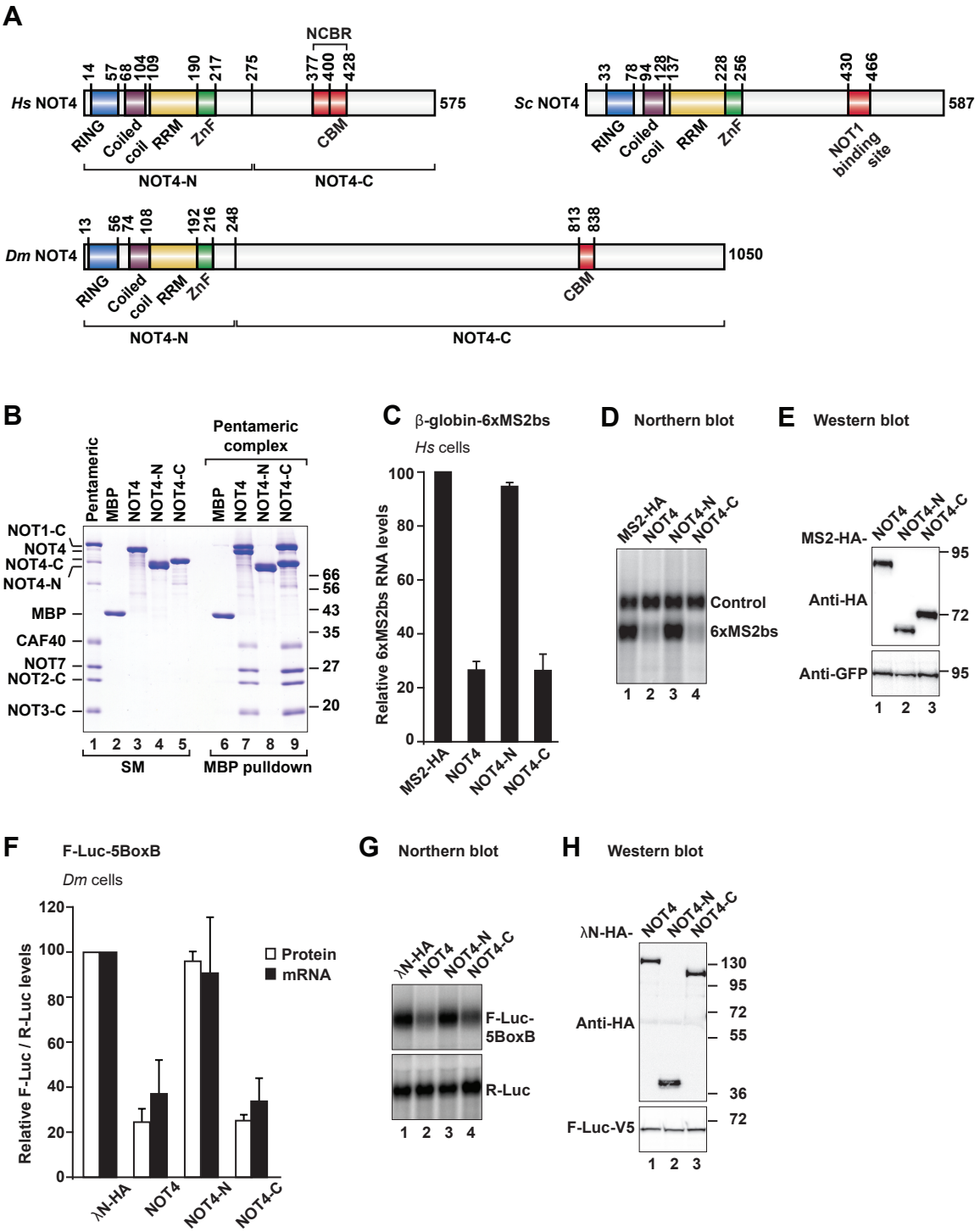


Figure 2. Keskeny et al.

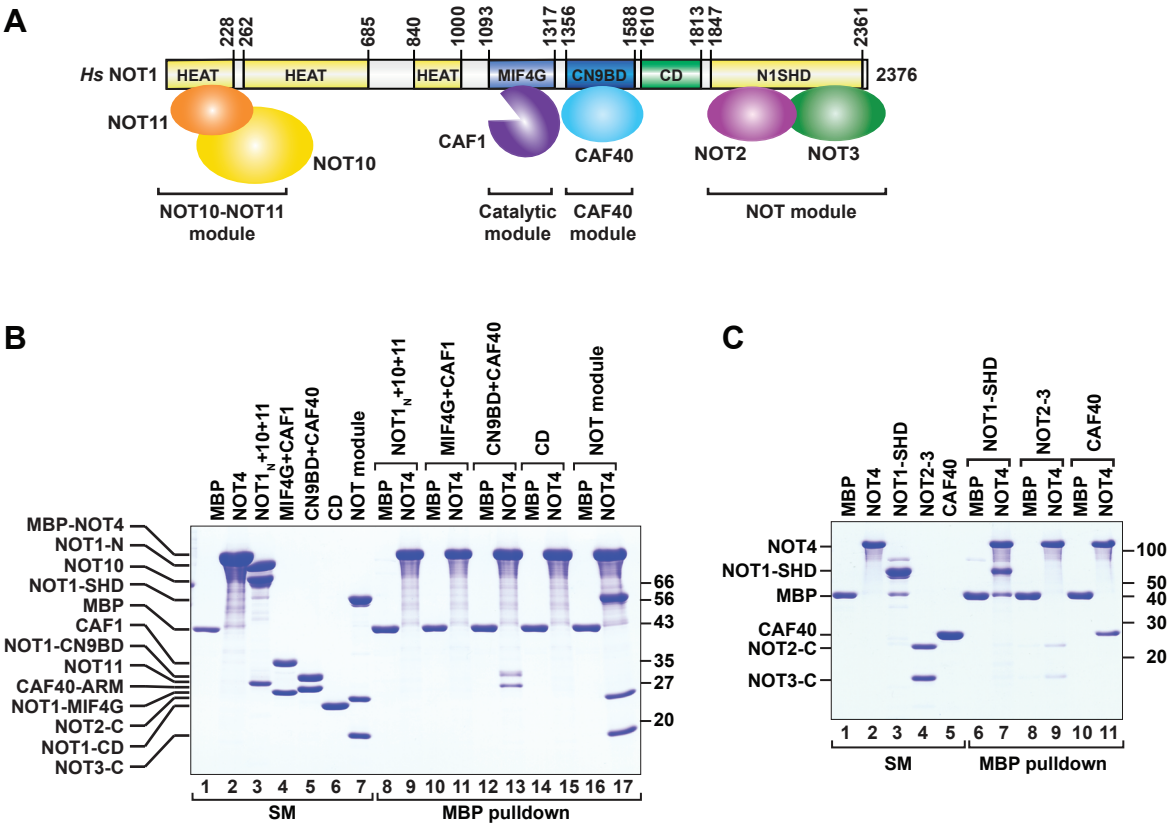


Figure 3. Keskeny et al.

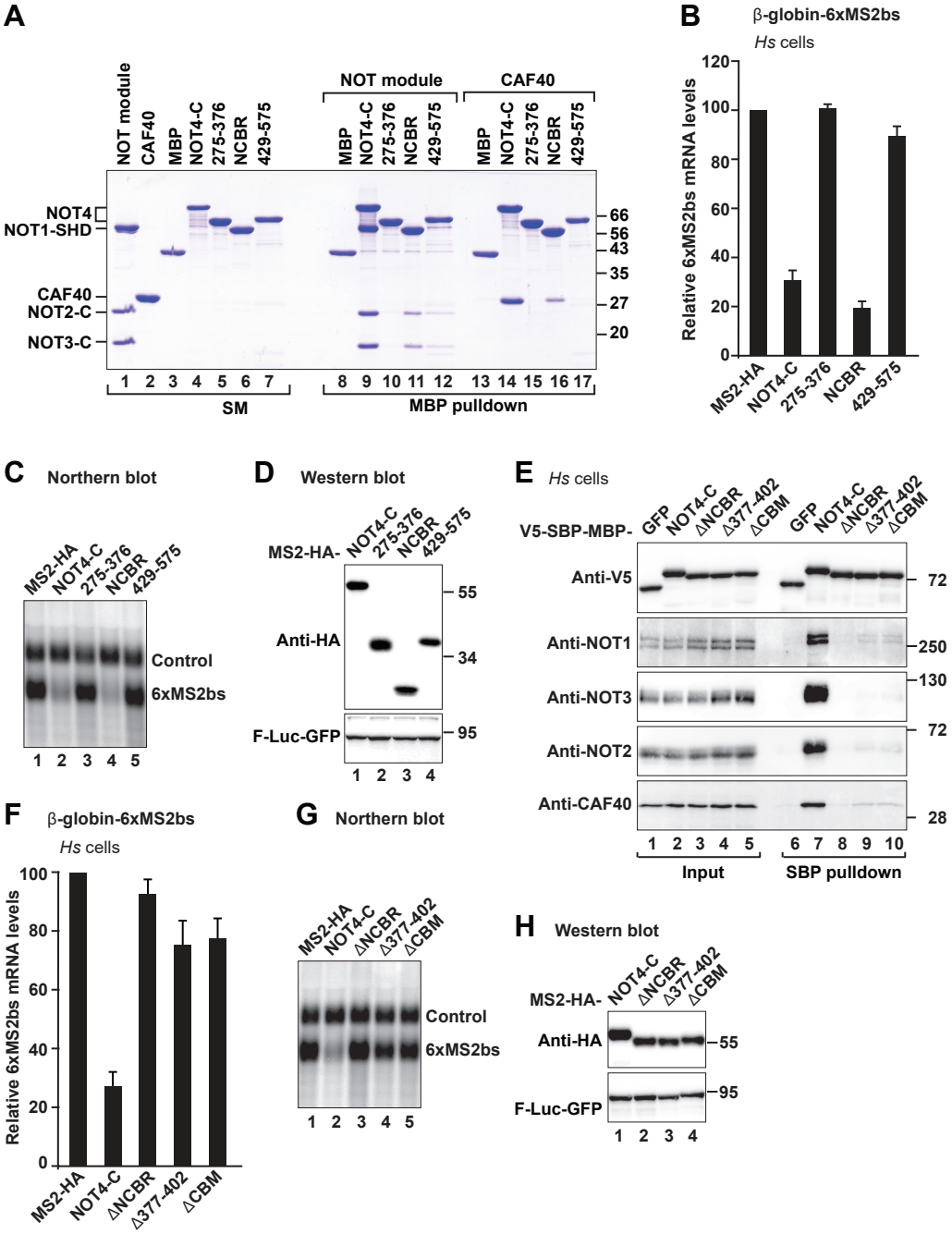
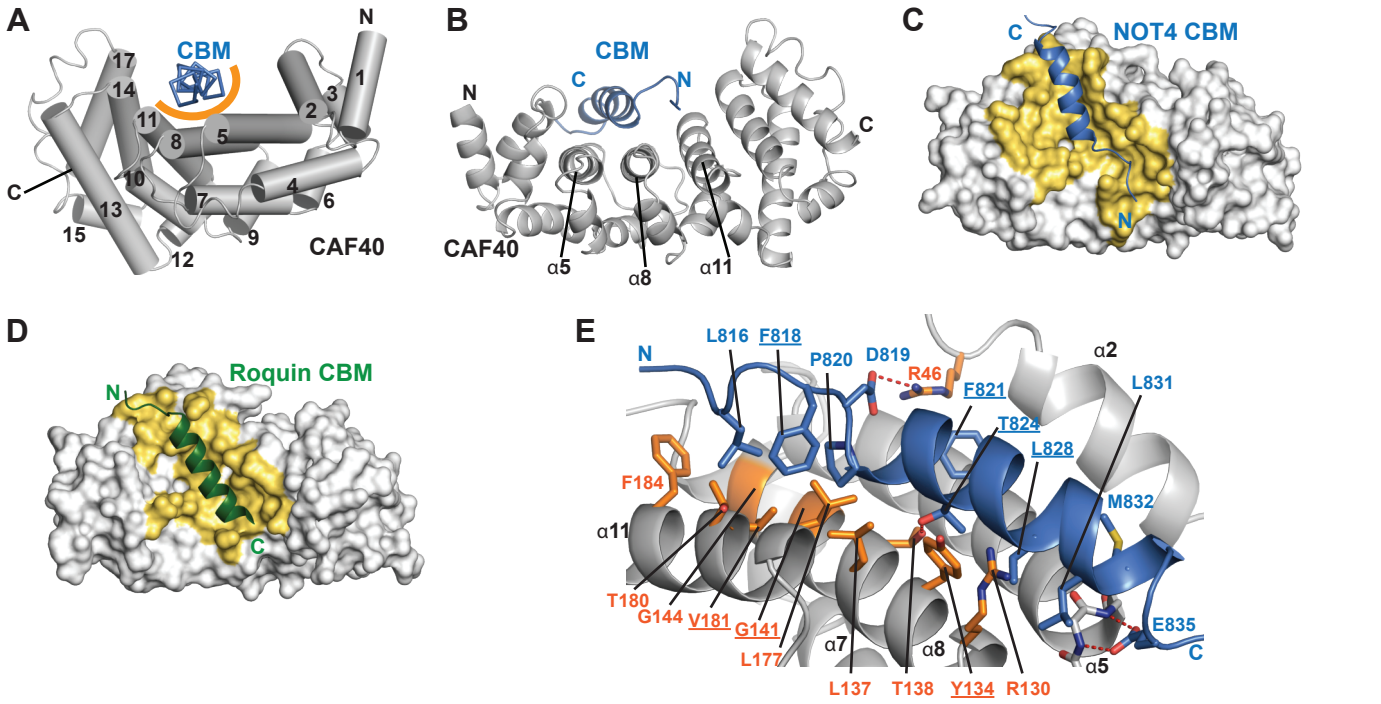
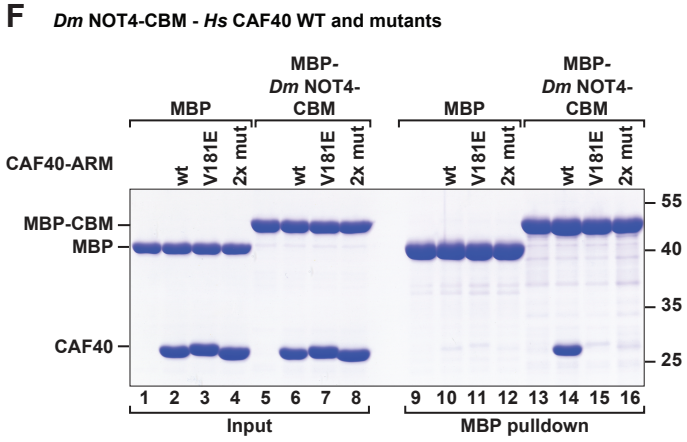


Figure 4. Keskeny et al.





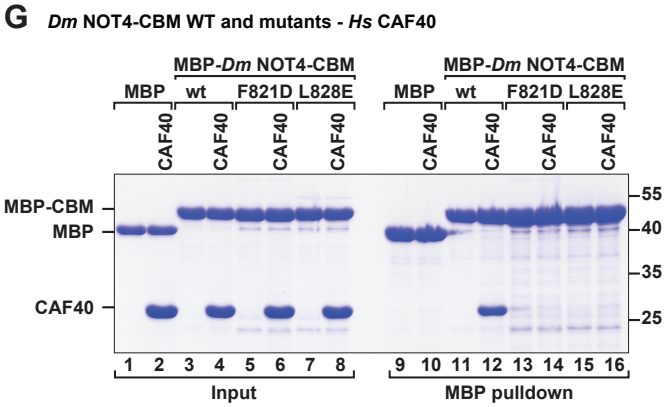


Figure 5. Keskeny et al.

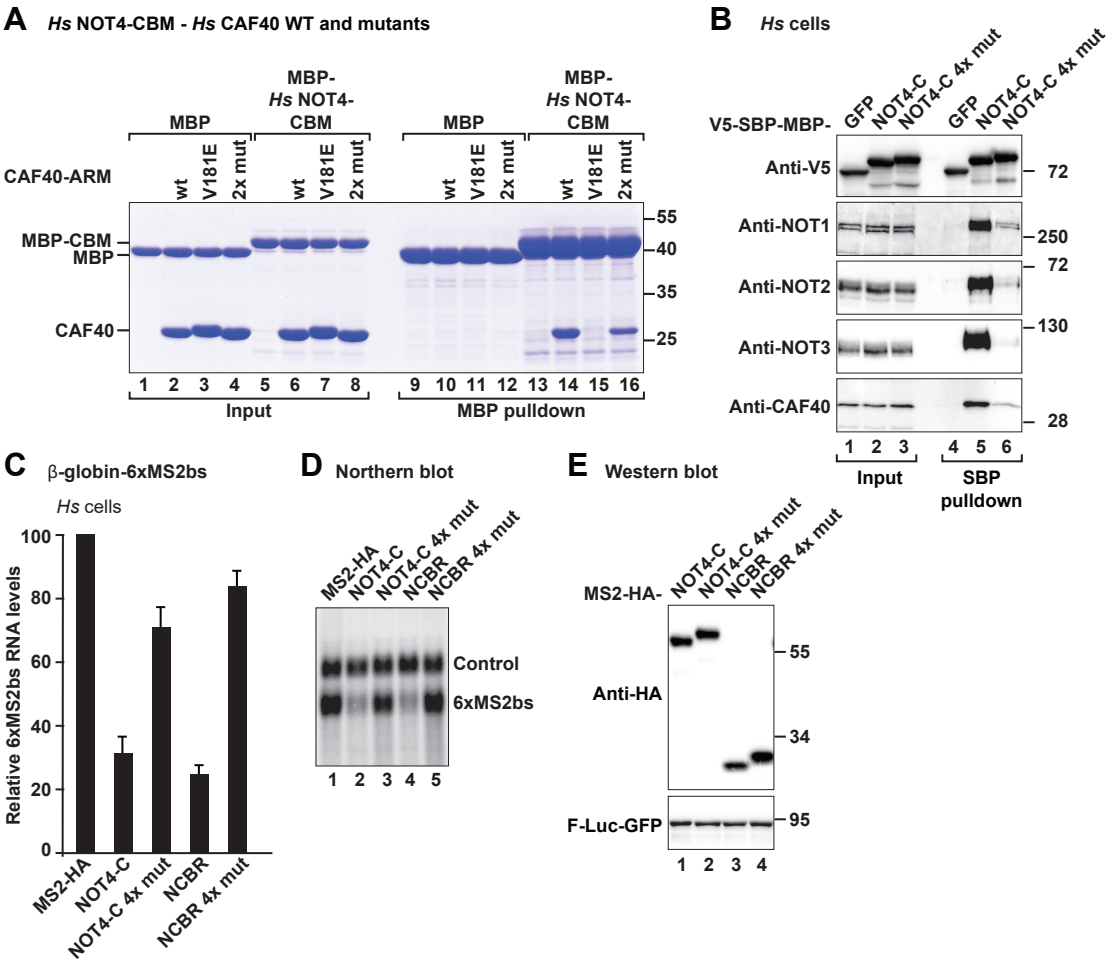


Figure 6. Keskeny et al.

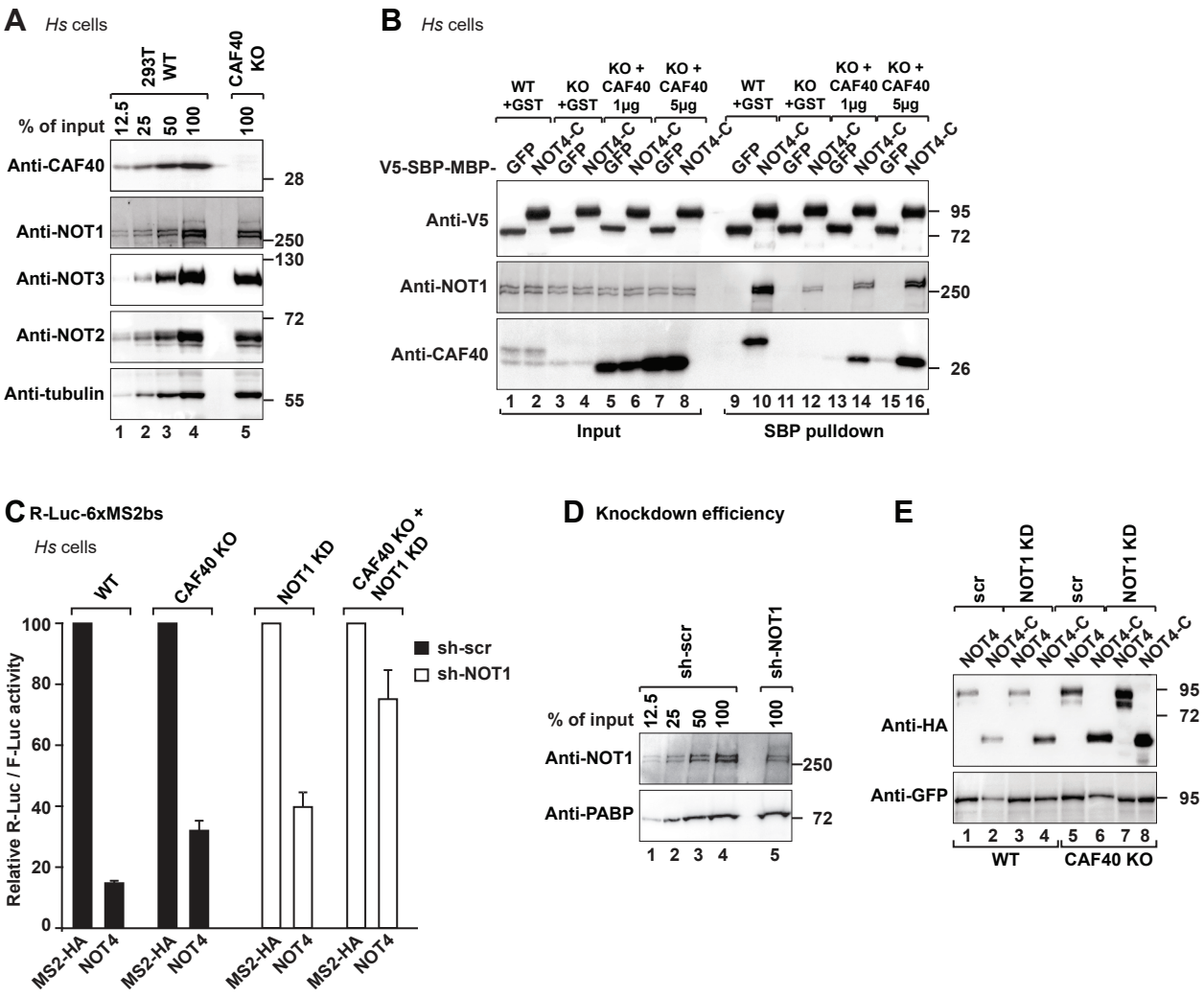


Figure S1. Keskeny et al.

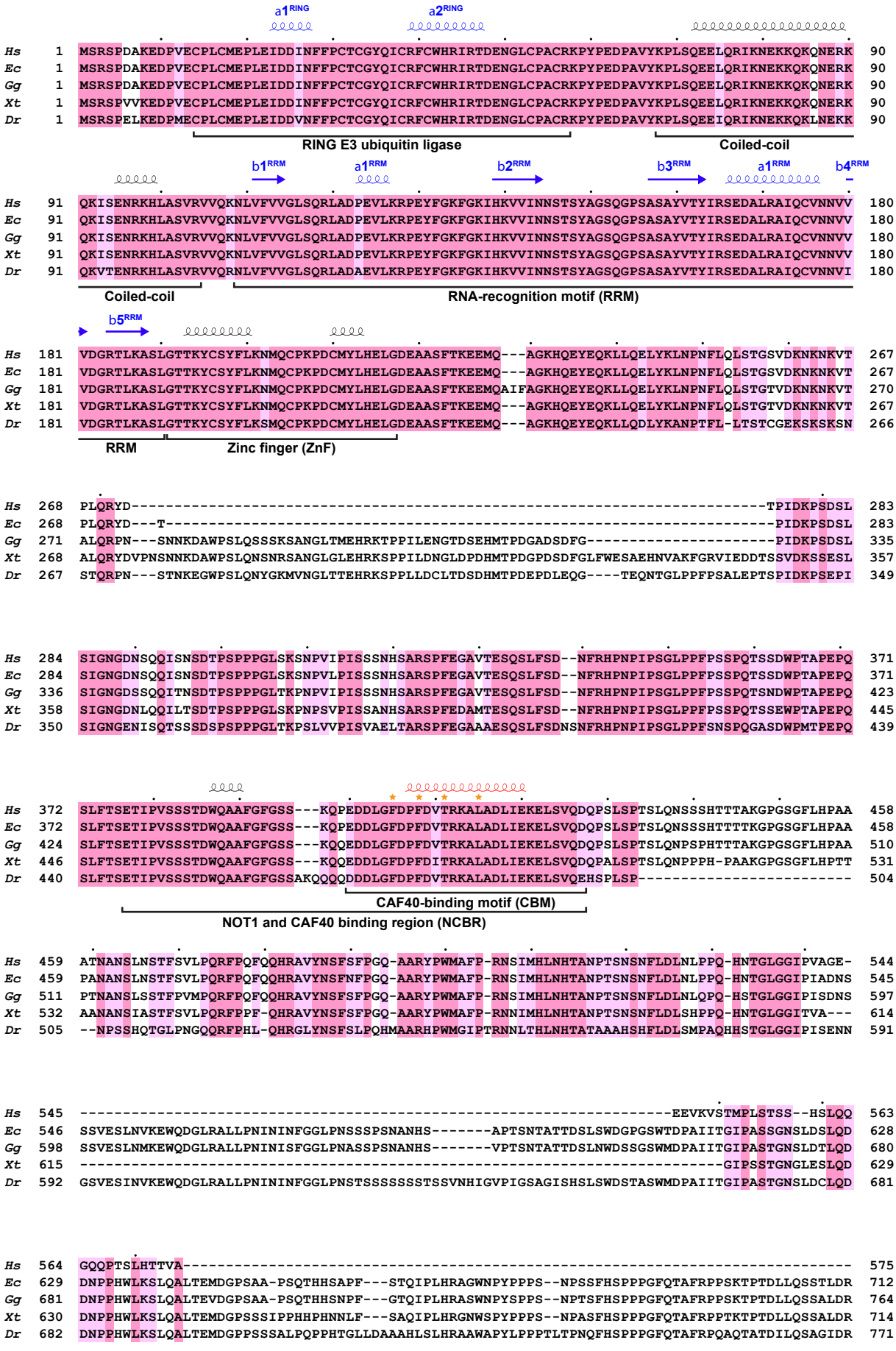


Figure S2. Keskeny et al.

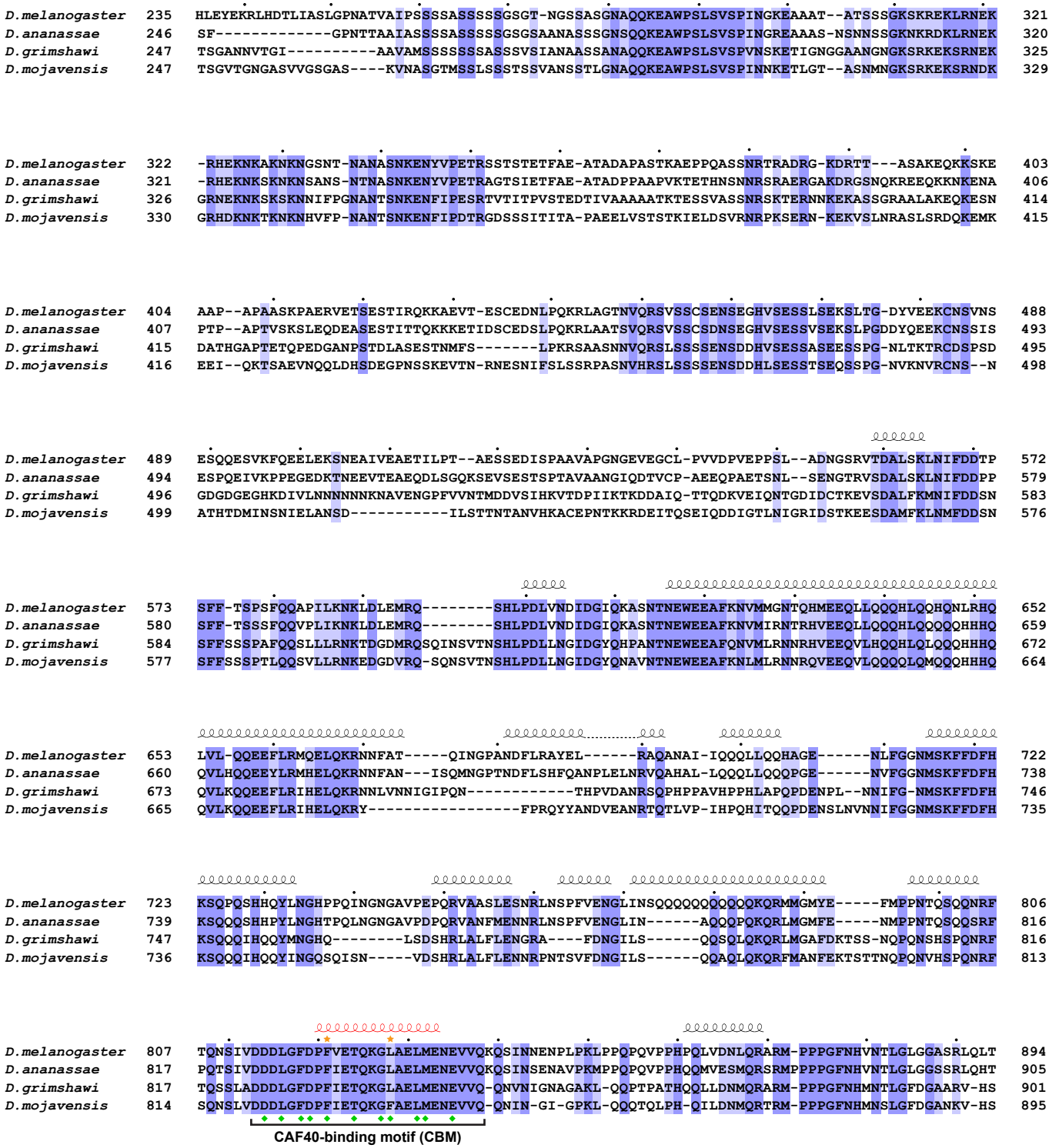


Figure S3. Keskeny et al.

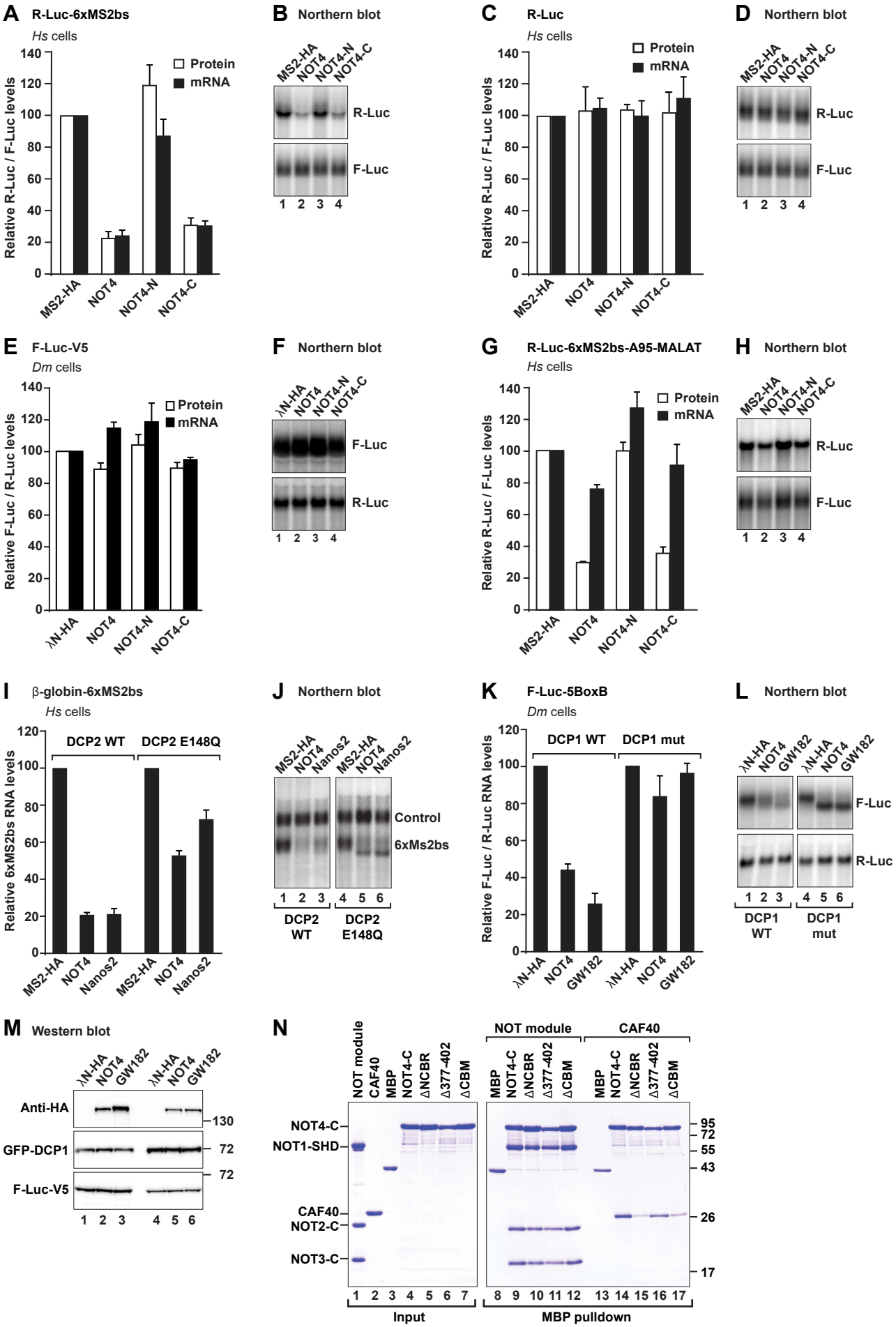
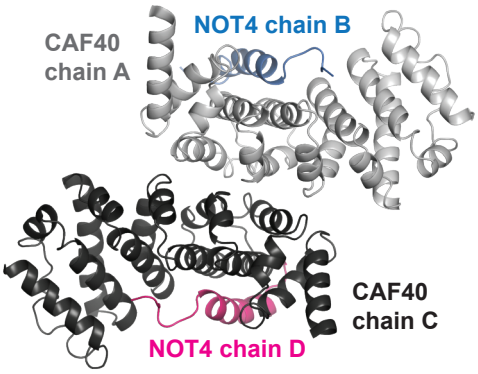
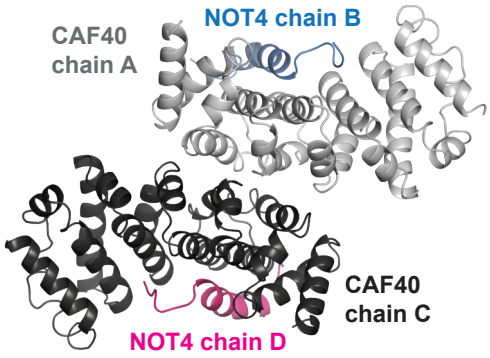


Figure S4. Keskeny et al.

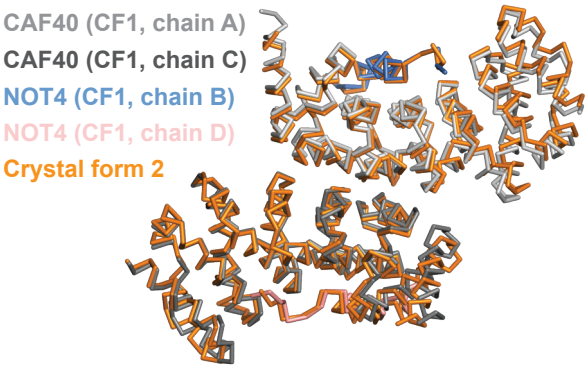
A Asymmetric unit crystal form 1



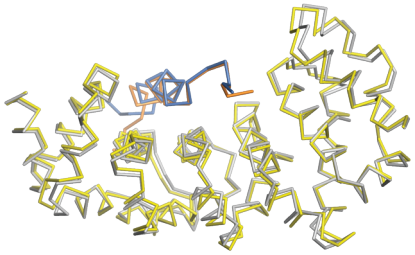
B Asymmetric unit crystal form 2



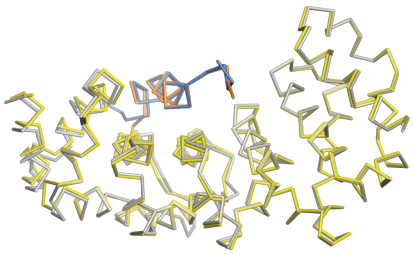
C Superposition crystal forms 1 and 2



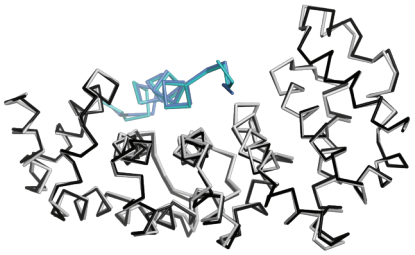
D Superposition dimers crystal form 1



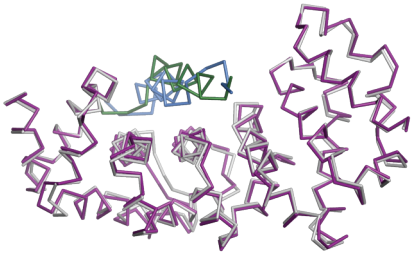
E Superposition dimers crystal form 2



F Superposition crystal forms 1 and 2



G Superposition NOT4 vs Roq



H

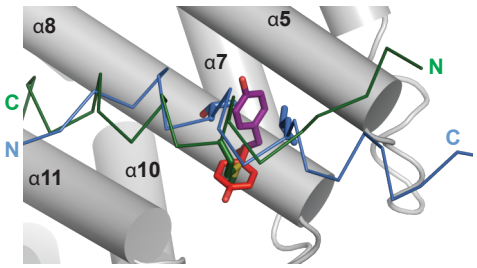


Figure S5. Keskeny et al.