

Inaugural dissertation
for
obtaining the doctoral degree
of the
Combined Faculty of Mathematics, Engineering and Natural Sciences of the
Ruprecht - Karls - University
Heidelberg

Presented by
(M.Sc.) Vaishali
born in: Delhi, India
Oral examination: 11 July 2022

Characterisation of RNA binding proteins and their roles in the
Drosophila germline

Referees: Prof. Dr. Michael Knop

Dr. Justin M. Crocker

E pur si muove
-Galileo Galilei

Table of contents

Acknowledgements.....	i
Summary.....	iii
Zusammenfassung.....	v
List of Abbreviations.....	vii
1. Introduction.....	1
1.1 Overview of RNA-binding proteins.....	1
2. FMR1 - a novel <i>oskar</i> mRNP component	5
2.1 Introduction	5
2.1.1 RNA localisation and localised translation in post-transcriptional gene regulation.....	5
2.1.2 Translation regulation of localised RNAs	6
2.1.3 <i>oskar</i> mRNP complex - a paradigm to study RNA localisation and localised translation in the <i>Drosophila</i> germline.....	8
2.1.4 Unbiased, comprehensive identification of the <i>oskar</i> RNP proteome.....	13
2.1.5 FMR1 regulates various facets of RNA function	14
2.2 Results	21
2.2.1 Validation and physical characterisation of <i>oskar</i> -FMR1 interaction.....	21
2.2.2 Functional analysis of <i>oskar</i> -FMR1 interaction	30
2.2.3 Mechanism of translational regulation by FMR1	36
2.3 Discussion	40
3. EB1: an unconventional RBP, or not?	43
3.1 Introduction	43
3.1.1 Identification of novel RBPs in <i>Drosophila melanogaster</i>	43
3.1.2 EB1 is a microtubule plus end binding protein	44
3.1.3 Preliminary investigation of RNA binding properties of EB1 revealed its putative RNA binding surface	46
3.2 Results	49
3.2.1 EB1 binds to RNA and microtubules in a mutually exclusive manner.....	49
3.2.2 RIP-seq to identify target RNAs of EB1 in the <i>Drosophila</i> oocyte	50
3.2.3 Co-localisation analysis of EB1 with the top hits.....	51
3.3 Discussion	54
4. Discussion and Outlook.....	56
5. Materials and methods	58
5.1 Materials.....	58
5.1.1 Antibodies.....	58
5.1.2 Fly stocks.....	59
5.1.3 Primers and Probes	60
5.1.4 Buffers and reaction mixes	64
5.1.5 Chemicals and reagents	67
5.1.6 Kits	69

5.1.7 Equipment.....	70
5.1.8 <i>In vitro</i> iCLIP Barcodes.....	70
5.2 Methods.....	71
Probe labelling for smFISH.....	71
Single molecule fluorescence in-situ hybridization (smFISH).....	71
Crosslinking and immunoprecipitation (CLIP) for ovaries.....	72
Recombinant protein purification from <i>E. coli</i>	72
<i>In vitro</i> Binding Assay.....	74
<i>In vivo</i> iCLIP	74
<i>In vitro</i> iCLIP.....	76
<i>In vitro</i> UV crosslinking assay	79
Western blot for ovaries	80
qRT-PCR for ovaries	80
ePAT assay	81
Pole cell analysis	82
Cuticle preparation for embryos	83
Hatching rate analysis.....	83
<i>In vitro</i> tethering assay	83
Co-sedimentation Assay	85
RNA-immunoprecipitation and sequencing (RIP-seq).....	85
References.....	87
List of figures.....	95
Appendix A: Supplementary Figures	97
Appendix B: Supplementary Materials and methods	98

Acknowledgements

My journey as a PhD student here at EMBL is one I will never forget. I still remember the day I visited this place for the first time for interviews. I had never left my town, let alone my country and the whole experience was so new for me, I felt out of place at times. But thanks to the wonderful people I met along the way, and the friends I made, Heidelberg became my second home.

I would first and foremost like to express my gratitude to my supervisor Anne, for giving me this chance to explore not just wonderful science, but also a new world. She has always been supportive of my work and my growth, giving me all the opportunities I sought to improve my science and skills. I would also like to thank my TAC members Michael Knop, Justin Crocker and Janosch Hennig for their feedback and assessments during my PhD that helped me make my work better. Thanks to the core facilities at EMBL, for helping me with experiments and suggestions, especially the Gene Core Facility, and the Electron Microscopy Core Facility. I would also like to thank my collaborators along the way- Julian König and Kathi Zarnack for the help with iCLIP experiments and analysis, and to Mandy Jeske for her help with the *in vitro* tethering assay.

I would also like to thank Frank Wippich and Lucy, who have been very instrumental in supporting my work through their wonderful guidance, discussions and contributions towards my two projects. Apart from being a wonderful colleague, Lucy has also been a great friend who has helped me not just professionally, but also personally. Thanks to all the present and past lab members for their feedback, discussions and suggestions all along: Simone, Masroor, Dylan, Imre, Zeljko, Matteo. Thanks to Anna for all her support and being a very cool bay-mate, and special thanks to Mainak for the thought-stimulating discussions over coffee breaks. Also, big thanks to Lucia for being a good friend, also for reading my thesis and her valuable feedback.

I made some lovely friends here, who have become like a family to me, sharing all the happiness and sadness together. Thanks to Janey, Daja and Huiting for always being there. For the wonderful trips, silly pictures, the inside jokes and the never ending memefest. Special thanks to Brana for being the best house-mate and the beautiful friend that she is.

Finally, to the people who I owe almost everything in my life, my parents, for always believing in me and reminding me that they have my back. And the biggest thanks to the person I dedicate my thesis to, my biggest cheerleader, my little brother Aaryan. He has always unconditionally supported me and my decisions, and he gives me the strength and the optimism to keep going. Thank you छोटी किरणी |

Summary

The important role of RNA binding proteins (RBPs) in regulating the fate and functions of RNAs has led to the development of transcript-specific as well as transcriptome-wide techniques allowing an unbiased and comprehensive identification of RBPs. These methods have extended our knowledge of the extent of RBPs in a cell, and studying the roles of these newly identified RBPs in cellular processes has provided us with novel insights into the RNA binding mechanisms, functions and regulation of RNA binding proteins.

For my PhD work, I assessed the RNA binding functions of two proteins identified in high-throughput screens. The first protein is the Fragile X Mental Retardation protein (FMR1), identified in a transcript-specific pulldown targeted at the *Drosophila* maternal mRNA *oskar*. I show that FMR1 is a bona fide component of the *oskar* RNA-protein complexes that interacts with the *oskar* 3'UTR *in vivo*. FMR1 positively regulates Oskar protein levels in the oocyte, without any effect on *oskar* RNA levels. Oskar protein nucleates germ plasm assembly and germ cell formation in the embryo, and the reduction in Oskar protein levels leads to a reduction in the number of pole cells formed in embryos knocked down for FMR1. Finally, I tried to determine how FMR1 regulates translation, with roles identified as both a repressor and activator of translation. FMR1 contains two types of RNA binding domains: two KH domains and a C-terminal RGG box. I show that, *in vitro*, FMR1 activates translation through the KH domains and requires the C-terminal RGG box for repression of translation. I have thus identified a new role of FMR1 in germline development in *Drosophila melanogaster*, and also a putative mechanism of how FMR1 performs antagonistic functions in translation regulation. The second protein I studied is the microtubule binding protein EB1, identified as a putative RNA binding protein in a transcriptome-wide RNA interactome capture study performed in *Drosophila* embryos. Preliminary data showed that EB1 binds to polyU₂₅ RNA *in vitro*, and uses the same binding surface for interacting with microtubules and RNA. I show that EB1 binds to microtubules and RNA in a mutually exclusive manner *in vitro*. Furthermore, I performed a RIP-seq experiment to identify the *in vivo* targets of EB1, but failed to validate the interaction of any of the top candidates with EB1 *in vivo*. This does not, however, negate a role of EB1 as an RNA binding protein altogether, as RNA might be regulating the functions of the protein, and this would require further investigation.

Zusammenfassung

Die wichtige Rolle von RNA-bindenden Proteinen (RBPs) bei der Regulierung der Bestimmung und der Funktionen von RNAs hat zur Entwicklung transkriptspezifischer sowie transkriptomweiter Techniken geführt, die eine unvoreingenommene und umfassende Identifizierung von RBPs ermöglichen. Diese Methoden haben unser Wissen über das Ausmaß von RBPs in einer Zelle erweitert, und die Untersuchung der Rollen dieser neu identifizierten RBPs in zellulären Prozessen hat uns neue Einblicke in die RNA-Bindungsmechanismen, Funktionen und die Regulation von RNA-bindenden Proteinen geliefert.

In meiner Doktorarbeit habe ich die RNA-Bindungsfunktionen von zwei Proteinen untersucht, die in Hochdurchsatz-Screenings identifiziert wurden. Das erste Protein ist das Fragile X Mental Retardation Protein (FMR1), welches in einem transkriptspezifischen Pulldown-Assay identifiziert wurde, der auf die mütterliche *Drosophila*-mRNA *oskar* abzielte. Ich zeige, dass FMR1 tatsächlich eine Komponente des *oskar*-RNA-Protein-Komplexes ist, der *in vivo* mit der *oskar*-3'UTR interagiert. FMR1 reguliert den *Oskar*-Proteingehalt in der Eizelle positiv, ohne Auswirkungen auf den *oskar*-RNA-Gehalt. Das *Oskar*-Protein stimuliert den Keimplasmaaufbau und die Bildung von Keimzellen im Embryo und die Verringerung der *Oskar*-Proteinkonzentration in FMR1-Knockdown Embryonen führt zu einer Verringerung der Anzahl der Polzellen, die in diesen Embryonen gebildet werden. Schließlich versuchte ich herauszufinden, wie FMR1 die Protein-Translation reguliert, wobei sowohl Repressor- als auch Aktivator-Funktionen für die Translation identifiziert wurden. FMR1 enthält zwei Arten von RNA-Bindungsdomänen: zwei KH-Domänen und eine C-terminale RGG-Box. Ich zeige, dass FMR1 *in vitro* die Translation durch die KH-Domänen aktiviert und die C-terminale RGG-Box für die Repression der Translation benötigt. Ich habe somit eine neue Rolle von FMR1 bei der Keimbahnentwicklung in *Drosophila melanogaster* identifiziert, sowie einen mutmaßlichen Mechanismus dafür, wie FMR1 antagonistische Funktionen bei der Translationsregulation ausübt.

Das zweite Protein, das ich untersucht habe, ist das Mikrotubuli-bindende Protein EB1, das in einer transkriptomweiten RNA-Interactome-Capture-Studie, die an *Drosophila*-Embryonen durchgeführt wurde, als mutmaßliches RNA-bindendes Protein identifiziert wurde. Vorläufige Daten zeigten, dass EB1 *in vitro* an polyU25-RNA bindet und die gleiche Bindungsfläche für die Wechselwirkung mit Mikrotubuli und RNA verwendet. Ich zeige,

dass EB1 in vitro auf sich gegenseitig ausschließende Weise an Mikrotubuli und RNA bindet. Darüber hinaus führte ich ein RIP-seq-Experiment durch, um die in-vivo-Ziele von EB1 zu identifizieren, konnte jedoch die Interaktion eines der Top-Kandidaten mit EB1 in vivo nicht validieren. Dies negiert jedoch nicht die Rolle von EB1 als RNA-bindendes Protein insgesamt, da RNA möglicherweise die Funktionen des Proteins reguliert und dies weitere Untersuchungen erfordern würde.

List of abbreviations

APC	Adenomatous polyposis coli
ARS	ATP Regenerating System
<i>ask1</i>	<i>Apoptosis signal-regulating kinase 1</i>
ATP	Adenosine-5'-triphosphate
<i>bcd</i>	<i>bicoid</i>
BRE	Bruno Response Element
CCR4-NOT	Carbon Catabolite Repression—Negative On TATA-less
<i>chc</i>	<i>Clathrin heavy chain</i>
CLEM	Corelative light and electron microscopy
CPEB	Polyadenylation Element Binding protein
CSP	Chemical Shift Perturbations
CTD	C-terminal Domain
CYFIP1	Cytoplasmic FMRP Interacting Protein
DAPI	4',6-diamidino-2-phenylindole
<i>dcr-1</i>	<i>dicer-1</i>
ddUTP	2',3'-Dideoxyuridine-5'-triphosphate
<i>dhc64c</i>	<i>Dynein heavy chain 64c</i>
Dnase	Deoxyribonuclease
DSHB	Developmental Studies Hybridoma Bank
DTT	Dithiothreitol
EB1	End Binding Protein 1
EBH	EB-Homology
eIF4A	Eukaryotic translation initiation factor 4A
eIF4E	Eukaryotic translation initiation factor 4E
eIF4E-BP	eIF4E Binding Protein
eIF4G	Eukaryotic translation initiation factor 4G
ePAT	extension polyA test
FL	full length
FMR1	Fragile X Mental Retardation Protein 1
FXAND	Fragile X-associated neuropsychiatric disorder
FXG	Fragile X Granules
FXPOI	Fragile X primary ovarian insufficiency
FXR1	Fragile X mental retardation syndrome-related protein 1
FXR2	Fragile X mental retardation syndrome-related protein 2
FXTAS	Fragile X-associated tremor and ataxia syndrome
<i>gcl</i>	<i>germ cell-less</i>
GFP	Green Fluorescent Protein
GTP	Guanosine-5'-triphosphate
HITS-CLIP	High-throughput sequencing of RNA isolated by crosslinking immunoprecipitation
hnRNPs	Heterogeneous nuclear ribonucleoproteins

HRP	Horse radish peroxidase
iCLIP	individual crosslinking and immunoprecipitation
Imp	IGF-II mRNA-binding protein
KH	K homology
luc	luciferase
m6A	N6-Methyladenosine
Me31B	maternal expression at 31B
mGluR	Metabotropic Glutamate Receptor
mRNP	mRNA-Protein complex
<i>msps</i>	<i>mini spindles</i>
Nos	Nanos
Orb	oo18 RNA-binding protein
Osk	Oskar protein
pAbp	polyA binding protein
PBS	Phosphate buffered saline
PCR	Polymerase Chain Reaction
<i>pgc</i>	<i>polar granule component</i>
PNK	Polynucleotide kinase
PTB	Polypyrimidine tract-binding protein
RAN	repeat-associated non-AUG
RBD	RNA Binding Domain
RBP	RNA Binding Protein
RIP-seq	RNA immunoprecipitation and sequencing
RNAi	RNA interference
RNase	Ribonuclease
RNP	RNA-Protein complex
RP	Ribosomal proteins
RPKM	Reads per kilo base per million mapped reads
RT	Room Temperature
<i>shd</i>	<i>shattered</i>
<i>shot</i>	<i>short stop</i>
smFISH	single molecule fluorescence in-situ hybridisation
SOLE	Spliced oskar localisation element
<i>synj</i>	<i>Synaptojanin</i>
TBS	Tris buffered saline
TdT	Terminal deoxynucleotidyl transferase
Tral	Trailer hitch
tRNA	transfer RNA
UTR	untranslated region
UV	ultraviolet
ZBP1	Zipcode Binding Protein 1

1. Introduction

In order to gain a better understanding of how cell fate and functions are controlled, much effort has been put in trying to discern the various networks, processes and interactions that take place inside a cell. Internal and external cues drive cascades of interactions important for regulating gene expression at the level of DNA and the process of transcription, or at the level of post-transcriptional and post-translational regulation affecting RNAs and proteins. A cell harbors several different kinds of RNAs, from protein coding messenger RNAs to ribosomal RNAs and transfer RNAs that aid in protein synthesis, as well as various small RNAs and long non-coding RNAs that regulate gene expression. Furthermore, recent evidence also points towards non-coding roles even of mRNAs, for instance, their roles in assembly and function of membraneless granules (Van Treeck et al. 2018). Regulating RNAs and their functions can, thus, help in regulating various facets of cellular function.

1.1 Overview of RNA-binding proteins

From the moment an RNA molecule is being transcribed in the nucleus, it starts to associate with RNA Binding Proteins (RBPs), forming RNA-protein complexes, that are responsible for regulating the various states and functions of the RNA. After transcription, most RNAs are processed by means of splicing, capping and polyadenylation, before being exported out of the nucleus into the cytoplasm. Once in the cytoplasm, the RNAs can be transported to a specific subcellular location before they are translated. The localisation of RNAs is especially important in large polarized cells, such as the neurons where the nucleus and cell body might be away from the site of action of particular proteins in the dendrites and axons. The local enrichment of RNA offers an efficient way to respond to local stimuli at the synapses (Das et al. 2019). The stability of the RNAs is also tightly regulated by RBPs (among other factors), protecting the RNAs from, or targeting them to degradation depending on the state of the cell (Boo and Kim 2020) (Fig 1.1).

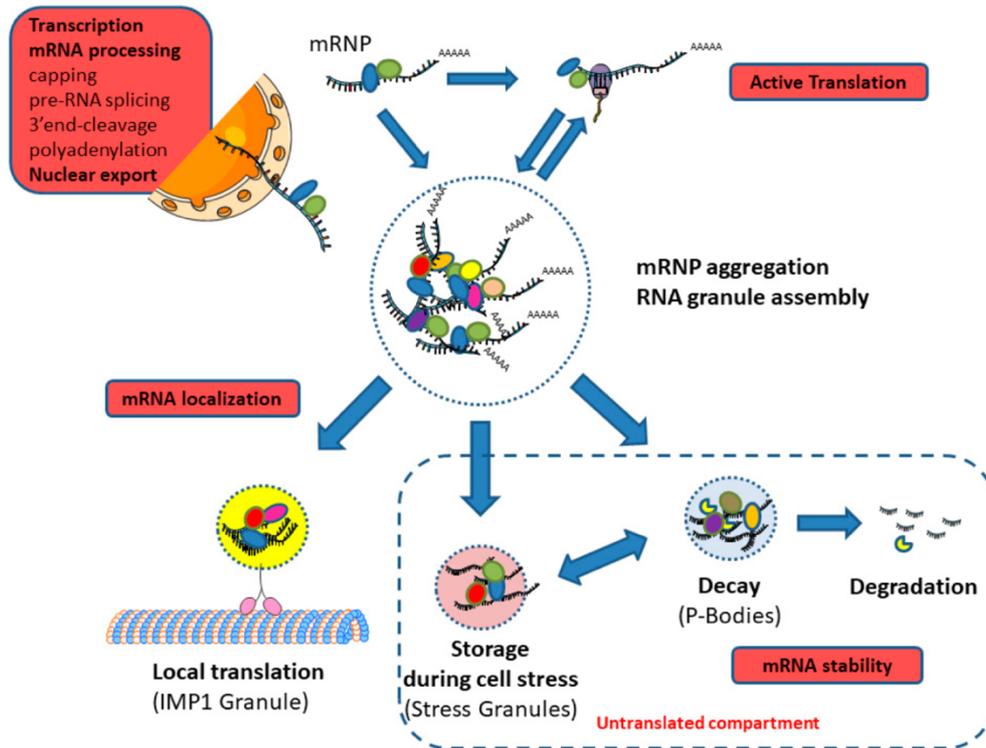


Fig. 1.1: RBPs play a crucial role in regulating various aspects of RNA function of the RNA life cycle, from transcription to degradation. (From (Coppin et al. 2018))

The proteome associated with a given RNA is quite dynamic, with proteins dissociating, and associating throughout the life of an RNA (Coppin et al. 2018). This dynamic change in the RNA bound proteome is crucial for regulating the RNA functions in response to the cell state changes as a result of both internal and external events. Studying the roles and functions of RBPs can thus provide insights into the mechanisms and processes of how post-transcriptional regulation takes place in a cell.

Because of the important roles of RBPs in a cell, several high throughput methods were developed in the last decade to identify the complete set of RBPs in a cell in an unbiased manner. Some of these were transcriptome-wide, aimed at identifying all the different RBPs in a cell (Castello et al. 2013; Sysoev et al. 2016; Wessels et al. 2016), while others were transcript-specific, aimed at identifying the repertoire of RBPs bound to a transcript of interest and dissecting the RBP network of a particular RNP complex (Wippich and Ephrussi 2020). These high throughput methods, however, offer little insight into the mechanisms of binding

or functions of these newly discovered RBPs. One, therefore, needs to systematically validate the identified candidates to confirm their physiological roles as RNA binding proteins.

The aim of this thesis was to study two proteins identified in high throughput screens, validating their RNA binding activities, and determining their cellular roles, if any, in regulating the functions of RNAs to which they bind. The first protein is the Fragile X mental Retardation Protein (FMR1), which was identified as a novel protein component of *oskar* RNP granules in an *oskar*-specific pulldown, performed by Frank Wippich, a former postdoctoral fellow in the lab. FMR1 is a known RNA binding protein that harbors two types of RNA binding domains: KH domains and an RGG box. FMR1 has been implicated in regulating several functions of RNAs, but its role in the germline as a component of *oskar* RNP granules has not been studied. The second protein, End-Binding Protein 1 (EB1), was identified as a novel, putative RBP in a transcriptome-wide analysis performed on *Drosophila* embryos by Vasily Sysoev, a former PhD student in the lab (Sysoev et al. 2016). EB1 is a known microtubule plus end binding protein. EB1 does not harbor any conventional RNA binding domains, and had no RNA related functions known to date. My investigation regarding EB1 has been published in (Vaishali et al. 2021).

2. FMR1 - a novel *oskar* mRNP component

2.1 Introduction

2.1.1 RNA localisation and localised translation in post-transcriptional gene regulation

RNA localisation and local translation are essential and well conserved mechanisms for the functional polarisation of a cell. RNAs are specifically targeted to subcellular compartments/locations, where the RNA itself or the protein it encodes are required for proper functioning of the cell (Fig. 2.1). In many organisms, RNA localisation is crucial during development, where the asymmetric localisation of RNAs creates asymmetry in cell division, a process important for germ layer specification in early development (Birsoy et al. 2006; Takatori et al. 2010), and for stem cell differentiation later (Broadus et al. 1998; Hughes et al. 2004). In the *Drosophila* oocyte, localisation of maternal RNAs is essential for embryonic polarity establishment (Medioni et al. 2012), a process in detail later. Furthermore, RNAs are also localised to distinct cellular compartments, such as the FMR1 regulated localisation of *centrocortin* mRNA to centrosomes, a process essential for proper spindle morphogenesis and genomic stability in *Drosophila* embryos (Ryder et al. 2020). The misregulation of RNA localisation can therefore disrupt developmental programs (Bashirullah et al. 1998), and cellular functions with implications in diseases including cancer and metastasis (Jakobsen et al. 2013).

A cell can employ various means to correctly localise RNAs to the target locations: 1) the RNAs can be actively transported to the site with the help of molecular motors along the cytoskeleton, 2) they can be selectively degraded at all other locations, except the site of action, and/or 3) RNAs can diffuse towards the site of action, and get entrapped by local proteins and complexes (Martin and Ephrussi 2009). Irrespective of the mechanism used for localisation, the RNAs to be localised are tightly regulated for correct spatial and temporal expression, and this is achieved by a combination of different *cis* acting elements and *trans* acting factors. The RNA sequence and structure often contains elements required for interactions with RBPs that are involved in targeting the RNA to the specific location in the cell. *oskar* mRNA in the *Drosophila* oocyte, for example, contains a stem loop structure called

the oocyte entry signal (OES) in its 3'UTR, which is important for the transport of *oskar* from the nurse cells into the oocyte. The RBPs also keep the translation status of the RNAs in check by repressing the translation during transport, and de-repressing/activating it only after receiving appropriate signals post localisation.

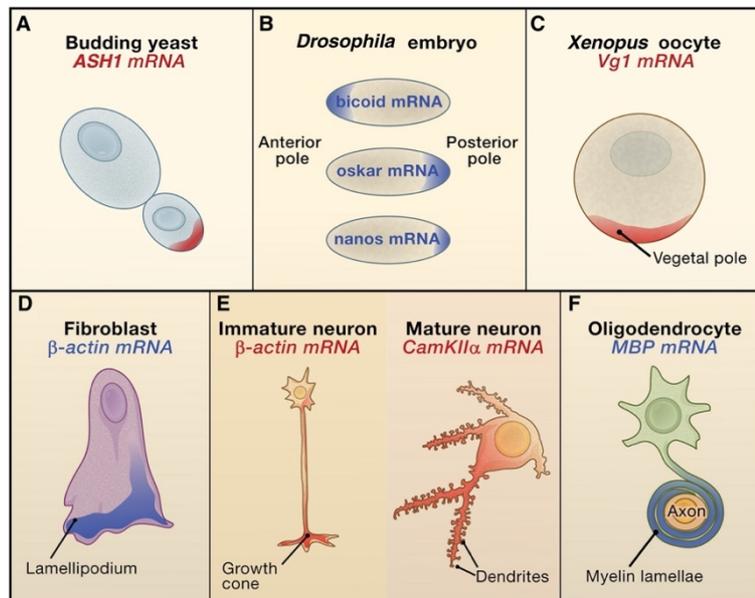


Fig. 2.1: Examples of mRNA localisation in different cell types and organisms. (from (Martin and Ephrussi 2009))

2.1.2 Translation regulation of localised RNAs

One of the main functional outcomes of RNA localisation is the synthesis of the protein product encoded by the RNA in a spatially controlled manner. As a single RNA can give rise to many protein molecules, localising RNAs is an energy efficient mechanism to increase the local concentration of proteins at the site of action. The translation status of the localising RNAs is therefore tightly regulated to ensure that the RNAs are translationally repressed during transport, and activated only upon successful localisation, so that the protein is formed only at the destined location. RBPs are key molecular players involved in regulating the translation of RNAs and different RBPs can contribute to repression or activation of translation through different mechanisms.

Mechanisms of translation repression by RBPs

RBPs can repress the translation of localising RNAs by different mechanisms. Cap-dependent translation relies on the binding of eIF4E to the 5' cap, and the recruitment of

eIF4G, eIF4A and other initiation factors to the RNA, forming the translation initiation complex (Fig. 2.2A). RBPs can inhibit cap-dependent translation initiation by directly binding to eIF4E, as does the Bicoid protein in the repression of *caudal* mRNA translation (Fig. 2.2B). RBPs can also recruit eIF4E-Binding Proteins (eIF4E-BP) and their homologs to the RNP complexes, that in turn bind to eIF4E, as is the case of Cup recruitment by Bruno to repress the translation of *oskar* mRNA in the *Drosophila* oocyte (Fig. 2.2B) (Stebbins-Boaz et al. 1999; Santini et al. 2017). RBPs can inhibit translation by binding to RNAs and preventing the 40S and 60S ribosomal subunits from assembling (Fig. 2.2C). ZBP1 binds to β -actin mRNA in fibroblasts and represses its translation using this mechanism. Upon localisation to the protrusions, ZBP1 is phosphorylated and the ZBP1- β -actin mRNA interaction is disrupted, leading to ribosome assembly and translation (Biswas et al. 2019). Translation inhibition can also be at the elongation step, manifested by binding of RBPs to the ribosomes directly and occluding tRNA binding site, as has been suggested for FMR1 (Fig. 2.2D) (Chen et al. 2014). Furthermore, granule formation by RNA-protein interactions can serve as a mechanism for translation repression by preventing, for instance, access of the translational machinery to the RNAs in the granules (Fig. 2.2E) (Chekulaeva et al. 2006; Kim Tae et al. 2019; Moissoglu et al. 2019).

Mechanisms of translation de-repression/activation by RBPs

Though a few mechanisms of translational repression of localising RNAs have been studied in detail, much less is known about the translation activation of RNAs once they are localised. Activation of translation at the correct location is dependent on the local milieu of proteins and factors. External and internal cues may lead to dynamic remodelling of the RNPs, leading to changes in the bound proteome that are more conducive to facilitating translation. Changes in local concentrations of competing RBPs, signalling molecules and proteins, and post-translational modifications, such as phosphorylation for ZBP1 mentioned above, can all play a role in RNP remodelling. Increased availability of ribosomes and other translation initiation and elongation factors can also create local translation “hubs” that can rapidly respond to de-repression and activation of translation. For instance, in mouse intestinal cells, ribosomal protein (RP) RNAs are enriched at the basal end during starvation. Upon refeeding, they are translocated to the translationally active apical region where the translation of these

RNAs is enhanced, possibly leading to increased ribosome synthesis and, in turn, protein synthesis at the apical region (Moor Andreas et al. 2017).

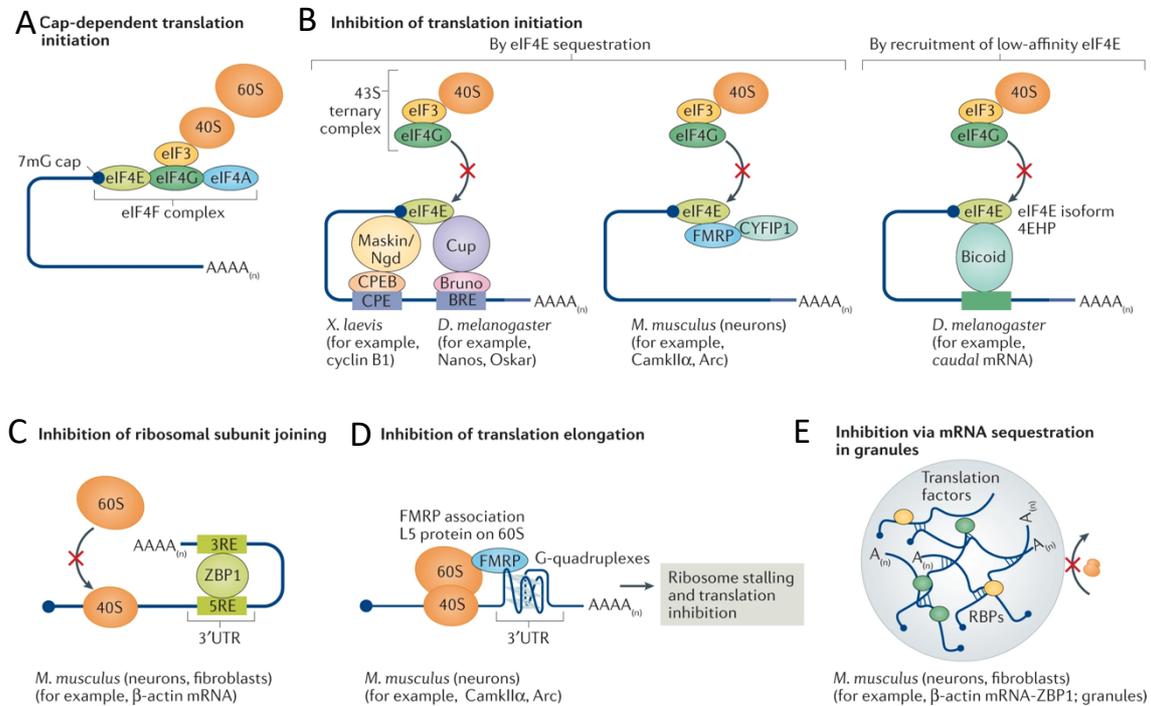


Fig. 2.2: Mechanisms of RNA translation inhibition. RNAs that rely on cap-dependent translation initiation can be inhibited by b) sequestering eIF4E via association with proteins that bind to it (4E-binding proteins) or by using low affinity eIF4E. Translation can also be inhibited by c) preventing the 40S and 60S ribosomal subunits from associating, or by d) inhibiting elongation. e) Granule formation can also inhibit translation by sequestering RNAs, and/or preventing access to ribosomes. (image adapted from (Das et al. 2021))

2.1.3 *oskar* mRNP complex - a paradigm to study RNA localisation and localised translation in the *Drosophila* germline

As mentioned earlier, in *Drosophila melanogaster*, embryonic axis establishment and patterning relies on the correct localisation of maternal RNAs in the oocyte (Fig. 2.3). The localisation of *bicoid* mRNA at the anterior and *oskar* mRNA at the posterior pole of the oocyte defines the anterior-posterior axis (Berleth et al. 1988; Ephrussi et al. 1991), whereas the *gurken* mRNA localisation at the antero-lateral cortex of the oocyte signals the induction of dorsal cell fate (González-Reyes et al. 1995). Function of localising RNAs is, however, not only limited to embryonic axis specification. High-throughput *in situ* hybridisation analysis

has revealed that as many as 71% of the 3370 RNAs studied exhibit some localisation pattern in the embryo (Lécuyer et al. 2007). Oogenesis and embryogenesis in *Drosophila melanogaster* has thus been studied extensively as a model to understand the process of RNA localisation and regulation at a molecular scale.

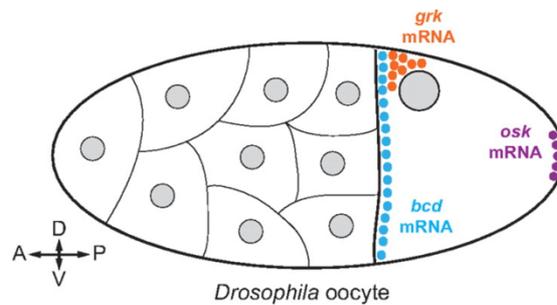


Fig. 2.3: Posterior localisation of *oskar* mRNA, anterior localisation of *bicoid* mRNA and antero-dorsal localisation of *gurken* mRNA in the *Drosophila* oocyte are essential for embryonic axis establishment (image from (Medioni et al. 2012))

In *Drosophila*, oogenesis takes place within a pair of ovaries, each of which consists of a string of developing oocytes called an ovariole (Fig 2.4). Egg chambers in different stages of development can be seen along the ovariole, with a germline stem cell-containing germarium at one end, and a fully developed oocyte at the other. Each egg chamber is a syncytium that consists of one oocyte and 15 supporting nurse cells connected to each other via ring canals, and surrounded by a layer of somatic follicular epithelium. During oogenesis, the oocyte is maintained in a transcriptionally silent state, and most of the transcripts and proteins needed by the oocyte during development are transcribed in the nurse cells and transported into the oocyte. During embryogenesis, there is an initial phase of 14 rapid nuclear divisions without cytokinesis, creating a syncytium of ~6000 nuclei. Zygotic transcription is activated only after nuclear division 8 (Laver et al. 2015), and therefore, the transcriptionally silent embryo relies on the maternally deposited RNAs and proteins for its development. The expression, localisation, translation and degradation of the maternal RNAs are thus very tightly regulated in the oocyte and the embryo.

One maternal effect gene crucial for proper embryonic patterning as well as germ cell formation in *Drosophila* is *oskar* (Nüsslein-Volhard et al. 1987; Ephrussi and Lehmann 1992). *oskar* mRNA is synthesised in the nurse cell nuclei, and bound by several RBPs forming *oskar*

RNP granules, which are then transported into the oocyte by active transport along microtubules in two phases: first, from nurse cells to the oocyte using the dynein motor for transport, and second, localisation to the posterior pole of the oocyte using the kinesin motor (Clark et al. 2007; Gáspár et al. 2017; Gáspár et al. 2021). Like most localising RNAs, *oskar* mRNA is also maintained in a translationally repressed state during transport, and de-repressed/activated once it is localised at the posterior pole (Fig. 2.4). *oskar* RNP granules are quite dynamic, and bound by a number of RBPs that regulate each step of the process.

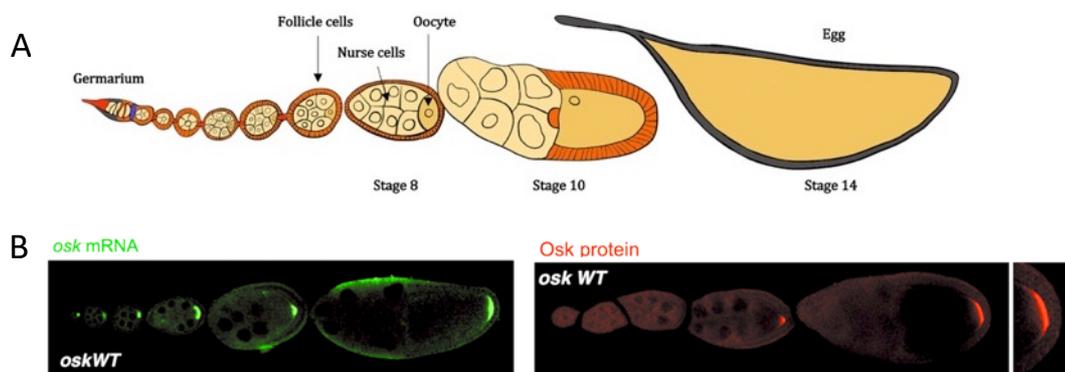


Fig. 2.4: A) Schematic representation of a *Drosophila* ovariole. B) *oskar* mRNA is localised to the posterior pole of the oocyte, and translated into Oskar protein only once correctly localised at stage 9. (image from (Vanzo and Ephrussi 2002))

Functions of Oskar protein

Oskar protein has two isoforms in the oocyte, Long Oskar, and Short Oskar translated from an alternative in-frame start (AUG) codon (Markussen et al. 1995). Short Oskar is more abundant, and sufficient for the pole plasm assembly and proper embryonic patterning. Long Oskar is not endowed with these functions, but is essential for stimulating endocytosis and actin nucleation at the posterior, and anchoring of *oskar* mRNA and Oskar protein (Fig. 2.5) (Vanzo and Ephrussi 2002; Vanzo et al. 2007). Several studies have been performed to understand the molecular interactions and basis for the functions of Oskar protein. Proper embryonic patterning in *Drosophila* is initiated by gradients of maternal morphogens across the anterior-posterior and dorsal-ventral axes. Nanos (Nos) protein acts as one such crucial

morphogen, accumulating at the posterior of the embryo, and the translational regulation of *nos* mRNA plays an important role in establishing this gradient. A protein, Smaug, binds to the *nos* 3'UTR in the embryo, and recruits the deadenylation complex CCR4-NOT, leading to destabilisation of the RNA and translational repression. At the posterior pole, Oskar protein prevents the binding of Smaug to *nos*, thereby preventing deadenylation and leading to translation activation of *nos* only at the posterior pole (Zaessinger et al. 2006). Furthermore, Oskar protein is known to nucleate the pole plasm assembly, and recruits several component proteins (such as Vasa, Tudor, Aubergine) and RNAs (such as *pgc* and *gcl*) to assemble the pole plasm (Anne 2010; Jeske et al. 2015). Ectopic expression of Oskar protein at the anterior of the oocyte using *bicoid* 3'UTR has the ability to induce ectopic pole plasm assembly, and head patterning defects (Ephrussi and Lehmann 1992). Therefore, precise spatial and temporal regulation of *oskar* mRNA is essential for development.

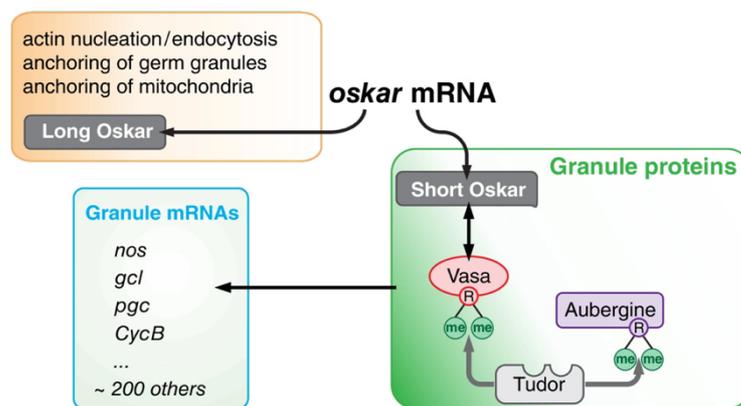


Fig. 2.5: *oskar* mRNA is translated into Long and Short Oskar at the posterior pole, and the two proteins have different functions: Short Oskar is required for germ granule assembly, and Long Oskar for endocytosis and germ plasm anchoring (image from (Treck and Lehmann 2019))

Translation regulation of *oskar* mRNA

The translation of *oskar* is tightly regulated so that it is translated only upon localisation at stage 9 of oogenesis. Several translation regulators of *oskar* mRNA have been identified over the years through biochemical and genetic studies. One of the main and most well characterised repressors of *oskar* translation is the Bruno protein (Kim-Ha et al. 1995). Bruno binds to *oskar* mRNA at sites called Bruno Response Elements (BREs) present at the 5' end (BRE AB) and 3' end (BRE C) of the *oskar* 3'UTR. Bruno represses the translation of *oskar*

through two mechanisms: one, by the recruitment of an eIF4E-Binding Protein, Cup, to *oskar* RNPs. Cup is known to directly interact with eIF4E, preventing its association with eIF4G, and thus inhibiting translation initiation (Nakamura et al. 2004). The second mechanism is a Cup independent mechanism, that involves Bruno-dependent mRNA oligomerisation and higher order complex formation of *oskar* (Chekulaeva et al. 2006). The formation of higher order granules seems to be an important mechanism used by *oskar* RNP particles to repress translation by preventing the access to ribosomes. The polypyrimidine tract-binding protein, PTB, is another translational repressor of *oskar*, and is crucial for the assembly of higher order *oskar* RNP complex formation *in vivo* (Besse et al. 2009). Owing to granule assembly, *oskar* can also exert translational control *in trans* on other *oskar* RNAs that lack *cis* regulatory elements (for examples, BRE site mutants), by recruiting the regulating factors to granules and influencing the regulation of mutant RNAs in the granule as well (Reveal et al. 2010; Macdonald et al. 2016). Several other RBPs known for their roles in translational regulation have also been identified as a part of *oskar* RNP granules. For instance, Hrp48, a member of hnRNPA/B family of RBPs (Yano et al. 2004), Imp (Geng and Macdonald Paul 2006) and Me31B (Nakamura et al. 2001) have all been studied and identified as translational repressors for *oskar*.

Though the mechanisms of translation de-repression/activation of *oskar* have not been studied in great detail, some *cis* and *trans* regulatory elements required for the process have been identified. As mentioned earlier, BRE AB and BRE C sites in the *oskar* 3'UTR are important for binding to Bruno and repressing the translation of *oskar*, but it was found that the BRE C site is also required for the activation of translation (Reveal et al. 2010). BRE C site overlaps with the binding site for an E3-ubiquitin ligase protein, Makorin 1 on *oskar* (Dold et al. 2020). Makorin 1 has been shown to compete with Bruno for binding at the site. In the presence of polyA binding protein (pAbp), the binding of Makorin 1 to BRE C is stabilised, leading to dissociation of Bruno and de-repression of *oskar* translation at the posterior pole (Dold et al. 2020).

Another RBP that has been identified for its role in stimulating the translation of *oskar* is Orb, which is a *Drosophila* homolog of Cytoplasmic Polyadenylation Element Binding (CPEB) protein (Chang et al. 1999) Studies have shown that the polyA tail length of *oskar* is crucial for its efficient translation *in vivo*, and evidence suggests that Orb helps in stimulating

the polyadenylation of *oskar* at the posterior pole. In the absence of Orb, there is an accumulation of *oskar* with short PolyA tail lengths, and reduced Oskar protein levels (Chang et al. 1999; Juge et al. 2002; Castagnetti and Ephrussi 2003).

Translation-independent role of *oskar* mRNA

Apart from the coding functions of *oskar* mRNA, there is evidence to show that *oskar* mRNA also has a translation-independent role in the process of oogenesis. Absence of *oskar* mRNA leads to an early oogenesis arrest, which can be rescued by expressing only the *oskar* 3'UTR (Jenny et al. 2006). It has been speculated that *oskar* might serve as scaffold or regulatory RNA using its 3'UTR, a region with binding sites for a number of RNA binding proteins. Indeed, a study has shown that one of the functions of *oskar* RNA is to sequester Bruno and limit its activity in the oocyte. The absence of *oskar* RNA also leads to altered distribution of proteins in the germline, with reduced enrichment in the nuage, or cytoplasmic granules such as sponge bodies (Kanke et al. 2015). It, therefore, shows that the interactions of *oskar* with RBPs, not only regulates the functions of *oskar*, but also that *oskar* might regulate the activity and functions of RBPs.

2.1.4 Unbiased, comprehensive identification of the *oskar* RNP proteome

Biochemical and genetic studies over the past years have led to the identification of several protein components of the *oskar* mRNP complex, expanding our understanding of the molecular players and mechanisms involved in the regulation of *oskar*. In order to further extend our knowledge of the repertoire of proteins that associate with the *oskar* RNP complex in an unbiased and comprehensive manner, a transcript specific pulldown was developed (Wippich and Ephrussi 2020), and performed on *Drosophila* oocytes by Frank Wippich, a former postdoc in our lab). A number of proteins, known to interact with *oskar* were enriched, along with many novel candidates previously not known to be a part of the complex (Fig. 2.6). One of such novel candidates was the synaptic functional regulator FMR1, known to regulate several aspects of RNA function.

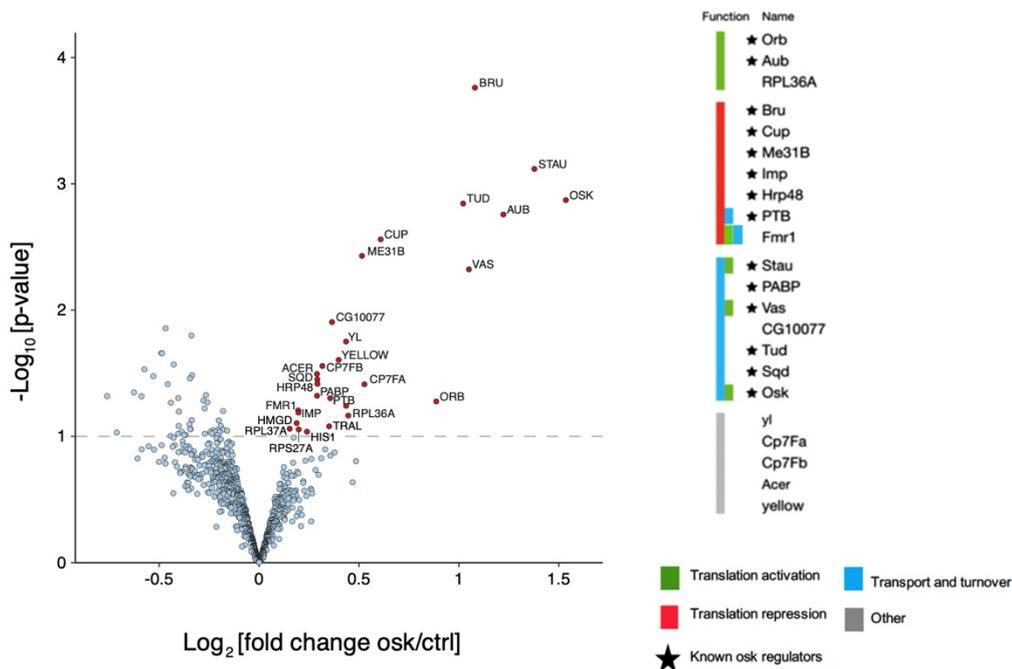


Fig. 2.6: The volcano plot shows the RBPs significantly enriched in the *oskar*-specific pulldown. On the right are listed some of the enriched RBPs and their known functions in RNA metabolism. The RBPs previously known to be components of *oskar*-RNP complexes are marked with a star. (Image courtesy: Frank Wippich)

2.1.5 FMR1 regulates various facets of RNA function

FMR1 is an RNA binding protein with a modular structure, and comprises of two types of RNA binding domains: two KH domains, and an RGG box in the mostly disordered C-terminal domain. It contains a Tudor (Agenet-like) domain in the N-terminus, important for homodimerization and protein-protein interactions. It also contains a nuclear localisation signal (NLS) and a nuclear export signal (NES) (Fig. 2.7). FMR1 is a highly conserved protein, with 70% similarity between the human and *Drosophila* orthologs. Mutations in the FMR1 gene are known to cause severe neurological disorders in humans. The most common mutation observed in the gene is a trinucleotide expansion (CGG) in the 5'UTR, and the length of the repeat expansion dictates the severity of the disease (Fig. 2.8) (Berman et al. 2014). Expansion of <55 repeats is considered normal, whereas 55-200 repeats are called a premutation, and can cause Fragile X primary ovarian insufficiency (FXPOI), Fragile X-associated tremor and ataxia syndrome (FXTAS) or Fragile X-associated neuropsychiatric disorder (FXAND). In this condition, there is an increased RNA production, but reduced protein levels of FMR1. One characteristic feature of FXTAS is the presence of intranuclear

inclusion bodies containing *fmr1* RNA, which binds to and sequesters several other RNA binding proteins such as Sam68, Pur α and hnRNP A2 in the granules, potentially making them unavailable for regulating other cellular processes (Iwahashi et al. 2006). Furthermore, there is also evidence that the premutation RNA can undergo repeat-associated non-AUG (RAN) mediated translation, producing toxic polyglycine containing polypeptides (Todd et al. 2013). When the trinucleotide repeat expansion is >200, it causes the fragile X syndrome, which is an inherited form of mental retardation, affecting 1 in 4000 males and 1 in 8000 females.

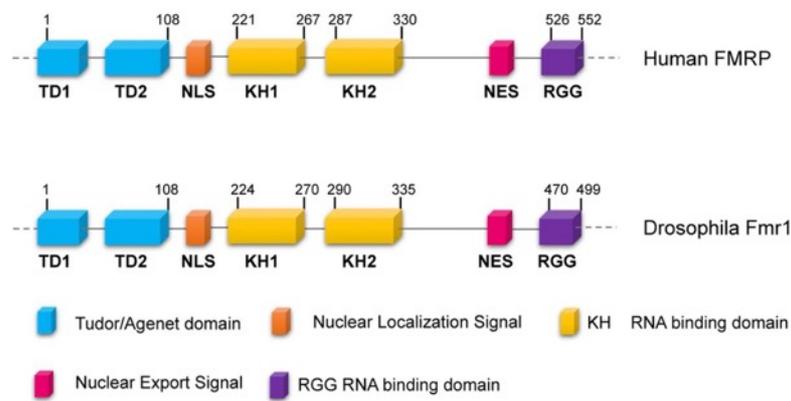


Fig. 2.7: Modular structure of human and *Drosophila* FMR1 proteins (image from (Specchia et al. 2019))

Studies to understand the functions of FMR1 have led to the identification of its roles in regulating multiple processes of post-transcriptional gene regulation. FMR1 plays an essential role in the assembly and regulation of several membraneless RNA-protein granules in the cell. It is a component of transport granules in neurons and regulates the efficient transport of mRNA granules in dendrites (Davidovic et al. 2007; Dichtenberg et al. 2008), a function conserved in *Drosophila* neurons as well (Estes et al. 2008). The molecular mechanism of how FMR1 contributes to the process is not very clear, but a study suggests that FMR1 might act as an adaptor between the RNA and the cytoskeletal motor proteins, kinesin in this case (Davidovic et al. 2007).

Apart from transport granules, FMR1 is also a bona fide component of stress granules (Mazroui et al. 2002; Gareau et al. 2013), P-bodies (Barbee et al. 2006) and Fragile X granules (FXG) (Chyung et al. 2018). These membraneless granules contain several other proteins as well as RNAs, and one of their main functions is the regulation of translation and degradation

of RNAs under different cellular conditions (Decker and Parker 2012). FMR1 has established roles as a translational regulator, predominantly as a repressor (Darnell et al. 2011; Chen et al. 2014). Studies in the past have shown that FMR1 can repress translation of target RNAs through different mechanisms (Fig. 2.9). The crystal structure of FMR1 with ribosomes revealed that it directly binds to ribosomes, at a site overlapping the tRNA binding site, and inhibits translation elongation by preventing tRNA binding (Chen et al. 2014). It can also directly inhibit cap-dependent translation initiation by interacting with Cytoplasmic FMRP Interacting Protein (CYFIP1), which is a eIF4E-BP that binds to eIF4E and prevents its association with eIF4G and thus the assembly of the initiation complex (Napoli et al. 2008) (Fig. 2.9B). There is evidence that FMR1 interacts with several RNAi and miRNA pathway proteins, another mechanism through which FMR1 contributes to translation repression of target RNAs (Ishizuka et al. 2002; Jin et al. 2004). In neurons, for example, the microRNA miR-125a is involved in the translation repression of *psd-95* RNA at the dendrites, and the assembly of the Ago2-miR-125a complex on *psd-95* is dependent on phosphorylated FMR1 (Fig. 2.9A). Upon mGluR stimulation, FMR1 is dephosphorylated and the Ago2-miRNA complex dissociates from the RNA, leading to the translation activation of *psd-95* (Muddashetty et al. 2011). Recently, it was discovered that FMR1 also has the ability to form phase separated granules using its C-terminal domain, and the granule formation is crucial for translation repression, possibly by excluding the translational machinery (Kim Tae et al. 2019; Tsang et al. 2019) (Fig. 2.9C).

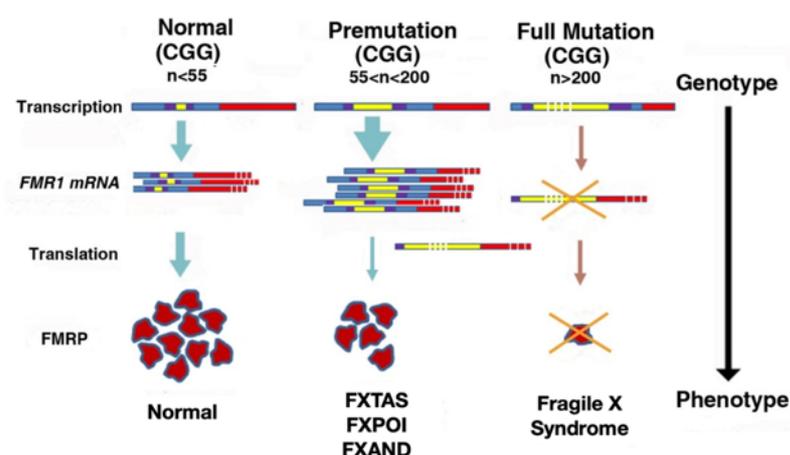


Fig. 2.8: The extent of trinucleotide repeat expansion in the FMR1 gene dictates the phenotype. Premutation leads to Fragile X primary ovarian insufficiency (FXPOI), Fragile X-associated tremor and ataxia syndrome (FXTAS) or Fragile X-associated neuropsychiatric disorder (FXAND), whereas a full mutation manifests as Fragile X syndrome, an inherited form of mental retardation (image adapted from (Berman et al. 2014))

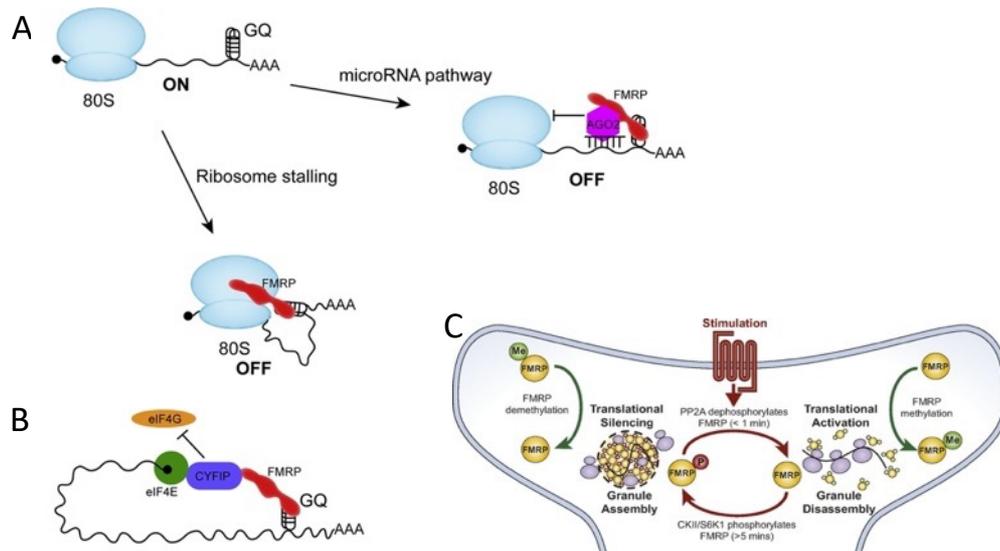


Fig. 2.9: Different mechanisms used by FMR1 for translation repression. A) FMR1 interacts with Ago and other components of the miRNA machinery, a mechanism though to inhibit translation. Furthermore, FMR1 can also inhibit elongation by directly binding to ribosomes at site overlapping the tRNA binding site. B) FMR1 in complex with CYFIP interacts with eIF4E, preventing its interaction with eIF4G and the assembly of initiation complex. (Image adapted from (Chen and Joseph 2015)) C) Phase separation mediated exclusion of translation machinery is another possible mechanism of repression by FMR1 (Image from (Tsang et al. 2019)).

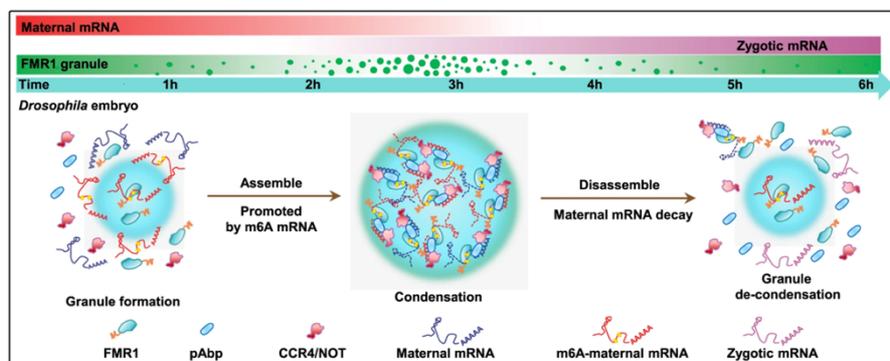


Fig. 2.10: FMR1 granules regulate m6A dependent maternal mRNA decay in mid-blastula transition during embryogenesis. Granule formation depends on the C-terminal domain of FMR1, as well as its RNA binding activity using KH2 domain (Image from (Zhang et al. 2022)).

FMR1 is also required for proper *Drosophila* embryogenesis by regulating aspects of RNA granule functions other than translation. FMR1 in this case is required for m6A methylation dependent decay of maternal mRNAs during the mid-blastula transition (Zhang et al. 2022). FMR1 binds to methylated RNAs and forms dynamic phase separated

condensates dependent on the FMR1 C-terminal domain. These condensates sequester other unmodified maternal RNAs as well, and targets them for degradation along with the methylated RNAs. This is followed by the de-condensation of FMR1 containing granules, and progression of normal embryogenesis (Fig. 2.10).

Though conventionally studied as a translation repressor, there have been studies that point towards a role of FMR1 in translation activation (Monzo et al. 2006; Bechara et al. 2009; Tabet et al. 2016; Greenblatt Ethan and Spradling Allan 2018). Ribosome profiling in *Drosophila* oocytes knocked down for FMR1 revealed that RNAs of 421 genes exhibited reduced translation as compared to wild type oocytes, and many of these RNAs were previously identified as FMR1 targets (Greenblatt Ethan and Spradling Allan 2018). Furthermore, during *Drosophila* embryogenesis, FMR1 is associated with dynamic ME31B and Tral containing cytoplasmic RNP bodies, and absence of FMR1 leads to a 2-fold reduction in Tral protein levels and affects cellularisation (Monzo et al. 2006). FMR1 also positively regulates the superoxide dismutase 1(*sod1*) RNA in mouse, and absence of FMR1 leads to reduced Sod1 protein levels. *Sod1* RNA contains an RNA structure called SoSLIP, that consists of three stem loop structures and acts in *cis* to activate translation. Binding of FMR1 to two of the stem loops induces structural changes in SoSLIP, and potentiates the translation activating effect of SoSLIP (Bechara et al. 2009).

Studies to determine the binding specificity of FMR1 have led to the identification of structures and motifs in RNAs that FMR1 preferentially binds to. One of the main structural motifs FMR1 has been found to interact with is the G-quadruplex structure in RNAs (Darnell et al. 2001; Ramos et al. 2003; Zhang et al. 2014). The interaction of FMR1 with G-quadruplexes is dependent on the C-terminal RGG box of FMR1, and is crucial for many functions of the protein, such as FMR1 dependent RNA localisation to neuronal projections (Goering et al. 2020). Binding motifs have also been identified for the KH domains of FMR1. A study showed that the KH2 domain of FMR1 has a binding affinity for ACUK and WGGA (K=G/U, and W=A/U) motifs in RNA (Ascano et al. 2012). Another study showed that the KH2 domain also exhibits binding to a complex RNA structure called “kissing complex” (loop-loop pseudoknots) (Fig. 2.11) (Darnell et al. 2005).

Aim:

The aim of this study was to investigate the role of FMR1 as an *oskar* mRNP component, and understand what aspects of *oskar* regulation it might be involved in. I undertook a biochemical and genetic approach in order to validate the interaction of FMR1 with *oskar*, understand the functional relevance of the interaction *in vivo*, and obtain a mechanistic insight into the functions of FMR1.

2.2 Results

2.2.1 Validation and physical characterisation of *oskar*-FMR1 interaction

a) *oskar* mRNA and FMR1-GFP co-localise *in vivo*

As previously mentioned, an *oskar*-specific pulldown performed by Frank Wippich led to the identification of FMR1 as a novel protein component of *oskar* RNP granules. In order to validate this putative interaction between *oskar* mRNA and FMR1, and confirm the association of FMR1 with *oskar* RNP granules, I performed a co-localisation analysis between *oskar* RNA and FMR1 in flies expressing FMR1-GFP. *oskar* mRNA was detected using single molecule fluorescence *in situ* hybridization (smFISH). I observed that FMR1-GFP co-localises with *oskar* mRNA in the nurse cells, the oocyte as well as the posterior pole of the oocyte where *oskar* localises (Fig. 2.12A and B). This showed that FMR1 might be a part of the *oskar* RNP granules *in vivo*, and that the protein associates with the granules already in the nurse cells.

b) FMR1-GFP directly binds to *oskar* mRNA

The transcript-specific pulldown performed to identify *oskar*-RNP components relies on UV crosslinking (that crosslinks direct protein-RNA interactions) and formaldehyde crosslinking (that crosslinks protein-protein interactions). It therefore enriches not only proteins that directly associate with *oskar*, but also indirect binding partners that are part of the granule via protein-protein interactions. Therefore it is necessary to confirm whether the enrichment of FMR1 with *oskar* mRNA is a result of direct binding of FMR1 to *oskar*, or indirect binding. To test this, a crosslinking and immunoprecipitation (CLIP) experiment was performed, by Frank Wippich, in flies expressing FMR1-GFP. The RNA-protein complexes were crosslinked *in vivo* using UV, creating a covalent bond between the bases and protein residues at zero distance. The RNA-FMR1-GFP complexes were then pulled down using anti-GFP antibody, and the enriched RNAs were analysed for the presence of *oskar* by qRT-PCR. As compared to the non-crosslinked sample, *oskar* mRNA was significantly enriched in the UV crosslinked samples (Fig. 2.12C), indicating that FMR1-GFP and *oskar* mRNA engage in a direct interaction *in vivo*.

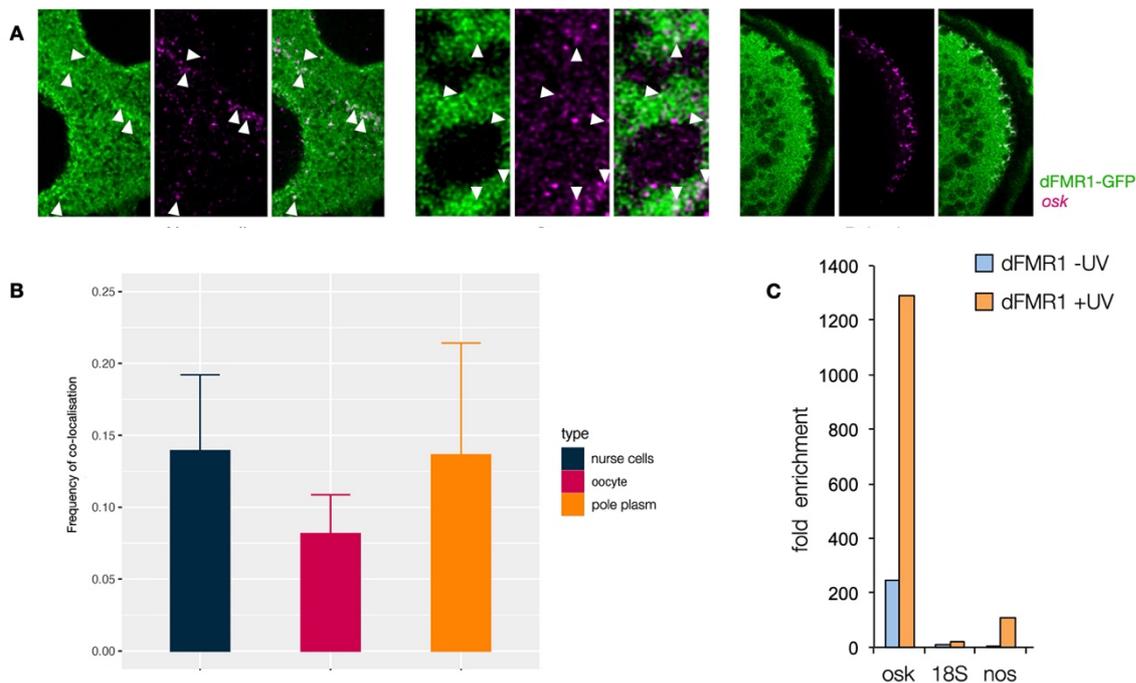


Fig. 2.12: A) FMR1-GFP and *oskar* mRNA significantly co-localize in the nurse cells, ooplasm and the posterior pole of the oocyte (the pole plasm). *oskar* mRNA was labelled by smFISH and FMR1-GFP was visualised via GFP fluorescence. B) The graph shows the frequency of co-localisation in each compartment. Error bars represent S.E. C) *oskar* mRNA is significantly enriched with FMR1-GFP in UV crosslinked samples as compared to non-crosslinked samples, indicating a direct interaction between the FMR1-GFP and *oskar* mRNA (performed by Frank Wippich). S.E. Standard error

c) *oskar* 3'UTR co-localises with FMR1 *in vivo*

The 3'UTR of *oskar* is a binding hub for several RBPs, important for regulating the functions of *oskar*. With the evidence that FMR1 directly interacts with *oskar* mRNA (Fig. 2.12C), I wanted to see if FMR1 interacts with the *oskar* 3'UTR *in vivo*. To this end, I performed a co-localisation analysis in flies expressing FMR1-GFP and *oskar* 3'UTR in an *oskar* RNA null background. A null background was crucial since *oskar* 3'UTR is known to dimerise. Transgenic *oskar* 3'UTR can thereby hitchhike with endogenous *oskar* RNA granules, making it difficult to distinguish between transgenic and endogenous *oskar* (Jambor et al. 2011). *oskar* 3'UTR was visualised using smFISH and FMR1 using GFP fluorescence, and the co-localisation was assessed in the nurse cells and the oocyte, but not at the posterior pole, since *oskar* 3'UTR is incapable of localising to the posterior due to the lack of SOLE, a splicing dependent localisation element present in the coding region of *oskar* required for posterior localisation (Ghosh et al. 2012). I observed that FMR1-GFP and *oskar* 3'UTR

significantly co-localised in the nurse cells as well as in the oocyte (Fig. 2.13). This suggests that FMR1-GFP possibly interacts with the *oskar* 3'UTR.

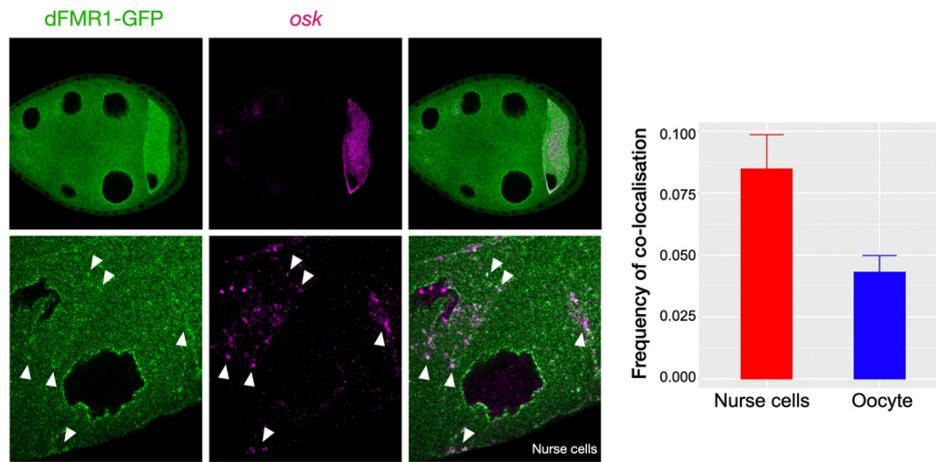


Fig. 2.13: FMR1-GFP and *oskar* 3'UTR significantly co-localise in the nurse cells and the oocyte, indicating an interaction between the two. Scale bar – 8 μ m. The graph shows the frequency of significant co-localisation. Error bars represent standard error.

d) The C-terminal domain of FMR1 is required for *oskar* 3'UTR binding *in vitro*

Next, I wanted to determine which of the two RNA binding domains of FMR1 (KH domains or RGG box) is required for binding to *oskar* mRNA. To this end, I generated and purified two GFP tagged recombinant proteins: KH(1_2)-GFP which contains KH domains 1 and 2, and KH(1_2)CTD-GFP which contains both the KH domains and the C-terminal domain including the RGG box. To test the ability of FMR1 domains to bind to the *oskar* 3'UTR, which shows co-localisation with FMR1 *in vivo*, I performed an *in vitro* binding assay. The *oskar* 3'UTR RNA and the protein were UV-crosslinked, after which a low concentration of RNase I was added to digest away the accessible portions of the RNA. The RNA-protein complexes were then pulled down using anti-GFP magnetic trap beads, and the samples were radioactively labelled using polynucleotide kinase, which transfers γ -P³² from γ -P³²-ATP to the 5'-OH of RNAs. Thus, only samples with bound RNAs are labeled.

I observed that only KH(1_2)CTD-GFP, when crosslinked with UV, showed a radioactive signal indicative of RNA binding (Fig. 2.14A). The KH domains alone (KH(1_2))

showed no binding to *oskar* 3'UTR RNA *in vitro*. This indicates that the C-terminal domain of FMR1 is required for binding to the *oskar* 3'UTR, at least *in vitro*, and that the KH domains alone are not sufficient.

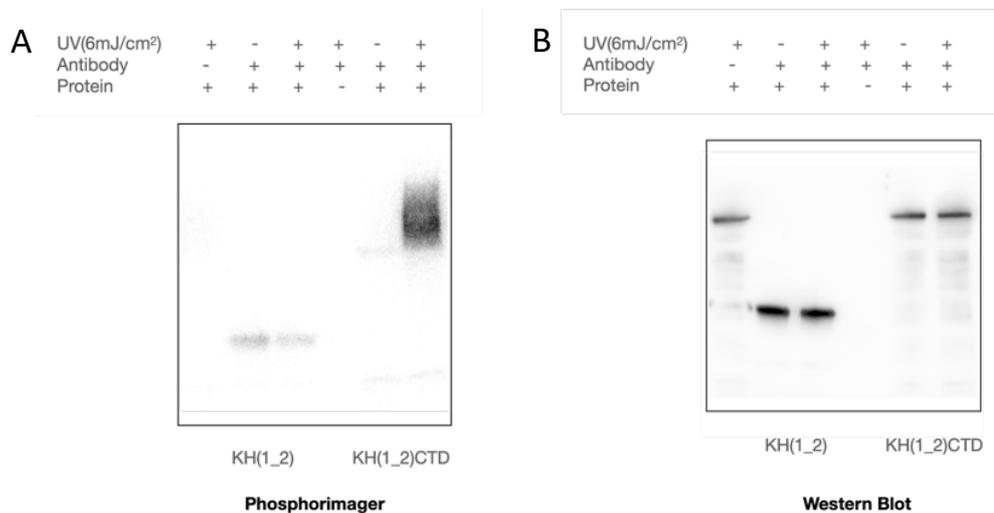


Fig. 2.14: A) KH(1_2)CTD shows a radioactive signal upon UV crosslinking indicative of RNA binding (lane 5 and 6). KH domains alone do not exhibit *oskar* 3'UTR binding *in vitro* (lane 2 and 3). No antibody and protein controls show no non-specific RNA binding (lane 1 and 4). B) Protein loading control for experiment in A).

e) iCLIP reveals the binding site dynamics of FMR1 on *oskar*

The pulldown assay (Fig. 2.12C) confirmed that there is a direct interaction between FMR1 and *oskar* mRNA *in vivo*. To identify the binding sites of FMR1 on *oskar* mRNA, an iCLIP (Individual nucleotide CrossLinking and ImmunoPrecipitation) experiment was performed by Matteo Bordi (a former PhD student in the lab) with the help of our collaborator Julian König at IMB, Mainz, Germany. iCLIP offers a nucleotide resolution view of binding sites of a protein on the transcripts with which it associates (Huppertz et al. 2014). iCLIP relies on the formation of a covalent bond between the protein and RNAs upon UV crosslinking, followed by Proteinase K digestion. This results in the presence of a peptide at the crosslink site that stalls the reverse transcriptase leading to the synthesis of cDNAs truncated at the crosslinked site. The position of the crosslink is then identified with respect to where the cDNA was truncated. The bioinformatic analysis for this study was performed by our collaborators Danilo Lüdke and Kathi Zarnack at BMLS, Goethe University Frankfurt. In order to identify the binding sites of FMR1 on target RNAs, peak calling analysis was performed using

PureCLIP as mentioned in (Busch et al. 2020). PureCLIP detects regions with enriched mapped reads, and the crosslink sites are identified as the ones where a significant fraction of reads accumulate around the same nucleotide as identified from the truncated cDNAs.

Particularly focusing on the binding pattern of FMR1 on *oskar* mRNA, two main observations could be made: one was the presence of FMR1 binding peaks throughout the coding region of *oskar* (Fig. 2.16A, top panel). This was not entirely surprising, as previous CLIP and Tribe analyses performed in mouse and *Drosophila*, respectively, showed that the majority fraction of peaks associated with FMR1 binding are located in the coding region of the target transcripts (McMahon et al. 2016) (Fig. 2.15A). Analysis of our iCLIP data also led to similar observations, wherein the majority fraction of peaks for FMR1 on target transcripts were also in the coding region (Fig. 2.15B). This consistent feature of FMR1 binding has been attributed to the role of FMR1 as a translational regulator. There is evidence of the direct association of FMR1 with ribosomes (Chen et al. 2014) and the binding peaks in the coding region are thought to be a consequence of this interaction. The presence of peaks in the coding region of *oskar*, therefore, also points towards a possible role of FMR1 in regulating the translation of *oskar*.

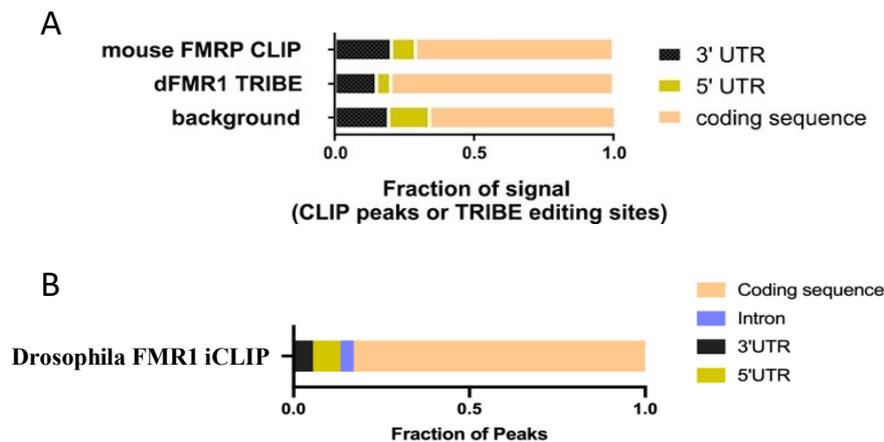


Fig. 2.15: A) Fraction of peaks in the 3'UTR, 5'UTR and coding sequence of target transcripts identified in published CLIP and Tribe experiments (adapted from (McMahon et al. 2016)). B) Fraction of peaks in the 3'UTR, 5'UTR, introns and coding sequence of target transcripts in the iCLIP experiment.

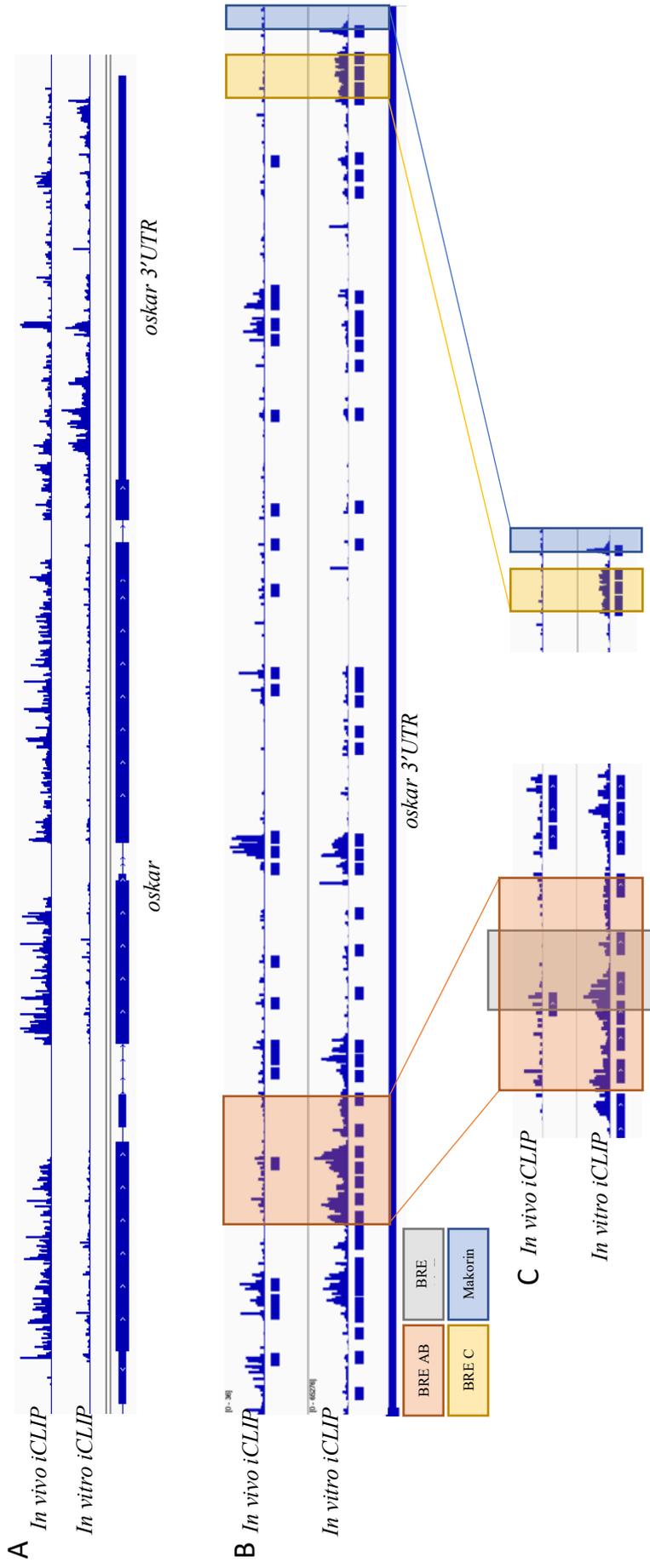


Fig. 2.16: A) The binding peaks of FMR1 on full length *oskar* mRNA as observed in *in vivo* and *in vitro* iCLIP. B) Zoom-in of the binding peaks in the *oskar* 3'UTR. Boxes below the peaks indicate the identified binding sites after peak calling. C) Zoom-in of binding peaks on BRE AB and BRE C in *oskar* 3'UTR

The second observation was that, apart from the coding region, binding peaks for FMR1 are also present in the 3'UTR of *oskar*. Peak calling analysis identified 21 potential binding sites for FMR1 in the *oskar* 3'UTR (Fig. 2.16B, top panel, boxes under the peaks indicate identified binding sites). This is also consistent with the co-localisation analysis that suggested that FMR1-GFP and *oskar* 3'UTR are closely associated *in vivo* (Fig. 2.13).

The binding peaks observed in the *in vivo* iCLIP experiments are a reflection of the presence of several other RNA binding proteins and factors present in the cell, and not the intrinsic binding capacity of the protein under study. To determine the intrinsic binding landscape of a given protein on an RNA of interest, the König Group at IMB, Mainz developed the *in vitro* iCLIP method, which identifies the binding sites of a recombinant protein on RNA *in vitro*, in the absence of all cellular factors. Comparing this to the *in vivo* landscape can provide insight into the binding dynamics of the protein on the RNA of interest (Sutandy et al. 2018). Studying this is especially interesting for *oskar* mRNA, since the *in vivo* data displayed extensive peaks in the coding region, which are thought to be due to the presence of ribosomes and not the protein itself. I therefore performed an *in vitro* iCLIP experiment using recombinant FMR1(KH1_2)CTD-GFP and *in vitro* transcribed full length *oskar* mRNA in collaboration with Julian König at IMB, Mainz. The cDNA library preparation from the extracted RNAs (see Methods) was performed by our collaborator Anna Orekhova from the König Group. The bioinformatic analysis for data quality control was performed by Anke Busch from IMB Mainz, and the peak calling analysis was performed by Danilo Lüdke and Kathi Zarnack from BMLS, Goethe University Frankfurt.

In vitro iCLIP revealed that the binding peaks in the coding region of *oskar* were absent *in vitro*, indicating that the presence of the binding peaks *in vivo* was indeed due to the presence of other cellular factors, probably ribosomes, and not due to the affinity and binding specificity of FMR1 (Fig. 2.16A, bottom panel). Furthermore, I observed that the binding pattern of FMR1 in the 3'UTR of *oskar* *in vitro* was quite similar to the *in vivo* pattern, except for two parts of the 3'UTR. One, at the Bruno Response Element (BRE) AB and second, at the BRE C site (Fig. 2.16B and C).

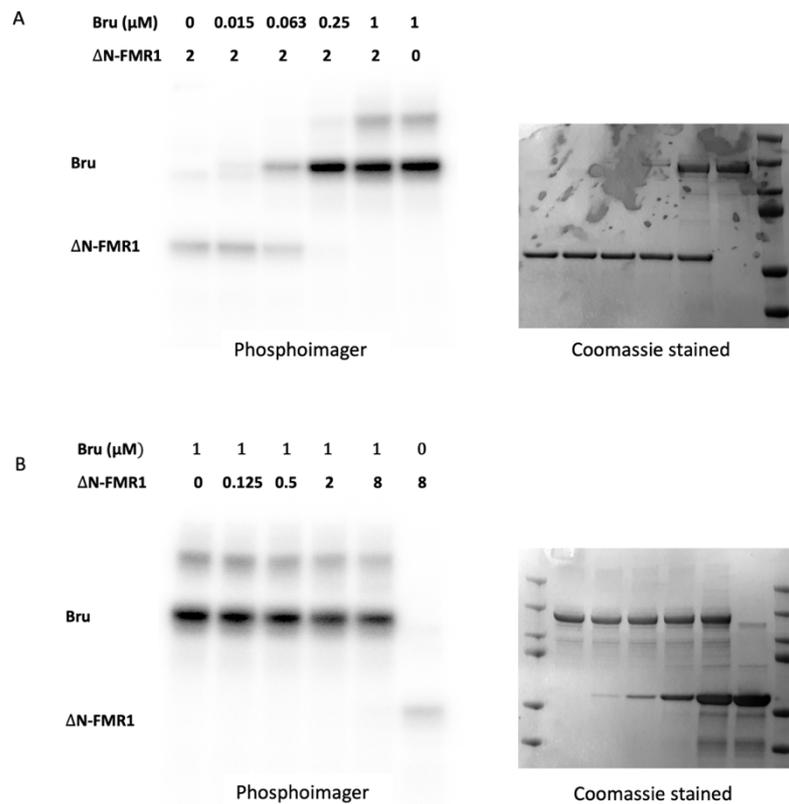


Fig. 2.17: A) With increasing concentrations of Bruno successfully outcompetes FMR1 binding to BRE A' II. B) FMR1 has much lower affinity for oligo BRE A' II as compared to Bruno and cannot outcompete Bruno for binding to the oligo even at 8 times higher concentration. The gels were stained with Coomassie blue after exposure to visualise protein amounts.

BREs are the binding sites for Bruno protein, one of the most important translational repressors of *oskar*. Bruno binds to *oskar* in the nurse cells and keeps the mRNA in a repressed state during its transport to the posterior pole. Increased binding of FMR1 *in vitro* to the BRE sites suggests that FMR1 has intrinsic binding affinity for those sites, but that it is unable to do so *in vivo*, possibly because of Bruno binding. Thus, to test if Bruno can prevent the binding of FMR1 to the BREs, I performed an *in vitro* UV crosslinking assay, adapted from (Zarnack et al. 2013), using an RNA oligonucleotide corresponding to the BRE A' II region of BRE AB (Fig. 2.16C), and the recombinant proteins $\Delta\text{N-FMR1}$ (which contains both KH domains and the RGG box, but lacks the N-terminal agenet-like domain), and MBP-GFP-Bruno. $\Delta\text{N-FMR1}$ and radioactively labeled RNA oligo were co-incubated, following which Bruno was added at different (increasing) concentrations. The samples were then UV crosslinked and run on an SDS-PAGE. I found that as the Bruno concentration was increased, the protein could successfully outcompete FMR1 binding to the oligo, even at low concentrations (Fig 2.17A).

On the contrary, when Δ N-FMR1 was added at increasing concentrations to the Bruno-oligo complex, Δ N-FMR1 was unable to displace Bruno even at concentrations 8-fold higher than Bruno (Fig. 2.17B). This indicates that Bruno has a much higher affinity for BRE A' _II than does FMR1, and indeed the presence of Bruno prevents FMR1 from binding to the BRE A' _II region.

2.2.2 Functional analysis of *oskar*-FMR1 interaction

a) Loss of FMR1 leads to reduced Oskar protein levels *in vivo*

The presence of binding peaks throughout the coding region, and the competition between FMR1 and the translational repressor Bruno for binding to the BRE AB region suggest a role of FMR1 in regulating the translation of *oskar* mRNA. Furthermore, competition between FMR1 and Bruno points towards possibly antagonistic roles of FMR1 and Bruno in translational regulation of *oskar*. Therefore, to check if absence of FMR1 has any effect on Oskar protein levels, I performed a western blot analysis of Oskar protein levels in FMR1 knock-down line and in FMR1 loss-of-function mutant line. To knock-down FMR1 specifically in the germline, I used *oskar*-Gal4 to drive the expression of shRNA in FMR1-TRiP line. To generate the FMR1 loss-of-function mutant line, I crossed the heterozygous deletion mutants $Fmr1^{\Delta 50}/Tm6B$ and $Fmr1^{\Delta 113}/Tm6B$ lines and used the $Fmr1^{\Delta 50}/Fmr1^{\Delta 113}$ progeny for the analysis. Both $Fmr1^{\Delta 50}$ and $Fmr1^{\Delta 113}$ are loss of function, amorphic alleles of FMR1. FMR1 protein levels in both the knock-down and loss-of-function mutants were analyzed using western blot (Fig. 2.18A).

I observed that for both the knock-down and loss-of-function mutants, the Oskar protein levels were reduced to half the wild type levels (Fig. 2.18A), and that the short and long isoforms of Oskar protein were similarly affected. To check if this might be a consequence of reduced *oskar* RNA levels, I performed a qRT-PCR analysis of *oskar* mRNA in wild type and FMR1 knock-down line. The analysis revealed no significant change in the *oskar* mRNA levels upon *fmr1* knock-down (Fig. 2.18C). These data suggest that FMR1 affects either mRNA translation of *oskar*, or Oskar protein stability.

Studies on the translation activation of *oskar* have shown that the length of the poly(A) tail of *oskar* is critical for the proper translation of *oskar* mRNA. Orb, a *Drosophila* homolog of Cytoplasmic Polyadenylation Element Binding protein (CPEB), plays a crucial role in stimulating the polyadenylation of *oskar* (Castagnetti and Ephrussi 2003) and thereby its translation. Therefore, I wished to determine if the low Oskar protein levels in the *fmr1* knock-down are due to an altered length of the poly(A) tail of *oskar*. To this end, I performed an ePAT assay (Jänicke et al. 2012) to check for any difference in the poly(A) tail length of *oskar* in *fmr1* knock-down line. This method relies on annealing a DNA oligo to adenylated RNA

b) Knockdown of FMR1 leads to reduced number of pole cells in embryos

Oskar protein is important for several molecular and cellular processes in the oocyte and the embryo. One of these functions is the assembly of the pole plasm, a specialized cytoplasm at the posterior pole of the oocyte required for the formation of pole cells, the progenitors for the future germ cells (Ephrussi and Lehmann 1992). Oskar protein nucleates the pole plasm assembly, and the amount of Oskar protein expressed in the oocyte directly affects the number of pole cells formed (Ephrussi and Lehmann 1992). Previous data from (Deshpande et al. 2006) has shown that there is a reduction in the pole cell numbers to almost half in *fmr1*³ protein null flies. I could also re-confirm the data for *fmr1* knock-down line (Fig. 4), where I stained the pole cells using antibodies against vasa, a protein marker for pole cells, and observed a roughly 50% reduction in the number of pole cells (Fig. 2.19). Since, FMR1 knockdown leads to reduced Oskar protein levels *in vivo*, this reduction in the number of pole cells is likely to be a direct consequence of reduced Oskar protein levels.

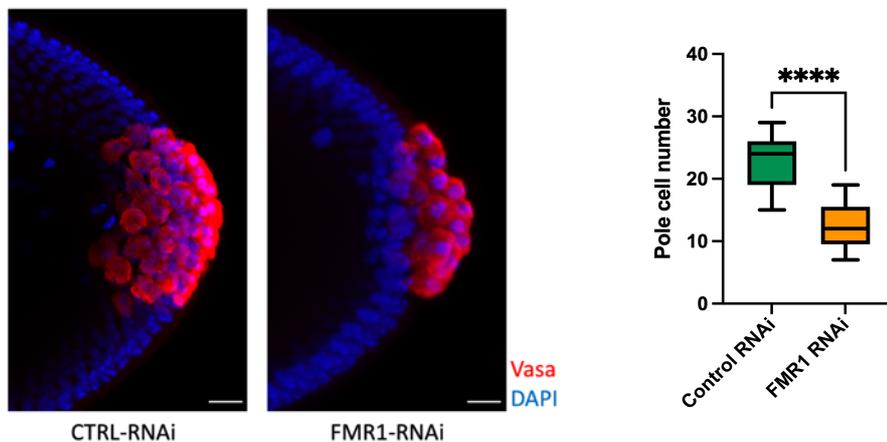


Fig. 2.19: FMR1 knockdown leads to an approximately 50% reduction in the number of pole cell as compared to the control knockdown line, scale bar – 10 μ m. T-test was used for statistical analysis, and p-value is <0.0001. Error bars represent standard deviation.

c) FMR1 knockdown does not affect abdominal patterning or hatching rate

Another important function of Oskar protein in the embryo is in proper abdominal patterning. To see whether the reduced Oskar protein levels upon FMR1 knock-down also

affect the abdominal patterning of embryos, I performed a cuticle pattern analysis in embryos knocked down for FMR1. Wild type embryos contain three thoracic segments and eight abdominal segments, and the reduction in Oskar protein levels can cause a series of defects in the abdominal segment formation, depending on the strength of the mutants. Strongest *oskar* mutant alleles lead to complete absence of abdominal segments, whereas weaker alleles disrupt the formation of middle abdominal segments (Nüsslein-Volhard et al. 1987). I observed no defects in patterning upon *fmr1* knock-down (Fig. 2.20A). This indicates that the amount of Oskar protein produced was sufficient for normal patterning. Furthermore, I also checked the hatching rate of FMR1 knock-down embryos as compared to w^{1118} , and again observed no difference (Fig. 2.20B).

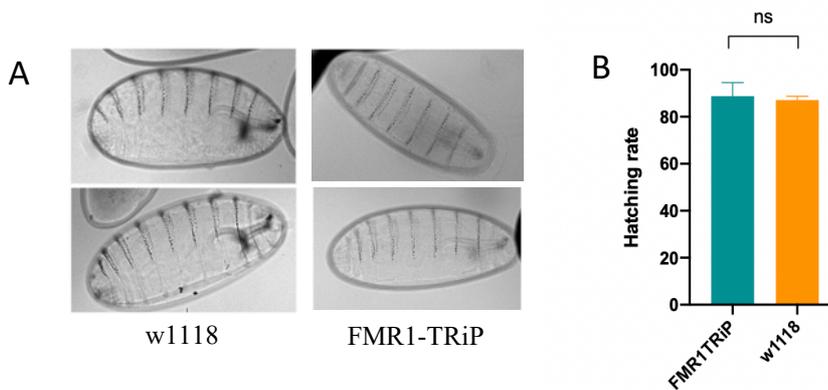


Fig. 2.20: A) Cuticle preparations for FMR1 knockdown and wild type (w^{1118}) line revealed no patterning defects in embryos. B) Hatching rate analysis of the FMR1 knockdown and the control (w^{1118}) line also revealed no significant difference (t-test used for statistical analysis). Error bars represent standard deviation.

d) FMR1 associates with *oskar* and Me31B in starvation induced P-bodies

The RNA-protein complexes in a cell are dynamic complexes with a constantly changing proteomic landscape. This is especially important for efficiently responding to changing environments and signals. Starvation, for instance, is one cue that leads to several molecular changes that allow a cell to better respond to the limiting conditions in order to survive. In the *Drosophila* germline, the starvation response is an important checkpoint during oocyte development. An amino acid deficient diet leads to the formation of starvation induced P-bodies in the mid-stage oocytes, and *oskar* mRNA is an essential component of these granules,

along with other known P-body markers like Me31B and Tral (Branislava Rankovic 2020, PhD Thesis DOI: 10.11588/heidok.00027585). Using flies expressing FMR1-GFP, I performed a combined smFISH and immunostaining to visualise *oskar* and Me31B respectively in starved oocytes, and found that FMR1 also closely associates with these granules (Fig. 2.21A). A co-localisation analysis performed by Frank Wippich to quantify the association of FMR1 with *oskar* in starvation vs well fed conditions (Fig. 2.21A) suggests an increased association of FMR1 with *oskar* upon starvation in the nurse cells, but a lower association at the posterior pole. This, however, might be due to the fact that the *oskar* localisation at the posterior pole is also disrupted upon starvation, as *oskar* associates with the P-bodies in the ooplasm.

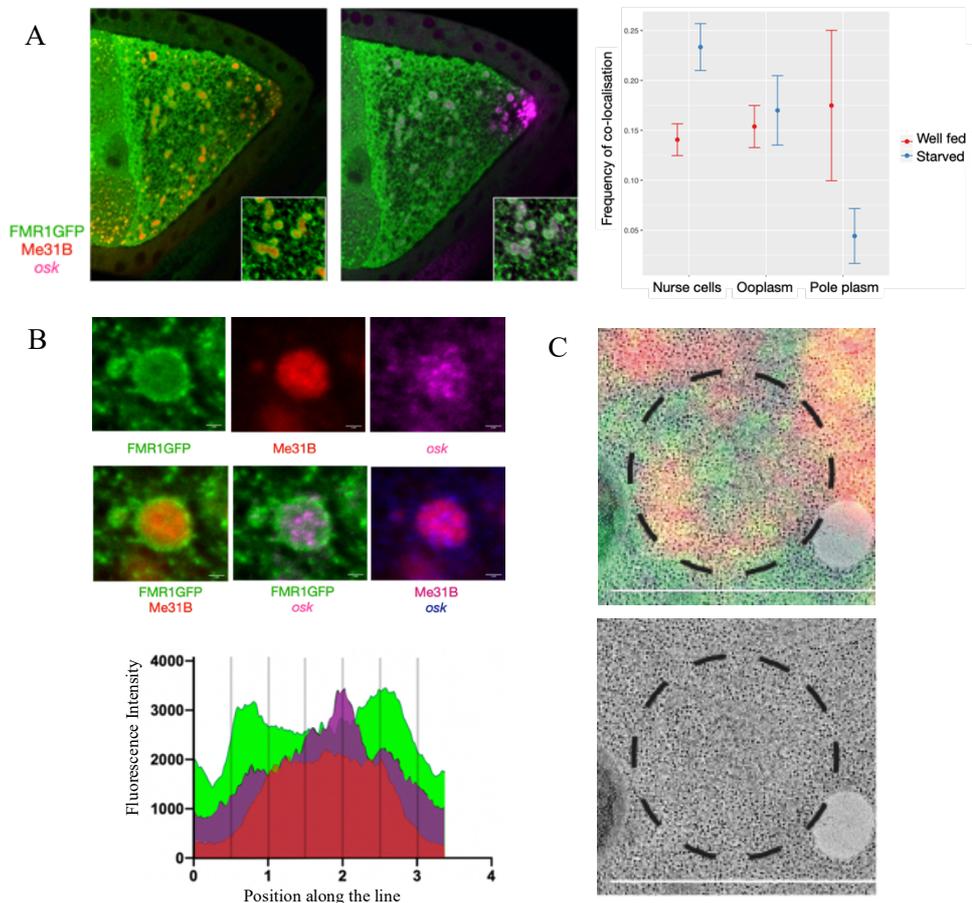


Fig. 2.21: A) FMR1-GFP localises to P-bodies with Me31B and *oskar* upon nutritional starvation. The graph shows changes in co-localisation frequency between FMR1-GFP and *oskar* mRNA in nurse cell, oocyte and the posterior pole in well-fed vs starved conditions (Graph by Frank Wippich). B) Co-staining of FMR1-GFP, Me31B and *oskar* mRNA in nutritionally starved flies shows that Me31B and *oskar* are present in the core of the granule, whereas FMR1-GFP is enriched peripherally. The graph shows the intensity profile for the three components along a line drawn through the granule. Scale bar – 1µm C) CLEM experiments show absence of any “zones” around the granules, indicating that FMR1 does not form any visible structural regions/limits for the granules. Scale bar - 2µm

Careful analysis of the starvation induced granules further revealed a specific spatial enrichment pattern for FMR1, *oskar* and Me31B within the granules. I observed that while *oskar* and Me31B are enriched at the core of the granules, FMR1-GFP exhibited a rather peripheral enrichment (Fig. 2.21B). In order to obtain an insight into the structure of the granules, I performed a Correlative Light and Electron Microscopy (CLEM) for the granules using flies expressing GFP-Me31B and *oskar*-Gal4 driven RFP-FMR1, with Paolo Ronchi from the Electron Microscopy Core Facility (EMBL). To identify the granules, fluorescence images were overlaid on the electron micrographs. I thus observed that the granules appear as spongy membraneless structures. There were however no visible boundaries or structural features within or around the granules at this scale (Fig. 2.21C).

2.2.3 Mechanism of translational regulation by FMR1

a) FMR1 uses KH domains for activation and RGG-CTD for repression

FMR1 has been implicated in repressing as well as activating translation. This raises the question of how a protein is able to perform two antagonizing roles. FMR1 contains two types of RNA binding domains: two KH domains, and an RGG box in the C-terminal domain (CTD). The KH domains are known to be important for binding directly to ribosomes (Chen et al. 2014) and together with the C-terminal domain inhibit translation elongation. However, recent evidence on human FMRP shows that the RGG-CTD alone is also able to interact with ribosomes/polysomes and is sufficient to inhibit translation (Athar and Joseph 2020; D'Souza et al. 2021). Therefore, I decided to perform an *in vitro* tethering assay to systematically test the effect of the different RNA binding domains of FMR1 on regulating the translation of the reporter RNAs in *Drosophila* embryo extracts. I used the lambda phage λ N-BoxB interaction to tether the FMR1 proteins to the reporters of interest. I generated three recombinant proteins N-terminally tagged with λ N-sfGFP: KH(1_2)CTD contains both the KH domains and the RGG-CTD, KH(1_2) contains only the KH domains, and CTD contains the RGG box and the C-terminal domain. I also *in vitro* transcribed three different luciferase reporter RNAs: lucBoxB which contains a luciferase reporter tagged with 5x BoxB loops, and luciferase reporters containing the *oskar* 3'UTR, either containing (lucoskBoxB) or lacking (lucosk) BoxB loops (Fig. 2.22A). The luciferase activity was used as a readout for the translation efficiency of the luciferase reporter RNAs. A 'no protein' control was used to determine the basal level translation of the reporter RNAs, and the effect of tethering FMR1 proteins on the translation of the reporters was determined as the relative luciferase activity upon tethering FMR1 as compared to the no protein control.

Inclusion of FMR1 KH(1_2)CTD and KH(1_2) at a concentration of 350nM led to a 6.7 and 8.2 fold increase respectively in the translation of lucBoxB (Fig. 2.22A), as compared to the no protein control. The stimulation of translation was lower for lucoskBoxB, which contains the *oskar* 3'UTR, and there was almost no change for lucosk reporter, which lacks the BoxB loops. The lower stimulation of translation of lucoskBoxB upon tethering KH(1_2)CTD and KH(1_2) could possibly be due to a regulatory effect of other factors in the

Drosophila embryo lysate on the *oskar* 3'UTR. The FMR1 CTD, on the other hand, had no effect on the translation of the reporters at the 350nM concentration (Fig. 2.22B).

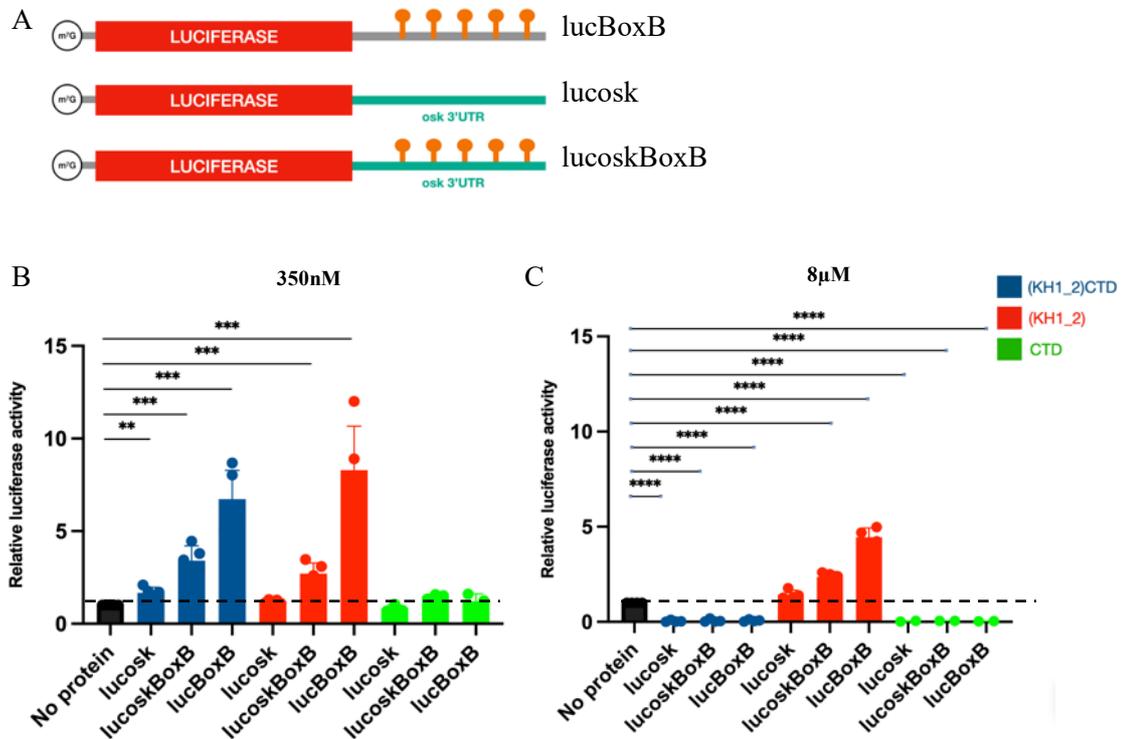


Fig. 2.22: A) Schematic representation of the reporters used for the tethering assay. Graphs show the relative luciferase activity of the three reporters upon tethering different FMR1 domains at a concentration of B) 350 nM and C) 8µM. Error bars represent standard deviation. P values: ****<0.0001, ***<0.0005, **<0.005

Next, I repeated the experiment using a protein concentration of 8µM, and found that at this concentration, both KH(1_2)CTD and CTD show a significantly repressed translation of all three reporters (Fig. 2.22C). KH(1_2), on the other hand, still stimulated the translation for BoxB containing reporters, though slightly less than at the 350nM concentration (Fig. 2.22C). This indicates that, as also found in the case of human FMRP, tethering of the C-terminal domain of FMR1 is sufficient to repress the translation of reporter RNAs *in vitro*. It also shows that KH domains of FMR1 do not have the ability to repress translation, but rather have an activating effect on the translation of reporter RNAs upon tethering. Furthermore, when both the domains are present (KH(1_2)CTD), there seems to be a protein concentration dependent effect on the translation of the reporters, such that at very high concentration KH(1_2)CTD represses the translation of reporter RNAs, but at lower concentrations, KH(1_2)CTD stimulates the translation of reporter RNAs. Recent data from (Kim Tae et al. 2019; Tsang et

al. 2019) show that the C-terminal domain of FMR1 has the ability to phase separate and that this property is essential for translation repression, by possibly excluding translation initiation factors such as eIF4E from the granules so formed. Therefore, I wished to determine if the repression observed at 8 μ M is also associated with phase separation of the FMR1 protein. Visualising the reaction mix under a confocal microscope, I observed that KH(1_2)CTD phase separates at the concentration of 8 μ M, but not at the concentration of 350nM (Fig. 2.23A), reinforcing the observations from (Kim Tae et al. 2019; Tsang et al. 2019) that repression of translation by FMR1 is associated with its property to phase separate.

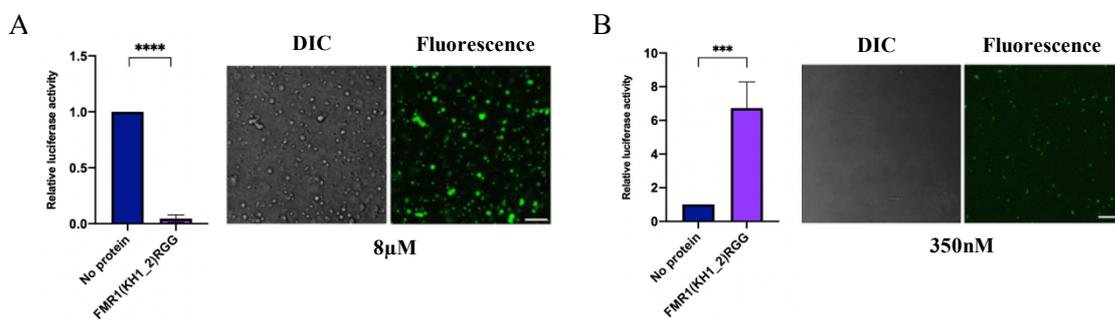


Fig. 2.23: A) At a concentration of 8 μ M, KH(1_2)CTD phase separates, and there is repression of translation of lucBoxB reporter. B) No phase separation of KH(1_2)CTD occurs at 350 nM protein concentration, and there is an activation of translation of the lucBoxB reporter. Error bars represent standard deviation. DIC: Differential Interference Contrast. Scale bar – 10 μ m. P values: ****<0.0001. ***<0.0005

b) Interaction of FMR1-KH domains with ribosomes is required for translation activation *in vitro*

There is evidence that both the KH domains, and the C-terminal domain of FMR1 can independently interact with ribosomes and polysomes (Chen et al. 2014; D'Souza et al. 2021). Interaction of KH domains with ribosomes was previously thought to be important for translation repression, but there is now ample evidence to indicate that the KH domains of FMR1 are not required for repression of translation (Fig. 2.22 B and C) (Athar and Joseph 2020; D'Souza et al. 2021). To study if the association of KH domains with ribosomes might instead be important for translation activation, I generated two KH domain mutants: KH(I244N) and KH(I307N). KH(I244N) mutant has been shown to exhibit a reduced binding of FMR1 to ribosomes *in vitro* (Chen et al. 2014), and KH(I307N) mutant has been shown to have reduced binding of FMR1 to polyribosomes *in vivo* in mammals (Feng et al. 1997).

I analysed the two mutants for their translation stimulating effect on the three reporter RNAs, and observed that the KH(I244N) mutant failed to activate the translation of the reporters (Fig. 2.24A, B and C). KH(I307N) mutant still activated translation, though, the stimulatory effect was significantly lower than that of the wild type KH domains. Taken together, these *in vitro* translation data suggest that the interaction of KH domains with the ribosomes, which was previously thought to be necessary for repression, might actually be required for translational activation by FMR1.

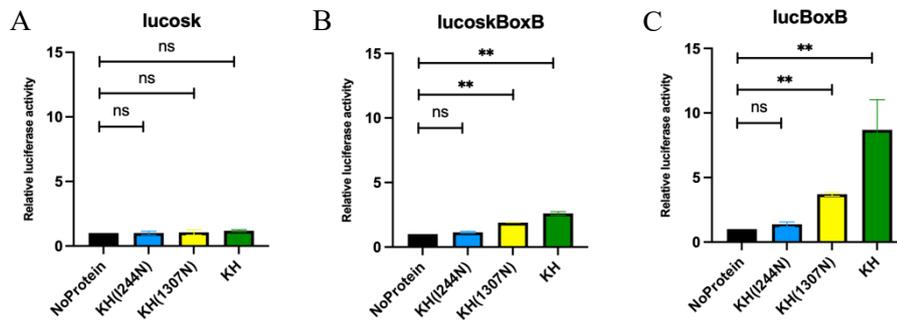


Fig. 2.24: Effect on translation of reporter RNAs A) lucosk B) lucoskBoxB, and C) lucBoxB upon tethering the FMR1 KH domain mutants KH(I244N) and KH(I307N). KH domain mutants have no/reduced ability to activate translation as compared to wild type KH domains. Error bars represent standard deviation. P value: **<0.005

2.3 Discussion

FMR1 is a newly identified protein component of the *oskar* mRNP granules in the *Drosophila* germline. It binds to *oskar* directly, and associates with the *oskar* mRNP granules already in the nurse cell cytoplasm. Absence of FMR1 leads to reduced Oskar protein levels *in vivo*, without any effect on RNA levels. The reduction in protein levels further leads to reduction in the number of pole cells, which is proportional to Oskar protein dosage. The reduction in protein levels is not due altered stability of the RNA, and is presumably a consequence of either reduced translation or reduced stability of the Oskar protein. Binding peaks of FMR1 in the *oskar* coding region, along with results of my *in vitro* translation assays of reporters containing *oskar* 3'UTR suggests that FMR1 might be involved in the regulation of translation of *oskar*, but this will have to be confirmed through further studies. Polysome fractionation would indicate if there is any change in the ribosome association with, or density on, *oskar* mRNA upon FMR1 knock-down.

iCLIP is a powerful technique to identify the binding sites of proteins on RNAs at a nucleotide resolution, and it revealed FMR1 binding peaks in the coding region as well as the 3'UTR of *oskar*. The coding region peaks, however, are not detected in the *in vitro* iCLIP indicating that other cellular factors are required for the association of FMR1 with the coding region. The binding sites in the 3'UTR, on the other hand, are also detected *in vitro*, suggesting that the 3'UTR harbors the binding sites of FMR1. However, one has to keep in mind that when using high throughput techniques like iCLIP, one loses spatial and temporal resolution, especially in tissues such as the *Drosophila* ovaries. *oskar* is transcribed in the nurse cells and transported into the spatially distinct oocyte, where it is translated only after stage 9. The binding peaks from an iCLIP experiment, therefore, represent the binding of FMR1 on *oskar* in all stages of oocyte development, on translationally repressed as well as de-repressed *oskar*. Any difference in binding of FMR1 in early versus late stage oocytes, or nurse cells versus posterior pole unfortunately cannot be deduced using the entire tissue, which contains egg chambers of all oogenetic stages.

The *in vitro* binding affinity of FMR1 for the BRE sites in the *oskar* 3'UTR, and the finding that Bruno outcompetes FMR1 for *oskar* RNA binding shows that there is a possible interplay between the two proteins in binding to the sites, and possibly regulating the

translation of *oskar*. iCLIP revealed that FMR1 does exhibit binding at the BRE AB site *in vivo*, albeit at a lower level when compared to the rest of the 3'UTR (Fig. 2.15C). At the posterior pole of the oocyte, where *oskar* is de-repressed, it is possible that Bruno binding is disrupted by dynamic remodeling, allowing FMR1 to bind *oskar*, and stimulate translation. The interplay between FMR1 and Bruno for regulating translation of *oskar* can be studied by generating BRE AB mutants, that have a disrupted binding of Bruno, but with no effect on the binding of FMR1. BRE AB mutant reporters show an increased translation due to a lack of repression by Bruno (Chekulaeva et al. 2006). Comparing the translation levels of these mutant reporters in the presence and absence of FMR1 could show if FMR1 contributes to this increased translation in the absence of Bruno binding.

In the case of *oskar*, FMR1 positively regulates Oskar protein levels *in vivo*, without exerting any effect on *oskar* RNA levels. There is now ample evidence to suggest a role of FMR1 as a translation activator. Along with FMRP, two other paralogs in humans - FXR1 and FXR2, have also been identified for their roles in positively regulating the translation of target RNAs (Vasudevan and Steitz 2007; Fernández et al. 2015). It is intriguing to understand how a protein can perform two such antagonizing functions. In this thesis, I have tried to address this using an *in vitro* tethering assay. Earlier studies had shown that the binding of KH domains close the tRNA binding site in ribosomes is important for the inhibition of translation elongation by FMR1 (Chen et al. 2014). However, recent data for human FMRP shows that the RGG box containing C-terminal domain is sufficient to repress translation of reporter RNAs (Tsang et al. 2019; Athar and Joseph 2020; D'Souza et al. 2021). I have shown that the C-terminal domain of *Drosophila* FMR1 is also sufficient for repressing translation of reporter RNAs *in vitro*, and that the KH domains might be involved in activating the translation of reporter RNAs. This effect seems to be dependent on ribosome binding, as KH mutants with disrupted binding to ribosomes exhibit reduced or no activating effect on translation. The *in vitro* binding experiments show that the C-terminal domain of FMR1 is essential for binding to *oskar* 3'UTR. It can thus be speculated that FMR1 might interact with *oskar* through the RGG box, and the KH domains might be necessary to regulate the translation of *oskar*. It will be interesting to see if the domain-dependency exists *in vivo*, and what would be the effect of deleting different RNA binding domains of FMR1 on its interaction with *oskar* and on Oskar protein levels. Of note, transgene expressing full length FMR1 CDS under UAS promoter was specifically expressed in the germline of flies knocked down for FMR1 using germline-

specific *oskar*-Gal4 driver. The transgene, however, could not rescue reduced Oskar protein levels in the oocyte (Fig. S1). This might be due to disrupted regulation of the transgene because of different expression pattern in the egg chambers under the *oskar*-Gal4>UAS promoter, and/or disrupted RNA regulation of the transgene, as it lacks the endogenous FMR1 introns and UTRs. Therefore, transgenes with genomic FMR1, or endogenous CRISPR deletions would have to be used for the analysis.

Another important outstanding question is how the two antagonistic activities of FMR1 are regulated *in vivo*. This would require extensive studies, but one can speculate that the different mechanisms in which the two domains are engaged, along with different binding partners in different temporal and spatial contexts, might play a role in regulating the opposing activities. The C-terminal domain of FMR1 is required for the repression of translation through the phase separation ability of the domain. In this case, the local concentration of FMR1 as well as other proteins in the immediate vicinity could promote the assembly and disassembly of the phase separated granules, leading to differential translation regulation. This might also be an interesting switch when responding to stress. Under stress conditions, FMR1 is known to contribute to the assembly of cytoplasmic granules through its RGG domain. These granules serve as sites of translationally repressing RNAs, with FMR1 also contributing to translation repression (Mazroui et al. 2002). It is possible that upon relief of stress, the granules are disassembled, and the function of FMR1 can now be switched, whereby it stimulates the translation of target RNAs through its KH domains.

3. EB1: an unconventional RBP, or not?

3.1 Introduction

RNA binding proteins (RBPs) are proteins that exhibit a physical interaction with RNAs, typically with the help of one or more RNA binding domains (RBDs), and regulate the functions of RNAs. RNA interactome capture studies, including one from our lab (Sysoev et al. 2016), have led to the identification of a multitude of novel RBPs in cells, many of which lack the conventional RBDs, and have no RNA related functions known till date. One of such novel, unconventional RBPs identified is the End-Binding Protein 1 (EB1), which is a master regulator of microtubule plus-end dynamics.

3.1.1 Identification of novel RBPs in *Drosophila melanogaster*

Over the years, several methods have been developed to identify RNA binding proteins in a cell. Most of these techniques were however, either targeted, and hence required prior information about the target RNAs and complexes, or relied on computational methods to identify novel RBPs by searching for the presence of known RNA binding domains in proteins. Such *in silico* methods, however, could not reveal RBPs with non-canonical RNA binding domains. To circumvent these issues, and with the aim of identifying novel RBPs in an unbiased and high-throughput manner, an mRNA interactome capture method was developed and applied to different cell lines and model systems (Castello et al. 2012; Ryder 2016; Sysoev et al. 2016; Wessels et al. 2016). The method relies on *in vivo* cross-linking the cells/embryos with UV light (254 nm), followed by pulling down the mRNA-protein complexes using oligo(dT) beads, and identifying the bound proteins using mass spectrometry.

Such studies led to the identification of several novel RNA binding proteins, many of which had no classical RNA binding domains or RNA related functions known. Studies on these putative, unconventional RBPs over the last few years, have extended our knowledge not only of the type of RBPs in a cell, but also of novel RNA binding domains and mechanisms of RNA binding. For example, Adenomatous polyposis coli (APC) is a microtubule plus end binding protein which was recently identified as an RBP that binds to RNA using a basic

unstructured region. 260 genes were identified as its *in vivo* targets using HITS-CLIP, including β 2B-tubulin mRNA, and APC is required for the localisation of β 2B-tubulin mRNA to dynamic microtubules (Preitner et al. 2014). Another protein identified in an interactome capture screen is the ubiquitin E3 ligase TRIM25. TRIM25 also uses a novel RNA binding domain (PRY/SPRY domain) to interact with RNA, and the RNA binding plays a role in regulating the ubiquitination activity of TRIM25 on self and another target protein ZAP (Choudhury et al. 2017). Studies on the newly identified RBPs have also shed light on novel mechanisms of regulating the functions of RBPs by virtue of their RNA binding activity, a process called riboregulation. A recent study has shown that binding of RNA by the glycolytic enzyme Enolase I can inhibit its enzymatic activity, a mechanism used to regulate glycolysis in embryonic stem cells (Huppertz et al. 2020).

RBPs regulate various aspects of RNA metabolism and their misregulation can lead to severe diseases. Many essential genes in humans have homologs in *Drosophila* and *Drosophila* is thus used as a model to study the molecular roles of proteins and RNAs (Pandey and Nichols 2011). It was therefore of interest to also identify the repertoire of RBPs in the *Drosophila* model system, to which end mRNA interactome capture was performed in *Drosophila* embryos by a former PhD student of our lab, Vasiliy Sysoev (Sysoev et al. 2016) and also by the Landthaler Lab in a contemporary study (Wessels et al. 2016). Both the studies led to the identification of EB1 as a novel RBP.

3.1.2 EB1 is a microtubule plus end binding protein

Microtubules constitute an important part of the cellular cytoskeleton. Formed by GTP dependent polymerisation of α/β tubulin dimers, the dynamic growth and shrinkage of microtubules by association and dissociation of tubulin, helps in regulating the various functions of microtubules such as providing scaffold and framework to the cell, providing tracks for molecular transport, and chromosome segregation during cell division among others. Microtubules are constituted of 13 protofilaments, assembled in a hollow core. Microtubules have an inherent polarity, with a fast growing end, called the plus end, where tubulin dimers associate and extend the microtubules, and a slow growing minus end. The polarity of microtubules is crucial for cellular processes such a transport using cytoskeletal motor proteins, as the motor proteins are plus-end or minus-end directed, transporting cargos

in different directions along microtubules (Gudimchuk and McIntosh 2021). Numerous microtubule binding proteins have been discovered and studied over the years that provide insight into how the microtubule dynamics are regulated *in vivo*. These proteins are recruited to the microtubules in a microtubule polarity dependent manner, with some binding to the plus ends (+TIP), and others to the minus ends (-TIP), regulating different aspects of growth, stability and disassembly of microtubules. One such evolutionarily conserved family of proteins is the End Binding or EB family of proteins. In humans, there are three EB protein homologs: EB1, EB2 and EB3 (Su and Qi 2001). In *Drosophila*, three genes CG3265, CG18190, and CG32371 were identified as closely related to human EB1, with CG3265 (also called DmEB1, referred to as simply EB1 hereafter) exhibiting the highest similarity (Elliott et al. 2005). EB1 is known to bind to the plus end of microtubules, and affect the growth dynamics by recruiting several other +TIP proteins such as CLIP 190 (Dzhinzhev et al. 2005). In mitotic *Drosophila* S2 cells as well as in the dividing embryo, depletion of EB1 leads to mislocalised spindle poles, malformed mitotic spindles and reduced astral microtubules (Rogers et al. 2002). Furthermore, in interphase S2 cells, EB1 depletion causes a severe reduction in microtubule dynamics of growth and shrinkage, without any effect on the organisation and distribution of microtubules (Rogers et al. 2002). In *Drosophila* axons, EB1 co-operatively interacts with two other microtubule binding proteins Tau and Msps and regulates the proper microtubule bundle formation and maintenance during axon development (Hahn et al. 2021).

From a structural point of view, EB1 is a 33kD protein with an N-terminal calponin homology domain, a flexible linker and a C-terminal EB-homology (EBH) domain (Fig. 3.1 A). The calponin homology domain is required for binding to microtubules, and there is also evidence that EB1 can directly bind to GTP through the same surface in the calponin homology domain that is required for microtubule binding (Gireesh et al. 2018). The C-terminal EBH domain is required for EB1 homodimerisation or heterodimerisation with the other EB family proteins through the coiled coil region (Fig. 3.1A), as well as for the interaction with other proteins through the EEY motif (Fig. 3.1A) (Akhmanova and Steinmetz 2008). As previously mentioned, the polymerisation of tubulin into microtubules is a GTP dependent process. GTP-loaded α/β tubulin dimers are polymerised at the growing plus end of the microtubules, followed by GTP hydrolysis to GDP-Pi, with a subsequent release of phosphate. Studies have shown that EB1 binds to microtubules in a nucleotide dependent

manner, preferentially binding to the GTP or GDP-Pi bound tubulin regions (Nehlig et al. 2017) (Fig 2.1 B).

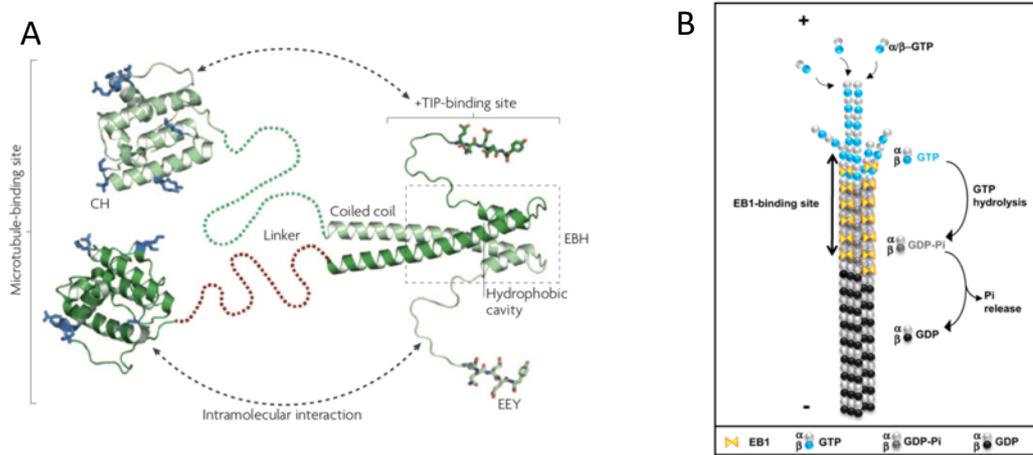


Fig. 3.1: A) Graphical representation of the domains of EB1 and their functions (from Akhmanova & Steinmetz, 2008). B) EB1 preferentially binds to microtubules in a nucleotide dependent manner, preferentially in the GTP or GDP-Pi states (from Nehlig et al., 2017)

3.1.3 Preliminary investigation of RNA binding properties of EB1 revealed its putative RNA binding surface

Identification of EB1 as a putative RNA binding protein warranted for an in-depth study of the possible RNA binding properties and functions of EB1. Preliminary biochemical assays were used to first confirm the RNA binding activity of EB1. To this end, a polynucleotide kinase (PNK) assay was performed by Vasily Sysoev in transgenic flies expressing EB1-GFP. Upon treating UV crosslinked RNA-EB1 complexes with increasing concentrations of RNase A (such that only a few nucleotides crosslinked to the protein remain), followed by labelling of the bound RNAs with ^{32}P using T4 PNK, discrete bands at the size expected for EB1-GFP (~60kD) could be seen, indicating a possible binding of EB1-GFP with RNAs *in vivo* (Sysoev et al. 2016) (Fig. 3.2A).

Furthermore, an Electrophoretic Mobility Shift Assay (EMSA) was performed by Lyudmila Dimitrova-Paternoga, a former postdoc in the lab, using recombinant EB1 and radioactively labelled oligonucleotide U₂₅. As the concentration of the protein was increased, there was a shift in the probe band, indicating the formation of EB1-oligo complex which runs

at a higher molecular weight than the probe alone (Fig. 3.2B) (Vaishali et al. 2021). This suggested that EB1 exhibits direct binding to oligo U₂₅ *in vitro*. EB1 however does not contain any of the known conventional RNA binding domains or motifs. Therefore, in order to identify what part of EB1 was involved in binding to RNAs *in vitro*, an NMR titration experiment was performed by Lyudmila Dimitrova-Paternoga using recombinant EB1 and oligo U₂₅. NMR spectroscopy is a sensitive method that can detect small changes in the surroundings of amino acids in a protein. If an amino acid residue plays a role in binding to an oligonucleotide, addition of the oligo to the protein solution causes a chemical shift in the peak of that residue on an NMR spectrum. These changes can then be mapped onto the structure of the protein to identify the binding surface for the oligo. Chemical shift perturbations detected in the NMR spectrum of EB1 in the presence of oligo U₂₅ revealed that the surface through which EB1 binds to RNA is the same surface that is required for binding to microtubules as well as GTP (Fig. 3.2D and E).

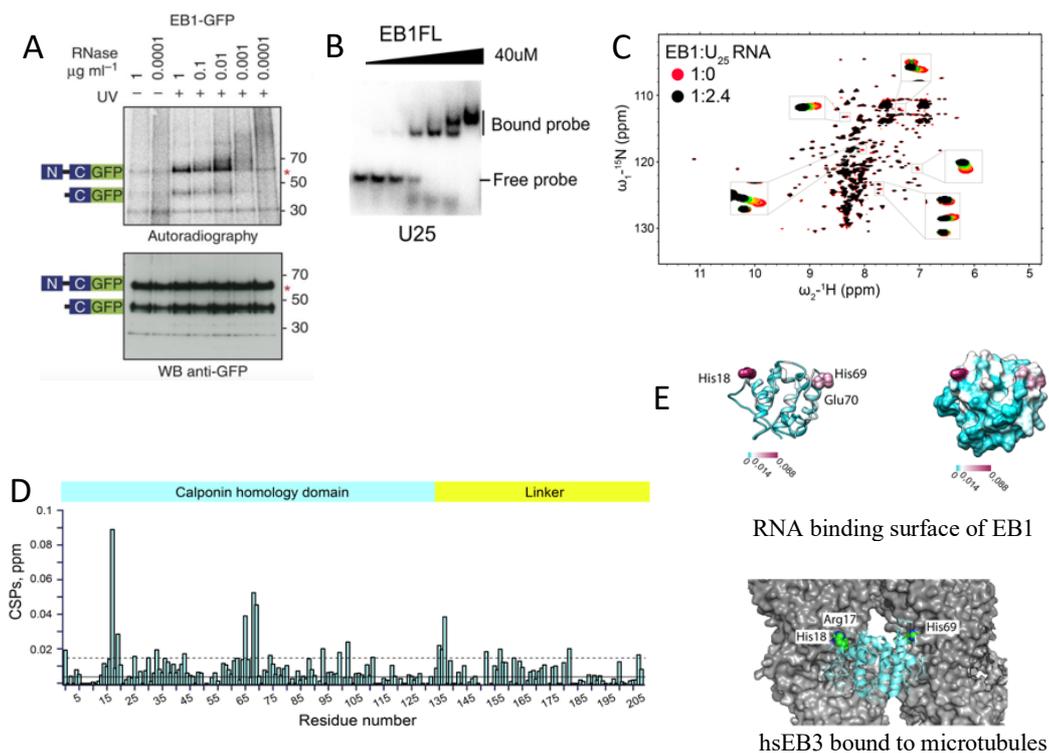


Fig. 3.2: A) PNK assay for EB1-GFP shows a discrete band at ~60kD, the molecular weight of EB1-GFP, upon high RNase treatment. Lower panel shows a western blot as a control for protein loading (image adapted from Sysoev et al., 2016). B) Electromobility Shift Assay (EMSA) for full length EB1 with U₂₅ RNA oligo shows a shift of the probe band as the concentration of EB1 is increased. C) NMR spectra for EB1 titrated with U₂₅ RNA oligo. Magnified boxes show shifts in certain amino acids. D) Graph showing the chemical shift perturbations (CSPs) along the N-terminal and linker region of the protein upon titration with U₂₅ RNA oligo. E) CSPs indicating RNA binding surface of N-terminal domain of EB1 plotted on its homology model (ribbon on left and surface on right, top panel) and the microtubule binding surface of human EB3 (bottom panel) (B-E images adapted from Vaishali et al., 2021)

Aim:

From the evidence presented above, it was clear that EB1 binds to RNA oligos at least *in vitro*, and the RNA binding surface of the protein could be mapped on to the microtubule binding surface. I therefore wanted to study if the RNA binding activity of EB1 also has a functional significance *in vivo*, and if so, to determine the RNAs with which EB1 interacts, and if and how EB1 might regulate their functions.

3.2 Results

3.2.1 EB1 binds to RNA and microtubules in a mutually exclusive manner

Based on the observation that EB1 uses the same surface for binding to microtubules and RNA, I wanted to see if the microtubules and RNA compete to bind to EB1. To test this, I performed an *in vitro* co-sedimentation assay (schematic Fig. 3.3A), wherein microtubules were polymerized using GTP γ S, a nonhydrolyzable analog of GTP, and the polymerized microtubules were incubated with EB1 in the presence or absence of RNA (U₂₅). GTP γ S was used to mimic the preferential state of GTP for EB1 binding. The extent of microtubule binding of EB1 was determined by analysing the amount of EB1 that pellets down with the microtubule fraction upon ultracentrifugation.

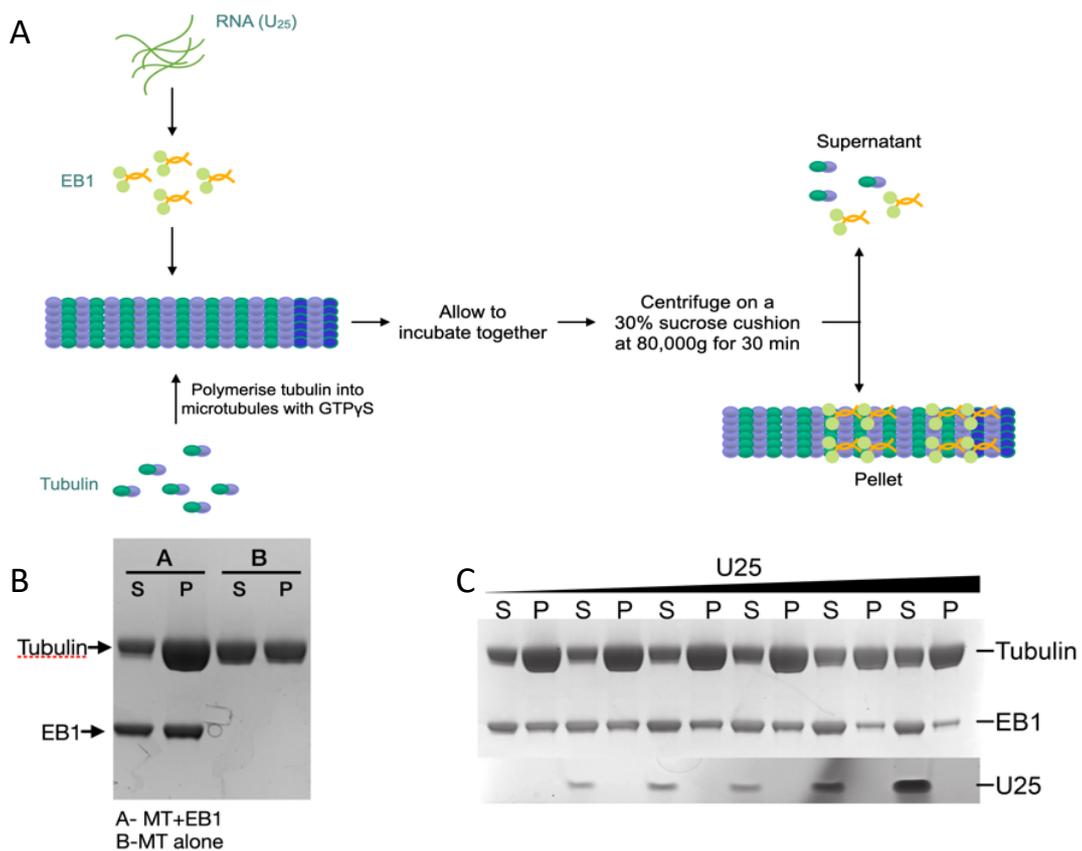


Fig. 3.3: A) Schematic of the co-sedimentation assay to study microtubule binding of proteins B) Presence of EB1 stimulates microtubule polymerisation. C) Increasing concentration of RNA (0-300 μ M) reduced the amount of EB1 pelleting with microtubules. (Adapted from (Vaishali et al. 2021)) MT: microtubules, S: supernatant, P: pellet

Performing the experiment initially without RNA, I observed that the presence of EB1 stimulates microtubule polymerization such that the extent of microtubules present in the pellet fraction as compared to supernatant is higher in the presence of EB1 than without (Fig. 3.3B). Next, I added an increasing concentration of RNA to EB1, and observed that the amount of EB1 that pelleted with the microtubules decreased with increasing the RNA concentration. This shows that EB1 binds to microtubules and RNA in a mutually exclusive manner, and that RNA can out-compete microtubule binding of EB1 to a certain extent. Furthermore, RNA was never detected in the pellet fraction indicating that RNA and microtubules exhibit no direct interaction.

3.2.2 RIP-seq to identify target RNAs of EB1 in the *Drosophila* oocyte

As the *in vitro* assays (EMSA and NMR) and PNK assay showed that EB1 exhibits RNA binding, I wanted to determine the RNA targets of EB1 *in vivo*, to get a functional insight into the RNA binding activity of EB1. For this, I performed an RNA immunoprecipitation and sequencing (RIP-seq) experiment in flies expressing EB1-GFP under the regulation of *nos* promoter. Flies expressing Flag-Myc-GFP were used as a control. The RNA-protein complexes were crosslinked *in vivo* using UV, and anti-GFP antibody was used to pull down EB1-GFP and any associated crosslinked RNAs. The enriched RNAs were then extracted and sequenced to determine the putative targets of EB1.

The analysis for identification of the differentially enriched genes in EB1-GFP vs GFP sample was performed with the help of DESeq2 (Love et al. 2014). Firstly, data quality assessment was performed via sample clustering to check for quality of the samples/replicates. Sample-to-sample distance was calculated, and a heatmap was generated to visualize the similarities (and dissimilarities) between different samples. The three EB1-GFP replicates clustered together, and so did the GFP replicates. Furthermore, the inter-sample distance between EB1-GFP and GFP samples was high, indicating their low similarity (Fig. 3.4A).

I enriched a total of 1017 genes with a adjusted p-value of <0.01 and >4 -fold enrichment when compared to the GFP sample (Fig. 3.4B) (Vaishali et al. 2021). As a control, I performed an expression level analysis of the enriched genes to check if I pulled down on highly expressed genes merely due to abundance. Using RNA expression profiling data from

modENCODE, I found that 220 genes out of the 1017 enriched ones (21.63%) were characterized as highly expressed (RPKM>51) (Gelbart and Emmert 2013).

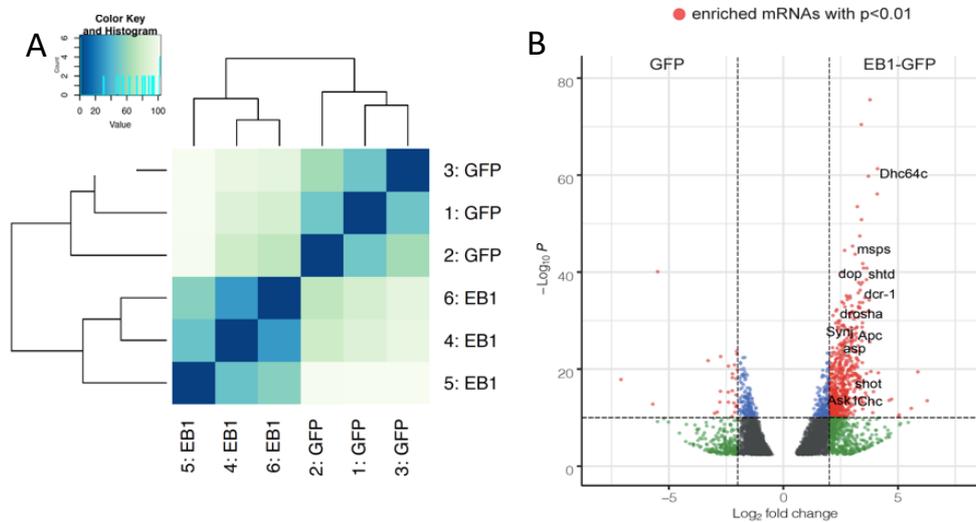


Fig. 3.4: A) Heatmap showing the similarities (and dissimilarities) between the different samples and replicates. B) Volcano plot showing genes enriched in EB1-GFP samples as compared to GFP samples. 1017 genes have a p-value < 0.01 and \log_2 fold change > 4 (adapted from (Vaishali et al. 2021)).

3.2.3 Co-localisation analysis of EB1 with the top hits

The RIP-seq analysis revealed 1017 genes significantly enriched in the EB1-GFP pulldown. To gain insights into the functional categories that were enriched for EB1 target RNAs, I decided to perform a GO term analysis of the identified genes, which revealed an enrichment of RNAs involved in transport/localisation/cytoskeleton related functions among other categories. I selected 12 of those genes (*apc*, *ask1*, *asp*, *chc*, *dcr-1*, *dhc64c*, *drosha*, *grip163*, *mmps*, *shot*, *shtd*, *synj*), and performed an *in vivo* co-localisation analysis of these target RNAs with EB1-GFP in the *Drosophila* egg chamber, using *oskar*, as a negative control as it was not enriched in the sample (Fig. 3.5A and B). When compared to the negative control, the frequency of co-localisation was generally very low for all the chosen targets, with *chc* exhibiting the highest frequency at $\sim 10\%$ (Fig.3.5A and B). In another approach, I transfected *Drosophila* S2 cells with EB1-GFP and analysed its co-localisation of EB1-GFP with three of the target RNAs that showed the highest frequency of co-localisation in the oocyte (*dhc64c*,

msps and *chc*). No co-localisation of EB1-GFP was observed with any of the three target RNAs in S2 cells (Fig. 3.5C).

While performing the co-localisation analysis for four of the target RNAs (*dhc64c*, *msps*, *asp* and *chc*) in the oocyte, I observed that they exhibit a specific localisation pattern (Fig. 3.5A and 3.6B). Since EB1 binds to microtubule plus ends and helps in regulating cytoskeletal dynamics, I wanted to assess if the protein has any role in the localisation of these putative target RNAs. To this end, I expressed EB1- RNAi in the oocyte, using a germline specific *oskar*-Gal4 driver and checked for any change in the localisation pattern of the targets. No visible change in the localisation pattern of any of the RNAs could be observed (Fig. 3.6B), indicating EB1 is not involved in the localisation of the RNAs.

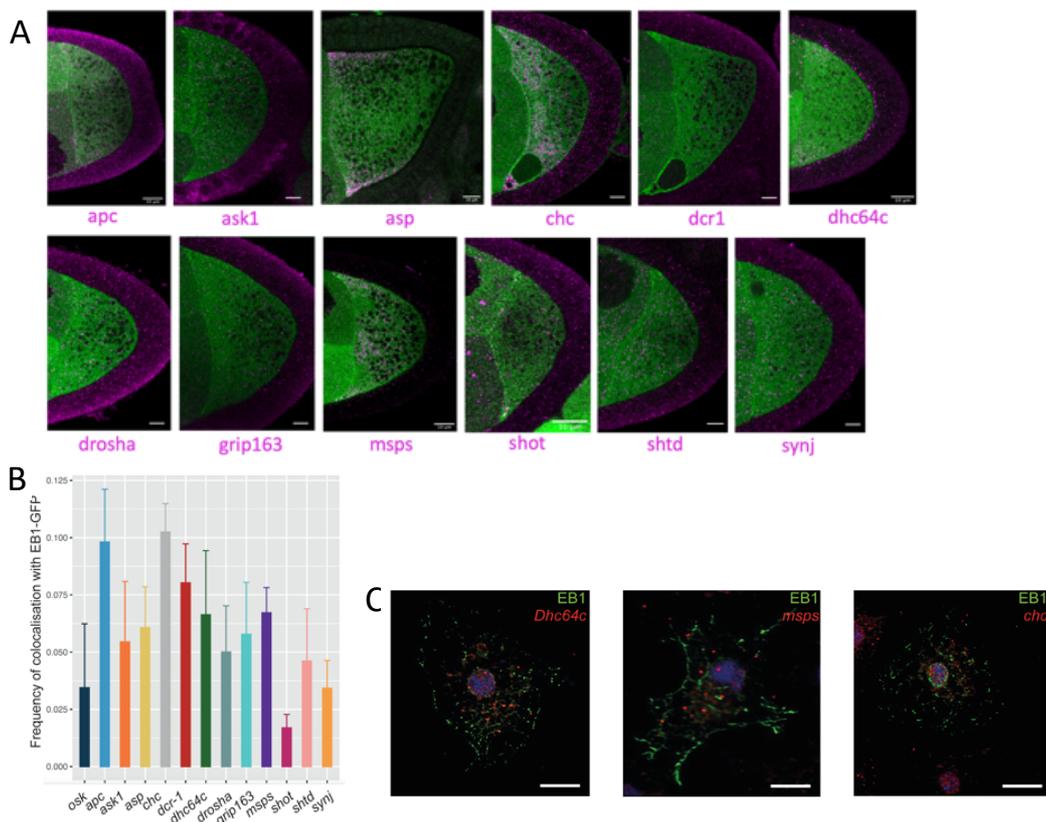


Fig. 3.5: A) Co-localisation analysis between EB1-GFP (in green) and respective target RNAs (in magenta) in the oocyte. Scale bar – 10 μ m B) Graph showing the frequency of co-localisation between EB1-GFP and targets from (A) (adapted from (Vaishali et al. 2021)). Error bars represent S.D. C) Co-localisation analysis between EB1-GFP and three of the target RNAs in Drosophila S2 cells Scale bar – 5 μ m (adapted from (Vaishali et al. 2021)).

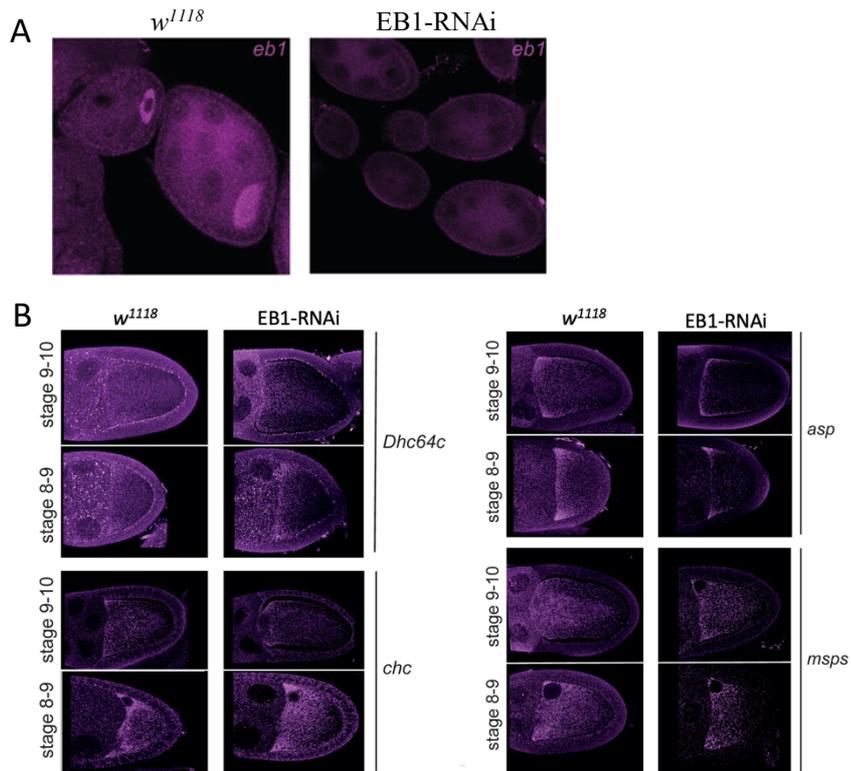


Fig. 3.6: A) smFISH for *eb1* RNA to check the knockdown efficiency in w^{1118} vs EB1-RNAi line, confirmed by loss of signal intensity in EB1-RNAi. B) Analysis of change in localisation pattern in any of the four target RNAs upon EB1 knockdown (A-B adapted from (Vaishali et al. 2021)).

3.3 Discussion

RNA interactome capture studies have been instrumental in expanding our knowledge about the extent of RNA binding proteins as well as RNA binding domains present in a cell. They offer an unbiased means of identifying novel RBPs, not restricted to specific RNAs or the presence of conventional RNA binding domains. However, just as with all high throughput studies, it is extremely important to consider the possibilities of false positives, and perform sensitive preliminary experiments to characterize the RNA binding activity of such proteins, before venturing into time and resource consuming *in vivo* experiments.

RNA interactome capture on *Drosophila* embryos identified EB1 as a novel unconventional RBP. Preliminary *in vitro* studies with EB1 revealed that it indeed has an RNA binding activity *in vitro*. Furthermore, the RNA binding surface of EB1 could also be successfully mapped to its microtubule binding surface. RIP-seq analysis of EB1 led to the identification of a large number of putative targets, the top hits of which, however, failed to exhibit significant co-localisation with EB1 when compared to a negative control. Furthermore, absence of observable phenotypes upon EB1 knockdown further complicated the analysis.

Failure in validating the binding of EB1 with some of the top hits of the RIP-seq experiment, suggests that the identification of EB1 as an RBP might be due to an opportunistic rather than specific interaction with RNA. The surface of EB1 mapped to bind to RNA is a positively charged surface, through which it also interacts with microtubules. The RNA binding activity of EB1 *in vitro*, thus, might be a result of electrostatic interaction with the negatively charged backbone of the RNA. Furthermore, there is also evidence from (Gireesh et al. 2018) that EB1 can also bind to GTP *in vitro* through the same surface as has been determined for microtubules and RNA. Not limited to EB1, several other GTPases have been identified repeatedly in such interactome capture studies (Fernandez-Chamorro et al. 2019; Liu et al. 2019), indicating that proteins with the ability to bind to nucleotides might also have an RNA binding activity, which in certain cases might have a functional significance.

Our study, however, does not entirely negate a possible role of EB1 as an RNA binding protein *in vivo*. We conventionally study the effect of RNA binding proteins on regulating

RNAs, but recent evidence suggests that RNA binding by a protein can also regulate the functions of the protein. For instance, (Horos et al. 2019) have shown that vault RNA can regulate autophagy by controlling the oligomeric state of protein p62. Furthermore, (Huppertz et al. 2020) show that RNA can regulate the process of glycolysis by regulating the enzymatic activity of Enolase 1. Since, we have not studied the effect of RNA binding on the functions of EB1, we cannot rule out the possibility that RNA binding might somehow regulate the activity of the protein. I have shown that EB1 binds to microtubules and RNA in a mutually exclusive manner. It is thus possible that RNA binding might regulate the microtubule binding activity of EB1. Assessing the regulation of EB1 by RNAs would require further study.

4. Discussion and Outlook

RNA and RNP regulation are indispensable for proper cellular functions. Recent technological advancements have extended our knowledge of the repertoire of RNA binding proteins in a cell, as well as in specific RNA granules. For my PhD project, I have tested the RNA binding ability and functions of two proteins identified in such high-throughput studies FMR1, a known RBP identified as a novel *oskar* RNP candidate, and EB1, identified as a novel, unconventional RBP.

For FMR1, I could successfully validate the protein's association with *oskar* RNP granules and its role in positively regulating Oskar protein levels *in vivo*. I could also identify the binding regions and domains of *oskar* RNA and FMR1 respectively required for the interaction. Furthermore, I also performed experiments to dissect the molecular basis for the dual functions of FMR1 in regulating translation. Additional studies though would be required to confirm the roles of the FMR1 domains *in vivo*.

In the case of EB1, although the *in vitro* data pointed towards RNA binding activity by the protein, none of the selected targets identified through RIP-seq could be validated *in vivo*. This does not exclude a possible RNA binding role of EB1, however. Considering that the RNA binding surface of EB1 is the same as the microtubule binding surface, RNA binding might regulate the functions of EB1 protein. Additional studies will need to be performed to determine if this is the case.

The high throughput methods to identify novel RNA binding proteins have definitely widened our knowledge of the extent of RBPs present in the cell. Hundreds of proteins in human cell lines are now identified as putatively RNA binding. Half of them, however, lack conventional RNA binding domains and their functions as RBPs has not yet been discovered. Nevertheless, as the importance and role of RNA metabolism in regulating cellular functions is gaining attention, it will be interesting to study how these RBPs play a role in maintaining cellular homeostasis. Older versions of the interactome capture methods had some limitations that led to the false identification of certain proteins (such as DNA binding proteins) as putative

RBPs (Perez-Perri et al. 2018). Newer improved versions of the methods have, however, greatly improved the quality of RBP identification. Studies on these novel, unconventional RBPs has shed light on new modes of RNA binding using non-canonical RBDs, new RNA related functions of known proteins such as moonlighting by enzymes, and also new modes of regulating protein activity by RNA binding called riboregulation (Hentze et al. 2018).

5. Materials and methods

5.1 Materials

5.1.1 Antibodies

<u>Name and type</u>	<u>Application and dilution</u>	<u>Source</u>
Mouse anti-FMRP, Primary	Western Blot, 1:200	DSHB 5A11
Rabbit anti-Oskar, Primary	Western Blot, 1:2500 Immunofluorescence, 1:3000	Made in-house
Rabbit anti-Histone H3, Primary	Western Blot, 1:2500	Abcam ab1791
Mouse anti-Tubulin, Primary	Western Blot, 1:2500	Sigma Aldrich T6074
Rabbit anti-GFP, Primary	Western Blot, 1:5000	Torrey Pines Biolabs TP401
Mouse anti-Me31B, Primary	Immunofluorescence, 1:200	Gift from A.Nakamura
Donkey ECL Anti - Rabbit Ig, HRP linked whole antibody, Secondary	Western Blot, 1:10000	GE Healthcare NA934
Sheep ECL Anti-Mouse IgG HRP linked whole antibody, Secondary	Western Blot, 1:10000	GE Healthcare NA931
Goat Alexa Fluor 647 anti-Rabbit, Secondary	Immunofluorescence, 1:1000	Jackson Immuno Research 712-605-153
Goat Alexa Flour 633 Anti-Mouse IgG (H+L), Secondary	Immunofluorescence, 1:1000	Invitrogen A21053

5.1.2 Fly stocks

<u>Genotype</u>	<u>Description</u>	<u>Source</u>
<i>w¹¹¹⁸</i>	<i>Wild type</i>	<i>Ephrussi Lab</i>
<i>;;nos_EB1[GFP]/Tm6,Tb</i>	<i>EB1-GFP expressed under nanos promoter, used for RIP-seq (section 2.2.2)</i>	<i>Ephrussi Lab</i>
<i>w⁻;P[UASp-FLAG::MYC::eGFP]/CyO; P(w⁺mC] = tubP-GAL4)LL7/TM3, Sb1Ser1</i>	<i>Flag-myc tagged eGFP expressed under UAS promoter driven by tubulin-Gal4, used for RIP-seq (section 2.2.2)</i>	<i>Ephrussi Lab</i>
<i>w⁻;Tg(dsRed-FMR1GFP)</i>	<i>Fosmid has FMR1 tagged with GFP at the C-terminus. dsRed is used as fosmid marker.</i>	<i>Sudhakaran et al. 2014</i>
<i>y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.GL00075}attP2</i>	<i>TRiP line on III chromosome, under UAS promoter to drive RNAi for FMR1 knockdown</i>	<i>Bloomington stock #35200</i>
<i>y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.HMS00248}attP2</i>	<i>TRiP line on III chromosome, under UAS promoter to drive RNAi for FMR1 knockdown</i>	<i>Bloomington stock #34944</i>
<i>w⁻; If/Cyo; Sb/TM3Ser</i>	<i>Double balancer line</i>	<i>Ephrussi Lab</i>
<i>w⁻; If/Cyo; TM2/TM6B,Tb</i>	<i>Double balancer line</i>	<i>Ephrussi Lab</i>
<i>w⁻; If/Cyo; oskGal4/Tm3,Ser</i>	<i>Expresses Gal4 under oskar promoter</i>	<i>Ephrussi Lab</i>
<i>w⁻; oskGal4/Cyo; Sb/Tm3,Ser</i>	<i>Expresses Gal4 under oskar promoter</i>	<i>Ephrussi Lab</i>
<i>w⁻; nosGal4-VP16/nosGal4-VP16</i>	<i>Expresses Gal4 under nanos promoter</i>	<i>Ephrussi Lab</i>
<i>oskGal4/Cyo; FMR1TRiP.GL00075/Tm3,Sb</i>	<i>Expresses FMR1 RNAi under oskGal4 driven UAS promoter</i>	<i>Stock generated for the experiment</i>
<i>w⁻; FMR1Δ50/Tm6,Tb</i>	<i>Loss of function, amorphic allele of FMR1</i>	<i>Bloomington stock #6930</i>

<i>w⁻; FMR1Δ113/Tm6,Tb</i>	<i>Loss of function, amorphic allele of FMR1</i>	<i>Bloomington stock #67403</i>
<i>y[1] w[1118]; PBac{y[+]-attP-9A}VK00018</i>	<i>Line with attP docking site on chromosome II for site specific transgenesis using phiC31 integrase. Referred to as VK18 hereafter.</i>	<i>Bloomington stock #9736</i>
<i>w⁻; UASp-GFP_FMR1/Cyo; Sb/Tm3,Ser</i>	<i>GFP_FMR1 full length expressed under UAS promoter. Integrated site-specifically on chromosome II in VK18 line.</i>	<i>Stock generated for the experiment</i>
<i>w⁻; UASp-GFP_FMR1(ΔKH)/Cyo; Sb/Tm3,Ser</i>	<i>GFP_FMR1 lacking the KH domains (residues -) expressed under UAS promoter. Integrated site-specifically on chromosome II in VK18 line.</i>	<i>Stock generated for the experiment</i>
<i>w⁻; UASp-GFP_FMR1(ΔCTD)/Cyo; Sb/Tm3,Ser</i>	<i>GFP_FMR1 lacking the C terminal domain (residues -) expressed under UAS promoter. Integrated site-specifically on chromosome II in VK18 line.</i>	<i>Stock generated for the experiment</i>

5.1.3 Primers and Probes

<u>Name</u>	<u>Sequence</u>	<u>Purpose</u>
qPCR-oskF	TATCACACAAACCTGCCACTTGA	qPCR- <i>oskar</i> forward primer
qPCR-oskR	CGTCTTTCTGTTTCCGTTTGCA	qPCR- <i>oskar</i> reverse primer
qPCR-nosF	AATCTCGGCGTGGGAATGGG	qPCR- <i>nos</i> forward primer
qPCR-nosR	AATCTCGGCGTGGGAATGGG	qPCR- <i>nos</i> reverse primer

qPCR-18SF	CGGAGAGGGAGCCTGAGAA	qPCR-18S forward primer
qPCR-18SR	AGCTGGGAGTGGGTAATTTACG	qPCR-18S reverse primer
IVT-FLoskF	GGATCACTTTCCTCCAAGCG	<i>In vitro</i> transcription, full length <i>oskar</i>
IVT-FLoskR	CCTATAACAAGCTGCAATGTAAA ATCC	<i>In vitro</i> transcription, full length <i>oskar</i>
IVT-3oskF	GTTGGGTTCTTAATCAAGATAC	<i>In vitro</i> transcription, <i>oskar</i> -3'UTR
IVT-3oskR	CCTATAACAAGCTGCAATGTAAA ATCC	<i>In vitro</i> transcription, <i>oskar</i> -3'UTR
ePAT-anc	GCGAGCTCCGCGGCCGCGTTTTT TTTTTT	ePAT-anchor primer
ePAT-uniR	GCGAGCTCCGCGGC	ePAT_universal reverse primer
ePAT-oskF	GCGCTTGTTTGTAGCACAG	ePAT_osk forward primer
ePAT-gapF	GCAGAGCAAGGACTAAACTAGC	ePAT_gapdh forward primer
L07clip2.0	/5Phos/NNNNCAGATCNNNNNAGA TCGGAAGAGCGTCGTG/3ddC/	iCLIP library preparation
L08clip2.0	/5Phos/NNNNACTTGANNNNNAGA TCGGAAGAGCGTCGTG/3ddC/	iCLIP library preparation
L09clip2.0	/5Phos/NNNNGATCAGNNNNNAGA TCGGAAGAGCGTCGTG/3ddC/	iCLIP library preparation
L10clip2.0	/5Phos/NNNNTAGCTTNNNNNAGAT CGGAAGAGCGTCGTG/3ddC/	iCLIP library preparation
L11clip2.0	/5Phos/NNNNATGAGCNNNNNAGA TCGGAAGAGCGTCGTG/3ddC/	iCLIP library preparation
L12clip2.0	/5Phos/NNNNCTTGTANNNNNAGAT CGGAAGAGCGTCGTG/3ddC/	iCLIP library preparation

L13clip2.0	/5Phos/NNNNAGTCAANNNNNAGA TCGGAAGAGCGTCGTG/3ddC/	iCLIP library preparation
L14clip2.0	/5Phos/NNNNAGTTCCNNNNNAGA TCGGAAGAGCGTCGTG/3ddC/	iCLIP library preparation
P5Solexa_s	ACACGACGCTCTTCCGATCT	iCLIP library preparation
P3Solexa_s	CTGAACCGCTCTTCCGATCT	iCLIP library preparation
P5Solexa	AATGATACGGCGACCACCGAGAT CTACACTCTTTCCTACACGACGC TCTTCCGATCT	iCLIP library preparation
P3Solexa	CAAGCAGAAGACGGCATAACGAGA TCGGTCTCGGCATTCCTGCTGAAC CGCTCTTCCGATCT	iCLIP library preparation
RT-oligo	GGATCCTGAACCGCT	iCLIP
L3-App	/rApp/AGATCGGAAGAGCGGTTCA G/ddC/	iCLIP
Primer IVT 1	GGGCGGAAAGTCCAAATTGTAAA GAGATCTCGATCCCGCGAAATTA AT	<i>In vitro</i> tethering assay
Primer IVT 2	TGCAGGTCTGACTCTAGAGGATCC CCTGGTAGGCCACCGCGG	<i>In vitro</i> tethering assay
Primer IVT 3	TCTCTACCAAACCTGGGGATT	<i>In vitro</i> tethering assay
Primer IVT 4	CTTTTACAATTTGGACTTTCC	<i>In vitro</i> tethering assay
Primer IVT 5	GGGCGGAAAGTCCAAATTGTAAA GAGATCTCGATCCCGCGAAATTA AT	<i>In vitro</i> tethering assay
Primer IVT 6	AATCCCCAGTTTGGTAGAGACCT GGTAGGCCAACCGCGGG	<i>In vitro</i> tethering assay
FMR1gateway_F	CACCACTACGTCTGGCGATATGG AAG	Cloning FMR1
FMR1gateway_R	GGACGTGCCATTGACCAG	Cloning FMR1
pET11-FMR1	AGACGGATCCATGGAAGATCTCC TCGTGGA (F) TCTGAGCTCTTAGGACGTGCCAT TGACCAGGCC (R)	Cloning full length FMR1 for protein purification in pET11 vector with HisSUMO tag.

pET11-ΔNFMR1	AGACGGATCCATGGGAAACTACG TTGAGGAG (F) TCTGAGCTCTTAGGACGTGCCATT GACCAGGCC (R)	Cloning FMR1 lacking the N-terminus (residues 220-681) for protein purification in pET11 vector with HisSUMO tag.
pET11-FMR1(KH1_2)	AGACGGATCCATGGGAAACTACG TTGAGGAG (F) TCTGAGCTCTTACTGATCAATCTC CATCTTCT (R)	Cloning FMR1-KH1KH2 domains (residues 220-378) for protein purification in pET11 vector with HisSUMO tag.
λN-FMR1	TTCTGTTCCAGGGGCCAGTATGG AAGATCTCCTCGTGGA (1) TCCGGTACCTCATTAAGTTTAAGT TTAGGACGTGCCATTGACCA (2)	Cloning FMR1 (residues 1-681) for protein purification in pMJ vector with λN tag, using infusion cloning.
λN-ΔNFMR1	GTTCTGTTCCAGGGGCCAGTATG G GAAACTACGTTGAGG (1) TCCGGTACCTCATTAAGTTTAAGT TT AGGACGTGCCATTGACCA (2)	Cloning FMR1 lacking the N-terminus (residues 220-681) for protein purification in pMJ vector with λN tag, using infusion cloning.
λN-FMR1(KH1_2)	GTTCTGTTCCAGGGGCCAGTATG GGAAACTACGTTGAGG (1) GTACCTCATTAAGTTTAAGTCTGA TCAATCTCCATCTTCT (2)	Cloning FMR1-KH1KH2 domains (residues 220-378) for protein purification in pMJ vector with λN tag, using infusion cloning.
λN-FMR1-CTD	CAACAGACCGGTGGATCCATGCA GCTTCGCGCCATCCAGGAA (1) TTCCTGGATGGCGCGAAGCTGCA TGGATCCACCGGTCTGTTG (2)	Cloning FMR1-C-terminal domain (residues 379-681) for protein purification in pMJ vector with λN

		tag, using infusion cloning.
--	--	------------------------------

5.1.4 Buffers and reaction mixes

<u>Buffer</u>	<u>Composition</u>	<u>Application</u>
PBT(1x)	PBS + 0.1% Triton X-100	smFISH
HYBEC	2x saline-sodium citrate (SSC), 15% ethylene carbonate, 1mM EDTA, 50µg/mL heparin, 100µg/mL salmon sperm DNA, 1% Triton X-100	smFISH
1x BRB80	80mM Pipes, 1mM MgCl ₂ , 1mM EGTA, pH-6.8 with KOH	Co-sedimentation assay
Lysis Buffer A	20mM Hepes (pH 7.5), 100mM KCl, 1mM MgCl ₂ , freshly added 80U/mL Ribolock, 0.05% NP-40 and 1x cComplete protease inhibitor cocktail.	RIP-seq
HS Buffer A	20mM Hepes (pH 7.5), 1 M NaCl, 1mM EDTA, 0.5% NP-40 freshly added 80U/mL Ribolock, 0.5mM DTT and 1x cComplete protease inhibitor cocktail.	RIP-seq
MS Buffer A	20mM Hepes (pH 7.5), 500mM NaCl, 1mM EDTA, 0.5% NP-40 freshly added 80U/mL Ribolock, 0.5mM DTT and 1x cComplete protease inhibitor cocktail.	RIP-seq
LS Buffer A	20mM Hepes (pH 7.5), 150 mM NaCl, 1mM EDTA, 0.5% NP-40 freshly added 80U/mL Ribolock, 0.5mM DTT and 1x cComplete protease inhibitor cocktail.	RIP-seq
Proteinsase K Buffer A	20mM Hepes (pH 7.5), 150 mM NaCl, 1% SDS	RIP-seq
UV CXL Buffer	1x PNK buffer A (thermo), 10µM DTT, 300 µCi γ- ³² P-ATP, 30 units PNK enzyme (thermo)	<i>In vitro</i> UV crosslinking assay
Binding buffer CXL	10mM Tris (pH 7.4), 100mM KCl, 2.5mM MgCl ₂	<i>In vitro</i> UV crosslinking assay

Lysis Buffer BA	50 mM Tris-HCl, pH 7.4, 100 mM NaCl, 1% Igepal CA-630, 0.1% SDS, 0.5% Sodium deoxycholate	<i>In vitro</i> binding assay, iCLIP
Binding Buffer BA	100mM Hepes (pH 7.2), 30mM MgCl ₂ , 30% glycerol, 10mM DTT, 100mM KCl	<i>In vitro</i> binding assay, iCLIP
High-salt wash Buffer BA	50 mM Tris-HCl, pH 7.4, 1 M NaCl, 1 mM EDTA pH 8.0, 1% Igepal CA-630, 0.1% SDS, 0.5% Sodium deoxycholate, freshly added 1mM DTT and 0.025mg/mL heparin	<i>In vitro</i> binding assay, iCLIP
PNK wash Buffer BA	20 mM Tris-HCl, pH 7.4, 10 mM MgCl ₂ , 0.2% Tween-20, freshly added 1mM DTT	<i>In vitro</i> binding assay, iCLIP
5X PNK Buffer, pH 6.5	350 mM Tris-HCl, pH 6.5, 50 mM MgCl ₂ , 5 mM Dithiothreitol	<i>In vitro</i> binding assay, iCLIP
PNK reaction mix	15 μ L water, 4 μ L 5X PNK buffer (pH 6.5), 0.5 μ L Ribolock, 0.5 μ L T4 PNK enzyme	<i>In vitro</i> binding assay, iCLIP
Adaptor ligation reaction mix	8 μ L water, 5 μ L 4x ligation buffer, 0.5 μ L Ribolock, 1.5 μ L Pre-adenylated L3-App (20 μ M), 4 μ L PEG400, 1 μ L T4 RNA Ligase	iCLIP
4X Ligation Buffer	200 mM Tris-HCl, pH 7.8, 40 mM MgCl ₂ , 4 mM Dithiothreitol	iCLIP
Hot PNK mix	6 μ L water, 0.8 μ L 10X PNK buffer, 0.8 μ L ³² P- γ ATP, 0.4 μ L T4 PNK (NEB)	iCLIP
Proteinase K (PK) Buffer B	100 mM Tris-HCl, pH 7.4, 50 mM NaCl, 10 mM EDTA	iCLIP
PK + urea buffer	100 mM Tris-HCl, pH 7.4, 50 mM NaCl, 10 mM EDTA, 7 M Urea	iCLIP
RT-CLP mix	7 μ L water, 4 μ L 5X RT buffer, 1 μ L 0.1M DTT, 0.5 μ L Ribolock, 0.5 μ L Superscript III	iCLIP
Lig-CLP mix	2 μ L 10X NEB RNA ligase buffer, 0.2 μ L 0.1M ATP, 9 μ L 50% PEG8000, 0.3 μ L water, 0.5 μ L high conc. RNA ligase	iCLIP
ePAT master mix (12 μ L)	4 μ L water, 4 μ L 5x Superscript III buffer, 1 μ L 100mM DTT, 1 μ L 10mM	ePAT

	NTPs, 1 μ L RNaseOUT, 1 μ L (5U) Klenow polymerase	
Lysis Buffer CLP	10 mM Tris-Cl pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 0.5 % NP-40, 0.5 % SDS, freshly added fresh Protease Inhibitor 1:100 & Ribolock 1:2000	CLIP
Dilution Buffer CLP	10 mM Tris-Cl pH 7.5, 150 mM NaCl, 0.5 mM EDTA, freshly added fresh Protease Inhibitor 1:100 & Ribolock 1:2000	CLIP
Wash Buffer CLP	10 mM Tris-Cl pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 0.5 % NP-40, 0.5 % SDS, 0.02 mg/ml Heparin, freshly added Protease Inhibitor 1:100 & Ribolock 1:2000	CLIP
High Salt Wash Buffer CLP	10 mM Tris-Cl pH 7.5, 750 mM NaCl, 0.5 mM EDTA, 0.5 % NP-40, 0.5 % SDS, 0.02 mg/ml Heparin, freshly added Protease Inhibitor 1:100 & Ribolock 1:2000	CLIP
IVT-A reaction mix	1X transcription buffer, 30 mM DTT, 1 mM rNTP mix (ACU), 50U Ribolock, 75U T3 RNA polymerase, 7mM 3'-O-Me-7mG(ppp)G	<i>In vitro</i> tethering assay
1% denaturing agarose gel	0.4g agarose, 1x MOPS, 6.5% formaldehyde	<i>In vitro</i> tethering assay
Master Mix (MM)-T	0.25g/l tRNA, 0.05M potassium acetate, 0.016M Hepes-KOH, pH 7.5, 0.08g/l Creatine kinase, 2nM reporter RNAs	<i>In vitro</i> tethering assay
ARS-T	0.1 mM amino acid mix, 0.02 M creatine phosphate, 0.8 mM AP	<i>In vitro</i> tethering assay
Lysis Buffer PP	500 mM NaCl, 20 mM Tris-Cl (pH 8.5), 5% Glycerol, 40 mM imidazole, freshly added 5mM beta-mercaptoethanol, 1X Roche protease inhibitor cocktail, 0.01% NP-40	Protein Purification
Elution Buffer Ni PP	500 mM NaCl, 20 mM Tris Cl (pH 7.5), 5% Glycerol, 600 mM Imidazole	Protein Purification

Gel Filtration Buffer PP	150 mM NaCl, 20 mM Tris Cl (pH 7.5), 2 mM MgCl ₂ , 5% Glycerol	Protein Purification
Salt Solution PC	0.4% NaCl, 0.3% Triton-X100	Pole cell analysis
Blocking Buffer PCA	PBS, 0.3% Triton X100, 0.5% BSA	Pole cell analysis
Blocking Buffer PCB	PBS, 0.1% Triton X100, 10% Normal Goat Serum	Pole cell analysis

5.1.5 Chemicals and reagents

<u>Chemical</u>	<u>Source</u>	<u>Application</u>
Terminal Deoxynucleotidyl Transferase (20U/ μ L)	Thermo Scientific, EP0161	Probe labelling for smFISH
Ethylene Carbonate, 98%	Sigma Aldrich, E26258	smFISH
20x Saline-sodium citrate	Ambion, AM9763	smFISH
Heparin	Sigma Aldrich, H3393	smFISH
Salmon sperm DNA	Invitrogen, 15632-011	smFISH
Triton X-100	Sigma, T2984	smFISH
T4 Polynucleotide Kinase (PNK) (10 U/ μ L)	Thermo Scientific	<i>In vitro</i> UV crosslinking assay
Igepal CA-630	Merck, I8896-50ML	<i>In vitro</i> binding assay, iCLIP
Tween-20	Merck, P9416-50ML	<i>In vitro</i> binding assay, iCLIP
Urea	Merck, U5378-1KG	iCLIP
Hepes	Merck, H0887-100ML	iCLIP
NaOH	Merck, S8045-500G	iCLIP
SuperScript III Reverse Transcriptase	Life Technologies, 18080085	iCLIP
10mM dNTP mix	NEB, N0447L	iCLIP

MyONE silane beads	Life technologies, 37002D	iCLIP
RLT buffer	Qiagen, 79216	iCLIP
High conc RNA ligase	NEB, M0437M	iCLIP
10X NEB RNA ligase buffer	NEB, B0216S	iCLIP
4X LDS sample buffer	Invitrogen, NP0007	Western Blots
NuPAGE MOPS SDS running buffer 20X	Invitrogen, NP0001	Western blot
NuPAGE Transfer Buffer 20X	Invitrogen, NP0006	Western blot
Precision Plus Protein Dual Color Ladder	BioRad, #1610374	Western Blots
Instant Blue Coomassie Stain	CBS Scientific	Protein gel staining
1,4-Dithiothreitol (DTT)	Roth, 6908	Multiple applications
Immobilien Western HRP substrate peroxide solution (ECL)	Milipore, WBKLSO500	Western blot
SYBR Green qPCR Master Mix	Applied Biosystems, 4309155	qRT-PCR
Trizol LS	Invitrogen, 10296028	qRT-PCR, CLIP
³² P-γ-ATP	Hartmann Analytics, SRP-301	
T4 Polynucleotide Kinase (PNK)	NEB, M0201S	<i>In vitro</i> binding assay, iCLIP
10X PNK Buffer	NEB, B0201S	<i>In vitro</i> binding assay, iCLIP
3'-O-Me-7mG(ppp)G RNA Cap Structure Analog (ARCA)	Jena Bioscience, NU-855S	<i>In vitro</i> tethering assay
DNase 1	NEB, M0303S	<i>In vitro</i> tethering assay
Phenol:chloroform:isoamyl alcohol, pH 4.8 (25:24:1)	Roth, X985.3	<i>In vitro</i> tethering assay
phenol:chloroform:isoamyl alcohol, pH 7.8/8.0 (25:24:1).	ITW, A0889,0100	<i>In vitro</i> tethering assay

Amino acid mix	Promega, L4461	<i>In vitro</i> tethering assay
Creatine kinase	Roche, 10127566001	<i>In vitro</i> tethering assay
Luciferase Assay System	Promega, E1500	<i>In vitro</i> tethering assay
Vectashield	Vector Laboratories, H-1000-10	Mounting media
Bleach (sodium hypochlorite 6-14%)	Supelco, 1056142500	Cuticle preparation, Hatching rate analysis
Porcine brain tubulin	Cytoskeleton, T-240	Co-sedimentation assay
GTP γ S	Roche, 10220647001	Co-sedimentation assay
Taxol	Cytoskeleton, T240	Co-sedimentation assay
4-12% Bis-Tris, precast SDS-PAGE mini gels	Invitrogen, NP0322BOX	Co-sedimentation assay, Western Blots, iCLIP, <i>In vitro</i> UV crosslinking assay
RiboLock	Thermo Scientific, EO0381	RIP-seq, iCLIP
cOmplete™ Protease Inhibitor Cocktail	Roche, 11697498001	RIP-seq
GFP-Trap® Magnetic Agarose (MA) beads	Chromotek, gtma-20	RIP-seq, iCLIP, <i>In vitro</i> binding assay
TRIzol™ LS Reagent	Invitrogen, 10296010	RIP-seq, RNA extraction from ovaries
Proteinase K solution (20mg/ml), RNA grade	Invitrogen, 25530049	RIP-seq, iCLIP

5.1.6 Kits

<u>Kit</u>	<u>Source</u>	<u>Application</u>
SENSE mRNA-Seq Library Prep Kit V2	Lexogen	RIP-seq
MEGAscript™ T7 Transcription Kit	Invitrogen, AMB13345	<i>In vitro</i> transcription

MEGAscript™ T7 Transcription Kit	Invitrogen, AMB1354	<i>In vitro</i> transcription
SuperScript™ III First-Strand Synthesis SuperMix for qRT-PCR	Invitrogen, 11752050	qRT-PCR
ProNex Chemistry	Promega, NG2001	iCLIP

5.1.7 Equipment

<u>Equipment</u>	<u>Source</u>	<u>Application</u>
G-25 columns	GE healthcare	<i>In vitro</i> UV crosslinking assay
Phase Lock Gel Heavy tubes	VWR International GmbH, 733-2478	iCLIP
TapeStation 2200	Agilent Technologies	iCLIP
Qubit 2.0 Fluorometer	Life Technologies	iCLIP
His-Trap HP columns	Merck Cytiva, 29-0510-21	Protein purification
HiLoad 16/600 Superdex 200 pg	Merck, Cytiva 28-9893-35	Protein purification
Typhoon FLA 9500 Biomolecular imager	GE healthcare, FLA 9500	<i>In vitro</i> binding assay, iCLIP, <i>In vitro</i> crosslinking assay
Step One Real Time PCR system	Applied Biosystems	CLIP, qRT-PCR
Mithras LB 940 plate reader	Berthold technologies	<i>In vitro</i> tethering assay
Kitchen Aid mixer	Kitchen Aid	CLIP, RIP-seq

5.1.8 *In vitro* iCLIP Barcodes

i7	L07clip2.0	GATCTG	G A T C T G
i8	L08clip2.0	TCAAGT	T C A A G T
i9	L09clip2.0	CTGATC	C T G A T C
i10	L10clip2.0	AAGCTA	A A G C T A
i11	L11clip2.0	GCTCAT	G C T C A T

i12	L12clip2.0	TACAAG	T A C A A G
i13	L13clip2.0	TTGACT	T T G A C T
i14	L14clip2.0	GGAACT	G G A A C T

5.2 Methods

Note: Recipes for all the buffers used are mentioned in section 5.1.4.

Probe labelling for smFISH

Probe labelling for single molecule FISH was performed as in (Gaspar et al. 2017). Probes for the following RNAs were labelled using ddUTP-Atto633: *oskar*, *eb1*, *apc*, *ask1*, *asp*, *chc*, *dcr-1*, *dhc64c*, *droscha*, *grip163*, *mmps*, *shot*, *shtd*, *synj*. For P-body co-localisation experiment, *oskar* probes were labeled with ddUTP-Atto565. 15µL of reaction mix was prepared containing 1000pmol of oligo mix, 24 units of terminal deoxynucleotidyl transferase (TdT), 3x molar excess than oligo mix of ddUTP-atto633 (or ddUTP-atto565) and 1x TdT buffer. The reaction mix was incubated at 37°C overnight. The next day, the probes were precipitated as follows: Sodium acetate (final 300mM) was added to the reaction mix to make up the volume to 100µL. 0.5µL of 5mg/ml linear acrylamide was added along with cold 100% ethanol to make up the volume to 500µL. The probes were incubated at -80°C for 1 hour, and centrifuged at 13,200 rpm at 4°C for 30 min. The supernatant was discarded and the pellet washed 2x with 1mL 80% ethanol, vortexed briefly and centrifuged at 13,200 rpm at 4°C for 10 min. The pellet was air dried at room temperature (RT) and resuspended in 30µL DNase/RNase free water. The OD₂₆₀ and absorbance (A₆₃₄ or A₅₇₀) of the probes were checked using a NanoDrop spectrophotometer and the concentration of the probes and degree of labelling were determined as described in (Gaspar et al. 2017).

Single molecule fluorescence in-situ hybridization (smFISH)

Ovaries were dissected in Phosphate Buffered Saline (PBS) solution and fixed in 2% Paraformaldehyde (in PBS + 0.1%Triton X100) for 20 minutes at room temperature on a nutating mixer. They were then washed with 750 µL PBT twice, 10 min each. 100µL of

HYBEC buffer was then added to the samples and incubated at 42°C for 15 min for pre-hybridization. 100µL of HYBEC buffer with 4nM per probe concentration was then added to the 100µL HYBEC containing sample, and incubated for 2 hours at 42°C. The samples were then washed as follows: 10 min with HYBEC buffer at 42°C, 10 min with HYBEC:PBT (1:1) at 42°C, 10 min with PBT at 42°C and 10 min with PBT at room temperature. 100µL of 80% 2,2'-thiodiethanol (TDE) in PBS was then added as mounting medium and left overnight at 4°C. Finally the samples were mounted and viewed using Leica SP8 confocal microscope. Object based co-localisation analysis was performed using the published R plugin – xsColoc as in (Gáspár et al. 2017).

Crosslinking and immunoprecipitation (CLIP) for ovaries

The following experiment was performed by Frank Wippich.

Ovaries were harvested in PBS using a Kitchen Aid, and resuspended in 600µL Lysis Buffer CLP. The sample was split into two: one was crosslinked with 0.3J/cm² UV in a Stratalinker, and the other was not subjected to crosslinking. Both the samples were homogenized using a pestle and centrifuged at 800 rpm for 5 min at 4°C. The supernatant was diluted 1:1 with dilution buffer CLP (300µL + 300µL). GFP-Trap MA beads were then added (5µL), and incubated for 2-3 hours on rotator at 4°C. Beads were washed twice with wash buffer CLP, twice with high salt wash buffer CLP and twice with wash buffer CLP again, 10 min each at 4°C on rotator. The beads were resuspended in 100µL Proteinase K Buffer A and incubated with 0.2mg/mL Proteinase K at 55°C for 45 min. The RNA was then extracted using Trizol LS reagent, following the manufacturer's instructions. First strand synthesis was performed using SuperScript™ III First-Strand Synthesis SuperMix for qRT-PCR using the manufacturer's instructions, and the cDNA thus prepared was used for qRT-PCR using SYBR green qPCR mix in Step One Real Time PCR system from Applied Biosystems.

Recombinant protein purification from *E. coli*

Cloning of expression vectors: RNA was extracted from ovaries and cDNA synthesis was performed using SuperScript™ III First-Strand Synthesis SuperMix for qRT-PCR following the manufacturer's instructions. FMR1 cDNA was amplified using the forward primer FMR1gateway_F and reverse primer FMR1gateway_R and cloned into pAW vector (pA-

FMR1) using the Gateway Cloning System (<https://emb.carnegiescience.edu/drosophila-gateway-vector-collection>). Cloning full length pET11-HisSUMO-FMR1: Forward and reverse primers pET11-FMR1 (see table 5.1.3) were used to amplify FMR1 from pA-FMR1. The amplicon and pET11 vector (from Protein Expression and Purification Core Facility, EMBL) were both digested with BamHI and SacI at 37°C for 30 min. The digested products were gel purified and ligated using T4 DNA ligase at 16°C overnight. The ligation mix was plated onto LB plus kanamycin plates and positives were screened by sequencing.

Cloning pET11-HisSUMO-ΔN-FMR1: FMR1 was amplified from pA-FMR1 using primers pET11-ΔNFMR1 (see table 5.1.3), and the amplicon and pET11 vector were both digested with BamHI and SacI at 37°C for 30 min. The digested products were gel purified and ligated using T4 DNA ligase at 16°C overnight. The ligation mix was plated onto LB plus kanamycin plates and positives were screened by sequencing.

Cloning pET11-HisSUMO-FMR1(KH1_2): FMR1 was amplified from pA-FMR1 using primers pET11-FMR1(KH1_2) (see table 5.1.3), and the amplicon and pET11 vector were both digested with BamHI and SacI at 37°C for 30 min. The digested products were gel purified and ligated using T4 DNA ligase at 16°C overnight. The ligation mix was plated onto LB plus kanamycin plates and positives were screened by sequencing.

Cloning λN-sfGFP tagged FMR1 constructs: Plasmid ‘pMJ-His-TEV-lambdaN-sfGFP-3C-ScaI-STOP’ was a gift from Mandy Jeske (Heidelberg University). The vector was linearised by digesting with ScaI. Different domains of FMR1 protein were amplified using the respective primer sets in table 5.1.3. The amplicons were integrated into the linearised vector using Takara In-Fusion HD cloning kit, following the manufacturer’s instructions.

Protein purification: Electrocompetent *E.coli* bacteria were transformed with expression vector (encoding the protein of interest under the control of lac promoter), and cultured overnight at 37°C in LB media with appropriate antibiotics (30μg/μL kanamycin plus 33 μg/μL chloramphenicol in this case). A fresh 1 L culture was then started with inoculum from the overnight culture and incubated at 37°C. When the culture reached the O.D. of 0.5, 0.2 mM IPTG was added to induce the expression of the protein. The culture was then incubated at 18°C overnight. The next day, cells were harvested by centrifuging the culture at 4500 rpm for 20 min at 4°C. The pellet was resuspended in 20mL water and centrifuged at 4000 rpm for 15 min at 4°C. The supernatant was removed. (The pellet can be snap-frozen in liquid nitrogen and stored at -80°C till used.) 20mL of lysis buffer PP was added to the bacterial pellet and

the cells were resuspended. A microfluidizer was then used to lyse the cells. The sample was centrifuged at 18000 rpm in a Beckman SS34 rotor for 20 min at 4°C. The supernatant was transferred to a new tube. For His-tagged proteins, the lysate was run through a His-Trap HP column, and the bound fraction containing the protein eluted in 2mL volumes, using a 0-100% gradient of imidazole (600mM) over a volume of 40mL using elution buffer Ni PP. The samples were then incubated with protease to cleave off the tag (if required) at 4°C overnight. The protein sample was then concentrated to 5 mL volume using Amicon Ultra centrifugal filters with the appropriate molecular weight cutoff, depending on the protein. The sample was then injected onto a gel filtration column HiLoad 16/600 Superdex 200 pg, using gel filtration buffer PP and 1mL fractions were collected. 10µL from each fraction was analysed for the presence and purity of the protein.

***In vitro* Binding Assay**

In vitro transcribed *oskar* 3'UTR RNA was heated to 70°C for 5 min and immediately put on ice. 500nM purified recombinant protein, KH(1_2)CTD-GFP or KH(1_2)_GFP, was incubated with 20nM RNA in binding buffer BA at 37°C for 10 min at 1100 rpm. The sample was then put on a sterile 10 x 10 cm plate and crosslinked with UV at 6mJ/cm² in a Stratalinker. The volume was brought to 900µL with lysis buffer BA, and 1:650 diluted RNase I and 1µL Turbo DNase were added to the sample. It was incubated at 37°C for another 3 min, followed by addition on 15 µL GFP-Trap MA beads and incubation for 1 hour at 4°C on rotator. The beads were then washed thrice with high-salt wash buffer BA, and thrice with PNK wash buffer, 10 min each. To radioactively label the bound RNAs, the beads were resuspended in 16µL of Hot PNK mix and incubated at 37 °C for 5 min at 1100 rpm. The beads were washed once with PNK wash buffer BA, followed by addition of 20 µL 1X LDS sample buffer, and incubated at 70°C for 5 min. The supernatant was then run on 4-12% Bis-Tris precast gel in MOPS running buffer. The gel was then exposed to a Fuji film phosphor screen and the screen was visualised using Typhoon FLA 9500 biomolecular imager from GE healthcare.

***In vivo* iCLIP**

The following experiment was performed by Matteo Bordi

Ovary extract preparation from FMR1 and Flag-Myc-GFP fly lines: Newly hatched flies were transferred to a fresh vial with yeast and fed for 2 days at room temperature. The flies were collected and ground using the Kitchen Aid in ice cold PBS. The sample was sieved using 400nm and 200nm sieves to remove other body parts such as legs, and the ovaries were collected using the 80nm sieve. The ovaries were transferred to a 15ml tube and centrifuged at 600rcf for 30 sec at 4°C. The supernatant was removed and 12 mL PBS was added to the ovaries and centrifuged again at 600rcf for 30'' at 4°C. The supernatant was removed and the ovaries were resuspended in 3ml ice cold PBS (Note: approx. 1.5ml ovaries / sample, from 75 ml dry volume of flies were used). The ovaries were transferred to a 10cm cell culture dish and placed in a container filled with ice.

The ovaries were crosslinked with UV using the Stratalinker. The ovaries were collected in a 15ml tube, washed with PBS and centrifuged at 600 rcf for 30 sec at 4°C. The supernatant was removed and the ovaries were resuspended in 3ml PBS and transferred to two 1.5ml protein low binding tubes. The ovaries were centrifuged at 600rcf for 30 sec at 4°C. The supernatant was removed and 200-300µl (depending on the amount of material collected) of iCLIP lysis buffer supplemented with DTT (1mM) and Ribolock (1:1000 v/v) was added and the ovaries were lysed using a pestle. The lystae was centrifuged at 16000rcf for 10 min at 4°C. The supernatant was recovered and the centrifugation step was repeated. The supernatant was collected, and snap-frozen for storage at -80°C.

Note: This should yield approx. 40 mg total protein samples (enough for one replicate).

Bead preparation: For each sample use 10µl of beads: for FMG CL α -GFP and FMR1 CL α -GFP use Magnetic agarose GFP Trap beads (Chromotek); for FMR1 CL beads-only use Magnetic agarose beads (Chromotek). The beads were washed twice with 1ml iCLIP high salt buffer supplemented with DTT (1mM), Ribolock (1:2000 v/v) and Heparin (0.02mg/ml) and resuspended in 1ml iCLIP lysis buffer supplemented with DTT (1mM) and Ribolock (1:2000 v/v).

RNase I mediated partial RNA digestion: For each sample in the experiment, 1030µl aliquots with 20mg of total protein content was used. iCLIP lysis buffer supplemented with DTT (1mM) and Ribolock (1:1000 v/v) was used for diluting the extracts. 5µl Turbo DNase was added along with 10µl RNase I (Invitrogen) diluted 1:20 and incubated at 37°C for 3 min at 1100rpm. The samples were incubated on ice for more than 3 min. 10 uL of the prepared beads were then added to the each sample.

Immunoprecipitation: The samples were Incubated for 2h in cold room on rotating wheel (11rpm). The samples were then washed thrice with 1mL iCLIP high salt buffer, transferred to new tubes, and then washed thrice with iCLIP PNK buffer.

The RNA 3'end dephosphorylation and the subsequent steps till sequencing were performed as mentioned in the "*in vitro* iCLIP" method.

***In vitro* iCLIP**

In vitro iCLIP was performed in two different conditions: 20nM RNA with 100nM KH(1_2)CTD-GFP protein and 20nM RNA with 250nM KH(1_2)CTD-GFP protein. Four replicates were prepared for each condition.

Sample preparation and immunoprecipitation: *In vitro* transcribed *oskar* RNA was heated to 70°C for 5 min and immediately put on ice. 250nM and 100nM purified recombinant protein KH(1_2)CTD-GFP was incubated with 20nM RNA in binding buffer BA at 37°C for 10 min at 1100 rpm. The sample was then put on sterile 10 x 10 cm plate and crosslinked with UV at 6mJ/cm² in a stratalinker. The volume was made up to 900µL with lysis buffer BA, and 1:650 diluted RNase I and 1µL Turbo DNase were added to the sample. It was incubated at 37°C for another 3 min, followed by addition on 15 µL GFP-Trap MA beads and incubation for 1 hour at 4°C on rotator (15rpm). The beads were then washed thrice with high-salt wash buffer BA, and thrice with PNK wash buffer BA, 10 min each on rotator (20rpm).

RNA 3'end dephosphorylation: The beads were then resuspended in 20µL of PNK reaction mix and incubated at 37°C for 20 min at 1100 rpm. They were then washed once with PNK wash buffer, twice with high salt wash buffer BA, and twice with PNK wash buffer BA, 10 min each on rotator (20rpm).

First adaptor ligation to the 3'end of RNA: The beads were then resuspended in 20µL adaptor ligation reaction mix and incubated overnight at 16°C at 1100 rpm. 500µL PNK wash buffer was then added and the beads washed twice with high salt wash buffer BA, and twice with PNK wash buffer BA (with tubes changed after first PNK wash buffer wash) for 10 min each, on rotator (20rpm).

Radioactive labelling of RNA 5'end: To radioactively label the bound RNAs, the beads were resuspended in 16µL of Hot PNK mix and incubated at 37 °C for 5 min at 1100 rpm. The beads were washed once with PNK wash buffer BA, followed by addition of 20 µL 1X LDS

sample buffer, and incubated at 70°C for 5 min. The supernatant was then run on a 4-12% Bis-Tris precast gel in MOPS running buffer. The gel was then exposed to Fuji film phosphor screen and the screen was visualised using a Typhoon FLA 9500 biomolecular imager from GE healthcare.

RNA extraction: Using the autoradiograph as mask, the membrane was cut and put into tubes to extract the RNA. 10µL of proteinase K (20mg/mL) in 200µL Proteinase (PK) buffer was added to the membrane, and incubated at 37°C for 20 min at 1100 rpm. 200µL of PK + urea buffer was then added and incubated additional 20 min at 37°C at 1100 rpm. The solution was put in a Phase Lock Gel Heavy tube along with 400 µL of phenol/chloroform (pH 7.8/8.0). The sample was incubated at 30°C for 5 min at 1100 rpm, and the phases were separated by centrifuging at 16,000 xg for 5 min at RT. The aqueous layer was transferred to a new tube and the RNA was precipitated by the addition of 0.75µL GlycoBlue (15mg/mL) and 40 µL 3M sodium acetate (pH 5.5), followed by 1 mL 100% ethanol. The solution was mixed by inverting the tube and incubated at -20°C overnight. The next day, the sample was centrifuged at 21,100 xg for 20 min at 4°C, washed with 80% ethanol, centrifuged for 5 min and the pellet was air dried for 3 min at the bench before resuspending it in 5 µL ultrapure water.

The following steps for library preparation were performed at IMB, Mainz by our collaborator Anna Orekhova from the König Group.

Reverse Transcription: 1µL of RT-oligo (0.5 pmol/µL) and 1µL of dNTP mix (10mM) were added to the resuspended pellet and incubated at 70°C for 5 min. This was followed by the addition of 13µL RT-CLP mix. The sample was run on the following RT program:

Temperature	Time (min)
25°C	5
42°C	20
50°C	40
80°C	5
4°C	hold

1.65 µL of 1M NaOH was then added and incubated at 98°C for 20 min, followed by addition of 20µL 1M Hepes-NaOH (pH 7.3).

MyONE Silane cleanup: 10µL MyONE silane beads per sample were washed and resuspended in 93µL RLT buffer and added to the sample. 112 µL 100% ethanol was then

added and mixed by pipetting, and incubated for 5 min at RT. The beads were again mixed and incubated for another 5 min at RT. The beads were magnetically separated, and supernatant was discarded. The beads were resuspended in 1mL of 80% ethanol and transferred to a new tube. The step was repeated twice and the beads were finally air dried for 5 min at RT and resuspended in 5 μ L ultrapure water, and incubated for 5 min at RT.

Second adaptor ligation: To the cDNA-bead solution, 2 μ L of L##clip2.0 (second adaptor) and 1 μ L 100% DMSO were added and incubated at 75°C for 2 min and immediately put on ice. Since there was a total of 8 samples (two conditions with four replicates each), 8 different second adaptors (L07clip2.0 to L14clip2.0) were used. 12 μ L of Lig-CLP mix was then added to each sample and mixed well, followed by the addition of 1 μ L RNA ligase. The samples were then incubated at RT overnight at 1100 rpm.

MyONE Silane cleanup: The next day, 5 μ L silane beads per sample were washed and resuspended in 60 μ L RLT buffer. The 60 μ L beads were then added to the sample along with 60 μ L 100% ethanol, mixed by pipetting and incubated at RT for 5 min. The step of mixing and incubation was repeated and the beads magnetically separated. The beads were then resuspended in 1 mL 80% ethanol and transferred to a new tube. This step was repeated two times and the beads were then air dried for 5 min at RT, resuspended in 23 μ L ultrapure water and incubated for 5 min at RT.

First PCR amplification: 2X Phusion high fidelity master mix (25 μ L) and P5Solexa_S and P3Solexa_S primer mix (2.5 μ L of 10 μ M stock) were mixed and added to 22.5 μ L cDNA. The following PCR program was then run:

Temperature	Time(s)	Cycles
98°C	30	1
98°C	10	6
65°C	30	
72°C	30	
72°C	180	1
16°C	hold	

First ProNex Size selection: ProNex chemistry was equilibrated to RT for 30 min and the beads were vortexed. Beads were then added to the sample in a 1:2.95 v/v ratio of sample:beads. The sample and beads were mixed by pipetting up and down 10 times and incubated at RT for 10 min. The sample was placed on a magnetic rack and was let stand for

2 min. The supernatant was removed and 500mL pf ProNex wash buffer was added to the beads while on the rack. The sample was incubated for 30-60 sec and the buffer was removed. The step was repeated once more and the samples were then air dried for 8 to 10 min until cracking of the beads begin. The sample was removed from the rack and 23μL water was added to elute the sample from the beads. The sample was incubated at RT for 5 min. The beads were then put back on the magnetic rack and the supernatant carefully transferred to a new tube.

Test PCR amplification: 9μL of PCR mix was prepared- 2X Phusion high fidelity master mix (5μL) and P5Solexa_S and P3Solexa_S primer mix (0.5μL of 10μM stock), and 3.5μL water; and added to 1μL cDNA. The following PCR program was then run:

Temperature	Time(s)	Cycles
98°C	30	1
98°C	10	6 and 9
65°C	30	
72°C	30	
72°C	180	1
16°C	hold	

2μL of the amplified library was run on capillary gel electrophoresis using the High Sensitivity D1000 kit in a TapeStation system, to test if the library looks good.

Preparative PCR: The same conditions of Test PCR amplification were repeated for 10μL cDNA. 2μL was again analysed using capillary gel electrophoresis, and if all looked well, the second half of the library was also amplified.

Second ProNex size selection: The ProNex size selection step was repeated, but with 1:2.4 v/v ratio of sample:beads, and the sample was eluted in 72μL of water.

The concentration of the library was calculated using Qubit hsDNA kit, and it was diluted to 10nM, 20μL of which was given for sequencing. The library was sequenced using NextSeq500 sequencing kit with Mid Output Flowcells.

***In vitro* UV crosslinking assay**

The assay was adapted from (Zarnack et al. 2013). 10% of oligonucleotide (BRE A' _II) was radioactively labeled with γ -³²P-ATP, incorporated using PNK as follows: 1μL of oligo

(10 μ M) was added to 39 μ L of UV CXL buffer. The reaction was incubated for 1 h at 37°C, followed by stopping the reaction at 95°C for 2 min. 9 μ L of unlabeled 10 μ M oligo was then added along with water to make up the volume to 100 μ L. The sample was passed through G-25 filter columns (GE healthcare) to remove unincorporated γ -³²P-ATP. 100nM of labeled probe was incubated with FMR1 (2 μ M). Reaction samples with Bruno in binding buffer CXL were prepared and added to the oligo and FMR1 mix such that final volume is 20 μ L. The reactions were incubated for 15 min at 37°C, and then UV crosslinked at 150mJ/cm² using Stratalinker 2400. 4x SDS loading dye (plus 100mM DTT) was added to the samples and heated to 95°C for 5 min. A NuPAGE Bis Tris 4-12% gel was run in MES running buffer and the gel was first exposed to Fuji film phosphor screen, and then stained with Coomassie. The phosphor screen was visualised using Typhoon FLA 9500 from GE healthcare.

Western blot for ovaries

Four pairs of ovaries were dissected in PBS and 80 μ L of 1X LDS sample buffer plus 10mM DTT was added. The ovaries were crushed using a pestle, and boiled at 95°C for 10 min. The sample was briefly centrifuged and then 20 μ L was loaded onto 4-12% bis-tris precast gel, and run at 180V for ~1 hour (till the dye reaches the bottom). The gel was removed, and transferred to nitrocellulose membrane using semi-dry blotting apparatus (Bio-Rad). The membrane was then blocked with 5% milk powder for 30 min at RT, followed by incubation in primary antibody (in blocking buffer) at 4°C overnight. The next day, membrane was washed in thrice, 10 min each at RT, followed by incubation in secondary antibody (in blocking buffer) for 2 hours at RT. The membrane was washed twice with blocking buffer, and once in PBST (PBS + 0.1% Tween-20) 10 min each at RT. The membrane was then developed using Immobilon Western HRP substrate peroxide solution (ECL).

For western blot of Oskar protein, all washes are done in TBST (Tris-buffered saline + 0.1% Tween-20) instead of blocking buffer or PBST.

qRT-PCR for ovaries

RNA extraction from ovaries: Three pairs of ovaries were dissected in ice cold PBS and 100 μ L of Trizol LS reagent was added to them. The ovaries were lysed using a pestle and 700 μ L

Trizol LS was further added (at this stage the sample can be stored at -80°C till use). 200 µL chloroform per 750 µL Trizol LS was added (210 µL in this case), vortexed and incubated on ice for 10 min. The sample was centrifuged at 13,200 rpm for 20 min at 4°C. The aqueous phase was extracted into a new tube and 0.5 µL glycoblue (stock 15 mg/ml) and 500 µL isopropanol per 750 µL Trizol LS (530 µL in this case) were added to the sample and incubated at RT for 10 min. It was centrifuged at 13,200 rpm for 20 min at 4°C. The supernatant was removed and 1 mL 70% ethanol was added to wash the pellet. It was again centrifuged at 13,200 rpm for 5 min at 4°C and the pellet was air dried and resuspended in 30 µL RNase/DNase free water. 0.1 volume (3 µL) of 10x DNase buffer was added to sample along with 1 µL Turbo DNase and incubated at 37°C for min. 500µL Trizol LS was then added and the RNA extraction steps from above repeated.

cDNA preparation: 1.5µg of RNA was used for cDNA synthesis using Superscript III First-Strand Synthesis Supermix kit, following the manufacturer's instructions. 2µL of this reaction was then used as template for qPCR using SYBR Green PCR mix. Step One Real Time PCR system from Applied Biosystems was used for qPCR using the Standard Quantitative conditions: Step 1- 95°C, 10 min; Step 2- 95°C, 15 sec; Step 3- 60°C, 1 min; Repeat cycles (step 2 and 3)- 40. 18S RNA was used to normalize the loading and Ct values were calculated. The experiment was performed in three replicates.

ePAT assay

8µL ePAT reaction was assembled in a 200µL PCR tube as follows: 1µg total RNA, 10mM ePAT-anchor primer and water upto 8µL. The reaction was incubated at 80°C for 5 min, and cooled to room temperature. 12 µL of ePAT master mix (4µL water, 4µL 5x Superscript III buffer (Life Technologies), 1µL 100mM DTT, 1µL 10mM NTPs, 1µL RNaseOUT (Life Technologies), 1µL (5U) Klenow polymerase (NEB)) was then added to the ePAT reaction, mixed thoroughly and incubated at 25°C for 1 h. The polymerase was inactivated by heating to 80°C for 10 min. The reaction was then cooled to 55°C for 1 min, and 1 µL (200U) of reverse transcriptase Superscript III (Life Technologies) was added and incubated at 55°C for 1 h. The reverse transcriptase was inactivated by heating to 80°C for 10 min. The cDNA thus prepared was used for PCR. It was diluted 1:6 by addition of 120µL dH₂O. 5µL of the diluted cDNA was used for PCR reactions of 20µL volume. A 2% agarose gel was run to detect the PCR amplicons. For analysis, Image J was used to get the plot profiles of the marker lanes.

These were then used to determine the standard curve between the log (molecular weight) of the DNA and distance from the wells. This was used to, in turn, determine the molecular weight of the amplicons.

Pole cell analysis

oskar-Gal4 driven control (*white*) RNAi and FMR1 RNAi flies were used for the pole cell analysis. Flies were fed for 48 hours on yeast at 25°C, and then transferred to cages with agarY plates (agar plus yeast) overnight. The next day, the plates were changed and the flies were allowed to pre-lay for 1 hour on fresh agarY plates, after which the plates were discarded. The flies were put on fresh agarY plates for 2 hours, followed by changing the plates and allowing the plates with laid eggs to incubate for further 2 hours at 25°C. Embryos from these plates were then collected and stored at -20°C after dehydrating in 100% methanol as follows: The embryos were dechorionated using 50% bleach (sodium hypochlorite 6-14%) in water for 2 min. They were then washed extensively and transferred to a tube. 500 µL of preheated (92°C) salt solution PC was then added, and the embryos were heat fixed at 92°C for 30 sec. 1 volume ice cold salt solution PC was then added. It was then removed and 1 volume heptane plus 1 volume methanol (500 µL plus 500 µL) were then added to the fixed embryos, and they were vortexed for 30 seconds. The embryos were allowed to sink to the bottom for 10 seconds and all the liquid with floating embryos was discarded. The sunken embryos were washed thrice with 100% methanol and then stored. The collections were performed till enough embryos could be collected.

The embryos were then rinsed thrice with PBST (PBS + 0.1% Tween-20), and then washed thrice with PBST for 15 min each at RT. They were then incubated in blocking buffer PC for 1 hour at RT. Primary antibody (anti-Vasa rat antibody) was diluted in blocking buffer and added to embryos. They were then incubated at 4°C overnight. The next day, the embryos were washed twice in blocking buffer PCA for 20 min each. They were again incubated in blocking buffer PCB for 1 hour at RT, then incubated in secondary antibody (anti-rat Alexa fluor 647) for 2-3 hours at RT on a shaker. The secondary antibody was removed and the embryos were incubated with DAPI (1:2500 in PBS) for 5 min. They were then washed twice with PBT(1x) for 20 min each. The buffer was removed and Vectashield was added to the

embryos as mounting media. The samples were then visualised under a Leica confocal SP8 at 20x and the pole cells were counted.

Cuticle preparation for embryos

oskar-Gal4 driven FMR1 RNAi and wild type flies were used for cuticle preparation. Flies were fed for 48 hours on yeast at 25°C. They were then transferred to cages with agarY plates (agar plus yeast) and incubated at 25 °C overnight. The next day, the laid eggs were aged for 24 hours at 25°C and then collected and dechorionated using 50% bleach in water (sodium hypochlorite 6-14%). They were then extensively washed with water and transferred to a glass slide, followed by mounting in Hoyer's medium and lactic acid (sigma). The embryos were then incubated overnight at 65°C and visualised under a bright field microscope.

Hatching rate analysis

oskar-Gal4 driven FMR1 RNAi and wild type (*w¹¹¹⁸*) flies were used for the analysis of the hatching rate. Flies were fed for 48 hours on yeast and transferred to cages with agarY plates (agar plus yeast) overnight. The next day, laid eggs were collected every 3 hours and counted as 'total eggs laid'. They were then incubated for 48 hours at 25°C, to allow hatching of the embryos, and unhatched eggs were counted. Hatched eggs were then calculated as the total number of eggs minus the unhatched eggs and the hatching was then determined. The experiment was performed in three replicates.

***In vitro* tethering assay**

Cloning of DNA plasmids for transcription of reporter RNAs: The DNA plasmid pFL-5xBoxB containing lucBoxB reporter under T3 promoter was a gift from Mandy Jeske (Heidelberg University). To clone lucosk reporter, pFL-5xBoxB was digested with BglII and BamHI to linearise. *oskar* 3'UTR was amplified from pUASp-*oskar* 3'UTR vector from (Jambor et al. 2014) using primer IVT 1 and primer IVT 2, and the amplicon was integrated into linearised pFL-5xBoxB using Takara In-Fusion HD cloning kit, following the manufacturer's instructions. To clone lucoskBoxB, pFL-5xBoxB was linearised using primer IVT 3 and primer IVT 4 and *oskar* 3'UTR was amplified from pUASp-*oskar* 3'UTR vector

using primer IVT 5 and primer IVT 6. The amplicon was integrated into the linearised pFL-5xBoxB using Takara In-Fusion HD cloning kit, following the manufacturer's instructions.

Preparation of RNA reporters: The plasmids were linearised by restriction digestion with BamH1 at the end of the reporter gene, in a 50 μ L reaction volume. 150 μ L of water was added to the sample and the digested DNA was then extracted using 200 μ L phenol:chloroform:isoamyl alcohol, pH 7.8/8.0 (25:24:1). Sample was vortexed for 2 min and centrifuged at 13,000 rpm for 30 min at RT. The top layer was collected into a new tube and 1 volume of chloroform was added to the sample. It was then vortexed, centrifuged for 5 min and the top layer was again collected. This step was repeated one more time. 1/9th volume of the aqueous layer of 3M sodium acetate was added to the sample along with 2.5x 100% ethanol. The sample was centrifuged at 13,000 rpm for 30 min at RT. The supernatant was discarded and 1mL 70% ethanol was added to the sample. It was centrifuged again for 5 min at 13,000 rpm at RT, the supernatant was discarded and the pellet was air dried and resuspended in 10-20 μ L water.

1 μ g DNA was then used for *in vitro* transcription by adding DNA to the IVT-A reaction mix for final volume 22.5 μ L and incubated at 37°C for 5 min. 2.5 μ L of GTP (10mM) was then added and incubated for another 1 hour at 37°C. 1.5 μ L of Dnase 1 was then added and incubated at 37°C for 20 min. The RNA was extracted using phenol chloroform as follows- The volume of the reaction was made up to 100 μ L with water, and 200 μ L of Phenol:chloroform:isoamyl alcohol, pH 4.8 (25:24:1) was added, vortexed thoroughly and centrifuged at 13,000 rpm for 30 min at RT. The aqueous phase was transferred to a new tube and 1 volume chloroform was added, vortexed and centrifuged at 13,000 rpm for 5 min at RT. The aqueous phase was extracted and the chloroform extraction step repeated. 1/3 the volume of aqueous phase of 10M ammonium acetate and 3.3x 100% ethanol was then added to the sample and centrifuged at 13,000 rpm for 30 min at RT. The supernatant was discarded and 1mL of 70% ethanol was added to wash the pellet at 13,000 rpm for 5 min at RT. The pellet was then air dried for 5 min, resuspended in 20 μ L water and incubated at 65°C for 15 min. 2 μ g of RNA was then loaded onto a 1% denaturing agarose gel, and stained with 1x SYBR Gold stain in 0.5x TBE, for 40 min on rotator, to check the quality of the RNAs.

Tethering assay: Master mix (MM)-T was prepared for each reporter RNA (2 μ L per reaction) and 2.5 μ L of protein (gel filtration buffer PP for no protein control) was added to the reaction along with 40% *Drosophila* embryo extract (4 μ L). The samples were pre-incubated at 22°C

for 20 min. 1.5 μ L ARS was then added to the reactions and samples were further incubated at 22°C for 60 min. The samples were then snap frozen in liquid nitrogen. A Mithras LB 940 plate reader was then used to measure the luciferase activity as a readout for the protein levels in each sample. 50 μ L of luciferase substrate was dispensed into 4 μ L of the reaction, it was shaken for 3 sec (2mm orbital), and the reading was taken for 7 sec. The values were normalized to the no protein control. All the samples were in three replicates each.

Prism 9 was used for statistical analysis. Unpaired t-test was used to compare the samples with no protein control.

Co-sedimentation Assay

60 μ M of porcine brain tubulin was polymerized into microtubules by incubating in 1xBRB80 in the presence of 1mM GTP γ S and 20 μ M Taxol at 37°C for 30 min. Meanwhile EB1-RNA (U₂₅) complexes were assembled by incubating 40 μ M EB1 with increasing concentration of RNA (0-300 μ M) on ice. The EB1-RNA complexes were then added to the polymerized microtubules such that the final concentration of tubulin was 30 μ M in each sample. The samples were further incubated at 37°C for 15 min, and then layered on top of a 30% sucrose cushion (in 1xBRB80 + 20 μ M Taxol). Samples were centrifuged at 80,000xg for 30 min in a Beckman SW55Ti rotor. The supernatant was removed and saved, and the pellet washed twice with 1xBRB80 (+Taxol) and resuspended in 50 μ L 1xBRB80. 5 μ L of supernatant and pellet samples were run on a 15% urea-PAGE to visualize the RNA (stained with methylene blue) and 5 μ L on 4-12% SDS-PAGE to visualize the proteins (stained with Coomassie Blue).

RNA-immunoprecipitation and sequencing (RIP-seq)

Ovaries from flies expressing EB1-GFP (or Flag-Myc-GFP as negative control) were harvested in PBS using a Kitchen Aid to grind the flies, and using sieves with sizes 400 μ and 200 μ to sieve the body parts and 80 μ sieve to finally collect the ovaries. The collected ovaries were lysed in lysis buffer A and cleared by centrifugation at 13,200 rpm at 4°C for 10 min. The supernatant was crosslinked with UV (254nm) in UV stratalinker 2400 at 0.3J energy. The lysate was then incubated with GFP-Trap beads at 4°C for 1.5h on rotator, to pulldown on EB1-GFP and the crosslinked RNAs. The beads were then washed with HS Buffer A, followed by MS Buffer A and finally LS Buffer A for 10 min each at 4°C on rotator. The

beads were resuspended in 100 μ L Proteinase K Buffer A and incubated with 0.2mg/mL Proteinase K at 55°C for 30 min. The RNA was then extracted using Trizol LS reagent, following the manufacturer's instructions. cDNA libraries were then prepared from the isolated RNAs using the kit SENSE mRNA-seq Library Prep V2. The libraries were sequenced using single end 50 sequencing on Illumina Hiseq2000 by the Gene Core Facility, EMBL, and the differential gene expression analysis was performed using DeSeq2 (Love et al. 2014).

References

- Akhmanova A, Steinmetz MO. 2008. Tracking the ends: a dynamic protein network controls the fate of microtubule tips. *Nature Reviews Molecular Cell Biology* **9**: 309-322.
- Anne J. 2010. Targeting and Anchoring Tudor in the Pole Plasm of the Drosophila Oocyte. *PLOS ONE* **5**: e14362.
- Ascano M, Jr., Mukherjee N, Bandaru P, Miller JB, Nusbaum JD, Corcoran DL, Langlois C, Munschauer M, Dewell S, Hafner M et al. 2012. FMRP targets distinct mRNA sequence elements to regulate protein expression. *Nature* **492**: 382-386.
- Athar YM, Joseph S. 2020. The Human Fragile X Mental Retardation Protein Inhibits the Elongation Step of Translation through Its RGG and C-Terminal Domains. *Biochemistry* **59**: 3813-3822.
- Barbee SA, Estes PS, Cziko A-M, Hillebrand J, Luedeman RA, Collier JM, Johnson N, Howlett IC, Geng C, Ueda R et al. 2006. Staufen- and FMRP-Containing Neuronal RNPs Are Structurally and Functionally Related to Somatic P Bodies. *Neuron* **52**: 997-1009.
- Bashirullah A, Cooperstock RL, Lipshitz HD. 1998. RNA localization in development. *Annu Rev Biochem* **67**: 335-394.
- Bechara EG, Didiot MC, Melko M, Davidovic L, Bensaid M, Martin P, Castets M, Pognonec P, Khandjian EW, Moine H et al. 2009. A Novel Function for Fragile X Mental Retardation Protein in Translational Activation. *PLOS Biology* **7**: e1000016.
- Berleth T, Burri M, Thoma G, Bopp D, Richstein S, Frigerio G, Noll M, Nüsslein-Volhard C. 1988. The role of localization of bicoid RNA in organizing the anterior pattern of the Drosophila embryo. *The EMBO Journal* **7**: 1749-1756.
- Berman RF, Buijsen RAM, Usdin K, Pintado E, Kooy F, Pretto D, Pessah IN, Nelson DL, Zalewski Z, Charlet-Bergeurand N et al. 2014. Mouse models of the fragile X premutation and fragile X-associated tremor/ataxia syndrome. *Journal of Neurodevelopmental Disorders* **6**: 25.
- Besse F, López de Quinto S, Marchand V, Trucco A, Ephrussi A. 2009. Drosophila PTB promotes formation of high-order RNP particles and represses oskar translation. *Genes & development* **23**: 195-207.
- Birsoy B, Kofron M, Schaible K, Wylie C, Heasman J. 2006. Vg1 is an essential signaling molecule in Xenopus development. *Development* **133**: 15-20.
- Biswas J, Nunez L, Das S, Yoon YJ, Eliscovich C, Singer RH. 2019. Zipcode Binding Protein 1 (ZBP1; IGF2BP1): A Model for Sequence-Specific RNA Regulation. *Cold Spring Harb Symp Quant Biol* **84**: 1-10.
- Boo SH, Kim YK. 2020. The emerging role of RNA modifications in the regulation of mRNA stability. *Experimental & Molecular Medicine* **52**: 400-408.
- Broadus J, Fuerstenberg S, Doe CQ. 1998. Staufen-dependent localization of prospero mRNA contributes to neuroblast daughter-cell fate. *Nature* **391**: 792-795.
- Busch A, Brüggemann M, Ebersberger S, Zarnack K. 2020. iCLIP data analysis: A complete pipeline from sequencing reads to RBP binding sites. *Methods* **178**: 49-62.

- Castagnetti S, Ephrussi A. 2003. Orb and a long poly(A) tail are required for efficient oskar translation at the posterior pole of the *Drosophila* oocyte. *Development* **130**: 835-843.
- Castello A, Fischer B, Eichelbaum K, Horos R, Beckmann BM, Strein C, Davey NE, Humphreys DT, Preiss T, Steinmetz LM et al. 2012. Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* **149**: 1393-1406.
- Castello A, Horos R, Strein C, Fischer B, Eichelbaum K, Steinmetz LM, Krijgsveld J, Hentze MW. 2013. System-wide identification of RNA-binding proteins by interactome capture. *Nature Protocols* **8**: 491-500.
- Chang JS, Tan L, Schedl P. 1999. The *Drosophila* CPEB homolog, orb, is required for oskar protein expression in oocytes. *Dev Biol* **215**: 91-106.
- Chekulaeva M, Hentze MW, Ephrussi A. 2006. Bruno Acts as a Dual Repressor of oskar Translation, Promoting mRNA Oligomerization and Formation of Silencing Particles. *Cell* **124**: 521-533.
- Chen E, Joseph S. 2015. Fragile X mental retardation protein: A paradigm for translational control by RNA-binding proteins. *Biochimie* **114**: 147-154.
- Chen E, Sharma MR, Shi X, Agrawal RK, Joseph S. 2014. Fragile X mental retardation protein regulates translation by binding directly to the ribosome. *Molecular cell* **54**: 407-417.
- Choudhury NR, Heikel G, Trubitsyna M, Kubik P, Nowak JS, Webb S, Granneman S, Spanos C, Rappsilber J, Castello A et al. 2017. RNA-binding activity of TRIM25 is mediated by its PRY/SPRY domain and is required for ubiquitination. *BMC Biol* **15**: 105.
- Chyung E, LeBlanc HF, Fallon JR, Akins MR. 2018. Fragile X granules are a family of axonal ribonucleoprotein particles with circuit-dependent protein composition and mRNA cargos. *J Comp Neurol* **526**: 96-108.
- Clark A, Meignin C, Davis I. 2007. A Dynein-dependent shortcut rapidly delivers axis determination transcripts into the *Drosophila* oocyte. *Development* **134**: 1955-1965.
- Coppin L, Leclerc J, Vincent A, Porchet N, Pigny P. 2018. Messenger RNA Life-Cycle in Cancer Cells: Emerging Role of Conventional and Non-Conventional RNA-Binding Proteins? *International Journal of Molecular Sciences* **19**.
- D'Souza MN, Ramakrishna S, Radhakrishna BK, Jhaveri V, Ravindran S, Yeramala L, Palakodeti D, Muddashetty RS. 2021. Function of FMRP domains in regulating distinct roles of neuronal protein synthesis. *bioRxiv*: 2021.2011.2015.468563.
- Darnell JC, Fraser CE, Mostovetsky O, Stefani G, Jones TA, Eddy SR, Darnell RB. 2005. Kissing complex RNAs mediate interaction between the Fragile-X mental retardation protein KH2 domain and brain polyribosomes. *Genes Dev* **19**: 903-918.
- Darnell JC, Jensen KB, Jin P, Brown V, Warren ST, Darnell RB. 2001. Fragile X Mental Retardation Protein Targets G Quartet mRNAs Important for Neuronal Function. *Cell* **107**: 489-499.
- Darnell Jennifer C, Van Driesche Sarah J, Zhang C, Hung Ka Ying S, Mele A, Fraser Claire E, Stone Elizabeth F, Chen C, Fak John J, Chi Sung W et al. 2011. FMRP Stalls Ribosomal Translocation on mRNAs Linked to Synaptic Function and Autism. *Cell* **146**: 247-261.
- Das S, Singer RH, Yoon YJ. 2019. The travels of mRNAs in neurons: do they know where they are going? *Curr Opin Neurobiol* **57**: 110-116.
- Das S, Vera M, Gandin V, Singer RH, Tutucci E. 2021. Intracellular mRNA transport and localized translation. *Nature Reviews Molecular Cell Biology* **22**: 483-504.

- Davidovic L, Jaglin XH, Lepagnol-Bestel A-M, Tremblay S, Simonneau M, Bardoni B, Khandjian EW. 2007. The fragile X mental retardation protein is a molecular adaptor between the neurospecific KIF3C kinesin and dendritic RNA granules. *Human Molecular Genetics* **16**: 3047-3058.
- Decker CJ, Parker R. 2012. P-bodies and stress granules: possible roles in the control of translation and mRNA degradation. *Cold Spring Harbor perspectives in biology* **4**: a012286-a012286.
- Deshpande G, Calhoun G, Schedl P. 2006. The drosophila fragile X protein dFMR1 is required during early embryogenesis for pole cell formation and rapid nuclear division cycles. *Genetics* **174**: 1287-1298.
- Dictenberg JB, Swanger SA, Antar LN, Singer RH, Bassell GJ. 2008. A direct role for FMRP in activity-dependent dendritic mRNA transport links filopodial-spine morphogenesis to fragile X syndrome. *Dev Cell* **14**: 926-939.
- Dold A, Han H, Liu N, Hildebrandt A, Brüggemann M, Rücklé C, Hänel H, Busch A, Beli P, Zarnack K et al. 2020. Makorin 1 controls embryonic patterning by alleviating Bruno1-mediated repression of oskar translation. *PLOS Genetics* **16**: e1008581.
- Dzhindzhev NS, Rogers SL, Vale RD, Ohkura H. 2005. Distinct mechanisms govern the localisation of Drosophila CLIP-190 to unattached kinetochores and microtubule plus-ends. *Journal of Cell Science* **118**: 3781-3790.
- Elliott SL, Cullen CF, Wrobel N, Kernan MJ, Ohkura H. 2005. EB1 is essential during Drosophila development and plays a crucial role in the integrity of chordotonal mechanosensory organs. *Mol Biol Cell* **16**: 891-901.
- Ephrussi A, Dickinson LK, Lehmann R. 1991. oskar organizes the germ plasm and directs localization of the posterior determinant nanos. *Cell* **66**: 37-50.
- Ephrussi A, Lehmann R. 1992. Induction of germ cell formation by oskar. *Nature* **358**: 387-392.
- Estes PS, O'Shea M, Clasen S, Zarnescu DC. 2008. Fragile X protein controls the efficacy of mRNA transport in Drosophila neurons. *Molecular and Cellular Neuroscience* **39**: 170-179.
- Feng Y, Absher D, Eberhart DE, Brown V, Malter HE, Warren ST. 1997. FMRP Associates with Polyribosomes as an mRNP, and the I304N Mutation of Severe Fragile X Syndrome Abolishes This Association. *Molecular Cell* **1**: 109-118.
- Fernández E, Li KW, Rajan N, De Rubeis S, Fiers M, Smit AB, Achsel T, Bagni C. 2015. FXR2P Exerts a Positive Translational Control and Is Required for the Activity-Dependent Increase of PSD95 Expression. *The Journal of Neuroscience* **35**: 9402.
- Fernandez-Chamorro J, Francisco-Velilla R, Ramajo J, Martinez-Salas E. 2019. Rab1b and ARF5 are novel RNA-binding proteins involved in FMDV IRES-driven RNA localization. *Life Sci Alliance* **2**.
- Gareau C, Houssin E, Martel D, Coudert L, Mellaoui S, Huot M-E, Laprise P, Mazroui R. 2013. Characterization of Fragile X Mental Retardation Protein Recruitment and Dynamics in Drosophila Stress Granules. *PLOS ONE* **8**: e55342.
- Gáspár I, Phea LJ, Ephrussi A. 2021. Egalitarian feeds forward to Staufen to inhibit Dynein during mRNP transport. *bioRxiv*: 2021.2004.2024.441269.
- Gáspár I, Sysoev V, Komissarov A, Ephrussi A. 2017. An RNA-binding atypical tropomyosin recruits kinesin-1 dynamically to oskar mRNPs. *The EMBO Journal* **36**: 319-333.

- Gaspar I, Wippich F, Ephrussi A. 2017. Enzymatic production of single-molecule FISH and RNA capture probes. *Rna* **23**: 1582-1591.
- Gelbart WM, Emmert DB. 2013. FlyBase High Throughput Expression Pattern Data.
- Geng C, Macdonald Paul M. 2006. Imp Associates with Squid and Hrp48 and Contributes to Localized Expression of *gurken* in the Oocyte. *Molecular and Cellular Biology* **26**: 9508-9516.
- Ghosh S, Marchand V, Gáspár I, Ephrussi A. 2012. Control of RNP motility and localization by a splicing-dependent structure in *oskar* mRNA. *Nat Struct Mol Biol* **19**: 441-449.
- Gireesh KK, Shine A, Lakshmi RB, Vijayan V, Manna TK. 2018. GTP-binding facilitates EB1 recruitment onto microtubules by relieving its auto-inhibition. *Sci Rep* **8**: 9792.
- Goering R, Hudish LI, Guzman BB, Raj N, Bassell GJ, Russ HA, Dominguez D, Taliaferro JM. 2020. FMRP promotes RNA localization to neuronal projections through interactions between its RGG domain and G-quadruplex RNA sequences. *eLife* **9**: e52621.
- González-Reyes A, Elliott H, St Johnston D. 1995. Polarization of both major body axes in *Drosophila* by *gurken-torpedo* signalling. *Nature* **375**: 654-658.
- Greenblatt Ethan J, Spradling Allan C. 2018. Fragile X mental retardation 1 gene enhances the translation of large autism-related proteins. *Science* **361**: 709-712.
- Gudimchuk NB, McIntosh JR. 2021. Regulation of microtubule dynamics, mechanics and function through the growing tip. *Nature Reviews Molecular Cell Biology* **22**: 777-795.
- Hahn I, Voelzmann A, Parkin J, Fülle JB, Slater PG, Lowery LA, Sanchez-Soriano N, Prokop A. 2021. Tau, XMAP215/Msps and Eb1 co-operate interdependently to regulate microtubule polymerisation and bundle formation in axons. *PLoS Genet* **17**: e1009647.
- Hentze MW, Castello A, Schwarzl T, Preiss T. 2018. A brave new world of RNA-binding proteins. *Nat Rev Mol Cell Biol* **19**: 327-341.
- Horos R, Buscher M, Sachse C, Hentze MW. 2019. Vault RNA emerges as a regulator of selective autophagy. *Autophagy* **15**: 1463-1464.
- Hughes JR, Bullock SL, Ish-Horowicz D. 2004. *inscuteable* mRNA Localization Is Dynein-Dependent and Regulates Apicobasal Polarity and Spindle Length in *Drosophila* Neuroblasts. *Current Biology* **14**: 1950-1956.
- Huppertz I, Attig J, D'Ambrogio A, Easton LE, Sibley CR, Sugimoto Y, Tajnik M, König J, Ule J. 2014. iCLIP: protein-RNA interactions at nucleotide resolution. *Methods* **65**: 274-287.
- Huppertz I, Perez-Perri JI, Mantas P, Sekaran T, Schwarzl T, Dimitrova-Paternoga L, Hennig J, Neveu PA, Hentze MW. 2020. RNA regulates Glycolysis and Embryonic Stem Cell Differentiation via Enolase 1. *bioRxiv*: 2020.2010.2014.337444.
- Ishizuka A, Siomi MC, Siomi H. 2002. A *Drosophila* fragile X protein interacts with components of RNAi and ribosomal proteins. *Genes & development* **16**: 2497-2508.
- Iwahashi CK, Yasui DH, An HJ, Greco CM, Tassone F, Nannen K, Babineau B, Lebrilla CB, Hagerman RJ, Hagerman PJ. 2006. Protein composition of the intranuclear inclusions of FXTAS. *Brain* **129**: 256-271.
- Jakobsen KR, Sørensen E, Brøndum KK, Dugaard TF, Thomsen R, Nielsen AL. 2013. Direct RNA sequencing mediated identification of mRNA localized in protrusions of human MDA-MB-231 metastatic breast cancer cells. *J Mol Signal* **8**: 9.
- Jambor H, Brunel C, Ephrussi A. 2011. Dimerization of *oskar* 3' UTRs promotes hitchhiking for RNA localization in the *Drosophila* oocyte. *RNA (New York, NY)* **17**: 2049-2057.

- Jambor H, Mueller S, Bullock SL, Ephrussi A. 2014. A stem-loop structure directs oskar mRNA to microtubule minus ends. *RNA (New York, NY)* **20**: 429-439.
- Jänicke A, Vancuylenberg J, Boag PR, Traven A, Beilharz TH. 2012. ePAT: a simple method to tag adenylated RNA to measure poly(A)-tail length and other 3' RACE applications. *RNA (New York, NY)* **18**: 1289-1295.
- Jenny A, Hachet O, Závorszky P, Cyrklaff A, Weston MD, Johnston DS, Erdélyi M, Ephrussi A. 2006. A translation-independent role of oskar RNA in early *Drosophila* oogenesis. *Development* **133**: 2827-2833.
- Jeske M, Bordi M, Glatt S, Müller S, Rybin V, Müller Christoph W, Ephrussi A. 2015. The Crystal Structure of the *Drosophila* Germline Inducer Oskar Identifies Two Domains with Distinct Vasa Helicase- and RNA-Binding Activities. *Cell Reports* **12**: 587-598.
- Jin P, Zarnescu DC, Ceman S, Nakamoto M, Mowrey J, Jongens TA, Nelson DL, Moses K, Warren ST. 2004. Biochemical and genetic interaction between the fragile X mental retardation protein and the microRNA pathway. *Nature Neuroscience* **7**: 113-117.
- Juge F, Zaessinger S, Temme C, Wahle E, Simonelig M. 2002. Control of poly(A) polymerase level is essential to cytoplasmic polyadenylation and early development in *Drosophila*. *The EMBO journal* **21**: 6603-6613.
- Kanke M, Jambor H, Reich J, Marches B, Gstir R, Ryu YH, Ephrussi A, Macdonald PM. 2015. oskar RNA plays multiple noncoding roles to support oogenesis and maintain integrity of the germline/soma distinction. *Rna* **21**: 1096-1109.
- Kim Tae H, Tsang B, Vernon Robert M, Sonenberg N, Kay Lewis E, Forman-Kay Julie D. 2019. Phospho-dependent phase separation of FMRP and CAPRIN1 recapitulates regulation of translation and deadenylation. *Science* **365**: 825-829.
- Kim-Ha J, Kerr K, Macdonald PM. 1995. Translational regulation of oskar mRNA by Bruno, an ovarian RNA-binding protein, is essential. *Cell* **81**: 403-412.
- Laver JD, Marsolais AJ, Smibert CA, Lipshitz HD. 2015. Chapter Two - Regulation and Function of Maternal Gene Products During the Maternal-to-Zygotic Transition in *Drosophila*. In *Current Topics in Developmental Biology*, Vol 113 (ed. HD Lipshitz), pp. 43-84. Academic Press.
- Lécuyer E, Yoshida H, Parthasarathy N, Alm C, Babak T, Cerovina T, Hughes TR, Tomancak P, Krause HM. 2007. Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. *Cell* **131**: 174-187.
- Liu L, Li T, Song G, He Q, Yin Y, Lu JY, Bi X, Wang K, Luo S, Chen YS et al. 2019. Insight into novel RNA-binding activities via large-scale analysis of lncRNA-bound proteome and IDH1-bound transcriptome. *Nucleic Acids Res* **47**: 2244-2262.
- Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**: 550.
- Macdonald PM, Kanke M, Kenny A. 2016. Community effects in regulation of translation. *eLife* **5**: e10965.
- Markussen FH, Michon AM, Breitwieser W, Ephrussi A. 1995. Translational control of oskar generates short OSK, the isoform that induces pole plasma assembly. *Development* **121**: 3723-3732.
- Martin KC, Ephrussi A. 2009. mRNA Localization: Gene Expression in the Spatial Dimension. *Cell* **136**: 719-730.

- Mazroui R, Huot M-E, Tremblay S, Filion C, Labelle Y, Khandjian EW. 2002. Trapping of messenger RNA by Fragile X Mental Retardation protein into cytoplasmic granules induces translation repression. *Human Molecular Genetics* **11**: 3007-3017.
- McMahon AC, Rahman R, Jin H, Shen JL, Fieldsend A, Luo W, Rosbash M. 2016. TRIBE: Hijacking an RNA-Editing Enzyme to Identify Cell-Specific Targets of RNA-Binding Proteins. *Cell* **165**: 742-753.
- Medioni C, Mowry K, Besse F. 2012. Principles and roles of mRNA localization in animal development. *Development* **139**: 3263-3276.
- Moissoglu K, Yasuda K, Wang T, Chrisafis G, Mili S. 2019. Translational regulation of protrusion-localized RNAs involves silencing and clustering after transport. *eLife* **8**: e44752.
- Monzo K, Papoulas O, Cantin Greg T, Wang Y, Yates John R, Sisson John C. 2006. Fragile X mental retardation protein controls trailer hitch expression and cleavage furrow formation in Drosophila embryos. *Proceedings of the National Academy of Sciences* **103**: 18160-18165.
- Moor Andreas E, Golan M, Massasa Efi E, Lemze D, Weizman T, Shenhav R, Baydatch S, Mizrahi O, Winkler R, Golani O et al. 2017. Global mRNA polarization regulates translation efficiency in the intestinal epithelium. *Science* **357**: 1299-1303.
- Muddashetty Ravi S, Nalavadi Vijayalaxmi C, Gross C, Yao X, Xing L, Laur O, Warren Stephen T, Bassell Gary J. 2011. Reversible Inhibition of PSD-95 mRNA Translation by miR-125a, FMRP Phosphorylation, and mGluR Signaling. *Molecular Cell* **42**: 673-688.
- Nakamura A, Amikura R, Hanyu K, Kobayashi S. 2001. Me31B silences translation of oocyte-localizing RNAs through the formation of cytoplasmic RNP complex during Drosophila oogenesis. *Development* **128**: 3233-3242.
- Nakamura A, Sato K, Hanyu-Nakamura K. 2004. Drosophila Cup Is an eIF4E Binding Protein that Associates with Bruno and Regulates oskar mRNA Translation in Oogenesis. *Developmental Cell* **6**: 69-78.
- Napoli I, Mercaldo V, Boyl PP, Eleuteri B, Zalfa F, De Rubeis S, Di Marino D, Mohr E, Massimi M, Falconi M et al. 2008. The Fragile X Syndrome Protein Represses Activity-Dependent Translation through CYFIP1, a New 4E-BP. *Cell* **134**: 1042-1054.
- Nehlig A, Molina A, Rodrigues-Ferreira S, Honoré S, Nahmias C. 2017. Regulation of end-binding protein EB1 in the control of microtubule dynamics. *Cellular and Molecular Life Sciences* **74**: 2381-2393.
- Nüsslein-Volhard C, Frohnhöfer Hans G, Lehmann R. 1987. Determination of Anteroposterior Polarity in Drosophila. *Science* **238**: 1675-1681.
- Pandey UB, Nichols CD. 2011. Human disease models in Drosophila melanogaster and the role of the fly in therapeutic drug discovery. *Pharmacol Rev* **63**: 411-436.
- Perez-Perri JI, Rogell B, Schwarzl T, Stein F, Zhou Y, Rettel M, Brosig A, Hentze MW. 2018. Discovery of RNA-binding proteins and characterization of their dynamic responses by enhanced RNA interactome capture. *Nat Commun* **9**: 4408.
- Preitner N, Quan J, Nowakowski DW, Hancock ML, Shi J, Tcherkezian J, Young-Pearse TL, Flanagan JG. 2014. APC is an RNA-binding protein, and its interactome provides a link to neural development and microtubule assembly. *Cell* **158**: 368-382.
- Ramos A, Hollingworth D, Pastore A. 2003. G-quartet-dependent recognition between the FMRP RGG box and RNA. *Rna* **9**: 1198-1207.

- Reveal B, Yan N, Snee MJ, Pai C-I, Gim Y, Macdonald PM. 2010. BREs Mediate Both Repression and Activation of oskar mRNA Translation and Act In trans. *Developmental Cell* **18**: 496-502.
- Rogers SL, Rogers GC, Sharp DJ, Vale RD. 2002. Drosophila EB1 is important for proper assembly, dynamics, and positioning of the mitotic spindle. *The Journal of Cell Biology* **158**: 873-884.
- Ryder PV, Fang J, Lerit DA. 2020. centrocortin RNA localization to centrosomes is regulated by FMRP and facilitates error-free mitosis. *Journal of Cell Biology* **219**: e202004101.
- Ryder SP. 2016. Protein-mRNA interactome capture: cartography of the mRNP landscape. *F1000Res* **5**: 2627.
- Santini E, Huynh TN, Longo F, Koo SY, Mojica E, D'Andrea L, Bagni C, Klann E. 2017. Reducing eIF4E-eIF4G interactions restores the balance between protein synthesis and actin dynamics in fragile X syndrome model mice. *Sci Signal* **10**.
- Specchia V, Puricella A, D'Attis S, Massari S, Giangrande A, Bozzetti MP. 2019. Drosophila melanogaster as a Model to Study the Multiple Phenotypes, Related to Genome Stability of the Fragile-X Syndrome. *Frontiers in Genetics* **10**.
- Stebbins-Boaz B, Cao Q, de Moor CH, Mendez R, Richter JD. 1999. Maskin is a CPEB-associated factor that transiently interacts with eIF-4E. *Molecular cell* **4**: 1017-1027.
- Su L-K, Qi Y. 2001. Characterization of Human MAPRE Genes and Their Proteins. *Genomics* **71**: 142-149.
- Sutandy FXR, Ebersberger S, Huang L, Busch A, Bach M, Kang HS, Fallmann J, Maticzka D, Backofen R, Stadler PF et al. 2018. In vitro iCLIP-based modeling uncovers how the splicing factor U2AF2 relies on regulation by cofactors. *Genome Res* **28**: 699-713.
- Sysoev VO, Fischer B, Frese CK, Gupta I, Krijgsveld J, Hentze MW, Castello A, Ephrussi A. 2016. Global changes of the RNA-bound proteome during the maternal-to-zygotic transition in Drosophila. *Nature Communications* **7**: 12128.
- Tabet R, Moutin E, Becker Jérôme AJ, Heintz D, Fouillen L, Flatter E, Krężel W, Alunni V, Koebel P, Dembélé D et al. 2016. Fragile X Mental Retardation Protein (FMRP) controls diacylglycerol kinase activity in neurons. *Proceedings of the National Academy of Sciences* **113**: E3619-E3628.
- Takatori N, Kumano G, Saiga H, Nishida H. 2010. Segregation of Germ Layer Fates by Nuclear Migration-Dependent Localization of Not mRNA. *Developmental Cell* **19**: 589-598.
- Todd PK, Oh SY, Krans A, He F, Sellier C, Frazer M, Renoux AJ, Chen KC, Scaglione KM, Basrur V et al. 2013. CGG repeat-associated translation mediates neurodegeneration in fragile X tremor ataxia syndrome. *Neuron* **78**: 440-455.
- Trcek T, Lehmann R. 2019. Germ granules in Drosophila. *Traffic* **20**: 650-660.
- Tsang B, Arsenault J, Vernon Robert M, Lin H, Sonenberg N, Wang L-Y, Bah A, Forman-Kay Julie D. 2019. Phosphoregulated FMRP phase separation models activity-dependent translation through bidirectional control of mRNA granule formation. *Proceedings of the National Academy of Sciences* **116**: 4218-4227.
- Vaishali, Dimitrova-Paternoga L, Haubrich K, Sun M, Ephrussi A, Hennig J. 2021. Validation and classification of RNA binding proteins identified by mRNA interactome capture. *Rna*.

- Van Treeck B, Protter DSW, Matheny T, Khong A, Link CD, Parker R. 2018. RNA self-assembly contributes to stress granule formation and defining the stress granule transcriptome. *Proceedings of the National Academy of Sciences* **115**: 2734.
- Vanzo N, Oprins A, Xanthakis D, Ephrussi A, Rabouille C. 2007. Stimulation of Endocytosis and Actin Dynamics by Oskar Polarizes the Drosophila Oocyte. *Developmental Cell* **12**: 543-555.
- Vanzo NF, Ephrussi A. 2002. Oskar anchoring restricts pole plasm formation to the posterior of the Drosophila oocyte. *Development* **129**: 3705-3714.
- Vasudevan S, Steitz JA. 2007. AU-rich-element-mediated upregulation of translation by FXR1 and Argonaute 2. *Cell* **128**: 1105-1118.
- Wessels H-H, Imami K, Baltz AG, Kolinski M, Beldovskaya A, Selbach M, Small S, Ohler U, Landthaler M. 2016. The mRNA-bound proteome of the early fly embryo. *Genome Res* **26**: 1000-1009.
- Wippich F, Ephrussi A. 2020. Transcript specific mRNP capture from Drosophila egg-chambers for proteomic analysis. *Methods* **178**: 83-88.
- Yano T, de Quinto SL, Matsui Y, Shevchenko A, Shevchenko A, Ephrussi A. 2004. Hrp48, a Drosophila hnRNPA/B Homolog, Binds and Regulates Translation of oskar mRNA. *Developmental Cell* **6**: 637-648.
- Zaessinger S, Busseau I, Simonelig M. 2006. Oskar allows nanos mRNA translation in Drosophila embryos by preventing its deadenylation by Smaug/CCR4. *Development* **133**: 4573-4583.
- Zarnack K, König J, Tajnik M, Martincorena I, Eustermann S, Stévant I, Reyes A, Anders S, Luscombe Nicholas M, Ule J. 2013. Direct Competition between hnRNP C and U2AF65 Protects the Transcriptome from the Exonization of Alu Elements. *Cell* **152**: 453-466.
- Zhang G, Xu Y, Wang X, Zhu Y, Wang L, Zhang W, Wang Y, Gao Y, Wu X, Cheng Y et al. 2022. Dynamic FMR1 granule phase switch instructed by m6A modification contributes to maternal RNA decay. *Nature Communications* **13**: 859.
- Zhang Y, Gaetano CM, Williams KR, Bassell GJ, Mihailescu MR. 2014. FMRP interacts with G-quadruplex structures in the 3'-UTR of its dendritic target Shank1 mRNA. *RNA Biology* **11**: 1364-1374.

List of figures

Figure 1.1: RNA-binding proteins play crucial roles in regulating RNA various functions	2
Figure 2.1: Examples of RNA localisation in different cell types across species	6
Figure 2.2: Mechanisms of translation inhibition used by RNA binding proteins	8
Figure 2.3: RNA localisation in the <i>Drosophila</i> oocyte is important for axes establishment	9
Figure 2.4: A) Schematic of a <i>Drosophila</i> ovariole. B) Localised <i>oskar</i> RNA and Oskar protein in the <i>Drosophila</i> germline	10
Figure 2.5: Functions of Short and Long Oskar protein at the posterior pole	11
Figure 2.6: RBPs identified as putative <i>oskar</i> -RNP complex components in the transcript-specific pulldown of <i>oskar</i>	14
Figure 2.7: Domain architecture of FMR1	15
Figure 2.8: Mutations in the FMR1 gene causes genetic diseases in humans	16
Figure 2.9: Different mechanisms used by FMR1 to repress the translation of its target RNAs	17
Figure 2.10: Role of FMR1 granules in maternal RNA decay during <i>Drosophila</i> embryogenesis	17
Figure 2.11: KH2 domain of FMR1 can bind to “kissing complex” RNA	19
Figure 2.12: A) and B) Co-localisation analysis between FMR1-GFP and <i>oskar</i> mRNA. C) CLIP results for FMR1-GFP	22
Figure 2.13: Co-localisation analysis between FMR1-GFP and <i>oskar</i> 3'UTR	23
Figure 2.14: FMR1 RGG box is required for binding to <i>oskar</i> -3'UTR <i>in vitro</i>	24
Figure 2.15: Fraction of FMR1 peaks in coding region, regions and UTRs of target RNAs as identified by iCLIP	25
Figure 2.16: FMR1 binding peaks on <i>oskar</i> RNA as identified in <i>in vitro</i> and <i>in vivo</i> iCLIP	26
Figure 2.17: In vitro crosslinking assay to study competition between FMR1 and Bru for binding to BRE A'_TypeII	28
Figure 2.18: A) and B) Western blot shows reduced Oskar protein levels in the absence of FMR1. C) qRT-PCR shows no change in RNA levels upon FMR1 knockdown. D) and E) ePAT assay shows no change in polyA length upon FMR1 knockdown	31

Figure 2.19: Pole cell analysis upon FMR1 knockdown	32
Figure 2.20: Cuticle preparation and hatching rate analysis upon FMR1 knockdown	33
Figure 2.21: FMR1 is a component of starvation induced P-bodies	34
Figure 2.22: <i>In vitro</i> tethering assay of FMR1 domains with reporter RNAs	37
Figure 2.23: Phase separation FMR1-Cterminal domain at 8 μ M concentration but not at 350nM	38
Figure 2.24: <i>In vitro</i> tethering assay with KH domain mutants	39
Figure 3.1: Domain architecture and nucleotide dependent microtubule binding of EB1	46
Figure 3.2: Preliminary data showing RNA binding ability of EB1 <i>in vitro</i>	47
Figure 3.3: Co-sedimentation assay of EB1 with microtubules in the presence and absence of RNA	49
Figure 3.4: RIP-seq of EB1-GFP to identify its <i>in vivo</i> targets	51
Figure 3.5: Co-localisation analysis of EB1-GFP with the top hits from RIP-seq	52
Figure 3.6: Analysis of localisation pattern of four of the putative EB1 targets upon EB1 knockdown	53
Figure S1: Oskar protein levels in flies expressing oskar-gal4 driven transgenes with FMR1-full length, FMR- Δ KH, and FMR1- Δ CTD	97

Appendix A: Supplementary Figures

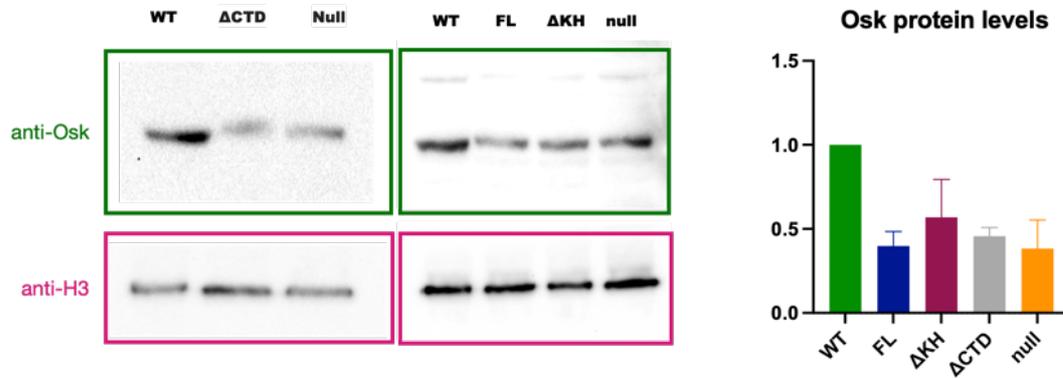


Fig. S1: *oskar*-Gal4 driven transgenes expressing full length FMR1 (FL), FMR1 lacking the C-terminal domain (Δ CTD) or FMR1 lacking the KH domains (Δ KH) failed to rescue Oskar protein levels when compared to wild type (WT) flies, and behaved like FMR1 null flies.

Appendix B: Supplementary Materials and methods

Oligos for smFISH probe synthesis

Name	Sequence : (5' to 3')		
asp_oligo1	ACTGTAGCGGATGATTCCG	asp_oligo34	GTCTCACGCCATCTCTTAG
asp_oligo2	CCATGTAAATGGCCTCTGG	asp_oligo35	TGCTGTAGAACGTAGCCCA
asp_oligo3	CTCACTGTCTTTGTCGCAC	asp_oligo36	TCACCAATGTTGGCGAGCA
asp_oligo4	GCGCAATAACATCTGGGCA	asp_oligo37	TGCGCTTCTGCTTAGCTTG
asp_oligo5	GTTACCCGTGGTGCTATTG	asp_oligo38	GTACTCTTCGGTGAGTGTG
asp_oligo6	GCGGGCCAAATTTAGGATG	asp_oligo39	GAGGATAAAGGTGCTGAGG
asp_oligo7	GCTTATCCACCTCTGATCG	asp_oligo40	TCTTCTCGCCAAAGACCAC
asp_oligo8	AACACCACATGAGCAGGTG	asp_oligo41	TGCGCACCTCGTTAAACAG
asp_oligo9	CTGATGCGATAAAGCGTCC	asp_oligo42	AAATCTGCGTCCAGGTCAG
asp_oligo10	ACTTTGCAGCGCACAGAG	asp_oligo43	GGTATGGAGACAAGGGCA
asp_oligo11	GTCGTCGGATTAGATAGCC	asp_oligo44	GTCAGCTTGATGACGATCC
asp_oligo12	TGGAATCGTTTTCGAGCCA	asp_oligo45	GTTGCTGCAAATGGATCCG
asp_oligo13	CAGGATCACAGAGGATTGG	asp_oligo46	CCGAGTCATAGAGCTTGAC
asp_oligo14	TTGGTAATCCTGTCTCGCC	asp_oligo47	AGGCTGAATGCTAGTGTC
asp_oligo15	TAGCATAGCACAGCGAGCA	asp_oligo48	TGCACTTCTGGGCAGAGTA
asp_oligo16	AAGTCGCACTTTGCGGATG	asp_oligo49	GTCCTCAACCCGTTGCTT
asp_oligo17	CGTATACTGCGCCACCAT	asp_oligo50	AGGAGGGGAAATGGCTACA
asp_oligo18	TAGCTGCCTGTTTCAGCTG	apc_oligo1	TCAAAGCCAGGGGACATTC
asp_oligo19	TGCTGCTTGACGCAGTTG	apc_oligo2	TCAGTTGGGTCTTGTACCG
asp_oligo20	GAAACTGTTTCTCTCCCG	apc_oligo3	GGCTATGGTGTAGGTCTTG
asp_oligo21	TTGAGCATGGAATCTGCGC	apc_oligo4	GCGATCCTTTGTGAGACG
asp_oligo22	AGATGTAGGTAATCCTGGC	apc_oligo5	CTTGTCTTTTCGGTGTGAGG
asp_oligo23	TCCTTTCGCATCTGCAGCA	apc_oligo6	ACGACGATTGGGTGATGGA
asp_oligo24	GTTCTTCTTTCTCGACGC	apc_oligo7	TTGTGGCATCCTCATCCTC
asp_oligo25	TGAAGCATTTCGCAGCCTC	apc_oligo8	GCCGTCTCATAGTTGGACA
asp_oligo26	ATTTTCCCTCTGTTCGCGC	apc_oligo9	GCTGTCCACACTGATCATC
asp_oligo27	TTCATGATGGCCCTCAGTC	apc_oligo10	TGACGGAGGCTTGATGTTG
asp_oligo28	ACATGATGAGGTCCCTCTG	apc_oligo11	ATCTCCGTGGTGTGCATCG
asp_oligo29	TCTGCCCAGCTGATAGCT	apc_oligo12	AGTCTTTGCGGCTCCAATC
asp_oligo30	TTTGCAAAGAGTCGTCGC	apc_oligo13	CTCCAGACGATCCTGAGTA
asp_oligo31	CCAGATCCGCTGAAGAAAG	apc_oligo14	ATCAGTTTCCCCTTCGAGC
asp_oligo32	GTCCGTCAACGATGTCTTG	apc_oligo15	TCCGGTTGCTGTTTCATTCG
asp_oligo33	ATTACCTCCATGACTCGGG	apc_oligo16	GCTGGGGAAACTCACATCA
		apc_oligo17	ACTCCATGGTCACTCTCTC
		apc_oligo18	CAGACTTCAGCATCAGGTC
		apc_oligo19	GATCCTTGTGCTTCTCGC
		apc_oligo20	TGAGAGCACTGTCCAACC
		apc_oligo21	TTTGAGCTGCTGGACTTC

apc_oligo22	TTGGCACGGCGTTGATCT	dhc_oligo6	TGTCCAATGTGTCGCAAGC
apc_oligo23	CGAGATGCCCCGATTGATG	dhc_oligo7	ACTCGTACTTCTTGCCCA
apc_oligo24	ATGCCCTTGAGGATGCACA	dhc_oligo8	CAAGAGGCACATATCGGAG
apc_oligo25	CGAAGAGCGACTTTGACTG	dhc_oligo9	AGGCCAGCAGGAAGTAAAG
apc_oligo26	GTTGGCTCTTGCGTACTG	dhc_oligo10	TGCGCAGCAGATTAGCTC
apc_oligo27	GTTGGTGTCTCTCCGAG	dhc_oligo11	TCTTCTCTCCAGCTGCAC
apc_oligo28	GACTTGCTCCTGTTGCCA	dhc_oligo12	CGTTCCTCAAGAGCACCCA
apc_oligo29	CTTCGCCTTCATCTCAGGA	dhc_oligo13	AGCTCGCTCAGCTTGATTG
apc_oligo30	CAGATAGCCAGGCAATTCG	dhc_oligo14	TCCTTCGGCTGAACCGATA
apc_oligo31	GGATCTGAAGTGCTTCAC	dhc_oligo15	GTTCAACTGCTTGTCCACC
apc_oligo32	GGGGTTTCTTCATGCCATC	dhc_oligo16	GGCGTTGGGCATAAAGTC
apc_oligo33	ATGAGCGTACATCTGTGCC	dhc_oligo17	GAAGCTTTCGGAAGATGGG
apc_oligo34	GTAGCTACGAATGGTCTCC	dhc_oligo18	CTTGACATAAGCAGGGCA
apc_oligo35	CCATGTTGATGCAGTTGGC	dhc_oligo19	GTGATCGGTTCTTCCTTCC
apc_oligo36	GGAGCTCATCTTCGTTGTC	dhc_oligo20	TAGCGAGTACTGGTAGAGG
apc_oligo37	TCGTCCACTATGCTCAGG	dhc_oligo21	CATCGTCAAGGATCTTGCC
apc_oligo38	ACTGAGCAGGTGTATGCTC	dhc_oligo22	GACTCTTTTCCAAGTGGCG
apc_oligo39	CGTTGAAGCTACTCAGGTC	dhc_oligo23	CAATATCCGGACGTTACAGC
apc_oligo40	CCTCGAAAACGCTGCGTAA	dhc_oligo24	TTAGCACCTGGTTGAGGCA
apc_oligo41	AACCGGTGGAGAATCGACA	dhc_oligo25	GCGAGTAGACAGGAAGATG
apc_oligo42	TTTGTGTGGGGGAATCTGG	dhc_oligo26	AAAGGACGGCGACAAATCG
apc_oligo43	CTCGGATGGGGATATTACC	dhc_oligo27	GACCAGCAATGGATTACCG
apc_oligo44	ACTAACCACCGAGCTCTTG	dhc_oligo28	ACGTAGTGCGGACTCCAAA
apc_oligo45	TGATATCCGCATCCGTTCC	dhc_oligo29	CTTGCGGAACGAATCATCC
apc_oligo46	CGTAGCGACTAAAGTACCC	dhc_oligo30	GAACGAGGTCTTCGTGATC
apc_oligo47	GGGTTTCTCTGGAGTGAC	dhc_oligo31	GCGCTTCAACATAATGGCG
apc_oligo48	TGGGAAGCTTTTTGGGAGC	dhc_oligo32	AAGATATTCAAGTGGGGCG
apc_oligo49	ACTTCTGGCTTACTTCCG	dhc_oligo33	AATGGTGGACATCTGCGAC
apc_oligo50	TCTGCATCCGCTTTAGCAG	dhc_oligo34	GAAAGTCTCGCTAGTGGAC
apc_oligo51	ACATAAGGGGTATCCTCCG	dhc_oligo35	CTCTCAGCAGTGAATGG
apc_oligo52	CGGATTTCTCGGGAAGTGG	dhc_oligo36	GAGCCTGGGAAATAAGCTG
apc_oligo53	CAGCAGGAGGAGTGATAG	dhc_oligo37	CATACTCCTTGTAGGC
apc_oligo54	CTGATATGTCCAGGTCGGA	dhc_oligo38	CAATACTGCGCTCCAGTTG
apc_oligo55	CCTGCTCCATTTCCAGATC	dhc_oligo39	GATCTGAGCAATGGCCCA
apc_oligo56	AGTCCGTGATCTTGGGCT	dhc_oligo40	CTGGCGGATTAACCTTC
apc_oligo57	CCTTTCGATGGGCATCTGA	dhc_oligo41	TTTCGCGCATAATCACGGC
apc_oligo58	CAGAGGCTGAGCGAATTAG	dhc_oligo42	CTTGCGGATTGACTTGACC
apc_oligo59	GATCGTAAGCGCAATCGGA	dhc_oligo43	GCTTGCGCATCAATGACAG
dhc_oligo1	AGCTGCTGCATCTTGAGCA	dhc_oligo44	GGTCAGCCAAGCGTATTTG
dhc_oligo2	GGTAAGCAACACCTTCTCC	dhc_oligo45	TCTGCTTGGCTTGGCGT
dhc_oligo3	GTTTCGGTAATCCAAGCCAC	dhc_oligo46	TCCTGCTTCTCACAGCCA
dhc_oligo4	TGCCATCGGGCATGGTAA	dhc_oligo47	GCTCAACAGTTTCGGCGA
dhc_oligo5	TTCTCTGGTGGCAGATTGG	dhc_oligo48	AGTCCAGATAGTGACGTGG

dhc_oligo49	TGGACATACACGCACGAG	dop_oligo16	CGTGGTATGCAGAGTCTTC
dhc_oligo50	AACGGATGGGAAGAAGTCC	dop_oligo17	TGCAGCTTGTGCTTGAGAC
dhc_oligo51	AAGTGGCAGCACGATCCT	dop_oligo18	TGTAGAGTGGATGGACGC
dhc_oligo52	GTCCATCGGTGGAAGGAT	dop_oligo19	GTTAGGCGAAGAGGATGAG
dhc_oligo53	CTTGTACAGCTCGTCACTG	dop_oligo20	CGACGTGGATTGCGATGAA
dhc_oligo54	GTCCAACATAAGACCTCG	dop_oligo21	AGGCGACAAATTTCTGGCC
dhc_oligo55	TCAGAGTGGTGTACTCGTC	dop_oligo22	TCTGCATTTGCACGCTTGC
dhc_oligo56	TCATACGCTCCAGGAAACC	dop_oligo23	ATTCACACCCTTCTTGGCC
dhc_oligo57	GGACAAGCCATTCATCCAG	dop_oligo24	ACTAAGAAGCAGCTGCAGC
dhc_oligo58	TCAGCAGCAAGTGACCTTG	dop_oligo25	CAGCCTCCCCATTAACATG
dhc_oligo59	CTTCGTGCAACAGTACGAG	dop_oligo26	TGGCGATTGAATGCCACAC
dhc_oligo60	AGAAGACCTTAAGTCGGGC	dop_oligo27	GACCGATGATGCTTCACAC
dhc_oligo61	GGACATATTCACGCAGCTC	dop_oligo28	TGGTTGATGCTGCAAAGCC
dhc_oligo62	CCTTGCTTAGCCAGTTGC	dop_oligo29	CGGAGATTGTTCTTGGC
dhc_oligo63	AAAGGATCGGACGCTGCAA	dop_oligo30	GCAAATGTCTCTGGCGCA
dhc_oligo64	AACGGGAAGGGAATCCAAC	dop_oligo31	CACCCTTTCGTTTCCGGT
dhc_oligo65	ACCATGGCGTTGGTCAAAAG	dop_oligo32	CGTTTGCGAACTCTCTGG
dhc_oligo66	CTGCATAACCACGAAGAGC	dop_oligo33	ATGCTTCAGCATTTCGACG
dhc_oligo67	TCAGAGATGTCTCTCCAGG	dop_oligo34	AGTGTTCATCCGTGTCATCG
dhc_oligo68	CAAGAACCAGTGCAGAAAGC	dop_oligo35	ATAGCGATCCATTCCGGGTG
dhc_oligo69	ATCGCTGGCCCTATAGAAG	dop_oligo36	CTGGTATCATCATCGTGGG
dhc_oligo70	GGAACGAGATTACACGCTG	dop_oligo37	AACTGCGGCACAACTCTG
dhc_oligo71	TGCCATAGCTGTCCATGTC	dop_oligo38	CTGCCGAGAAGAGAATTC
dhc_oligo72	TCTTGCGGTACTCGCAGTA	dop_oligo39	CTCCTTCATCTCCAGTGCA
dhc_oligo73	TCAGGCCTACCACTTCCA	dop_oligo40	AGTTCCTCAGTGGTTTCGC
dhc_oligo74	CGCAGAGCACCAATGGTT	dop_oligo41	AAACGGCACACAACCGATG
dhc_oligo75	GCTCAGCCAGCCATGTATA	dop_oligo42	AGCCTTGGCGCAATATCAC
dhc_oligo76	CACCACTCATGTTGACCTC	dop_oligo43	AAAACCTGCCTCGTCTCCGA
dop_oligo1	CTACTTCTCTGCTTGGCA	dop_oligo44	GATCAAAAGGTTGTCCGGC
dop_oligo2	AGAAGTGGCTGCAGATCC	dop_oligo45	CCAGGCACAAGTGCTTCT
dop_oligo3	TGATCCAGAGGATCCTCCA	dop_oligo46	CGAAAGAATGTCACGCTCC
dop_oligo4	TGTGGGGCTTAGTTTTGGG	dop_oligo47	GTGGTCTTGTGCTTCACCA
dop_oligo5	GAAGATGAGGATTCCCCC	dop_oligo48	TAGACGGCTCCATAAGCAC
dop_oligo6	AAATCCCTCGGACACACTG	dop_oligo49	TCGTTTTCTTGGGGTGAGG
dop_oligo7	GCTGTCTTGGGTACTAGTG	dop_oligo50	GAACTCGATCCTCCTCCAA
dop_oligo8	AGACCACTGTTGCCATAGC	dop_oligo51	GCACTACTTGATCCTCCC
dop_oligo9	ACTGAAGGAGCCCCAAAA	dop_oligo52	TGCTTGGCTGCACTGGTTA
dop_oligo10	ATAGGCGGCGAACAGCATA	dop_oligo53	GGCGAGTTGAGTAGCAAAC
dop_oligo11	GGCGATATCTTTGTCTCCG	dop_oligo54	CTGCAGTTCAGCGATTGGA
dop_oligo12	TGGTTCGTGCAAATCCTGC	dop_oligo55	CTAGGCATTCCAGCAATCG
dop_oligo13	CTTCGTGTTTGCCTTACC	alphaspec_oligo1	GTGACGCTTTAGGATTGCC
dop_oligo14	GATTGTGTGGTGTGTTGAGG	alphaspec_oligo2	CAGTTGGTCTTAAGTGCC
dop_oligo15	ATGTGACCCACAGACTTCC	alphaspec_oligo3	GATGTTGTGTATGCCCTCC

alphaspec_oligo4	CTCGAAGGCATTGAGACC	alphaspec_oligo3	TTCTGAACGGAGGTCAGG
alphaspec_oligo5	ACTCTCCGTGGTGGTTCT	alphaspec_oligo4	AATTGACCCTCGATCTCGC
alphaspec_oligo6	CGCAGATCAAAGAACAGCG	alphaspec_oligo4	AGCCACAGCTCAATGTCC
alphaspec_oligo7	CAGCCAGGTTCTTTAGCTC	alphaspec_oligo4	ATGGTGCGGTTGAACTGC
alphaspec_oligo8	TTCTTGAGGTTCTGGACGC	alphaspec_oligo4	TGTTGGCATGCCTCGTTGA
alphaspec_oligo9	GTCAGATCGCGACCATAG	alphaspec_oligo4	TGTTGGCATTGAGCTCGTG
alphaspec_oligo1	TCTCCTTGATCCAGCTCTC	alphaspec_oligo4	GAATCCTCCAGCAATTGGC
alphaspec_oligo1	ATCGGCCTGGTTGTTTCATG	alphaspec_oligo4	AGCTTCTGTCCGGATTGAG
alphaspec_oligo1	GTCCTTGAGCTCTCTTCG	alphaspec_oligo4	TCCTGGTGACGTTCCAGAA
alphaspec_oligo1	TTCTCGATCAGACCCTCC	alphaspec_oligo4	GCCTTCATGCCATTGATCC
alphaspec_oligo1	GTCCTCGTGCTTCTTGATC	alphaspec_oligo4	GGACTCATCTAGCTTCTGC
alphaspec_oligo1	TCGCAGTCACGCATGTAGA	alphaspec_oligo5	GCCAGTAGTTGGCAATCTC
alphaspec_oligo1	GCAGATCCAAGTTCTGCTC	alphaspec_oligo5	CCTGTTTGTACGGATCTG
alphaspec_oligo1	CTCCTTGATTCAGGCGAC	shot_oligo1	ATTCTTCCGGTGCGCATG
alphaspec_oligo1	GTTCCGGTGTTCCTGATGAC	shot_oligo2	CACTTGTTCCGTTTCTCGC
alphaspec_oligo1	TCTGGTCCCACATCTCGT	shot_oligo3	TACTGCATCATGCCACTCC
alphaspec_oligo2	TCTCCTTCTGCTTGGCGTA	shot_oligo4	TTGTAGCCATGGTCCAGAG
alphaspec_oligo2	CATTGGCCTTCTCAGCGA	shot_oligo5	CACTCGTTCCCAACGATG
alphaspec_oligo2	CCTTGGTCTCGTCAATGTC	shot_oligo6	CAAGATCACGTCCTGCTTC
alphaspec_oligo2	TGTTTGCCAGACGCACTTC	shot_oligo7	TCTTGTCCAGCAGCAGCA
alphaspec_oligo2	CTCCCACAACATCAGCGA	shot_oligo8	TTCCAGAACGCGTGAGAC
alphaspec_oligo2	AGAACGTACGCCTTCACAG	shot_oligo9	CAAAGGCTTGCAGTGTGTC
alphaspec_oligo2	CGTTTAGCTTGTCTGGCG	shot_oligo10	TCTTCTGTCCGGATGCACG
alphaspec_oligo2	GCGAGTTGATCTGGTTCTG	shot_oligo11	CTGTAAGTTGCTGGACGAG
alphaspec_oligo2	TTCATCGAGACCTCACGAG	shot_oligo12	GGGTCTGTTTTTCATCCAGC
alphaspec_oligo2	GATATCCACAACAGGGGTC	shot_oligo13	CCAGTTGCTCAGTAGCAG
alphaspec_oligo3	CCAGTTGCAATCGGTTCC	shot_oligo14	CCAGCCATCCCAGAATATC
alphaspec_oligo3	TTGGCGTCGGCAAAGTAC	shot_oligo15	GGACCTCGGCAATATATCC
alphaspec_oligo3	TGTGCCTGGAGAGAGTCA	shot_oligo16	TGGCCTTTTGCAGAAGCTG
alphaspec_oligo3	GACTTCTCCTCAGCGTG	shot_oligo17	AGCTTCTCCTGGGTAGTAG
alphaspec_oligo3	TTCTCGCCGAAGTATGATCA	shot_oligo18	CATCGTTCAGTGTGTCCAC
alphaspec_oligo3	TTGTTGATCTCGGCCAGCA	shot_oligo19	AGTTCTGGTGGGCTGAA
alphaspec_oligo3	GCACACCAATCAAGTCACG	shot_oligo20	GGCCAATTCAACCTCTAGC
alphaspec_oligo3	GTTCCTTCTCAGGATCCA	shot_oligo21	TTGCGGATCTCCAGGAATC
alphaspec_oligo3	CCCTCCTTGTGCGAATG	shot_oligo22	CTTCAGTTGGTCCAGTGTG
		shot_oligo23	GGAATTGACCCAAGTGCAG
		shot_oligo24	TCAATGTCCGAACCCACAG
		shot_oligo25	CCCAGATGGGCAAAGTTC
		shot_oligo26	TCGGCCTTGGTCAAGAAC

shot_oligo27	AGGAGGTTCTGCAGTGTC	ask1_oligo10	TTCTGATCAGCGTCGACC
shot_oligo28	TGGAACTCCATGGCCTTC	ask1_oligo11	GCTCGTCTCACTAATGCTC
shot_oligo29	TGGCATCGATGAGGTTCTC	ask1_oligo12	GGGAATTTTCGCCGTTCAAG
shot_oligo30	TTGTCCCATGCGTTGTCCA	ask1_oligo13	TAAACTGGTTGCTGGGCAG
shot_oligo31	CTCCAGCTTGGCCTTAATG	ask1_oligo14	GCGTTATGTCCAATTCCGG
shot_oligo32	GTTGCCAGCCTTAGCGATA	ask1_oligo15	TTGTGGGCGTATTGCAGG
shot_oligo33	CTGCTCACGTATCTTCTGC	ask1_oligo16	TCAGTATTGATGGGCAGGG
shot_oligo34	TGATTGACCTGATCCACGG	ask1_oligo17	GTAGGAAGGGATCTTCCAG
shot_oligo35	ACCAATGCACCCAAAGCC	ask1_oligo18	CGTGTTAACACAGCACGTTG
shot_oligo36	GTTTCGTTCTTGTCTGCTG	ask1_oligo19	CGCGATGAACAATGTCCTG
shot_oligo37	TCTGCTGCTCCATGTCATC	ask1_oligo20	CGAGGATCTGCTTGGAGTA
shot_oligo38	GGCGAAATGTTTCAGACAGG	ask1_oligo21	CTCTAAGAGGTCGAAAGC
shot_oligo39	AGCAACAGTGTGACGATC	ask1_oligo22	TTGAAGAAGCCGTTCTCCG
shot_oligo40	CATCTCGTACACGTCCCT	ask1_oligo23	TTCCGGCACCTCCTTAATG
shot_oligo41	CTCCGCCTTGC GTTGAAT	ask1_oligo24	ACCTTTGCCAGCACCAT
shot_oligo42	ACTGTCCAGAATAGGACGC	ask1_oligo25	GCGATTCTGCTCATCGTAG
shot_oligo43	ATCAGCCAATTCACCTCC	ask1_oligo26	GATCTGCAGTCATCTCCAG
shot_oligo44	CCGAAGCACATCAGCTTGA	ask1_oligo27	CTTGTAAGGCTGACGGAC
shot_oligo45	ACGGTTGTGAGCTTCATGG	ask1_oligo28	CTTTGCGTGAGCCAAACAG
shot_oligo46	TGCTGCACCAATGGTAGTG	ask1_oligo29	GCTGTTCTTAACCTCCTCG
shot_oligo47	TCTGTGACTCTTCAGCAG	ask1_oligo30	GTCCATCCAGAACTGGAAC
shot_oligo48	TCCACATTCGAAGCGTGAC	ask1_oligo31	TCAGATACCAGTTGGGTGG
shot_oligo49	GCTTTTCAGCGTAAACGGG	ask1_oligo32	AGCTTGAACATGCACTCGG
shot_oligo50	AGACCAAGTGACGCAATCC	ask1_oligo33	TCGTTGGCCAAAACGGAGA
shot_oligo51	CCCAGATTAACGGCTTCC	ask1_oligo34	CAAAGAAGGTGGCCACATC
shot_oligo52	AGTCGGGCTTCTTTCCCAA	ask1_oligo35	ACTCACTCAGCGAGGATAG
shot_oligo53	TTTCGTCGTGTAGACGGTC	chc_oligo1	CTACAAGTAGGGATAGCCC
shot_oligo54	CGTGTTGTCCAGCCAAAG	chc_oligo2	TCAGTTGTGGCTCCATCTG
shot_oligo55	TCGCTTAGCCACTTCCATG	chc_oligo3	GTGTATTTCGCGCAGAACC
shot_oligo56	CCTCAAGATCCAACCTCAG	chc_oligo4	GGAACCAACCCAACAACCTC
shot_oligo57	GCAACTTCCTTGCCGAGA	chc_oligo5	GTACAAGTAGGCGGCAATC
shot_oligo58	GCTAGAGAGCCCATAGAG	chc_oligo6	TCGAGTTTCTGTGCCAACG
shot_oligo59	GCTGTGCCTTTACCACCT	chc_oligo7	TTAAGAGTCCGTTTCAGGGC
shot_oligo60	GATTAGCCACCATCTCCTC	chc_oligo8	GAGAGATTGACTGAACGC
ask1_oligo1	TTCCGGGAGTTGTGGTTG	chc_oligo9	ATAAGGCTTGACGAGTGCC
ask1_oligo2	GCGCAAAAGAGCTATTGCC	chc_oligo10	ACAGCACGAGTGTGATCCA
ask1_oligo3	GGTGGACTGTCCACACAA	chc_oligo11	TTCAACAGCAGCGGCTTGA
ask1_oligo4	CCTTCATTCAGGATGAGCG	chc_oligo12	GGAATGTTAACGCGAGACC
ask1_oligo5	TCTGCTTTTCTGGCTCTC	chc_oligo13	AACAGCTCCAAGTGTTCGC
ask1_oligo6	TCCAGACAGGTCTTTGTCC	chc_oligo14	CAGATCCTCTAGCTCATCG
ask1_oligo7	GAGAAAATAGTGACCGCCG	chc_oligo15	ATGCACCACAATGTGCAGG
ask1_oligo8	TGATGCAATGGTACCGCAG	chc_oligo16	GCACATCTGAGCTAGCCTA
ask1_oligo9	TCCTCGGATTAGAACCTCC	chc_oligo17	GGTAGACCAAAGTGATGGC

chc_oligo18	ATGAACTCCTCCAGATCGG	dcr-1_oligo16	ATCCGGAATGCTGTGTTGC
chc_oligo19	CAGATAGCGAACGAGGTCA	dcr-1_oligo17	GCATCGGCTTTTTTCGCGAA
chc_oligo20	TGTACGAGTTCGATAGCCTC	dcr-1_oligo18	CTCGCAGATTTGCACACTG
chc_oligo21	CTTGATCGATGAGCACCTG	dcr-1_oligo19	AGCTTCCAATGGTGAGTGG
chc_oligo22	TGGCCGATGTGTTACATC	dcr-1_oligo20	TCTCTAGCTCCTTTGGCAC
chc_oligo23	GTAGTTCTCCAGCCGGTTA	dcr-1_oligo21	GTCGTGCGGCTCGAATTTA
chc_oligo24	ATCAGCCTTGATGGCTGTG	dcr-1_oligo22	TCACCCAGTCTCTTACGTC
chc_oligo25	GACAAAGCGGTCTGTACCA	dcr-1_oligo23	GCTCCAGATTGATGCCATC
chc_oligo26	TGATCGATCAACTGGCGC	dcr-1_oligo24	GCATTGGACATTGTGAGGG
chc_oligo27	TGGATTGCTCTCCGAAAGG	dcr-1_oligo25	CGGTTGTCGATCGAAACTG
chc_oligo28	GTGCTTCGCTCTTGAACAG	dcr-1_oligo26	GCCAGCTGGTAATATGGC
chc_oligo29	CTCGTTACAAACGGCGATC	dcr-1_oligo27	GGTTATGGCCTTCGTGGAA
chc_oligo30	CAGGGAAGGAGAAGCTTGA	dcr-1_oligo28	CATTGTAGGTCTCCTGCTC
chc_oligo31	ACCAGTTCGTCGGTTGAGA	dcr-1_oligo29	GCTCTACCAAATCCGACAG
chc_oligo32	TGTCCCTTGACCACGAGAA	dcr-1_oligo30	TAGAACCAACTGCTGGTGC
chc_oligo33	ATGGTAGCTGATCCGTCAG	dcr-1_oligo31	CTGATACTGCCGCTCGTTA
chc_oligo34	GGCCTCCTTCAAGAAGTTC	dcr-1_oligo32	ATTGGTCGCTGGATGATG
chc_oligo35	AGTTCTGGCGCAAATTCGC	dcr-1_oligo33	CACTTCCTGTGCGGTTTCA
chc_oligo36	CGGCCCGCTTAATGTCATA	dcr-1_oligo34	GTCTCCGCTTCATTCTTGG
chc_oligo37	GGGTGAACATAGCATTGCC	dcr-1_oligo35	TGTAATTGAGCGGACCTGC
chc_oligo38	CAGCAGGAACACGTAATCG	dcr-1_oligo36	ATCGTCGTCATCGTAGCTG
chc_oligo39	TAGATGGACAGGGCAAGTG	dcr-1_oligo37	CCCTGCTTATTCTGACTGC
chc_oligo40	CCAACTTCTCCTCCTTCAG	dcr-1_oligo38	AAACTGGGAGATGGAAGGC
chc_oligo41	ACTTCTCGCACAGCTGCT	dcr-1_oligo39	ACTAGCTATATCGTCGGCC
chc_oligo42	ACTTGTTAGCTTGCCCTG	dcr-1_oligo40	GAAGTCATCAGCCTCTTGC
chc_oligo43	GAAACGCTGGATCGTCTG	dcr-1_oligo41	CTTCTCCTCCTCCCTCAATG
chc_oligo44	TGCCAGGGCAGCAACTTTA	dcr-1_oligo42	TTGTTTGGACTCCCAGCA
chc_oligo45	AATTGTTGCGCACGGCCA	dcr-1_oligo43	CTTGAGCACCTCCGATAGA
dcr-1_oligo1	CAATGCGATAGTTGCGTCC	dcr-1_oligo44	GGCGTGAGGAAGTTTAACC
dcr-1_oligo2	TGCCACGGAAGGTTCCCTT	dcr-1_oligo45	AATAGCGGCTGCGAGGTA
dcr-1_oligo3	GAAGACATCCACGGTAACG	dcr-1_oligo46	GTAGTTGTCACCAGGGAAG
dcr-1_oligo4	GTCCAGCGACATGTTTGAG	dcr-1_oligo47	GCTGAGTGGGGATAGATG
dcr-1_oligo5	GAGAAAAATGGCACCTGCG	dcr-1_oligo48	AAGCGTTGCGGATCAAACG
dcr-1_oligo6	TGTCCATTCTCCTGCTGGA	dcr-1_oligo49	TCCATTCGCTTGATCAGC
dcr-1_oligo7	CCGCAAATCCGTTAATGCG	dcr-1_oligo50	CAGAAACTGCCAGTCGATG
dcr-1_oligo8	ATCTAGAACAGCATCGCCC	dcr-1_oligo51	ACTCCAGGGAAACCTTCAC
dcr-1_oligo9	GAACTCCAGACGCTGATAG	dcr-1_oligo52	CACCGGAACGCGTGAATA
dcr-1_oligo10	ATCCGTCAATCGATTGGGC	dcr-1_oligo53	AGAAAGCACTCAGCTTGGG
dcr-1_oligo11	TGCAACAGGTACGACCGA	dcr-1_oligo54	AATCCCTGCTGCGCATCT
dcr-1_oligo12	CTCAAACCTCCTCAAAGCCG	dcr-1_oligo55	GGCGGATAAATCTTGCGTC
dcr-1_oligo13	TAGTAACTGGTCCAGCTCC	dcr-1_oligo56	ATCGGACATTGGAGCGTCA
dcr-1_oligo14	TTGTGCTACCGGGTATTTCG	dcr-1_oligo57	CCGATAGGCTGCAACTGA
dcr-1_oligo15	CGACGCAATCGGCAATAGA	dcr-1_oligo58	CTAACTCACCGATCCTGTG

dcr-1_oligo59	GTCATGCTTCAATGGCGAG	grip163_oligo7	TGAAATCGTCTCCGAGCTG
dcr-1_oligo60	TGAAACAGGGTCACTCCAG	grip163_oligo8	GGACAGGAAGTTGAGCGTA
drosha_oligo1	TTGCACTTCGGAGAAGAGG	grip163_oligo9	CGGGATGGTGGTACAATTG
drosha_oligo2	TTGGGAGCAGTGCTTTCTGA	grip163_oligo10	GGTAAGGTTCTCCGATAACC
drosha_oligo3	TCTTCGTCTCCGGCAAC	grip163_oligo11	CGTTTCATCCGCCGAACAA
drosha_oligo4	CGCTGGAAATGTTCTCGTG	grip163_oligo12	CGCCAGCGCATTGGTTAAA
drosha_oligo5	TGTCCAATTCGTTGCCCG	grip163_oligo13	CCCGCCCTTATTCTTCCTA
drosha_oligo6	CAGGTCCCTTGAATTCTCC	grip163_oligo14	TCGTTTCGTGAGTGATGCTC
drosha_oligo7	TGCAGCTTCGACTTGGA	grip163_oligo15	AATTGACCGTCCACCAGGA
drosha_oligo8	GGCACGATCTTTGGTCTTC	grip163_oligo16	GTAGTTTCTGAGCTTGCGG
drosha_oligo9	ATCAGCCTTGGGATTGGCA	grip163_oligo17	ATGCTCGTAGATACGCTGC
drosha_oligo10	AATCATCGCAAACCACCGC	grip163_oligo18	TTTGGAACCTCCGGAGACA
drosha_oligo11	GGTACTCCGAGCAAATCAG	grip163_oligo19	TCGACGTGAAGCTTGTCTGA
drosha_oligo12	ACTCAGCAGTTTCTCGTCC	grip163_oligo20	GTCGATTTCTGCCACGTTT
drosha_oligo13	ACCTCATCTGCCACCTTGA	grip163_oligo21	AGTTCGGTCAGCTCTTTTCG
drosha_oligo14	CCACCATCCAATAGAAGCG	grip163_oligo22	CCGGGCAATTGAGATTAGC
drosha_oligo15	AGCCAAGTGCTGGTTTTGG	grip163_oligo23	CTGGCGGTTTCGATTGAAC
drosha_oligo16	AATTGCCGCTCGGTAAGTG	grip163_oligo24	TGCTTGGGTTGTTGGATGC
drosha_oligo17	CAGATGGATTGAGCTGAGG	grip163_oligo25	GCTTATCCTCCATCCGTTT
drosha_oligo18	CTACTACAGCATCACCGAG	grip163_oligo26	CCCGCTGAAATGCAAGTTC
drosha_oligo19	TGTGTTGATACCCCGCTTG	grip163_oligo27	TAACAATGTCCTGACGCCG
drosha_oligo20	TGATCCGGATTGGTACCG	grip163_oligo28	TTGGAGTTCACGCTGCTTC
drosha_oligo21	CTCCTTGTATGAGGGATGC	grip163_oligo29	ATGCGGTCTCGAAAAATGC
drosha_oligo22	TACTCTCCTCTAGCAGGTC	grip163_oligo30	GTTGCTCCGTTAGGCAAAC
drosha_oligo23	GCGACTTGTGAAAGCGAAG	grip163_oligo31	CAATATCCGGGTACTGCAC
drosha_oligo24	ATTTTCGTTTCATGCGCCCC	grip163_oligo32	ATATGCCCTTAGCACCAGG
drosha_oligo25	GTTTCGCATCTCCTGAAGAC	grip163_oligo33	GACGGTGAAGTGTTTCAGAC
drosha_oligo26	TGCTCCTTCTCCTCTAGC	grip163_oligo34	CGACTCGTCTAGTTCACCA
drosha_oligo27	TCGACATGTTGGCCATCAG	grip163_oligo35	ATTACGGGATATGCAGCGC
drosha_oligo28	TTCCAGCGTAGCTTAGCTG	grip163_oligo36	GTCTTTGTTTCCGCAGGTG
drosha_oligo29	GCTGGATGTGAGATTCCAG	grip163_oligo37	CAGGAAGGAGCCAGTTATC
drosha_oligo30	TTGCCACTCGCTCTGACTA	grip163_oligo38	ATCGGCTTGCAAAGAGAGC
drosha_oligo31	TATCCAGCAAGTAGCGGAG	grip163_oligo39	AAACTGCCGAGTGGTGCT
drosha_oligo32	GAAAAGCTGGGCAGGATTC	grip163_oligo40	GTCCATAATCGTCTCTCGC
drosha_oligo33	ATGTCTAGCTCGTGGATGG	grip163_oligo41	GGCTCGATTTAGAGGCCT
drosha_oligo34	GTTCTCTGGCATCTTCTCC	grip163_oligo42	GTTCCATTTGGTGC GGAA
drosha_oligo35	GTTGAAGCGGATCACCTTG	grip163_oligo43	CTCCAACTTATCTGCCGAC
grip163_oligo1	CCTGCTGCAGACGTTTGAA	grip163_oligo44	GGTGATCGATTAGGGTTGC
grip163_oligo2	CCCGCTTTAAGTAGGCAAC	grip163_oligo45	CCGAAACGAAGTGCTTCTC
grip163_oligo3	TGGAATCGGCTGTGCACAA	shtd_oligo1	GCTCAAGTCCACTCCATTC
grip163_oligo4	CCTCCTTGAAGCTCTTCCA	shtd_oligo2	GTGGTCACATAGCTGGGA
grip163_oligo5	CATCTTGTGGCGCATCATC	shtd_oligo3	TTGGATAAGCTCCACGGAG
grip163_oligo6	GCCCATTGCTGCAGATAC	shtd_oligo4	GGCATCGATGAGCTTCATC

shtd_oligo5	GGACTTGGTTACGTGATCG	shtd_oligo48	GTGCACCTTGGATATGAGC
shtd_oligo6	TAGGGTTAGGTAAACCGGC	shtd_oligo49	GGTCTGACCCACTAAATCC
shtd_oligo7	CAGCCCCTTCGCTTCTTAA	shtd_oligo50	GTGGATAAATGCTCGGGTG
shtd_oligo8	CAATCGGTGCACTCATCTC	shtd_oligo51	CATCTCGCAGAACTCTCG
shtd_oligo9	CACTGGCCAATAGTTCTCG	shtd_oligo52	GTTCCGTAGATGTTCTCCG
shtd_oligo10	AGAGATGTCTCAGCGCTTG	shtd_oligo53	ATCCTTCACGGACATGGAC
shtd_oligo11	ACTCCGGTGTCTGGGAAA	shtd_oligo54	CTGGCTAATACTGCTCTGC
shtd_oligo12	AGATGAGCAGCACCATCAG	shtd_oligo55	TTGGCTCAATGGAGTGCC
shtd_oligo13	TGGCAAAGGCCACCAGAT	synj_oligo1	TCGTAGGCGGTATCAGATG
shtd_oligo14	AATGCAGAAGGCAGCTCCA	synj_oligo2	GCGAAGTCTTTCCCGAAGA
shtd_oligo15	TCCTCGGGTGCATACTCA	synj_oligo3	ACGTCGAAGAACTAGCGGA
shtd_oligo16	CGCACCATGTCCAGTAGA	synj_oligo4	CAAGTTCGGTTGGCGATTG
shtd_oligo17	TGTTGACGGTATCGCCTTC	synj_oligo5	ATTTGGAGGATGGCGTGG
shtd_oligo18	GAAGTTGGAACGAGGCTAG	synj_oligo6	AGGTTGAGATTGCGGTTGC
shtd_oligo19	CCGACCATGTAGTAGTGCA	synj_oligo7	CTACAGGATTAGCGCACAG
shtd_oligo20	AGCAGACCAACGTTGGTCA	synj_oligo8	GTATGGTGTTCTGTTGGTCA
shtd_oligo21	GACACTCATGAAGGACAGC	synj_oligo9	GATGCTGCCAGTCCCTTTGA
shtd_oligo22	CCAGGGCATTATGCGAATG	synj_oligo10	AAAGCCGATTCTCCATCCC
shtd_oligo23	TGGAGTCTATGTCCTGAGC	synj_oligo11	ATGTGACCCACATGGTGTC
shtd_oligo24	GGCACAATTTGGGCATGG	synj_oligo12	CCACATAGCGCACCAAAG
shtd_oligo25	GTGAACATACTCTGCCCA	synj_oligo13	CCTCTCCCATCTTGGATAG
shtd_oligo26	CGCAGACGCAGTAGTTTTG	synj_oligo14	CGGTAATCAGGGCTGACA
shtd_oligo27	CCATGTTGTCCATGTCTGTC	synj_oligo15	CATCCTCATCTCCAGTTGC
shtd_oligo28	AGTTCATAGGTGGCCATGC	synj_oligo16	AGACTCCAGGACATGTACC
shtd_oligo29	AATGTCTCGCCGTGTCAAG	synj_oligo17	CTGCTCGAATACAGCACG
shtd_oligo30	CATGGCCAGAAGTAGCTG	synj_oligo18	CGTAGTCGTCGCTGAACAA
shtd_oligo31	ATCCGCCGTGATATTACGC	synj_oligo19	CATACTGTACGTCGGGTC
shtd_oligo32	TGAGTCACCAGCGATACGA	synj_oligo20	AAGTGATCTCTCCCTCGAG
shtd_oligo33	AACGCACTCCAAGAAGGTG	synj_oligo21	TCGCCAAACACATTGCCAG
shtd_oligo34	ATGTGCTCGAGTTGGGCAA	synj_oligo22	TCCTGCTCCTTGCGCAAT
shtd_oligo35	TCCTGGTGCAGCATCATAG	synj_oligo23	TTCGTCTTCTCCATGTCTG
shtd_oligo36	GATGTTACGCTGTAGCTG	synj_oligo24	CCAAAACACCCAGTCGTG
shtd_oligo37	GGAAGTCAACTAGAACCG	synj_oligo25	TTTTAGCGTCTGCCCATC
shtd_oligo38	ACTCGATTCCGTCGCAGTA	synj_oligo26	TGCTGGTATGTGAGCAGCA
shtd_oligo39	CAGCCATTCTGCTCTATG	synj_oligo27	TAGTCATTGTCCCGCGAGA
shtd_oligo40	TCCTGTCTGCTATACCAACG	synj_oligo28	CACAAGTGGCATTGTCCGG
shtd_oligo41	ACAAAGTCCAGGAACTGGG	synj_oligo29	ACACAATGCTGCGGAAGTG
shtd_oligo42	GGACTTAGAACCTGCCGTA	synj_oligo30	TGCCGCCGTTGACGTTATA
shtd_oligo43	GTAGCCACACATCTGGTGA	synj_oligo31	CATATTGGAGGGCAGTAGG
shtd_oligo44	TTCTGAGCATTCTGCCCG	synj_oligo32	GGCTTCTGCTTTGAGTTG
shtd_oligo45	GGATCTCTCAGGGATTTGC	synj_oligo33	TCTGTCCGGAAGATCTCCTC
shtd_oligo46	TCCACGAAAGAAGTCTGTC	synj_oligo34	CCGCGAAATATTCTGCTGC
shtd_oligo47	AGCAGCTGCCATTGGTGT	synj_oligo35	TAGGCAATTCGTGCGCACA

eb1 oligo1	AATACTCCTCGTCCTCTGG	osk oligo1	GATCCATCAGCGTAAATCG
eb1 oligo2	CCTTTTCCAATCCCTCCAG	osk oligo2	CCAACCTAATACTCCAGACTCG
eb1 oligo3	TGCGCATATCCATCACCTG	osk oligo3	CCAGAACAGATAGGGTTCC
eb1 oligo4	GTGCAGAGCTCCTCGATT	osk oligo4	TCGTTGATTAGACAGGAGTG
eb1 oligo5	CTGCCATCGTAATTGGCATC	osk oligo5	ACAATAGTTGCCAGCGG
eb1 oligo6	GAACTGACAGTAAGCTGCAC	osk oligo6	TTTGTTAGAATCGGCACCAA
eb1 oligo7	GTGGAGTAGACGTTTACAGC	osk oligo7	GCATATTGTGCATCTCCTTGA
eb1 oligo8	CTTTTGTATGATCGGATGCGC	osk oligo8	CTCGATCTGAACCAAAGGC
eb1 oligo9	CTCGAAATTGTCTTGAAGCG	osk oligo9	ATAATGTCCACCGATCCGA
eb1 oligo10	CCACAGACATCTTCTGAAGC	osk oligo10	GACGATGATCTGAGTACCC
eb1 oligo11	GAATTGGGAAACAGCATGTCC	osk oligo11	AGTCCGGATACACAAAGTCC
eb1 oligo12	ATCGTCAGGCGGCGCAAAA	osk oligo12	CATTCCGGGCGAGATATAGCA
eb1 oligo13	CGTCATCGGCTTCTTGGA	osk oligo13	CATCGCCATAAGCGGAAAG
eb1 oligo14	CGTTCTTCTTGACCGTGCC	osk oligo14	AGATAGGCATCGTAATCCGAG
eb1 oligo15	TGCTGTTGGCACAGGCGT	osk oligo15	TCGTCAGCAGAGAATCGTTG
eb1 oligo16	CGGCAGCACCTTGATAC	osk oligo16	GTCATTTCTGTGGCGTCTCT
eb1 oligo17	GGCTTCACTGCAGACGTCA	osk oligo17	GCTTTGGGTCTGACAGCT
eb1 oligo18	GTGCCGGGCAGTGACTTTA	osk oligo18	GAGCCAAATTGATTGGTTCCTC
eb1 oligo19	TCCCGATCCGAAGCCCAT	osk oligo19	GCTGTAGATGTTGATGGG
eb1 oligo20	ACCGCGCTGGCATCGTAA	osk oligo20	GCATTTACGCTGGCTTGC
eb1 oligo21	GAGAATTGCGACTGGAGGC	osk oligo21	AATTATCCTGGTAGCACCAG
eb1 oligo22	GGCGCGAGAGATTCTCTGA	osk oligo22	GTTTGAAGGGATTCTTCCAG
eb1 oligo23	CCTCAGTCGCATATAAGATGTC	osk oligo23	AGGTGCTCGTGGTATGTTG
eb1 oligo24	GATTGTTGACCGAATTGCTCAC	osk oligo24	TAGTCGCTGGTGCCTCT
eb1 oligo25	GAAGTTCTGTATGTAATCGTGC	osk oligo25	AGCACCATATCCAGGAGG
eb1 oligo26	CGTTAACCCAAGCTAGCATATC	osk oligo26	CGTTCTTCAGGCTCGCTT
eb1 oligo27	CGAAATTTGACACGCTTTACTGGC	osk oligo27	AAGATCCGCTTACCGGAC
eb1 oligo28	TTGCCGATGGGCCACGTCGATA	osk oligo28	CTGCACTCAGCGGTCACA
eb1 oligo29	CCCGCAACTTAGAGAAGTAAAAGTCTCG	osk oligo29	GGAATGGTCAGCAGGAAA
eb1 oligo30	TTCTGCTCGCTGGGGCTGCGTTGT	osk oligo30	CGTCACGTTGTCTGTCAG
		osk oligo31	AAATGGATTGCCCGTCAG
		osk oligo32	CTTGATGCTCGATATCGTGA
		osk oligo33	TGGGCGTGGCTCAGCAATA
		osk oligo34	CGCGCACCTCACTATCTA
		osk oligo35	ATATTCCTCGCGCACGGA
		osk oligo36	ATAGTTGCTCTCGATGATGG
		osk oligo37	TGTTCTCGCTGGTGTGTC
		osk oligo38	GTTGTAGGTGATTTCCCTTGG
		osk oligo39	TCTGAGTGGACGAGAAGAG
		osk oligo40	GCTACGACTTGCAACTGC
		osk oligo41	GAGTTCATGGGCCACCAA
		osk oligo42	CTTCCACAACCTCCGGCAA