mRNA localization in metazoans: A structural perspective

Daniela Lazzaretti and Fulvia Bono

Max Planck Institute for Developmental Biology, Tübingen, Germany

ABSTRACT
Asymmetric localization of mRNAs is a widespread gene regulatory mechanism that is crucial for many cellular processes. The localization of a transcript involves multiple steps and requires several protein factors to mediate transport, anchoring and translational repression of the mRNA. Specific recognition of the localizing transcript is a key step that depends on linear or structured localization signals, which are bound by RNA-binding proteins. Genetic studies have identified many components involved in mRNA localization. However, mechanistic aspects of the pathway are still poorly understood. Here we provide an overview of structural studies that contributed to our understanding of the mechanisms underlying mRNA localization, highlighting open questions and future challenges.

Introduction
The localization of mRNAs is a gene regulatory mechanism, widespread in eukaryotes, that is crucial for many processes, including patterning of embryonic axes, asymmetric cell division, cell migration and synaptic plasticity (reviewed in Holt and Bullock, 2009; Martin and Ephrussi, 2009). Among metazoans, the fruit fly Drosophila melanogaster has been most extensively studied, given the genetic accessibility of this system. During Drosophila oogenesis and early development, hundreds of mRNAs have been shown to localize in very defined subcellular patterns. Among vertebrates, a model system that has been classically used is Xenopus laevis. Cell cultures of mammalian neuronal cells and fibroblasts are also widely used to explore how mRNA localization contributes to neuronal function and cell motility, respectively. Cultured cells-based systems have allowed transcriptome-wide characterization of mRNA localization. For example, several mRNAs have been shown to be localized in dendrites, in axons or in cell protrusions in these systems. Many mRNA-localizing factors initially identified in Drosophila have homologs in vertebrates, including Staufen, Bruno, Bicaudal D (BicD), ZBP1 and Pumilio (Pum). In vertebrates some of the proteins that are involved in mRNA localization in Drosophila are also involved in mRNA localization, for example, mammalian Staufen and Fragile X Mental Retardation Protein (FMRP) are required for mRNA localization in neurons, and the Exon Junction Complex (EJC) component eIF4AIII is involved in dendritic mRNA localization. Genetic studies in Drosophila have provided detailed part-lists of mRNA localization components. However, our mechanistic understanding lags behind the genetic characterization of this pathway.

Here we review the contribution of structural studies to the mechanistic understanding of mRNA localization. We will discuss how target mRNAs are recognized, repressed and transported to their target site within the cell. We will only briefly touch on RNA recognition mediated by canonical RNA-binding domains (such as KH, RRM, dsRBD), for which we refer to more focused reviews. We will also not address the mechanisms of motor function, mRNA localization studies in fungi, nor provide an in-depth description of pathways intersecting with mRNA localization (e.g. mRNA decay), which have been reviewed elsewhere.

Steps in mRNA localization
The key step in mRNA localization is the specific recognition of the transcript, which depends on cis-acting elements, generally found in 3’ untranslated regions (3’UTR) of mRNAs (reviewed in Beisse and Ephrussi, 2008; Meignin and Davis, 2010; Medioni et al., 2011). These elements are recognized by trans-acting factors, or RNA-binding proteins (RBPs). Many RBPs contain several domains, that can contribute to specific target recognition or mediate the recruitment of additional protein factors. In the case of active localization, mRNAs are transported along the cytoskeleton either by dynein or kinesin motor proteins (mediating microtubule minus end- or plus end-directed transport, respectively) or myosins (which move along actin filaments). Alternatively, localized expression can be achieved by selective translational repression or degradation of the target mRNA in parts of the cell. Often, the localized messenger ribonucleoprotein particle (mRNP) is maintained at the destination site by anchoring to the cytoskeleton, generally actin filaments. Translational repression of the mRNP before localization is fundamental for both active and passive

ARTICLE HISTORY
Received 28 March 2017
Revised 15 May 2017
Accepted 31 May 2017

KEYWORDS
Drosophila; gene expression regulation; molecular mechanism; Post-transcriptional regulation; RBP; RNA-protein complexes; mRNA localization; structure; trans-acting factors

CONTACT Fulvia Bono; fulvia.bono@tuebingen.mpg.de

© 2017 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
localization mechanisms, to prevent protein expression at ectopic sites in the cell. Such repression may persist at the destination site (e.g., during Drosophila oogenesis), until it is relieved by an extrinsic cue. This translational activation is currently one of the least understood steps of the mRNA localization pathway.

RNA localization signals
A variety of mRNA localization signals has been identified to date, ranging from linear sequences to defined secondary structures. These signals are critical for providing recognition sites for proteins of the localization machinery.32-34

Recognition of linear signals
One of the best characterized examples of a linear localization signal is the zipcode element on β-actin mRNA, recognized by the Zipcode-Binding Protein 1 (ZBP1).13,35 ZBP1 regulates local translation of β-actin mRNA at the base of activated dendritic spines in hippocampal neurons.36 The crystal structure of human ZBP1 RNA-binding domains (hnRNP K homology, or KH domains) revealed an intramolecular pseudo-dimer arrangement that positions the RNA-binding surfaces on opposing faces of the protein. This induces the looping of the bipartite zipcode element around ZBP1 (Fig. 1, 2a). Binding of both zipcode elements enhances the affinity of ZBP1 for β-actin mRNA. Since the 2 zipcode elements can only bind if they are separated by a spacer sequence of defined length, this ensures specificity in mRNA recognition.37 The occurrence of multiple domains to obtain high RNA-binding affinity and specificity, often combined with the remodelling of the target mRNA (as in the β-actin mRNA and ZBP1), is a recurrent feature of RNA-protein complexes (reviewed in Lunde et al., 2007).31

Another example is provided by the CUG-binding protein 1 (CUGBP1). The human CUGBP1 protein is a member of the CUGBP1 and ETR-like factors (CELF) protein family, which includes the Drosophila translational repressor Bruno (Bru). CELF family members share a conserved domain architecture, characterized by two adjacent RNA recognition motifs (RRMs) in the N-terminal region, and a third C-terminal RRM, which is separated from the others by a long, non-conserved, linker.10 The two N-terminal RRM domains of CUGBP1 bind with similar affinities to their respective target sequences (UGUU/G). NMR studies indicate that the two domains tumble independently in solution in the absence of the RNA substrate.38 However, in the presence of an RNA molecule comprising two recognition signals the two RRM domains are found in a compact and rigid arrangement (Fig. 2b), which explains the observed binding cooperativity.38 While the short linker between RRM1 and RRM2 of CUGBP1 favors cooperative binding of the two domains on adjacent sequences on the target mRNA, the longer, and probably disordered, linker that connects RRM3 could allow for recognition of a more distant sequence element, perhaps belonging to a different mRNA molecule. This mechanism could help in creating a large messenger ribonucleoprotein particle (mRNP) for transport, and would also contribute to maintaining the translational repression of the target mRNA, as has been proposed for Drosophila Bru in oskar (osk) mRNA regulation.39 In addition, such unstructured regions can cooperate in RNA binding: the linker region N-terminal to the third RRM in both human CUGBP1 and Drosophila Bru extends the binding surface of the RNA-recognition motif, which increases RRM3 affinity for its target mRNA.40,41

Another way to achieve recognition of an RNA sequence element with high affinity and specificity is by combining multiple copies of a simple repeated structural motif within a protein domain. This approach is used by members of the Pumilio
and fem-3 mRNA binding factor (Puf) protein family, of which *Drosophila* Pumilio (Pum) is the founder (reviewed in Wickens et al., 2002). Puf proteins are characterized by the presence of a sequence-specific RNA-binding domain known as the Pum homology domain (Pum-HD), comprising 8 sequence repeats of three helices each, plus N- and C-terminal flanking regions. This helical domain is reminiscent of the Armadillo repeat domain, a protein-protein interaction module found in a wide range of proteins with diverse functions, including nucleo-cytoplasmic transport, intracellular signaling, and cytoskeletal organization (reviewed in Hatzfeld et al., 1999; Coates, 2003). The Pum-HD domain contacts RNA through its concave surface, where each nucleotide is recognized by a triumvirate of amino acid side chains at conserved positions within the helical repeats (Fig. 2c). In addition, the Pum-HD domain can interact with the CCR4:NOT deadenylation complex and promote shortening of the poly(A) tail on the target mRNA. In this way, Puf proteins control stability and translation of a variety of different mRNAs, recognized through a Pum response element (PRE) in their 3′ UTRs.

**Recognition of structured signals**

Given the difficulty in predicting tertiary or even secondary structures of RNAs with high reliability, the characterization of structured localization signals has proven more challenging and relies on experimental verification. This is well exemplified by the *Drosophila* *fs(1)K10* Transport and Localization Signal (*K10* TLS), which mediates minus-end-directed, dynein-dependent transport along the microtubules during early *Drosophila* development. The structure of this localization element, derived by NMR spectroscopy, has revealed a stem-loop conformation in which purine-purine stacking within the double stranded (ds) region forces the stem to adopt an unusual A′-form conformation. This conformation has 2 widened major grooves which are oriented at 90° to one another in contrast to the standard A-form dsRNA where major and minor grooves have similar widths (Fig. 3). This specific arrangement (reminiscent of B-form dsDNA) is required for the localization of *K10* TLS-containing mRNAs in vivo. Other mRNAs that are transported at various stages during *Drosophila* oogenesis and early development toward the minus end of microtubules also contain localization signals that could form hairpins of similar structure but of different primary sequence as the *K10* TLS. This suggests that widened major grooves in dsRNA, rather than being used for primary sequence recognition, represent an unusual structural feature that can be sensed by a common localization machinery. Indeed, several of these elements, including the *K10* TLS, are recognized by the protein Egalitarian (Egl). Mutation of the putative catalytic residues within Egl Exo-domain does not affect mRNA localization in vivo or RNA-binding activity in vitro indicating that, if the protein is a functional exonuclease, this activity is dispensable for protein function. Similarly, the protein Exuperantia (Exu) uses a catalytically inactive Exo-domain to bind and localize *bicoid* (*bcd*) mRNA during early oogenesis in *Drosophila*. Both Egl and Exu recognize mRNA signals with defined secondary structures, though the molecular details of the interaction are not known.

To date, there are very few examples of structured localization signals in complex with RNA-binding proteins. One is provided by the RNA-binding domain of yeast Vts1 in complex
with the Smaug (Smg) Recognition Element (SRE) RNA (Fig. 3c). Vts1 binds RNA through its Sterile α motif (SAM), a domain consisting of 5 α-helices arranged in a globular bundle, that was previously thought to be solely a protein-protein interaction motif. The SRE consists of a stem-loop structure, but only the 5 residues of the loop (CUGGC) are directly contacted by a shallow, positively-charged surface patch on the SAM domain. The only residue that is specifically recognized is the G at position 3 within the pentaloop, while the other protein-RNA interactions involve non-sequence specific contacts with the RNA phosphate oxygens. However, the formation or stabilization of a base pair within the pentaloop upon protein binding also seems to be important, suggesting that the G3 nucleotide base within the loop is recognized within a specific structural context. The sequence of the stem, as long as base-pairing is preserved, does not influence Vts1 binding affinity. Yeast Vts1 is a homolog of Drosophila Smg, a protein that contributes to anterior-posterior axis determination in the Drosophila embryo. Smg represses nanos (nos) mRNA translation everywhere in the embryo but at the posterior pole plasm. This generates a gradient of Nos protein emanating from the posterior of the embryo. Such gradient is required to trigger a series of downstream events that result in proper abdominal segmentation. Smg also binds mRNA by its SAM domain, suggesting evolutionary conservation, but the details of this interaction are not known. A SAM-like domain is also present in Exu and mutational analysis suggests that it is important for RNA binding. However, in Exu there is an additional surface not present in Smg or Vts1 that contributes to mRNA binding.
Another conserved mRNA-localization protein, Staufen (Stau), presents a distinct RNA-binding mechanism. Members of the Stau protein family are important for mRNA localization in many organisms and contain several copies of dsRNA-binding domain (dsRBD), a domain that was first identified in the Drosophila Stau protein.7 In Drosophila, Stau is required for oskar (osk) mRNA localization at the posterior pole during oogenesis, for bcd mRNA anchoring at the anterior pole in late oogenesis and early embryogenesis, and for prospero (pros) mRNA transport during neuroblast asymmetric cell division.60-62 Structural characterization of the third dsRBD of Drosophila Staufen in complex with a non-physiologic stem-loop RNA sequence showed that the dsRBD3 interacts with the RNA sugar-phosphate backbone, without making direct contacts with the RNA bases. Some sequence specificity could be provided by the interaction of the first α-helix with the nucleotide bases in the hairpin loop67,68 (Fig. 3d), although structural studies on a physiologic substrate would be required to confirm this.

For other localized mRNAs, the protein factors required for the recognition of their structured localization signals are still unidentified. This is the case of the Spliced osk Localization Element (SOLE)69,70 (Fig. 3b). The SOLE is generated upon splicing of the first intron of osk and it is required for plus end-directed, posterior localization of osk mRNA during oogenesis in Drosophila.69 Osk localization also requires the deposition of the EJC upstream of the first exon-exon junction. The isolated SOLE assumes a stem-loop structure capped by a 5′ nucleotide loop, and, like the K10 TLS, displays a widened major groove.70 A protein binding partner of the SOLE has yet to be identified.

**Combinatorial regulation mechanisms (The importance of a bigger picture)**

In many of the examples described above, the specificity of the RNA-binding proteins (such as Vts1,26, Egl,50 Stau,23 Exu23) appears to be less stringent in vitro than in vivo.71,72 This could be due to the use of isolated RBDs in in vitro assays. The inclusion of multiple domains, where present, would be important in future studies to better approximate in vivo conditions. In other cases, the binding affinity and specificity of an RNA-binding protein (RBP) for its target mRNA is modulated by the interaction with a protein partner. Such cases highlight the importance of structural studies focusing on multiprotein complexes together with their bound mRNA targets. The effects of partner proteins on mRNA binding could be mediated by partner-dependent stabilization of RBPs, as has been proposed for Drosophila Egl and Bicaudal D (BicD).50,72,73 Alternatively, binding partners could alter RBP binding specificity either by promoting conformational changes or by contributing additional surfaces for binding. A striking example is provided by the complex of Pum and Nanos (Nos). These proteins co-regulate several mRNAs in Drosophila, including repression of maternal hunchback (hb) mRNA at the posterior pole of the embryo, and of Cyclin B (CycB) mRNA in primordial germ cells and germline stem cells.74,75 The crystal structure of the Pum: Nos complex bound to hb and CycB recognition elements revealed that Nos not only induces conformational changes in Pum that increase its affinity for RNA, but also directly contacts the RNA upstream of the Pum canonical recognition sequence76 (Fig. 4a–b). In this way, Nos adds upstream recognition specificity, thereby relaxing the requirement for a perfect Pum-binding consensus sequence. The Pum:Nos complex can thus regulate a broader range of mRNA targets than Pum alone. Cooperativity is also exhibited by the proteins Sex-lethal (Sxl) and Upstream of N-Ras (Unr). In Drosophila females, the two proteins bind adjacent sequences in the 3′UTR of male-specific lethal 2 (msl2) mRNA, which encodes for the limiting component of the dosage compensation complex and represses its translation.77 The crystal structure of the Sxl:Unr:msl2 mRNA ternary complex showed how the 5′ end of the RNA sequence is recognized by Sxl, while the 3′ end is sandwiched by both proteins. The formation of such a “triple zipper” allows additional protein:RNA contacts that would not be possible with either protein alone78 (Fig. 4b–c).

**Transport and anchoring**

Long-range active transport usually depends on microtubules, with kinesin and dynein motors mediating plus end- and minus-end directed movement, respectively. There are very few instances where all biochemical links between the localizing mRNA and the molecular motor have been identified. These include the complex linking ASH1 mRNA to myosin 4 motor during localization to the bud of Saccharomyces cerevisiae (reviewed in Niedner et al., 2014)26 and Drosophila BicD, a protein linking localizing mRNAs to the dynein motor during minus end-directed mRNA localization (reviewed in Hoogenraad and Akhmanova, 2016)79 (Fig. 5a). BicD is the founder of a conserved family of motor adaptor proteins, which, in addition to coupling the dynein motor to various cargoes, stimulates processive dynein motility by stabilizing dynein interaction with its constitutive cofactor dynactin.80-82 While the N-terminal 2 coiled coil domains of BicD associate simultaneously with dynexin and dynactin, stabilizing their interaction,82 the C-terminal part of BicD can recognize different protein partners, that in turn mediate the transport of specific cargoes.83 In Drosophila, one of the most important partners of BicD is Egl, which is required for dynein-dependent localization of a variety of transcripts.50 Another BicD-associated factor is FMRP, which has an important role in mRNA localization in neurons and is mutated in the most common inherited form of cognitive deficiency in humans.84 All existing data point to BicD binding to one partner at a time, and in some cases through overlapping interaction surfaces.85 Though we can draw a link between localizing mRNPs and dynein-mediated transport, little is known about kinesin-mediated transport.

Another key adaptor protein in mRNA localization is Miranda (Mira). Mira is required for the segregation of the cell fate determinants Brain tumor (Brat), Prospero (Pros) and Numb to the basal cortex during the asymmetric cell division of neuroblasts in Drosophila.86,87 In addition, Mira directly interacts with the dsRNA-binding protein Stau to localize pros mRNA in an actin-dependent manner both in neuroblasts and epithelial cells.86,88,89 (Fig. 5b). In this case, the connection with the cytoskeletal motor is still unknown.

The mechanism by which osk mRNA is anchored at the posterior cortex during Drosophila oogenesis and early
embryogenesis is also poorly understood. At the posterior pole, two Osk protein isoforms are translated from alternative start codons: Short Osk is necessary and sufficient to induce pole cell formation and posterior patterning in the embryo, while Long Osk is required for proper anchoring of both osk mRNA and Short Osk to the posterior cortex of the oocyte.92-95 Short Osk recruits the DEAD box helicase Vasa through its N-terminal LOTUS domain to initiate germ plasm assembly.94,96 In addition, short Osk associates with germ plasm localized mRNAs, including its own and nos mRNA.96,97 Analogous to Egl and Exu, Osk binds RNA through an enzyme fold: the C-terminal OSK domain, which resembles a SGNH hydrolase, but lacks the catalytic residues.96,97 The RNA sequence specificity of Osk has not been yet determined. In addition to the LOTUS and OSK domains, Long Osk contains an N-terminal extension, which is required, but not sufficient, for the posterior anchoring of osk mRNA.95 Despite the presence of the LOTUS domain, Long Osk cannot recruit Vasa, suggesting that the N-terminal extension is somehow modulating the function of the other Osk domains.92,94,96 Instead, Long Osk stimulates clathrin-mediated endocytosis and contributes to the organization of the actin cytoskeleton at the posterior of the oocyte, promoting its own and Short Osk maintenance at the posterior pole through a poorly understood mechanism.98

Translational repression and selective degradation

Translational repression mechanisms frequently target translation initiation, which is generally rate-limiting. During initiation, the scaffolding factor eIF4G interacts with the cap-binding protein eIF4E and the Poly(A)-Binding Protein (PABP), to connect 5’ and 3’ end of the mRNA in a closed loop conformation that is thought to stimulate translation; in addition, eIF4G interacts with eIF3, which in turn recruits the small ribosomal subunit (reviewed in Jackson et al., 2010).99 The eIF4E:eIF4G interaction is targeted by a conserved class of translational repressors, the eIF4E binding proteins (4E-BPs) (reviewed in Topisirovic et al., 2011).100 A member of this family is the Drosophila protein Cup, which is required during
oogenesis and early development for repressing translation of the axis-determining factors osk, nos, and gurken (grk). Cup does not contact RNA, but is recruited to its mRNA targets by different RBPs (Bru in the case of osk and grk; Smg for nos). The crystal structure of an N-terminal fragment of Cup in complex with eIF4E revealed how Cup interacts through two binding motifs (one canonical and one non-canonical), and engages the same surfaces of eIF4E responsible for interaction with eIF4G (Fig. 6). In addition, Cup binding stabilizes eIF4E and increases its grip on the target mRNA, thus protecting it from decapping and degradation. Alternatively, translation initiation can be inhibited by preventing the recruitment of the large ribosomal subunit, as described for ZBP1.

Other proteins repress translation of their target mRNA by recruiting the CCR4:NOT deadenylation complex: examples include Cup, Bicaudal C (BicC), Smg, Pum and Nos. For the recruitment of the CCR4:NOT complex, at least some of these proteins rely on short linear motifs embedded in peptide regions of predicted disorder (e.g., Nos). Interestingly, many RNA- and DNA-binding proteins show a significant enrichment in low complexity regions when compared with the entire proteome. In addition to providing quickly-evolving interaction surfaces, these regions have the potential for forming reversible amyloid-like fibers in vitro, in vivo, the same sequences could mediate oligomerization and assembly of large particles that would reinforce translational repression by steric exclusion of the ribosomes, as proposed for osk mRNA.

Indeed, to ensure efficient repression of the localizing mRNA, translation is usually inhibited at multiple steps. Complete repression of nos expression, for instance, requires both Cup-mediated inhibition at the initiation step and Smg-mediated deadenylation.

**Translational activation**

How mRNA recognition factors, translational repressors and localization machinery could dissociate from the localized mRNP is a critical but poorly understood step in the mRNA localization pathway. Release could be mediated by targeting the localizing factor for degradation when it is no longer needed, as has been shown for Mira. Another mechanism involves dissociation of the RBP from the recognition element on the mRNA. For example, the majority of nos mRNA is...
evenly distributed throughout the *Drosophila* embryo and translationally repressed in a Smg-dependent manner.\textsuperscript{104,109,114} However, at the posterior pole, Osk protein prevents Smg binding to *nos* mRNA, thereby allowing its translation.\textsuperscript{114-116} The molecular details of this regulation are still unclear. Osk could bind Smg directly, and perhaps alter its affinity for RNA,\textsuperscript{115} alternatively, Osk and Smg could compete for the same binding site on *nos* mRNA, as suggested by the discovery that Osk is also an RNA-binding protein.\textsuperscript{96,97}

Post-translational modifications, such as phosphorylation, also seem to play an important role. The affinity of 4E-BPs for eIF4E, for example, is decreased by the phosphorylation of 4E-BPs at multiple sites.\textsuperscript{117,118} Analogously, Src-dependent phosphorylation of ZBP1 relieves translational repression of *β-actin* mRNA by disrupting ZBP1:RNA interaction,\textsuperscript{36} though the molecular details of this regulation are still unclear.

**Future perspectives**

Structural studies have greatly advanced our understanding of the mechanisms underlying mRNA localization. However, many issues remain. The challenge for the future is to understand the assembly and dynamics of the large multiprotein-RNA complexes involved in mRNA localization.

First, structural studies will have to tackle larger multimolecular assemblies. RBPs often contain multiple domains that combine protein- and RNA-binding activities, and, due to their complexity and flexibility, offer a challenge for structural studies. Moreover, many localizing mRNAs assemble in multimeric complexes, through both protein- (e.g., Exu, Bru, Osk, BicD, Mira) and RNA-mediated dimerization (e.g., *bcd*). The increasing number of examples in which combinations of structural motifs, protein domains and protein interaction partners are exploited to modulate RNA-binding affinity and specificity, however, prompts for an effort in the characterization of more physiologic complexes. It will be challenging to characterize stable assemblies and to purify them in suitable amounts for structural studies. Cryo-EM approaches hold promise to solve large assemblies at high resolution and with the need of less material. However, inherent flexibility and heterogeneous composition of the targets remains an issue across techniques when tackling large assemblies.

Localizing mRNP composition changes at the different steps of gene expression. Studies of single molecule dynamics are particularly promising to gain mechanistic insights of mRNA localization at high resolution. These studies provide a direct readout of the spatial and temporal details of localization. It will be particularly challenging to integrate information from these kinds of studies with the static information derived from structural studies. In this respect, structural snapshots of the same component/complex in different cellular states are very informative although technically very challenging.

A third avenue will be to integrate the wealth of genetic information with structural and single molecule studies. Genetic studies, especially in *Drosophila*, provide much data on the identity and hierarchy of the factors required for mRNA localization, and represent an excellent starting point for a more detailed biochemical characterization. Furthermore, novel genome editing techniques will facilitate the *in vivo* validation of functional hypotheses from structural data in reverse genetic approaches.

It will also be essential to chart the complement of RBPs and RNA target motives by large-scale studies. New methods of mRNA-protein interaction\textsuperscript{119,120} including CRAC,\textsuperscript{121} CLIP,\textsuperscript{122} PAR-CLIP,\textsuperscript{123} iCLIP,\textsuperscript{124} hiCLIP\textsuperscript{125} provide a rich source of RBPs and RNA targets. Those will be the starting points for structural studies of complexes to understand RNA binding modes.

The variety of canonical and non-canonical RNA-binding domains used by RBPs, together with the divergence in RNA signal sequences, makes it difficult to derive rules for target mRNA selection. Bioinformatics predictions and studies on isolated RNA elements can provide valuable information. However, experimental validation of the protein:RNA interaction remains essential; this is especially true for cases in which the RNA secondary structure changes following post-transcriptional modifications of the nucleotide bases (e.g., N\textsuperscript{6}-methyladenosine-dependent RNA structural switches\textsuperscript{126}), or interaction with the protein partner (e.g., structural rearrangements in the E3 localization element from yeast *ASH1* mRNA bound to the localization machinery\textsuperscript{127}).

**Disclosure of potential conflicts of interest**

No potential conflicts of interest were disclosed.

**Acknowledgments**

We would like to thank Atlanta Cook and Gáspár Jékel for critical reading of the manuscript. The authors are funded by the Max Planck Gesellschaft, the European Research Council under the European Union’s Seventh Framework Program (FP7/2007-2013), ERC grant agreement no. 310957 and the Deutsche Forschungsgemeinschaft (BO3588/2-1 to F.B.).

We apologize to those investigators whose work was inadvertently overlooked or omitted for space constraints.

**ORCID**

Fulvia Bono http://orcid.org/0000-0003-0563-1840

**References**

1. Holt CE, Bullock SL. Subcellular mRNA Localization in Animal Cells and Why It Matters. Science 2009; 326:1212-6; PMID:19965463; https://doi.org/10.1126/science.1176488
2. Martin KC, Ephrussi A. mRNA localization: Gene expression in the spatial dimension. Cell 2009; 136:719-30; PMID:19239891; https://doi.org/10.1016/j.cell.2009.01.044
3. Lécuyer E, Yoshida H, Parthasarathy N, Alm C, Babak T, Cerovina T, Hughes TR, Tomancak P, Krause HM. Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. Cell 2007; 131:174-87; PMID:17923096; https://doi.org/10.1016/j.cell.2007.08.003
4. Jambor H, Surendranath V, Kalinka AT, Meijstrik P, Saalfeld S, Tomancak P. Systematic imaging reveals features and changing localization of mRNAs in Drosophila development. eLife 2015; 4:e05003; PMID:25838129; https://doi.org/10.7554/eLife.05003
5. Wilk R, Hu J, Blotsky D, Krause HM. Diverse and pervasive subcellular distributions for both coding and long noncoding RNAs. Genes Dev 2016; 30:594-609; PMID:26944682; https://doi.org/10.1101/gad.276931.115
6. Medioni C, Mowry K, Besse F. Principles and roles of mRNA localization in animal development. Development 2012; 139:3263-76; PMID:22921410; https://doi.org/10.1242/dev.078626

7. St Johnston D, Brown NH, Gail JG, Jantsch M. A conserved double-stranded RNA-binding domain. Proc Natl Acad Sci U S A 1992; 89:10979-83; PMID:1438302; https://doi.org/10.1073/pnas.89.22.10979

8. Suzuki H, Maegawa S, Nishibu T, Sugiyama T, Yasuda K, Inoue K. Flamingo in Drosophila melanogaster: Role in asymmetric cell division. J Cell Sci 2007; 120:3129-39; PMID:17931335; https://doi.org/10.1242/jcs.037008

9. Marión RM, Fortes P, Beloso A, Dotti C, Ortin J. A human sequence homologue of Staufen is an RNA-binding protein that is associated with polysomes and localizes to the rough endoplasmic reticulum. Mol Cell Biol 1999; 19:2221-9; PMID:10022908; https://doi.org/10.1128/MCB.19.3.2220

10. Good PJ, Chen Q, Warner SJ, Herring DC. A family of human RNA–protein binding domains. Mol Cell 1999; 2:719-21; PMID:10988972; https://doi.org/10.1016/S0960-9822(99)80007-0

11. Suzuki H, Maegawa S, Nishibu T, Sugiyama T, Yasuda K, Inoue K. Flamingo in Drosophila melanogaster: Role in asymmetric cell division. J Cell Sci 2007; 120:3129-39; PMID:17931335; https://doi.org/10.1242/jcs.037008

12. Deshler JO, Hightett MI, Schnapp BJ. Localization of Xenopus Vg1 mRNA by Vera protein and the endoplasmic reticulum. Science 1997; 276:1128-31; PMID:9488009; https://doi.org/10.1126/science.276.5315.1128

13. Boylan KLM, Mische S, Li M, Marqués G, Morin X, Chia W, Hays TS. Motility screen identifies Drosophila IGF-II mRNA-binding protein–zipcode-binding protein acting in oogenesis and synaptogenesis. Proc Natl Acad Sci U S A 2008; 105:18281-112; PMID:18371317; https://doi.org/10.1073/pnas.1000403

14. Wiens M, Bernstein DS, Kimberle J, Parker R. A PUF family portrait: A PUF family portrait: A family of human RNA–protein binding proteins related to the Drosophila Bruno translational regulator. J Biol Chem 2000; 275:28583-92; PMID:10893231; https://doi.org/10.1074/jbc.M003083200

15. Wickham L, Ducha İ, Nabi IR, DesGroseillers L. Mam-malian staufen is a double-stranded RNA- and tubulin-binding protein which localizes to the rough endoplasmic reticulum. Mol Cell Biol 1999; 19:2212-9; PMID:10022908; https://doi.org/10.1128/MCB.19.3.2212

16. Boylan KLM, Mische S, Li M, Marqués G, Morin X, Chia W, Hays TS. Motility screen identifies Drosophila IGF-II mRNA-binding protein–zipcode-binding protein acting in oogenesis and synaptogenesis. Proc Natl Acad Sci U S A 2008; 105:18281-112; PMID:18371317; https://doi.org/10.1073/pnas.1000403

17. Singer RH. RNA zipcodes for cytoplasmic addresses. Curr Bio 1993; 3:719-21; PMID:15335871; https://doi.org/10.1016/0960-9822(93)90079-4

18. Bullock SL, Ringel I, Ish-Horowicz D, Lukavsky P, A’-form RNA helices are required for cytoplasmic mRNA transport in Drosophila. Nat Publ Group 2010; 17:703-9; PMID:20473315; https://doi.org/10.1038/nrm.1823; https://doi.org/10.1038/nrm.1823

19. Dictenberg JB, Swanger SA, Antar LN, Singer RH, Bassell GJ, Condeelis J, Singer RH. Spatial regulation of beta-actin translation by Src-dependent phosphorylation of ZBP1. Nature 2005; 438:512-5; PMID:16306994; https://doi.org/10.1038/nature04115

20. Chao IA, Patskovsky V, Patel V, Levy M, Almo SC, Singer RH. ZBP1 recognition of beta-actin zipcode induces RNA looping. Genes Dev 2010; 24:148-58; PMID:20080952; https://doi.org/10.1101/gad.1862910

21. Teplov P, Song J, Gaw HY, Teplov A, Patel DJ, Levy M. Structural insights into RNA recognition by the alternate-splicing regulator CUG-Binding protein 1. Structure 2010; 18:1364-77; PMID:20947024; https://doi.org/10.1016/j.str.2010.06.018

22. Chekulaeva M, Hentze MW, Ephrussi A. Bruno acts as a dual repressor of oskar translation, promoting mRNA oligomerization and formation of silencing particles. Cell 2006; 124:521-33; PMID:16469 699; https://doi.org/10.1016/j.cell.2006.01.031

23. Lyon AM, Revel BS, MacDonald PM, Hoffman DW. Bruno protein contains an expanded RNA recognition motif. Biochemistry 2009; 48:12202-12; PMID:19919093; https://doi.org/10.1021/bi900624j

24. Tsuda K, Kuschakowski K, Takashita M, Someya T, Inoue M, Terada T, Kobayashi N, Shirouzu M, Kigawa T, Tanaka A, et al. Structural basis for the sequence-specific RNA-recognition mechanism of human CUG-BP1 RRM3. Nucleic Acids Res 2009; 37:5151-66; PMID:19553194; https://doi.org/10.1093/nar/gkp546

25. Edwards TA, Pyle SE, Wharton RP, Aggarwal AK. Structure of Pumilio reveals similarity between RNA and peptide binding motifs. Cell
62. St Johnston D, Driever W, Berleth T, Richstein S, Nüsslein-Volhard C. oskar mRNA is localized to the posterior pole of the Drosophila oocyte. Cell 1991; 66:33-5; PMID:2070417; https://doi.org/10.1002/0922-8674(91)00318-X

63. Wang X, Zamore PD, Hall TM. Crystal structure of a Pumilio homology domain. Mol Cell 2001; 7:835-65; PMID:11336708; https://doi.org/10.1016/S1097-2765(01)00229-5

64. Hatfield M. The armadillo family of structural proteins. Int Rev Cytol 1999; 186:179-224; PMID:9773030; https://doi.org/10.1016/S0074-7696(08)61054-2

65. Kim-Ha J, Smith JL, Macdonald PM. oskar mRNA is localized to the germ plasm and directs localization of the posterior determinant nanos. Nature 1994; 369:315-8; PMID:7514276; https://doi.org/10.1038/369315a0

66. Schuldt AJ, Adams JH, Davidson CM, Micklem DR, Haseloff J, St Johnston D, Brand AH. Miranda mediates asymmetric protein and RNA localization in the developing nervous system. Genes Dev 1998; 12:1847-57; PMID:9637686; https://doi.org/10.1101/gad.12.12.1847

67. Ramos A, Bayer P, Varani G. Determination of the structure of the RNA complex of a double-stranded RNA-binding domain from Drosophila Staufen protein. Biopolymers 1999; 52:181-96; PMID:11295750; https://doi.org/10.1002/1097-0282(1999)52:4%3e181:AID-BIP1003%3e3.0.CO;2-5

68. Nakwowosky E, Harris ME. Specification and nonspecificity in RNA-protein interactions. Nat Rev Mol Cell Biol 2016; 15:334-44; PMID:26285679; https://doi.org/10.1038/nrm4032

69. Weidmann CA, Qiu C, Arvola RM, Lou TF, Killingsworth J, Campbell CJ, Frenette S, Bycroft M, St Johnston D, Varani G. RNA recognition by the Staufen double-stranded RNA-binding domain. EMBO J 2000; 19:997-1009; PMID:10698941; https://doi.org/10.1093/emboj/19.5.997

70. Ghosh S, Marchand V, Gaspar I, Ephrussi A. Control of RNP motility and localization by a splicing-dependent structure in oskar mRNA. Nat Publ Group 2012; 19:441-9; PMID:22426546; https://doi.org/10.1038/nm.2257

71. Simon B, Masiewicz P, Ephrussi A, Carlomagno T. The structure of an RNA recognition motif in Staufen. EMBO Rep 2003; 4:223-7; PMID:12619324; https://doi.org/10.1038/sj.embor.7400153

72. Weidenma CA, Qiu C, Arvola RM, Lou TF, Killingsworth J, Campbell CJ, Frenette S, Bycroft M, St Johnston D, Varani G. RNA recognition by the Staufen double-stranded RNA-binding domain. EMBO J 2000; 19:997-1009; PMID:10698941; https://doi.org/10.1093/emboj/19.5.997

73. Vazquez-Pianzola P, Schaller B, Colombo M, Beuchle D, Neuenschwander S. Structural basis for the assembly of the BicD/Egl transport machinery. RNA Biology 2017; 14:73-89; PMID:27801632; https://doi.org/10.1080/15476286.2016.1251542

74. Arvola RM, Cameron MD, Lou TF, Killingsworth J, Campbell CJ, Frenette S, Bycroft M, St Johnston D, Varani G. RNA recognition by the Staufen double-stranded RNA-binding domain. EMBO J 2000; 19:997-1009; PMID:10698941; https://doi.org/10.1093/emboj/19.5.997

75. Weidenma CA, Qiu C, Arvola RM, Lou TF, Killingsworth J, Campbell CJ, Frenette S, Bycroft M, St Johnston D, Varani G. RNA recognition by the Staufen double-stranded RNA-binding domain. EMBO J 2000; 19:997-1009; PMID:10698941; https://doi.org/10.1093/emboj/19.5.997

76. Weidenma CA, Qiu C, Arvola RM, Lou TF, Killingsworth J, Campbell CJ, Frenette S, Bycroft M, St Johnston D, Varani G. RNA recognition by the Staufen double-stranded RNA-binding domain. EMBO J 2000; 19:997-1009; PMID:10698941; https://doi.org/10.1093/emboj/19.5.997

77. Graindorge A, Militti C, Gebauer F. Functional analysis of X-chromosome dosage compensation. Wiley Interdiscip Rev 2011; 3:534-45; PMID:21957042; https://doi.org/10.1002/wirn.75

78. Hennig J, Milioti C, Popowicz GM, Wang I, Sonntag M, Geerlof A, Gabel F, Gebauer F, Sattler M. Structural basis for the assembly of the Sxl-Urn translation regulatory complex. Nature 2014; 515:287-90; PMID:25209665; https://doi.org/10.1038/nature13693
79. Hoogenraad CC, Akhmanova A. Bicaudal D family of motor adaptors: Linking dynein motility to cargo binding. Trends Biol 2016; 26:327-40; PMID:26822037; https://doi.org/10.1016/j.tcb.2016.01.001
80. Schlager MA, Serra-Marques A, Grigoriev I, Gmyt LF, Esteves da Silva M, Wulf PS, Akhmanova A, Hoogenraad CC. Bicaudal family adaptor proteins control the velocity of Dynein-based movements. Cep Ecol Rep 2014; 1:8124-56; PMID:25176647; https://doi.org/10.1101/gad.212381.112
81. McKenney RJ, Holding AN, Johnson CM, Stephens E, Lukavsky PJ, Walshaw J, Bullock SL. Bicaudal-D uses a parallel, homodimeric coiled coil with heterotypic registry to coordinate recruitment of cargos to dynein. Genes Dev 2013; 27:1223-46; PMID:23723415; https://doi.org/10.1101/gad.212381.112
82. Bianco A, Dienstbier M, Salter HK, Gatto G, Bullock SL. Bicaudal-D regulates fragile X mental retardation protein levels, motility, and function during neuronal morphogenesis. Curr Biol 2010; 20:1487-92; PMID:20691595; https://doi.org/10.1016/j.cub.2010.07.016
83. Ikehisa-Kataoka H, Sheath JB, Nabeshima Y, Doe CQ, Matsuzaki F. Miranda directs Prospero to a daughter cell during Drosophila asymmetric divisions. Nature 1997; 390:625-9; PMID:9403694; https://doi.org/10.1038/37641
84. Nelson MR, Leidal AM, Smibert CA. Drosophila Cup is an eIF4E-binding protein required for both the translational repression of oskar and the recruitment of Barentsz. J Cell Biol 2003; 163:1197-204; PMID:14691132; https://doi.org/10.1083/jcb.200309088
85. Zappavigna V, Piccioni F, Villaescusa JC, Verrotti AC. Cup is a nucleocytoplasmic shuttling protein that interacts with the eukaryotic translation initiation factor 4E to modulate Drosophila ovasp development. Proc Natl Acad Sci U S A 2004; 101:14800-5; PMID:15465908; https://doi.org/10.1073/pnas.0406451101
86. Nelson MR, Leidal AM, Smibert CA. Drosophila Cup is an eIF4E binding protein that functions in Smaug-mediated translational repression. EMBO J 2004; 23:150-9; PMID:14685270; https://doi.org/10.1093/sj/ejbo.7600026
87. Markussen FH, Michon AM, Breitwieser W, Ephrussi A. Translational control of oskar generates short OSK, the isoform that induces asymmetric divisions. Nature 1998; 393:4089-98; PMID:9735369.
113. Kato M, Han TW, Xie S, Shi K, Du X, Wu LC, Mirzaei H, Goldsmith EJ, Longgood J, Pei J, et al. Cell-free formation of RNA granules: Low complexity sequence domains form dynamic fibers within hydrogels. Cell 2012; 149:753-67; PMID:22579281; https://doi.org/10.1016/j.cell.2012.04.017

114. Zaessinger S, Busseau I, Simonelig M. Oskar allows nanos mRNA translation in Drosophila embryos by preventing its deadenylation by Smaug/CCR4. Development 2012; 139:4573-83; PMID:17050620; https://doi.org/10.1242/dev.02649

115. Dahanukar A, Walker JA, Wharton RP. Smaug, a novel RNA-binding protein that operates a translational switch in Drosophila. Mol Biol Cell 1999; 4:209-18; PMID:10488336; https://doi.org/10.1016/S1097-2765(00)80368-8

116. Jeske M, Moritz B, Anders A, Wahle E. Smaug assembles an ATP-dependent stable complex repressing nanos mRNA translation at multiple levels. EMBO J 2010; 30:90-103; PMID:2081899; https://doi.org/10.1038/emboj.2010.283

117. Gingras AC, Gygi SP, Raught B, Polakiewicz RD, Abraham RT, Hoeckstra MF, Aebersold R, Sonenberg N. Regulation of 4E-BP1 phosphorylation: A novel two-step mechanism. Genes Dev 1999; 13:1422-37; PMID:10364159; https://doi.org/10.1101/gad.13.11.1422

118. Gingras AC, Gygi SP, Raught B, Polakiewicz RD, Wyslouch-Cieszynska A, Aebersold R, Sonenberg N. Hierarchical phosphorylation of the translation inhibitor 4E-BP1. Genes Dev 2001; 15:2852-64. PMID:11691836; https://doi.org/10.1101/gad.912401

119. Hentze MW, Preiss T. The REM phase of gene regulation. Trends Biochem Sci 2010; 35:423-6; PMID:20554447; https://doi.org/10.1016/j.tibs.2010.05.009

120. Castello A, Horos R, Strein C, Fischer B, Eichelbaum K, Steinmetz LM, Krijgsveld J, Hentze MW. System-wide identification of RNA-binding proteins by interactome capture. Nat Protoc 2013; 8:491-500; PMID:23411631; https://doi.org/10.1038/nprot.2013.020

121. Granneman S, Kudla G, Petfalski E, Tollervey D. Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. Proc Natl Acad Sci U S A 2009; 106:9613-8; PMID:19482942; https://doi.org/10.1073/pnas.0901997106

122. Wang Z, Tollervey J, Briese M, Turner D, Ule J. CLIP: Construction of cDNA libraries for high-throughput sequencing from RNAs cross-linked to proteins in vivo. Methods 2009; 48:287-93; PMID:19272451; https://doi.org/10.1016/j.ymeth.2009.02.021

123. Hafner M, Landthaler M, Burger L, Khorshid M, Hauesser J, Berninger P, Rothballer A, Ascano M Jr, Jungkamp AC, Munschauer M, et al. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell 2010; 141:129-41; PMID:20371350; https://doi.org/10.1016/j.cell.2010.03.009

124. König J, Zarnack K, Rot G, Cuk T, Kayikci M, Zupan B, Turner DJ, Luscombe NM, Ule J. hiCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. Nat Publ Group 2010; 17:909-15; PMID:20601959; https://doi.org/10.1038/nbib383

125. Sugimoto Y, Vigilante A, Darbo E, Zirra A, Militti C, D’Ambrogio A, Luscombe NM, Ule J. hicLIP reveals the in vivo atlas of mRNA secondary structures recognized by Staufen I. Nature 2015; 519:491-4; PMID:25799884; https://doi.org/10.1038/nature14280

126. Liu N, Pan T. Probing RNA modification status at single-nucleotide resolution in total RNA. Methods Enzymol 2015; 506:149-59. PMID:26253970; https://doi.org/10.1016/bs.meie.2015.03.005

127. Edelmann FT, Schlundt A, Heym RG, Jenner A, Niedner-Boblenz A, Syed MI, Paillart JC, Stehle R, Janowski R, Sattler M, et al. Molecular architecture and dynamics of ASH1 mRNA recognition by its mRNA-transport complex. Nat Publ Group 2017; 24:152-61; PMID:28092367; https://doi.org/10.1038/nbib3351

128. Nicastro G, Candel AM, Uhl M, Oregioni A, Hollingworth D, Backofen R, Martin SR, Ramos A. Mechanism of β-actin mRNA Recognition by ZBP1. Cell Rep 2017; 18:1187-99; PMID:28147274; https://doi.org/10.1016/j.celrep.2016.12.091

129. Emsley P, Cowtan K. Coot: Model-building tools for molecular graphics. Acta Crystallogr Sec D Biol Crystallogr 2004; 60:2126-32; PMID:15572765; https://doi.org/10.1107/S0907444904019158

130. Handa N, Nureki O, Kurimoto K, Kim I, Sakamoto H, Shimura Y, Muto Y, Yokoyama S. Structural basis for recognition of the trna precursor by the Sex-lethal protein. Nature 1999; 398:579-85; PMID:10217141; https://doi.org/10.1038/19242

131. Gruner S, Peter D, Weber R, Wohlbald L, Chung MY, Weichenrieder O, Valkov E, Igreja C, Izaaralde E. The structures of eIF4E-eIF4G complexes reveal an extended interface to regulate translation initiation. Mol Cell 2016; 64:467-79; PMID:27773676; https://doi.org/10.1016/j.molcel.2016.09.020