RNA assemblages orchestrate complex cellular processes

Nielsen, Finn Cilius; Hansen, Heidi Theil; Christiansen, Jan

Published in:
BioEssays

DOI:
10.1002/bies.201500175

Publication date:
2016

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY-NC-ND

Citation for published version (APA):
RNA assemblages orchestrate complex cellular processes

Finn Cilius Nielsen1), Heidi Theil Hansen2) and Jan Christiansen2)*

Eukaryotic mRNAs are monocistronic, and therefore mechanisms exist that coordinate the synthesis of multi-protein complexes in order to obtain proper stoichiometry at the appropriate intracellular locations. RNA-binding proteins containing low-complexity sequences are prone to generate liquid droplets via liquid-liquid phase separation, and in this way create cytoplasmic assemblages of functionally related mRNAs. In a recent iCLIP study, we showed that the Drosophila RNA-binding protein Imp, which exhibits a C-terminal low-complexity sequence, increases the formation of F-actin by binding to 3′ untranslated regions of mRNAs encoding components participating in F-actin biogenesis. We hypothesize that phase transition is a mechanism the cell employs to increase the local mRNA concentration considerably, and in this way synchronize protein production in cytoplasmic territories, as discussed in the present review.

Keywords:
- post-transcriptional RNA regulon; RNA assemblage; RNA-binding protein; RNP granule; liquid droplet; low-complexity sequence

Introduction

Since 1961, when Jacob and Monod suggested their operon model, the concept of the polycistronic mRNA encoding functionally related proteins has been a paradigm in bacterial and archaeal gene regulation. It is assumed that this mode of co-regulation ensures proportional synthesis of components in multiprotein complexes by translational control, and recent ribosome profiling analysis of global gene expression in Escherichia coli by Weissman and co-workers [1] has shown that this is generally the case. Proportional synthesis avoids dominant-negative effects of excess components that will have to be eliminated by proteolytic quality control. An added bonus of protein synthesis from polycistronic mRNAs is the proximity of the resulting protein products, thus facilitating expedient macromolecular assembly or catalysis.

The prokaryotic mode of creating multiprotein complexes is in sharp contrast to the way the considerably larger eukaryotic cell, encoding monocistronic mRNAs, achieves proximity and stoichiometry, although classical cases of pyrimidine and purine biosynthesis have evolved multifunctional enzymes and enzyme clustering [2, 3]. Faced with this conundrum, Jack Keene [4] suggested that the coordination of gene expression to some extent was delegated from the DNA to the RNA level via the formation of post-transcriptional RNA regulons. The salient feature behind this hypothesis is the pivotal role of a common RNA-binding protein (RBP) associating with mRNAs encoding functionally related proteins and in this way dictating a common fate.

In this review, we describe examples of low-complexity RBPs coordinating RNA metabolism in both mono- and multicellular organisms with an emphasis on cytosolic ribonucleoprotein (RNP) granules (Table 1). We present the available evidence regarding their status as liquid droplets, including our own recent study of the participation of the Drosophila RBP Imp in F-actin formation.

Modular RBPs have a multitude of RNA targets

It is estimated that mammalian cells contain between 800 and 1,600 different RBPs, and about half of these encompasses common RNA-binding modules such as RRMs, KH domains,
CSDs, dsRBDs, RGG motifs, and zinc fingers in various combinations [5–7]. The other half does not contain an easily recognisable module – usually illustrated by the dual behavior of cytosolic aconitase, which switches between high-affinity RNA-binding and catalysis depending on the intracellular iron concentration [8]. From X-ray and NMR studies, a good understanding of the structure of the single RNA-binding module in complex with RNA has emerged [9, 10]. Moreover, several structural examples of truncated multi-domain RBPs in complex with RNA targets are available [11, 12].

There has been a flurry of high-throughput CLIP data on the targetome of full-length RBPs and attempts at defining recognition elements. Besides the well-characterized interaction between Pumilio and its target revealing an 8-nt recognition “code” [13, 14], many RBPs are promiscuous and exhibit low specificity, although the combination of modules within a single RBP will increase the specificity considerably [15]. Because a single RBP may literally have thousands of different interactions with the transcriptome, and hundreds of RBPs may behave in this manner, the complexity is daunting. Nevertheless, the combinatorial binding or competition among RBPs determines the fate of the individual transcript [16, 17].

### Are RNP granules membrane-less liquid droplets?

The cell is compartmentalized, and in recent years the participation of RBPs in the formation of membrane-less organelles by phase transition has been examined extensively. Seminal studies were carried out with germ granules (termed P granules) in *Caenorhabditis elegans*, identifying them as liquid droplets formed by phase separation from the cytoplasm [18], but the concept of phase transition of low-complexity proteins was described earlier for FG-rich nucleoporins [19]. The fortuitous use – by McKnight and co-workers – of a biotinylated isoxazole derivative to stabilize β-strand conformations in intrinsically disordered RBPs paved the way for the biochemical isolation of material within these membrane-less organelles. It turned out that especially low-complexity RBPs such as FUS (Fused in Sarcoma), hnRNPA1 and hnRNPA2B1 were abundant in the droplets/hydrogels [20, 21](Fig. 1). Intriguingly, mRNAs in the hydrogels exhibited a preponderance of long 3′UTRs, which provide a platform for RBP interplay, and could mean that mRNAs in liquid droplets are likely to be regulated mRNA species. An emerging theme in mRNA studies is that extensive 3′UTRs act as scaffolds for post-transcriptional regulatory purposes, and that physiologic responses depend on competing and/or cooperating trans-acting factors [22–24].

### P bodies and stress granules contain low-complexity RBPs

P bodies and stress granules are cytoplasmic RNP bodies that exhibit droplet properties. P bodies have been examined in studies of mRNA decay and translational repression, and the marker is usually the decapping enzyme Dcp2 [25]. Stress granules are formed in response to various cellular stresses such as arsenite and glucose deprivation, and are considered to be a pool of mRNPs stalled during translation initiation with TIA-1 as a common marker [26, 27]. However, a clear distinction between the two types of RNP bodies, based on the presence of a particular RBP, is not straightforward [28], and neither type seems to be the cause, but rather a consequence, of polysome disassembly [29–31].

Both types of RNP bodies incorporate low-complexity RBPs. In particular stress granules are regarded as pivotal in the pathogenesis of the neurodegenerative diseases amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) via fibrillization of RBPs such as hnRNPA1, hnRNPA2B1, FUS and TDP-43 [32–34]. Although the low-complexity sequence from hnRNPA1 is sufficient to mediate liquid-liquid phase separation, the inclusion of the two RRMs in full-length hnRNPA1, and the presence of RNA, favor reversibility rather than fibrillization [35]. This study also showed that fluorescence recovery after photobleaching (FRAP) analysis of hnRNPA1 within liquid droplets exhibited kinetics on the timescale of seconds, both in vitro and in vivo. This is in striking contrast to the near absence of recovery within 15 min in cross-β fibrillar hydrogels, and suggests a more rigid incorporation of hnRNPA1 in the latter. If a zipper motif within the low-complexity sequence of hnRNPA1 is removed (termed delta-hexa mutant), fibrillization is abrogated without affecting liquid droplet formation, thereby uncoupling the two phenomena mechanistically. Moreover, the presence of RNA reduced the necessary critical hnRNPA1 concentration for liquid-liquid phase separation. Finally, a disease-causing hnRNPA1 D262V mutation [36] did not appear to alter molecular interactions that drive phase separation, but increased the propensity toward amyloid-like fibrillization subsequently [35].

Parker and co-workers [37] also examined the propensity of full-length hnRNPA1 and the delta-hexa and D262V mutants to form liquid droplets in vitro, and obtained similar results, namely that a low salt concentration (37.5 mM) and high

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Organism</th>
<th>RNA-binding domains</th>
<th>Low-complexity sequence</th>
<th>RefSeq</th>
</tr>
</thead>
<tbody>
<tr>
<td>hnRNPA1</td>
<td>human</td>
<td>2 RRMs</td>
<td>Gly and Ser</td>
<td>NP_002127</td>
</tr>
<tr>
<td>hnRNPA2B1</td>
<td>human</td>
<td>2 RRMs</td>
<td>Gly</td>
<td>NP_115333</td>
</tr>
<tr>
<td>FUS</td>
<td>human</td>
<td>2 RRMs</td>
<td>Gly, Ser, and Tyr</td>
<td>NP_004951</td>
</tr>
<tr>
<td>Whi3p</td>
<td>yeast</td>
<td>1 RRM</td>
<td>Gin</td>
<td>NP_014202</td>
</tr>
<tr>
<td>PuF3p</td>
<td>yeast</td>
<td>8 pumilio repeats</td>
<td>Asn and Gln</td>
<td>NP_013088</td>
</tr>
<tr>
<td>Imp</td>
<td>Drosophila</td>
<td>4 KH domains</td>
<td>Gln</td>
<td>NP_511111</td>
</tr>
</tbody>
</table>

Table 1. RNA-binding proteins discussed in this review
Problems & Paradigms

of FUS are seen in relation to the neurodegenerative diseases between nucleus and cytoplasm, and importantly, mutations in terminal low-complexity sequence, an ability to shuttle to several facts: High intracellular concentration, extensive N-liquid-like compartments both in vitro and in vivo. This is due pending issue.

less, the physiological relevance of fibrous hydrogels is a

676

Bioessays 38: 674–681, © 2016 The Authors BioEssays Published by WILEY Periodicals, Inc.

An emerging theme in studies of liquid droplets is that they evolve into more solid-like entities, and that this transition is greatly accelerated by disease mutations. A controversial issue has been whether conformations of low-complexity sequences in liquid droplets are different from those appearing in more fibrous hydrogels. However, a chemical footprinting experiment with N-acetyl-imidazole has been unable to identify differences in side-chain accessibility of the low-complexity region in hnRNPA2B1, regardless of its presence in liquid droplets, hydrogels – or nuclei for that matter – and suggests cross-β polymerization as a unifying principle [39]. Nevertheless, the physiological relevance of fibrous hydrogels is a pending issue.

Fused in Sarcoma (FUS) is a favorite RBP in studies of liquid-like compartments both in vitro and in vivo. This is due to several facts: High intracellular concentration, extensive N-terminal low-complexity sequence, an ability to shuttle between nucleus and cytoplasm, and importantly, mutations of FUS are seen in relation to the neurodegenerative diseases (NLS) is present, the kinetics of fibrillization is increased. In the latter case, inhibition of nuclear localization increases the cytoplasmic concentration of FUS, thereby facilitating phase separation and the likelihood of subsequent aggregation. This maturation phenomenon is essentially a conversion from a metastable liquid-like state to a thermodynamically more stable solid-like state, and takes place within 8 h in vitro. However, it is much slower in vivo, probably because it is efficiently counteracted by chaperones and disaggregases [41, 42]; hence cytoplasmic inclusions will first become a severe phenotype with age [40].

A similar conclusion was reached in a C. elegans model of FUS-dependent neurodegeneration [43]. In this study of C-terminal NLS mutations, the conversion of liquid droplets into fibrillar hydrogels was sufficient to mediate neurotoxicity. An important consequence of the transition from dynamic liquid droplets to more static hydrogels was impairment of on-site translation in Xenopus retinal neurons, due to entrapped mRNPs. In the ongoing discussion about the physiological significance of hydrogels, a distinction between reversible and irreversible hydrogels may be more fruitful, because the difference in viscosity is slight between the former and liquid droplets [44]. Therefore, partial polymerization into a loose yet reversible fibrous mesh could be appropriate for the physiology of longer-lived RNP granules and the nuclear pore matrix [45].

Whi3p droplet properties depend on RNA identity

In multinucleate large Ashbya cells, the G1 cyclin CLN3 transcript shows a nonrandom spatial clustering due to a glutamine-rich region in the RBP Whi3p, and this inhomogeneous distribution is crucial for cell-cycle timing variability and asynchrony of the nuclei [32]. An additional mRNA target for Whi3p is the BN1 transcript encoding formin, which is important for establishing polarity [46]. Recombinant Whi3p (28 μM) is able to form protein-only liquid droplets at 75 mM KCl in vitro, but not at physiological salt concentrations. In contrast, the liquid-liquid phase separation of protein-RNA droplets is promoted at physiological salt concentration and
Although entity in terms of high-affinity RNA-binding, granular RNP assembly, and RNA phylogenetic and experimental standpoint, the four KH domains constitute a functional RBPs, the 2\( ^{\text{nd}} \) RNA recognition motifs (RRM1 and 2), and four C-terminal hnRNPK homology amino acids and exhibits six characteristic RNA-binding modules, namely two N-terminal low-complexity sequences in intrinsically disordered regions of the RBP. By breaking protein-protein, rather than RNA-protein, interactions, mRNA could be released in an accessible form to the translational apparatus.

**Puf3p phosphorylation activates mRNAs encoding mitochondrial proteins**

Low-complexity sequences exhibit enrichment of serine and arginine residues [54], thereby providing ample opportunity for post-translational modifications such as phosphorylations and methylations, respectively, that could regulate the reversibility of mRNAs within RNP granules [55]. For example, chemical inhibition of the dual specificity kinase DYRK3 in HeLa cells affects the dissolution of stress granules by preventing autophosphorylation of its own low-complexity N-terminal and phosphorylation of RBPs [56]. Moreover, phosphorylation of intrinsically disordered MEG substrates by the C. elegans DYRK3 homologue drives the dissolution of P granules in embryos, whereas condensation is mediated by a PP2A phosphatase [57]. In the next section, we describe one of the most clear-cut examples of a post-transcriptional RNA regulon that is regulated by environmental cues.

**RNP granules encompass RNA assemblages**

The RNP granule can be regarded as the physical manifestation of the post-transcriptional RNA regulon, because low-complexity sequences in intrinsically disordered regions of RBPs would have an ability to interact both homo- and heterotypically in a dynamic fashion. The concept of an RNP granule is appealing, because this would allow mRNAs encoding functionally related proteins to be stored/transported in a repressed form and be translated locally in a synchronous manner. The RNP granule concept has been widely accepted among neurobiologists, because mRNAs are transported over considerable distances to obtain local protein synthesis in axons and dendrites [48].

Despite the widespread acclaim of the RNP granule as a mode of harnessing and transporting mRNAs, there are few studies that have identified the molecular motors involved [49–51]. It is generally assumed that long-range movements of RNP granules take place on microtubules and that short-range transport will take place on microfilaments, a concept that is based on disturbances following nocodazole and cytochalasin D treatments, respectively. However, because optical resolution has improved, it is not clear whether RNP granules are co-localizing with microtubules or are tethered via ER/endosomes [52, 53]. An additional question related to RNP granules – at least if based on a solid-like structure – is that of reversibility, i.e. how is the mRNA deployed at the site of translation? The most obvious solution is a signal transduction event leading to a post-translational modification or proteolysis of the RBP.

**Problems & Paradigms**

200 nM protein by the presence of RNA in the guise of either CLN3 or BNII mRNAs [47]. Moreover, the strict dependence on the RRM domain for Whi3p droplet formation reinforces the important role of RNA. Although both mRNAs encompass five putative UGCAU attachment sites for Whi3p, the viscosity of the two types of protein-RNA droplets, and their propensity to fuse, are different. This implies that the identity of the participating mRNA component could tune the functional specificity and encompass information beyond mere coding ability. Therefore, a combination of low-complexity sequences and multivalent interactions appears to be critical for intracellular phase transition, although the effect of the spacing pattern of Whi3p attachment sites on mRNA targets is unknown. During droplet maturation, the transition from a liquid to a more fibrillar state of glutamine-rich RBPs, such as Whi3p, is diminished considerably by the presence of RNA [40, 47]. In fact, the negatively charged RNA influences the viscoelastic properties of liquid droplets, as measured by microheology and FRAP, and gives rise to decreased viscosity within the droplet termed RNA-induced fluidization [44, 47]. The presence of mRNA within these intracellular reactors is therefore crucial for reversibility and, by inference, its subsequent recruitment to the translational apparatus.

**Bioessays 38: 674–681, © 2016 The Authors BioEssays Published by WILEY Periodicals, Inc.**

---

**Figure 2.** Schematic representation of Drosophila Imp and human IMP1. Numbers refer to amino acid positions bordering various domains. Human IMP1 is composed of 579 amino acids and exhibits six characteristic RNA-binding modules, namely two N-terminal RNA recognition motifs (RRM1 and 2), and four C-terminal hnRNPK homology domains (KH1–4). Whereas two RRM or four KH domains can be found in other RBPs, the 2 + 4 modular architecture seen in vertebrate IMPs is unique. From both a phylogenetic and experimental standpoint, the four KH domains constitute a functional entity in terms of high-affinity RNA-binding, granular RNP assembly, and RNA localization [77]. Although Drosophila Imp exhibits rudimentary features of RRM in minor isoforms, the major isoform of 566 amino acid lacks RRM. Instead, Drosophila Imp contains a large C-terminal low-complexity glutamine-rich sequence (LCS).
Problems & Paradigms

cross-linking sites on the entire transcriptome regardless of experimental approach allows the identification of in vivo followed by immunoprecipitation of lysates [64]. The iCLIP Drosophila an RNA assemblage within RNP granules [63], we subjected coordinates the expression of a subset of mRNAs by creating [20]. To address the possibility that this particular RBP of hydrogels in a biotinylated isoxazole precipitation assay passes a C-terminal glutamine-rich low-complexity sequence homologue, we noticed that the latter encom-

In a comparison of the oncofetal RBP IMP family with its Drosophila Imp family, we noticed that the latter encompasses a C-terminal glutamine-rich low-complexity sequence (Fig. 2). Moreover, Drosophila Imp has been shown to be part of hydrogels in a biotinylated isoxazole precipitation assay [20]. To address the possibility that this particular RBP coordinates the expression of a subset of mRNAs by creating an RNA assemblage within RNP granules [63], we subjected Drosophila S2 cells to either short or long-wave UV irradiation followed by immunoprecipitation of lysates [64]. The iCLIP experimental approach allows the identification of in vivo cross-linking sites on the entire transcriptome regardless of the global S2 transcriptome, and that the total cytosolic actin monomer concentration was unchanged. We interpret this as an effect of an RBP on local cytoplasmic events rather than on overall cellular post-transcriptional regulation. A simple model would be that Imp defines an RNA assemblage ensuring that functionally related mRNAs – i.e. a post-transcriptional RNA regulon – would be in the vicinity of each other, and produce components at a higher local concentration than possible from dispersed protein synthesis in the cytoplasm.

Imp-dependent stimulation of F-actin formation has developmental consequences, especially for processes relying on F-actin dynamics such a neuronal growth cone steering and synaptogenesis [67]. Imp-deficient flies exhibit a spectrum of neurogenesis defects – almost to the point of being stochastic – and even survivors at the pharate adult stage are unable to eclose, suggesting locomotory failure. The broad spectrum of defects is compatible with what was observed in the cell-line: namely that all components appear to be formed in the appropriate amount – at least at the transcript level – regardless of Imp. However, the possibility of producing the components in high local concentrations facilitating macromolecular assembly is jeopardized to varying degrees. A study by Besse and co-workers [68] supports the role of Drosophila Imp in facilitating F-actin formation during neurogenesis,
because defects in mushroom body (somewhat reminiscent of the hippocampus) neurogenesis in Imp-deficient animals can be partially rescued by Chickadee. The latter is the Drosophila homologue of profilin, which is the main facilitator of F-actin formation from G-actin. One would also expect transcripts encoding Rho GTPases, moesin, ena/VASP, and cytosolic actin itself in a post-transcriptional RNA regulon coordinated by Imp and conveying cues from the membrane to the cytoskeleton, and this was actually observed [64]. A recent translational profiling study of the conversion of growth cones into presynaptic terminals in Drosophila photoreceptor R cells has shown that prior to differentiation there is a 40-fold upregulation of Imp. Moreover, all of the actin regulatory proteins encoded by mRNAs, identified in our iCLIP study, are expressed in R cells during synapse formation [69]. Taken together, a picture emerges where one phase transition, namely the partitioning of RNPs into liquid droplets, facilitates a subsequent transition of monomeric actin into F-actin [70].

**The advantages in clustering functionally related mRNAs within an RNA assemblage**

Compared with solid-like cytoplasmic inclusions, liquid droplets offer some intrinsic advantages especially in terms of dynamics, because material can enter and exit the droplet faster but still be concentrated in a small volume [71]. The latter is important, because concomitant release of a high local concentration of related mRNAs to the translational apparatus will facilitate a high local protein concentration of newly synthesized interacting species (Fig. 4). Based on RNP body condensation/dissolution behavior following translational inhibition of either initiation or elongation, respectively [29, 72], it is generally assumed that the translational apparatus does not have access to the droplet interior. However, it should be recalled that ribosomal proteins themselves exhibit intrinsically disordered regions termed “waggly tails” [73]. A few studies have addressed the presence of translating ribosomes in the vicinity of RNP granules [74, 75], and the study by Mili and co-workers [76] actually showed translation within cytoplasmic FUS granules generating local protein production from APC-RNPs in NIH 3T3 protrusions. In general, the understanding of the location and the identity of the ribosomes carrying out the local synthesis is lagging behind, so this should be an important line of enquiry in upcoming studies.

**Conclusions and prospects**

In this essay, we have presented the rationale behind clustering functionally related mRNAs in assemblages, with an emphasis on coordinated and accelerated recruitment to the translational apparatus. An additional bonus of partitioning is fidelity, allowing much less volume and time for aberrant interactions among the resulting protein products. However, this mechanism does not come without a cost, because phase separation of an RBP is a fine balance between physiological assembly and pathological fibrillization.

Recent genome-wide screenings strongly support the concept of RNA assemblages segregated by particular RBPs. To understand the physiological significance of RNA assemblages, we need a deeper understanding of their dynamic behavior and molecular composition. The crucial role of RNA, salt and temperature for droplet dynamics has been obtained from studies in vitro focusing on the biophysical behavior of tagged low-complexity sequences. However, visualization of the interplay between RNP granules and the translational apparatus in vivo is strongly needed. Microscopes are now approaching single RNP particle resolution, so studies should be directed toward endogenous components rather than reporters with various tags. Moreover, low-complexity sequences in homologous RBPs exhibit low levels of conservation, so comparative studies may provide clues to the rationale behind these intrinsically disordered regions in terms of evolutionary rewiring of post-transcriptional regulation.
A number of granular RBPs have been implicated in neurological diseases. So far we know little about the potential druggability of the conditions, but the concept of modifying particular RNA assemblages is an appealing one, because a complex process might be influenced by a single drug. Blockmirs may be employed to up-regulate RBPs, whereas another possibility would be to prevent granule formation by modulating post-translational modifications, such as phosphorylations, or by introducing RNA sponges. Considering the many ongoing human sequencing initiatives, we envisage that RBPs will be implicated in a number of both common and rare diseases, which will stimulate the development of new therapeutic strategies.

References


Bioessays 38: 674–681, © 2016 The Authors BioEssays Published by WILEY Periodicals, Inc. 681