Unconventional RNA-binding proteins step into the virus–host battlefront

Manuel Garcia-Moreno* | Aino I. Järvelin* | Alfredo Castello

Department of Biochemistry, University of Oxford, Oxford, UK

Correspondence
Alfredo Castello, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK.
Email: alfredo.castello@bioch.ox.ac.uk

Funding information
Marie-Sklodowska-Curie Grant, Grant/Award Number: 700184; European Commission; University of Oxford through the John Fell Fund programme; MRC: Medical Research Council (UK), Grant/Award Number: MR/L019434/1

The crucial participation of cellular RNA-binding proteins (RBPs) in virtually all steps of virus infection has been known for decades. However, most of the studies characterizing this phenomenon have focused on well-established RBPs harboring classical RNA-binding domains (RBDs). Recent proteome-wide approaches have greatly expanded the census of RBPs, discovering hundreds of proteins that interact with RNA through unconventional RBDs. These domains include protein–protein interaction platforms, enzymatic cores, and intrinsically disordered regions. Here, we compared the experimentally determined census of RBPs to gene ontology terms and literature, finding that 472 proteins have previous links with viruses. We discuss what these proteins are and what their roles in infection might be. We also review some of the pioneering examples of unorthodox RBPs whose RNA-binding activity has been shown to be critical for virus infection. Finally, we highlight the potential of these proteins for host-based therapies against viruses.

This article is categorized under:
- RNA Interactions with Proteins and Other Molecules > Protein–RNA Interactions: Functional Implications
- RNA in Disease and Development > RNA in Disease
- RNA Interactions with Proteins and Other Molecules > RNA–Protein Complexes

KEYWORDS
infection, RBPome, RNA, RNA-binding domain, RNA-binding protein, RNA interactome capture, RNA metabolism, translation, virus

1 | INTRODUCTION

As obligate intracellular pathogens with small-sized genomes, viruses rely on host cell resources for the completion of their biological cycle and have developed sophisticated mechanisms to hijack key host machineries (Davey, Trave, & Gibson, 2011). On the other hand, the host cell activates a complex response to generate an “antiviral state”, transitioning from a virus permissive to a hostile environment (Barbalat, Ewald, Mouchess, & Barton, 2011; Rehwinkel & Reis e Sousa, 2013). The infected cell thus becomes a battlefield where the virus and the host cell deploy their molecular “ arsenals” to take control of the processes required for infection. RNA is a central molecule in virus biology, because it functions as a genome and template for replication and transcription in RNA viruses, and also as a messenger (m)RNA for the synthesis of viral proteins in both DNA and RNA viruses (Baltimore, 1971). While viral RNA-binding proteins (RBPs) guide the viral (v)RNA through the series of RNA metabolic steps, including synthesis, splicing, localization, translation, and packaging, cellular RBPs are also

*Manuel Garcia-Moreno and Aino I. Järvelin contributed equally to this work.
necessary for these processes to occur. Work accumulated over the last two decades has shown that cellular RBPs are indeed involved in virtually every stage of vRNA metabolism in the cycle of both RNA and DNA viruses. Moreover, RBPs are also essential for antiviral defense, because many sensors of the innate immunity recognize unusual RNA products of viral replication, such as double-stranded (ds)RNA and triphosphate ends (Barbalat et al., 2011; Vladimer, Gorna, & Superti-Furga, 2014).

Many of the present-day pandemics involve viruses, including human immunodeficiency virus (HIV), hepatitis C virus (HCV), and hepatitis B virus (HBV), leading to millions of deaths every year. However, most virus infections still lack effective treatments or vaccines. So far, antiviral treatments have mainly focused on the use of compounds acting directly on viral enzymes. However, drug resistance can arise due to the high mutation rate of viral polymerases and the co-existence of genetically diverse viral populations called quasispecies (Irwin, Renzette, Kowalik, & Jensen, 2016; Sanjuan & Domingo-Calap, 2016). The selective pressure driven by the drug promotes the expansion of resistant viral strains, shifting the viral population from drug sensitive to resistant (Clutter, Jordan, Bertagnolio, & Shafer, 2016; Garbelli, Riva, Crespan, & Maga, 2017). To prevent the emergence of resistant strains, novel antiviral treatments combine multiple drugs targeting different viral enzymes (Broder, 2010; Emery & Feld, 2017; Valuev-Elliston & Kochetkov, 2017). However, a broad catalogue of antivirals is required to make this strategy feasible, and this is only available for a small number of viruses such as HIV. Although some drugs can inhibit multiple viruses (e.g., nucleoside analogues with activity against both HIV and HBV [Debing, Neyts, & Delang, 2015; Mendes-Correia & Nunez, 2010]), viral enzyme-targeting drugs usually fail at protecting even against closely related virus species. By contrast, inhibition of cellular factors required for virus infection is expected to provide a higher barrier to the emergence of resistance. This is due to two key factors: (a) the higher stability of human genomes and (b) the difficulty for the virus to develop new strategies to utilize alternative host factors or pathways (Law, Korth, Benecke, & Katze, 2013). One of the best examples of targeting host factors is alispoviril (also known as Debio 025), an inhibitor of cyclophilin A (PPIA) that strongly reduces HCV infection (Flisiak et al., 2008; Paehuyse et al., 2006). Because different viruses often exploit common cellular pathways, host proteins are promising targets for the development of broad-spectrum antiviral therapies (Martinez, Sasse, Bronstrup, Diez, & Meyerhans, 2015; Zhu, Meng, Wang, & Wang, 2015). Cellular kinases (Schor & Einav, 2018), cyclophilins (Dawar, Tu, Khattak, Mei, & Lin, 2017; Debing et al., 2015), and heat-shock proteins (Geller, Taguwa, & Frydman, 2012; Y. Wang et al., 2017) represent potential candidates for such antiviral intervention.

A substantial proportion of the host proteins that benefit or hinder viral replication are RBPs, and many efforts have been undertaken in the last years to characterize the roles of protein–RNA interactions in infected cells. Many cellular RBPs containing well-established RNA-binding domains (RBDs) are known to be critical for the infection of specific viruses. Next generation, proteome-wide methods have discovered hundreds of unconventional RBPs in the past years (Baltz et al., 2012; Castello et al., 2012; Hentze, Castello, Schwarzl, & Preiss, 2018), raising the possibility of their potential involvement in virus infection as well. Some of these novel RBPs have already been linked to diverse virus-related processes, but in most cases, their ability to interact with RNA was unknown. Here, we present a compendium of RBPs involved in virus infection and review several pioneering studies about the roles of unconventional RBPs in infection.

## 2 THE ERA OF THE RNA INTERACTOMES: NEW INSIGHTS INTO VIRUS INFECTION

### 2.1 Identifying “virus-related” RBPs

In the past, proteins were classified as RBPs largely based on the presence of well-established, classical RBDs such as the RNA recognition motif (RRM), K-homology (KH) domain, DEAD box helicase domain, dsRNA motif (DSRM), zinc-fingers (ZF), and others, as reviewed in detail by Anantharaman, Koonin, and Aravind (2002) and Lunde, Moore, and Varani (2007). Most of these classical RBDs have been well characterized biochemically and structurally, and they fold into well-established stable secondary and tertiary structures (Lunde et al., 2007). RBPs are usually modular, combining multiple RBDs that cooperate to recognize different sequence and/or structural features in target RNAs (Gronland & Ramos, 2017; Lunde et al., 2007). This modularity enables interactions that can range from very selective to promiscuous, supporting a wide diversity of functions (Hentze et al., 2018).

A recently developed proteome-wide method called “RNA interactome capture” (RNA-IC) has greatly expanded the census of cellular RBPs (Baltz et al., 2012; Castello et al., 2012; Hentze et al., 2018). RNA-IC employs ultraviolet (UV) crosslinking of living cells to induce covalent bonds between proteins and RNA placed at zero distance. Cells are then lysed under denaturing conditions, and polyadenylated [poly(A)] RNA and proteins covalently linked to it are captured under very stringent conditions with oligo(dT) magnetic beads. After elution, RBPs are released by RNase treatment and analyzed by quantitative proteomics. This method has been applied to numerous cell lines and organisms, and has revealed RBPs acting in living systems in a comprehensive manner (Hentze et al., 2018). Proteins identified by this method are highly enriched in

---

**References**

- Barbalat et al., 2011
- Vladimer, Gorna, & Superti-Furga, 2014
- Flisiak et al., 2008
- Paehuyse et al., 2006
- Martinez, Sasse, Bronstrup, Diez, & Meyerhans, 2015
- Zhu, Meng, Wang, & Wang, 2015
- Schor & Einav, 2018
- Debing et al., 2015
- Hentze, Castello, Schwarzl, & Preiss, 2018
- Anantharaman, Koonin, and Aravind, 2002
- Lunde, Moore, and Varani, 2007
- Hentze et al., 2018
- Baltz et al., 2012
- Castello et al., 2012
- Hentze, Castello, Schwarzl, & Preiss, 2018
- Gronland & Ramos, 2017
- Lunde et al., 2007
- Clutter, Jordan, Bertagnolio, & Shafer, 2016
- Garbelli, Riva, Crespan, & Maga, 2017
classical RBDs, as expected. However, hundreds of unorthodox RBPs were also reported to bind RNA. A recent review compiled the scope of RBPs identified by all the RNA-IC studies so far, classifying 1,392 proteins as RBPs in human cells (Hentze et al., 2018). Among them, 278 RBPs harbor at least one classical RBD, and 1,114 can be classified as unorthodox RBPs based on lack of such domains (Figure 1a and File S1, Supporting Information).

The newly established RNA-binding proteome (RBPome) represents a unique resource to explore the interconnections between host RBPs and viruses. In order to do so, we first analyzed how many of these experimentally identified RBPs have

**FIGURE 1** RNA-binding proteome (RBPome) implicated in virus infection and immunity. (a) Classification by RNA-binding domain (i.e., classical, nonclassical, and other) of RNA-binding proteins (RBPs) in different datasets: human RNA interactome capture (RNA-IC); RNA-IC linked to virus by gene ontology (GO) terms; RNA-IC linked to virus by literature mining (LM); RNA-IC linked to virus by GO and LM. (b) Word cloud of protein domains present in all virus-linked RBPs. (c) STRING protein network showing connections between all virus-linked RBPs. ER, endoplasmic reticulum
been annotated by gene ontology (GO) terms related to viruses, infection, and immunity (File S1), leading to the identification of 342 “virus-linked” RBPs. Furthermore, complementary literature mining analysis revealed 130 additional RBPs with more than five publications in PubMed linking them to viruses. Taking together, GO and literature mining lead to a superset of 472 virus-linked RBPs (Table S1 and Figure 1a).

Among the compendium of virus-linked RBPs, the most prevalent domains are classical RBDs including RRM, DEAD box helicase, KH, DSRM, and ZF (Figure 1b), as expected due to the frequent occurrence of these domains in the human RBPome. Virus-linked RBPs are involved in every step of posttranscriptional control of gene expression, including splicing and RNA transport, and translation (Figure 1c). Strikingly, 378 virus-linked RBPs lack canonical RBDs, suggesting that many newly discovered RBPs are involved in virus infection. We further grouped these unorthodox RBPs into “noncanonical” and “other” based on the presence or absence of nonclassical domains previously reported to bind RNA, respectively [Castello et al., 2012; Castello et al., 2016]. We discuss these unconventional RBPs in detail below.

### 2.2 Well-known RBPs meet viruses

Well-established RBPs, in particular classical RBPs, have been linked to virtually every step of virus infection (Table 1). Even though most RNA viruses encode their own RNA polymerases to synthesize vRNAs, host RBPs participate in: (a) selective

| RNA-related process | Examples of RBPs | References |
|---------------------|------------------|------------|
| Recognition of vRNA template | EEF1A1, PCBP1, PTBP1, NCL | (Z. Li & Nagy, 2011; Nagy & Pogany, 2011; R. Y. Wang & Li, 2012) |
| Recruitment of vRNA to replication factory | HNRNPA1, HNRNPC, EEF1A1 | (Z. Li & Nagy, 2011; Nagy & Pogany, 2011; R. Y. Wang & Li, 2012) |
| Switch from translation to replication | PTBP1, PCBP2, PABPC1, SSb, ILF3 | (Z. Li & Nagy, 2011; Nagy & Pogany, 2011; R. Y. Wang & Li, 2012) |
| vRNA synthesis | HNRNPC, TIA1, PTBP1, HNRNPK, POLR2, NONO, SFPP, RBM14, EEF1A1 | (Harwig, Landick, & Berkhout, 2017; Z. Li & Nagy, 2011; Nagy & Pogany, 2011; R. Y. Wang & Li, 2012) |
| Host mRNA transcription shutoff | CPSF4, PABPN1, POLR2 | (Herbert & Nagy, 2016) |
| Splicing | HNRNPA1, HNRNP2B1, HNRNP3, HNRNPH, HNRNPC, HNRNP4M, U2AF1, U2AF2, SR proteins | (Meyer, 2016; Stoltzfus & Madsen, 2006; Tazi et al., 2010) |
| RNA editing/reading of edited RNA | ADAR, ADAR1B, ADAR2B, RBM15, ALKBH5, YTHDF1, YTHDF3 | (Gonzales-van Horn & Sarnow, 2017; Samuel, 2011) |
| Nucleus–cytoplasm shuttling | PTBP1, PCBP2, SSb, CSDE1, HNRNPC, HNRNPA1, HNRNPK, HNRNP4M, DHX9, SYNCRIP, TIA1, TIAL1, G3BP1 | (Lloyd, 2015) |
| vRNA export from nucleus | NFX1, ALYREF, RAN, SR proteins, DDX1, DDX3X, DDX5, HNRNP2B1, SFPP, MATR3 | (Kuss, Mata, Zhang, & Fontoura, 2013; Stake, Bann, Kaddis, & Parent, 2013) |
| vRNA trafficking and packaging | HNRNPA2B1, EEF1A1, STAUA1 | (Cochrane, McNally, & Mouland, 2006; Kaddis Maldonado & Parent, 2016; Stake et al., 2013) |
| Extracellular trafficking | HNRNPA2B1, AGO2, SYNCRIP, YBX1 | (Kouwaki, Okamoto, Tsukamoto, Fukushima, & Oshiumi, 2017) |
| vRNA stability | ELAVL1, PCBP2, HNRNP2D, YBX1, ILF3 | (Dickson & Wilusz, 2011; Moon & Wilusz, 2013) |
| 5′→3′ RNA degradation | XRN1, PATL1, XRN2 | (Molleston & Cherry, 2017; Moon, Barnhart, & Wilusz, 2012; Narayan & Makino, 2013; Oshiumi, Mifsud, et al., 2016; Rigby & Rehwinkel, 2015) |
| 3′→5′ RNA degradation | DDX17, ZC3HAV1, UPF1, UPF3, RRP6 | (Molleston & Cherry, 2017; Moon, Barnhart, & Wilusz, 2012; Narayan & Makino, 2013; Oshiumi, Mifsud, et al., 2016; Rigby & Rehwinkel, 2015) |
| Translation | PABPC1, EIF4E, EIF4G, EIF4A, EIF2, EIF3, ribosomal proteins, EIF5B, EEF1A1, EEF2, ETF1 | (McCormick & Khapersky, 2017; Smith & Gray, 2010; Walsh & Mohr, 2011) |
| RNP granules | G3BP1, G3BP2, EIF2AK2, LSM14A, XRN1, TIA1, TIAL1, STAUA1, DDX3X, DDX6 | (Beckham & Parker, 2008; Lloyd, 2013; McCormick & Khapersky, 2017; Onomoto, Yoneyama, Fung, Kato, & Fujita, 2014; Reineke & Lloyd, 2013; Yoneyama, Jogi, & Onomoto, 2016) |
| Antiviral sensors and cofactors | TLR3, DDX58, IFI1H1, ADAR, ADAR1B, EIF2AK2, OAS, DHX9, DDX1, DDX3X, DHX29, DHX36 | (Beachboard & Homer, 2016; Gelinas, Clerzio, Shaw, & Gatignol, 2011; Mastial, Barraud, & Allain, 2013; Oshiumi, Kouwaki, et al., 2016; Peisley & Hur, 2013; Ranji & Boris-Lawrie, 2010; Steiner & Klostermeier, 2012) |

*Note.* mRNA = messenger RNA; RBP = RNA-binding protein; vRNA = viral RNA.
recognition of the viral template RNA; (b) recruitment of vRNA to the membranous compartments where viral replication takes place and activation of viral replicase complexes; (c) switch from translation to replication; and (d) synthesis of negative, positive, and subgenomic vRNAs (Z. Li & Nagy, 2011; Nagy & Pogany, 2011; R. Y. Wang & Li, 2012). Other viruses, such as HIV, rely on the host RNA transcription machinery, mainly RNA polymerase II (POLR2). RBPs such as non-POU domain containing octamer binding (NONO) or splicing factor proline and glutamine rich (SFPQ) can be recruited co-translationally to regulate vRNA synthesis, as described for the DNA virus Epstein–Barr (Harwig et al., 2017). Furthermore, many viruses fine-tune or simply shut off host transcription to evade antiviral defenses. To do so, viruses often target cellular RBPs such as cleavage and polyadenylation specific factor 4 (CPSF4), poly(A) binding protein nuclear 1 (PABPN1), transcription factors, or, directly, POLR2 (Herbert & Nag, 2016).

Viruses also hijack RBPs involved in co/post-transcriptional RNA processing. For example, several RNA and DNA viruses usurp the cellular capping and polyadenylation machineries to add a 5’ cap structure and 3’ poly(A) tail to their mRNAs through a broad variety of molecular mechanisms (Decroly, Ferron, Lescar, & Canard, 2011; Schrom, Moschall, Schuch, & Bodem, 2013). The spliceosome and regulators of splicing are hijacked by viruses that need to splice their RNAs (Meyer, 2016; Stoltzfuß & Madsen, 2006; Tazi et al., 2010). Several RBPs involved in RNA editing (e.g., 5’-methylcytosine, N6-methyladenosine, or adenosine-to-inosine modifications) or recognition of edited RNA can positively or negatively modulate virus infection, affecting, for example, vRNA stability, translation, and antiviral sensing (Gonzales-van Horn & Sarnow, 2017; Samuel, 2011). Interestingly, many of the RBPs involved in the aforementioned processes have a nuclear distribution, but cytoplasmic RNA viruses have the capacity to alter their subcellular localization and co-opt them for new functions (Lloyd, 2015). Heterogeneous nuclear ribonucleoproteins (HNRNPs) are typically identified as interactors of vRNAs with an altered localization in infected cells. While these interactions are likely to be functionally relevant for different steps of infection (Table 1), HNRNPs are very abundant proteins and thus easy to detect in proteomic-based analyses, probably representing the tip of the iceberg.

Host RBPs involved in RNA trafficking are also hijacked by viruses and contribute to vRNA export from the nucleus, transport throughout the cytoplasm, and packaging into viral capsids (Cochrane et al., 2006; Kaddis Maldonado & Parent, 2016; Stake et al., 2013). On the other hand, viruses can impair host RNA trafficking by targeting RBPs, including the nuclear RNA export factor 1 (NXF1) and GTP-binding nuclear protein Ran (RAN). By doing so, viruses prevent nuclear export of cellular mRNAs encoding antiviral factors and cytokines and block import into the nucleus of relevant transcription factors (Castello, Alvarez, & Carrasco, 2011; Kuss et al., 2013).

Cellular RBPs that modulate RNA stability, such as ELAV-like RBP 1 (ELAVL1 or HuR), can be utilized by viruses to preserve the integrity of their vRNAs (Dickson & Wilusz, 2011; Moon & Wilusz, 2013). On the other hand, viruses can promote sequestration of cellular RBPs that enhance RNA stability, leading to destabilization of host transcripts. Cellular exonucleases are also key players in infection. Viruses have developed multiple ways of evading the cellular 5’→3’ and 3’→5’ RNA degradation machineries that would otherwise target vRNAs for decay (Molleston & Cherry, 2017; Moon et al., 2012; Narayanan & Makino, 2013; Oshiumi, Mifsud, & Daito, 2016; Rigby & Rehwinkel, 2015). Viruses have also evolved strategies to use the 5’→3’ mRNA decay machinery, including decapping activators (Jungfleisch, Blasco-Moreno, & Diez, 2016) and the nonsense-mediated decay machinery (Balistreri, Bognanni, & Muhlemann, 2017), to the benefit of vRNA translation and replication.

Viruses strongly rely on the host translation machinery to synthesize viral proteins (McCormick & Khaperskyy, 2017; Walsh & Mohr, 2011). Nevertheless, many viruses disrupt specific host RBPs critical for translation initiation, such as poly(A) binding protein cytoplasmic 1 (PABPC1) (Smith & Gray, 2010), eukaryotic translation initiation factor (EIF)4E, EIF4G, EIF2, or EIF3; translation elongation (EIF5B, EEF1A1, or EEF2) and termination (e.g., ETF1) as well as ribosomal proteins. vRNA can bypass host translation shutoff through noncanonical translation mechanisms (Au & Jan, 2014; Firth & Brierley, 2012; Mailliot & Martin, 2018), possibly by recruiting specific RBPs and specialized ribosomes. For example, picornaviruses and other viruses employ cellular trans-acting factors (ITAFs) to facilitate internal ribosome entry site (IRES)-mediated, cap-independent translation (Mailliot & Martin, 2018).

RNA helicases are a family of proteins often implicated in virus infection. They can promote different steps of viral replication or mediate the activation of the antiviral response (Ranji & Boris-Lawrie, 2010; Steimer & Klostermeier, 2012). Several antiviral sensors and effectors, such as the retinoic acid-inducible gene I protein (RIG-I, also known as DDX58), harbor classical helicase cores to recognize replication intermediates (Barbalat et al., 2011; Rehwinkel & Reis e Sousa, 2013). These antiviral sensors cooperate with different RNA helicases, including DExH-box helicase (DHX)9, DEAD-box helicase 3 X-linked (DDX3X), or DHX36, that act as accessory factors in antiviral pathways (Oshiumi, Kowaki, & Seya, 2016).
3 | UNORTHODOX RBPs: NEW PLAYERS IN THE VIRUS–HOST BATTLEFIELD

3.1 | The discovery of unorthodox RBPs

Classical RBDs do not account for all RNA-binding activities observed in the cell (Hentze et al., 2018). Initially, dozens of unorthodox RBPs were discovered in a somewhat sporadic manner. Further mechanistic studies demonstrated that these proteins have pivotal functions in RNA metabolism and gene expression (Arif et al., 2018; Castello, Hentze, & Preiss, 2015; Chang et al., 2013; Muckenthaler, Galy, & Hentze, 2008; Mukhopadhyay, Jia, Arif, Ray, & Fox, 2009). The scope of unconventional RBPs has been greatly expanded by RNA-IC, and other recently developed techniques have uncovered in a proteome-wide scale their modes of RNA binding (Box 1) (Castello et al., 2016; He et al., 2016; Kramer et al., 2014). A number of conclusions arose from these studies: (a) half of the RNA-binding regions within RBPs map to intrinsically disordered regions (IDRs) (Box 2); (b) globular domains with diverse known functions (e.g., protein–protein interaction domains, DNA-binding domains, and enzymatic cores) and protein folds with unknown function can also interact with RNA; (c) RBPs are hotspots for posttranslational modifications (PTMs), including phosphorylation, acetylation, and methylation, which are expected to regulate RBP activity; (d) novel RNA-binding regions are more evolutionary conserved than adjacent regions that

---

**BOX 1**

**IDENTIFICATION OF RNA-BINDING SITES BY RBDmap**

RBDmap is a novel system-wide method built on RNA-IC for the comprehensive in vivo identification of RNA-interacting sites within RBPs (Castello et al., 2016; Castello et al., 2017). Living cells are irradiated with UV light at 254 nm to induce covalent bonds between nucleotides and amino acids interacting directly (zero distance) in physiological conditions. After cell lysis, poly(A) RNAs are captured by hybridization with oligo(dT) magnetic beads under very stringent denaturing conditions. Because protein–RNA interactions are covalently preserved, noncovalent binders (i.e., protein–protein interactors) are removed and only direct RNA binders are purified together with the RNA. RBPs are digested with a protease that cleaves every 17 amino acids on average and poly(A) RNAs are purified via a second oligo(dT) capture. In this way, the protein regions still bound to RNA (termed “RBDpeps”) are isolated from the released peptides, which represent the protein regions that do not interact with RNA. RBDpeps are crosslinked to (at least) one nucleotide, making their identification by mass spectrometry challenging. To circumvent this problem, RBDpeps are digested with trypsin that cleaves every seven to nine amino acids on average, thus releasing a peptide still crosslinked to RNA and another peptide with native mass that is referred to as “neighboring peptide”. These latter peptides are readily identified by proteomics and the original RBDpep is reconstructed in silico. RNA-binding sites are identified by comparing the peptides released upon the protease treatment (released fraction) against those remaining bound to the RNA (RNA-bound fraction) (Castello et al., 2017). RBDmap revealed 1,174 high-confidence RNA-binding sites within 529 proteins in human cells (Castello et al., 2016) and 568 sites within 368 RBPs in murine cardiomyocytes (Liao et al., 2016). Forty-five percent of the identified RNA-binding sites mapped to well-established RBDS, while 55% mapped to previously unknown or poorly characterized RBDs and IDRs.

---

**BOX 2**

**INTRINSICALLY DISORDERED REGIONS IN RNA BINDING**

Approximately 25% of the experimentally identified cellular RBPs have classical RBDS in human and murine cell lines, and other 25% contain nonclassical but known RBDs (Castello et al., 2012; Castello et al., 2016). However, about half of the human and mouse RBPomes lack known RBDs (Castello et al., 2012; Castello et al., 2016; Liao et al., 2016). In many cases, the regions responsible for RNA binding are intrinsically disordered, lacking stable 3D structure and containing multiple interaction surfaces (Castello et al., 2016; Jarvelin, Noerenberg, Davis, & Castello, 2016). These RNA-binding IDRs are rich in serine (S), glycine (G), proline (P), lysine (K), arginine (R), and tyrosine (Y) amino acids, and their contribution to protein–RNA interactions has recently been reviewed (Jarvelin et al., 2016). Interestingly, IDRs show a modular organization in RBPs as is typical for classical RBPs: sequences of different length are repeated in a nonrandom mode, including RS dipeptides, RGG boxes, YGG boxes (also called [G/S]Y[G/S]), and poly(K) motifs, among others. In many cases, they can mediate both specific and nonspecific interactions with RNA in cooperation with globular RBDS. It is still unclear how specificity is generated by disordered RNA-binding regions.
do not interact with RNA; and (e) Mendelian mutations linked to disease and mapping to RBPs occur predominantly at RNA-binding disordered regions and PTM sites, highlighting their regulatory relevance (Castello et al., 2016; Castello, Fischer, Hentze, & Preiss, 2013; Moore, Jarvelin, Davis, Bond, & Castello, 2017).

3.2 Unorthodox globular RBPs in “virus-linked” RBPs

To analyze the scope of newly discovered RBPs linked to infection, we first removed all RBPs harboring a classical RBD. The resulting set of unconventional RBPs was then divided into two groups: (a) “nonclassical” RBPs if the protein lacks classical RBDs but harbors other domains described experimentally to bind RNA (Castello et al., 2012; Castello et al., 2016; File S1) and (b) “other” RBPs when lacking both classical and nonclassical RBDs. Ribosomal proteins were prevalent among nonclassical virus-linked RBPs (Figure 2a,b), which is expected as viral genomes cannot encode the large amount of protein and RNA components required to build this complex machinery. Another group of nonclassical RBPs linked to virus infection comprises protein chaperones from the heat shock protein (HSP)90, HSP70, and HSP20 families, known to harbor discrete RNA-binding regions conserved across homologous proteins (Castello et al., 2016; Moore et al., 2017). These proteins are important for the co-translational folding of proteins and the assembly and remodeling of cellular ribonucleoproteins (RNPs) (Iwasaki et al., 2010; Willmund et al., 2013). The involvement of protein chaperones in virus infection has been known for a long time; however, the relevance of their newly described RNA-binding activity for their recruitment and function remains unknown (Geller et al., 2012; Y. Wang et al., 2017). Peptidyl-prolyl cis–trans isomerases from the FKBPs and PPI families are represented in the noncanonical RBP group (Figure 2a,b). This finding expands the well-documented impact of PPIA on different viruses (D. Zhou, Mei, Li, & He, 2012) to other members of the PPI and FKBPs families. Proteins containing thioredoxin enzymatic domain, 14-3-3 protein–protein interaction domain, high mobility (HGMB) protein family, cytoskeleton-related proteins, and metabolic enzymes are also present in the group of nonclassical, virus-linked RBPs.

In the “other” group of virus-linked RBPs, we find proteins with clear roles in RNA biology such as translation initiation factors as well as proteins involved in splicing and RNA transport, but also kinases, metabolic enzymes, ubiquitin-binding proteins (including E3 ubiquitin ligases), proteasome components, and others (Figure 2c,d). Future work should focus on elucidating the poorly understood connections between virus infection and the RNA-binding capacity of these proteins.

3.3 Disordered regions in RBPs: Roles in virus infection?

Cellular RNP granules, such as stress granules and processing (P)-bodies, are complex aggregates of RNA and proteins that regulate RNA translation, turnover, and storage. Cytoplasmic granules have primarily an antiviral nature and many viruses inhibit their formation by targeting mainly G3BP stress granule assembly factor 1 (G3BP1), G3BP2, and EIF2AK2 (PKR) (Lloyd, 2013; McCormick & Khaperskyy, 2017; Onomoto et al., 2014; Yoneyama et al., 2016). On the other hand, many of the host proteins found in stress granules and/or P-bodies can be co-opted by viruses to enhance replication (Beckham & Parker, 2008; Reineke & Lloyd, 2013).

Transition from a soluble phase (dispersed proteins and RNA) to a condensed state (concentrated components) is responsible for RNP granule formation. Different lines of evidence have shown that IDRs within RBPs are important for this process to occur (Jarvelin et al., 2016). IDRs with specific sequence features (Box 2) have been shown to bind RNA (Basu & Bahadur, 2016; Castello et al., 2016; Jarvelin et al., 2016). In 170 out of 529 RBPs identified by RBDmap (Box 1) the only detectable RNA-binding sites were located at IDRs, suggesting that they are sufficient to mediate interactions with RNA (Castello et al., 2016). IDRs involved in RNA binding are present in both classical and unorthodox RBPs (Table S1), and many of these proteins display different roles in virus infection. For example, interleukin enhancer binding factor 3 (ILF3, 76% disordered) harboring a GQSY-rich region at the C-terminus (Figure 3a), G3BP1 (76% disordered), G3BP2 (80% disordered), HNRNPA1 (80% disordered), and cell cycle associated protein 1 (CAPRIN1, 86% disordered, Figure S1a) contain RNA-binding IDRs. The importance in virus infection of IDRs and their properties to bind RNA and promote phase transitions deserves further attention.

Strikingly, viruses mimic short disordered motifs to regulate key processes for virus infection (Davey et al., 2011). For example, dengue virus (DENV) and HCV IDRs are involved in protein–protein binding as well as in interaction with nucleic acids, metals, and other small ligands (Fan et al., 2014; Meng et al., 2015). These viral IDRs are phosphorylation hotspots with regulatory roles. The control of protein–RNA interactions and viral replication by host kinases phosphorylating viral proteins has been reviewed by Jakubiec and Jupin (2007) and Keck, Atya, Amaya, Bailey, and Narayanan (2015). Several viral RBPs interact with RNA through disordered regions, including HIV Tat and Rev that harbor an arginine-rich motif with nucleic acid-binding capacity (Jarvelin et al., 2016; Xue, Mizianty, Kurgan, & Uversky, 2012). These short, disordered peptides co-fold with the target RNA forming stable structures (Daugherty, Liu, & Frankel, 2010; Puglisi, Tan, Calnan, Frankel, & Williamson, 1992). Other arginine-rich RNA-binding motifs have been described in Ebola virus (EBOV) VP30 (John et al., 2012).
FIGURE 2  Virus-linked RNA-binding proteins (RBPs) with nonclassical or other RNA-binding domains (RBDs). (a and c) Word cloud of protein domains present in nonclassical (a) and other (c) virus-linked RBPs. (b and d) STRING protein network showing connections between nonclassical (b) and other (d) virus-linked RBPs. ER, endoplasmic reticulum; F-bP, fructose-biphosphate; PPI, peptidyl-prolyl cis–trans isomerase.
2007), Flaviviridae capsid (Ivanyi-Nagy, Lavergne, Gabus, Ficheux, & Darlix, 2008), and Nodamura virus RNA-dependent RNA polymerase (Gitlin, Hagai, LaBarbera, Solovey, & Andino, 2014). Short motifs confer viral proteins the ability to interact with many different viral and host components despite the limited coding capacity of viral genomes (Xue et al., 2014). This is a cost-effective RNA (and protein)-binding solution that requires only a dozen amino acids (Jarvelin et al., 2016; Phan et al., 2011), while the most common classical RBDs (e.g., RRM) contain 75–85 amino acids (Lunde et al., 2007).

4 | FUNCTIONS OF NOVEL RBPs IN VIRUS INFECTION

4.1 | Ubiquitination routes in antiviral responses mediated by tripartite motif proteins

Virus infection activates diverse immune signaling pathways involving many RBPs that are highly regulated through PTMs such as phosphorylation or ubiquitination (Chiang & Gack, 2017; Y. Zhou, He, Wang, & Ge, 2017). Ubiquitination of antiviral sensors and effectors can modulate protein function in many ways and requires three different classes of enzymes: ubiquitin-activating (E1), -conjugating (E2), and -ligating (E3). Tripartite motif (TRIM) is a large family of proteins that harbor E3 ligase activity and restrict virus infection through direct and indirect mechanisms (van Tol, Hage, Giraldo, Bharaj, & Rajsbaum, 2017; Versteeg et al., 2013). TRIM proteins have broad roles in the cell, but these were not linked to RNA biology until recently. Among the upward of 80 TRIM family members, four (TRIM25, TRIM28, TRIM56, and TRIM71) were found to interact directly with RNA in living cells by RNA-IC experiments (Hentze et al., 2018). The RNA-binding activity of TRIM25 and TRIM71 was further validated by orthogonal approaches in mouse embryonic stem cells (Kwon et al., 2013). TRIM25 carries the critical function of ubiquitinating RIG-I, a cellular receptor that recognizes triphosphate ends in vRNAs and triggers the type I interferon response (Gack et al., 2007). The high-precision proteome-wide method to identify RBDs, RBDmap (Box 1), revealed that TRIM25 interacts with RNA via its PRY/SPRY domain (Figure 3b) (Castello et al., 2016), which had been previously implicated in protein–protein interactions. This finding was confirmed and further characterized by orthogonal in vitro and in vivo methods (Choudhury et al., 2017). Importantly, this study showed that the E3-ubiquitin ligase activity of TRIM25 is regulated by RNA binding. For example, in vitro experiments revealed that auto-ubiquitination of TRIM25 and ubiquitination of other targets such as zinc finger CCCH-type containing antiviral 1 (ZC3HAV1 or ZAP) is impaired in mutants lacking the RNA-binding region identified by RBDmap, or in the presence of RNases. DENV2 was recently shown to prevent TRIM25 antiviral potential through an ingenious mechanism (Manokaran et al., 2015). The exonuclease XRN1 degrades the DENV2 genomic RNA up to a pseudoknot that stalls it. This produces an abundant noncoding...
subgenomic flavivirus RNA (sfRNA) with 5′ monophosphate end, and thereby it is not recognized by intracellular sensors. This sfRNA is bound (directly or indirectly) by TRIM25 and sequesters it, thus reducing RIG-I ubiquitination and the subsequent stimulation of type I interferon synthesis (Manokaran et al., 2015). ZC3HAV1 is a broad but not universal antiviral RBP that promotes translation inhibition and degradation of target RNAs (Todorova, Bock, & Chang, 2015). TRIM25 can ubiquitinate ZC3HAV1, but this is not directly responsible for the increase in ZC3HAV1 antiviral activity against Sindbis virus (SINV) (M. M. Li et al., 2017). The interaction of TRIM25 with other host factors (mainly involved in RNA metabolism [Choudhury et al., 2017]) and/or their ubiquitination may contribute to ZC3HAV1 antiviral effects. Interestingly, the RNA-binding SPRY domain of TRIM25 is critical for the interaction with ZC3HAV1 (M. M. Li et al., 2017) possibly through the formation of an RNA bridge. Many other RBPs (including ZC3HAV1) have been proven to interact with TRIM25 in an RNA-dependent manner (Choudhury et al., 2017). Therefore, RNA may not only modulate the E3 ligase activity of TRIM25 but could also serve as a scaffold to concentrate substrates in the proximity of TRIM25-binding sites. Other viral strategies to usurp the host ubiquitylation and SUMOylation machineries have been reviewed by Wimmer and Schreiner (2015).

TRIM56 was initially linked to the host response against DNA viruses, contributing to the cGAS-STING pathway in a ubiquitin-ligase-dependent manner (Tsuchida et al., 2010). TRIM56 also regulates the TLR3/TRIF signaling pathway for dsRNA sensing but, in this case, its ubiquitin ligase activity is not required (Shen et al., 2012). Furthermore, TRIM56 can specifically restrict several RNA viruses through diverse antiviral mechanisms (Kane et al., 2016; Liu et al., 2016; Liu et al., 2014; J. Wang et al., 2011). Distinct regions of TRIM56 have been linked to different viruses and stages of the viral cycle: (a) both the N-terminal E3 ubiquitin ligase and the integrity of the C-terminus are required for the suppression of bovine viral diarrhea virus (BVDV), yellow fever virus, and DENV2 RNA replication; (b) the E3 ubiquitin ligase activity is sufficient to inhibit late steps (i.e., packaging and/or egress) of human coronavirus infection; and (c) the C-terminus suffices to impede influenza A virus (IAV) and influenza B virus (IBV) RNA synthesis. The C-terminus of TRIM56 has no known domain architecture, but it is known to interact with different cellular (STING and TRIF) and viral (BVDV N-protease) proteins. TRIM56 C-terminal part is also believed to function as a scaffold for the formation of multiprotein complexes that are essential for TLR3 signaling (Shen et al., 2012). The recent discovery that TRIM56 also interacts with RNA in living cells suggests that its E3 ubiquitin ligase activity may be regulated by RNA (Hentze et al., 2018), in analogy to TRIM25 (Choudhury et al., 2017). However, TRIM56 lacks the PRY/SPRY domain present in TRIM25 and thus must employ a different yet unknown RBD. TRIM56 antiviral activities in BVDV, IAV, and IBV infections rely on the last 130 amino acids of the C-terminal part. This sequence has been proposed to interact with vRNA through conserved residues at the NHL repeats (Liu et al., 2016; J. Wang et al., 2011). However, the proposed ability of the C-terminus of TRIM56 to interact with RNA should be experimentally assessed.

4.2 | Recognition of vRNA 5′ ends by interferon-induced proteins

The minimal cellular cap structure, named “cap0”, consists of a N7-methylated guanosine (m7G) linked to the 5′ end of the mRNA by a 5′ triphosphate bond. This structure controls mRNA export, maturation, stability, and translation. The first and second nucleotides of cellular mRNAs are further methylated at the 2′-O positions, leading to the formation of “cap1” and “cap2”. Several antiviral sensors interact with the 5′ end of vRNAs, which usually bear unusual cap structures (Decroly et al., 2011) or triphosphate ends, in contrast to uncapped host RNAs that have 5′ monophosphate ends. Incubation of cellular extracts with immobilized RNA probes harboring 5′ triphosphate ends coupled to proteomics led to the identification of a family of proteins with high affinity for this unusual molecular pattern, called IFIT1-5 or interferon-induced protein with tetratricopeptide repeats (TPRs) (Pichlmair et al., 2011). The expression of these proteins is triggered rapidly after activation of the host antiviral defenses. IFIT proteins target nonself, single-stranded (ss)RNAs in a 5′-end-dependent manner (sensing 5′ triphosphate and cap0 RNAs) to inhibit infection by different mechanisms, mainly by affecting viral translation (Habjan et al., 2013; Kumar et al., 2014). The unconventional RNA-binding activity of IFIT1, 2, and 5 has been recently unraveled and resides in the TPRs, which usually mediate protein–protein interactions. These repeats generate an α-helix-rich fold with a positively charged, small pocket that, together with the participation of acidic residues that coordinate divalent cations, embraces the negatively charged triphosphate ends (Abbás, Pichlmair, Gorna, Superti-Furga, & Nagar, 2013). The different proteins of the IFIT family appear to have developed different roles and modes of action. Both IFIT1 and IFIT5 interact only with ssRNA. IFIT5 binds RNA with cap0, 5′ monophosphate, or 5′ triphosphate ends with similar affinities and do not interact with cap1 (Katibah et al., 2014). IFIT1 forms part of a multiprotein complex together with IFIT2 and IFIT3 (Pichlmair et al., 2011) and can bind cap0 with very high affinity in addition to 5′ triphosphate RNAs (with lower affinity than IFIT5) (Kumar et al., 2014). Functionally, IFIT1 can outcompete the binding of EIF4F complex to the cap0, thus inhibiting viral translation initiation, while IFIT5 can inhibit 48S translation initiation complex formation on mRNAs with 5′ triphosphate ends (Fleith et al., 2018; Kumar et al., 2014). IFIT1 and IFIT2 can also block viral protein synthesis by interacting with the EIF3 complex (Diamond & Farzan, 2013; Vladimer et al., 2014), which has a central role in translation initiation of cellular and viral
transcripts (Cate, 2017). IFIT3 has no known RNA-binding capacity, but it has been recently described to allosterically modulate the IFIT1 RNA-binding pocket thus enhancing its affinity for cap0 (Fleith et al., 2018; Johnson et al., 2018). IFIT2 is an RBP that can selectively bind AU-rich RNAs and might recognize specific sequences in vRNAs (Yang et al., 2012). Some viruses such as flavivirus, poxvirus, and coronavirus can escape from IFIT1 and IFIT2 by methylating their vRNA 5′ cap (Daffis et al., 2010). Other vRNAs lack a 2′-O methylated cap but can still evade IFIT1 restriction. For example, a stable stem-loop structure located in the 5′ untranslated region (UTR) of the genomic RNA of alphaviruses antagonizes IFIT1 antiviral activity (Hyde et al., 2014).

4.3 | The multiple pro- and antiviral roles of interleukin enhancer factors

Interleukin enhancer-binding factor 3 (ILF3) is a multifunctional protein with broad roles in cellular and viral RNA metabolism, including regulation of transcription and translation, mRNA stabilization and localization, and noncoding RNA biogenesis (Castella, Bernard, Corno, Fradin, & Larcher, 2015; Patino, Haenni, & Urcuqui-Inchima, 2015). ILF3 is a well-established RBP that harbors three RBDS: two DSRM in positions 417–478 and 540–601, and an intrinsically disordered RGG box in position 653–669 that binds single-stranded RNA and DNA (Figure 3a). These RBDS cooperate in an intricate dynamic mechanism to recognize diverse RNA substrates (Schmidt et al., 2016). ILF3 can act alone as homodimer or form an RNA regulatory complex with ILF2 and other factors, promoting or restricting virus infection depending on the context. ILF3/ILF2 heterodimerization is mediated by the domain associated with ZF (DZF), which is a nucleotidyltransferase domain that lacks enzymatic activity (Wolkowicz & Cook, 2012). RBDmap identified a novel RNA-binding activity within this DZF domain, which is present in both ILF3 and ILF2 (Figure 3a,c) and was further validated by orthogonal assays (Castello et al., 2016). However, the target RNAs and the biological importance of this interaction are still unknown. PKR is a dsRNA-dependent protein kinase that phosphorylates the threonine (Thr)188 and Thr315 of ILF3, inducing ILF3/ILF2 complex disassembly and shuttling of ILF3 to the cytoplasm. Once there, ILF3 binds to viral transcripts and blocks viral translation (Harashima, Guettouche, & Barber, 2010). Thr315 resides in the novel RNA-binding site, but this is also a protein-binding interface; therefore, it is still uncertain which of these two activities is affected by Thr315 phosphorylation. One possibility is that disruption of the ILF3/ILF2 complex (or the interactions with other proteins) affects vRNA recognition. In this regard, ILF2 can modulate in vitro the RNA-binding specificity and efficiency of ILF3 (Schmidt, Friedrich, Golbik, & Behrens, 2017; Schmidt, Knick, et al., 2017), and this may also occur in cells. ILF3 and ILF2 may have additional roles in RNA metabolism as homodimers. Accordingly, ILF2 binds to AU-rich IRES and enhances translation of cellular and viral RNAs (Faye et al., 2013; Graber, Baird, Kao, Mathews, & Holcik, 2010). The IRES trans-acting factor (ITAF) activity of ILF2 does not involve ILF3 and, in fact, it is even possible that ILF3 reduces ILF2 ITAF activity when forming the heterodimer (Graber et al., 2010). ILF3 associates preferentially with 25–30 nt AU-rich motifs present at the 3′ UTRs of cellular RNAs, blocking translation (Kuwano et al., 2010). Strikingly, ILF2 and ILF3, together with the helicase DHX9, bind to AU-rich sequences present in the 5′ and 3′ UTRs of HCV and BVDV RNAs triggering their circularization in vitro (Isken et al., 2007). These proteins are recruited to HCV replication complexes in infected cells. Deletion of their binding sites in HCV UTRs completely blocks viral replication, suggesting that the circularization favors infection. In a reconstituted in vitro system, ILF2 interaction with ILF3 enhances ILF3 RNA chaperone activity and stimulates the first step of HCV RNA synthesis mediated by the viral polymerase NS5B (Schmidt, Friedrich, et al., 2017). In the related BVDV, ILF2/ILF3-mediated circularization of the vRNA may also be important for IRES translation (Isken et al., 2003; Isken, Grassmann, Yu, & Behrens, 2004). Finally, ILF3 and ILF2 promote HIV infection by interacting directly with vRNA. Deletion of the DSRMs of ILF3 or the N-terminal RGG domain of ILF2 disrupts binding to HIV RNA (T. Li et al., 2016), thus confirming their importance for RNA binding (Figure 3a,c). It is believed that ILF3 stabilizes vRNA while ILF2 upregulates HIV gene expression without affecting RNA stability, but the exact mechanism is still not well documented.

There are two main isoforms of ILF3, NF90 and NF110. NF110 contains an extended C-terminus that is intrinsically disordered and is enriched in GQSY motifs that mediate protein–protein interactions and increase the RNA-binding activity of the DSRMs (Reichman & Mathews, 2003). RBDmap assigns RNA-binding activity to the GQSY region (Figure 3a), which suggests that NF110 may have different RNA-binding properties than the NF90 isoform. Recent reports showed that both NF90 and NF110 promote RNA circularization in the nucleus and associate with mature circular RNAs in the cytoplasm (X. Li et al., 2017). Circular RNAs have regulatory roles in the antiviral immune response and are generated by back-splicing from exons with flanking complementary Alu repeats. Both NF90 and NF110 bind dsRNA formed by inverted Alu pairing; however, NF110 also interacts with single-stranded Alu RNA (X. Li et al., 2017). NF90 binds preferentially to circularized RNA compared to linear RNA with the same sequence, while NF110 strongly associates with both. NF90/NF110 binds vescicular stomatitis virus RNAs in the cytoplasm to inhibit viral replication, and simultaneously causes a reduction in host circular RNAs (X. Li et al., 2017). The specificity (if any) of the GQSY-rich region of NF110 and the biological importance of this region in infected cells still requires further study.
ILF3 and ILF2 have also been implicated in the regulation of stress granules. The GQSY-rich region of the NF110 isoform of ILF3 interacts with CAPRIN1 (Figure 3a and Figure S1a) and promotes stress granule formation. This region is similar in sequence and functionally exchangeable with the IDR present in RBP FUS that can trigger RNP aggregation (Shiina & Nakayama, 2014). This region might be sufficient but not essential to promote the formation of RNA granules, because the short ILF3 isoform (i.e., NF90) that lacks it can also induce the assembly of stress granules in response to dsRNA (Wen et al., 2014). Furthermore, ILF proteins directly influence the fate of viral proteins. For example, the nucleoprotein of IAV is sequestered in stress granules by a mechanism that involves NF90, and this harms viral replication. This process is antagonized by the multifunctional IAV NS1 protein (Khaperskyy, Hatchette, & McCormick, 2012; Mok et al., 2012).

4.4 Hijacking the ribosome: The EIF3 complex

The cap structure and the poly(A) tail, present in most cellular mRNAs, are recognized by EIF4E and PABPC1, respectively. These two proteins interact with EIF4G to form a closed loop between the 5′ and 3′ ends of the mRNA. To recruit ribosomes, EIF4G interacts with a complex formed by 13 subunits, called EIF3, which serves as a molecular bridge between the ribosomes and the RNA (Hinnebusch, 2014). Interestingly, virtually all EIF3 subunits have been linked to virus infection.

While EIF3B and EIF3G harbor RRMs, other subunits employ unconventional modes of RNA binding (Table S1 and Figure S1b–d). For example, RBDmap revealed a region between the amino acids 112–131 of EIF3D that interact with RNA (Figure S1b), in agreement with previous in vitro experiments (Asano et al., 1997). Strikingly, HIV protease cleaves EIF3D in positions 114–115, and the knockdown of this EIF3 subunit, but not others, increases HIV infectivity (Jager et al., 2011). The latter study proposed that EIF3D binds the vRNA and inhibits reverse transcription, and that this antiviral mechanism is counteracted by cleaving EIF3D (Jager et al., 2011). Notably, it has been recently described that EIF3D is a novel cap-binding protein that interacts with a specific pool of cellular mRNAs that do not employ the EIF4F complex to initiate translation (Lee, Kranzusch, Doudna, & Cate, 2016). The translation mechanism proposed implicates both a specific RNA structure that opens the EIF3D cap-binding pocket and another RNA element that prevents EIF4F recruitment. The possibility that similar structures are present in vRNAs has been proposed, but requires experimental confirmation (Carrasco, Sanz, & Gonzalez-Almela, 2018). SINV subgenomic mRNA follows a noncanonical translation initiation mechanism that employs a cap-dependent “scanning” but lacks the participation of several EIFs, including the EIF4E, EIF4A, and EIF2α, and intact EIF4G (Castello, Sanz, Molina, & Carrasco, 2006; Garcia-Moreno, Sanz, & Carrasco, 2015; Garcia-Moreno, Sanz, Pelletier, & Carrasco, 2013). A stem-loop structure present at the 5′ end of this mRNA is required for efficient translation (McInerney, Kedersha, Kaufman, Anderson, & Liljestrom, 2005; Ventoso et al., 2006). Thus, novel cap-binding proteins such as EIF3D may replace the EIF4F complex and promote viral translation after activation by the hairpin structure (Carrasco et al., 2018).

Other vRNAs lack a 5′ cap structure, but instead a viral protein (VPg) is covalently linked to the 5′ end of the RNA. The calicivirus VPg interacts in vitro and in vivo with EIF3D as well as with other EIFs and ribosomal proteins (Chung et al., 2014; Daughenbaugh, Fraser, Hershey, & Hardy, 2003; Daughenbaugh, Wobus, & Hardy, 2006). The role of EIF3D in calicivirus translation is not yet well understood, but this interaction points toward its importance in cap-independent translation initiation by binding viral proteins covalently linked to the 5′ end of vRNAs.

4.5 Utilizing the cellular transport network: Viruses and dynein

4.5.1 RNA transport mediated by the dynein molecular motor

RNA localization over long distances often involves the active bidirectional transport along microtubules by molecular motors, mainly dynein and kinesin proteins and adapter proteins (Marchand, Gaspar, & Ephrussi, 2012; Tekotte & Davis, 2002). Recently, RNA-IC experiments in different cells and organisms revealed that numerous dynein and kinesin proteins are endowed with RNA-binding activity (Table S1) (Hentze et al., 2018), and RBDmap in human cells identified potential RNA-binding sites in some members of the dynein family (Figure S2a–e). Interestingly, Drosophila dynein light chain 1 (highly conserved in humans) can strongly bind in vitro to a particular region in the 3′ UTR of gurken RNA, and mediates its correct localization during oogenesis in vivo (Rom, Faicevici, Almog, & Neuman-Silberberg, 2007). Active transport has also been linked to P-body and stress granule formation in mammalian and Drosophila cells because silencing of dynein cytoplasmic 1 heavy chain 1 (DYNC1H1) leads to a dramatic impairment of their formation (Loshi, Leishman, Berardone, & Boccaccio, 2009). Most viruses can subvert cellular transport motors. For example, dyneins and kinesins facilitate uncoating, movement of preintegration complexes, and transport of retrovirus proteins (e.g., Gag and Env) as well as the movement of the vRNA to the perinuclear region for virion assembly (Arriagada, 2017).

Movement of newly synthesized retroviral RNA throughout the cytoplasm may require the dynein motor complex, although results are conflicting. Some studies suggest that, after nuclear export, the HIV unspliced RNA moves toward the microtubule organizing center (MTOC) where Gag interacts with the psi packaging signal. Gag also associates with
endoosomes that are transported to the virion assembly sites. The dynein complex, including members such as DYNC1H1 and dynactin subunit 2 (DCTN2), would mediate the transport of both the vRNA to the MTOC and the vesicles containing viral Gag/RNA RNP (Lehmann et al., 2009; Molle et al., 2009). However, another study suggests that HIV unspliced RNA traffics throughout the cytoplasm by diffusion after leaving the nucleus en masse through a Rev-dependent mechanism (Pocock, Becker, Swanson, Ahlquist, & Sherer, 2016). Further studies are required to clarify the vRNA transport mechanisms followed by retroviruses and the implications of the newly discovered RNA-binding activity of dyneins in this process.

4.5.2 The role of dyneins in HIV uncoating and reverse transcription

HIV uncoating is an incompletely understood multistep process during which the capsid shell enclosing the genome is disassembled. Recent studies suggest that dyneins promote HIV uncoating and this process is linked to the intracellular movement toward the nucleus (Lukic, Dharan, Fricke, Diaz-Griffero, & Campbell, 2014; Pawlica & Berthoux, 2014). Accordingly, DYNC1H1 knockdown or disruption of the dynein complex by overexpressing DCTN2 interferes with uncoating (Pawlica & Berthoux, 2014). Whether dynein-mediated uncoating affects reverse transcription is unclear, because microtubule network disruption with nocodazole does not inhibit HIV reverse transcription (Lukic et al., 2014), and perturbation of the dynein complex only causes a transient defect in reverse transcription (Pawlica & Berthoux, 2014). A different study showed that cytoplasmic dynein light chain 1 (DYNLL1), acting outside the microtubule-associated dynein complex, is required for both HIV uncoating and reverse transcription (Jayappa et al., 2015).

The novel RNA-binding activities of several dynein and adapter proteins such as DYNC1H1, dynein cytoplasmic 1 intermediate chain 2 (DYNC1I2), dynein cytoplasmic 1 light intermediate chain 1 (DYNCL1I1), and DCTN2 may contribute to both early and late events of HIV infection by binding directly to vRNA. In agreement, DYNC1H1 was identified in a recent approach aimed at deciphering the in vivo HIV RNA interactome (Knoener, Becker, Scalf, Sherer, & Smith, 2017).

4.5.3 Involvement of dyneins in viral replication factory assembly

Most cytoplasmic viruses induce the remodeling of intracellular membranes to assemble stable compartments (referred to as “viral factories”) for vRNA replication, translation, and packaging protected from host antiviral defenses. In HCV-infected cells, the viral core protein directs a dynein-mediated redistribution of lipid droplets around the MTOC (Boulant et al., 2008). HCV nonstructural protein 5A (NS5A) is essential for the regulation of vRNA synthesis and the formation of endoplasmic reticulum (ER)-derived viral factories. Newly synthesized NS5A associates with core-coated lipid droplets and facilitates viral genome transfer, thus linking HCV replication with virion assembly (Eyre et al., 2014). Interestingly, microtubule polymerization inhibitors and DYNC1H1 knockdown impair NS5A movement and HCV RNA replication (Eyre et al., 2014). The dynein complex may also be involved in the movement of DENV2 proteins (and possibly vRNA) to the ER and from the ER to the Golgi apparatus for viral replication and virion assembly (Shrivastava, Sripada, Kaur, Shah, & Cecilia, 2011). Moreover, dynein complex disruption by DCTN2 overexpression and depolymerization of microtubules with specific drugs affect the distribution of viral nucleocapsid and inhibit vRNA synthesis in Hantavirus-infected cells (Ramanathan et al., 2007). The roles of dyneins and their RNA-binding activities in viral factory biogenesis should be assessed in the future.

4.6 RBPs resident in the ER affect virus infection

The protein transport protein Sec31A (SEC31A) is a component of the coat protein complex II, which mediates the budding of transport vesicles from the ER, and is endowed with RNA-binding activity (Figure 4a) as identified by RNA-IC and RBDmap (Castello et al., 2016; Hentze et al., 2018). SEC31A redistributes from the perinuclear region to the entire cytoplasm and is degraded as poliovirus infection progresses, suggesting that it may have an important antiviral role (Richards, Soares-Martins, Riddell, & Jackson, 2014). Future investigations will shed light on the role of SEC31A and its RNA-binding activity in virus infection.

Several proteins of the ER translocation machinery such as Sec61 translocon alpha 1 subunit (SEC61A1), Sec61 translocon beta subunit (SEC61B, Figure 4b), SEC62 homolog preprotein translocon factor (SEC62, Figure 4c), or SEC63 homolog protein translocation regulator (SEC63) were described to bind RNA in different RNA-IC experiments (Hentze et al., 2018). Orthogonal proteomic studies aimed at identifying the native ER-associated RNA interactome confirmed that several ER integral membrane proteins have noncanonical RNA-binding activity, including SEC61A1 and SEC61B (SEC61 complex) (Jagannathan et al., 2014). They are involved in a novel mechanism that selectively anchors mRNAs encoding endomembrane proteins to the ER (Jagannathan et al., 2014). The SEC61 complex may play multiple roles in virus infection. Zika virus promotes the formation of cytoplasmic vacuoles, and viral RNA and E protein (that contains a transmembrane domain) accumulate in their borders (Monel et al., 2017). Inhibition of SEC61A1 with specific drugs impairs both vacuole formation and virus production. The SEC61 complex is also essential for glycoprotein proteostasis of IAV hemagglutinin and neuraminidase proteins and HIV gp120 protein, as well as for DENV RNA replication (Heaton et al., 2016). SEC61A1 also interacts and co-
localizes with the membrane-associated VP24 protein of EBOV, and is proposed to regulate transcription and replication of the viral genome (Iwasa et al., 2011). The SEC61B subunit, endowed with noncanonical RNA-binding capacity (Figure 4b), was also identified as an important factor for flavivirus replication in siRNA (Savidis et al., 2016) and CRISPR screens (Zhang et al., 2016). Whether the SEC61 complex anchors vRNA directly to ER-derived membranous compartments, as described for specific cellular mRNAs (Jagannathan et al., 2014), remains unknown.

The dimeric SEC62/SEC63 complex interacts with the SEC61 complex and acts in posttranslational translocation into the ER. SEC62 novel RNA-binding site maps to the disordered cytosolic N-terminal part (Figure 4c), while SEC63 RBD is still unknown. SEC62 abundance increases after FMDV infection and its depletion reduces FMDV titre (Guo et al., 2015). Moreover, SEC63 knockout decreases virus yield in both mammalian and insect cells infected with different flaviviruses (Zhang et al., 2016). Although SEC62 and SEC63 are important for viral fitness, their exact roles in infection and the potential of their RNA-binding activity to promote or restrict viral replication remain unknown.

5 | OUTLOOK

RBPs are key regulators of cellular pathways that respond to infection or are hijacked by viruses. Functions of many classical RBPs have been characterized in the context of infection. However, recent RNA-IC experiments have expanded the census of RBPs and identified unanticipated modes of RNA binding (Hentze et al., 2018). About 200 unorthodox RBPs have been linked to virus infection or immunity but, in many cases, their RNA-binding properties had remained unnoticed. In this review, we have discussed several examples of unconventional RBPs that either support viral replication or mediate the host antiviral response. Most virus-linked, unorthodox RBPs remain poorly understood. Importantly, many of these newly discovered RBPs have enzymatic activities such as chaperone, kinase, isomerase, E3 ubiquitin ligase, and others, and are therefore susceptible to inhibition by small compounds in host-based antiviral therapies. Further research should aim at characterizing their roles in infection to better understand the interplay between unorthodox RBPs and viruses.

System-wide approaches to discover the vRNA–host RBP landscape have contributed to our understanding of the role of host RBPs in infected cells (Knoener et al., 2017; LaPointe, Gebhart, Meller, Hardy, & Sokoloski, 2018; Lenarcic, Landry, Greco, Cristea, & Thompson, 2013; Phillips, Soderblom, Bradrick, & Garcia-Blanco, 2016; Viktorovskaya, Greco, Cristea, & Thompson, 2016). Although the currently available methods have technical limitations, we are confident that global analyses of protein–RNA interactions can determine the arsenal of RBPs deployed in the host–virus battlefield. Such approaches must be complemented with techniques such as crosslinking immunoprecipitation followed by sequencing (CLIP-seq) and its variants (Wheeler, Van Nostrand, & Yeo, 2018) to determine in a global scale the footprints of RBPs on viral and host RNAs. We believe that there is a whole universe of host–virus interactions to be discovered that will not only further our understanding of virus and cell biology, but also lead to the development of novel antiviral approaches.
ACKNOWLEDGMENTS
A.C. is funded by MRC Career Development Award (number MR/L019434/1) and John Fell funds from the University of Oxford. M.G.M. is funded by the European Union’s Horizon 2020 Research and Innovation programme under the Marie-Skłodowska-Curie Grant agreement number 700184.

CONFLICT OF INTEREST
The authors have declared no conflicts of interest for this article.

RELATED WIREs ARTICLE
Advances in the characterization of RNA-binding proteins

ORCID
	Manuel Garcia-Moreno http://orcid.org/0000-0001-7181-5459
	Aino I. Järvelin http://orcid.org/0000-0002-1225-4396
	Alfredo Castello http://orcid.org/0000-0002-1499-4662

REFERENCES

Abbas, Y. M., Pichlmair, A., Gorna, M. W., Superti-Furga, G., & Nagar, B. (2013). Structural basis for viral 5'-PPP-RNA recognition by human IFIT proteins. Nature, 494(7435), 60–64. https://doi.org/10.1038/nature11783

Anantharaman, V., Koonin, E. V., & Aravind, L. (2002). Comparative genomics and evolution of proteins involved in RNA metabolism. Nucleic Acids Research, 30(7), 1427–1464.

Arif, A., Yao, P., Terenzi, F., Jia, J., Ray, P. S., & Fox, P. L. (2018). The GAIT translational control system. WIREs RNA, 9(2), e1441. https://doi.org/10.1002/wrna.1441

Arriagada, G. (2017). Retroviruses and microtubule-associated motor proteins. Cellular Microbiology, 19(9), e12759. https://doi.org/10.1111/cmi.12759

Asano, K., Vornlocher, H. P., Richter-Cook, N. J., Merrick, W. C., Hinnebusch, A. G., & Hershey, J. W. (1997). Structure of cDNAs encoding human eukaryotic initiation factor 3 subunits. Possible roles in RNA binding and macromolecular assembly. The Journal of Biological Chemistry, 272(43), 27042–27052.

Au, H. H., & Jan, E. (2014). Novel viral translation strategies. WIREs RNA, 5(6), 779–801. https://doi.org/10.1002/wrna.1246

Balisterri, G., Bogmann, C., & Mühlmann, O. (2017). Virus escape and manipulation of cellular nonsense-mediated mRNA decay. Viruses, 9(1), 24. https://doi.org/10.3390/v9010024

Baltimore, D. (1971). Expression of animal virus genomes. Bacteriological Reviews, 35(3), 235–241.

Baltz, A. G., Munschauer, M., Schwanhausser, B., Vasile, A., Murakawa, Y., Schueler, M., … Barton, G. M. (2011). Nucleic acid recognition by the innate immune system. Annual Review of Immunology, 29, 185–214. https://doi.org/10.1146/annurev-immunol-031210-101340

Barbalat, R., Ewald, S. E., Mouchess, M. L., & Barton, G. M. (2011). Nucleic acid recognition by the innate immune system. Annual Review of Immunology, 29, 185–214. https://doi.org/10.1146/annurev-immunol-031210-101340

Basu, S., & Bahadur, R. P. (2016). A structural perspective of RNA recognition by intrinsically disordered proteins. Cellular and Molecular Life Sciences, 73(21), 4075–4084. https://doi.org/10.1007/s00018-016-2283-1

Beachboard, D. C., & Horner, S. M. (2016). Innate immune evasion strategies of DNA and RNA viruses. Current Opinion in Microbiology, 32, 113–119. https://doi.org/10.1016/j.mib.2016.05.015

Beckham, C. J., & Parker, R. (2008). P bodies, stress granules, and viral life cycles. Cell Host & Microbe, 3(4), 206–212. https://doi.org/10.1016/j.chom.2008.03.004

Boulant, S., Douglas, M. W., Moody, L., Budkowska, A., Targett-Adams, P., … McLauchlan, J. (2008). Hepatitis C virus core protein induces lipid droplet redistribution in a microtubule- and dynein-dependent manner. Traffic, 9(8), 1268–1282. https://doi.org/10.1111/j.1600-0854.2008.00767.x

Broder, S. (2010). The development of antiretroviral therapy and its impact on the HIV-1/AIDS pandemic. Antiviral Research, 85(1), 1–18. https://doi.org/10.1016/j.antiviral.2009.10.002

Carrasco, L., Sanz, M. A., & Gonzalez-Almela, E. (2018). The regulation of translation in alphanvirus-infected cells. Viruses, 10(2), 70. https://doi.org/10.3390/v10020070

Castella, S., Bernard, R., Corno, M., Fradin, A., & Larcher, J. C. (2015). Iif3 and NF90 functions in RNA biology. WIREs RNA, 6(2), 243–256. https://doi.org/10.1002/wrna.1270

Castello, A., Alvarez, E., & Carrasco, L. (2011). The multifaceted poliovirus 2A protease: Regulation of gene expression by picornavirus proteases. Journal of Biomedicine & Biotechnology, 2011, 369648. https://doi.org/10.1155/2011/369648

Castello, A., Fischer, B., Eichelkaer, K., Horos, R., Beckmann, B. M., Strein, C., … Hentze, M. W. (2012). Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell, 149(6), 1393–1406.

Castello, A., Fischer, B., Frese, C. K., Horos, R., Alleaume, A. M., Foehr, S., … Hentze, M. W. (2016). Comprehensive identification of RNA-binding domains in human cells. Molecular Cell, 63(4), 696–710. https://doi.org/10.1016/j.molcel.2016.06.029

Castello, A., Fischer, B., Horos, R., & Preiss, T. (2013). RNA-binding proteins in Mendelian disease. Trends in Genetics, 29(5), 318–327.

Castello, A., Frese, C. K., Fischer, B., Jarvelin, A. L., Horos, R., Alleaume, A. M., … Hentze, M. W. (2017). Identification of RNA-binding domains of RNA-binding proteins in cultured cells on a system-wide scale with RBDmap. Nature Protocols, 12(12), 2447–2464. https://doi.org/10.1038/nprot.2017.106

Castello, A., Hentze, M. W., & Preiss, T. (2015). Metabolic enzymes enjoying new partnerships as RNA-binding proteins. Trends in Endocrinology and Metabolism, 26(12), 746–757. https://doi.org/10.1016/j.tem.2015.09.012

Castello, A., Sanz, M. A., Molina, S., & Carrasco, L. (2006). Translation of Sindbis virus 26S mRNA does not require intact eukaryotic initiation factor 4G. Journal of Molecular Biology, 355(5), 942–956. https://doi.org/10.1016/j.jmb.2005.11.024

Cate, J. H. (2017). Human eIF3: From ‘biodology’ to biological insight. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 372(1716), 20160176. https://doi.org/10.1098/rstb.2016.0176
