The new (dis)order in RNA regulation

Aino I. Järvelin, Marko Noerenberg, Ilan Davis and Alfredo Castello*

Abstract

RNA-binding proteins play a key role in the regulation of all aspects of RNA metabolism, from the synthesis of RNA to its decay. Protein-RNA interactions have been thought to be mostly mediated by canonical RNA-binding domains that form stable secondary and tertiary structures. However, a number of pioneering studies over the past decades, together with recent proteome-wide data, have challenged this view, revealing surprising roles for intrinsically disordered protein regions in RNA binding. Here, we discuss how disordered protein regions can mediate protein-RNA interactions, conceptually grouping these regions into RS-rich, RG-rich, and other basic sequences, that can mediate both specific and non-specific interactions with RNA. Disordered regions can also influence RNA metabolism through protein aggregation and hydrogel formation. Importantly, protein-RNA interactions mediated by disordered regions can influence nearly all aspects of co- and post-transcriptional RNA processes and, consequently, their disruption can cause disease. Despite growing interest in disordered protein regions and their roles in RNA biology, their mechanisms of binding, regulation, and physiological consequences remain poorly understood. In the coming years, the study of these unorthodox interactions will yield important insights into RNA regulation in cellular homeostasis and disease.

Keywords: RNA-binding protein, Intrinsically disordered protein, Co- and post-transcriptional RNA regulation, RS repeat, RGG-box, GAR repeat, Basic patch, Poly-K patch, Arginine-rich motif, RNA granule

Plain English summary

DNA is well known as the molecule that stores genetic information. RNA, a close chemical cousin of DNA, acts as a molecular messenger to execute a set of genetic instructions (genes) encoded in the DNA, which come to life when genes are activated. First, the genetic information stored in DNA has to be copied, or transcribed, into RNA in the cell nucleus and then the information contained in RNA must be interpreted in the cytoplasm to build proteins through a process known as translation. Rather than being a simple process, the path from transcription to translation entails many steps of regulation that make crucial contributions to accurate gene control. This regulation is in large part orchestrated by proteins that bind to RNA and alter its localisation, structure, stability, and translational efficiency. The current paradigm of RNA-binding protein function is that they contain regions, or domains, that fold tightly into an ordered interaction platform that specifies how and where the interaction with RNA will occur. In this review, we describe how this paradigm has been challenged by studies showing that other, hitherto neglected regions in RNA-binding proteins, which in spite of being intrinsically disordered, can play key functional roles in protein-RNA interactions. Proteins harbouring such disordered regions are involved in virtually every step of RNA regulation and, in some instances, have been implicated in disease. Based on exciting recent discoveries that indicate their unexpectedly pervasive role in RNA binding, we propose that the systematic study of disordered regions within RNA-binding proteins will shed light on poorly understood aspects of RNA biology and their implications in health and disease.

Background

Structural requirements for RNA-protein interactions

RNA-binding proteins (RBPs) assemble with RNA into dynamic ribonucleoprotein (RNP) complexes that mediate all aspects of RNA metabolism [1, 2]. Due to the prominent role that RBPs play in RNA biology, it is not surprising that mutations in these proteins cause major diseases, in particular neurological disorders, muscular atrophies and cancer [3–7]. Until recently, our understanding of how
RBPs interact with RNA was based on a limited number of globular RNA-binding domains (RBDS), which include RNA-recognition motif (RRM), K-homology domain (KH), double-stranded RBD (dsRBD), zinc fingers (Znf), DEAD box helicase domain, and others (for recent reviews, see [8–10]). Each of these RBDS interacts with RNA following distinct mechanisms and differ in specificity and affinity for their target RNA. Promiscuous RNA binding is often mediated by interactions with the phosphate-sugar backbone, whereas sequence-specificity builds on interactions with the nucleotide base and shape complementarity between protein and RNA interfaces. While the most common RBDS interact with short (4–8 nt) sequences, others display lower or complete lack of sequence selectivity, recognising either the RNA molecule itself or secondary and three-dimensional structures [8, 11]. As the affinity and specificity of a single RBD is often insufficient to provide selective binding in vivo, RBPs typically have a modular architecture containing multiple RNA-interacting regions [8]. RNA-binding proteins are typically conserved, abundant, and ubiquitously expressed, reflecting the core importance of RNA metabolism in cellular physiology [12, 13].

The coming of age for RNA-binding proteins — the emerging role of protein disorder

Early on, it was recognised that not all RNA-binding activities could be attributed to classical RBDS. Computational predictions based on transcriptome complexity suggested that 3–11% of a given proteome should be dedicated to RNA binding, whereas only a fraction of this number could be identified by homology-based searches for classical RBDS [14, 15]. Moreover, there were several reports of RNA-binding activities within protein domains lacking similarities to any classical RBD [16, 17]. A number of studies showed that intrinsically disordered regions, lacking any stable tertiary structure in their native state, could contribute to RNA binding. For example, the flexible linker regions that separate the two RRM of the poly(A)-binding protein (PABP) and polypyrimidine tract binding protein 1 (PTBP1), not only orientate the domains with respect each other, but also mediate RNA binding [18]. Flexible regions in RBPs rich in serine and arginine (S/R) and arginine and glycine (R/G) were found to contribute, or even to account for, RNA-binding activities [19, 20]. Furthermore, early computational analyses revealed that proteins involved in transcription and RNAs processing are enriched in disordered protein regions [21, 22], hinting on a broader role for protein disorder in RNA metabolism.

Recently, the development of proteome-wide approaches for comprehensive determination of the RBP repertoire within the cell (RBPome) has substantially increased the number of known unorthodox RBPs. In vitro studies in yeast identified dozens of proteins lacking classical RBDS as putative RBPs, including metabolic enzymes and DNA-binding proteins [23, 24]. Two recent studies that employed in vivo UV crosslinking, poly(A)-RNA capture, and mass spectrometry, identified more than a thousand proteins interacting with RNA, discovering hundreds of novel RBPs [25, 26]. Strikingly, both known and novel RBPs were significantly enriched in disordered regions compared with the total human proteome. Approximately 20% of the identified mammalian RBPs (~170 proteins) were disordered by over 80% [25, 27]. Apart from the disorder-promoting amino acids such as serine (S), glycine (G), and proline (P), these disordered regions were enriched in positively (K,R) and negatively (D, E) charged residues as well as tyrosine (Y) [25], amino acids often found at RNA-interacting surfaces in classical RBDS [8]. Disordered amino acid sequences in RBPs form recognisable patterns that include previously reported motifs such as RG-and RS-repeats as well as new kinds of motifs, such as K or R-rich basic patches (Fig. 1). As with classical RBDS, disordered regions also occur in a modular manner in RBPs, repeating multiple times in a non-random manner across a given protein and, in some instances, combining with globular domains [25]. Taken together, these observations suggest that disordered regions 1) contribute to RBP function; 2) combine in a modular manner with classical RBDS suggesting functional cooperation; and 3) may play diverse biological roles, including RNA binding. Supporting this, a recent report has shown that globular RBDS are on average well conserved in number and sequence across evolution, while disordered regions of RBPs have expanded correlating with the increased complexity of transcriptomes [13]. What is the contribution and functional significance of protein disorder in RNA-protein interactions? Below, we will discuss what is known about disordered regions in RNA binding and metabolism, as well as physiology and disease, based on accumulating literature (Table 1, Additional file 1: Figure S1).

Review

Disordered RS repeats put RNA splicing in order

Disordered, arginine and serine (RS) repeat containing regions occur in a number of human proteins referred to as SR proteins and SR-like proteins (reviewed in [28, 29]). SR proteins are best known for their roles in enhancing splicing but have been ascribed functions in other RNA processes from export, translation, and stability to maintenance of genome stability (e.g. [30, 31] for reviews). There are twelve SR proteins in human that contain 1–2 classical RRMs and an RS repetitive motif of varying length [30]. Classical SR proteins bind exonic splicing enhancers in nascent RNA through their RRMs
and promote splicing of adjacent introns [32, 33]. The RS repeat enhances splicing in a length-dependent manner [34]. RS repeats are predicted to be intrinsically disordered [35] (Table 1), but phosphorylation promotes a transition towards a less flexible, arch-like structure with an influence on RNA binding in the serine/arginine-rich splicing factor 1 (SRSF1) [36] (Fig. 1). RS repeats have been shown to directly bind RNA during multiple steps of splicing [19, 37–39] and to contribute to binding affinity of RRMs for RNA by inducing a higher affinity form of the RRM [40]. RS repeats can also mediate protein-protein interactions [28, 33], hence their association with RNA can also be indirect. RS-mediated protein binding seems to be compatible with RNA binding [33, 41], suggesting that protein and RNA binding could take place simultaneously or sequentially. RNA-binding RS repeat sequences can also be found in viral proteins, such as the nucleocapsid of severe acute respiratory syndrome coronavirus (SARS-CoV), causative agent of the alike-named disease. This protein employs RS-rich disordered region, in cooperation with other RNA-binding regions, to capture viral RNA and package it into virions [46]. Taken together, these reports suggest that RS repeats have broader roles in RNA-binding than previously anticipated.

Fig. 1 Three classes of disordered protein regions involved in direct RNA-interactions. Blue oval indicates the disordered region of each protein involved in RNA binding. Sequence is shown below the protein model, and typical sequence characteristics are indicated by boxes. Disorder profile was calculated using IUPred [172]. Values above 0.4 are considered disordered.
Table 1: Examples of RNA binding proteins where a disordered, non-classical region is involved in direct RNA binding. Additional details for each protein are presented in Additional file 1: Figure S1. Disorder prediction was calculated using IUPred [172]

| Protein                  | ID          | Name                              | Aliases            | Species       | Canonical domains | Function                                      | Class     | Sequence | Disorder assignment | Target RNA preference | Regulation at disordered region | Interaction with other biomolecules | Ref       |
|--------------------------|-------------|-----------------------------------|--------------------|---------------|------------------|------------------------------------------------|-----------|----------|---------------------|-------------------------------|----------------------------------|----------------------------------|-----------|
| SRSF1                    |            | Serine/arginine-rich splicing factor 1 | ASF, SF2, SF2P33, SFRS1 | Homo sapiens  | 2xRRM            | RNA splicing, Essential for heart development. | RS        | 110-258 | GPRSP5YGRGSR5R5R5SRSRSRSNSNSRS5SRSY227 | Experimental | Serine phosphorylated, becomes more structured upon phosphorylation, alternatively spliced. | Protein [36, 39, 173–175] |         |
| U2AF2                    |            | Splicing factor U2AF 65 kDa subunit | U2AF65             | Homo sapiens  | 3xRRM            | RNA splicing. | RS        | 1-201 | 1-MSDFDEFERQLNENKQERK6ENRIRKSHSRSRSRSDKRRRRR RSKPRPTLSLRGAEEOHGG LIRSPREKKKKVRKYWDVPPPG | Predicted | No specificity | Serine phosphorylation, lysine acetylation, lysine hydroxylation | Protein [19, 176] |         |
| NKAP                     |            | NF-kappa-B-activating protein      | -                  | Homo sapiens  | None             | RNA splicing, transcriptional repression. | RS        | 1-201 | 1-MSDFDEFERQLNENKQERK6ENRIRKSHSRSRSRSDKRRRRR RSKPRPTLSLRGAEEOHGG LIRSPREKKKKVRKYWDVPPPG | Predicted | poly (U) | Lysine acetylation | Protein [43] |         |
| Nucleocapsid protein     |            | Nucleoprotein, NC, N              | Severe acute respiratory syndrome coronavirus (SARS-CoV) | Homo sapiens  | None             | Major structural component of virions that associates with genomic RNA to form a long, flexible, helical nucleo-capsid. | Other, RS, poly/K, other | 1-201 | 1-MSDFDEFERQLNENKQERK6ENRIRKSHSRSRSRSDKRRRRR RSKPRPTLSLRGAEEOHGG LIRSPREKKKKVRKYWDVPPPG | Experimental | poly (U) ssRNA | - | Protein [46, 177] |         |
| ALYREF                   |            | Aly/REF export factor 2           | Alyref             | Mus musculus  | 1xRRM            | RNA export. | RG        | 27-VNRGGPGRNRPAPAIAA RGRRNRPAP5YSR | Experimental | TAP displaces RNA from ALYREF | Protein [54, 55, 57, 178] |         |
| Aven                     |            | Cell death regulator Aven         | -                  | Homo sapiens  | None             | Positive translational regulator. | RG        | 1-201 | 1-MQAFARGGARGGRRRGRGGRPGGGRGDRHSERPG AAAAVARGGGGGGGGG DGGGGGRRGGRGFR GARAGGGGGGAPRR | Predicted | RNA G-quadruplex | Methylation (no influence on RNA binding, influences protein interactions and polysome association), alternative transcript (mouse) | Protein [179, 180] |         |
| Caprin-1                 |            | -                                 | None               | None          | None             | - | RG        | Predicted | - | - | - | - | | | |
| Protein | Function | Disorder Prediction | Interaction | Modifiable Region |
|---------|----------|---------------------|-------------|------------------|
| GPIAP1, GRP137, M1151, RNG105 | Regulation of localised translation, synaptic plasticity, cell proliferation and migration. | GQGGARGARQLMGNGYR GPPANGFRGGYDGYRGP | Homo sapiens, Xenopus | RGG1, GQGGARGARQLMGNGYR |
| DDX4 | Vasa | RNA helicase | Experimental | Single-stranded DNA. |
| EWS | RNA-binding protein EWS | 1xRBM | Predicted | G-quadruplex (RGG3, not RGG1 or RGG2) |
| FMRP | Fragile X mental retardation protein 1 | 2xKH | Experimental | G quadruplex, Arginine methylation. |

Disorder prediction was calculated using IUPpred [172].
Table 1 Examples of RNA binding proteins where a disordered, non-classical region is involved in direct RNA binding. Additional details for each protein are presented in Additional file 1: Figure S1. Disorder prediction was calculated using IUPred [172] (Continued)

| Protein | Function | Disorder prediction | Nucleic acid interactions | Protein interactions |
|---------|----------|---------------------|---------------------------|----------------------|
| FUS | RNA-binding protein FUS | TLD | Homo sapiens, Drosophila melanogaster | Splicing, polyadenylation. | Predicted G-quadruplex | Arginine methylation. |
| TLS | - | Homo sapiens | 213-RRGPRGGG-220, 241-PGRGGGGGGGGGCGG-253, 377-RRGGNGRGRGGRGGGMPRGRGGYGGGGGGGGGRRGG-409, 412-RGGPRGGYDRGGYRGRGGDRG-425 | - | [190-193] |
| hnRNPU | Heterogeneous nuclear ribonucleoprotein U | SAFA, U21.1 | Homo sapiens | RNA stability, U2 snRNP maturation, DNA binding. | Predicted Poly (U and poly (G)) homopolymers, UGUG | - | DNA |
| ICP27 | - | Herpes simplex virus | 138-RRGGPRGRRRGRG-152 | Predicted Poly (G) and poly (U) homopolymers, GC-rich sequences | Methylated | - | [194-196] |
| LAF1 | - | DDX3 | C. elegans | RNA helicase. | Experimental | - | Region 43-106 containing RG-repeat is alternative. |
| NXF1 | Nuclear RNA export factor 1 | TAP | Mus musculus, homo sapiens | Nuclear export. | Predicted Non-specific | - | Protein. Overlaps a nuclear localisation and export signals. |
| Nucleolin | - | NCL, Protein C23 | Hamster | Chromatin decondensation, pre-RNA transcription, ribosome assembly. | Predicted | - | Protein (in human) |
| RBMX | RNA-binding motif protein, X chromosome | HNRPG, RBMXP1 | Homo sapiens, Xenopus laevis | Regulation of transcription, splicing. | Predicted C-terminal regions binds structured (hairpin) RNA | Identical C-terminal sequence is mouse RBMX is alternatively spliced. | - | [201-206] |
| Protein                              | Organism | Disorder prediction | FTx| Additional details for each protein are presented in Additional file 1: Figure S1. Disorder prediction was calculated using IUPred [172] (Continued) |
|-------------------------------------|-----------|---------------------|----|----------------------------------------------------------------------------------|
| Foamy virus Gag                     | Human foamy virus | None              | RG [46] RPSRGRGRGQN [49] | Predicted | - | - | [207–210] |
| TERF2 Telomeric repeat-binding factor 2 | Homo sapiens | None              | RG [41] MAGGGGSSDGGRAG RRSRSSGRRGRRGHEPGL GGPAERGAG [46] | Predicted | G-rich, TERRA Arginine methylation Protein | [211–214] |
| XTUT7                               | Xenopus laevis | Zinc finger | Basic patch (poly R) [453] MRRNRVRRRNNENAG NQR [45] | Predicted | - | - | [215] |
| Tat Transactivating regulatory protein | Human immuno-deficiency virus (HIV) | None | transcriptional activator, transcription elongation. | Basic patch (poly R) [40] RKKRRQRRR [57] | Experimental | Structured RNA (HIV-1 Trans-activation response element, TAR) Arginine methylation. Protein | [85, 88–91, 93, 216–223] |
| Rev Regulator of expression of viral proteins | Human immuno-deficiency virus (HIV) | None | RNA export. | Basic patch (poly R) [34] TROARRNRRRRWRER QR [30] | Experimental | Structured RNA (HIV-1 Rev response element, RRE) Arginine methylation. Protein | [96–101, 103, 104, 153, 154, 224] |
| Tat Transactivating regulatory protein | S ORF, bTat | Bovine immuno-deficiency virus | None | Transcriptional activator | Basic patch (poly R) [70] RGTRGKGRIRR [81] | Experimental | Structured RNA (TAR) | Protein | [91] |
| Coat protein                        | Alfalfa mosaic virus | None | Capos protein, viral RNA, Translation initiation. | Basic patch (poly K) [6] KKAGGKAKPTKRSQ NYAAALRK [27] | Experimental | - | - | [225, 226] |
| PAPDS Non-canonical poly (A) RNA polymerase | Homo sapiens | None | RNA oligoadenylation, RNA stability | Basic patch (poly K) [557] KKRRH[55] | Predicted | May have a preference for structured RNA Alternative splicing | - | [109] |
| SDAD1 Protein SDA1 homolog         | Homo sapiens | None | Protein transport, ribosomal large subunit export from nucleus. | Basic patch (poly K) [244] RDLVQYATGXXKK NKKKLVAMKVLKIQ KKKKPEVPNIK [280] | Predicted | - | - | [58] |
| HMGA1                               | None | (e) AT | Experimental | DNA | | | | | |
Table 1 Examples of RNA binding proteins where a disordered, non-classical region is involved in direct RNA binding. Additional details for each protein are presented in Additional file 1: Figure S1. Disorder prediction was calculated using IUPred [172] (Continued)

| Protein Name | Disorder Type | Additional Details |
|--------------|---------------|--------------------|
| High mobility group protein HMGB-1/WW-Y | Homo sapiens | Binds structured RNA. Arginine methylation. |
| Tip5 | Homo sapiens | None | Epigenetic rRNA gene silencing. Experimental Exhibits preferential binding towards dsRNA |
| Bromodomain adjacent to zinc finger domain protein 2A | Homo sapiens | Binds structured RNA. Arginine methylation. [121, 124, 125, 127] |
| PTOV1 | Homo sapiens | None | Regulation of transcription. Experimental Exhibits preferential binding towards dsRNA |
| Tip5 Bromodomain adjacent to zinc finger domain protein 2A | Homo sapiens | None | Epigenetic rRNA gene silencing. Arginine methylation. [121, 124, 125, 127] |
| PTOV1 | Homo sapiens | None | Regulation of transcription. Experimental Exhibits preferential binding towards dsRNA |
| SRSF2 | Homo sapiens | 1xRRM | RNA splicing. Other experimental proteins. |
| Tra2-β1 Transformer-2 protein homolog beta | Homo sapiens | 1xRRM | RNA splicing. Other experimental proteins. |
| hnRNPA1 Heterogeneous nuclear ribonucleoprotein A1 | Human, Xenopus tropical | 2xRRM | hnRNP particle formation, nucleo-cytoplasmic transport, splicing. Other/ RG Mutations affecting the RG- and FG-repeat peptides is alternatively spliced. RG-region may mediate RNA binding. The entire region is involved in hnRNPA1 aggregation and includes a nuclear targeting sequence. Region containing the RG- and FG-repeat peptides is alternatively spliced. RG-region may mediate RNA binding. The entire region is involved in hnRNPA1 aggregation and includes a nuclear targeting sequence. |
| LUZP4 Leucine zipper protein 4 | Homo sapiens | None | Nuclear export. Experimental Protein |
| ORF57 | Homo sapiens | None | Other Experimental Protein |

(Järvelin et al. Cell Communication and Signaling 2016)
Table 1 Examples of RNA binding proteins where a disordered, non-classical region is involved in direct RNA binding. Additional details for each protein are presented in Additional file 1: Figure S1. Disorder prediction was calculated using IUPred [172] (Continued)

| Protein                        | Function                                | Disorder prediction | RNA binding region |
|--------------------------------|-----------------------------------------|---------------------|-------------------|
| **52 kDa immediate-early phosphoprotein, mRNA export factor ICP27 homolog** | Viral RNA regulation. RVYRSPSPWRHGKRPR RISSNTRSCXTSW ADRVEAAQRK**120** | **Viral RNA: GAAGAGG, CAGUGCG AAGAGG** | **RNA binding region partially overlaps with ALYREF binding site.** |
| APC Adenomatous polyposis coli protein | Microtubule binding, negative regulator of Wnt signaling. **2223**-**SISRGRTMIHIPGLR NSSSSTSTPVSKKRP LKTPASPKSPEGPG A**TTSPRIGTPKGA**SGSEL SPIRTQTSQGSGSNK GSSRSGRSDSTPSRP TQQLPSRPAMQSPGQR NISPVRNGGIPNPRL SQLPRITSPSTASSTK SGGSKKMYSTSPPGQ RLSOQLNTQASLS5KN ASSIPRESASAKGLN OMSNGNGNSNKRVELS RMSSTKSSGSGEDSS ERPAVRQSTTFKEA P**SPTLR**RKLEESASF ESLSPPSRDPSPTRS QAQTPVSLPSLPGMS LSTHPVSQAGGWGRRL PPNLSPITRYNDGRP TKRHDRIARHSEPS RLPIRNAGTVKREH5 KHESSLPRVSTWRT GSSSSLSASSE**2579** | **Predicted G-rich motif** | **-** |
| CTCF Transcriptional repressor CTCF | None | **Predicted** | **-** |
| Df31 Decondensation factor 31 Anon1A4 D. melanogaster | Regulation of higher-order chromatin structure, maintenance of open chromatin. | ** experimental** | **Non-specific but does not bind ssDNA or dsDNA. Preferentially binds snoRNA.** |
Table 1  Examples of RNA binding proteins where a disordered, non-classical region is involved in direct RNA binding. Additional details for each protein are presented in Additional file 1: Figure S1. Disorder prediction was calculated using IUPred [172] (Continued)

| Protein                  | Species       | Disorder Prediction | Additional Details |
|--------------------------|---------------|---------------------|--------------------|
| EzH2 Histone-lysine N-methyltransferase EZH2 | Mus musculus | Polycomb group protein, involved in H3 methylation (H3K9me and H3K27me). | Predicted | May have a preference for RNA stem loops. 1st Thr is phosphorylated in a cell cycle dependent manner. Phosphorylation increases RNA binding. This region overlaps a region involved in protein-protein interactions in human, however, RNA and protein binding regions may be distinct from one another. |
| Enx1h | Mus musculus | None | ASTKDEVQNGAEASE VAA | [241–243] |
| Nrep Neuronal regeneration-related protein | P311 Mus musculus | None | experimental | - - Protein |
| Gemin5 Gem-associated protein 5 | Homo sapiens | None | snRNP assembly, splicing, IRES-mediated translation initiation. | Other 27 | - - |
| Nup153 | Homo sapiens | None | Component of the nucleopore, RNA trafficking. | Other 250 | - - |
| SCML2 Sex comb on midleg-like protein 2 | Homo sapiens | None | Binds Polycomb Repressive | Other 256 | - - |

Järvelin et al. Cell Communication and Signaling (2016) 14:9
Table 1 Examples of RNA binding proteins where a disordered, non-classical region is involved in direct RNA binding. Additional details for each protein are presented in Additional file 1: Figure S1. Disorder prediction was calculated using IUPred [172] (Continued)

| Protein | Function | Disorder Location | Disorder Prediction | Disorder Features | Binding Site | Serine Phosphorylation |
|---------|----------|-------------------|---------------------|-------------------|--------------|------------------------|
| KDM4D   | Lysine-specific demethylase 4D | Homo sapiens | Complex 1 and histones. Involved in epigenetic silencing. | VKNIIPKRPGPNSQK KEKPLPVQTSASG | Between RNA and DNA. | Serine phosphorylation * |
| JMJD2D  | None     | Synthetic         | None                | MEPRVPASQELSTQK 348-351 | Experimental | -                      |
| -       | -        | Synthetic         | None                | SRRSRARRRRRARRRRR | Experimental | Structured RNA (HIV-1 Rev response element, RRE) | [105] |

* According to uniprot, from a large-scale study but no detailed experimental confirmation available.
RG-rich repeats — The swiss-army knife of protein-RNA interactions

A commonly occurring disordered RNA–binding motif in RBPs consists of repeats of arginine and glycine, termed RGG-boxes or GAR repeats. These sequences are heterogeneous both in number of repeats and in their spacing. A recent analysis divided these RG-rich regions into di- and tri-RG and -RGG boxes, and identified instances of such repeats in order of tens (di- and tri-RGG) to hundreds (tri-RG) and nearly two thousand (di-RG) proteins [47]. Proteins containing such repeats are enriched in RNA metabolic functions [47]. However, it is not currently clear whether the different repeat architectures provide distinct functional signatures.

The RGG box was first identified in the heterogeneous nuclear ribonucleoprotein protein U (hnRNPU, also known as SAF-A) as a region sufficient and required for RNA binding (Table 1, Fig. 1). hnRNPU lacks canonical RBDS, but has semi-structured SAP domain involved in DNA binding [48–50]. hnRNPU has been found to target hundreds of non-coding RNAs, including small nuclear (sn)RNAs involved in RNA splicing, and a number of long non-coding (lnc)RNAs, in an RGG-box-dependent manner [51]. RGG-mediated interaction of hnRNPU with the lncRNAs Xist [52] and PANDA [53] has been implicated in epigenetic regulation.

RG(G) -mediated RNA binding also plays a role in nuclear RNA export, as illustrated by the nuclear RNA export factor 1 (NXF1). While NXF1 harbours an RRM capable of binding RNA [54], most of the in vivo RNA-binding capacity is attributed to the RGG-containing, N-terminal region [55] (Table 1). The arginines in this motif play a key role in the interaction with RNA, which has been shown to be sequence-independent but necessary for RNA export [55]. NXF1 overall affinity for RNA is low [55, 56], and requires the cooperation with the export adapter ALY/REF [57]. ALY/REF also bears an N-terminal disordered arginine-rich region that resembles an RGG-box [57] and mediates both RNA binding [54, 58, 59] and the interaction with NXF1 [60]. The activation of NXF1 is proposed to be triggered by the formation of a ternary complex between ALY/REF and NXF1, in which their RG-rich disordered regions play a central role. Analogous sequences has been identified in viral proteins and also facilitate viral RNA export by bypassing canonical nuclear export pathways (Table 1).

Fragile X mental retardation protein (FMRP) is another RBP with a well-characterised, RNA-binding RGG-box (Fig. 1). Involved in translation repression in the brain [61], loss of FMRP activity leads to changes in synaptic connectivity [62], mental retardation [63–65], and may also promote onset of neurodegenerative diseases [66]. In addition to its RGG-box, FMRP contains two KH domains that contribute to RNA binding. The RGG-box of FMRP has been shown to interact with high affinity with G-quadruplex RNA structures [67–77]. The RGG-box is unstructured in its unbound state [70, 78], but folds upon binding to a guanine-rich, structured G-quadruplex in target RNA [78] (Fig. 2). Both arginines and glycines play a key role in the function of the RGG-box and replacement of these amino acids impair RNA binding [78]. The arginine residues used to interact with RNA vary depending on the target RNA [70, 76, 78]. The FMRP RGG-box targets its own mRNA at an G-quadruplex structure that encodes the RGG-box [69]. This binding regulates alternative splicing of FMRP mRNA proximal to the G-quartet, suggesting it may auto-regulate the balance of FRMP isoforms [74]. Surprisingly, a recent transcriptome-wide study of polysome-associated FMRP found no enrichment for predicted G-quadruplex structures in the 842 high-confidence target mRNAs [79]. Another study identified FMRP binding sites enriched in specific sequence motifs, where the KH2 domains emerged as the major specificity determinants [80]. These results suggest that the role of RGG-box in this RBP could be limited to increase the overall binding affinity of the protein, supporting the sequence-specific interactions mediated by the KH2 domains. However, we cannot exclude the possibility of differential UV crosslinking efficiency of the KH2 domains and the RGG-box, which could result in biased binding signatures in CLIP studies.

A number of other RBPs use an RGG-repeat region to target G-rich and structured RNA targets and are implicated in neurological disease as well as cancer (Table 1). These RG-rich regions can mediate both unselective and specific interactions with RNA and can be involved in varied RNA metabolic processes.

Catching the RNA with a basic arm

Basic residues often cluster in RBPs to form basic patches that can contribute to RNA-binding. Analysis of mammalian RNA-binding proteomes showed that such motifs are abundant among unorthodox RBPs [25, 27]. Basic patches are normally composed of 4–8 lysines (K) or, less frequently, arginines (R), forming a highly positive and exposed interface with potential to mediate molecular interactions [25]. Basic patches can occur at multiple positions within an RBP forming islands that often flank globular domains. This suggests functional cooperation between natively structured and unstructured regions [25]. Many RBPs contain alternating basic and acidic tracts that form highly repetitive patterns with unknown function [25]. Since acidic regions are not thought to interact with RNA [58], they may be involved in other intra- or intermolecular interactions, or contribute to accessibility and compaction of the region [81].

Arginine rich motifs (ARMs) (Table 1) are probably best characterised in viral proteins. These motifs tend to
be disordered, and when bound to RNA, range from completely disordered to ordered but flexible. Although simple in terms of amino acid composition, ARMs seem to be able to target RNAs quite diversely and often specifically [82]. Lentiviral Tat proteins (Trans-Activator of Transcription) are key regulator of viral biological cycle by promoting viral gene expression upon binding to an RNA structure present at the 5' end of the nascent viral RNA (called the trans-activation response element, TAR) [83]. Human immunodeficiency virus (HIV) Tat ARM is intrinsically disordered in its free-state [84–87]. Only one key arginine, flanked by basic amino acids, is required for specific interaction with TAR [88, 89]. Differences in the flanking basic amino acids contribute to selectivity between TARs from different viruses [90]. ARMs can accommodate different binding conformations depending on their target RNA. For example, bovine immunodeficiency virus (BIV) Tat ARM forms a beta-turn conformation upon binding to TAR [91] (Fig. 2c). Jembrana disease virus (JDV) Tat ARM can bind both HIV and BIV TARs, as well as its own TAR, but does so adopting different conformations and using different amino acids for recognition [92]. The RNA-binding disordered region of HIV Tat also mediates protein-protein interactions required for nuclear localisation [93]. Structural flexibility required to engage in diverse simultaneous or sequential RNA and protein interactions might explain why the native ARM-RNA interactions do not display very high affinity [92].

Similar to Tat proteins, lentiviral Rev auxiliary protein binds a structured RNA element (the Rev response element, RRE) present in partially-spliced and unspliced viral RNAs to facilitate nuclear export of viral RNA [94, 95]. The HIV Rev ARM was experimentally shown to be intrinsically disordered when unbound in physiological conditions [96–98] (Table 1, Fig. 1). Disorder-to-structure transition correlates with RNA binding and the RRE-bound Rev folds into an alpha-helical structure that maintains some structural flexibility [96–100]. Rev oligomerises and binds the multiple stems of the RRE using diverse arginine contacts, which results in a high-affinity ribonucleoprotein that allows efficient nuclear export of unspliced HIV RNAs [101–103]. Interestingly, Rev can also bind in an extended conformation to in vitro selected RNA aptamers [104], highlighting the role of RNA secondary and tertiary structure in the conformation that Rev adopts. The RRE can also be recognised by several different in vitro selected R-rich peptides that include additional serine, glycine, and glutamic acid residues [105–107] — these peptides are predicted to be disordered (Table 1). A simple, single nucleotide base changes in the RRE can direct affinity towards a particular ARM [108]. These features highlight the structural malleability of the Rev ARM, and suggest that some structural flexibility is relevant for in vivo binding.

The basic amino acid lysine can form disordered polylysine peptides that interact with RNA. 47 proteins identified in the human RNA-binding proteome have a long poly-K patch but lack known RBDs, suggesting these motifs are good candidates for RNA binding [25]. The K-rich C-terminal tail of protein SDA1 homolog (SDAD1) is composed of 45 amino acids, including 15 K, one R, two glutamines (Q) and two asparagines (N) (Table 1, Fig. 1). It binds RNA in vivo with similar efficiency as a canonical domain such as RRM [58]. The human non-canonical poly(A) polymerase PAPD5, that is involved in oligoadenylylating aberrant rRNAs to target them for degradation [109, 110], also lacks canonical RBDs, but its C-terminal basic patch is directly involved in binding RNA (Fig. 1, Table 1). Removal or mutation of this sequence results in impaired RNA binding and reduced catalytic activity [109].
Basic tails in RBPs share physicochemical similarities with analogous sequences in DNA-binding proteins (DBPs) [111]. In DNA-binding context, basic patches are known to endow faster association with DNA due to increased ‘capture radius’ as well as to promote hopping and sliding movements along DNA molecules [112–118]. DNA binding through basic tails seems to be sequence-independent [119] and structural studies have shown that basic residues are projected into the minor groove of the double stranded DNA helix, establishing numerous electrostatic interactions with the phosphate-sugar backbone [116, 120]. Basic patches in RBPs may modulate RNA searching and binding avidity in a similar manner.

One open question is whether basic tails can distinguish between DNA and RNA. The AT-hook, defined as G-R-P core flanked by basic arginine and/or lysine residues, binds DNA and is found in many nuclear, DNA-binding proteins [121, 122]. However, this motif has been recently shown also to bind RNA [123–126]. Furthermore, an extended AT-hook (Table 1), occurring in tens of mouse and human proteins, binds RNA with higher affinity than DNA [127]. This motif from Prostate Tumor Overexpressed 1 (PTOV1) was shown to bind structured RNA, in agreement with the previously known property of basic tails to bind in the minor groove of double stranded DNA [116, 120]. Therefore, different types of disordered sequences may be able to recognise both RNA and DNA, albeit they may have preference for one.

A role for disordered regions of RBPs in retaining RNA in membraneless granules

RNA processing and storage is often undertaken in the context of dynamic, membraneless organelles that vary in size, composition, and function. These organelles include the nucleolus, PML bodies, nuclear speckles and Cajal bodies in the nucleus as well as P-bodies, stress and germ granules in the cytoplasm [128–130]. RNA granule formation relies on a spatiotemporally controlled transition from disperse “soluble” RNA and protein state to a condensed phase [131, 132]. The lack of a membrane allows a direct, dynamic and reversible exchange of components between the cytoplasm and the granule [131]. The rate of exchange and localisation of a protein within a granule can be markedly different depending on granule composition and the intrinsic properties of the protein [133–136]. RNA granules have roles in RNA localisation, stability, and translation, and perturbations in their homeostasis are hallmarks of numerous neurological disorders [137, 138].

Several recent studies have shown that disordered, low complexity regions in a number RBP have a capacity to form such granules [131, 139–141]. Different low complexity regions can promote RNA granule formation. For example, the disordered RG-rich sequence of LAF-1 (DDX3) was demonstrated to be both necessary and sufficient to promote P-granule formation in C. elegans [142]. Similarly, the RG/GR and FG/GF disordered tail of human RNA helicase DDX4 (aka Vasa) aggregates in vivo and in vitro [130]. Furthermore, the [G/S]YG/G/S and poly glutamine (polyQ) motifs, which are present in a broad spectrum of RBPs, are necessary and sufficient to cause aggregation in vitro and in vivo [139, 140, 143–146]. It remains unclear how RNA binding by these sequences influences granule formation. Illustrating this idea, the RG-rich region of LAF-1 displays direct RNA–binding activity in addition to granule formation capacity. While RNA is not required for LAF-1 driven aggregation, it increases the internal dynamics of these LAF-1 droplets, making them more fluid [142]. In yeast, formation of P-body-like granules by the Lsm4 disordered region requires the presence of RNA [147]. Notably, the biophysical properties of RBP droplets can be altered by the presence of different RNA species [148]. A recent work reports an additional layer of complexity in the interplay between nucleic acids and granules. While single-stranded DNA is retained in DDX4-induced granules, double-stranded DNA is excluded, suggesting some degree of nucleic acid selectivity [130]. Given the biophysical similarities between DNA and RNA, it is possible that granules formed by analogous low complexity sequences also retain single stranded over double stranded RNA.

Interestingly, different types of low complexity sequences may help to form different types of aggregates and ways to embed RNA. A recent study showed that while low complexity sequences promote formation of both P-bodies and stress granules in yeast, these granules differ in their dynamic properties, P-bodies displaying more dynamic/fluid phase transition than more solid-like stress granules [147]. Granule structure, composition, and age can affect the biophysical properties of the granules [135, 136]. There is considerable overlap in the composition of different RNA granules [149]. Different proportions of such components may lead to the existence of a continuum of granule types with increasingly distinct physicochemical properties. In summary, it is clear that protein disorder has a role in formation of RNA granules. The importance of direct interaction between disordered regions and RNA in the context of granules remains to be determined.

Modulating interactions between disordered regions and RNA

Post-translational modifications can modulate protein’s interaction properties [150]. A number of disordered RNA-
binding regions are known to be post-translationally modified (Table 1, Additional file 1: Figure S1) and some of these modifications can modulate RNA-binding affinity or cause local structural changes. For example, methylation of arginines of the RNA-binding RGG-box in the RNA export adapter ALY/REF reduces its affinity for RNA [151]. Arginine methylation of the RGG-box of the translational regulator FMRP affects interaction with target RNA as well as its polyribosome association [76, 152]. Also the RNA-binding basic patch of HIV protein Rev is methylated, which changes its interaction dynamics with its target RNA [153, 154]. Serine phosphorylation at the RNA-binding RS repeats of SRSF1 and DDX23 has been shown to induce a partial structuring of this region, which may impact their RNA-binding properties [36]. Assembly of RNA granules can also be modified by phosphorylation or methylation of the low complexity region [130, 155, 156]. In summary, occurrence of post-translational modifications at disordered regions represents an additional layer of regulation of RNA binding and metabolism (Fig. 3).

In other contexts, it is known that alternative splicing can alter the sequence and function of proteins. Several global analyses have reported that short, regulatory sequences such as sites for post-translational modifications and protein-protein interactions are often subjected to alternative splicing [157–159]. Could protein-RNA interactions be regulated in a similar manner? A number alternative isoform variants catalogued in large-scale studies affect RNA-binding disordered regions (Table 1, Additional file 1: Figure S1). As an illustrative example, alternative splicing of mouse ALY/REF selectively includes or excludes the RNA-binding RG-rich region, resulting in changes in its targeting to nuclear speckles and an increased cytoplasmic distribution [57, 60]. Alternative splicing affecting a region adjacent to the FMRP RGG-box influences the protein’s RNA-binding activity [160], reduces its ability to associate with

---

**Fig. 3** Models for properties of protein disorder in RNA binding. **a** Attributes of disordered protein regions in RNA interactions. **b** Post-translational modification and alternative splicing can modulate RNA-binding.
polyribosomes [161], and can also impact RGG-box methylation [162]. Another splice isoform results in ablation of the RGG-box as a result of a translational frameshift, which induces nuclear distribution of the protein [163]. Also RNA granule formation can be differentially regulated in different tissues though selective splicing isoforms including or excluding granule-forming low complexity regions [164]. Although to our knowledge a genome-wide analysis is still outstanding, these anecdotal examples hint that alternative splicing may operate to alter disorder-RNA interactions in a global manner (Fig. 3).

RNA-binding activity can also be modulated by competitive or cooperative interactions (Table 1, Fig. 3). The ability of some disordered regions to mediate protein-protein or protein-DNA interactions in addition to protein-RNA interactions could provide additional means to regulate RBP function. Therefore, disordered regions, although neglected for decades, have the potential to emerge as dynamic mediators of RNA biology.

Conclusions
Why disorder?
We have discussed the contribution of RS-, RG-, and K/ R-rich, disordered regions to RNA interactions, and given examples of how they participate in co- and post-transcriptional regulation of RNA metabolism; how defects in these interactions can lead to disease; and how disorder in RBPs can be utilised by viruses during their infection cycle. Disordered regions are emerging as malleable, often multifunctional RNA-binding modules whose interactions with RNA range from non-specific to highly selective with defined target sequence or structural requirements (Fig. 3). How specificity is generated for RNA sequences or structures by disordered RNA-binding regions remains to be determined. Specific interactions with defined RNA structures have been demonstrated in some instances. It seems likely that specificity and affinity can be increased by oligomerisation and through the combinatorial modular architecture of RBPs. Disorder may be a spatially cost-effective way of encoding general affinity for RNA and/or structural flexibility to enable co-folding in presence of the target RNA, thus allowing multiple binding solutions not easily achievable by structured domains. Because disorder-mediated interaction with RNA typically relies on physicochemical properties of short stretches of sequence, they can be easily regulated through post-translational modifications. Disorder may also endow special properties such as propensity to form RNA granules and interact with other RBPs. Here we have grouped the RNA-binding disordered regions based on their amino acid composition. It is possible that other functional RNA-binding motifs with unobvious sequence patterns remain to be discovered.

Outstanding questions
Much remains to be learnt about disorder-mediated protein-RNA interactions. How do disordered regions interact with RNA? How many functionally relevant disorder-RNA interactions exist? Can more refined motifs be identified among the different classes of RNA-binding disordered regions? Are there further subclasses of motifs within RS-, RG-, basic, and other RNA-binding disordered regions with distinct binding characteristics? How is RNA binding regulated post-translationally, by alternative splicing, or by competitive interactions with other biomolecules? How do mutations in disordered regions involved in RNA binding cause disease? Fundamental principles of disorder-RNA interactions are likely to have close parallels to what has been elucidated for protein-protein and protein-DNA interactions, where disorder-mediated regulation has received much more attention over the past decade [111, 165–170]. Thus, the conceptual framework to start answering questions on the role of protein disorder in RNA binding already has a firm foundation.

Concluding statement
Structure-to-function paradigm [171] has persisted long in the field of protein-RNA interactions. In this review, we have highlighted the important role that disordered regions play in RNA binding and regulation. Indeed, the recent studies on mammalian RNA-binding proteomes place disordered regions at the centre of the still expanding universe of RNA-protein interactions. It is thus time to embark on a more systematic quest of discovery for the elusive functions of disordered protein regions in RNA biology.

Additional file
Additional file 1: Figure S1. Properties of RNA-binding, disordered proteins. Disorder and charge profiles for proteins listed in Table 1. The disordered, RNA-binding regions (RBR) are marked in blue in the left panel, and their sequence given in the right panel. Amino acid sequence, GO terms, and annotations for protein domains, isoforms, and post-translational modifications (PTMs) were extracted from Uniprot [250]. Disorder was calculated using IUPred [172] using default values. Score above 0.4 indicates the region is intrinsically disordered (in physiological conditions). Charge was calculated using EMBOSS charge [251] using default values. PTMs: A, acetylation; M, methylation; P, phosphorylation; O, other. See Table 1 for literature references for each protein. (PDF 904 kb)

Abbreviations
ARM: arginine-rich motif; dsRBD: double-stranded RNA-binding domain; GAR repeat: glycine-arginine-rich repeat; KH domain: K-homology domain; RBD: RNA-binding domain; RBP: RNA-binding protein; RGG-box: arginine-glycine-glycine-box; RRm: RNA recognition motif; RS repeat: arginine-serine repeat.

Competing interests
The authors declare that they have no competing interests.
References
1. Mitchell SF, Parker R. Principles and properties of eukaryotic mRNA. Mol Cell. 2014;54:547–58.
2. Müller-Mänicke M, Neugebauer KM. How cells get the message: dynamic assembly and function of mRNA-protein complexes. Nat Rev Genet. 2013;14:275–87.
3. Lukong KE, Chang K-W, Khandjian EW, Richard S. RNA-binding proteins in human genetic disease. Trends Genet. 2009;25:44–55.
4. Licatalosi DD, Darnell RB. Splicing regulation in neurologic disease. Neuron. 2006;52:93–101.
5. Wurth L, Gebauer F. RNA-binding proteins, multifaceted translational regulators in cancer. Biochim Biophys Acta. 1849;2015:881–6.
6. Neelamraju Y, Hashemikhabir S, Janga SC. The human RBPome: From genes and proteins to human disease. J Proteomics. 2015;127(Pt A):61–70.
7. Castello A, Fischer B, Hentze MW, Preiss T. RNA-binding proteins in Mendelian disease. Trends Genet. 2013;29:318–27.
8. Lunde BM, Moore C, Varani G. RNA-binding proteins: modular design for efficient function. Nat Rev Mol Cell Biol. 2007;8:479–90.
9. Nicastro G, Taylor IA, Ramos A. KH-RNA interactions: back in the groove. Curr Opin Struct Biol. 2015;30:62–70.
10. Cikry A, Aliaín FH-T. From structure to function of rna binding domains. In: Madame Curie Bioscience Database [Internet]. Austin (TX): Landes Bioscience; 2000-2013. http://www.ncbi.nlm.nih.gov/books/NBK653528/.
11. Abbas YM, Pichlmair A, Görø MW, Superti-Furga G, Nagar B. Structural basis for viral 3’-PPP-RNA recognition by human IFIT proteins. Nature. 2013;494:60–4.
12. Gerstberger S, Hafner M, Tuschl T. A census of human RNA-binding proteins. Nat Rev Genet. 2008;9:44–55.
13. Beckmann BM, Horos R, Fischer B, Castello A, Eichelbaum K, Allekamp A-M, et al. The RNA-binding proteomes from yeast to man harbours conserved enigmRBPs. Nat Commun. 2015;6:10127.
14. Anantharaman V, Koonin EV, Aravind L. Comparative genomics and evolution of proteins involved in RNA metabolism. Nucleic Acids Res. 2002;30:1427–64.
15. Hogan DJ, Rordan DP, Gerber AP, Herschlag D, Brown PO. Diverse RNA-binding proteins interact with functionally related sets of mRNAs, suggesting an extensive regulatory system. PLoS Biol. 2008;6:e255.
16. Muckenthaler MU, Galy B, Hentze MW. Systemic iron homeostasis and the iron-responsive element/iron-regulatory protein (IRE/IRP) regulatory system. PLoS Biol. 2008;6:e255.
17. Walden WE, Seleznева AI, Dupuy J, Volbeda A, Fontecilla Camps JC, Theil L019434/1) granted to A.C., and by a Wellcome Trust senior research fellowship to I.D. (grant number 096144). We thank Prof. Matthias Hentze for discussion.
18. Oberstrass FC, Auweter SD, Erat M, Hargous Y, Henning A, Wenter P, et al. Cell Communication and Signaling
19. Kiledjian M, Dreyfuss G. Primary structure and binding activity of the hnRNP L protein: binding RNA through RGG box. EMBO J. 1992;11:2655–64.
20. Xie H, Vucetic S, Iakoucheva LM, Oldfield CJ, Dunker AK, Uversky VN, et al. Functional anthology of intrinsic disorder. 1. Biological processes and functions of proteins with long disordered regions. J Proteome Res. 2007;6:1882–98.
21. Lobley A, Swindells MB, Orenco CA, Jones DT. Inferring function using patterns of native disorder in proteins. PLoS Comput Biol. 2007;3:e162.
22. Tsvetanova NG, Klass DM, Salzman J, Brown PO. Proteome-wide search reveals unexpected RNA-binding proteins in Saccharomyces cerevisiae. PLoS One. 2010;5:e12671.
23. Scherrett T, Mittal N, Janga SC, Gerber AP. A screen for RNA-binding proteins in yeast indicates dual functions for many enzymes. PLoS One. 2010;5:e15496.
24. Castello A, Fischer B, Eichelbaum K, Horos R, Beckmann BM, Strein C, Davey NE, Humphreys DT, Preiss T, Steinmetz LM, Krijgsveld J, Hentze MW. Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell. 2012;149:393–406.
25. Balitz AG, Munschauer M, Schwanhäusser B, Vasilé A, Murakawa Y, Schuele M, et al. The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. Mol Cell. 2012;46:674–90.
26. Kwon SC, YH, Eichelbaum K, Föhr S, Fischer B, You KT, Castello A, Krijgsveld J, Hentze MW, Kim VN. The RNA-binding protein repertoire of embryonic stem cells. Nat Struct Mol Biol. 2013;20:112–20.
27. Shepard PJ, Hertel KJ. The SR protein family. Genome Biol. 2009;10:10;242. doi:10.1186/gb-2009-10-10-242. Epub 2009 Oct 27. PMID:19857271.
28. Long JC, Caceres JF. The SR protein family of splicing factors: master regulators of gene expression. Biochem J. 2009;417(1):15-27. doi:10.1042/BJ20081501. PMID:19061484.
29. Howard JM, Sanford JR. The RNAiassence family: SR proteins as multifaceted regulators of gene expression. Wiley Interdiscip Rev. RNA. 2015;6:93–110.
30. Zheng XY, Wang P, Han J, Rosenfeld MG, Fu XD. SR proteins in vertical integration of gene expression from transcription to RNA processing to translation. Mol Cell. 2009;35:1–10.
31. Chen M, Marley J. Mechanisms of alternative splicing: insights from molecular and genetics approaches. Nat Rev Mol Cell Biol. 2009;10:741–54.
32. Gravelle BR, Hertel KJ, Maniatis T. A systematic analysis of the factors that determine the strength of pre-mRNA splicing enhancers. EMBO J. 1998;17:6747–56.
33. Haynes C, Iakoucheva LM. Serine/arginine-rich splicing factors belong to a class of intrinsically disordered proteins. Nucleic Acids Res. 2006;34:305–12.
34. Xiang S, Gapsys V, Kim H-Y, Bessonov S, Hisao H-H, Möhlmann S, Klaukien V, Fischer R, Becker S, Urlaub H, Lührmann R, de Groot B, Zweckstetter M. Phosphorylation drives a dynamic switch in serine/arginine-rich proteins. Structure. 2013;21:2163–74.
35. Shen H, Han JLC, Green MR. Argininerich-seine-rich domains bound at splicing enhancers contact the branchpoint to promote prespliceosome assembly. Mol Cell. 2004;13(3):367–76. PMID:14967144.
36. Shen H, Green MR. A pathway of sequential arginine-serine-rich domain-splicing signal interactions during mammalian spliceosome assembly. Mol Cell. 2004;16:263–73.
37. Shen H, Green MR. RS domain-splicing signal interactions in splicing of U12-type and U2-type introns. Nat Struct Mol Biol. 2007;14:597–603.
38. Jamros MA, Aubol BE, Keshwani MM, Zhang Z, Stamm S, Adams JA. Intra-domain cross-talk regulates serine-arginine protein kinase 1-dependent phosphorylation and splicing function of transformer 2B. J Biol Chem. 2015;290:17269–81.
39. Gravelle BR. A protein interaction domain contacts RNA in the prespliceosome. Mol Cell. 2004;13:302–4.
40. Boucher L, Ouzounis CA, Enright AJ, Blencowe BJ. A genome-wide survey of RS domain proteins. RNA. 2001;7:1693–701.
41. Burgute BD, Peche VS, Stockelberg A-L, Göckner G, Gaßten B, Gehring NH, Noegel AA. NKAP is a novel RS-related protein that interacts with RNA and RNA binding proteins. Nucleic Acids Res. 2014;42:3177–93.
42. Chen DJ, Li Z, Yang Q, Zhang J, Zhai Z, Shu H. Identification of a nuclear protein that promotes NF-kappaB activation. Biochem Biophys Res Commun. 2003;310:720–4.
43. Pajerowski AG, Nguyen C, Aghajanian H, Shapiro MI, Shapiro VS. NKAP is a transcriptional repressor of notch signaling and is required for T cell development. Immunity. 2009;30:696–707.
44. Chang CK, Hsu Y-L, Chang Y-H, Chao F-A, Wu M-C, Huang Y-S, Hu C-K, Chang C-K, Hsu Y-L, Chang Y-H, Chao F-A, Wu M-C, Huang Y-S, Hu C-K, Chang C-K, Hsu Y-L, Chang Y-H, Chao F-A, Wu M-C, Huang Y-S, Hu C-K, Chang C-K, Hsu Y-L, Chang Y-H, Chao F-A, Wu M-C, Huang Y-S, Hu C-K, Chang C-K, Hsu Y-L, Chang Y-H, Chao F-A, Wu M-C, Huang Y-S, Hu C-K, Chang C-K, Hsu Y-L, Chang Y-H, Chao F-A, Wu M-C, Huang Y-S, Hu C-K, Chang C-K, Hsu Y-L, Chang Y-H, Chao F-A, Wu M-C, Huang Y-S, Hu C-K, Chang C-K, Hsu Y-L, Chang Y-H, Chao F-A, Wu M-C, Huang Y-S, Hu C-K, Chang C-K, Hsu Y-L, Chang Y-H, Chao F-A, Wu M-C, Huang Y-S, Hu C-K, Chang C-K, Hsu Y-L, Chang Y-H, Chao F-A, Wu M-C, Huang Y-S, Hu C-K, Chang C-K, Hsu Y-L, Chang Y-H, Chao F-A, Wu M-C, Huang Y-S, Hu C-K, Chang C-K.
Schaeffer C, Bardoni B, Mandel JL, Ehresmann B, Ehresmann C, Moine H. The characterization of SAF-A, a novel nuclear DNA binding protein from HeLa cells with high affinity for nuclear matrix/scaffold attachment DNA elements. EMBO J. 1992;11:3431–40.

Rong H, Fackeldegger FO, Renz A, Rampserger U, Richter A. Characterization of SAF-A, a novel nuclear DNA binding protein from HeLa cells with high affinity for nuclear matrix/scaffold attachment DNA elements. EMBO J. 2002;21:112–4.

Xiao R, Tang P, Yang B, Huang J, Zhou Y, Shao C, Li H, Sun H, Zhang Y, Fu X-D. Nuclear matrix factor hnRNP U/SAF-A et al. directs a global control of alternative splicing by regulating U2 snRNA maturation. Mol Cell. 2012;45:6556–68.

Hasegawa Y, Brockdoff N, Kawano S, Tsutsui K, Tsutui K, Nakagawa S. The matrix protein hnrnp U is required for chromosomal localization of Xist RNA. Dev Cell. 2010;19:469–76.

Puvvula PK, Desetty RD, Pineau P, Marchio A, Moon A, Dejean A, Bischof O. Long noncoding RNA PANDA and scaffold-attachment-factor SAFα control senescence entry and exit. Nat Commun. 2014;5:5323.

Golovanov AP, Hautbergue GM, Tintaru AM, Lian LY, Wilson SA. The solution structure of REF1-2 reveals interdomain interactions and regions involved in binding mRNA export factors and RNA. RNA. 2006;12:1933–48.

Hautbergue GM, Hung M-L, Golovanov AP, Lian L-Y, Wilson SA. Mutually exclusive interactions drive handover of mRNA from export adaptors to TAP. Proc Natl Acad Sci U S A. 2008;105:5154–9.

Braun IC, Rohrbach E, Schmitt C, Izaurralde E. TAP binds to the constitutive transport element (CTE) through a novel mRNA-binding motif that is sufficient to promote CTE-dependent RNA export from the nucleus. EMBO J. 1999;18:1953–65.

Stutz F, Bach A, Doerk T, Braun IC, Séraphin B, Wilm M, Bork P, Izaurralde E. REF, an evolutionary conserved family of hnRNp-like proteins, interacts with TAP/Mex67p and participate in mRNA nuclear export. RNA. 2000;6:638–50.

Stein C, Alleaume A-M, Rothbauer U, Hentze MW, Castello A. A versatile assay for RNA-binding proteins in living cells. RNA. 2014;20:721–31.

Zenklusen D, Vinciguerra P, Strahm Y, Stutz F. The yeast hnRNp L/Hp recognizes scaffold attachment region DNA. Mol Cell Biol. 2000;20:7480–1.

Rodrigues JP, Rode M, Gatfield D, Blencowe BJ, Carmo-Fonseca M, Izaurralde E. REF proteins mediate the export of spliced and unspliced mRNAs from the nucleus. Proc Natl Acad Sci U S A. 2001;98:1030–5.

Chen E, Joseph S. Fragile X mental retardation protein: A paradigm for translational control by RNA-binding proteins. Biochimie. 2015;114:147–54.

Pfeffer BE, Huber KM. The state of synapses in fragile X syndrome. Dev Cell. 2011;18:796–804.

Zhang Y, Gaetano CM, Williams KR, Bassell GJ, Mihalescu M-R. FMRP interacts with G-quadruplex structures in the 3′-UTR of its dendritic target Shank1 mRNA. RNA Biol. 2014;11:1364–74.

Phan AT, Kuryavyi V, Darnell JC, Sargentov A, Majumdar A, Inlin S, Raslin T, Polonskaia A, Chen C, Clain D, Ramb D, Patel DJ. Structure-function studies of FMRP RGG peptide recognition of an RNA duplex-quadruplex junction. Nat Struct Mol Biol. 2011;18:467–71.

Darnell JC, Van Driesche SJ, Zhang C, Hung KYS, Mele A, Fraser CE, Stone EF, Chen C, Fak JJ, Chi SW, Licatalosi DD, Richter JD, Ramb D. FMRP stalls ribosomal translation on mRNAs linked to synaptic function and autism. Cell. 2011;146:247–61.

Ascano M, Mukherjee N, Bandaru P, Miller JB, Nusbaum JD, Corcoran DL, et al. FMRP targets distinct mRNA sequence elements to regulate protein expression, promote CTE-dependent RNA export from the nucleus. Nature. 2012;492:382–6.

Das RK, Ruff KM. Pappu RV. Relating sequence encoded information to form and function of intrinsically disordered proteins. Curr Opin Struct Biol. 2015;32:102–12.

Bayer TS, Booth LN, Knudsen SM, Ellington AD. Arginine-rich motifs present multiple interfaces for specific binding by RNA. RNA. 2005;11:1848–57.

Ott M, Geyer M, Zhou Q. The control of HIV transcription: keeping RNA polymerase II on track. Cell Host Microbe. 2011;10:426–35.

Bayer P, Kraft M, Echtar A, Westendorp M, Frank R, Rösch P. Structural studies of HIV-1 Tat protein. J Mol Biol. 1995;247:529–35.

Calhan BJ, Bicanalana S, Hudson D, Frankel AD. Analysis of arginine-rich peptides from the HIV-1 Tat protein reveals unusual features of RNA-protein recognition. Genes Dev. 1995;9:201–10.

Long KS, Crotters DM. Interaction of human immunodeficiency virus type 1 Tat-derived peptides with TAR RNA. Biochemistry. 1995;34:8885–9.

Aboul-ela F, Karn J, Varani G. The structure of the human immunodeficiency virus type-1 TAR RNA reveals principles of RNA recognition by Tat protein. J Mol Biol. 1995;253:313–32.

Calhan BJ, Tidor B, Bicanalana S, Hudson D, Frankel AD. Arginine-mediated RNA recognition: the arginine fork. Science. 1991;252:878–9.

Das RK, Ruff KM. Pappu RV. Relating sequence encoded information to form and function of intrinsically disordered proteins. Curr Opin Struct Biol. 2015;32:102–12.

Bayer TS, Booth LN, Knudsen SM, Ellington AD. Arginine-rich motifs present multiple interfaces for specific binding by RNA. RNA. 2005;11:1848–57.

Ott M, Geyer M, Zhou Q. The control of HIV transcription: keeping RNA polymerase II on track. Cell Host Microbe. 2011;10:426–35.

Bayer P, Kraft M, Echtar A, Westendorp M, Frank R, Rösch P. Structural studies of HIV-1 Tat protein. J Mol Biol. 1995;247:529–35.

Calhan BJ, Bicanalana S, Hudson D, Frankel AD. Analysis of arginine-rich peptides from the HIV-1 Tat protein reveals unusual features of RNA-protein recognition. Genes Dev. 1995;9:201–10.

Long KS, Crotters DM. Interaction of human immunodeficiency virus type 1 Tat-derived peptides with TAR RNA. Biochemistry. 1995;34:8885–9.

Aboul-ela F, Karn J, Varani G. The structure of the human immunodeficiency virus type-1 TAR RNA reveals principles of RNA recognition by Tat protein. J Mol Biol. 1995;253:313–32.

Tao J, Frankel AD. Electrostatic interactions modulate the RNA-binding and transcriptional specificities of the human immunodeficiency virus and simian immunodeficiency virus Tat proteins. Proc Natl Acad Sci U S A. 2005;102:6849–54.

Puglisi JD, Chen L, Blanchard S, Frankel AD. Solution structure of a bovine immunodeficiency virus Tat-TAR peptide-RNA complex. Science. 1995;270:200–3.

Smith CA, Calabro V, Frankel AD. An RNA-binding chameleon. Mol Cell. 2000;5:1067–76.

Truant R, Cullen BR. The arginine-rich domains present in human immunodeficiency virus type 1 Tat and Rev function as direct importer beta-dependent nuclear localization signals. Mol Cell Biol. 1999;19:1210–7.

Kuzumbayeva M, Dilley K, Sardo L, Hu W-S. Life of p53: how full-length HIV-1 RNAs become packaged genomes in the viral particles. Virology. 2014;454–455:362–70.

Rausch JW, Le Grice SFJ. HIV Rev assembly on the rev response element (RRE): a structural perspective. Viruses. 2015;7:3053–75.

Casu F, Duggan BM, Henning M. The arginine-rich binding motif of HIV-1 Rev is intrinsically disordered and folds upon RRE binding. Biophys J. 2013;105:1004–17.

Tan R, Chen L, Buettner JA, Hudson D, Frankel AD. RNA recognition by an isolated alpha helix. Cell. 1993;73:1031–40.

Tan R, Frankel AD. Costabilization of peptide and RNA structure in an HIV Rev peptide-RRE complex. Biochemistry. 1994;33:14579–85.
99. Battiste JL, Mao H, Rao NS, Tan R, Muhandiram DR, Kay LE, Frankel AD, Williamson JR. Alpha helix-RNA major groove recognition in an HIV-1 Rev peptide-RRE RNA complex. Science. 1996;273:1547–51.

100. Wilkinson TA, Zhu L, Hu W, Chen Y. Retention of conformational flexibility in HIV-1 Rev-RNA complexes. Biochemistry. 2004;43:16153–60.

101. Daugheidy MD, D’Orso J, Frankel AD. A solution to limited genomic capacity: using adaptable binding surfaces to assemble the functional HIV Rev oligomer on RNA. Mol Cell. 2008;31:824–34.

102. Jayaraman B, Crosby DC, Horner C, Ribiero I, Mavro D, Frankel AD. RNA-directed remodeling of the HIV-1 protein Rev orchestrales assembly of the Rev-Rev response element complex. Elife. 2014;3:e04120.

103. Bai Y, Tambe A, Zhou K, Doudna JA. RNA-guided assembly of Rev-RRE nuclear export complexes. Elife. 2014;3:e03656.

104. Ye X, Gorin A, Frederick R, Hu W, Majumdar A, Xu W, McGlendon G, Ellington A, Patel DJ. RNA architecture dictates the conformations of a bound peptide. Chem Biol. 1999;6:657–69.

105. Harada K, Martin SS, Frankel AD. Selection of RNA-binding peptides in vivo. Nature. 1996;380:175–9.

106. Tan R, Frankel AD. A novel glutamine-RNA interaction identified by screening libraries in mammalian cells. Proc Natl Acad Sci U S A. 1998;95:4247–52.

107. Harada K, Martin SS, Tan R, Frankel AD. Molding a peptide into an RNA site by in vivo peptide evolution. Proc Natl Acad Sci U S A. 1999;94:11887–92.

108. Iwazaki T, Li X, Harada K. Evolvability of the mode of peptide binding by an RNA. RNA. 2005;11:1:364–73.

109. Rammelt C, Bilen B, Zavolan M, Keller Y, Breslauer KJ, Remeta DP, Minetti J, Privalov PL, Dragan AI, Crane-Robinson C. The extended arms of DNA-RNA binding domains: a tale of tails. Trends Biochem Sci. 2006;31:547–50.

110. Iwahara J, Zweckstetter M, Clore GM. NMR structural and kinetic mechanism. Proc Natl Acad Sci U S A. 2010;107:2746–50.

111. Vazquez D, Lev Y. Intrinsically disordered regions as affinity tuners in protein-DNA interactions. Mol Biosyst. 2012;8:457–67.

112. Shoemaker BA, Portman JJ, Wolynes PG. Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. Proc Natl Acad Sci U S A. 2000;97:8869–73.

113. Trizic E, Lev Y, Wolynes PG. Capillarity theory for the fly-casting mechanism. Proc Natl Acad Sci U S A. 2010;107:2746–50.

114. Halford SE, Szczelinski MD. How to get from A to B: strategies for analysing protein motion on DNA. Eur Biophys J. 2002;31:251–67.

115. El J, Li GW, Xie XS. Probing transcription factor dynamics at the single-molecule level in a living cell. Science. 2007;316:1191–4.

116. Iwahara J, Zweckstetter M, Clore GM. NMR structural and kinetic mechanism. Proc Natl Acad Sci U S A. 2010;107:2746–50.

117. Gowens DM, Wilson GJ, Halford SE. Measurement of the contributions of 1D and 3D pathways to the translocation of a protein along DNA. Proc Natl Acad Sci U S A. 2005;102:15885–8.

118. Vazquez D, Lev Y. DNA search efficiency is modulated by charge composition and distribution in the intrinsically disordered tail. Proc Natl Acad Sci U S A. 2010;107:21004–9.

119. Crane-Robinson C, Dragan AJ, Privalov PL. The extended arms of DNA-binding domains: a tale of tails. Trends Biochem Sci. 2006;31:547–50.

120. Privalov PL, Dragan AI, Crane-Robinson C, Breslauer KJ, Remeta DP, Minetti J, Privalov PL. The extended AT-hook is a novel RNA binding motif. RNA Biol. 2015;12:864–76.

121. Brangwynne CP, Mitchson TJ, Hyman AA. Active liquid-like behavior of nucleoli determines their size and shape in Xenopus laevis oocytes. Proc Natl Acad Sci U S A. 2011;108:4334–9.

122. Benecke AG, Elebrecht S. RNA-Mediated Regulation of HMGA1 Function. Biogeosciences. 2015;12:943–57.

123. Filarsky M, Ziller K, Araya I, Villar-Garea A, Merkl R, Längst G, Németh A. The extended AT-hook is a novel RNA binding motif. RNA Biol. 2015;12:864–76.

124. Brangwynne CP, Mitchson TJ, Hyman AA. Beyond stereospecificity: liquids and mesoscale organization of cytoplasm. Dev Cell. 2011;21:14–6.

125. Nott T, Petsalaki E, Farber P, Jervis D, Fussner E, Plchowietz A, Crapps TD, Bazett-Jones DP, Pavson T, Forman-Kay JD, Baldwin AJ. Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. Mol Cell. 2015;57:936–47.

126. Weber SC, Brangwynne CP. Getting RNA and protein in phase. Cell. 2012;149:1188–91.

127. Hyman AA, Weber CA, Jülicher F. Liquid-liquid phase separation in biology. Annu Rev Cell Dev Biol. 2014;30:39–58.

128. Erickson SL, Lykke-Andersen J. Cytoplasmic mRNP granules at a glance. J Cell Sci. 2011;124:293–7.

129. Weil TT, Parson RM, Herpers B, Soetaert J, Veeneendaal T, Xanthakis D, Dobbie IM, Halstead JM, Hayashi R, Rabouille C, Davis I. Drosophila patterning is established by differential association of mRNAs with P bodies. Nat Cell Biol. 2012;14:1305–13.

130. Jain S, Wheeler JR, Walters RW, Agrawal A, Bascic A, Parker R. ATPase-modulated stress granules contain a diverse proteome and substructure. Cell. 2013;154:478–89.

131. Lin Y, Pottorl DS, Rosen MK, Parker R. Formation and maturation of phase-separated liquid droplets by RNA-binding proteins. Mol Cell. 2013;50:208–19.

132. Buchan JR. mRNP granules. Assembly, function, and connections with disease. RNA Biol. 2014;11:1019–30.

133. Ramassamy M, Taylor JP, Parker R. Altered ribostasis: RNA-protein granules in degenerative disorders. Cell. 2013;154:727–36.

134. Kato M, Han TW, Xie S, Shi K, Du X, Wu LC, et al. Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. Cell. 2012;149:753–67.

135. Han TW, Kato M, Xie S, Wu LC, Mi-raezi H, Pei J, et al. Cell-free formation of RNA granules: bound RNAs identify features and components of cellular assemblies. Cell. 2012;149:768–79.

136. Uversky VN, Kuznetsova IA, Tuvorovov KK, Zaslavsky B. Intrinsically disordered proteins as crucial constituents of cellular aqueous two phase systems and coacervates. FEBS Lett. 2015;589:15–22.

137. Elbaum-Garfinkle S, Kim Y, Szczepaniak K, Chen CC-H, Eckmann CR, Myong SH, Brangwynne CP. The disordered P granule protein LAP-1 drives phase separation into droplets with tunable viscosity and dynamics. Proc Natl Acad Sci U S A. 2015;112:1788–94.

138. Lee C, Zhang H, Baker AE, Occhipinti P, Borsuk ME, Gladfelter AS. Protein aggregation behavior regulates cyclin transcript localization and cell-cycle control. Dev Cell. 2013;25:572–84.

139. King OD, Gitter AD, Shorer J. The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease. Brain Res. 2012;1462:61–80.

140. Han TW, Kato M, Xie S, Wu LC, Mi-raezi H, Pei J, et al. Cell-free formation of RNA granules: bound RNAs identify features and components of cellular assemblies. Cell. 2012;149:753–67.

141. Alberts S, Halfmann R, King Q, Kapila A, Lindquist S. A systematic survey identifies prions and illuminates sequence features of prionogenic proteins. Cell. 2009;137:146–58.

142. Reijns MAM, Alexander RD, Spiller MP, Beggs JD. A role for Q/N-rich aggregation-prone regions in P-body localization. J Cell Sci. 2008;121:2463–72.

143. Kosjivakov S, Maharana S, Mateju D, Malinovska L, Nüske E, Poser I, Richter D, Alberti S. Promiscuous interactions and protein disguises determine the material state of stress-inducible RNP granules. Elife. 2015;4:e06807.

144. Zhang H, Elbaum-Garfinkle S, Langdon EM, Taylor N, Occhipinti P, Bridges AA, et al. RNA Controls polyQ protein phase transitions. Mol Cell. 2015;59:220–30.

145. Buchan JR, Parker R. Eukaryotic stress granules: the ins and outs of translation. Mol Cell. 2009;36:397–407.

146. Van Roey K, Gibson TJ, Davey NE. Motif switches: decision-making in cell systems and coacervates. FEBS Lett. 2015;589:15–22.
ine rich motifs? J Biochem. 2010;147:

Dyson HJ. Roles of intrinsic disorder in protein-nucleic acid interactions. Mol 

Fuxreiter M, Simon I, Bondos S. Dynamic protein-DNA recognition: beyond 

Tompa P, Davey NE, Gibson TJ. Linear motifs confer functional diversity 

Shiina N, Nakayama K. RNA granule assembly and disassembly modulated 

Dolzhanskaya N, Merz G, Denman RB. Alternative splicing modulates protein 

Sitar A, Devys D, Weber C, Mandel JL. Alternative splicing of exon 14 determines 

Weatheritt RJ, Davey NE, Gibson TJ. Linear motifs confer functional diversity 

Denman RB, Sung YJ. Species-specific and isoform-specific RNA binding of 

Romero PR, Zaidi S, Fang YY, Uversky VN, Radijcovic P, Oldfield CJ, Cortese 

Uversky VN, Dave V, Iakoucheva LM, Malaney P, Metello SJ, Pathak RR, et al. 

Van Roey K, Uyar B, Weatheritt RJ, Dinkel H, Seiler M, Budz A, Gibson TJ, 

Tompa P, Davey NE, Gibson TJ, Babu MM. A million peptide motifs for the 

Bosysiot 2012;897–104. 

Wright PE, Dyson HJ. Intrinsically unstructured proteins: re-assessing the 

Dobosz½nyi Z, Ciszomk V, Tompa P, Simon I. The pairwise energy content 

Xu X, Yang D, Ding J-H, Wang W, Chu P-H, Dalton ND, Wang H-Y, 

Naylor O, Strating W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

David CJ, Boyne AR, Millhouse SR, Manley JL. The RNA polymerase II C- 

Chang C-K, Sue S-C, Yu T-H, Hsieh C-M, Tsai C-K, Chiang Y-C, Lee S-J, Hsiao 

Tonnicliffe RB, Hauertgme GM, Wilson SA, Kalra P, Golovanov AP. 

Thandapani P, Song J, Gandin V, Cai Y, Rouleau SG, Garant J-M, Boisvert F-M, 

Ina S, Tsunekawa N, Nakamura A, Noce T. Expression of the mouse Aven 

Sobue G, Ohno K. Position-specific binding of FUS to nascent RNA regulates 

Sobue G, Ohno K, Yasuura H, Tarami M, Kato S, Kato H, Ichihara S, Ishihara 

Sharma DJ, Mose T, Todd AG, Eggleton P, Lorson CL, Young PJ. Identification 

Thandapani P, Song J, Gandin V, Cai Y, Rouleau SG, Garant J-M, Boisvert F-M, 

Vatsan O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strätling W, Bourquin JP, Stajglar L, Lindemann L, Jasper H, et al. 

Jayvelin O, Strä
198. Viphakone N, Hautbergue GM, Walsh M, Chang C-T, Holland A, Folco EG, Reed R, Wilson SA. TREX exposes the RNA-binding domain of Nfs1 to enable miRNA export. Nat Commun. 2012;3:1006.

199. Ghisolfi L, Joseph G, Amalic F, Eard M. The glycin-rich domain of nucleolin has an unusual supersedimentary structure responsible for its RNA-helix-stabilizing properties. J Biol Chem. 1992;267:2959–9.

200. Bouvet P, Diaz JI, Kindbinder K, Madjar JI, Amalic F. Nucleolin interacts with several ribosomal proteins through its RGG domain. J Biol Chem. 1998;273:19025–9.

201. Kanhoush R, Beenders B, Perrin C, Moreau J, Bellini M, Penrad-Mobayed M. Novel domains in the hnrNP G/ RBMX protein with distinct roles in RNA binding and targeting nascent transcripts. Nucleus. 2010;1:109–22.

202. Dichmann DS, Fletcher RB, Harland RM. Expression cloning in Xenopus identifies RNA-binding proteins as regulators of embryogenesis and RBmx as necessary for neural and muscle development. Dev Dyn. 2008;237:1755–66.

203. Adamson B, Smogorzewska A, Sigiloff FD, King RW, Elledge SJ. A genome-wide homologous recombination screen identifies the RNA-binding protein RBMX as a component of the DNA-damage response. Nat Cell Biol. 2012;14:318–28.

204. Tseng-Asayih E, O'Sullivan LA, Grützer FS, Onnebo SWN, Lewis RS, Delbridge ML, Marshall Graves JA, Ward AC. RBMX gene is essential for development in zebrafish. Dev Dyn. 2005;234:682–8.

205. Shashi V, Xie P, Schoch K, Goldstein DB, Howard TD, Fris B, Schwartz CE, Corini K, Sliva S, Allen A, Need AC. The RBMX gene as a candidate for the Shashi X-linked intellectual disability syndrome. Clin Genet. 2015;88:890–6.

206. Mazeyrat S, Saut N, Mattei MG, Mitchell MJ. RBMX evolved on the Y chromosome from a ubiquitously transcribed X-Y identical gene. Nat Genet. 1999;22:224–6.

207. Schlepke AW, Rethwilm A. Nuclear localization of foamy virus Gag precursor protein. J Virol. 1994;68:4946–54.

208. Yu SF, Edelmann K, Baer C, Weiss M, Gu Z, Schipper M, Kerner R, Sullmann H, Provenzano M, Yaso M-H, Brons B, Korbel J, Schlimont T, Sauter G, Ellis P, Riss C, Santoro R. BAZ2A (TIP60) is involved in epigenetic alterations in prostate cancer and its overexpression predicts disease recurrence. Nat Genet. 2014;37:22–30.

209. Zhang J, Lieu YK, Ali AM, Pensno A, Reggio KS, Rabadan R, et al. Disease-associated mutation in SRSF2 misregulates splicing by altering RNA-binding affinities. Proc Natl Acad Sci U S A. 2015;112:E276–32.

210. Kanhoush R, Beenders B, Perrin C, Moreau J, Bellini M, Penrad-Mobayed M. Novel domains in the hnrNP G/ RBMX protein with distinct roles in RNA binding and targeting nascent transcripts. Nucleus. 2010;1:109–22.

211. Zhang P, Abdelmohsen K, Liu Y, Tominaga-Yamanaka K, Yoon J-H, Ioannis M, Mullers E, Uhlig T, Stirnnagel K, Fiebig U, Zentgraf H, Lindemann D. Novel functional and conformational differences of the arginine methylation of the human immunodeficiency virus type 1 Tat protein by PRMT6 negatively affects Tat Interactions with both cyclin T1 and the Tat transactivation region. J Biol Chem. 2007;282:4226–34.

212. Deng L, La Fuente De C, Fu P, Wang L, Donnelly R, Wade JD, Lambert P, Li H, Lee CG, Kashanchi F. Acetylation of HIV-1 Tat by CBP/P300 increases transcription of integrated HIV-1 genome and enhances binding to core histones. Virology. 2000;277:278–95.

213. D’Oroso I, Frankel AD. Tat acetylation modulates assembly of a viral-host RNA-protein transcription complex. Proc Natl Acad Sci U S A. 2009;106:3101–6.

214. Zhang P, Casaday-Potts R, Precht P, Schneezer M, Loret E, Cole PA, Verdin E, Ott M. Acetylation of Tat defines a cyclinT1-independent step in HIV transcription. Mol Cell. 2003;12:167–76.

215. Lapointe CP, Wickens M. The nucleic acid-binding domain and translational...
241. Kaneko S, Li G, Son J, Xu CF, Margueron R, Neubert TA, Reinberg D. Phosphorylation of the PRC2 component Ezh2 is cell cycle-regulated and up-regulates its binding to ncRNA. Genes Dev. 2010;24:2615–20.

242. Zhang Y, Yang X, Gui B, Xie G, Zhang D, Shang Y, Liang J. Corepressor protein CDYL functions as a molecular bridge between polycomb repressor complex 2 and repressive chromatin mark trimethylated histone lysine 27. J Biol Chem. 2011;286:42414–25.

243. Zhao J, Sun BK, Erwin JA, Song JJ, Lee JT. Polycomb proteins targeted by a short repeat ma to the mouse X chromosome. Science. 2008;322:750–6.

244. Yue MM, Lv K, Meredith SC, Martindale JL, Gorospe M, Schugger L. Novel RNA-binding protein P311 binds eukaryotic translation initiation factor 3 subunit b (eIF3b) to promote translation of transforming growth factor β1-3 (TGF-β1-3). J Biol Chem. 2014;289:33971–83.

245. Fernandez-Chamorro J, Pineiro D, Gordon JMB, Ramajo J, Francisco-Velilla R, Macias MJ, et al. Identification of novel non-canonical RNA-binding sites in Gemin5 involved in internal initiation of translation. Nucleic Acids Res. 2014;42:5742–54.

246. Dimaano C. RNA Association defines a functionally conserved domain in the nuclear pore protein Nup153. J Biol Chem. 2001;276:45349–57.

247. Ball JR. The RNA binding domain within the nucleoporin Nup153 associates preferentially with single-stranded RNA. RNA. 2004;10:19–27.

248. Bonasio R, Lecona E, Narendra V, Voigt P, Parisi F, Kluger Y, Reinberg D. Interactions with RNA direct the Polycomb group protein SCML2 to chromatin where it represses target genes. Elife. 2014;3:e02637.

249. Zoabi M, Nadar-Ponniah PT, Khouy-Haddad H, Usaj M, Budowski-Tal I, Haran T, Henn A, Mandel-Gutfreund Y, Ayoub N. RNA-dependent chromatin localization of KDM4D lysine demethylase promotes H3K9me3 demethylation. Nucleic Acids Res. 2014;42:3026–38.

250. UniProt Consortium: UniProt: a hub for protein information. Nucleic Acids Res 2015, 43(Database issue):D204–12.

251. Rice P, Longden I, Bleasby A: EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet 2000, 16:276–277.