

THE DOUBLE-STRANDED-RNA-BINDING MOTIF: INTERFERENCE AND MUCH MORE

Bin Tian*, Philip C. Bevilacqua[‡], Amy Diegelman-Parente[§] and Michael B. Mathews*

Abstract | RNA duplexes have been catapulted into the spotlight by the discovery of RNA interference and related phenomena. But double-stranded and highly structured RNAs have long been recognized as key players in cell processes ranging from RNA maturation and localization to the antiviral response in higher organisms. Penetrating insights into the metabolism and functions of such RNAs have come from the identification and study of proteins that contain the double-stranded-RNA-binding motif.

RNA molecules assume a wide variety of secondary and tertiary structures owing to interactions between their bases and with other ligands (FIG. 1). Base pairing of the four nucleobases ranges from standard Watson–Crick pairs to various non-canonical pairs including wobble, Hoogsteen, reverse Hoogsteen, reverse Watson–Crick and sheared pairs^{1,2}. RNA secondary and tertiary structures provide the basis for interactions with myriad proteins. Proteins can recognize RNA by binding to single-stranded RNA, perfectly duplexed RNA helices, internal loops, bulges, hairpin loops, pseudoknots and a plethora of complex tertiary structures^{3–8}. The simplest of these structures, duplexed RNA — or double-stranded (ds)RNA — has profound regulatory effects *in vivo*. Its recognition by proteins, the roles and evolution of these proteins, and the consequences of the dsRNA–protein interactions, are the subject of this article.

Characteristically, structural and/or sequence features of RNA molecules interact with specific modules of protein structure known as domains or motifs. RNA-binding motifs include larger motifs of 50–100 amino acids, such as the dsRNA-binding motif (dsRBM), the zinc finger and the ribonucleoprotein (RNP) or RNA-RECOGNITION MOTIF (RRM), as well as shorter motifs of 10–25 amino acids, such as the arginine-rich motif (ARM), K-homology (KH) motif and arginine–glycine–glycine (RGG) box^{3,6–10}. In recent years, increasing

numbers of structures of RNA–protein complexes have been solved by X-ray crystallography and NMR techniques. Perhaps most notably, high-resolution structures of the small and large subunits of the ribosome have provided the structural details for nearly 50 proteins that interact with RNA^{11–14}. These structures delineate an array of RNA-binding motifs that complement the array of RNA secondary and tertiary motifs.

First recognized in 1992 (REFS 15–17), the dsRBM (also known as dsRNA-binding domain (dsRBD)) is responsible for many, although not all (supplementary information S1 (table)), interactions of proteins with RNA duplexes. Here, we discuss the structural features of the dsRBM, its interaction with RNAs, its distribution in major species and the functions of those proteins that contain the dsRBM.

The dsRBM defined
The dsRBM adopts an α – β – β – α topology structure, where α signifies a stretch of α -helix and β signifies a β -strand (FIG. 2a). It forms an α/β sandwich global fold, which is similar to the ubiquitous RRM RNA-binding motif⁸. Equivalent three-dimensional structures have been determined for dsRBMs from widely divergent proteins. These include *Escherichia coli* ribonuclease RNASE III (REF. 18), the *Drosophila melanogaster* protein **Staufen**^{19,20}, the human RNA-dependent protein kinase **PKR**²¹ and the *Saccharomyces cerevisiae* RNase-III

*Department of Biochemistry and Molecular Biology, New Jersey Medical School, University of Medicine and Dentistry of New Jersey, 185 South Orange Avenue, PO Box 1709, Newark, New Jersey 07101-1709, USA.

[‡]Department of Chemistry, The Pennsylvania State University, 104 Chemistry Building, University Park, Pennsylvania 16802, USA.

[§]Division of Mathematics and Natural Sciences, Altoona College, The Pennsylvania State University, 3000 Ivyside Park, Altoona, Pennsylvania 16601, USA.

Correspondence to M.B.M.
e-mail: mathews@umdnj.edu
doi:10.1038/nrm1528

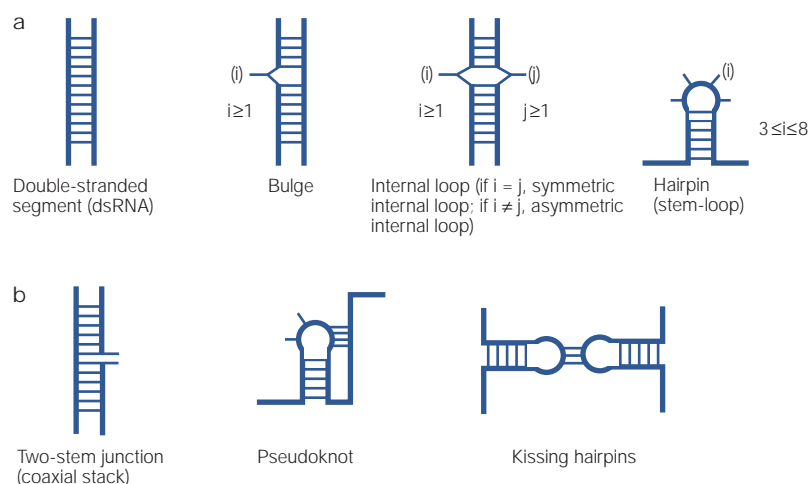


Figure 1 | Secondary and tertiary structural elements in RNA. **a** | Secondary structural elements. Secondary structural elements are defined by the pattern of paired and unpaired bases and are typified by double helices, bulges, internal loops and hairpin loops. Bulges have one or more unpaired nucleotides on one strand only, and the bulge can either stack into or out of the strand depending on the sequence context¹²⁵. Internal loops have one or more unpaired nucleotides on both strands, although the 'unpaired' nucleotides often engage in non-Watson-Crick base pairing: if the number of unpaired nucleotides (indicated by 'i' or 'j') on the two strands are equal, the loop is symmetrical, otherwise it is asymmetrical. Hairpins, or stem-loops, have unpaired nucleotides at the apex of a double-stranded stem, which itself can have structural defects. **b** | Tertiary structural elements. The production of complex RNA structures by folding is typically hierarchical, with higher-order structures assembling from preformed secondary structures^{126,127}. Two-stem junctions, or coaxial stacks, occur when two stems co-linearize; furthermore, three- and four-stem junctions, in which two of the stems co-linearize, are also found (not shown). Pseudoknots involve base pairing between unpaired nucleotides bounded within a stem-loop (that is, either a hairpin loop, a bulge or an internal loop) and nucleotides that flank the stem-loop. Kissing hairpins involve base pairing between two hairpin loops. Furthermore, there are also many other RNA tertiary interactions, including base triples and quadruples, and interactions with 2'-hydroxyls and metal ions.

endonuclease Rnt1 (REF. 22), which have all been examined by NMR. X-ray crystallographic structures have been published for a dsRBM of *Xenopus laevis* RNA-binding protein A (Xlrpba)²³ as well as the entire RNase III from *Aquifex aeolicus*²⁴. Residues that are important for the structure and function of the dsRBM are conserved, particularly in the C-terminal third of the motif (FIG. 3). For example, conserved hydrophobic residues in the α -helices pack along one side of the three-stranded anti-parallel β -sheet to maintain the overall structure, and the three RNA contact regions (FIG. 3 and see below) show high conservation of hydrophilic residues that support the dsRBM-dsRNA interactions. Although the motif is typically 65–70 amino acids in length, insertions and deletions can be accommodated, particularly in the β -strands and the linking loops between helix 1 and strand 1 (loop 1), strands 1 and 2 (loop 2), and strands 2 and 3 (loop 3). Therefore, some dsRBMs are much larger than average. For example, the second dsRBM of *D. melanogaster* Staufen is 83 amino-acids long, with a proline-rich insertion sequence in loop 2 (the insertion seems to affect the properties of the domain; see below). Of the three RNA contact regions into which the dsRBM structure can be divided, as described below, all of the sequence length variability resides in region 2.

RNA-RECOGNITION MOTIF (RRM). This motif is among the most common in eukaryotic proteins. It usually comprises 80–90 amino acids, forming a β - α - β - β - α - β structure. Many proteins that contain RBMs bind RNA in a sequence-specific manner.

RNASE III
An endoribonuclease that cleaves RNA substrates containing regular double-helical or stem-loop structures.

dsRBM-RNA interaction

The mode of interaction of the dsRBM with dsRNA has been elucidated through X-ray crystallographic analyses of a complex that comprises a short synthetic 10-bp RNA duplex and a dsRBM of Xlrpba²³ or RNase III (REF. 24), and by NMR analyses of RNA complexes with the dsRBMs of Staufen or Rnt1 (REFS 20,22). In all cases, the dsRBM binds exclusively to one face of the helix without wrapping around it, as illustrated in FIG. 2b. The two dsRBM motifs that are found in a single unit cell of the crystal span 16 bp, but bury only 1600–1700 Å² of the RNA chain, which reflects the limited contact surface of protein residues with the RNA^{23,24}. One consequence of this restricted interface is that it is possible to pack multiple dsRBMs tightly along an extended helix, as has been shown biochemically with PKR^{25,26}. Although, at present, we do not have any high-resolution structures of a multiple-dsRBM-containing protein bound to dsRNA, NMR studies show that the 15-residue linker that connects the 2 dsRBMs of PKR is highly flexible²¹, which indicates that it should be possible to position the 2 dsRBMs in a close-packed fashion on the same dsRNA molecule. The recent RNase-III-dsRNA structure²⁴ also shows how a dsRBM is accommodated in a protein that contains other functional domains — in this case, an endonuclease domain. A structurally flexible seven-residue linker is found between the dsRBM and the endonuclease domain, leading to a highly modular protein.

In the dsRBM-RNA complex, the dsRNA is in the A-form. The compact A-form RNA helix is characteristic of dsRNA, whereas dsDNA usually adopts the more open B-form. Three regions of the dsRBM make contact with the dsRNA. Region 1 contacts involve dsRBM helix α 1, region 2 contacts involve loop 2, and region 3 contacts involve loop 4 and helix α 2. Travelling along the dsRNA helix, these regions occur in the order 1–3–2 (REFS 20,22–24) (FIG. 2a). As the protein binds to only one face of the dsRNA helix and spans 16 bp, which represents about 1.5 turns of the helix, the following pattern is found: region-1 interactions occur at the minor groove of the helix, region-3 interactions at the major groove and region-2 interactions at the minor groove, which has returned to the same face of the helix after translating 16 bp. Regions 1 and 2 supply the most extensive set of contacts with the RNA and these primarily involve the 2'-hydroxyls of the ribose sugar. Region 3 makes six direct or water-mediated contacts with non-bridging oxygen residues of the phosphodiester backbone. On a related note, an unusually small number of ionic interactions — only one (PKR) or two (Xlrpba) — were found between the dsRBM and a model 20-bp dsRNA^{23,26}. Collectively, the absence of a strong ionic binding component, key interactions with 2'-hydroxyl groups of RNA and a preference for the groove dimensions of A-form rather than B-form helices prevent promiscuous binding to dsDNA or RNA-DNA hybrids^{23,26,27}.

The A-form shape of dsRNA remains largely unchanged after protein binding^{22–24}, which is unusual as RNA-protein binding is often accompanied by

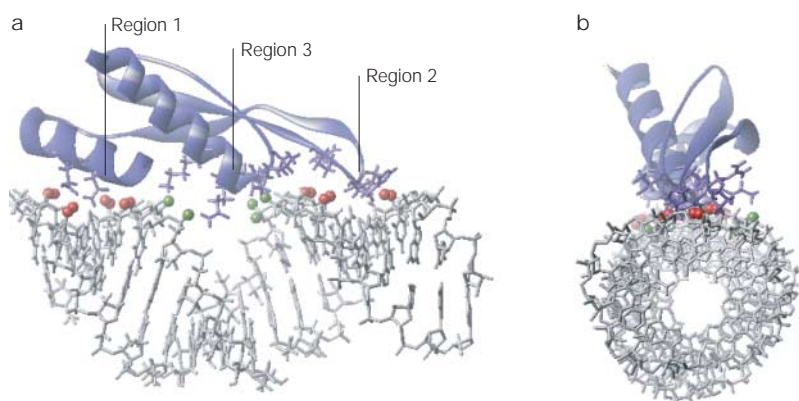


Figure 2 | Two views of the dsRBM structure. **a** | Side-on view showing the three major regions of contact with double-stranded (ds)RNA. Regions 1 and 2 of the dsRNA-binding motif (dsRBM) contact the minor groove, whereas region 3 contacts the major groove. Ryter and Schultz²³ identified important 2'-hydroxyls (red) and patterns of phosphates (green), which help to discriminate against dsDNA and RNA–DNA hybrids. Note that regions 1 and 2 involve protein contacts with 2'-hydroxyls in the minor groove, whereas region 3 involves contacts with phosphates in the major groove. Protein residues that are important for binding are shown as sticks. Some of the contacts to the RNA are water mediated. **b** | End-on view showing that all contacts are located on one face of the dsRNA helix. Both views were generated from the crystal structure of the complex between dsRBM2 of *Xenopus laevis* RNA-binding protein A (Xlrpba) and dsRNA²³. Protein database accession number 1DI2; image generated using WebLab ViewerLite version 3.2.

conformational changes, or induced fit, of the RNA, the protein, or both^{6,28,29}. Part of the reason for this might lie in the rigidity of dsRNA^{30,31}. It should be noted that the RNA bending reported in the NMR structure of a Staufen dsRBM–hairpin complex²⁰ has been recently attributed to the dynamics induced by a very short dsRNA stem, and does not seem to be a general phenomenon²². On the other hand, dsRBMs from several proteins interact with RNAs that harbour significant secondary structural imperfections. PKR, for example, which has two dsRBMs, can bind oligonucleotide APTAMERS that contain bulges, internal loops, hairpin loops and multi-stem junctions³², as well as natural cellular and viral RNAs that contain a wealth of secondary and tertiary features^{33–35} including pseudoknots^{36,37}. Staufen, which has five dsRBMs, interacts with mRNA 3' UNTRANSLATED REGIONS (3' UTRs) that contain bulges and internal loops²⁰. Dicer, which has a single dsRBM, binds and cleaves precursor microRNAs (pre-miRNAs), which are RNA stem-loops that also contain bulges and internal loops^{38–40}.

But how do dsRBM proteins bind to RNAs with structural imperfections? The definitive answer awaits high-resolution structural studies of such complexes. However, experiments on bent RNAs that contain bulges indicate that the RNA might undergo a conformational change — a feature of most RNAs in RNA–protein complexes^{6,28,29}. Gel mobility-shift experiments on HELICAL CHIMAERAS⁴¹ indicate that the RNA straightens to adopt the more linear, dsRNA-like state that the protein has evolved to bind. In this complex, the imperfections are probably merely tolerated⁴¹ (see BOX 1). The observation that the helical defects in pre-miRNAs processed by Dicer do not seem to be conserved in

sequence, size or position further supports this idea^{42–44}. The fact that the dsRBM binds to one face of the dsRNA (FIG. 2b) suggests a simple model in which the helical imperfection might merely be extruded away from the RNA–protein binding interface⁴¹. If so, an array of structural defects could be readily accommodated within the context of a dsRBM–RNA complex. In some situations, the imperfections could act as inhibitors, or ANTIDETERMINANTS, of enzyme function, as found in RNase III (REFS 24,45,46). Such antideterminants, which are present as helical imperfections or flanking RNA sequence, might explain why certain dsRNAs do not trigger RNA interference (RNAi; see BOX 2). Conversely, the dsRBM retains considerable residual flexibility in the RNA–dsRBM complex⁴⁷. The high degree of protein flexibility in the final complex seems to facilitate binding by reducing the entropic penalty from rigidification of protein side chains.

The sequence-specificity paradox

Some dsRBM proteins, such as Staufen and the ADENOSINE DEAMINASES that act on RNA (ADARs), exhibit a high degree of specificity in their interactions with mRNAs. For example, mRNA editing by ADAR2 is both highly specific and essential for life in mice⁴⁸. In the case of Staufen, binding requires a 'kissing' interaction between two molecules of the bicoid 3' UTR, thereby indirectly imparting specificity to complex formation⁴⁹. From a structural perspective, although there are several minor groove contacts between the dsRBM and bases from the RNA — which might seem to offer opportunities for sequence-specific recognition^{23,24} — most of these contacts are indirect, water-mediated interactions, and the dsRBM proteins are able to adjust their hydrogen-bonding patterns to any sequence. The few direct protein–base interactions are with minor groove hydrogen-bond acceptors that would be present in any of the four bases and would, therefore, provide little or no sequence information. In Xlrpba, however, there is a Pro140 contact that might be specific for a GC base pair²³. In RNase III that is complexed with the same 10-bp dsRNA, a further 15 hydrogen-bonding interactions with the RNA were described, but only 1 possible sequence-specific interaction, between a guanine and Gln161, was noted²⁴. So, it is more probable that any sequence-specific recognition is accomplished through a distortion of the dsRNA helix, which might allow amino-acid side-chain entry into the major groove (BOX 1), a situation that could occur for imperfect dsRNAs.

On the other hand, so far, there is no strong evidence to indicate that interactions with perfect or imperfect dsRNAs are sequence specific. One exception, a high-resolution structure in which the dsRBM makes several contacts with the minor groove face of a UUCG tetraloop, has been reported, although this might be a red herring²⁰. A stem-loop in which the stem was shortened to 12 bp led to the interaction of a physiologically non-relevant UUCG hairpin loop with a dsRBM from Staufen. However, no sequence-specific recognition was found for the stem, and the exceptionally short helix — typically 16 bp is required (see above) — might have

APTAMER

An RNA, either engineered or natural, that forms a precise three-dimensional structure and selectively binds a target molecule, for example a dsRBM-containing protein.

3' UNTRANSLATED REGION

(3' UTR). This is the sequence of a messenger RNA that is located downstream of the stop codon.

HELICAL CHIMAERA

A helix in which one of the strands has a mixture of ribose and deoxyribose nucleotides.

ANTIDETERMINANTS

Sequences that block the binding of a protein to an otherwise suitable site, first defined in tRNAs and later in RNase-III substrates.

ADENOSINE DEAMINASE

An enzyme that catalyses adenosine-to-inosine conversion in an RNA substrate, a process also known as RNA editing.

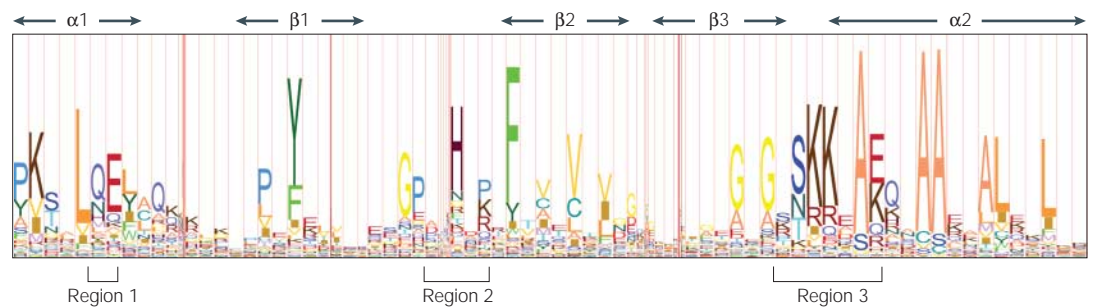


Figure 3 | Sequence logo of the dsRBM. All double-stranded-RNA-binding motifs (dsRBMs) in the InterPro database (a total of 1,428 from various species) were used to generate a hidden Markov model (HMM), which was converted to a sequence logo using HMM logos¹²⁸. Adjacent residues are separated by a line, the thickness of which indicates the probability of an insertion of random residues. The width of each residue indicates the probability of having a residue at that position; that is, narrow residues tend to be deleted at their positions. The height of a residue indicates how conserved it is among all dsRBMs. Secondary structures of the dsRBM are marked above the sequence logo. Three dsRNA interaction regions (regions 1, 2 and 3) are marked below the sequence logo. Intervening sequences between α -helices (α) and β -strands (β) are known as loops: loop 1 is between helix 1 and strand 1, loop 2 between strand 1 and strand 2, loop 3 between strand 2 and strand 3, and loop 4 between strand 3 and helix 2.

forced the interaction with the loop^{20,22}. As noted above, biochemical and structural information indicates that direct interactions of the dsRBM with RNA do not discriminate among dsRNA sequences^{20,23,24,26}. This notion is further supported by a recent report that describes a complex between Rnt1 and a physiologically relevant AGNN-tetraloop RNA-stem-loop substrate²². Strikingly, neither the highly conserved A nor the universally conserved G of the loop⁵⁰ is recognized by hydrogen bonding to Rnt1; instead, the N-terminal helix interacts with the minor groove of the tetraloop and makes non-sequence-specific contacts with the sugar–phosphate backbone and the two nonconserved loop bases. Indeed, the lack of sequence specificity could make dsRBM-containing proteins, such as PKR and Dicer, ideally suited to participate in the innate immune response^{51–53}. In summary, sequence- or structure-specific discrimination by the dsRBM seems unlikely, although it is not ruled out. The sequence selectivity seen in certain instances could be conferred by other motifs in the dsRBM protein, by other proteins that participate in the interactions or by structural imperfections in the RNA⁵⁴.

Evolution of the dsRBM proteins
dsRBMs are present, or predicted to be present, in proteins from all eukaryotes and most eubacteria, but in only one archaeon — and this is probably a result of horizontal transfer from a eubacterium. Many animal viruses, including viruses with dsDNA, dsRNA and single-stranded RNA (ssRNA) genomes^{55,56}, also encode dsRBM-containing proteins. So, dsRBM-containing proteins are widespread, although not universal, in nature. They often contain other protein domains (TABLE 1), which confer various functions and in some cases exhibit species specificity. For example, several protein domains, such as the adenosine deaminase (ADAR) domain, are associated with the dsRBM exclusively in the animal kingdom, whereas several other domains, such as the NLI-interacting-factor-like phosphatase and metallo-phosphoesterase, are associated

with the dsRBM exclusively in plants. Humans and mice have several dsRBM-containing proteins that are not found in *D. melanogaster* or *Caenorhabditis elegans*.

Among these associated domains, the RNase-III domain coexists with the dsRBM to the broadest evolutionary extent (TABLE 1). *E. coli* RNase III was the first dsRBM-containing protein to be purified and studied, and the mechanism of action of these enzymes is still an active topic of research^{54,57,58}. In bacteria, the dsRBM generally coexists with the RNase-III domain in RNase-III proteins. Exceptionally, in two species of mycoplasma, *Mycoplasma genitalium* and *Mycoplasma pneumoniae*, the RNase-III protein lacks a dsRBM. As these species have very small genomes (*M. genitalium* has the smallest known bacterial genome of only 580 kb)⁵⁹, it is conceivable that dsRBMs were lost during the reductive evolution of these parasitic bacteria. Evidently, the dsRBM is dispensable for the functions of RNase III in these species, which is consistent with the finding that the *E. coli* enzyme is active, albeit weakly, in the absence of its dsRBM⁶⁰. Similarly, trypanosomes contain RNase-III-like proteins that lack dsRBMs (REF. 61; B. T., unpublished observations).

The dsRBM has been found in only one of over a dozen completely sequenced archaeal genomes — that of *Methanosarcina mazei* — and the RNase-III domain is notably absent from all sequenced archaeal genomes. *M. mazei* has adopted a large number of genes, as much as one third of its genome, by lateral gene transfer from bacteria⁶², indicating that dsRBMs probably did not evolve in the Archaea. Interestingly, the solitary *M. mazei* dsRBM resides in a protein that contains an RNase-II (or RNB) domain, rather than an RNase-III domain. The RNase-II domain has exoribonuclease function and is present in bacteria, eukaryotes and one other archaeon (*Halobacterium* NRC-1). Possibly, the coupling of the dsRBM and RNase-II domains furnishes *M. mazei* with an enzyme activity that is similar to that of RNase III. This could be considered an example of convergent molecular evolution.

Box 1 | Themes for sequence-specific RNA recognition

Perfectly double-stranded (ds)RNA has a wide and shallow minor groove, which makes it an obvious target for protein binding, and a narrow and deep major groove, which is inaccessible to amino acids. As the major groove harbours the majority of the sequence-specific information in RNA, perfect dsRNA is recalcitrant to sequence-specific recognition¹¹⁶. However, secondary structure imperfections, such as bulges and internal loops (FIG. 1), widen the major groove allowing penetration by amino acids and the possibility of sequence-specific recognition¹¹⁷. It should also be noted that, although sequence-specific recognition of dsRNA in the minor groove is less common than in the major groove, it is also possible^{118,119}.

The wide spectrum of RNA–protein structures illustrates the importance of structural imperfections in allowing sequence-specific recognition^{3,5–10}. Several specific themes for RNA–protein recognition have been suggested, including ‘ β -sheet binders’, in which a β -sheet surface interacts with completely or partially single-stranded RNA, and ‘groove binders’, of which the dsRBM is a member^{20,22–24}, in which a protein structure is inserted into a groove of the RNA⁶. The ‘groove binder’ class often involves sequence-specific recognition in the major groove; in these instances, the major groove is typically widened by a secondary structural defect. Three simple principles of molecular recognition of RNA emerge from the examination of these structures. First, structural features of the RNA can attract features of the protein through hydrogen bonding and van der Waals interactions. Second, structural features of the RNA can sterically exclude features of the protein by preventing a large protein structure from entering a crevice on the RNA¹²⁰. Third, structural features of the RNA can be tolerated within the complex with the protein; that is, they neither attract nor exclude protein. In all cases, the RNA and/or the protein often undergo conformational changes to arrive at the final bound conformation^{6,28,29}.

Multiple dsRBMs in a single protein

As with many protein domains, evolutionarily advanced organisms possess more dsRBM-containing proteins⁵⁶. Concomitantly, in higher eukaryotes, these proteins have a wider variety of domains, as well as multiple dsRBMs (TABLE 1). At the other extreme, known viral proteins contain only one dsRBM, which reflects the economy of viral genomes. Evidently, a single dsRBM suffices for these proteins to carry out their tasks in the viral life cycle, most notably in the evasion of cellular anti-viral mechanisms (see below).

The presence of multiple dsRBMs can impact on protein function in two main ways. First, it can lead to higher dsRNA-binding affinity as a result of cooperation between dsRBMs. For example, the second dsRBM of PKR has only marginal dsRNA-binding activity, but it cooperates with the first dsRBM to achieve better binding, and might help to define the minimum length of dsRNA that activates PKR⁶³. Second, it allows for functional divergence of the motifs. Indeed, many of the dsRBMs in proteins that contain multiple dsRBMs do not have strong affinity for dsRNA and instead undertake other functions. The dsRBM can participate in protein–protein interactions, which can be intramolecular, as in PKR, or intermolecular. The third dsRBM of PACT (a protein activator of PKR) does not bind dsRNA, but has a vital role in activating PKR^{64,65}. The second and fifth dsRBMs of Staufen, which also do not bind dsRNA, are involved in microtubule-dependent mRNA localization and de-repression of protein translation, respectively⁶⁶.

When multiple dsRBMs reside on the same protein, they are often positioned close to each other. Whereas the intervening sequence is important for the dsRNA-binding activity of some dsRBM-containing proteins, its length varies from protein to protein. For orthologous proteins that contain multiple dsRBMs, corresponding dsRBMs show greater sequence similarities to their counterparts in the orthologues than to neighbouring dsRBMs in the same protein, indicating descent from common protein ancestors that contain multiple dsRBMs⁵⁶. For paralogous genes that contain multiple dsRBMs, the resemblance between corresponding dsRBMs can extend to the splicing pattern in the transcripts that encode them. For example, the location of dsRBM introns is well conserved between corresponding dsRBMs of PACT and TRBP (TAR RNA-binding protein), and between those of NF90 (nuclear factor-90) and SPNR (spindle perinuclear protein).

Conversely, several human dsRBM-containing proteins, such as ADAR and NRF (NF- κ B-repressing factor), show a high degree of sequence similarity among the dsRBMs on the same protein, and the splicing patterns of these dsRBMs are identical. In such cases, it seems that the multiple dsRBMs of each of these proteins originate from a single common ancestral dsRBM⁵⁶ and were derived by domain duplication within the same gene. Although it is difficult to generalize about all proteins that contain multiple dsRBMs as many of them are so ancient that their evolutionary paths are muddled, this finding seems to indicate that the need to have multiple dsRBMs located close to each other is a driving force that leads to the duplication of dsRBMs.

Functions of the dsRBM protein family

Proteins that contain dsRBMs are functionally diverse (see FIG. 4 for representative proteins from bacteria, yeast and metazoa). Whereas many of these proteins have transactions with dsRNA in which the RNA duplex is a substrate for enzymatic action or a cargo for transport, in others, the ligand is a regulator (TABLE 2). Several dsRBM proteins cluster in small, functionally related families with biochemical roles in transcription, RNA processing, mRNA localization and translation. A detailed list of their functions is given in **supplementary information S2 (table)**. As discussed below, several proteins are implicated in cellular defense systems against foreign RNA. A second loosely related group serves as regulators of cellular and viral gene expression, whereas the functions of other proteins can only be extrapolated from the motifs that they contain in addition to the dsRBM.

Cellular defense systems. One recurrent theme is RNA cleavage and RNAi. Judging by its wide distribution in present-day species, the function that is most primitively associated with the dsRBM is that of the structure-specific ribonuclease RNase III, which is present in almost all organisms except the Archaea (see above). RNase-III enzymes can be divided into three classes

Box 2 | RNA interference

One of the biggest recent breakthroughs in biology has been the discovery of RNA interference (RNAi), which contributes to a wide range of developmental, cellular-defensive and regulatory processes (reviewed in REF. 121). RNAi is a cellular process in which small dsRNAs (21–25 nt) induce sequence-specific degradation of cognate mRNAs. First identified in *Caenorhabditis elegans*²², RNAi has since been seen in virtually every eukaryotic system apart from *Saccharomyces cerevisiae* and some trypanosomes (reviewed in REF. 123). Associated functions include the transcriptional regulation of gene expression, heterochromatin formation and centromere maintenance. Despite its evolutionary conservation, important distinctions are apparent among RNAi pathways in different species. First, different proteins, including dsRBM proteins, are involved; for example, no R2D2/RDE-4 homologue has been found in humans. Second, in some species, such as plants and worms, RNAi entails an amplification step that is lacking in humans and fruitflies. This seems to be attributable to the presence or absence of an RNA-dependent RNA polymerase (RdRP). Third, RNAi in plants and worms can have systemic effects, which result from the transport of the signal across cells.

RNAi seems to be a key element in cellular defenses against viral infection and parasitic DNAs that reside in cell genomes. Cells exploit another, related process to control gene expression at the level of translation in a developmental and cell-type-specific manner (reviewed in REF. 124). This process involves microRNAs (miRNAs), which are small dsRNAs with imperfect duplexes. Genetic and biochemical evidence indicates that the RNAi pathway is intimately involved in the actions of miRNA. The decision to undergo RNAi-induced RNA degradation or translational repression might lie in the degree of complementarity between the RNA and its target.

based on their domain structures (reviewed in REFS 58,67). In bacteria and single-celled eukaryotes such as the yeast *S. cerevisiae*, there is only one dsRBM and RNase-III motif in the class-I RNase III. In higher eukaryotes, the dsRBM is found in two further types of ribonuclease, **Drosha** (class II) and **Dicer** (class III), both of which have two RNase-III motifs and are involved in RNAi. Furthermore, in eukaryotes, a mitochondrial ribosomal protein (**MRPL44** in humans) resembles the class-I RNase-III enzyme. So far, its function is unknown.

Until now, at least three dsRBM-containing proteins have been implicated in RNAi. Dicer, which is the most extensively studied of these proteins, is a nuclease that cuts perfect (and imperfect; see below) dsRNA into suitable segments of 21–24 bp, which are known as small interfering RNAs (siRNAs)^{68,69}. These are incorporated into an RNA-induced silencing complex (RISC), which then cleaves mRNA in a sequence-specific fashion. Drosha has recently been reported to participate in this phenomenon by carrying out the first steps in generating miRNAs^{70,71}, which have an important role in RNAi-related post-transcriptional regulation. In this process, Drosha cleaves the primary nuclear transcripts that contain miRNA sequences (pri-miRNAs), yielding precursor stem-loops of ~70 nucleotides (nt) with helical imperfections (pre-miRNAs) that are further cleaved by Dicer to give rise to the 22-nt effector miRNA molecules. In addition to these two nucleases belonging to the RNase-III family, a third type of dsRBM protein, which is represented by RDE-4 in *C. elegans* and **R2D2** in *D. melanogaster*, seems to function as an adaptor in RNAi⁵⁸. Despite overall sequence similarity and common interacting proteins, RDE-4 and R2D2 seem to behave differently in the RNAi pathway. Whereas RDE-4 was implicated in the initiation of RNAi in *C. elegans*, R2D2 was shown to have a role in bridging the initiation and effector steps in the *D. melanogaster* RNAi pathway^{72,73}. Putatively, RDE-4 and R2D2 could be prototypes for proteins that confer specificity on enzymatic reactions that are catalysed by other enzymes, such as RISC.

Antiviral defenses that are induced by interferon constitute a related theme⁷⁴. Two proteins, **ADAR1** and **PKR**, are mediators of the pathway. They are found in humans and mice but not in *D. melanogaster* or *C. elegans*. The adenosine deaminases represent another class of dsRBM-containing proteins for which the dsRNA ligand is a substrate (TABLE 2). These enzymes convert A residues to I (inosine, which pairs like guanine) and are responsible for the editing of many cellular and viral RNAs. Remarkably, they are essential for the proper protein coding of some mRNAs in the central nervous system⁷⁵. There are four ADAR genes in mammals, one of which is inducible by interferon (**ADAR1**). The functional differentiation among the ADARs is not yet entirely clear, but they seem to have evolved from adenosine deaminases that act on tRNAs (ADATs), which are present in many species, including bacteria, yeasts and metazoans⁷⁵. In the case of PKR, the dsRNA is a modulator rather than a substrate (TABLE 2). PKR is a dsRNA-regulated protein kinase that controls translation by phosphorylating the initiation factor eIF2, and it probably also modulates signal transduction and other processes^{76,77}. PKR is activated by viral dsRNA, either the viral genome itself or double-stranded intermediates or transcripts that are generated during infection. Furthermore, both ADARs and PKR have been reported to interact with Alu transcripts in human cells^{78,79}.

Whereas RNAi has been shown to confer sequence-specific RNA-based intracellular immunity in cells^{80,81}, interferon-induced genes constitute an innate immune response that is protein based and RNA-sequence non-specific. For organisms that possess both pathways, discrimination between the two has traditionally been ascribed to differences in the RNA-length requirement. Conventional wisdom holds that RNAi and related phenomena (see BOX 2) involve short dsRNAs of <30 nt, whereas interferon-induced anti-viral defense systems are triggered by longer duplexes^{82,83}. It is not known whether this distinction is conferred by the number of dsRBMs that a protein contains (Drosha and Dicer have one dsRBM, whereas PKR and ADARs have two or

Table 1 | Domains associated with dsRBMs in several species

Domains associated with dsRBMs*	Hs	Mm	Dm	Ce	At	Os	Sc	Ec	Ar	Vi
Ribonuclease III domain	√	√	√	√	√	√	√	√		√
DEAD/DEAH-box helicase domain	√	√	√	√	√	√				
PAZ domain	√	√	√	√	√	√				
DUF283 domain†	√	√	√	√	√	√				
Adenosine deaminase/editase domain	√	√	√	√						
FHA domain	√	√	√	√						
D111/G-patch domain	√	√	√	√						
DUS domain	√	√	√	√						
WW/Rsp5/WWP domain	√	√	√	√						
Z-α-DNA-binding domain	√	√								√
Serine/threonine protein kinase domain	√	√								
DZF domain	√	√								
Single-stranded, nucleic-acid-binding R3H domain	√	√								
NLI-interacting-factor-like phosphatase domain					√	√				
Metallo-phosphoesterase domain					√					
RNA-helicase domain (found in single-stranded viruses)										√
RdRP domain										√
Rotavirus non-structural-protein NSP3 domain										√
Ribonuclease-II domain									√	
dsRBM (>1 copy)	√	√	√	√	√	√				

The presence of proteins containing both dsRBMs and other specified domains is indicated by a tick, √. The bottom row shows the presence of proteins that contain more than one dsRBM. *Only domains in the protein families (Pfam) database are listed here. †The function of this domain is unknown. Ar, Archaea (only *Methanosarcina mazei* strain Goe1); At, *Arabidopsis thaliana*; Ce, *Caenorhabditis elegans*; Dm, *Drosophila melanogaster*; dsRBM, double-stranded-RNA-binding motif; DUS, dihydrouridine synthase; DZF, associated with dsRBM or zinc-finger; Ec, *Escherichia coli*; FHA, forkhead-associated; Hs, *Homo sapiens*; Mm, *Mus musculus*; NLI, nuclear LIM interactor; Os, *Oryza sativa*; PAZ, Piwi, Argonaute and Zwillie protein; RdRP, RNA-directed RNA polymerase; Sc, *Saccharomyces cerevisiae*; Vi, viruses.

more dsRBMs). However, things might not be quite so simple, as siRNA and short hairpin RNAs (shRNAs) have been reported to elicit a cellular immune response in human cells, albeit distinct from that induced by interferon or long dsRNA^{84–86}.

Several dsRBM proteins have regulatory roles in these two pathways. For example, PKR activity can be regulated by the cellular dsRBM-containing proteins PACT, TRBP and NF90, as well as by viral proteins exemplified by vaccinia virus E3L. The E3L protein, which contains a dsRBM and a Z-α-DOMAIN (also present in the interferon-inducible form of ADAR), sequesters dsRNA as an antidote to host antiviral defense mechanisms by blocking PKR activation and also by inhibiting the RNAi pathway^{87,88}. ADAR proteins (which are not interferon inducible) in *C. elegans* and *D. melanogaster* have been shown to antagonize the RNAi pathway by modifying dsRNAs that are involved in RNAi^{89,90}, as A-to-I modification can destabilize RNA duplex structures. These data hint at the possible interplay between RNAi and the innate immune response.

Other functions of dsRBM proteins. Some dsRBM proteins are regulators of gene expression. They function at the levels of transcription, translation or mRNA localization (TABLE 2), and might be essential for development and survival. For example, RNA helicase A (RHA), NF90, NRF and the protein SON, which binds the negative regulatory element of the hepatitis B virus,

are involved in the transcriptional regulation of both cellular and viral genes (REFS 91–94; T. W. Reichman and B. T., unpublished observations), although in many cases their precise modes of action are unclear. RHA also associates with topoisomerase IIα, which indicates a role in regulating chromatin structure⁹⁵. Staufen has been well studied in *D. melanogaster*, where it helps to establish the polarity of the embryo by localizing particular mRNAs to the poles of the oocyte, thereby establishing concentration gradients of the encoded proteins⁹⁶. One of its dsRBMs also contributes to translational control, although it apparently functions through protein–protein, rather than protein–RNA, interactions⁹⁶. Staufen is also implicated in long-term memory in flies⁹⁷ and RNA trafficking in mammalian neuronal cells⁹⁸. There are indications that it functions in the localization and transport of specific mRNAs in several metazoan cell types.

Whereas many viruses exploit the dsRBM to neutralize cell functions, others seem to use it in different ways. For example, the *D. melanogaster* C virus — an insect virus that contains a ssRNA genome — encodes a protein with a dsRBM, an RNA-directed RNA polymerase domain and an RNA helicase domain, which presumably comprises part of the viral replication or transcription machinery⁹⁶.

As dsRBM proteins function in many aspects of RNA metabolism, it is not surprising that these proteins are present in various subcellular locations.

Z-α-DOMAIN

A protein domain that binds left-handed Z-form DNA, which is believed to occur transiently in the cell during gene transcription.

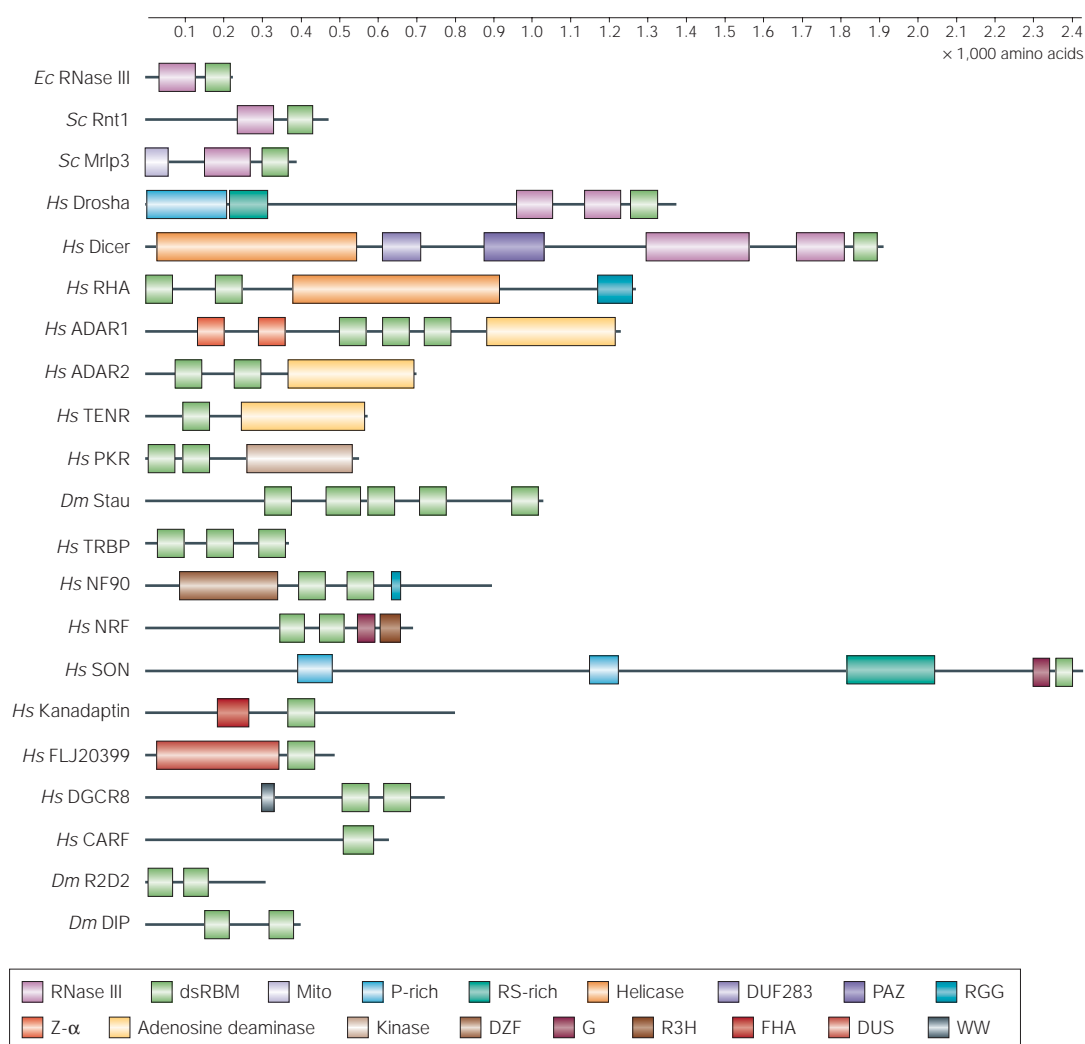


Figure 4 | **Domain structures of dsRBM proteins.** Representative double-stranded-RNA-binding motif (dsRBM)-containing proteins are selected from human (*Hs*, *Homo sapiens*), fruitfly (*Dm*, *Drosophila melanogaster*), baker's yeast (*Sc*, *Saccharomyces cerevisiae*) and *Escherichia coli* (*Ec*). Protein sequences and domains are drawn approximately to scale. ADAR, adenosine deaminase that acts on RNA; CARF, collaborator of alternate reading frame (ARF) protein; DGCR8, DiGeorge syndrome critical region gene-8; DIP, disco-interacting protein; DUS, dihydrouridine synthase; DZF, associated with dsRBM or zinc-finger; FHA, forkhead-associated; Mito, mitochondrial transit peptide; NF90, nuclear factor-90; NF-κB, nuclear factor-κB; NRF, NF-κB-repressing factor; PAZ, Piwi, Argonaute and Zwillie protein; P-rich, proline-rich region; PKR, protein kinase, interferon-inducible double-stranded-RNA dependent; RGG, arginine-glycine-glycine box; RHA, RNA helicase A; RNase, ribonuclease; RS-rich, arginine/serine-rich region; Stau, Staufen; TENR, testis nuclear RNA-binding protein; TRBP, TAR RNA-binding protein. Other domains are listed in TABLE 1.

A large proportion of dsRBM proteins contain a bipartite nuclear localization signal (NLS), and are therefore located in the nucleus⁵⁵. Interestingly, several studies indicate that the dsRBM alone can direct a protein to the nucleolus^{63,99}, where ribosome biogenesis takes place. Indeed, it has been suggested that nucleolar sequestration might modulate the activity of ADAR proteins^{99,100}. As ribosomal RNAs (rRNAs) are highly structured, it is tempting to speculate that dsRBM proteins are sequestered in the nucleolus by the interaction between the dsRBM and rRNAs. Similarly, some dsRBM proteins, such as PKR, NF90 and Xlrpba, associate with the ribosomal fraction of the cell^{101–103}. Furthermore, several dsRBM proteins, including ADAR1 (REF. 104) and NF90 (REF. 105), shuttle between

the nucleus and the cytoplasm. Whereas ADAR1 contains both canonical nuclear export and import signals^{106,107} and its nuclear export is mediated by the nuclear-export receptor CRM1 (REF. 106), the shuttling of NF90 is attributable to its association with exportin 5, a protein that also exports pre-miRNAs, some tRNAs and viral hairpin RNAs out of the nucleus^{108–112}. The second dsRBM of NF90 seems to be required for this association, and the interaction can be antagonized by dsRNAs¹⁰⁸. In fact, exportin 5 can also interact with dsRBMs from other proteins, raising the possibility that it has a general role in localizing dsRBM proteins^{108,113}. Conversely, dsRBM proteins might use this mechanism to export their RNA cargos out of the nucleus.

Table 2 | Functions of dsRBM proteins and the roles of their RNA ligands*

Role of dsRNA	Transcription	RNA processing	mRNA localization	Translation	Other	Unknown
Substrate	RHA	RNase III Drosha Dicer ADARs		ADARs Staufen		
Modulator	NF90			PKR		
Cargo			Staufen NF90			
Unknown	NRF SON	R2D2/RDE-4 NF90		TRBP PACT NF90 SPNR MRPL44	CARF SPNR	TENR Kanadapin FLJ20399 DGCR8 DIP

*Plant proteins are not included. ADAR, adenosine deaminase that acts on RNA; CARF, collaborator of alternate reading frame (ARF) protein; DGCR8, DiGeorge syndrome critical region gene-8; DIP, disco-interacting protein; dsRBM, double-stranded-RNA-binding motif; MRPL44, mitochondrial ribosomal protein L44; NF90, nuclear factor-90; NF- κ B, nuclear factor- κ B; NRF, NF- κ B-repressing factor; PACT, protein activator of PKR; PKR, protein kinase, interferon-inducible double-stranded-RNA dependent; RDE-4, RNAi defective gene-4; RHA, RNA helicase A; RNase, ribonuclease; SPNR, spindle perinuclear protein; TENR, testis nuclear RNA-binding protein; TRBP, TAR RNA-binding protein.

Whither the dsRBM?

The dsRBM is a remarkably versatile molecular tool. It is exploited by a large fraction of the organisms on the planet for a range of purposes. Its biochemical properties, including dsRNA binding, RNA annealing^{63,114} and protein–protein interactions, are used in proteins that have enzymatic functions as well as proteins that do not seem to be conventional enzymes, such as TRBP, PACT and Staufen.

The interaction between the dsRBM and RNA duplexes has been studied for over a decade, but there is still much to learn. Although a wide variety of structured RNAs have been found to interact with dsRBMs, including cellular and viral as well as physiological and pathological RNAs, new types of dsRBM-interacting RNAs will surely continue to come to light. The possibility that some dsRBMs carry out sequence-specific functions has not been entirely excluded; however, if dsRBM–RNA interactions truly are sequence independent, then the source of the specificity has to be uncovered. Some dsRBM proteins contain other RNA-binding domains, such as the RGG box in RHA and NF90, which might confer sequence specificity on the protein, but it is still unknown how these RNA-binding domains function in conjunction with the dsRBM. These observations indicate the need for further high-resolution structures of dsRBM–dsRNA complexes. In particular, structures of multiple-dsRBM-containing proteins complexed with dsRNA are needed, as are the structures of dsRBMs complexed with RNA targets of cellular and viral origin, especially those containing dsRNA imperfections.

Whether dsRBM–RNA interactions can be regulated is an open question. Protein phosphorylation has been shown to modulate the RNA-binding affinity of the dsRBMs of PKR¹¹⁵. Similar studies on other dsRBM proteins have not been reported, although several dsRBM proteins can be phosphorylated by PKR *in vitro*. The possibility that other kinases can phosphorylate and modulate the RNA-binding activity of the dsRBM has yet to be investigated. Alternative splicing

can generate a protein isoform of SON that lacks a significant portion of its dsRBM, presumably abolishing its RNA-binding activity. However, the biological significance of this is unclear. The genes that encode most dsRBMs in higher organisms contain introns within the domain sequence, and it will be interesting to see if modulation of the dsRBM by alternative splicing is widespread.

Many genes function by engaging in enzymatic complexes (sometimes called ‘somes’) and through signalling pathways. Some dsRBM proteins seem to function in a concerted manner. For example, PKR and ADAR1 can both be induced by interferon in the cellular antiviral response, and Dicer, Drosha and R2D2/RDE-4 have roles in RNA interference and related cellular processes. Whether other dsRBM proteins are involved in these or other complexes or pathways is yet to be explored, and clues might be supplied by studies that correlate the coexistence of proteins in complexes with the expression patterns of the corresponding genes.

Finally, although it is unlikely that many dsRBM proteins have been overlooked in the organisms in which genomes have been scrutinized, plants contain many under-studied dsRBM protein family members. Their existence indicates that other organisms might yield further examples of dsRBM proteins. Conspicuously, so far, many dsRBM proteins do not have clearly assigned functions ([supplementary information S2 \(table\)](#)), including proteins in metazoans (NRF, CARF, SON, and so on), Archaea and, especially, plants. The discovery of their functions is a daunting task for the years to come. Even greater challenges lie in the characterization of their RNA ligands and protein partners, and the functional consequences of these interactions. Looking back over the past few years, during which three dsRBM proteins of previously unknown functions were found to participate in RNAi — a biological process of fundamental importance — there is every reason to believe that scrutiny of the dsRBM, its ligands and the proteins that carry it, will continue to reap rich rewards.

1. Burkard, M. E., Turner, D. H. & Tinoco, I. Jr in *The RNA World* (ed. Atkins, J. F.) 675–685 (Cold Spring Harbor Laboratory Press, New York, USA, 1999).
2. Leontis, N. B., Stombaugh, J. & Westhof, E. The non-Watson–Crick base pairs and their associated isostericity matrices. *Nucleic Acids Res.* **30**, 3497–3531 (2002).
3. Draper, D. E. Protein–RNA recognition. *Annu. Rev. Biochem.* **64**, 593–620 (1995).
4. Nagai, K. RNA–protein interactions. *Curr. Opin. Struct. Biol.* **2**, 131–137 (1992).
5. Varani, G. RNA–protein intermolecular recognition. *Acc. Chem. Res.* **30**, 189–195 (1997).
6. Draper, D. E. Themes in RNA–protein recognition. *J. Mol. Biol.* **293**, 255–270 (1999).
7. Perez-Canadillas, J. M. & Varani, G. Recent advances in RNA–protein recognition. *Curr. Opin. Struct. Biol.* **11**, 53–58 (2001).
8. Hall, K. B. RNA–protein interactions. *Curr. Opin. Struct. Biol.* **12**, 283–288 (2002).
9. Burd, C. G. & Dreyfuss, G. Conserved structures and diversity of functions of RNA-binding proteins. *Science* **265**, 615–621 (1994).
10. Nagai, K. RNA–protein complexes. *Curr. Opin. Struct. Biol.* **6**, 53–61 (1996).
11. Ban, N., Nissen, P., Hansen, J., Moore, P. B. & Steitz, T. A. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* **289**, 905–920 (2000).
12. Wimberly, B. T. *et al.* Structure of the 30S ribosomal subunit. *Nature* **407**, 327–339 (2000).
13. Yusupov, M. M. *et al.* Crystal structure of the ribosome at 5.5 Å resolution. *Science* **292**, 883–896 (2001).
14. Harms, J. *et al.* High resolution structure of the large ribosomal subunit from a mesophilic eubacterium. *Cell* **107**, 679–688 (2001).
15. St Johnston, D., Brown, N. H., Gall, J. G. & Jantsch, M. A conserved double-stranded RNA-binding domain. *Proc. Natl Acad. Sci. USA* **89**, 10979–10983 (1992).
Describes the first biochemical identification and sequence alignment of the double-stranded-RNA-binding motif.
16. McCormack, S. J., Thomis, D. C. & Samuel, C. E. Mechanism of interferon action: identification of a RNA binding domain within the N-terminal region of the human RNA-dependent P1/eIF-2α protein kinase. *Virology* **188**, 47–56 (1992).
17. Green, S. R. & Mathews, M. B. Two RNA binding motifs in the double-stranded RNA activated protein kinase, DAI. *Genes Dev.* **6**, 2478–2490 (1992).
18. Kharrat, A., Macias, M. J., Gibson, T. J., Nilges, M. & Pastore, A. Structure of the dsRNA binding domain of *E. coli* RNase III. *EMBO J.* **14**, 3572–3584 (1995).
19. Bycroft, M., Grunert, S., Murzin, A. G., Proctor, M. & St Johnston, D. NMR solution structure of a dsRNA binding domain from *Drosophila* staufer protein reveals homology to the N-terminal domain of ribosomal protein S5. *EMBO J.* **14**, 3563–3571 (1995).
References 18 and 19 are two back-to-back papers that first characterized the structure of the dsRBM.
20. Ramos, A. *et al.* RNA recognition by a Staufer double-stranded RNA-binding domain. *EMBO J.* **19**, 997–1009 (2000).
21. Nanduri, S., Carpick, B. W., Yang, Y., Williams, B. R. & Qin, J. Structure of the double-stranded RNA-binding domain of the protein kinase PKR reveals the molecular basis of its dsRNA-mediated activation. *EMBO J.* **17**, 5458–5465 (1998).
22. Wu, H., Henras, A., Chanfreau, G. & Feigon, J. Structural basis for recognition of the AGNN tetraloop RNA fold by the double-stranded RNA-binding domain of Rnt1p RNase III. *Proc. Natl Acad. Sci. USA* **101**, 8307–8312 (2004).
23. Ryter, J. M. & Schultz, S. C. Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA. *EMBO J.* **17**, 7505–7513 (1998).
Atomic-level description of interactions of the dsRBM with RNA.
24. Blaszczak, J. *et al.* Noncatalytic assembly of ribonuclease III with double-stranded RNA. *Structure* **12**, 457–466 (2004).
25. Manche, L., Green, S. R., Schmidt, C. & Mathews, M. B. Interactions between double-stranded RNA regulators and the protein kinase DAI. *Mol. Cell. Biol.* **12**, 5238–5248 (1992).
Defines the dsRNA-length dependence for PKR binding and activation.
26. Bevilacqua, P. C. & Cech, T. R. Minor-groove recognition of double-stranded RNA by the double-stranded RNA-binding domain from the RNA-activated protein kinase PKR. *Biochemistry* **35**, 9983–9994 (1996).
27. Hung, M. L., Chao, P. & Chang, K. Y. dsRBM1 and a proline-rich domain of RNA helicase A can form a composite binder to recognize a specific dsDNA. *Nucleic Acids Res.* **31**, 5741–5753 (2003).
28. Williamson, J. R. Induced fit in RNA-protein recognition. *Nature Struct. Biol.* **7**, 834–837 (2000).
29. Leulliot, N. & Varani, G. Current topics in RNA-protein recognition: control of specificity and biological function through induced fit and conformational capture. *Biochemistry* **40**, 7947–7956 (2001).
30. Kebbekus, P., Draper, D. E. & Hagerman, P. Persistence length of RNA. *Biochemistry* **34**, 4354–4357 (1995).
31. Auffinger, P. & Westhof, E. Water and ion binding around r(UpA)12 and d(TpA)12 oligomers — comparison with RNA and DNA (CpG)12 duplexes. *J. Mol. Biol.* **305**, 1057–1072 (2001).
32. Bevilacqua, P. C., George, C. X., Samuel, C. E. & Cech, T. R. Binding of the protein kinase PKR to RNAs with secondary structure defects: role of the tandem A–G mismatch and noncontiguous helices. *Biochemistry* **37**, 6303–6316 (1998).
Describes the isolation, through an *in vitro* selection technique, of a family of structured RNAs that can interact with the dsRBMs of PKR.
33. Clarke, P. A., Sharp, N. A. & Clemens, M. J. Translational control by the Epstein-Barr virus small RNA *EBER-1*. *Eur. J. Biochem.* **193**, 635–641 (1990).
34. Vuysich, M., Spangord, R. J. & Beal, P. A. The binding site of the RNA-dependent protein kinase (PKR) on EB1 RNA from Epstein-Barr virus. *EMBO Rep.* **3**, 622–627 (2002).
35. Tian, B. *et al.* Expanded CUG repeat RNAs form hairpins that activate the double-stranded-RNA-dependent protein kinase PKR. *RNA* **6**, 79–87 (2000).
36. Ma, Y. & Mathews, M. B. Secondary and tertiary structure in the central domain of adenovirus type 2 VA RNA I. *RNA* **2**, 937–951 (1996).
37. Ben-Asouli, Y., Banai, Y., Pel-Or, Y., Shir, A. & Kaempfer, R. Human interferon-γ mRNA autoregulates its translation through a pseudoknot that activates the interferon-inducible protein kinase PKR. *Cell* **108**, 221–232 (2002).
38. Ketting, R. F. *et al.* Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev.* **15**, 2654–2659 (2001).
39. Knight, S. W. & Bass, B. L. A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in *Caenorhabditis elegans*. *Science* **293**, 2269–2271 (2001).
40. Lee, Y., Jeon, K., Lee, J. T., Kim, S. & Kim, V. N. MicroRNA maturation: stepwise processing and subcellular localization. *EMBO J.* **21**, 4663–4670 (2002).
41. Zheng, X. & Bevilacqua, P. C. Straightening of bulged RNA by the double-stranded RNA-binding domain from the protein kinase PKR. *Proc. Natl Acad. Sci. USA* **97**, 14162–14167 (2000).
42. Lagos-Quintana, M., Rauhut, R., Lendeckel, W. & Tuschl, T. Identification of novel genes coding for small expressed RNAs. *Science* **294**, 853–858 (2001).
43. Lau, N. C., Lim, L. P., Weinstein, E. G. & Bartel, D. P. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* **294**, 858–862 (2001).
44. Lee, R. C. & Ambros, V. An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* **294**, 862–864 (2001).
45. Calin, J.-G., Negrini, M., & Croce, C. M. MicroRNA: a new paradigm for cancer diagnosis, prognosis and therapy. *Cell* **127**, 426–438 (2006).
46. Zhang, K. & Nicholson, A. W. Regulation of ribonuclease III processing by double-helical sequence antiterminants. *Proc. Natl Acad. Sci. USA* **94**, 13437–13441 (1997).
47. Castrignano, T., Chillemi, G., Varani, G. & Desideri, A. Molecular dynamics simulation of the RNA complex of a double-stranded RNA-binding domain reveals dynamic features of the intermolecular interface and its hydration. *Biophys. J.* **83**, 3542–3552 (2002).
48. Higuchi, M. *et al.* Point mutation in an AMPA receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2. *Nature* **406**, 78–81 (2000).
Shows that the ADAR2 knockout causes a lethal phenotype in mice because of defective editing of a specific ion-channel mRNA.
49. Ferrandon, D., Elphick, L., Nusslein-Volhard, C. & St Johnston, D. Staufer protein associates with the 3'UTR of bicoid mRNA to form particles that move in a microtubule-dependent manner. *Cell* **79**, 1221–1232 (1994).
50. Nagel, R. & Ares, M. Jr Substrate recognition by a eukaryotic RNase III: the double-stranded RNA-binding domain of Rnt1p selectively binds RNA containing a 5'-AGNN-3' tetraloop. *RNA* **6**, 1142–1156 (2000).
51. Balachandran, S. *et al.* Essential role for the dsRNA-dependent protein kinase PKR in innate immunity to viral infection. *Immunity* **13**, 129–141 (2000).
52. Plasterk, R. H. RNA silencing: the genome's immune system. *Science* **296**, 1263–1265 (2002).
53. Zamore, P. D. Ancient pathways programmed by small RNAs. *Science* **296**, 1265–1269 (2002).
54. Lamontagne, B. & Elela, S. A. Evaluation of the RNA determinants for bacterial and yeast RNase III binding and cleavage. *J. Biol. Chem.* **279**, 2231–2241 (2004).
55. Saunders, L. R. & Barber, G. N. The dsRNA binding protein family: critical roles, diverse cellular functions. *FASEB J.* **17**, 961–983 (2003).
56. Tian, B. & Mathews, M. B. Phylogenetics and functions of the double-stranded RNA-binding motif: a genomic survey. *Prog. Nucleic Acid Res. Mol. Biol.* **74**, 123–158 (2003).
Survey of the genomes of organisms from several important taxa for dsRBM-containing proteins using a bioinformatics approach.
57. Nicholson, A. W. Function, mechanism and regulation of bacterial ribonucleases. *FEMS Microbiol. Rev.* **23**, 371–390 (1999).
58. Carmell, M. A. & Hannon, G. J. RNase III enzymes and the initiation of gene silencing. *Nature Struct. Mol. Biol.* **11**, 214–218 (2004).
59. Fraser, C. M. *et al.* The minimal gene complement of *Mycoplasma genitalium*. *Science* **270**, 397–403 (1995).
60. Sun, W., Jun, E. & Nicholson, A. W. Intrinsic double-stranded-RNA processing activity of *Escherichia coli* ribonuclease III lacking the dsRNA-binding domain. *Biochemistry* **40**, 14976–14984 (2001).
61. Aphasizhev, R. *et al.* Isolation of a U-insertion/deletion editing complex from *Leishmania tarentolae* mitochondria. *EMBO J.* **22**, 913–924 (2003).
62. Deppenmeier, U. *et al.* The genome of *Methanosarcina mazei*: evidence for lateral gene transfer between bacteria and archaea. *J. Mol. Microbiol. Biotechnol.* **4**, 453–461 (2002).
63. Tian, B. & Mathews, M. B. Functional characterization of and cooperation between the double-stranded RNA-binding motifs of the protein kinase PKR. *J. Biol. Chem.* **276**, 9936–9944 (2001).
64. Patel, R. C. & Sen, G. C. PACT, a protein activator of the interferon-induced protein kinase, PKR. *EMBO J.* **17**, 4379–4390 (1998).
Describes the dsRBM-containing protein PACT, which binds and activates another dsRBM-containing protein, PKR.
65. Peters, G. A., Hartmann, R., Qin, J. & Sen, G. C. Modular structure of PACT: distinct domains for binding and activating PKR. *Mol. Cell. Biol.* **21**, 1908–1920 (2001).
66. Micklem, D. R., Adams, J., Grunert, S. & St Johnston, D. Distinct roles of two conserved Staufer domains in oskar mRNA localization and translation. *EMBO J.* **19**, 1366–1377 (2000).
Exploration of the biological and biochemical roles of the Staufer dsRBMs.
67. Zamore, P. D. Thirty-three years later, a glimpse at the ribonuclease III active site. *Mol. Cell* **8**, 1158–1160 (2001).
68. Bernstein, E., Caudy, A. A., Hammond, S. M. & Hannon, G. J. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409**, 363–366 (2001).
Shows the role of Dicer in RNAi.
69. Hutvagner, G. *et al.* A cellular function for the RNA-interference enzyme Dicer in the maturation of the *let-7* small ribosomal RNA. *Science* **293**, 834–838 (2001).
70. Filippov, V., Solovjev, V., Filippova, M. & Gill, S. S. A novel type of RNase III family proteins in eukaryotes. *Gene* **245**, 213–221 (2000).
71. Lee, Y. *et al.* The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**, 415–419 (2003).
72. Liu, Q. *et al.* R2D2, a bridge between the initiation and effector steps of the *Drosophila* RNAi pathway. *Science* **301**, 1921–1925 (2003).
73. Tabara, H., Yigit, E., Siomi, H. & Mello, C. C. The dsRNA binding protein RDE-4 interacts with RDE-1, DCR-1, and a DEXH-box helicase to direct RNAi in *C. elegans*. *Cell* **109**, 861–871 (2002).
References 72 and 73 show that the dsRBM-containing proteins R2D2 and RDE-4 function in RNAi.
74. Samuel, C. E. Antiviral actions of interferons. *Clin. Microbiol. Rev.* **14**, 778–809 (2001).
75. Keegan, L. P., Leroy, A., Sproul, D. & O'Connell, M. A. Adenosine deaminases acting on RNA (ADARs): RNA-editing enzymes. *Genome Biol.* **5**, 209 (2004).
76. Kaufman, R. J. in *Translational Control of Gene Expression* (ed. Mathews, M. B.) 503–528 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 2000).
77. Williams, B. R. Signal integration via PKR. *Sci. STKE* **2001**, RE2 (2001).
78. Levanon, E. Y. *et al.* Systematic identification of abundant A-to-I editing sites in the human transcriptome. *Nature Biotechnol.* **22**, 1001–1005 (2004).
79. Rubin, C. M., Kimura, R. H. & Schmid, C. W. Selective stimulation of translational expression by Alu RNA. *Nucleic Acids Res.* **30**, 3253–3261 (2002).

80. Li, H., Li, W. X. & Ding, S. W. Induction and suppression of RNA silencing by an animal virus. *Science* **296**, 1319–1321 (2002).
81. Adelman, Z. N. *et al.* RNA silencing of dengue virus type 2 replication in transformed C6/36 mosquito cells transcribing an inverted-repeat RNA derived from the virus genome. *J. Virol.* **76**, 12925–12933 (2002).
82. Semizarov, D. *et al.* Specificity of short interfering RNA determined through gene expression signatures. *Proc. Natl Acad. Sci. USA* **100**, 6347–6352 (2003).
83. Chi, J. T. *et al.* Genomewide view of gene silencing by small interfering RNAs. *Proc. Natl Acad. Sci. USA* **100**, 6343–6346 (2003).
84. Bridge, A. J., Pebernard, S., Ducraux, A., Nicoulaz, A. L. & Iggo, R. Induction of an interferon response by RNAi vectors in mammalian cells. *Nature Genet.* **34**, 263–264 (2003).
85. Sledz, C. A., Holko, M., de Veer, M. J., Silverman, R. H. & Williams, B. R. Activation of the interferon system by short-interfering RNAs. *Nature Cell Biol.* **5**, 834–839 (2003).
86. Persengiev, S. P., Zhu, X. & Green, M. R. Nonspecific, concentration-dependent stimulation and repression of mammalian gene expression by small interfering RNAs (siRNAs). *RNA* **10**, 12–18 (2004).
87. Davies, M. V., Chang, H. W., Jacobs, B. L. & Kaufman, R. J. The E3L and K3L vaccinia virus gene products stimulate translation through inhibition of the double-stranded RNA-dependent protein kinase by different mechanisms. *J. Virol.* **67**, 1688–1692 (1993).
88. Kim, Y. G., Lowenhaupt, K., Oh, D. B., Kim, K. K. & Rich, A. Evidence that vaccinia virulence factor E3L binds to Z-DNA *in vivo*: Implications for development of a therapy for poxvirus infection. *Proc. Natl Acad. Sci. USA* **101**, 1514–1518 (2004).
89. Tonkin, L. A. & Bass, B. L. Mutations in RNAi rescue aberrant chemotaxis of ADAR mutants. *Science* **302**, 1725 (2003).
90. Scadden, A. D. & Smith, C. W. RNAi is antagonized by A→I hyper-editing. *EMBO Rep.* **2**, 1107–1111 (2001).
91. Reichman, T. W., Muniz, L. C. & Mathews, M. B. The RNA binding protein nuclear factor 90 functions as both a positive and negative regulator of gene expression in mammalian cells. *Mol. Cell Biol.* **22**, 343–356 (2002).
92. Reichman, T. W. & Mathews, M. B. in *Handbook of Cell Signaling* Vol. 3 (eds Bradshaw, R. A. & Dennis, E. A.) 335–342 (Academic Press, San Diego, USA, 2003).
93. Sun, C. T. *et al.* Transcription repression of human hepatitis B virus genes by negative regulatory element-binding protein/SO_N. *J. Biol. Chem.* **276**, 24059–24067 (2001).
94. Nourbakhsh, M. & Hauser, H. Constitutive silencing of *IFN-β* promoter is mediated by NRF (NF-κB-repressing factor), a nuclear inhibitor of NF-κB. *EMBO J.* **18**, 6415–6425 (1999).
95. Zhou, K. *et al.* RNA helicase A interacts with dsDNA and topoisomerase IIα. *Nucleic Acids Res.* **31**, 2253–2260 (2003).
96. St Johnston, D., Beuchle, D. & Nusslein-Volhard, C. Staufeu, a gene required to localize maternal RNAs in the *Drosophila* egg. *Cell* **66**, 51–63 (1991).
97. Dubnau, J. *et al.* The staufeu/pumilio pathway is involved in *Drosophila* long-term memory. *Curr. Biol.* **13**, 286–296 (2003).
98. Mallardo, M. *et al.* Isolation and characterization of Staufeu-containing ribonucleoprotein particles from rat brain. *Proc. Natl Acad. Sci. USA* **100**, 2100–2105 (2003).
99. Desterro, J. M. *et al.* Dynamic association of RNA-editing enzymes with the nucleolus. *J. Cell Sci.* **116**, 1805–1818 (2003).
100. Sansam, C. L., Wells, K. S. & Emeson, R. B. Modulation of RNA editing by functional nucleolar sequestration of ADAR2. *Proc. Natl Acad. Sci. USA* **100**, 14018–14023 (2003).
101. Kostura, M. & Mathews, M. B. Purification and activation of the double-stranded RNA-dependent eIF-2 kinase DAI. *Mol. Cell Biol.* **9**, 1576–1586 (1989).
102. Liao, H. J., Kobayashi, R. & Mathews, M. B. Activities of adenovirus virus-associated RNAs: purification and characterization of RNA binding proteins. *Proc. Natl Acad. Sci. USA* **95**, 8514–8519 (1998).
103. Eckmann, C. R. & Jantsch, M. F. Xlrpba, a double-stranded RNA-binding protein associated with ribosomes and heterogeneous nuclear RNPs. *J. Cell Biol.* **138**, 239–253 (1997).
104. Eckmann, C. R., Neunteufl, A., Pfaffstetter, L. & Jantsch, M. F. The human but not the *Xenopus* RNA-editing enzyme ADAR1 has an atypical nuclear localization signal and displays the characteristics of a shuttling protein. *Mol. Biol. Cell* **12**, 1911–1924 (2001).
105. Shim, J., Lim, H., Yates, J. R. & Karin, M. Nuclear export of NF90 is required for interleukin-2 mRNA stabilization. *Mol. Cell* **10**, 1331–1344 (2002).
106. Poulsen, H., Nilsson, J., Damgaard, C. K., Egebjerg, J. & Kjems, J. CRM1 mediates the export of ADAR1 through a nuclear export signal within the Z-DNA binding domain. *Mol. Cell Biol.* **21**, 7862–7771 (2001).
107. Nie, Y., Zhao, O., Su, Y. & Yang, J. H. Subcellular distribution of ADAR1 isoforms is synergistically determined by three nuclear discrimination signals and a regulatory motif. *J. Biol. Chem.* **279**, 13249–13255 (2004).
108. Brownawell, A. M. & Macara, I. G. Exportin-5, a novel karyopherin, mediates nuclear export of double-stranded RNA binding proteins. *J. Cell Biol.* **156**, 53–64 (2002). **Shows that the nuclear export protein exportin 5 interacts with dsRBMs from several proteins.**
109. Yi, R., Qin, Y., Macara, I. G. & Cullen, B. R. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev.* **17**, 3011–3016 (2003).
110. Lund, E., Guttinger, S., Calado, A., Dahlberg, J. E. & Kutay, U. Nuclear export of microRNA precursors. *Science* **303**, 95–98 (2004).
111. Bohnsack, M. T., Czaplinski, K. & Gorlich, D. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA* **10**, 185–191 (2004).
112. Bohnsack, M. T. *et al.* Exp5 exports eEF1A via tRNA from nuclei and synergizes with other transport pathways to confine translation to the cytoplasm. *EMBO J.* **21**, 6205–6215 (2002).
113. Macchi, P. *et al.* The brain-specific double-stranded RNA-binding protein Staufeu2: nucleolar accumulation and isoform-specific exportin-5-dependent export. *J. Biol. Chem.* **279**, 31440–31444 (2004).
114. Hitti, E., Neunteufl, A. & Jantsch, M. F. The double-stranded RNA-binding protein X1rbpa promotes RNA strand annealing. *Nucleic Acids Res.* **26**, 4382–4388 (1998).
115. Jammi, N. V. & Beal, P. A. Phosphorylation of the RNA-dependent protein kinase regulates its RNA-binding activity. *Nucleic Acids Res.* **29**, 3020–3029 (2001).
116. Saenger, W. *Principles of Nucleic Acid Structure* (ed. Cantor, C. R.) (Springer, New York, 1984).
117. Weeks, K. M. & Crothers, D. M. Major groove accessibility of RNA. *Science* **261**, 1574–1577 (1993).
118. Seeman, N. C., Rosenberg, J. M. & Rich, A. Sequence-specific recognition of double helical nucleic acids by proteins. *Proc. Natl Acad. Sci. USA* **73**, 804–808 (1976).
119. Kielkopf, C. L. *et al.* A structural basis for recognition of A-T and T-A base pairs in the minor groove of B-DNA. *Science* **282**, 111–115 (1998).
120. Kool, E. T. Hydrogen bonding, base stacking, and steric effects in DNA replication. *Annu. Rev. Biophys. Biomol. Struct.* **30**, 1–22 (2001).
121. Novina, C. D. & Sharp, P. A. The RNAi revolution. *Nature* **430**, 161–164 (2004).
122. Fire, A. *et al.* Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806–811 (1998).
123. Ullu, E., Tschudi, C. & Chakraborty, T. RNA interference in protozoan parasites. *Cell Microbiol.* **6**, 509–519 (2004).
124. Bartel, D. P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281–297 (2004).
125. Turner, D. H. Thermodynamics of base pairing. *Curr. Opin. Struct. Biol.* **6**, 299–304 (1996).
126. Brion, P. & Westhof, E. Hierarchy and dynamics of RNA folding. *Annu. Rev. Biophys. Biomol. Struct.* **26**, 113–137 (1997).
127. Tinoco, I. Jr & Bustamante, C. How RNA folds. *J. Mol. Biol.* **293**, 271–281 (1999).
128. Schuster-Bockler, B., Schultz, J. & Rahmann, S. HMM Logos for visualization of protein families. *BMC Bioinformatics* **5**, 7 (2004).

Acknowledgements

We thank B. Golden at Purdue University for assistance with figure 2 and H. Zhang at New Jersey Medical School for assistance with figure 4. Support from the following funding agencies is acknowledged: from the National Institutes of Health to P.C.B. and M.B.M. and from the American Cancer Society to A.D.P.

Competing interests statement

The authors declare that they have no competing financial interests.

Online links

DATABASES

The following terms in this article are linked online to:
Flybase: <http://flybase.bio.indiana.edu/>
Drosna | R2D2 | Staufeu
Interpro: <http://www.ebi.ac.uk/interpro/dsRBD/RNB>
SwissProt: <http://www.ca.expasy.ch>
ADAR1 | ADAR2 | Dicer | MRPL44 | NF90 | NRF | PACT | PKR | TRBP | Xlrpba

FURTHER INFORMATION

Michael Mathews' laboratory: <http://www.umdj.edu/biochweb/Mathews/htm>
Bin Tian's laboratory: <http://exon.umdj.edu>
 Access to this links box is available online.