EMBO reports

review

RNA sequence- and shape-dependent recognition by proteins in the ribonucleoprotein particle

Richard Stefl, Lenka Skrisovska & Frédéric H.-T. Allain[+]
Swiss Federal Institute of Technology Zürich, Zürich, Switzerland

At all stages of its life (from transcription to translation), an RNA transcript interacts with many different RNA-binding proteins. The composition of this supramolecular assembly, known as a ribonucleoprotein particle, is diverse and highly dynamic. RNA-binding proteins control the generation, maturation and lifespan of the RNA transcript and thus regulate and influence the cellular function of the encoded gene. Here, we review our current understanding of protein–RNA recognition mediated by the two most abundant RNA-binding domains (the RNA-recognition motif and the double-stranded RNA-binding motif) plus the zinc-finger motif, the most abundant nucleic-acid-binding domain. In addition, we discuss how not only the sequence but also the shape of the RNA are recognized by these three classes of RNA-binding protein.

Keywords: double-stranded RNA-binding motif; RNA-binding proteins; RNA recognition; RNA-recognition motif; zinc-finger motif
EMBO reports (2005) **6,** 33–38. doi:10.1038/sj.embor.7400325

Introduction

The association of RNA-binding proteins (RBPs) with RNA transcripts begins during transcription. Some of these early-binding RBPs remain bound to the RNA until it is degraded, whereas others recognize and transiently bind to RNA at later stages for specific processes such as splicing, processing, transport and localization (Dreyfuss *et al*, 2002). The RBPs cover the RNA transcripts and control their fate. Some RBPs function as RNA chaperones (Lorsch, 2002) by helping the RNA, which is initially single-stranded, to form various secondary or tertiary structures. When folded, these structured RNAs, together with specific RNA sequences, act as a signal for other RBPs that mediate gene regulation. Here, we review our current structural understanding of protein–RNA recognition mediated by the two most abundant RNA-binding domains, the RNA-recognition motif (RRM) and the double-stranded RNA-binding motif (dsRBM), and by the most abundant nucleic-acid-binding motif, the CCHH-type zinc-finger domain. We discuss how these three small domains recognize RNA: some bind single-stranded RNA by direct readout of the primary sequence, whereas others recognize primarily the shape of the RNA or both the sequence and the shape. Other types of RNA-binding domains, such as the K-homology (KH) domain or the oligonucleotide/oligosaccharide-binding (OB) fold, have recently been reviewed and are not discussed here (Messias & Sattler, 2004).

Institute for Molecular Biology and Biophysics, Swiss Federal Institute of Technology Zürich, ETH-Hönggerberg, CH-8093 Zürich, Switzerland
[+]Corresponding author. Tel: +41 (0)1 63 33940; Fax: +41 (0)1 63 31294;
E-mail: allain@mol.biol.ethz.ch

Submitted 22 September 2004; accepted 26 November 2004

RNA shape-dependent recognition by double-stranded RBM

The dsRBM is a 70–75 amino-acid domain with a conserved αββββα protein topology in which the two α-helices are packed along one face of a three-stranded anti-parallel β-sheet (Fig 1; Fierro-Monti & Mathews, 2000; St Johnston *et al*, 1992). These domains occur mostly in multiple copies (up to five) and have so far been found in 388 eukaryotic proteins, 72 of which are human (data taken from the SMART database; Letunic *et al*, 2004). These proteins have an essential role in RNA interference, RNA processing, RNA localization, RNA editing and translational repression (Doyle & Jantsch, 2002; Saunders & Barber, 2003).

So far, only three structures of dsRBMs in complex with dsRNA have been determined (Table 1): a 1.9 Å crystal structure of the second dsRBM of *Xenopus laevis* RNA-binding protein A (Xlrbpa2) bound to two coaxially stacked dsRNA molecules, each 10 bp long (Ryter & Schultz, 1998); a nuclear magnetic resonance (NMR) structure of the third dsRBM from the *Drosophila* Staufen protein in complex with a symmetrical GC-rich 12-bp duplex capped by a UUCG tetraloop (Ramos *et al*, 2000); and an NMR structure of the dsRBM of Rnt1p (an RNase III homologue from budding yeast) bound to a 14-bp RNA duplex capped by an AGAA tetraloop (Wu *et al*, 2004). All three structures have several common features that reveal how a dsRBM is able to bind to any dsRNA but not to dsDNA, regardless of its base composition. The dsRBMs interact along one face of the RNA duplex through both α-helices and their β1–β2 loop (Fig 1). The contacts with the RNA cover 15 bp that span two consecutive minor grooves separated by a major groove. In all three structures, the contacts to the sugar-phosphate backbone of the major groove and of one minor groove (Fig 1) are mediated by the β1–β2 loop and the amino-terminal part of α-helix 2. These interactions are non-sequence-specific as they involve 2′-hydroxyls and phosphate oxygens and are perfectly adapted to the shape of an RNA double helix. By contrast, the interactions mediated by α-helix 1 are different in all three complexes. In the dsRBM of Xlrbpa2, α-helix 1 interacts non-specifically with the other minor groove of the RNA (Fig 1A), with a few contacts to the bases. In the dsRBM of Staufen, α-helix 1

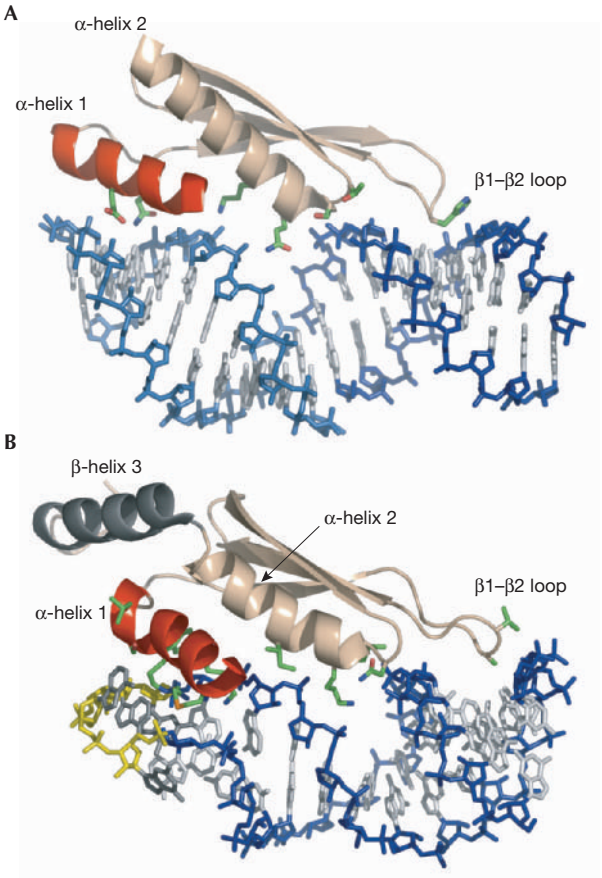


Fig 1 | Double-stranded RNA recognition by double-stranded RNA-binding motifs. (**A**) The double-stranded RNA-binding motif (dsRBM) of Xlrbpa2 bound to dsRNA (Ryter & Schultz, 1998). The α-helix 1 (in red), amino-terminal part of α-helix 2, and β1–β2 loop recognize non-sequence-specifically the shape of dsRNA. Backbones are coloured blue and light blue for two co-axially stacked duplexes. (**B**) The dsRBM of Rnt1p bound to the stem–loop closed by the AGNN tetraloop (backbone in yellow and bases in black; Wu *et al*, 2004). The α-helix 1, a key element for shape-specific recognition by dsRBMs, is highlighted in red. The dsRBM of Rnt1p has an additional carboxy-terminal α-helix 3 (in black), that modulates the conformation of α-helix 1 (Leulliot *et al*, 2004). Side chains involved in intermolecular interactions are shown.

interacts with a UUCG tetraloop that caps the RNA double helix. Although the UUCG tetraloop is not a natural substrate of Staufen, this finding led to the proposal that α-helix 1 modulates the specificity of individual dsRBMs (Ramos *et al*, 2000). Indeed, this was recently confirmed by the structure of the dsRBM of Rnt1p bound to its natural RNA substrate (Fig 1B), in which α-helix 1 recognizes the specific shape of the minor groove created by the conserved AGNN tetraloop (Wu *et al*, 2004). The α-helix 1, the conformation of which is stabilized by an additional carboxy-terminal α-helix 3 (Fig 1B; Leulliot *et al*, 2004), is tightly inserted into the RNA minor groove and contacts the sugar-phosphate backbone and the two non-conserved tetraloop bases, whereas the conserved A and G bases are not involved in the interactions (Wu *et al*, 2004). This structure illustrates how this dsRBM recognizes the specific shape of its RNA target but not its sequence. dsRBMs are highly conserved and have the same structural framework, but are chemically distinct through variations in key residues. The structure of the dsRBM of Rnt1p in complex with RNA highlights the essential role of the α-helix 1 in the recognition of structured elements that deviate from regular dsRNA. The α-helix 1 is the least-conserved secondary structure element among various dsRBMs and seems to have a different spatial arrangement relative to the rest of the domain in different dsRBMs. This variability may be an important factor as many biochemical experiments have shown that dsRBM-containing proteins have binding specificity for a variety of RNA structures, such as stem–loops, internal loops, bulges or helices with mismatches (Doyle & Jantsch, 2002; Fierro-Monti & Mathews, 2000; Ohman *et al*, 2000; Stephens *et al*, 2004). Clearly, further structures are needed to decipher the extent of RNA shape-dependent recognition by dsRBMs.

RNA sequence- and shape-dependent recognition by an RRM

The RRM is the most common RNA-binding motif. It is a small protein domain of 75–85 amino acids with a typical βαββαβ topology that forms a four-stranded β-sheet packed against two α-helices (Mattaj, 1993). RRMs are found in about 0.5%–1% of human genes (Venter *et al*, 2001) and are often present in multiple copies (up to six per protein). RRM-domain-containing proteins are involved in many cellular functions, particularly messenger RNA and ribosomal RNA processing, splicing and translation regulation, RNA export and RNA stability (Dreyfuss *et al*, 2002).

So far, ten structures of an RRM in complex with RNA have been determined using either NMR spectroscopy or X-ray crystallography (Table 1). These structures reveal the complexity of protein–RNA recognition mediated by the RRM, which often involves not only protein–RNA interactions but also RNA–RNA and protein–protein interactions. All ten structures reveal some common features. The main protein surface of the RRM involved in the interaction with the RNA is the four-stranded β-sheet, which usually contacts two or three nucleotides (exemplified here by the RRM1 of sex-lethal; Fig 2A; Handa *et al*, 1999). The nucleotides are located on the surface of the β-sheet, with the bases oriented parallel to the β-sheet plane and often packed against conserved hydrophobic side-chains (usually aromatics). These two or three nucleotides are recognized sequence-specifically by interactions with the protein side-chains of the β-sheet and with the main-chain and side-chains of the residues carboxy-terminal to the β-sheet. Interestingly, it seems that almost all possible sequences (doublets or triplets) can be accommodated on such a surface as the RNA sequences are different in each structure (Table 1).

Often, RRM-containing proteins bind more than three nucleotides and recognize longer single-stranded RNA (for example, poly(A)-binding protein (PAPB; Deo *et al*, 1999), sex-lethal (Handa *et al*, 1999), Hu protein D (HuD; Wang & Hall, 2001), heterogeneous nuclear RNP A1 (hnRNP A1; Ding *et al*, 1999), nucleolin (Allain *et al*, 2000; Johansson *et al*, 2004), RNA stem–loops U1A (Oubridge *et al*, 1994), U2B″ (Price *et al*, 1998), nucleolin (Allain *et al*, 2000)) or even internal loops (U1A; Allain *et al*, 1997; Varani *et al*, 2000), all with high affinity ($K_d \approx 10^{-9}M^{-1}$). In U1A, U2B″, nucleolin and sex-lethal, two loops between the secondary-structure elements of the RRM (the β2–β3 loop and the β1–α1 loop) are essential for additional contacts with the RNA (Fig 2B). These loops vary significantly in size and amino-acid sequence between the different RRMs. In the RRM of CBP20, the C- and N-terminal extensions (which are stabilized by the cognate

Table 1 | Various structures of RNA-binding proteins bound to RNA

Complex	RNA secondary structure	No. of RBDs in structures vs in full-length protein	RNA sequence recognized specifically by RRM β-sheet	Function	Reference; Protein Data Bank (PDB) ID
dsRBM type					
Second dsRBM of Xlrbpa2	Duplex[a]	1/3	–	hnRNP association, translation repression	Ryter & Schultz, 1998; 1DI2
Third dsRBM of Staufen	Stem–loop[b]	1/5	–	mRNA localization, translation control	Ramos *et al*, 2000; 1EKZ
dsRBM of Rnt1p	Stem–loop	1/1	–	RNA processing	Wu *et al*, 2004; 1T4L
RRM type					
N-terminal RRM of U2B" (in U2B"–U2A'–RNA complex)	Stem–loop	1/2	CAG	Pre-mRNA splicing	Price *et al*, 1998; 1A9N
N-terminal RRM of U1A	Stem–loop	1/2	CAC	Pre-mRNA splicing	Oubridge *et al*, 1994; 1URN
N-terminal RRM of U1A	Internal loop	1/2	CAC	Pre-mRNA splicing	Allain *et al*, 1997; 1AUD, Varani *et al*, 2000; 1DZ5
Two N-terminal RRMs of nucleolin	Stem–loop	2/4	RRM1-CG RRM2-UC	Ribosome biogenesis	Allain *et al*, 2000; 1FJE, Johansson *et al*, 2004; 1RKJ
Two N-terminal RRM of PABP	Single strand	2/4	RRM1-AAA RRM2-AAA	Translation initiation	Deo *et al*, 1999; 1CVJ
Two RRMs of sex-lethal	Single strand	2/2	RRM1-UUU RRM2-UGU	Alternative splicing	Handa *et al*, 1999; 1H2T
Two N-terminal RRMs of HuD	Single strand	2/3	RRM1-UUU RRM2-UU	mRNA stability, translation regulation	Wang & Hall, 2001; 1FXL, 1G2E
RRM of CBP20	Single strand	1/1	m7GpppG	Maturation of pre-mRNA and U-rich snRNA	Mazza *et al*, 2002; 1H2V
Zinc finger					
Fourth, fifth and sixth zinc fingers (CCHH-type) of TFIIIA	Truncated 5S RNA[c]	3/9	–	Transcription regulation	Lu *et al*, 2003; 1UN6
First and second zinc fingers (CCCH-type) of TIS11d	Single strand	2/2	–	RNA processing and degradation	Hudson *et al*, 2004; 1RGO

[a]Two coaxially stacked dsRNA (each 10-bp long). [b]12-bp duplex capped by a non-physiologically relevant UUCG tetraloop. [c]Consists of loop A, loop E and helices I, IV and V.
CBP20, cap-binding protein 20; dsRBM, double-stranded RNA-binding motif; hnRNP, heterogeneous nuclear ribonucleoprotein; HuD, Hu protein D; PABP, poly(A)-binding protein; RBD, RNA-binding domain; Rnt1p, RNase III homologue; RRM, RNA-recognition motif; TFIIIA, transcription factor IIIA; Xlrbpa2, *Xenopus laevis* RNA-binding protein A.

protein CBP80) provide a tight binding pocket for the 5′ capped RNAs (7-methyl-G(5′)ppp(5′)N, where N is any nucleotide; Fig 2C; Mazza *et al*, 2002). In proteins that contain several RRMs, high-affinity binding can only be achieved by the cooperative binding of at least two RRMs to the RNA (for example, in nucleolin (Fig 2D), PABP and sex-lethal). In addition to the β-sheet–RNA contacts, interactions between the inter-domain linker and the RNA and between the RRMs themselves contribute to the marked increase in affinity compared with the binding of the individual domain alone. These structures show that the RRM is a platform with a large capacity for variation in order to achieve high RNA-binding affinity and specificity. For example, it is remarkable that a single domain like nucleolin RRM2 contacts only two nucleotides, whereas U1A RRM1 contacts 12 nucleotides and the RRM of Y14 (Fribourg *et al*, 2003) does not contact RNA but rather another protein. This fascinating plasticity of the RRM explains why it is so abundant and why it is involved in so many different biological functions; however, this plasticity makes it difficult to predict how the RRM achieves RNA recognition.

RNA recognition by zinc fingers

CCHH-type zinc-finger domains are the most common DNA-binding domain found in eukaryotic genomes. Typically, several fingers are used in a modular fashion to achieve high sequence-specific recognition of DNA (Miller *et al*, 1985). Each finger displays a ββα protein fold in which a β-hairpin and an α-helix are pinned together by a Zn^{2+} ion. DNA-sequence-specific recognition is achieved by the interactions between protein side-chains of the α-helix (at position –1, 2, 3 and 6, for the canonical arrangement) and the DNA bases in the major groove (Fig 3A; Wolfe *et al*, 2000). However, there is increasing evidence that zinc fingers are also used to recognize RNA (Finerty & Bass, 1997; Mendez-Vidal *et al*, 2002; Picard & Wegnez, 1979; Theunissen *et al*, 1992). The crystal structure of three zinc fingers (fingers 4–6) of transcription factor IIIA (TFIIIA) in complex with a 61-nucleotide fragment of the 5S RNA (Lu *et al*, 2003) provided the first insight into RNA recognition by CCHH-type zinc fingers. In this structure, finger 4 binds to loop E, finger 5 to helix V, and finger 6 to loop A (Fig 3B). Finger 4 recognizes loop E by specifically interacting with a bulged guanosine (Fig 3C) and, similarly, finger 6 recognizes loop A by specifically

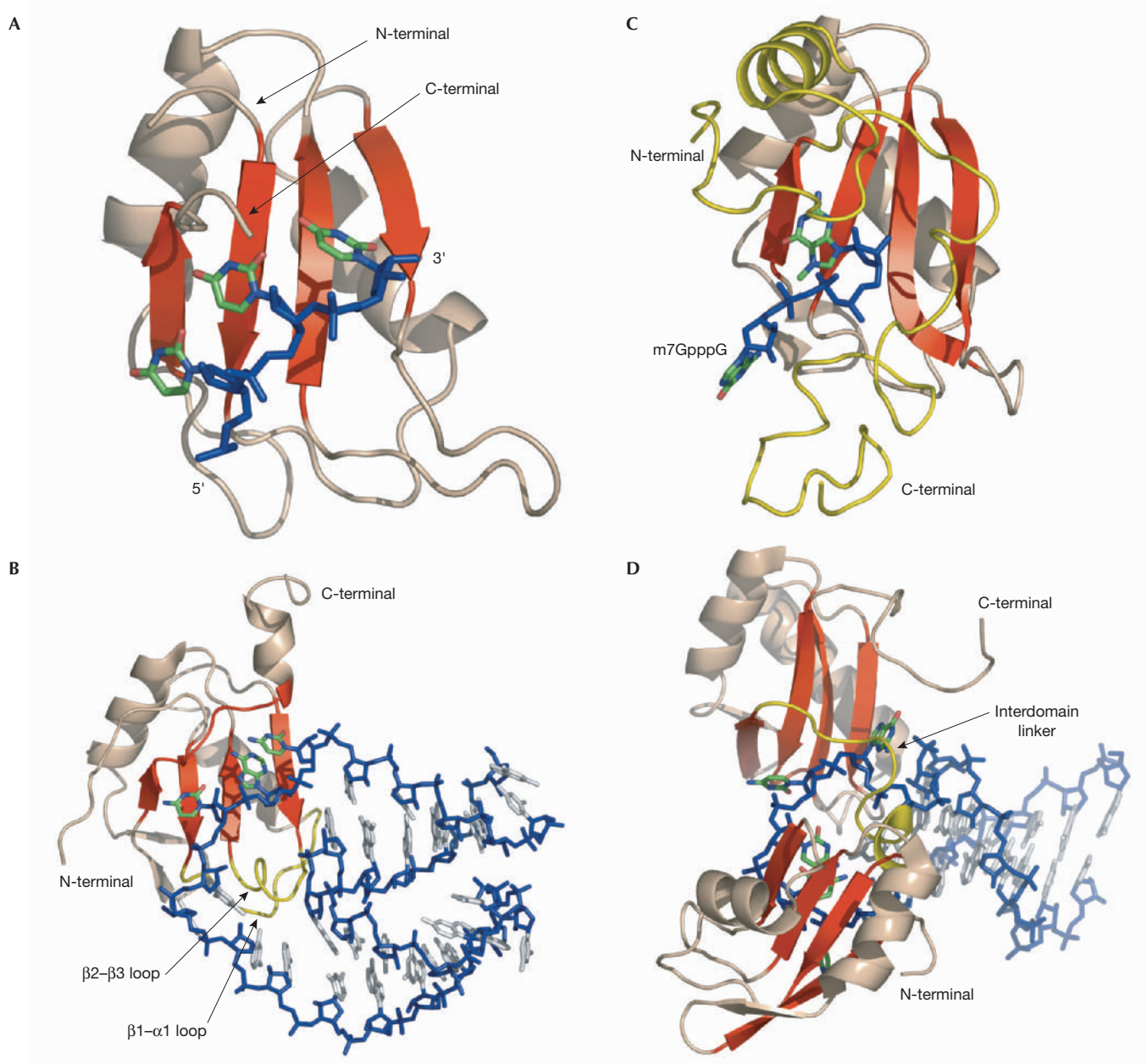


Fig 2 | RNA recognition by RNA-recognition motifs. The similarities and differences are highlighted in red and yellow, respectively. (**A**) RNA-recognition motif 1 (RRM1) of sex-lethal (shown as a ribbon model) interacts with the triplet UUU (shown as a stick model; Handa *et al*, 1999). The four-stranded β-sheet (in red) recognizes three nucleotides—this is the canonical mode of RRM–RNA interaction. (**B**) RRM1 of U1A bound to an RNA internal loop (Allain *et al*, 1997). The four-stranded β-sheet (in red) recognizes three nucleotides. The β2–β3 loop and the β1–α1 loop (in yellow) contact additional RNA residues. (**C**) RRM of CBP20 complexed with m^7GpppG (7-methyl-G(5′)ppp(5′)G; Mazza *et al*, 2002). C- and N-terminal extensions (in yellow) provide additional protein–RNA contacts that creates a specific binding pocket for 5′ capped RNAs. (**D**) RRMs 1 and 2 of nucleolin bound to an RNA stem–loop (Allain *et al*, 2000). The four-stranded β-sheet of RRM2 (in red) recognizes only two nucleotides. The interdomain linker (in yellow) participates in the recognition of the hairpin architecture.

interacting with two bases (an adenine and a cytosine) that also bulge out from the rest of the RNA (Fig 3B). The specific recognition of the RNA by both fingers 4 and 6 is achieved by side-chain contacts from the N-terminal parts of the α-helix (at position –1, 1 and 2; Fig 3C). The interaction of finger 5 with helix V differs from the ones made by fingers 4 and 6. In this case, finger 5 recognizes a short RNA double helix by multiple contacts between basic amino acids of the α-helix and the RNA sugar-phosphate backbone (Fig 3D).

In contrast to the above-mentioned CCHH zinc fingers, another class of zinc fingers (CCCH-type) was recently found to adopt a different fold and to recognize sequence-specifically single-stranded RNA (Hudson *et al*, 2004). In this NMR structure, sequence-specific

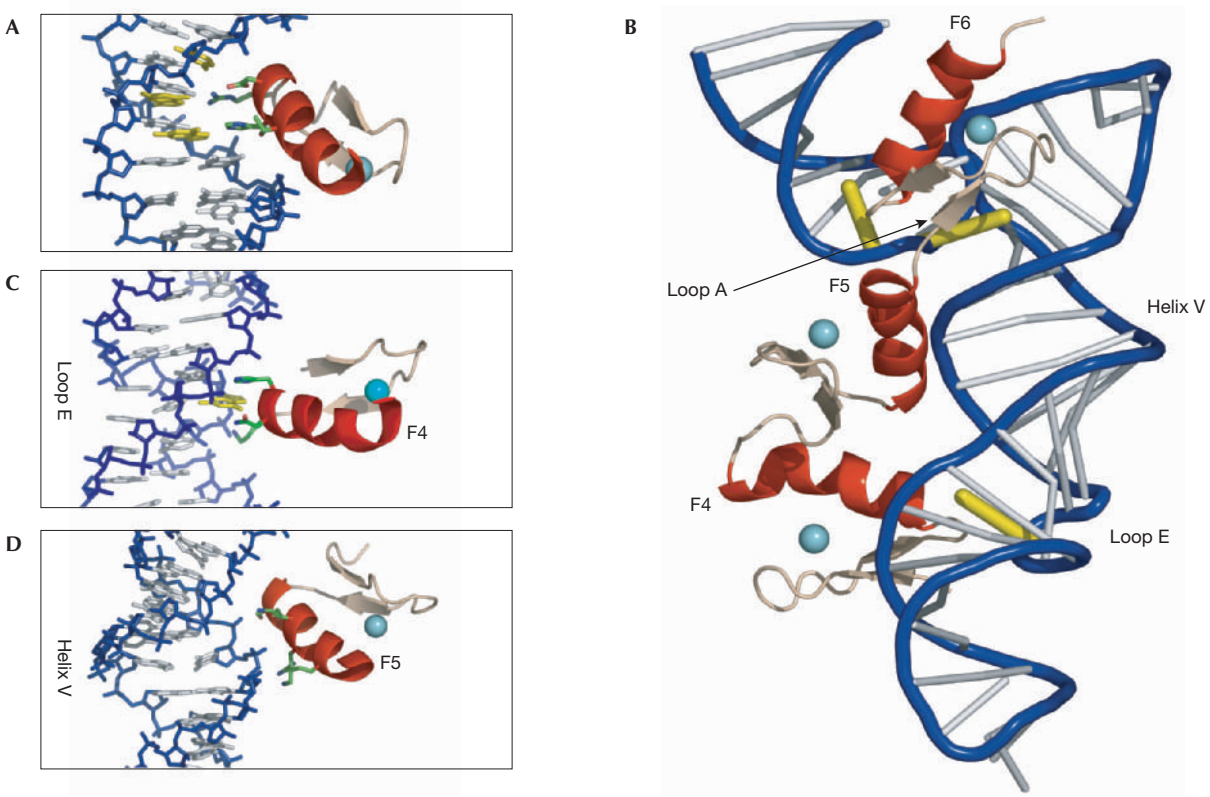


Fig 3 | DNA vs RNA recognition by CCHH-type zinc fingers. (**A**) Zinc finger 2 of Zif268 bound to double-stranded DNA (Pavletich & Pabo, 1991). The α-helix of the zinc finger (in red) inserts into the DNA major groove; base contacts are made from positions –1, 2, 3 and 6 of the α-helix (the protein side-chains are shown as an element-type coloured stick model). The DNA bases that are recognized by the finger are coloured yellow. (**B**) Overall view of the complex of transcription factor IIIA (TFIIIA) fingers 4–6 (F4–F6) and 61-nucleotide 5S RNA (Lu *et al*, 2003). The protein and RNA are represented as ribbon models. The bulged bases involved in the recognitions are highlighted in yellow. Cyan balls represent zinc ions. (**C**) TFIIIA finger 4 (F4) bound to loop E (Lu *et al*, 2003). The α-helix (in red) of the finger 4 specifically interacts with a guanosine base that bulges out (in yellow); the base contacts are made from the side-chain at position –1, 1 and 2 of the α-helix. (**D**) TFIIIA finger 5 (F5) bound to helix V (Lu *et al*, 2003). The α-helix (in red) of finger 5 recognizes the dsRNA shape by non-sequence-specific contacts to the RNA sugar-phosphate backbone.

RNA recognition is achieved by a network of intermolecular hydrogen bonds between the protein main-chain functional groups and the Watson–Crick edges of the bases (Hudson *et al*, 2004). These structures reveal that zinc fingers bind to RNA differently to the way they do to DNA. The CCHH-type zinc fingers have two modes of RNA binding. First, the zinc fingers interact non-specifically with the backbone of a double helix, and second, the zinc fingers specifically recognize individual bases that bulge out of a structurally rigid element. The CCCH-type zinc fingers show a third mode of RNA binding, in which the single-stranded RNA is recognized in a sequence-specific manner. Taken together, zinc fingers represent a unique class of nucleic-acid-binding proteins that are capable of a direct readout of the DNA sequence within a DNA double helix, a direct readout of the RNA sequence within single-stranded RNA, and an indirect readout of the RNA as they recognize the shape of the RNA rather than its sequence. Of course, more structures of CCHH-type and CCCH-type zinc fingers in complex with RNA will need to be determined to generalize their mode of RNA recognition.

Conclusions

Proteins that contain RNA-binding domains and their interactions with RNA have important roles in all aspects of gene expression and regulation. The enormous diversity of interactions observed in protein–RNA complexes indicates that a simple recognition code is unlikely to exist in the world of protein–RNA interactions. However, two unifying themes may be inferred from the known complexes: the recognition of the primary RNA sequence and/or the recognition of the RNA shape by individual RBPs. In a simplistic view, the RRMs, dsRBMs and CCHH-type zinc fingers seem to be shaped to recognize single-stranded RNA, double-stranded RNA and RNA bulges, respectively. However, we have shown here by reviewing several recent protein–RNA complex structures that, the RRM and, to a lesser extent, the dsRBMs and the CCHH-type zinc fingers have evolved to recognize specifically a rich repertoire of RNAs in terms of length, sequence and structure. This is achieved in three ways: first, by the subtle amino-acid change in variable regions of the domains, namely the β2–β3 and the β1–α1 loops in the RRM, α-helix 1 in the dsRBM and the α-helix in the zinc fingers; second, by multiplication of the domains to achieve higher affinity through cooperative binding; and third, by extension of the protein domain. Although more structures still need to be determined, it might soon be possible to predict which RBP binds to which RNA, and how it recognizes its target. As a consequence, post-transcriptional gene expression and its regulation could be understood and controlled at the atomic level.

ACKNOWLEDGEMENTS

We apologize to authors whose work could not be cited due to space constraints. The authors are supported by the Swiss National Science Foundation (No. 31-67098.01), the Roche Research Fund for Biology at the ETH Zurich (F.H.-T.A.), and the European Molecular Biology Organization and the Human Frontier Science Program postdoctoral fellowships (R.S.).

REFERENCES

Allain FH, Howe PW, Neuhaus D, Varani G (1997) Structural basis of the RNA-binding specificity of human U1A protein. *EMBO J* **16:** 5764–5772

Allain FH, Bouvet P, Dieckmann T, Feigon J (2000) Molecular basis of sequence-specific recognition of pre-ribosomal RNA by nucleolin. *EMBO J* **19:** 6870–6881

Deo RC, Bonanno JB, Sonenberg N, Burley SK (1999) Recognition of polyadenylate RNA by the poly(A)-binding protein. *Cell* **98:** 835–845

Ding J, Hayashi MK, Zhang Y, Manche L, Krainer AR, Xu RM (1999) Crystal structure of the two-RRM domain of hnRNP A1 (UP1) complexed with single-stranded telomeric DNA. *Genes Dev* **13:** 1102–1115

Doyle M, Jantsch MF (2002) New and old roles of the double-stranded RNA-binding domain. *J Struct Biol* **140:** 147–153

Dreyfuss G, Kim VN, Kataoka N (2002) Messenger-RNA-binding proteins and the messages they carry. *Nat Rev Mol Cell Biol* **3:** 195–205

Fierro-Monti I, Mathews MB (2000) Proteins binding to duplexed RNA: one motif, multiple functions. *Trends Biochem Sci* **25:** 241–246

Finerty PJ, Bass BL (1997) A *Xenopus* zinc finger protein that specifically binds dsRNA and RNA–DNA hybrids. *J Mol Biol* **271:** 195–208

Fribourg S, Gatfield D, Izaurralde E, Conti E (2003) A novel mode of RBD-protein recognition in the Y14–Mago complex. *Nat Struct Biol* **10:** 433–439

Handa N, Nureki O, Kurimoto K, Kim I, Sakamoto H, Shimura Y, Muto Y, Yokoyama S (1999) Structural basis for recognition of the *tra* mRNA precursor by the Sex-lethal protein. *Nature* **398:** 579–585

Hudson BP, Martinez-Yamout MA, Dyson HJ, Wright PE (2004) Recognition of the mRNA AU-rich element by the zinc finger domain of TIS11d. *Nat Struct Mol Biol* **11:** 257–264

Johansson C, Finger LD, Trantirek L, Mueller TD, Kim S, Laird-Offringa IA, Feigon J (2004) Solution structure of the complex formed by the two N-terminal RNA-binding domains of nucleolin and a pre-rRNA target. *J Mol Biol* **337:** 799–816

Letunic I, Copley RR, Schmidt S, Ciccarelli FD, Doerks T, Schultz J, Ponting CP, Bork P (2004) SMART 4.0: towards genomic data integration. *Nucleic Acids Res* **32:** D142–D144

Leulliot N, Quevillon-Cheruel S, Graille M, Van Tilbeurgh H, Leeper TC, Godin KS, Edwards TE, Sigurdsson ST, Rozenkrats N, Nagel RJ, Ares M, Varani G (2004) A new α-helical extension promotes RNA binding by the dsRBD of Rnt1p RNAse III. *EMBO J* **23:** 2468–2477

Lorsch JR (2002) RNA chaperones exist and DEAD box proteins get a life. *Cell* **109:** 797–800

Lu D, Searles MA, Klug A (2003) Crystal structure of a zinc-finger—RNA complex reveals two modes of molecular recognition. *Nature* **426:** 96–100

Mattaj IW (1993) RNA recognition: a family matter? *Cell* **73:** 837–840

Mazza C, Segref A, Mattaj IW, Cusack S (2002) Large-scale induced fit recognition of an m(7)GpppG cap analogue by the human nuclear cap-binding complex. *EMBO J* **21:** 5548–5557

Mendez-Vidal C, Wilhelm MT, Hellborg F, Qian W, Wiman KG (2002) The p53-induced mouse zinc finger protein wig-1 binds double-stranded RNA with high affinity. *Nucleic Acids Res* **30:** 1991–1996

Messias AC, Sattler M (2004) Structural basis of single-stranded RNA recognition. *Acc Chem Res* **37:** 279–287

Miller J, Mclachlan AD, Klug A (1985) Repetitive zinc-binding domains in the protein transcription factor Iiia from *Xenopus* oocytes. *EMBO J* **4:** 1609–1614

Ohman M, Kallman AM, Bass BL (2000) *In vitro* analysis of the binding of ADAR2 to the pre-mRNA encoding the GluR-B R/G site. *RNA* **6:** 687–697

Oubridge C, Ito N, Evans PR, Teo CH, Nagai K (1994) Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature* **372:** 432–438

Pavletich NP, Pabo CO (1991) Zinc finger–DNA recognition: crystal structure of a Zif68-DNA complex at 2.1 Å. *Science* **252:** 809–817

Picard B, Wegnez M (1979) Isolation of a 7s particle from *Xenopus laevis* oocytes—5s RNA–protein complex. *Proc Natl Acad Sci USA* **76:** 241–245

Price SR, Evans PR, Nagai K (1998) Crystal structure of the spliceosomal U2B″–U2A′ protein complex bound to a fragment of U2 small nuclear RNA. *Nature* **394:** 645–650

Ramos A, Grunert S, Adams J, Micklem DR, Proctor MR, Freund S, Bycroft M, St Johnston D, Varani G (2000) RNA recognition by a Staufen double-stranded RNA-binding domain. *EMBO J* **19:** 997–1009

Ryter JM, Schultz SC (1998) Molecular basis of double-stranded RNA–protein interactions: structure of a dsRNA-binding domain complexed with dsRNA. *EMBO J* **17:** 7505–7513

Saunders LR, Barber GN (2003) The dsRNA binding protein family: critical roles, diverse cellular functions. *FASEB J* **17:** 961–983

St Johnston D, Brown NH, Gall JG, Jantsch M (1992) A conserved double-stranded RNA-binding domain. *Proc Natl Acad Sci USA* **89:** 10979–10983

Stephens OM, Haudenschild BL, Beal PA (2004) The binding selectivity of ADAR2's dsRBMs contributes to RNA-editing selectivity. *Chem Biol* **11:** 1239–1250

Theunissen O, Rudt F, Guddat U, Mentzel H, Pieler T (1992) RNA and DNA-binding zinc fingers in *Xenopus* Tfiiia. *Cell* **71:** 679–690

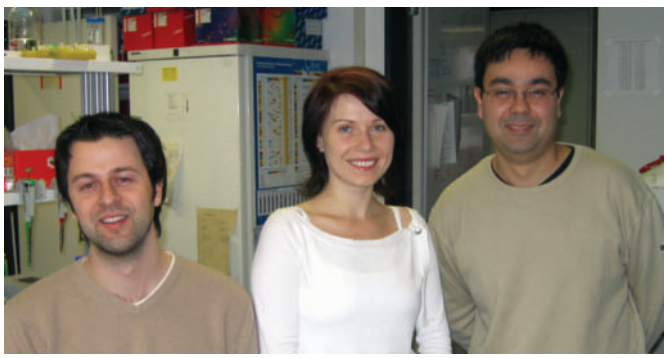
Varani L, Gunderson SI, Mattaj IW, Kay LE, Neuhaus D, Varani G (2000) The NMR structure of the 38 kDa U1A protein–PIE RNA complex reveals the basis of cooperativity in regulation of polyadenylation by human U1A protein. *Nat Struct Biol* **7:** 329–335

Venter JC *et al* (2001) The sequence of the human genome. *Science* **291:** 1304–1351

Wang XQ, Hall TMT (2001) Structural basis for recognition of AU-rich element RNA by the HuD protein. *Nat Struct Biol* **8:** 141–145

Wolfe SA, Nekludova L, Pabo CO (2000) DNA recognition by Cys(2)His(2) zinc finger proteins. *Annu Rev Biophys Biomol Struct* **29:** 183–212

Wu H, Henras A, Chanfreau G, Feigon J (2004) 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



Richard Stefl, Lenka Skrisovska & Frédéric H.-T. Allain, who is an EMBO Young Investigator