Noncanonical G recognition mediates KSRP regulation of let-7 biogenesis

Giuseppe Nicastro¹, María Flor García-Mayoral^{1,2}, David Hollingworth¹, Geoff Kelly³, Stephen R Martin⁴, Paola Briata⁵, Roberto Gherzi⁵ & Andres Ramos¹

Let-7 is an important tumor-suppressive microRNA (miRNA) that acts as an on-off switch for cellular differentiation and regulates the expression of a set of human oncogenes. Binding of the human KSRP protein to let-7 miRNA precursors positively regulates their processing to mature let-7, thereby contributing to control of cell proliferation, apoptosis and differentiation. Here we analyze the molecular basis for KSRP–let-7 precursor selectivity and show how the third KH domain of the protein recognizes a G-rich sequence in the pre-let-7 terminal loop and dominates the interaction. The structure of the KH3–RNA complex explains the protein recognition of this noncanonical KH target sequence, and we demonstrate that the specificity of this binding is crucial for the functional interaction between the protein and the miRNA precursor.

miRNAs regulate the expression of more than half of all human genes by acting on the stability and translation of mRNA targets, and the cellular concentration of miRNAs is tightly controlled both at the transcriptional and post-transcriptional level^{1,2}. The ~22-nucleotide miRNA is transcribed as part of a much larger RNA molecule where the mature miRNA sequence base-pairs with a quasicomplementary sequence to form a hairpin-like structure. The two-step processing of the initial transcript to mature miRNA can be selectively regulated by the interaction between effector proteins and the terminal loops of precursor (pre-miRNA) and primary miRNA (pri-miRNA) hairpins (reviewed in ref. 1). The molecular information available on the processing complexes and the high sequence conservation observed for some but not all miRNA terminal loops support the proposed role of terminal loops as a platform for the binding of effector proteins and the regulation of miRNA biogenesis³.

Originally discovered in *Caenorhabditis elegans*, where it has an important role in development, let-7 was the first miRNA to be identified in humans. Let-7 provides an on-off switch for cell differentiation and has been shown to inhibit the expression of a set of important oncogenes in humans. Let-7 functions as a major tumor suppressor, and its cellular concentration must be tightly regulated⁴. At the post-transcriptional level, this regulation relies on proteins that bind to the terminal loop of the let-7 precursors and either increase (KSRP) or decrease (lin-28 and hnRNPA1) the efficiency of let-7 biogenesis^{5–9}. The key to establishing the selectivity of these post-transcriptional regulatory mechanisms is the recognition of the RNA terminal loop by different protein effectors.

Two recent studies have clarified the molecular basis of the recognition of the let-7 terminal loop by the negative regulator lin-28,

which interacts with the RNA by using a CCHC double zinc-finger domain and a cold-shock RNA-binding domain^{10,11}. However, the recognition between let-7 and the activator KSRP, which has a very different domain composition, is still to be elucidated^{8,9}. KSRP is a multifunctional protein that interacts with its nucleic acid targets by using four sequential KH domains (Fig. 1), and it has been shown to have a role in the decay, splicing and localization of selected mRNAs¹². In miRNA biogenesis, KSRP interacts with an AGGGU sequence at the 5' end of the let-7 terminal loop (Fig. 1c) and stimulates pri- and pre-let-7 processing⁹. How KSRP recognizes this G-rich sequence is unclear, as the current structural understanding of KH–RNA recognition advocates a strict requirement for A or C nucleobases in the central positions of the target RNA sequences¹³; nor is it clear what role the different KH domains of KSRP have in the interaction.

Here we set out to explain the molecular bases of KSRP recognition of the let-7 miRNA precursor. We clarify that the different KH domains of KSRP have very different roles in the interaction. We then analyze the interaction between the key KH3 domain and its pre-miRNA target sequence and explain how KSRP can use a conserved KH scaffold to recognize a noncanonical G-rich sequence. Our data highlight that even moderate changes in this recognition have very important effects on the ability of KSRP to promote let-7 biogenesis.

RESULTS

KSRP KH3 dominates the KSRP–let-7 precursor RNA interaction

To clarify the role of the four KH domains of KSRP in pre-let-7a recognition, we have used biophysical binding assays and a set of KSRP

¹Molecular Structure Division, Medical Research Council National Institute for Medical Research, London, UK. ²Departamento de Química Física Biológica, Instituto de Química Física Rocasolano, Consejo Superior de Investigaciones Científicas, Madrid, Spain. ³Medical Research Council Biomedical NMR Centre, Medical Research Council National Institute for Medical Research, London, UK. ⁴Physical Biochemistry Division, Medical Research Council National Institute for Medical Research, London, UK. ⁵Gene Expression Regulation Laboratory IRCCS Azienda Ospedaliera Universitaria San Martino–IST, Genova, Italy. Correspondence should be addressed to A.R. (aramos@nimr.mrc.ac.uk) or R.G. (rgherzi@ucsd.edu).

Received 23 March; accepted 24 September; published online 11 November 2012; doi:10.1038/nsmb.2427

Figure 1 KSRP, pre-let-7a and their interaction. (a) Domain organization of KSRP. (b) Sequence alignment of the four KH domains of KSRP. Conserved amino acids are in red (identity) or other colors (similarity). The short continuous line below the sequence alignment highlights the position of the conserved GXXG loop, and the asterisk marks the position of the K368R mutant. (c) Comparison of sequence and secondary structure of wild-type pre-let-7a miRNA (let-7 WT) and a mutant with a shorter GC-rich stem (let-7-GC), which binds with the same affinity as the wild type and has been used in BLI studies. The terminal loop is shown inside the box, the wild-type pre-let-7a in black and the mutations used for the short stable let-7-GC in red. (d) The contribution of different KH domains to KSRP–let-7 binding, measured with BLI, in which a streptavidin-coated sensor was derivatized with biotinylated let-7-GC and exposed to different concentrations of KSRP (wild type and mutants). Response units are shown plotted against protein concentration, with K_d calculated for the different mutants. KH3 binding was too weak for data fitting, and only a lower limit for the K_d could be estimated.

point mutants. We had previously proposed that KSRP recognizes its various RNA targets by combinatorial and diverse use of its KH domains¹⁴. More recently, we have established that by systematically mutating the conserved GXXG loop in each of the KH domains of KSRP to GDDG we can eliminate RNA binding by a single domain in the context of the full-length protein without altering the structural framework of recognition¹⁵. This allows an accurate evaluation of the importance of the different domains in the interaction. Here we test the binding of the four GDDG KSRP mutants to the let-7 terminal loop by using biolayer interferometry (BLI). To reduce nonspecific RNA absorption to the surface of the sensor used in the BLI measurements, we have substituted the wild-type pre-let-7a stem with a shorter stable GC-rich stem (let-7-GC, Fig. 1c), after having confirmed by CD that this change does not affect KSRP binding to the terminal loop (Supplementary Fig. 1). BLI data show that wild-type KSRP and the KH1 mutant have equivalent affinities for the let-7-GC ($K_d \sim 250$ nM), suggesting that KH1 does not contribute to binding, whereas the KH2 and KH4 mutants bind with an affinity equal to half that of the wild-type protein (Fig. 1d and Supplementary Fig. 1). In marked contrast, we did not detect any interaction between the KH3 mutant and the RNA even at the highest tested concentration (600 nM), which indicated a $K_d > 6,000$ nM. Hence, our biophysical binding assays show that KH3 dominates the KSRP–let-7 precursor interaction in the context of the full-length protein, although both KH2 and KH4 provide a small contribution to protein-binding affinity. The modest contribution of the KH2 and KH4 domains to binding is consistent with the difference in affinity we measure between the KSRP–let-7-GC and the KH3–AGGGU interactions. Further, KH3 binds with the same affinity to AGGGU and let-7-GC RNAs, and binding to the two RNAs results in analogous chemical-shift changes in the KH3 ¹⁵N-¹H NMR spectrum (Supplementary Fig. 2), which indicates that no appreciable interaction is taking place between KH3 and nucleotides outside the AGGGU sequence. Finally, the comparative analysis of methyl resonances in the ¹³C-¹H spectrum of KSRP, alone and in complex with the AGGGU RNA, shows that only resonances of KH3 are affected by AGGGU binding and confirms that

KH2 and KH4 do not make appreciable interactions with the short KH3 target sequence (Supplementary Fig. 2).

KH3 recognition of the target pre-let-7 sequence

The different contributions of the four KH domains of KSRP to the interaction with pre-let-7 RNA clarify that, in order to identify the key determinants of KSRP–let-7 recognition and to understand regulation of let-7 biogenesis, the KH3–RNA interaction must be dissected. KH domains can interact with five or more nucleotides, but only four nucleobases make contact with the hydrophobic groove of the protein¹³. Of these, the two central bases (positions 2 and 3) are recognized with higher selectivity and are either an A or a C in all the published structures^{13,16,17}. However, we have previously shown that the isolated KSRP KH3 domain prefers Gs in these central positions, which is consistent with the identified AGGGU target sequence on pre-let-7a^{9,14}, and we have now confirmed that the contacts that KH3 makes with the AGGGU RNA recapitulate the interaction with the larger pre-let-7 terminal loop. To explain how KSRP KH3 can achieve the recognition of noncanonical G nucleobases within a canonical KH structural framework, we have determined the structure of the KH3–AGGGU complex (Table 1 and Fig. 2 and Supplementary Fig. 3), which we describe below with a focus on the two central and more specific positions.

Our NMR structure shows that AGGGU RNA (referred to below as nucleobases A1, G2, G3, G4 and U5, in positions 0, 1, 2, 3 and 4 of the specifically recognized sequence) interacts with the nucleic acid–recognition groove of KSRP KH3. Although the A1 nucleobase makes limited contacts with the protein (Fig. 2 and Supplementary Fig. 3) and its position is not well defined in our solution structure, the position and the relative arrangement of the four RNA nucleobases in the hydrophobic groove (G2, G3, G4 and U5) are similar to those reported for the equivalent nucleobases of other KH–RNA complexes (Fig. 3)^{16,17}. However, the hydrophobic groove of KH3 is longer than

Table 1 NMR and refinement statistics for the KSRP KH3–AGGGU RNA complex

	Protein	RNA
NMR distance and dihedral constraints		
Distance restraints		
Total NOE	2,366	47
Unambiguous	2,296	
Ambiguous	70	
Intra-residue	895	33
Inter-residue	1,471	14
Sequential ($ i - j = 1$)	486	14
Non-sequential ($ i - j > 1$)	985	
Hydrogen bonds	23	
Protein–RNA intermolecular		67
Total dihedral angle restraints		
Sugar pucker		5
Backbone		10
Structure statistics		
Violations (mean $\pm$ s.d.)		
Distance constraints (>0.3 Å)	1	
Max. distance constraint violation (Å)	0.367	
Deviations from idealized geometry		
Bond lengths (Å)	0.002 ± 0.000	
Bond angles ($^{\circ}$)	0.363 ± 0.003	
Impropers ($^{\circ}$)	0.249 ± 0.006	
Average pairwise r.m.s. deviation ^a (Å)		
Protein		
Heavy (residues 11–46, 55–82)	1.26 ± 0.18	
Backbone (residues 11–46, 55–82)	0.48 ± 0.10	
RNA		
All RNA heavy (residues 11–46, 55–82, RNA)		0.74 ± 0.17
Complex		
All complex heavy (C, N, O, P) (residues 11–46, 55–82, RNA)	1.21 ± 0.17	

^aStatistics from the final 20 water-minimized structures.

the grooves of the Nova-1 and hnRNPK KH3 domains and broadens in a wide pocket that allows the protein to accommodate a G (G2 and G3) in positions 1 and 2 (Figs. 2 and 3 and Supplementary Fig. 3). G2 rests on a broad hydrophobic platform comprising the Val334, Gly335 and Val336 amino acids (Fig. 2 and Supplementary Fig. 3), and its Watson-Crick edge does not engage in hydrogen-bonding in our structure, which is consistent with the lack of specificity that we have previously observed by scaffold-independent analysis¹⁴.

Figure 2 Structure of the KH3–RNA complex. (a) The protein surface of the lowest-energy conformer is in gray (hydrophobic amino acids in yellow), and the RNA heavy atom representation is color coded (blue, nitrogen; green, carbon; red, oxygen). The RNA nucleotides are labeled. (b) Two enlargements of the structure highlight the hydrogen bonds observed between the RNA G3 in position 2 (top) and G4 in position 3 (bottom) and the protein. (c) Protein–RNA contacts observed in the KH3–AGGGU structure. Hydrogen bonds are in red, hydrophobic interactions in blue. The position of the solvent-exposed side chain of the GXXG arginine residue (Arg340) is not well defined, and an interaction with an RNA phosphate group is possible.

Several of the hydrophobic amino acids contacting G2 also make contact with G3, but G3–protein contacts include a hydrogen bond between the Watson-Crick edge of the base and a protein side chain (Fig. 2). KH3 shows a limited specificity in position 2 and can tolerate mutation of G3 to A and U but not C (Supplementary Fig. 4 and Supplementary Table 1), most likely because of a steric clash between the C amino group and a methyl group of Val334. In contrast to G3, G4 forms four intermolecular hydrogen bonds, three of which are with the protein backbone. This creates a recognition pattern that is very specific but also different from the canonical pattern observed in the known KH–RNA structures (Figs. 2 and 3 and Supplementary Fig. 4). Indeed, rather than binding the backbone amide and carboxy group of the same amino acid, this RNA nucleobase recognizes the carboxy and amide groups of two separate amino acids, Ile356 and Phe358 (Figs. 2 and 3). NMR binding assays confirm our structural observations, indicating that G4 recognition is highly specific and that mutation of the nucleobase in this position causes a drop in affinity of between one and two orders of magnitude (Supplementary Fig. 4 and Supplementary Table 1). Finally, the U5 nucleobase stacks onto G4 (Fig. 2 and Supplementary Fig. 3), a common arrangement in KH–nucleic acid complexes, and it is possible that further contacts take place with the surrounding solvent-exposed protein side chains. Globally, the KH3–AGGGU structure explains how the canonical KH–RNA recognition can be adapted to select a G-rich sequence.

KH3–AGGGU recognition and let-7 processing

Next, we assessed the functional role and contribution of KH3–AGGGU recognition to the regulation of let-7 biogenesis. To evaluate how sensitive the system is to changes in the KH3–RNA interaction, we tested the processing activity of two KSRP mutants, one with a fully compromised RNA-binding capability of KH3 (the KH3 GDDG mutant) and the other where the KH3 affinity for the target RNA is reduced to one-third. The second mutant was designed on the basis of the new structural information provided by the KH3–AGGGU complex. The structure shows that Lys368, which is solvent exposed in the free protein and whose mutation does not change KH3 structure and stability (Supplementary Fig. 5), forms a hydrogen bond with G3 O6. In most other canonical KH–RNA complexes, an arginine residue is present in this position and is associated with the specific recognition of a C (Fig. 3 and Supplementary Fig. 5). We mutated Lys368 to arginine and, using NMR and isothermal titration calorimetry (ITC), compared the binding of wild-type KH3 and the stable K368R mutant to the AGGGU RNA (and to the three-nucleobase permutations of G3; Supplementary Table 1 and Supplementary Fig. 5). The lysine-to-arginine mutation results in a decrease in affinity for the target (from

Figure 3 Comparison of the KSRP KH3–AGGGU complex with three representative KH–nucleic acid complexes (that is, the Nova KH3, SF1 KH3 and hnRNP KH3 complexes with nucleic acid). Left, surface representations of the three proteins (gray) with hydrophobic residues (L, I, V, A, F) in red. The nucleic acids are in yellow. KSRP has a broader hydrophobic groove that accommodates a G in positions 1 and 2 (G2 and G3 of the AGGGU). Middle, the specific recognition of a C in position 2 by Nova1 KH3 and hnRNP KH3 is mediated by a set of hydrogen bonds to the Watson-Crick edge of the base. Here we display the two hydrogen bonds formed between the guanidinium group of an arginine in the protein β 3 strand and the O2 and N3 groups of the base. The equivalent amino acid in KSRP KH3 is a lysine (Lys368), and only one hydrogen bond is observed, with the O3 of the G. The different degree of specificity of the interactions is consistent with our published scaffold-independent analysis data as well as with the binding data reported in **Supplementary Table 1**, that indicate that the protein tolerates a U, an A and a G in position 2. Right, in all the complexes the nucleobase in position 3 is recognized specifically by a network of intermolecular hydrogen bonds, but in KSRP KH3 the pattern of hydrogen bonds is different from that observed for other KH–nucleic acid complexes where the amide and carboxy group of the same amino acid are interacting with the nucleobase.

5 to 19 μ M), a result that we have validated also in the context of the pre-let-7 terminal loop (from 6 to 16 μ M; **Supplementary Fig. 5**).

We then explored in cell extract whether KSRP–let-7 precursor association is affected by the mutation. We used ribonucleoprotein immunoprecipitation (RIP) assays to assess the association of wild-type KSRP and KH3 GDDG and K368R mutants to pre-let-7a (Fig. 4a). The results indicate that K368R mutant association to pre-let-7a-1 is strongly affected by the mutation. Finally we tested the efficiency of pri-let-7a processing by wild-type KSRP and by the KH3 GDDG and KH3 K368R mutants in HEK-293 cells and found that although pri-let-7a was efficiently processed by extracts from cells expressing wild-type KSRP, its processing was greatly reduced by both KH3 GDDG and KH3 K368R mutations (Fig. 4b,c and **Supplementary Fig. 5**). These results suggest that a relatively small decrease in affinity of the KH K368R mutant for RNA has a strong effect on KSRP activity and demonstrate that KH3 recognition of the G-rich target sequence is key to the role of KSRP in regulating let-7 biogenesis.

DISCUSSION

The interaction of protein regulators with the terminal loop of the let-7 miRNA precursors determines the concentration of the mature let-7 miRNA, but understanding of this interaction is still incomplete. Here we define the contribution of the individual KH domains of the protein KSRP, a positive regulator of let-7 biogenesis, to its interaction with the

let-7 miRNA precursor. We show that, contrary to what we observed in KSRP–TNF α AU-rich–element recognition¹⁵, one of the domains of the protein (KH3) has a dominant role in the interaction. The different use of KH domains in KSRP recognition of the two TNF α AU-rich elements and pre-let-7 RNAs emphasizes the ability of the protein to target a broad range of RNAs and to act at different post-transcriptional regulatory steps. We determine the structure of the functionally well-characterized KH3–AGGGU complex, which encompasses the key recognition elements of the KSRP–let-7 precursor interaction, and we show that even a moderate impairment in the ability of KH3 to recognize a specific G-rich sequence has a strong effect on KSRP's ability to promote pri-let-7 processing. This assay both provides a functional validation of our molecular data and highlights the exquisite sensitivity of recognition in this multicomponent system.

In this work, we dissect the structural contacts between KSRP KH3 and its G-rich target sequence. The structures of several KH domains

Figure 4 KH3 interaction with an AGGGU RNA in the let-7-precursor hairpin domain is essential for KSRP function in let-7 biogenesis. (a) Immunoprecipitation assay of pre-let-7a-1 using extracts from HEK-293 cells transfected with either pTAG2B-KSRP, pTAG2B-KH3GDDG or pTAG2B-KSRPK368R. The values shown are averages ($\pm$ s.e.m.) of three independent experiments. (b) *In vitro* pri-let-7a-1 processing assays performed using total extracts from HEK-293 cells transiently transfected with either pTAG2B-KSRP (lanes 1–4), pTAG2B-KH3GDDG (lanes 5–8) or pTAG2B-KSRPK368R (lanes 9–12). Internally ³²P-labeled RNA substrate was added and used to monitor pri-let-7a-1 processing. (c) The intensity of the bands corresponding to pri-let-7a-1 and pre-let-7a-1, quantified with ImageJ software (<http://rsb.info.nih.gov/ij/index.html>) and expressed as percentage of pre-let-7a-1 generated from pri-let-7a-1 at each time point. To avoid signal saturation, the quantification was performed on underexposed autoradiograms.

in complex with nucleic acids have been solved, and the analysis of the KH-RNA (and KH-DNA) contacts advocates for a requirement for A or C nucleobases in the central positions of the bound nucleic acid¹³. However, the KH3-AGGGU structure shows that relatively minor changes in the geometry and the hydrophobicity of the nucleic acid-recognition groove of a KH domain facilitates the binding and recognition of G nucleobases. Particularly striking is how the previously described use of a Watson-Crick-like recognition of the base in position 3 is here maintained by shifting the contacts made by the RNA base along the protein β -strand. This finding provides a more general understanding of KH-RNA recognition and facilitates the design of KH domains with different sequence specificities.

The regulation of let-7 biogenesis involves several proteins, including hnRNP A1, lin-28 and KSRP, and it is a prototype for the post-transcriptional regulatory mechanisms that selectively control miRNA concentration. We have previously shown that high levels of lin-28 can counteract KSRP-mediated upregulation of let-7 biogenesis in P19 cells⁹. However, in the recently reported structure of the lin-28-let-7 complex¹¹, the AGGGU sequence target of KH3 does not directly interact with lin-28. We expect that future work will define the molecular interplay between these factors at the molecular level, providing further insight into how let-7 biogenesis is regulated.

METHODS

Methods and any associated references are available in the [online version of the paper](#).

Accession codes. Atomic coordinates for the NMR ensemble have been deposited in the Protein Data Bank, with accession code [4B8T](#), and the corresponding chemical-shift assignments have been deposited in the Biological Magnetic Resonance Data Bank, with accession code [18702](#).

Note: Supplementary information is available in the [online version of the paper](#).

ACKNOWLEDGMENTS

All the NMR experiments were recorded at the Medical Research Council Biomedical NMR Centre at the Medical Research Council National Institute for Medical Research, and we thank T. Frenkiel and A. Oregioni for assistance. We thank S. Howell for mass spectrometry analysis of protein samples. The work of A.R. has been funded by the Medical Research Council (MRC) grant U117574558 and the Wellcome Trust Grant WT082088MA. The work of S.R.M. and G.K. has been funded by MRC grants U117570592 and U117533887, respectively. The work of P.B. and R.G. has been supported by grants from the Association for International Cancer Research (10-0527) and Associazione Italiana per la Ricerca sul Cancro (I.G. 10090). We thank B. Faust for help in the expression and purification of the KH3 K368R mutant.

AUTHOR CONTRIBUTIONS

Cloning and RNA synthesis were performed by D.H. Protein purification was performed by G.N. and D.H. Recording and analysis of the NMR spectra were performed by G.N., M.F.G.-M. and G.K. Structure calculations and recording of the isothermal titration calorimetry, CD and BLI data were performed by G.N. Analysis of the CD and isothermal titration calorimetry data was performed by S.R.M. and G.N. RIP and pri-miRNA processing assays were performed by P.B. and R.G. The manuscript was written by A.R. and G.N.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at <http://www.nature.com/doi/10.1038/nsmb.2427>.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

- Kim, V.N., Han, J. & Siomi, M.C. Biogenesis of small RNAs in animals. *Nat. Rev. Mol. Cell Biol.* **10**, 126–139 (2009).
- Krol, J., Loedige, I. & Filipowicz, W. The widespread regulation of microRNA biogenesis, function and decay. *Nat. Rev. Genet.* **11**, 597–610 (2010).
- Michlewski, G., Guil, S., Semple, C.A. & Cáceres, J.F. Posttranscriptional regulation of miRNAs harboring conserved terminal loops. *Mol. Cell* **32**, 383–393 (2008).
- Büssing, I., Slack, F.J. & Grosshans, H. Let-7 microRNAs in development, stem cells and cancer. *Trends Mol. Med.* **14**, 400–409 (2008).
- Viswanathan, S.R., Daley, G.Q. & Gregory, R.I. Selective blockade of microRNA processing by Lin28. *Science* **320**, 97–100 (2008).
- Heo, I. *et al.* TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. *Cell* **138**, 696–708 (2009).
- Hagan, J.P., Piskounova, E. & Gregory, R.I. Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. *Nat. Struct. Mol. Biol.* **16**, 1021–1025 (2009).
- Michlewski, G. & Cáceres, J.F. Antagonistic role of hnRNP A1 and KSRP in the regulation of let-7a biogenesis. *Nat. Struct. Mol. Biol.* **17**, 1011–1018 (2010).
- Trabucchi, M. *et al.* The RNA-binding protein KSRP promotes the biogenesis of a subset of microRNAs. *Nature* **459**, 1010–1014 (2009).
- Loughlin, F.E. *et al.* Structural basis of pre-let-7 miRNA recognition by the zinc knuckles of pluripotency factor Lin28. *Nat. Struct. Mol. Biol.* **19**, 84–89 (2012).
- Nam, Y., Chen, C., Gregory, R.I., Chou, J.J. & Sliz, P. Molecular basis for interaction of let-7 microRNAs with Lin28. *Cell* **147**, 1080–1091 (2011).
- Gherzi, R., Chen, C.Y., Trabucchi, M., Ramos, A. & Briata, P. The role of KSRP in mRNA decay and microRNA precursor maturation. *Wiley Interdiscip. Rev. RNA* **1**, 230–239 (2010).
- Valverde, R., Edwards, L. & Regan, L. Structure and function of KH domains. *FEBS J.* **275**, 2712–2726 (2008).
- García-Mayoral, M.F., Díaz-Moreno, I., Hollingworth, D. & Ramos, A. The sequence selectivity of KSRP explains its flexibility in the recognition of the RNA targets. *Nucleic Acids Res.* **36**, 5290–5296 (2008).
- Hollingworth, D. *et al.* KH domains with impaired nucleic acid binding as a tool for functional analysis. *Nucleic Acids Res.* **40**, 6873–6886 (2012).
- Lewis, H.A. *et al.* Sequence-specific RNA binding by a Nova KH domain: implications for paraneoplastic disease and the fragile X syndrome. *Cell* **100**, 323–332 (2000).
- Backe, P.H., Messias, A.C., Ravelli, R.B., Sattler, M. & Cusack, S. X-ray crystallographic and NMR studies of the third KH domain of hnRNP K in complex with single-stranded nucleic acids. *Structure* **13**, 1055–1067 (2005).

ONLINE METHODS

RNA and protein samples. The different KSRP constructs were cloned into a modified pETM-30 vector (European Molecular Biology Laboratory Heidelberg, protein-expression facility) and expressed as His-GST fusion proteins. Unlabeled and uniformly ^{15}N - and ^{13}C -labeled samples of KSRP KH3 domain were obtained as previously described¹⁸. Briefly, a His-GST fusion protein comprising amino acids 423 to 525 of human KSRP protein (NM_003685) was expressed in *E. coli* and purified by Ni-IMAC. The tag was removed by overnight TEV protease digestion at 4 °C, followed by a second Ni-IMAC step and by size-exclusion chromatography (Superdex 75 16/60 column, Pharmacia). The KH1–KH4 KSRP construct (amino acids G68–Q525) which is used in our biophysical assays and is referred to in the paper as ‘KSRP’ was expressed and purified as above, except that a MonoQ 5/50GL anion-exchange column (GE Healthcare) followed by a 1 ml HiTrap Heparin column (GE Healthcare) was used. The final proteins were stored in 10 mM Tris-HCl, pH 7.4, 50 mM NaCl (or LiCl), 2 mM Tris(2-carboxyethyl)phosphine (TCEP), 0.05% (w/v) NaN_3 . The concentration of the protein samples was determined from their 280-nm absorbance, and their molecular weight and purity were confirmed by electrospray mass spectrometry. KSRP K368R and GXXG-to-GDDG constructs were prepared by using the QuikChange Site-Directed Mutagenesis Kit (Stratagene) according to manufacturer's instructions as described previously¹⁵ and were expressed and purified as the wild-type protein. Wild-type Let-7-SL RNA was synthesized by run-off transcription and purified by using denaturing acrylamide gels, as previously described¹⁹. All other nonbiotinylated or biotinylated oligonucleotides were purchased from Dharmacon and IDT.

NMR spectroscopy. NMR experiments were recorded at 25 °C, 15 °C and 5 °C on Bruker Avance and Varian Inova spectrometers operating at 700, 600 and 800 MHz ^1H frequency. Protein and RNA samples were in buffer of 10 mM Tris, pH 7.4, 50 mM NaCl (or LiCl). Backbone and side chain resonance assignments were obtained by using 2D ^1H - ^{15}N HSQC, 2D ^1H - ^{13}C HSQC, 3D HNCA, 3D CBCACONH, 3D HCCH-TOCSY, 3D ^{15}N NOESY-HSQC experiments as well as 3D ^{13}C NOESY-HSQC experiments optimized for either aliphatic or aromatic resonances. Assignment of the resonances of the free and KH3-bound AGGGU RNA were obtained by using 2D ^1H - ^1H TOCSY and 2D ^1H - ^1H NOESY experiments. NOESY experiments were recorded with mixing times between 100 ms and 250 ms, and TOCSY experiments with mixing times of 30, 50, 60, 70 and 100 ms. Intramolecular restraints were derived from ^{15}N - and ^{13}C -NOESYHSQC spectra with a mixing time of 100 ms at 25 °C. Intermolecular NOEs were obtained from the analysis of decoupled and nondecoupled 2D ^1H - ^1H NOESY (mixing times, 50, 100, 200 and 250 ms), 3D ^{15}N NOESY-HSQC (mixing time, 100 ms), 3D ^{13}C NOESY-HSQC (mixing times, 100 and 120 ms) and 3D-filtered ^{13}C NOESY on samples of unlabeled RNA and ^{15}N - and ^{13}C -labeled protein (mixing times, 100 and 150 ms) at 25 °C. The temperature dependence of backbone amide chemical shift was calculated as the ratio between chemical shift and temperatures in ^{15}N - ^1H correlated spectra recorded at 10, 20, 24, 30, 35, 40 and 45 °C. NMR spectra were processed by using the NMRpipe suite of programs²⁰ and analyzed by using the Sparky²¹ and XEASY²² programs.

Structure calculations. The structure of the KH3–AGGGU complex was calculated by using a semiautomated ARIA 1.2–based protocol²³. Experimental distance restraints were obtained from the integration of NOE peaks in 3D and 2D NOESY spectra by using the XEASY program²². All intraprotein NOE cross-peaks were calibrated automatically and assigned iteratively within ARIA, whereas the peaks arising from RNA proton resonances were calibrated manually in a semiquantitative fashion²⁴. Protein angle restraints were derived from chemical-shift analysis by using the program TALOS. RNA angle restraints (α , ζ and δ) were derived from ^{31}P - ^1H correlation spectra and ^1H - ^1H TOCSY spectra as described previously²⁴. Structures were initially calculated without the use of hydrogen bond restraints; only if a proton was hydrogen-bonded in at least 50% of this initial set of structures was the corresponding hydrogen bond restraint added in the final set of calculations. Intermolecular hydrogen bond constraints were validated by a downfield shift of the proton resonance upon RNA binding²⁵ and a low temperature dependence of the chemical shifts^{26,27}. ARIA 1.2 was used to calculate 100 conformers of the complex (iterations 0–7). The 40 conformers with the lowest restraint energies were refined in a shell of explicit water. The 20 conformers with the lowest restraint energies, restraint

violations and r.m.s. deviations from the ideal covalent geometry were taken as representative of the converged structures and selected for structural analysis. The structures were analyzed by using the programs Molmol²⁸, InsightII (Accelrys) and PyMOL (<http://www.pymol.org/>). A Ramachandran analysis of the structure shows 83.8%, 14.9%, 0.8% and 0.8% of the protein residues in the most favored, additional, generously and disallowed regions, respectively.

ITC binding assays. ITC experiments were recorded in 10 mM Tris, pH 7.4, 50 mM NaCl (or LiCl for the AGGGU RNA) by titration of a concentrated RNA solution into a cell containing KSRP KH3. For the AGGGU RNA, which can form G tetrads at high concentrations, a concentrated solution of protein was added to the RNA. Protein concentration was optimized according to the expected binding affinities of the different RNAs, as derived from NMR experiments. Dissociation constants were determined by fitting the measured heat of reaction at 25 °C by using the Origin analysis package and a 1:1 binding model. The heat of dilution of the different RNAs was evaluated in control experiments where RNA solutions were injected into buffer.

NMR binding assays. ^1H - ^{15}N correlation experiments (typically SOFAST ^{15}N HMQC spectra) were recorded at the different steps of a titration of ^{15}N -labeled samples of KH3 with the different RNA oligonucleotides. The titrations were performed in 10 mM Tris, pH 7.4, 50 mM NaCl, 2 mM TCEP, at protein concentration between 50 and 100 μM . To obtain K_d values, the chemical-shift changes ($\Delta\delta_{\text{ave}} = [(\Delta\text{NH})^2 + (\Delta^{15}\text{N}/10)^2]^{1/2}$) of five to ten peaks in the fast regime of exchange were measured and plotted against increasing RNA:protein ratios. A nonlinear two-parameter least-square fit procedure was applied to each data set by using the program Origin and a single-site binding model²⁹. K_d values are reported as mean $\pm$ 2 s.d.

CD binding assays. CD spectra were recorded on a Jasco J-715 spectropolarimeter equipped with a Peltier system for temperature control. The titrations were performed in 10 mM Tris, pH 7.4, 150 mM NaCl and 2 mM TCEP by adding increasing amounts of protein to RNA samples. RNA concentration varied between 1.5 μM and 3 μM . Experiments were recorded at 5 °C to maximize signal to noise, and the CD signal at 260 nm was fitted to the protein:RNA ratio by using nonlinear regression and in-house software as described in ref. 30. Thermal denaturation of the KH3 and KH K368R mutant was monitored by using the ellipticity at 220 nm as a function of temperature between 5 °C and 90 °C at a rate of 1 °C per min. The data were fitted to a two-state model with in-house software, as reported in ref. 18.

BLI binding assays. BLI experiments were recorded in 10 mM Tris, pH 7.4, 150 mM NaCl, 2 mM TCEP, 0.5 mg/ml BSA, 0.005% TWEEN 20 at 25 °C on a Octet Red instrument (ForteBio, Inc). Briefly, the biotinylated RNAs were immobilized on Streptavidin-coated sensors and exposed to different concentrations of the wild-type and mutant four-domain proteins. Dissociation constants were derived by fitting the increase in BLI signal as a function of protein concentration (Fig. 1 and Supplementary Fig. 1) by using nonlinear regression and in-house software, as described in ref. 31.

Pri-miRNA *in vitro* processing assays. Pri-let-7a-1 processing assays were performed essentially as reported in ref. 32. Briefly, total cell extracts from HEK-293 cells were prepared in 50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.5% Triton X-100, 1 $\times$ Complete Protease Inhibitor Cocktail (Roche) and 10% glycerol and incubated (typically 40 μg per 25 μl reaction at 37 °C for the indicated times) with *in vitro*-synthesized and uniformly labeled pri-Let-7a-1 (5 fmol) in processing buffer containing 100 mM potassium acetate, 2 mM magnesium acetate, 10 mM Tris-Cl, pH 7.6, 2 mM DTT, 10 mM creatine phosphate, 1 μg creatine phosphokinase, 1 mM ATP, 0.4 mM GTP, 0.1 mM spermine and 2 units Stop RNase Inhibitor (5 PRIME). For some experiments, 4 μg of HEK-293 total cell extracts were preincubated (1 h at 16 °C under continuous agitation in a Thermomixer (Eppendorf) with immunocomplexes derived from anti-Flag immunoprecipitation of 250- μg aliquots of total extracts of HEK-293 cells transiently transfected (Lipofectamine Plus, Invitrogen) with either empty pTAG2B vector (Stratagene) or pTAG2B-KSRP, pTAG2B-KH3GDDG or pTAG2B-KSRPK368R. Immediately after preincubations, pri-let-7a-1 processing assays were performed.

Ribonucleoprotein-complex immunoprecipitation (RIP) assays. RIP assays were performed as previously described in ref. 32, with the following modifications. Briefly, cell lysates were immunoprecipitated with Protein G-coupled magnetic beads, cross-linked to anti-Flag (M2, Sigma, St. Louis, MO, USA) mouse monoclonal antibody, at 4 °C overnight. Pellets were washed four times with 50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.5% Triton X-100. Total RNA was prepared from immunocomplexes by using miRNeasy Mini Kit (Qiagen), retrotranscribed and amplified by qPCR by using miScript Precursor Assays for pre-let-7a-1 (Qiagen).

18. García-Mayoral, M.F. *et al.* The structure of the C-terminal KH domains of KSRP reveals a noncanonical motif important for mRNA degradation. *Structure* **15**, 485–498 (2007).
19. Nikonowicz, E.P. *et al.* Preparation of ^{13}C and ^{15}N labelled RNAs for heteronuclear multi-dimensional NMR studies. *Nucleic Acids Res.* **20**, 4507–4513 (1992).
20. Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
21. Pettersen, E.F. *et al.* UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
22. Bartels, C., Xia, T., Billeter, M., Guntert, P. & Wuthrich, K. The program XEASY for computer supported NMR spectral analysis of biological macromolecules. *J. Biomol. NMR* **6**, 1–10 (1995).
23. Linge, J.P., Habeck, M., Rieping, W. & Nilges, M. ARIA: automated NOE assignment and NMR structure calculation. *Bioinformatics* **19**, 315–316 (2003).
24. Varani, G., Aboul-ela, F. & Allain, F.H.-T. NMR investigation of RNA structure. *Prog. Nucl. Magn. Reson. Spectrosc.* **29**, 51–127 (1996).
25. Liu, Z. *et al.* Structural basis for recognition of the intron branch site RNA by splicing factor 1. *Science* **294**, 1098–1102 (2001).
26. Baxter, N.J. & Williamson, M.P. Temperature dependence of ^1H chemical shifts in proteins. *J. Biomol. NMR* **9**, 359–369 (1997).
27. Cierpicki, T., Zhukov, I., Byrd, R.A. & Otlewski, J. Hydrogen bonds in human ubiquitin reflected in temperature coefficients of amide protons. *J. Magn. Reson.* **157**, 178–180 (2002).
28. Koradi, R., Billeter, M. & Wuthrich, K. MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 29–32 (1996).
29. García-Mayoral, M.F., Díaz-Moreno, I., Hollingworth, D. & Ramos, A. The sequence selectivity of KSRP explains its flexibility in the recognition of the RNA targets. *Nucleic Acids Res.* **36**, 5290–5296 (2008).
30. Martin, S.R., Masino, L. & Bayley, P.M. Enhancement by Mg^{2+} of domain specificity in Ca^{2+} -dependent interactions of calmodulin with target sequences. *Protein Sci.* **9**, 2477–2488 (2000).
31. Cukier, C.D. *et al.* Molecular basis of FIR-mediated c-myc transcriptional control. *Nat. Struct. Mol. Biol.* **17**, 1058–1064 (2010).
32. Briata, P. *et al.* PI3K/AKT signaling determines a switch between distinct KSRP functions favoring skeletal myogenesis. *Cell Death Differ.* **19**, 478–487 (2012).