ZBP1 recognition of β-actin zipcode induces RNA looping

Jeffrey A. Chao,[1] Yury Patskovsky,[2] Vivek Patel,[1] Matthew Levy,[2] Steven C. Almo,[2] and Robert H. Singer[1,3]

[1]Department of Anatomy and Structural Biology, Albert Einstein College of Medicine, Bronx, New York 10461, USA; [2]Department of Biochemistry, Albert Einstein College of Medicine, Bronx, New York 10461, USA

ZBP1 (zipcode-binding protein 1) was originally discovered as a *trans*-acting factor for the "zipcode" in the 3′ untranslated region (UTR) of the β-actin mRNA that is important for its localization and translational regulation. Subsequently, ZBP1 has been found to be a multifunctional regulator of RNA metabolism that controls aspects of localization, stability, and translation for many mRNAs. To reveal how ZBP1 recognizes its RNA targets, we biochemically characterized the interaction between ZBP1 and the β-actin zipcode. The third and fourth KH (hnRNP K homology) domains of ZBP1 specifically recognize a bipartite RNA element located within the first 28 nucleotides of the zipcode. The spacing between the RNA sequences is consistent with the structure of IMP1 KH34, the human ortholog of ZBP1, that we solved by X-ray crystallography. The tandem KH domains are arranged in an intramolecular anti-parallel pseudodimer conformation with the canonical RNA-binding surfaces at opposite ends of the molecule. This orientation of the KH domains requires that the RNA backbone must undergo an ~180° change in direction in order for both KH domains to contact the RNA simultaneously. The RNA looping induced by ZBP1 binding provides a mechanism for specific recognition and may facilitate the assembly of post-transcriptional regulatory complexes by remodeling the bound transcript.

[*Keywords*: ZBP1; RNA-binding protein; KH domain; RNA localization]

Supplemental material is available at http://www.genesdev.org.

Received September 10, 2009; revised version accepted November 23, 2009.

Localization of messenger RNA (mRNA) into distinct subcellular compartments allows for the spatial regulation of gene expression that is required for the establishment and maintenance of cell polarity (for review, see Martin and Ephrussi 2009). A global study using *Drosophila* embryos found that >71% of all the mRNAs characterized (3370 genes) were localized and, furthermore, these mRNA could be grouped into 35 unique localization patterns (Lecuyer et al. 2007). A second genome-wide study that isolated mRNAs from fibroblast cell protrusions identified >50 mRNAs that were specifically localized to pseudopodia (Mili et al. 2008). These recent studies underscore the fundamental role of mRNA localization in diverse cellular and developmental processes and, while technological advances have increased the number of mRNAs that been shown to localize, the underlying mechanisms that give rise to these asymmetric distributions have remained elusive.

The localization of β-actin mRNA to the leading edge of chicken embryo fibroblasts was one of the earliest transcripts identified to be subcellularly localized and has served as a model system for understanding the process (Lawrence and Singer 1986). Asymmetric sorting of the β-actin transcript is achieved by transport along both microtubule and actin microfilaments, and is delocalized in myosin II-B knockout fibroblasts (Latham et al. 1994; Fusco et al. 2003; Oleynikov and Singer 2003). A 54-nucleotide (nt) *cis*-acting element, termed the zipcode, positioned directly following the termination codon in the 3′ untranslated region (UTR) of β-actin mRNA was shown to be necessary and sufficient for targeting reporter RNA constructs to the cellular periphery (Kislauskis et al. 1994). The *trans*-acting factor Zipcode-binding protein 1 (ZBP1) was identified based on its ability to interact with the zipcode, and its knockdown results in impaired invadopodia formation, cytoplasmic spreading, and cell adhesion (Ross et al. 1997; Vikesaa et al. 2006).

ZBP1 is the founding member of a highly conserved family (termed VICKZ in reference to the founding members:Vg1RBP/Vera, IMP1-3, CRD-BP, KOC, and ZBP1) of RNA-binding proteins that have been implicated in the post-transcriptional regulation of several different RNAs (Yisraeli 2005). In *Xenopus laevis*, Vg1RBP/Vera is required for the localization of Vg1 mRNA to the vegetal cortex of oocytes and also the localization of β-actin mRNA in axons (Deshler et al.

[3]**Corresponding author.**
E-MAIL robert.singer@einstein.yu.edu; FAX (718) 430-8697.
Article is online at http://www.genesdev.org/cgi/doi/10.1101/gad.1862910.

1998; Havin et al. 1998; Leung et al. 2006; Yao et al. 2006). Humans contain three paralogs (IMP1–3) that were originally identified because of their ability to regulate insulin-like growth factor II (Igf-II) mRNA translation, but have since been found to promote the localization of H19 and tau mRNAs as well as stabilize CD44 and β-TrCP1 mRNAs (Nielsen et al. 1999; Runge et al. 2000; Atlas et al. 2004; Vikesaa et al. 2006; Elcheva et al. 2009). In mice, coding region determinant-binding protein (CRD-BP) was shown to protect the c-myc mRNA from endonucleolytic cleavage, thereby stabilizing the transcript (Doyle et al. 1998). This family's ability to broadly regulate RNA metabolism can lead to adverse cellular effects, as evidenced by their overexpression and correlation with poor prognosis in several types of cancers (Hammer et al. 2005; Dimitriadis et al. 2007; Jiang et al. 2008; Kobel et al. 2009).

VICKZ family members share a characteristic arrangement of six canonical RNA-binding modules with two RNA recognition motifs (RRM) followed by four hnRNP-K homology (KH) domains. Sequence alignments of the proteins show that conserved residues are clustered into three didomains (RRM12, KH12, and KH34), which suggests that these evolutionarily conserved regions may function in concert (Fig. 1A; Git and Standart 2002). Interestingly, ZBP1 and IMP1 share considerable sequence identity (>94%), yet studies of their interactions with their respective RNA targets have failed to produce a unified understanding of the requirements for specific RNA recognition (Runge et al. 2000; Farina et al. 2003; Nielsen et al. 2004; Patel and Bag 2006; Atlas et al. 2007; Jonson et al. 2007). Models for RNA recognition differ with regard to both the domains and oligomerization state of ZBP1 required for binding as well as the proposed RNA determinants, which range from a minimal 5′-ACACCC-3′ sequence to RNA binding being entirely sequence-independent (Farina et al. 2003; Nielsen et al. 2004; Atlas et al. 2007; Oberman et al. 2007).

Here we present a biochemical characterization of ZBP1 recognition of the β-actin zipcode RNA. The ZBP1 KH34 monomer binds to two nonsequential stretches of RNA located within the proximal portion of the zipcode. This bipartite recognition element is consistent with our crystal structure of IMP KH34 (98% sequence identity with ZBP1 KH34), the first structural data for this family of proteins, that positions the KH domains in an anti-parallel arrangement with their putative RNA-binding surfaces located at opposite ends of the molecule. This orientation of the KH domains explains both the sequence and distance dependence of RNA binding that were determined biochemically. The residues that link KH3 to KH4 were also shown to play a role in RNA binding, demonstrating that the KH34 domain functions as a single unit whose precise geometry dictates its interaction with RNA.

Results

ZBP1 recognition of zipcode RNA

We took the interaction between ZBP1 and the first 54 nt of the 3′ UTR of β-actin mRNA (zipcode[1–54]) as a starting point to further investigate both the protein and RNA contributions to specific recognition. A polyacrylamide gel electrophoretic mobility shift assay (EMSA) was used to resolve fluorescein-labeled zipcode[1–54] in complex

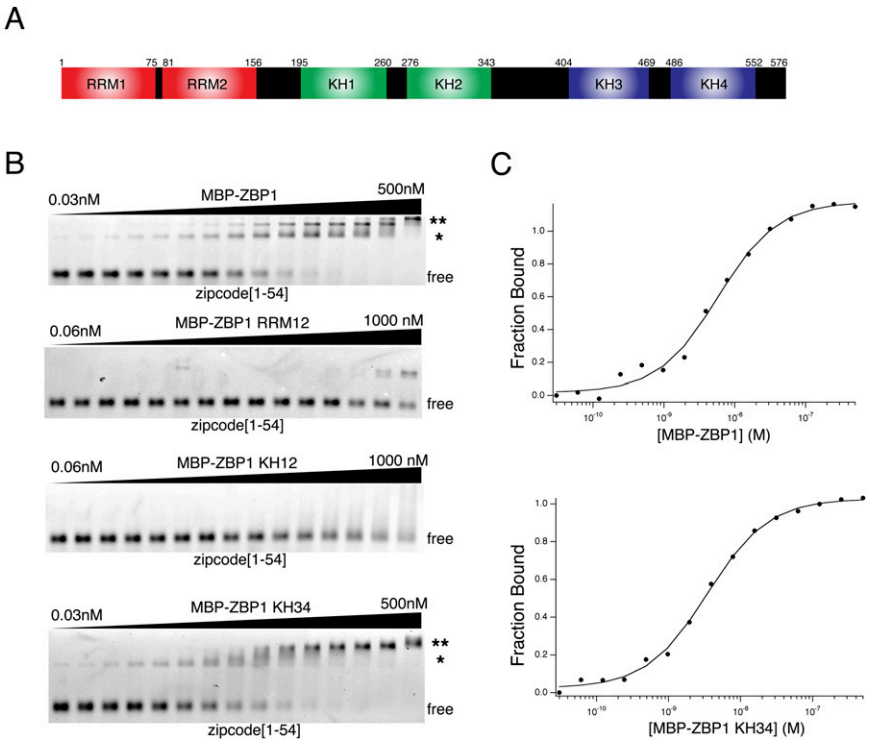

Figure 1. The KH34 didomain of ZBP1 is responsible for recognition of zipcode[1–54] RNA. (*A*) Schematic diagram of ZBP1 showing conserved didomain organization. (*B*) Representative EMSA results for full-length ZBP1, RRM12, KH12, and KH34 binding to zipcode[1–54] RNA. The filled triangle represents a 1:1 serial dilution of recombinant protein. Free RNA (*) and RNA–protein complexes (**) are labeled. (*C*) Quantification of the fraction of RNA bound in EMSA data for ZBP1 and KH34 were fit to the Hill equation to measure the $K_{d,\ \mathrm{app}}$ and Hill coefficient.

with ZBP1 from free RNA. The fraction of bound RNA was measured as a function of protein concentration, and the affinity of recombinant ZBP1 for the zipcode[1–54] was determined by fitting the data to the Hill equation (Ryder et al. 2008). ZBP1 binds tightly to this RNA with an apparent disassociation constant ($K_{d,\ app}$) of 7.3 ± 3.0 nM and a Hill coefficient of 0.9 ± 0.2, which is consistent with previous studies of the ZBP1–zipcode[1–54] interaction measured by nitrocellulose filter binding (Fig. 1B,C; Farina et al. 2003).

In order to determine which of the putative RNA-binding domains of ZBP1 were responsible for recognition of the zipcode[1–54], truncations of ZBP1 that contain either RRM12, KH12, or KH34 were generated. Both ZBP1 RRM12 and KH12 do not bind the zipcode[1–54] with high affinity (Fig. 1B). ZBP1 KH34, however, binds the zipcode[1–54] with a similar affinity as the full-length protein ($K_{d,\ app}$ = 4.7 ± 2.0 nM, Hill coefficient of 0.9 ± 0.2), indicating that recognition of the zipcode is dependent on this domain. Recombinant ZBP1 constructs that contained the individual KH3 or KH4 domains did not bind the zipcode[1–54] with high affinity, which is consistent with previous studies of isolated KH domains having affinities for their targets in the micromolar range (Supplemental Fig. 1; Valverde et al. 2008). These experiments demonstrate that the two KH domains function together for high-affinity RNA binding.

Interestingly, at higher protein concentrations of both full-length ZBP1 and the KH34 domain alone, higher-molecular-weight complexes can be resolved by EMSA (Fig. 1B). While previous experiments performed with the human and frog homologs of ZBP1 (IMP1 and Vg1RBP) concluded that KH34 contains a dimerization motif, we find that both full-length ZBP1 and the KH34 domain exist predominantly as monomeric species in solution, as assayed by size exclusion chromatography at concentrations (>20 μM) above those used in the EMSA experiments (Supplemental Fig. 2; Git and Standart 2002; Nielsen et al. 2004; Oberman et al. 2007). Furthermore, binding to the zipcode was not found to be cooperative (Hill coefficient, ~1) for both full-length ZBP1 and the KH34 domain, indicating that a potential ZBP1 dimer is not stabilized by RNA binding. Based on these data, we hypothesize that the higher-molecular-weight complexes likely result from ZBP1 monomers binding independently to alternative lower-affinity binding sites within the zipcode[1–54]. Previous studies of isolated KH domains have found that their RNA-binding sites are usually comprised of four consecutive single-stranded nucleotides, so it seems plausible that the 54-nt zipcode element could contain multiple binding sites for ZBP1 KH34 (Auweter et al. 2006; Valverde et al. 2008).

Identification of the ZBP1 KH34 RNA-binding site within the zipcode

Since the zipcode[1–54] was found to potentially contain more than one binding site for ZBP1 KH34, fragments of this RNA were synthesized to identify the positions of the nucleotides recognized by ZBP1. A fragment of the zipcode containing nucleotides 1–28 (zipcode[1–28]) was bound by ZBP1 KH34 with similar affinity ($K_{d,\ app}$ = 3.6 ± 0.2 nM, Hill coefficient of 0.9 ± 0.2) as the entire zipcode[1–54] sequence, indicating that the high-affinity binding site is located within this region (Fig. 2A). Importantly, even at the highest protein concentration (500 nM), only a single RNA–protein complex was observed, providing a more tractable system for further biochemical characterization. A second ZBP1-binding site within the zipcode[1–54], however, could not be identified by EMSA. Experiments using a zipcode fragment containing nucleotides 29–54 of the 3′ UTR of β-actin mRNA (zipcode[29–54]) did not result in a well-resolved shifted complex (Fig. 2C). Perhaps in the context

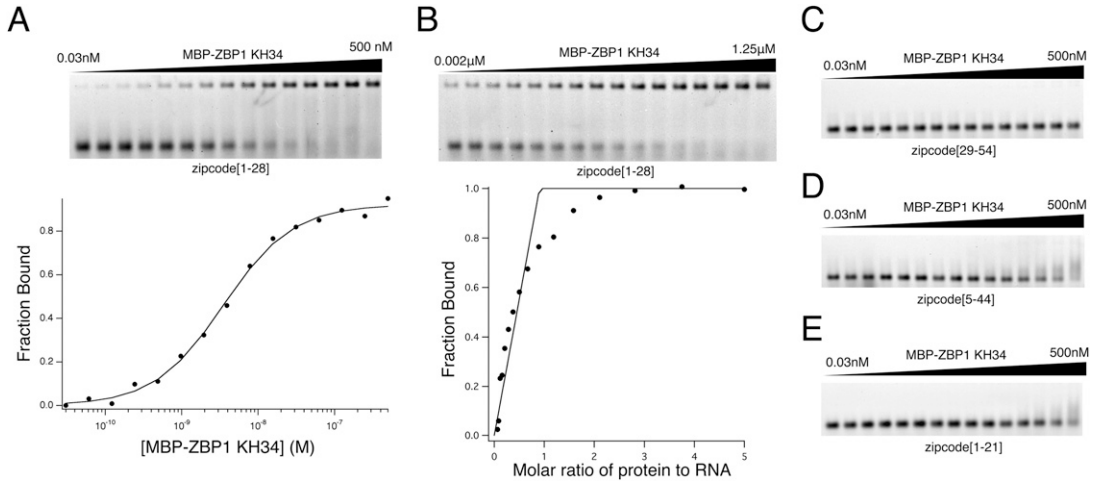

Figure 2. Binding site for ZBP1 KH34 is within the first 28 nt of the zipcode. (*A*) Representative EMSA results for ZBP1 KH34 binding to zipcode[1–28] RNA with fit of data to the Hill equation. (*B*) Representative stoichiometry binding assay for ZBP1 KH34 binding to zipcode[1–28] with fit to quadratic model of saturable ligand binding. (*C–E*) Representative EMSA results for ZBP1 KH34 binding to zipcode[29–54], zipcode[5–44], and zipcode[1–21] RNAs. The filled triangle represents a 1:1 serial dilution or 3:1 serial dilution (for stoichiometry experiment) of recombinant protein.

of a shorter RNA, the lower-affinity binding site within the zipcode is beyond the detection limit of the EMSA. Further experiments will be required to determine if these secondary interaction sites have any biological significance, or if ZBP1's function is mediated by only the high-affinity binding site.

To further clarify the composition of the RNA–protein complex, the stoichiometry of the ZBP1 KH34–zipcode[1–28] interaction was determined by EMSA, where the concentration of the unlabeled zipcode[1–28] fragment (250 nM) was fixed well above the measured $K_{d,\ app}$ (3.6 nM) for this interaction. The data were fit to a quadratic model of saturable ligand binding, resulting in a stoichiometric equivalence point of 1.0 ± 0.1 (Fig. 2B; Rambo and Doudna 2004). The ZBP1 KH34–zipcode[1–28] was also found to have a stoichiometry of 1:1 by size exclusion chromatography (Supplemental Fig. 2). Taken as a whole, the RNA-binding and stoichiometry experiments, in conjunction with the size exclusion chromatography data, demonstrate that ZBP1 KH34 can bind RNA sequences within the zipcode as a monomer, and that multimerization is not required for stable RNA recognition, as suggested previously (Nielsen et al. 2004).

To better define the interaction between ZBP1 KH34 and the zipcode[1–28] RNA, fragments that contained deletions at the 5′ end (zipcode[5–44]) and 3′ end (zipcode[1–21]) of zipcode[1–28] were used in the EMSA. Both of these RNAs failed to bind ZBP1 KH34 with high affinity, demonstrating that nucleotides at both the 5′ and 3′ ends of the zipcode[1–28] are necessary for recognition (Fig. 2D,E). Interestingly, nucleotides 16–21 and 22–27 of the zipcode share an identical nucleotide sequence (ACACCC), yet zipcode[1–21] could not bind to ZBP1 KH34. This suggests that the spacing between the nucleotides with which ZBP1 KH34 interacts at the 5′ and 3′ ends of zipcode[1–28] is an important factor for recognition. It should also be noted that zipcode[5–44] (40 nt) is longer than the zipcode[1–28] (28 nt), indicating that ZBP1 KH34 binding to RNA is not simply length-dependent, but requires some component of RNA sequence specificity. This data also demonstrates that the ACACCC "motif" is not sufficient for ZBP1 recognition, although a subset of its nucleotides may form the bindings site for one of the KH domains.

The previous experiments implicated nucleotides at both the 5′ and 3′ ends of zipcode[1–28] as potentially being recognized by ZBP1 KH34, but the identity of these nucleotides could not be determined. To better understand the sequence requirements for recognition, a selection for binding to ZBP1 KH34 was performed using a degenerate RNA library where, at each position within zipcode[1–28] sequence, the nucleotide's identity had the probability of being 85% wild-type and 5% of each non-wild-type nucleotide. After three rounds of selection and amplification, the RNA population was cloned and sequenced. This approach identified two stretches of nucleotides—5′-GGACU-3′ (4–8) and 5′-ACA-3′ (22–24)—that were highly conserved compared with the initial population of RNAs that retained binding to ZBP1 KH34 (Fig. 3A). The 3′ UTRs of human β-actin mRNA also share

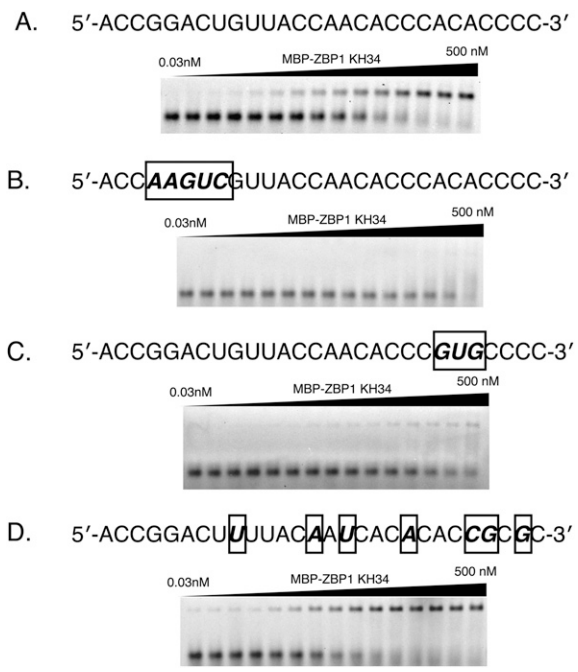

Figure 3. Selection for ZBP1 KH34 binding to doped library identifies two RNA elements required for recognition. (*A*) Sequence of zipcode[1–28] that was used to generate a degenerate RNA library that was designed to be 85% wild type. EMSA of ZBP1 KH34 binding to RNA pool that resulted from three rounds of selection. (*B*,*C*) Mutation of either conserved RNA element reduces affinity for ZBP1 KH34. (*D*) One clone isolated from the degenerate selection that harbors seven mutations binds as well as the wild-type zipcode[1–28] RNA. The filled triangle represents a 1:1 serial dilution of recombinant protein.

these sequence elements, but are spaced 2 nt further apart than in the chicken zipcode. Mutation of either one of these sequences resulted in a dramatic reduction in RNA binding (Fig. 3B,C). In comparison, one cloned RNA that harbors seven point mutations binds to ZBP1 KH34 as well as the wild-type zipcode[1–28] sequence (Fig. 3D). In this particular clone, the 5′ RNA element is not mutated and the 3′ RNA element has been shifted 2 nt upstream. While further biochemical characterization will be required to determine the precise nucleotides that form the consensus RNA-binding site of ZBP1 KH34, the results of this study support a model of recognition where the individual KH3 and KH4 domains each bind to unique nonsequential sequences within zipcode[1–28].

Overall structure of IMP1 KH34

The KH34 domain of IMP1, the human ortholog of ZBP1, shares 98% sequence identity (four amino acid substitutions: M407T, Q409H, I455V, and D533E) with ZBP1 KH34 and behaves identically in RNA-binding experiments using the zipcode[1–54], suggesting that this domain has been structurally and functionally conserved in vertebrates (Supplemental Fig. 3). Hexagonal crystals of IMP1 KH34 were obtained that diffract to 2.75 Å resolution, and the structure was determined by molecular replacement (Table 1). There are three molecules of IMP1

Table 1. *Crystallographic data and structure refinement*

	IMP1 KH34
Data collection	
Space group	$P6_5$
Cell dimensions	
a, *b*, *c*	103.5 Å, 103.5 Å, 131.6 Å
α, β, γ	90°, 90°, 120°
Resolution	53.0 Å–2.75 Å (2.90 Å–2.75 Å)
R_{merge}	0.064 (0.806)
I/σ I	21.9(2.9)
Completeness	99.9% (100.0%)
Multiplicity	9.7 (9.9)
Refinement	
Resolution	50.0 Å–2.75 Å
Number of reflections	20,305
R_{work}/R_{free}	19.8/23.9
Number of atoms	
Protein	3650
B factors	
Protein	97.5 Å^2
RMS deviations	
Bond lengths	0.012 Å
Bond angles	1.23°

Values in parentheses are for highest-resolution shell.

KH34 within the asymmetric unit and the independent copies are almost identical, having a root-mean-square deviation (RMSD) of 0.35 Å for all 157 C_α. Interestingly, two of the three IMP1 KH34 domains pack against symmetry-related molecules via interactions between their second and third α helices in a manner similar to what was previously observed for Nova KH3 (Lewis et al. 1999). These arrangements, however, are likely an artifact of crystallization and do not represent a stable dimerization, based on calculations of the free energy of disassociation (Krissinel and Henrick 2007). While we cannot exclude the possibility that such an interaction may form in the context of an RNA granule, the structure of IMP1 KH34 is consistent with our biochemical data showing that this domain recognizes RNA as a monomer.

In the structure of IMP1 KH34, both KH3 and KH4 adopt the $\beta_1\alpha_1\alpha_2\beta_2\beta_3\alpha_3$ topology seen in other type I eukaryotic KH domains. The KH3 and KH4 domains have the classical KH fold, with three α helices that pack against one face of anti-parallel β sheet formed by the three β strands (Fig. 4A). The conserved GXXG loop, which connects α_1 and α_2, is formed by residues 422–425 (GKKG) for KH3 and residues 504–507 (GKGG) for KH4, and the variable loop, located between β_2 and β_3, is formed by residues 442–450 for KH3 and residues 524–543 for KH4. A short linker (residues 479–487) connects α_3 of KH3 to β_1 of KH4. The structures of the KH3 and KH4 domains relative to each other are similar, with an RMSD of 1.0 Å for 69 C_α, with the largest differences located in the variable loops.

IMP1 KH34 forms an intramolecular anti-parallel pseudodimer

The KH3 and KH4 domains of IMP1 KH34 form an intramolecular pseudodimer (Fig. 4A). The KH domains are arranged in an anti-parallel manner that forms an extended six-stranded β-sheet surface. This interaction is stabilized by a number of residues located on the β_1 strands and α_3 helices of both KH domains, and also by residues located within the linker between KH3 and KH4 (Fig. 4B). Formation of the pseudodimer buries a large combined buried surface area (~2250 Å^2), suggesting that this arrangement is the functional conformation of IMP1 KH34 in vivo.

There are several large hydrophobic residues located within this intramolecular interface that likely contribute to its stability (Fig. 4B). Alternating residues located

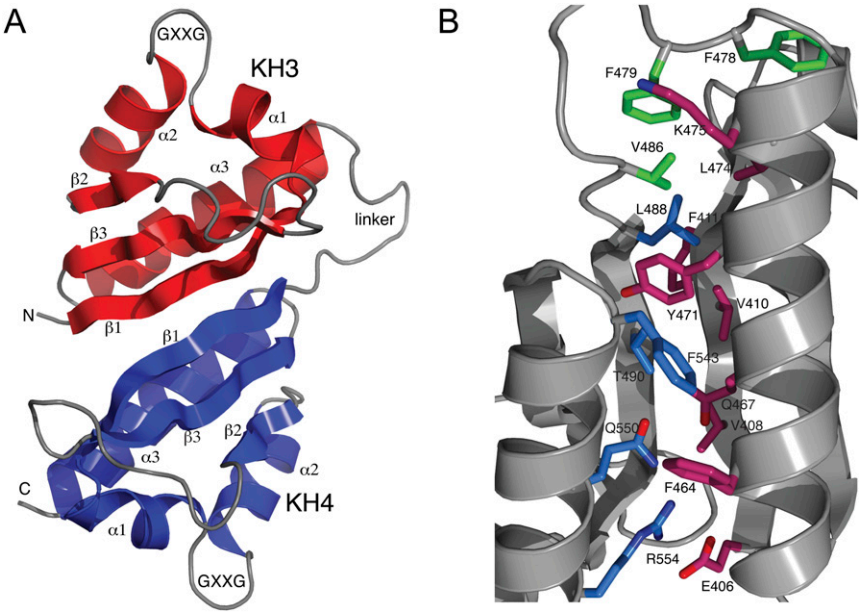

Figure 4. Structure of IMP1 KH34 pseudodimer. (*A*) KH3 (red) and KH4 (blue) with N and C termini and secondary structures labeled. The anti-parallel arrangement of KH3 and KH4 places the putative RNA-binding surfaces on opposite ends of the molecule. (*B*) Detailed view of intramolecular interface. The pseudodimer is stabilized by the packing of several hydrophobic residues as well as hydrogen bonds and electrostatic interactions.

on the α_3 helices of KH3 (Phe464 and Tyr471) and KH4 (F543) form the crux of the interaction that packs against a hydrophobic platform comprised of residues on the β_1 strands of KH3 (Val8 and Val10) and KH4 (L488 and Thr490). This hydrophobic surface is extended into the linker by Val486 and Phe479, as well as a second phenylalanine in the linker (Phe478) that packs against α1 of KH3.

While hydrophobic interactions likely dominate formation of the pseudodimer, hydrogen bonds and electrostatic interactions may also contribute to its stability. The anti-parallel arrangement of the KH3 and KH4 β_1 strands forms two interdomain hydrogen bonds between backbone amide and carbonyl groups (Gln409/His491 and Phe411/Glu489). Potential hydrogen bonds are also formed by two opposing glutamine residues on the α_3 helices of KH3 (Gln467) and KH4 (Gln550) and the ε amino group of Lys475 with the backbone carbonyl of Glu484. Favorable electrostatic interactions may also be mediated by the guanidino group of Arg554 with the side chain carboxylate group of Glu406.

Implications of IMP1 KH34 pseudodimer for RNA binding

The pseudodimer arrangement of IMP KH34 positions the putative RNA-binding surfaces of KH3 and KH4 at opposing ends of the molecule, with a distance of ~50 Å between the GXXG loops of KH3 and KH4 (Fig. 5A). Furthermore, the RNA-binding surfaces are oriented in opposite directions, which would require the RNA backbone to undergo almost a 180° change in direction in order to interact with both KH domains simultaneously as predicted by the biochemical data. Modeling of the KH34–RNA complex, based on structures of other KH domains bound to their nucleic acid targets, positions the 5′ and 3′ ends of the RNA~35 Å apart (Backe et al. 2005; Du et al. 2005, 2007). This distance could be spanned by a minimum of 5–6 nt; however, such a tight looping of the RNA is likely to be energetically unfavorable and would require the RNA elements recognized by the KH34 domain to be separated further.

To better determine the number of nucleotides that can separate the two halves of the bipartite RNA recognition element, nucleotides were systematically deleted from the zipcode[1–28] and their affinity for ZBP1 KH34 was measured by EMSA. Deletion of 2 and 4 nt did not have an appreciable effect on these RNAs' ability to interact with ZBP1 KH34 (Fig. 5B). However, when 6 nt were deleted from zipcode[1–28], the affinity of this RNA for ZBP1 KH34 was reduced by almost an order of magnitude, and deletion of 8 nt completely abolished RNA binding. This precipitous drop in RNA-binding affinity is consistent with the positioning of the KH domains in the structure of IMP1 KH34. While an upper limit for binding has not been determined, if the spacing becomes too large, the RNA elements will likely function independently, which will also lead to a reduction in apparent affinity.

Conservation of pseudodimer in VICKZ KH34 family members

The high degree of sequence identity (>80%) and conservation of hydrophobic residues at the intramolecular interface between IMP1 KH34 and its homologs suggests that all VICKZ family members adopt the same pseudodimer arrangement. A crystal structure of IMP2 KH34 confirms that the organization of the KH domains is conserved for this particular IMP1 paralog (J Chao, Y Patskovsky, S Almo, and R Singer, in prep.). Consequently, the bipartite RNA recognition element is also likely to be conserved between VICKZ family members, although it remains to be determined if the KH domains of individual family members have distinct RNA sequence preferences. Although amino acid differences between the homologs do not cluster around the putative RNA-binding surfaces, preliminary data show that the IMP1–3 paralogs do not recognize the zipcode[1–54] with the same affinity, suggesting that the family members are not functionally redundant (Supplemental Fig. 3).

Structural homology with other tandem KH domains

Despite the existence of numerous RNA-binding proteins with multiple KH domains, most of the structural data related to the function of these proteins have been gleaned from studies of isolated KH domains. There are only three structures of tandem type I KH domains that have been reported to date: far-upstream element-binding

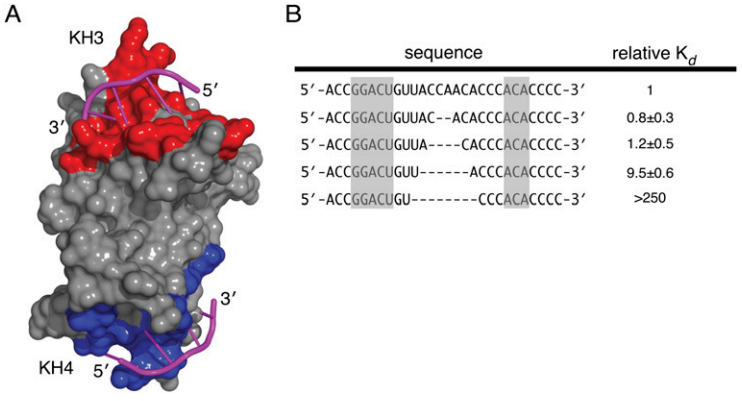

Figure 5. Model of RNA binding by IMP1 KH34 and spacing between bound RNA sequences. (*A*) The putative RNA-binding surfaces of KH3 (red) and KH4 (blue) are shown with modeled RNA tetranucleotide (magenta). Anti-parallel arrangement of KH domains requires RNA backbone to undergo an ~180° change in direction to interact with both domains at the same time, which induces looping of the RNA. (*B*) Systematic deletion of nucleotides within zipcode[1–28]. Changes in RNA-binding affinity are shown relative to zipcode[1–28].

protein (FBP) KH34 in complex with ssDNA, fragile X mental retardation protein (FMRP) KH12, and poly-C-binding protein 2 (PCBP2) KH12 (Braddock et al. 2002; Valverde et al. 2007; Du et al. 2008). In the case of FBP KH34, nuclear magnetic resonance (NMR) experiments found that the two KH domains, which are separated by a 30-residue glycine-rich linker, function independently (Braddock et al. 2002). A similar situation was observed by NMR for the third and fourth KH domains of K-homology splicing protein (KSRP), although a full structure determination of the tandem domains was not performed (Garcia-Mayoral et al. 2007). In the X-ray crystal structure of FMRP KH12, which has a single residue linking KH1 to KH2, the KH domains are positioned at ~60° relative to one another, resulting in a limited intramolecular interaction surface between the two KH domains (Valverde et al. 2007). PCBP2 KH12, however, has a 14-amino-acid linker that connects its two KH domains and adopts an anti-parallel pseudodimer topology similar to IMP1 KH34 (Du et al. 2008).

IMP1 KH34 and PCBP2 KH12 share only 29% sequence identity, yet adopt similar overall structures that have an RMSD of 3.1 Å for 149 C_α (Fig. 6A). Interestingly, only one aromatic residue (Phe464-IMP1, Phe69-PCBP2) is conserved between IMP1 KH34 and PCBP2 KH12 at the pseudodimer interface, suggesting that this arrangement can be stabilized by multiple complementary interactions. Also, few residues involved in RNA binding are conserved between IMP1 KH34 and PCBP2 KH12, and only Arg57 that makes a cytosine-specific contact to the second nucleotide in the PCBP2 KH1–RNA complex is conserved in IMP1 KH3 (Arg452) (Du et al. 2007). Interestingly, this base-specific readout by an arginine residue at this position is also found in the complexes between hnRNP K KH3 and Nova-2 KH3 with their nucleic acid-binding partners (Lewis et al. 2000; Backe et al. 2005). This suggests that IMP1 KH3 may recognize the 3′ RNA element of zipcode[1–28], but the lack of sequence conservation makes homology modeling difficult, and it is unclear if this particular interaction is maintained between these structures.

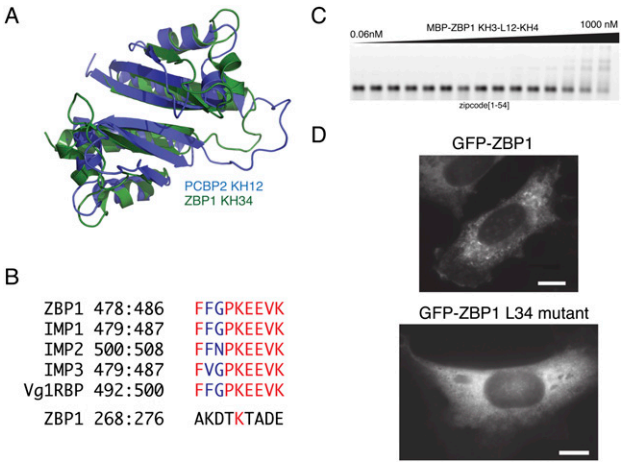

Figure 6. Conserved residues that link KH3 and KH4 function in RNA binding and granule formation. (*A*) IMP1 KH34 and PCBP2 KH12 adopt similar structures, with the largest deviation coming from the residues that link the KH domains. (*B*) Sequence alignment of amino acids that connect KH3 and KH4 from ZBP1, IMP1-3, and VgRBP1 shown with ZBP1 KH12 linker. (*C*) Mutation of KH34 linker to residues in KH12 linker reduces RNA binding to zipcode[1–54]. The filled triangle represents a 1:1 serial dilution of recombinant protein. (*D*) Expression of GFP-ZBP1 with mutant KH34 linker results in a change in subcellular localization of the fusion protein from granular to diffuse cytoplasmic.

Role of conserved linker between KH3 and KH4

While IMP1 KH34 and PCBP2 KH12 share a similar organization of their respective KH domains, the linkers that connect the domains adopt very different conformations (Fig. 6A). In the PCBP2 KH12 structure, the linker contains a number of polar residues and its structure is poorly defined due to lack of experimental NMR restraints, indicating that this region is more dynamic than the rest of the molecule (Du et al. 2008). In contrast, the linker in the IMP1 KH34 structure forms a compact structure that helps to stabilize the pseudodimer interface by the van der Waals interactions made by the two large aromatic residues (Phe479 and Phe480). The linker residues are highly conserved in the VICKZ family of RNA-binding proteins, which has led previously to the suggestion that this region may be functionally important (Fig. 6B; Git and Standart 2002).

To determine if this linker has a functional role, a mutant ZBP1 KH34 protein was generated that replaced the linker residues with the amino acids that connect KH1 and KH2 within ZBP1 (Fig. 6B). This mutation only changes the identity of the residues in the linker, but not its length. The RNA-binding ability of this ZBP1 KH3–L12–KH4 mutant protein was measured by EMSA using the zipcode[1–54] RNA. ZBP1 KH3–L12–KH4 mutant binds to the zipcode[1–54] considerably weaker ($K_{d,\ app}$ ~ 500 nM) than ZBP1 KH34 ($K_{d,\ app}$ = 3.6 nM) (Figs. 2A, 6C). This reduction in RNA-binding affinity could arise from a destabilization of the KH34 pseudodimer, indicating that the precise orientation of the domains is required for RNA recognition, or residues in the linker could make direct interactions with the zipcode that were lost in the ZBP1 KH3–L12–KH4 mutant.

In order to see if this linker mutation could affect the physiology of ZBP1 in cells, the mutation was generated within the full-length protein and fused to the C terminus of green fluorescent protein (GFP). Wild-type ZBP1, when fused to GFP, was found to form large cytoplasmic granules when expressed in mouse embryonic fibroblasts (MEFs), similar to what was observed in previous studies (Fig. 6D; Nielsen et al. 2002; Farina et al. 2003). The GFP-ZBP1 L34 linker mutant protein, however, exhibits a diffuse cytoplasmic appearance and reduction of granule formation, indicating that the impaired RNA-binding affinity observed in vitro alters the subcellular localization of ZBP1 in vivo (Fig. 6D). While the precise function this linker has remains to be determined, the data presented here show that residues located outside of the

canonical KH domain RNA-binding surface can affect RNA recognition, and that the ZBP1 KH34 domains acts as a single functional unit.

Discussion

The architecture of the KH domain only allows for the recognition of short stretches of nucleic acid with relatively weak binding affinity. Therefore, many nucleic acid-binding proteins contain multiple copies of these domains in order to increase both affinity and specificity. Furthermore, these domains can be arranged in precise orientations that can provide additional restrictions on their interacting partners and, also, insight into the mechanism by which these proteins accomplish their biological function.

The third and fourth KH domains of ZBP1/IMP1 function as a single module by adopting an anti-parallel pseudodimer arrangement that places the putative RNA-binding surfaces on opposite ends of the molecule. This structural organization of the KH domains allows the protein to recognize its targets through sequence-specific contacts distributed over two stretches of RNA that must be separated by a certain number of nucleotides. By fashioning the KH3 and KH4 domains together in this orientation, highly specific RNA binding can be achieved from two low-affinity interactions.

While the precise identity of the nucleotides that comprise the consensus bipartite RNA-binding site remains to be determined, a model describing the interaction between KH34 and its target RNA can be proposed. Due to the ambiguity of the orientation of the KH domains to the RNA, two distinct models can be constructed that differ in the position of the extruded nucleotide loop, but maintain the same polarity of the RNA to the KH domains (Fig. 7A,B). Currently, it is unclear which of the models is correct, and whether or not residues outside of the canonical RNA-binding surfaces can participate in discriminating between the two modes of binding. If both models are possible, then, in principle, the position of the two RNA elements within the transcript can also be swapped. Further experiments will be required to clarify these features of the RNA–protein interface, and whether or not the other domains of ZBP1 can influence proper RNA target recognition.

A consequence, however, of both models is that binding of KH34 to the RNA will induce looping of the transcript. This conformation of the RNA brings sequences that are distant from one another in the transcript into closer proximity, which may create the RNA-binding sites for additional factors (Fig. 7C). In this way, the VICKZ family of RNA-binding proteins may nucleate the assembly of higher-order complexes without directly interacting with other proteins. This result may help to explain the lack of direct protein–protein partners for IMP1 identified by immunopurification and mass spectrometry characterization of IMP1-containing granules (Jonson et al. 2007; Weidensdorfer et al. 2009). RNA looping has also been proposed to occur when polypyrimidine tract-binding protein 1 (PTB1) binds to the

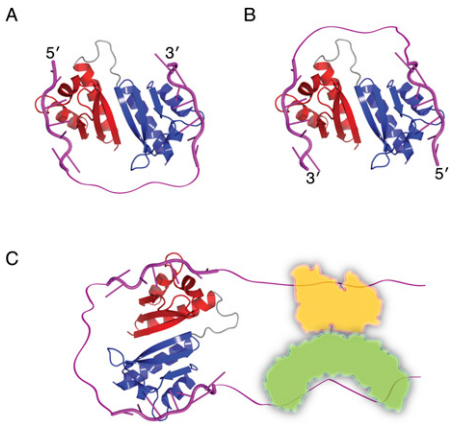

Figure 7. Model for RNA looping induced by binding of the ZBP1 KH34 domain and the assembly of RNA–protein complexes. (*A*,*B*) Current data allows two distinct modes of RNA binding to be modeled based on which sequences within the bipartite RNA element the individual KH domains interact with. Polarity of the RNA to the KH domains is conserved in both models. (*C*) Looping of the transcript may form the binding sites for additional RNA-binding proteins and nucleate the assembly of larger RNPs.

γ-aminobutyric acid–γ2 pre-mRNA, and is thought to contribute to PTB1's function in repressing the splicing of alternative exons (Oberstrass et al. 2005). We propose that this feature may be a hallmark of RNA-binding proteins that play a primary role in the regulation of RNA–protein complexes by remodeling the transcript so that additional factors can assemble.

Materials and methods

Protein preparation

Full-length ZBP1 and truncations (RRM12 1–157, KH12 195–354, KH34 404–561, KH3 404–478, KH4 487–576, and KH3–L12–KH4 404–561 with 478–486 replaced by 268–276) were cloned by PCR into a derivative of pMalc (New England BioLabs) that contains a Tobacco Etch virus (TEV) protease site after the maltose-binding protein (MBP). A C-terminal His_6 tag was added by PCR to both ZBP1 full-length and KH34 constructs to ensure purification of the intact fusion proteins. These constructs were transformed into *Escherichia coli* strain Rosetta2 (EMD Biosciences), and recombinant protein was induced with 1 mM IPTG for 4 h at 37°C. Cell pellets were resuspended in lysis buffer (50 mM Tris at pH 7.5, 1.5 M NaCl, 1 mM EDTA, 1 mM DTT) supplemented with one Complete EDTA-free protease inhibitor tablet (Roche), and were lysed by sonication. Cell debris was removed by centrifugation, and the soluble fusion protein was purified by amylose affinity chromatography (New England BioLabs) followed by either TALON affinity (Clontech) or anion exchange (GE Healthcare) chromatography. Protein concentrations were calculated by measuring the absorbance at 280 nm and using extinction coefficients determined by ProtParam (Gasteiger et al. 2005).

IMP1 KH34 (404–566) was cloned by PCR into pET22HT as an N-terminal His_6 tag followed by a TEV cleavage site (Chao et al. 2005). This construct was expressed and purified using conditions identical to those described for MBP–ZBP1 except that His_6-IMP1 KH34 was lysed in a buffer compatible with IMAC

(50 mM sodium phosphate, 1.5 M NaCl, 10 mM imidazole). TALON affinity chromatography was used as the first step of purification followed by TEV cleavage. A second TALON affinity chromatography step was performed to remove His_6-TEV protease and uncleaved recombinant His_6-IMP1 KH34.

EMSA

The complexes between recombinant ZBP1 constructs and zipcode RNA fragments were monitored by an EMSA. Zipcode RNA fragments with fluorescein modifications were prepared by chemical synthesis (Dharmacon) and deprotected, lyophilized, and stored according to the manufacturer's protocol. The sequences of the RNAs used in these experimentswere 1–54, 5′-ACCGGACUGUUACCAACACCCACACCCCUGUGAUGAAACAAAACCCAUAAAUGC-Fl-3′; 1–28, 5′-Fl-ACCGGACUGUUACCAACACCCACACCCC-3′; 29–54, 5′-Fl-UGUGAUGAAACAAAACCCAUAAAUGC-3′; 1–21, 5′-Fl-ACCGGACUGUUACCAACACCC-3′; and 5–44, 5′-Fl-GACUGUUACCAACACCCACACCCCUGUGAUGAAACAAAAC-3′. RNAs (100 pM) were equilibrated with a twofold serial dilution of recombinant proteins in a buffer containing 10 mM Tris (pH 7.5), 100 mM NaCl, 0.1 mM EDTA, 0.01 mg mL^{-1} tRNA, 50 μg mL^{-1} heparin, and 0.01% IGEPAL CA630 for approximately 3 h in order to ensure the binding reaction reached equilibrium. Protein–RNA complexes were resolved from unbound RNA by native polyacrylamide gel electrophoresis (5% [w/v] 29:1 acrylamide/Bis-acrylamide, 0.5× TBE) run at 80 V at 4°C. Gels were scanned using a fluorescent gel imager (Typhoon 9400, GE Healthcare) with 488 nm excitation and 520 BP 40 filter. The fraction of bound RNA was determined by either the disappearance of the free RNA or a ratio of the free and bound RNA, and then fit to a modified version of the Hill equation as described previously to determine the apparent dissociation constant (Ryder et al. 2008).

Stoichiometry experiments were performed similarly as described above except that the binding reaction contained unlabeled zipcode[1–28] RNA at 250 nM, and MBP–ZBP1 KH34 was serial-diluted 3:1 in order to obtain more data points in the transition from free RNA to bound. The data were fit to a quadratic model of saturable ligand binding resulting in determination of the stoichiometric equivalence point (Rambo and Doudna 2004).

Analytical size exclusion chromatography

The apparent molecular weights of ZBP1, ZBP1 KH34, and the ZBP1 KH34–zipcode[1–28] complex were determined using a HiLoad 16/60 Superdex 200 column (GE Healthcare) gel filtration column. Recombinant protein samples (~20 μM) were loaded onto a column equilibrated in a buffer of 50 mM Tris (pH 7.5), 300 mM NaCl, and 1 mM DTT, and run at 1 mL min^{-1}. The ZBP1 KH34–zipcode[1–28] complex (~100 μM) was loaded onto a column equilibrated in a buffer containing 50 mM Tris (pH 7.2) and 150 mM NaCl and run at 1 mL min^{-1}. The A280 absorbance was monitored to determine the retention time of the protein, and the apparent molecular weight was estimated by comparison with known protein standards: γ-globulin (158 kDa), ovalbumin (44 kDa), myoglobin (17 kDa), and vitamin B_{12} (1.35 kDa) (Bio-Rad Laboratories).

RNA selection using degenerate zipcode[1–28] library

An antisense degenerate zipcode[1–28] library with the sequence 5′-AAGCTTCGTCAAGTCTGCAGTGAA*GGGGTGTGGGTGTTGGTAACAGTCCGGT*TCGTAGATGTGGATCCATTCCC-3′ was prepared by chemical synthesis (IDT). Italicized nucleotides indicate positions that were doped at 85% wild-type and 5% of each non-wild-type nucleotide. The initial DNA library was gel-purified by denaturing polyacrylamide gel electrophoresis (8% [w/v] 29:1 acrylamide/Bis-acrylamide, 0.5× TBE), and was eluted by crush and soak in 300 mM NaCl. The library (200 pmol) was converted to dsDNA by reverse transcription (Invitrogen) using the 41.30 primer (5′-GATAATACGACTCACTATAGGGAATGGATCCACATCTACGA-3′). The RNA pool used for selection was generated by T7 transcription (Ambion) of the DNA library, gel-purified, and resuspended in TE. In the first round of selection, MBP–ZBP1 KH34 (~10 nM) was bound to amylose resin and then equilibrated with the RNA pool (~75 nM) for 1 h in a buffer containing 10 mM Tris (pH 7.5), 200 mM NaCl, 0.1 mM EDTA, and 0.01 mg mL^{-1} tRNA. The binding reaction was transferred to a 0.45-μm centrifugal filter tube (Millipore) and washed extensively (10 mM Tris at pH 7.5, 200 mM NaCl) before elution with wash buffer supplemented with 10 mM maltose. After phenol/chloroform extraction, the RNA was converted to cDNA by reverse transcription using the 24.30 primer (5′-AAGCTTCGTCAAGTCTGCAGTGAA-3′). The resulting cDNA library was amplified by PCR using the 41.30 and 24.30 primers and then transcribed into RNA for the next round of selection. Subsequent rounds of selection were performed similarly, except that a negative selection step was included by incubating the RNA pool with amylose resin in the absence of protein to remove any RNAs with nonspecific affinity for the amylose resin. RNA from the third round of selection was fluorescein-labeled, and its affinity for MBP–ZBP1 KH34 was quantified by EMSA (Pagano et al. 2007).

Cell culture, transfection, and microscopy

GFP-ZBP1 with the mutant L12 linker between KH3 and KH4 was generated by PCR by ligating overlapping fragments containing the mutation. This GFP-ZBP1 L34 mutant and GFP-ZBP1 wild type were transfected into MEFs that were seeded onto coverslips using FuGENE 6 (Roche). After transfection, the MEFs were incubated at 37°C in Dulbecco's modified Eagle's medium (DMEM; HyClone) supplemented with 10% fetal bovine serum (FBS) for 24 h prior to washing with PBS and fixation using 4% paraformaldehyde (PFA) in PBS for 20 min. Cells were washed twice with PBS prior to mounting onto glass slides using ProLong Gold Anti-fade reagent (Invitrogen). Cells were visualized on a BX61 microscope (Olympus) with an UPlanApo 100×, 1.35NA objective (Olympus) coupled to an X-Cite 120 PC metal-halide light source (EXFO Life Science) and filter sets for GFP (Chroma Technology). Digital images were collected with a CoolSNAP HQ camera (Roper Scientific) using IPLab (Windows version 4.0, BD Biosciences).

Crystallization and structure determination of IMP1 KH34

IMP1 KH34 (0.5 mM) was crystallized using sitting-drop vapor diffusion at 22°C by mixing equal volumes of the protein and reservoir solution (1.65 M ammonium citrate, 100 mM Tris at pH 7.8). Crystals were cryoprotected by soaking in reservoir solution supplemented with 25% glycerol before flash-cooling in liquid nitrogen. Data were collected to 2.75 Å resolution from a single crystal at the National Synchotron Light Source X29a beamline at a wavelength of 0.979 Å. The diffraction data were indexed, integrated, and scaled using MOSFLM and the CCP4 suite of programs (Collaborative Computational Project, Number 4 1994).

The structure of IMP1 KH34 was determined by molecular replacement with Phaser using the structure of IMP2 KH34 (J Chao, Y Patskovsky, S Almo, and R Singer, in prep.) as a search

model (Read 2001). The crystal contained three independent copies of IMP1 KH34 in the asymmetric unit. Rounds of refinement and model building were carried out with Phenix and Coot (Adams et al. 2002; Emsley and Cowtan 2004). Noncrystallographic symmetry (NCS) restraints were applied during all stages of refinement, and TLS-refinement was performed after initial rounds of refinement with domains corresponding to KH3 and KH4. Protein stereochemistry was checked using Molprobity (Davis et al. 2007). The final model contained residues 405–562 for chain A, 405–481; 484–562 for chain B and 405–482; 484–562 for chain C. The side chain of K450 of all three IMP1 KH34 copies was truncated to the C_β due to lack of electron density.

Accession codes

Protein Data Bank: Coordinates and structure factors were deposited with accession codes 3KRM (IMP1 KH34).

Acknowledgments

We thank the staff at the NSLS X29a beamline for assistance with data collection, and G. Arenas, U. Meier, S. Nguyen, D. Rueda, and S. Ryder for helpful discussions. This work was supported by the U.S. National Institutes of Health (GM084364 to R.H.S.) and NRSA individual fellowship support (F32GM083430 to J.A.C.) and the Albert Einstein Cancer Center.

References

Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger TR, McCoy AJ, Moriarty NW, Read RJ, Sacchettini JC, Sauter NK, Terwilliger TC. 2002. PHENIX: Building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* **58:** 1948–1954.

Atlas R, Behar L, Elliott E, Ginzburg I. 2004. The insulin-like growth factor mRNA binding-protein IMP-1 and the Ras-regulatory protein G3BP associate with tau mRNA and HuD protein in differentiated P19 neuronal cells. *J Neurochem* **89:** 613–626.

Atlas R, Behar L, Sapoznik S, Ginzburg I. 2007. Dynamic association with polysomes during P19 neuronal differentiation and an untranslated-region-dependent translation regulation of the tau mRNA by the tau mRNA-associated proteins IMP1, HuD, and G3BP1. *J Neurosci Res* **85:** 173–183.

Auweter SD, Oberstrass FC, Allain FH. 2006. Sequence-specific binding of single-stranded RNA: Is there a code for recognition? *Nucleic Acids Res* **34:** 4943–4959.

Backe PH, Messias AC, Ravelli RB, Sattler M, Cusack S. 2005. 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.

Braddock DT, Louis JM, Baber JL, Levens D, Clore GM. 2002. Structure and dynamics of KH domains from FBP bound to single-stranded DNA. *Nature* **415:** 1051–1056.

Chao JA, Lee JH, Chapados BR, Debler EW, Schneemann A, Williamson JR. 2005. Dual modes of RNA-silencing suppression by Flock House virus protein B2. *Nat Struct Mol Biol* **12:** 952–957.

Collaborative Computational Project, Number 4. 1994. The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50:** 760–763.

Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB III, Snoeyink J, Richardson JS, et al. 2007. MolProbity: All-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* **35:** W375–W383. doi: 10.1093/nar/gkm216.

Deshler JO, Highett MI, Abramson T, Schnapp BJ. 1998. A highly conserved RNA-binding protein for cytoplasmic mRNA localization in vertebrates. *Curr Biol* **8:** 489–496.

Dimitriadis E, Trangas T, Milatos S, Foukas PG, Gioulbasanis I, Courtis N, Nielsen FC, Pandis N, Dafni U, Bardi G, et al. 2007. Expression of oncofetal RNA-binding protein CRD-BP/IMP1 predicts clinical outcome in colon cancer. *Int J Cancer* **121:** 486–494.

Doyle GA, Betz NA, Leeds PF, Fleisig AJ, Prokipcak RD, Ross J. 1998. The c-myc coding region determinant-binding protein: A member of a family of KH domain RNA-binding proteins. *Nucleic Acids Res* **26:** 5036–5044.

Du Z, Lee JK, Tjhen R, Li S, Pan H, Stroud RM, James TL. 2005. Crystal structure of the first KH domain of human poly(C)-binding protein-2 in complex with a C-rich strand of human telomeric DNA at 1.7 A. *J Biol Chem* **280:** 38823–38830.

Du Z, Lee JK, Fenn S, Tjhen R, Stroud RM, James TL. 2007. X-ray crystallographic and NMR studies of protein–protein and protein–nucleic acid interactions involving the KH domains from human poly(C)-binding protein-2. *RNA* **13:** 1043–1051.

Du Z, Fenn S, Tjhen R, James TL. 2008. Structure of a construct of a human poly(C)-binding protein containing the first and second KH domains reveals insights into its regulatory mechanisms. *J Biol Chem* **283:** 28757–28766.

Elcheva I, Goswami S, Noubissi FK, Spiegelman VS. 2009. CRD-BP protects the coding region of βTrCP1 mRNA from miR-183-mediated degradation. *Mol Cell* **35:** 240–246.

Emsley P, Cowtan K. 2004. Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60:** 2126–2132.

Farina KL, Huttelmaier S, Musunuru K, Darnell R, Singer RH. 2003. Two ZBP1 KH domains facilitate β-actin mRNA localization, granule formation, and cytoskeletal attachment. *J Cell Biol* **160:** 77–87.

Fusco D, Accornero N, Lavoie B, Shenoy SM, Blanchard JM, Singer RH, Bertrand E. 2003. Single mRNA molecules demonstrate probabilistic movement in living mammalian cells. *Curr Biol* **13:** 161–167.

Garcia-Mayoral MF, Hollingworth D, Masino L, Diaz-Moreno I, Kelly G, Gherzi R, Chou CF, Chen CY, Ramos A. 2007. The structure of the C-terminal KH domains of KSRP reveals a noncanonical motif important for mRNA degradation. *Structure* **15:** 485–498.

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. 2005. Protein identification and analysis tools on the ExPASy server. In *The proteomics protocols handbook* (ed. JM Walker), pp. 571–607. Humana Press, Totawa, NJ.

Git A, Standart N. 2002. The KH domains of Xenopus Vg1RBP mediate RNA binding and self-association. *RNA* **8:** 1319–1333.

Hammer NA, Hansen TO, Byskov AG, Rajpert-De Meyts E, Grondahl ML, Bredkjaer HE, Wewer UM, Christiansen J, Nielsen FC. 2005. Expression of IGF-II mRNA-binding proteins (IMPs) in gonads and testicular cancer. *Reproduction* **130:** 203–212.

Havin L, Git A, Elisha Z, Oberman F, Yaniv K, Schwartz SP, Standart N, Yisraeli JK. 1998. RNA-binding protein conserved in both microtubule- and microfilament-based RNA localization. *Genes & Dev* **12:** 1593–1598.

Jiang Z, Lohse CM, Chu PG, Wu CL, Woda BA, Rock KL, Kwon ED. 2008. Oncofetal protein IMP3: A novel molecular marker that predicts metastasis of papillary and chromophobe renal cell carcinomas. *Cancer* **112:** 2676–2682.

Jonson L, Vikesaa J, Krogh A, Nielsen LK, Hansen T, Borup R, Johnsen AH, Christiansen J, Nielsen FC. 2007. Molecular

composition of IMP1 ribonucleoprotein granules. *Mol Cell Proteomics* **6:** 798–811.
Kislauskis EH, Zhu X, Singer RH. 1994. Sequences responsible for intracellular localization of β-actin messenger RNA also affect cell phenotype. *J Cell Biol* **127:** 441–451.
Kobel M, Xu H, Bourne PA, Spaulding BO, Shih Ie M, Mao TL, Soslow RA, Ewanowich CA, Kalloger SE, Mehl E, et al. 2009. IGF2BP3 (IMP3) expression is a marker of unfavorable prognosis in ovarian carcinoma of clear cell subtype. *Mod Pathol* **22:** 469–475.
Krissinel E, Henrick K. 2007. Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372:** 774–797.
Latham VM Jr, Kislauskis EH, Singer RH, Ross AF. 1994. β-Actin mRNA localization is regulated by signal transduction mechanisms. *J Cell Biol* **126:** 1211–1219.
Lawrence JB, Singer RH. 1986. Intracellular localization of messenger RNAs for cytoskeletal proteins. *Cell* **45:** 407–415.
Lecuyer E, Yoshida H, Parthasarathy N, Alm C, Babak T, Cerovina T, Hughes TR, Tomancak P, Krause HM. 2007. Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. *Cell* **131:** 174–187.
Leung KM, van Horck FP, Lin AC, Allison R, Standart N, Holt CE. 2006. Asymmetrical β-actin mRNA translation in growth cones mediates attractive turning to netrin-1. *Nat Neurosci* **9:** 1247–1256.
Lewis HA, Chen H, Edo C, Buckanovich RJ, Yang YY, Musunuru K, Zhong R, Darnell RB, Burley SK. 1999. Crystal structures of Nova-1 and Nova-2 K-homology RNA-binding domains. *Structure* **7:** 191–203.
Lewis HA, Musunuru K, Jensen KB, Edo C, Chen H, Darnell RB, Burley SK. 2000. Sequence-specific RNA binding by a Nova KH domain: Implications for paraneoplastic disease and the fragile X syndrome. *Cell* **100:** 323–332.
Martin KC, Ephrussi A. 2009. mRNA localization: Gene expression in the spatial dimension. *Cell* **136:** 719–730.
Mili S, Moissoglu K, Macara IG. 2008. Genome-wide screen reveals APC-associated RNAs enriched in cell protrusions. *Nature* **453:** 115–119.
Nielsen J, Christiansen J, Lykke-Andersen J, Johnsen AH, Wewer UM, Nielsen FC. 1999. A family of insulin-like growth factor II mRNA-binding proteins represses translation in late development. *Mol Cell Biol* **19:** 1262–1270.
Nielsen FC, Nielsen J, Kristensen MA, Koch G, Christiansen J. 2002. Cytoplasmic trafficking of IGF-II mRNA-binding protein by conserved KH domains. *J Cell Sci* **115:** 2087–2097.
Nielsen J, Kristensen MA, Willemoes M, Nielsen FC, Christiansen J. 2004. Sequential dimerization of human zipcode-binding protein IMP1 on RNA: A cooperative mechanism providing RNP stability. *Nucleic Acids Res* **32:** 4368–4376.
Oberman F, Rand K, Maizels Y, Rubinstein AM, Yisraeli JK. 2007. VICKZ proteins mediate cell migration via their RNA binding activity. *RNA* **13:** 1558–1569.
Oberstrass FC, Auweter SD, Erat M, Hargous Y, Henning A, Wenter P, Reymond L, Amir-Ahmady B, Pitsch S, Black DL, et al. 2005. Structure of PTB bound to RNA: Specific binding and implications for splicing regulation. *Science* **309:** 2054–2057.
Oleynikov Y, Singer RH. 2003. Real-time visualization of ZBP1 association with β-actin mRNA during transcription and localization. *Curr Biol* **13:** 199–207.
Pagano JM, Farley BM, McCoig LM, Ryder SP. 2007. Molecular basis of RNA recognition by the embryonic polarity determinant MEX-5. *J Biol Chem* **282:** 8883–8894.
Patel GP, Bag J. 2006. IMP1 interacts with poly(A)-binding protein (PABP) and the autoregulatory translational control element of PABP-mRNA through the KH III-IV domain. *FEBS J* **273:** 5678–5690.
Rambo RP, Doudna JA. 2004. Assembly of an active group II intron-maturase complex by protein dimerization. *Biochemistry* **43:** 6486–6497.
Read RJ. 2001. Pushing the boundaries of molecular replacement with maximum likelihood. *Acta Crystallogr D Biol Crystallogr* **57:** 1373–1382.
Ross AF, Oleynikov Y, Kislauskis EH, Taneja KL, Singer RH. 1997. Characterization of a β-actin mRNA zipcode-binding protein. *Mol Cell Biol* **17:** 2158–2165.
Runge S, Nielsen FC, Nielsen J, Lykke-Andersen J, Wewer UM, Christiansen J. 2000. H19 RNA binds four molecules of insulin-like growth factor II mRNA-binding protein. *J Biol Chem* **275:** 29562–29569.
Ryder SP, Recht MI, Williamson JR. 2008. Quantitative analysis of protein-RNA interactions by gel mobility shift. *Methods Mol Biol* **488:** 99–115.
Valverde R, Pozdnyakova I, Kajander T, Venkatraman J, Regan L. 2007. Fragile X mental retardation syndrome: Structure of the KH1–KH2 domains of fragile X mental retardation protein. *Structure* **15:** 1090–1098.
Valverde R, Edwards L, Regan L. 2008. Structure and function of KH domains. *FEBS J* **275:** 2712–2726.
Vikesaa J, Hansen TV, Jonson L, Borup R, Wewer UM, Christiansen J, Nielsen FC. 2006. RNA-binding IMPs promote cell adhesion and invadopodia formation. *EMBO J* **25:** 1456–1468.
Weidensdorfer D, Stohr N, Baude A, Lederer M, Kohn M, Schierhorn A, Buchmeier S, Wahle E, Huttelmaier S. 2009. Control of c-myc mRNA stability by IGF2BP1-associated cytoplasmic RNPs. *RNA* **15:** 104–115.
Yao J, Sasaki Y, Wen Z, Bassell GJ, Zheng JQ. 2006. An essential role for β-actin mRNA localization and translation in Ca^{2+}-dependent growth cone guidance. *Nat Neurosci* **9:** 1265–1273.
Yisraeli JK. 2005. VICKZ proteins: A multi-talented family of regulatory RNA-binding proteins. *Biol Cell* **97:** 87–96.