Published in final edited form as:

J Mol Biol. 2010 September 17; 402(2): 412–427. doi:10.1016/j.jmb.2010.07.040.

Structure of the RNA binding domain of a DEAD-box helicase bound to its ribosomal RNA target reveals a novel mode of recognition by an RNA recognition motif

John W. Hardin¹, YaoXiong Hu³, and David B. McKay^{1,2,*}

¹Department of Chemistry and Biochemistry, University of Colorado, Boulder, CO 80301, USA

²Howard Hughes Medical Institute, University of Colorado, Boulder, CO 80301, USA

³Department of Structural Biology, Stanford University School of Medicine, Stanford, CA, 94305, USA

Abstract

DEAD-box RNA helicases of the bacterial DbpA subfamily are localized to their biological substrate when a carboxy-terminal RNA recognition motif (RRM) domain binds tightly and specifically to a segment of 23S ribosomal RNA (rRNA) that includes hairpin 92 of the peptidyl transferase center. A complex between a fragment of 23S rRNA and the RNA binding domain (RBD) of the *Bacillus subtilis* DbpA protein YxiN was crystallized and its structure determined to 2.9 Å resolution, revealing an RNA recognition mode that differs from those observed with other RRM. The RBD is bound between two RNA strands at a three-way junction. Multiple phosphates of the RNA backbone interact with an electropositive band generated by lysines of the RBD. Nucleotides of the single-stranded loop of hairpin 92 interact with the RBD, including the guanosine base of G2553, which forms three hydrogen bonds with the peptide backbone. A G2553U mutation reduces the RNA binding affinity by two orders of magnitude, confirming that G2553 is a sequence specificity determinant in RNA binding. Binding of the RBD to 23S rRNA in the late stages of ribosome subunit maturation would position the ATP-binding duplex destabilization fragment of the protein for interaction with rRNA in the peptidyl transferase cleft of the subunit, allowing it to “melt out” unstable secondary structures and allow proper folding.

Keywords

DEAD-box RNA helicase; RNA recognition motif; RNA binding; ribosomal RNA folding; x-ray crystallography

© 2010 Elsevier Ltd. All rights reserved.

*Corresponding author at: Department of Chemistry and Biochemistry, Mail Code 215, University of Colorado, Boulder, CO 80301-0215, USA, tel. 1-303-492-0970 fax 1-303-492-6194, David.McKay@Colorado.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Protein Data Bank (PDB) accession code

Coordinates and structure factors have been deposited with the Research Collaboratory for Structural Biology (www.rcsb.org), identifier 3MOJ.

Introduction

Proteins of the DEx(D/H)-box RNA helicase family participate in a diversity of activities involving rearrangement of RNA structure and modulation of protein-RNA interactions^{1, 2}. The proteins all have in common a ~400 amino acid residue “catalytic core” consisting of two RecA-related structural domains, one of which binds ATP, and both of which participate in binding and destabilizing RNA duplexes in an ATPase-dependent reaction³⁻⁵. Beyond the catalytic core, many of the proteins have additional domains, or interact with accessory proteins, which are thought to confer specificity in biochemical target and biological function to the catalytic fragments.

Within the DEx(D/H)-box helicase family, the *E. coli* DbpA (DEAD-box protein A), participates in biogenesis of the large ribosomal subunit; it has been shown that overexpression of a mutant DbpA bereft of ATPase activity confers a dominant negative phenotype and produces immature large subunit complexes that are stalled in a late stage of assembly^{6, 7}. It has been known for some time that DbpA binds 23S ribosomal RNA (rRNA) in the vicinity of hairpin 92, which lies in the peptidyl transferase site in the mature ribosome⁸⁻¹¹. Full-length DbpA protein binds a series of rRNA fragments that include stems H90, H91 and H92 with 0.03-1.0 nM affinity in the presence of AMPPNP, and 0.3-1.0 nM in absence of AMPPNP. When one strand of H90 is deleted, making it single-stranded, the affinity decreases to ~10 nM in absence of AMPPNP¹¹. The ~10 nM RNA binding affinity of DbpA can be accomplished with short oligonucleotides that include hairpin 92 plus a short single-strand extension¹⁰.

The primary structure of proteins in the bacterial DbpA helicase subfamily consists of the minimal catalytic core plus a small, subfamily-specific 70- to 80-residue carboxy-terminal domain that is tethered to the helicase core by a peptide linker that ranges from 5 to 35 amino acid residues in length, depending on the species¹². For the *Bacillus subtilis* YxiN protein, an ortholog of *E. coli* DbpA, it has been shown that a 76-residue carboxy-terminal fragment (residues 404-479) binds short oligonucleotides derived from 23S rRNA with essentially the same affinity and sequence specificity as the full-length protein¹³. The catalytic core of YxiN, residues 1-368, has an ATPase activity and a weak, sequence-nonspecific RNA binding activity similar to those of eukaryotic initiation factor 4A (eIF4A), a minimal DEAD-box helicase consisting of only the two catalytic core domains¹⁴⁻¹⁶. These data confirm a modular structure for the DbpA proteins, in which the carboxy-terminal domain is an RNA-binding domain (RBD) which recognizes a specific site in 23S rRNA, thereby localizing the proteins to their biological target, and the catalytic core exercises a generic, ATP-dependent RNA duplex destabilization activity in the vicinity of the specific rRNA binding site.

The crystal structure of the *B. subtilis* YxiN RBD has been solved at 1.7 Å resolution¹⁷. The RBD has the tertiary fold of an RNA recognition motif (RRM), which was first described for the spliceosomal U1A protein¹⁸. RRM have a 4-strand antiparallel beta sheet with helices on one face in a $\beta\alpha\beta\beta\alpha\beta$ topology. Similar to U1A, the RBD has two aromatic residues on the exposed face of the beta sheet. In U1A, ssRNA binds in a manner that stacks bases of nucleotides on the side chains of these aromatic residues¹⁹⁻²¹; mutating either of these aromatic residues results in substantial reduction of RNA affinity^{22, 23}. However, mutating either of these residues to alanine in the YxiN RBD (Y407A, Y447A mutations) has negligible effect on the binding of short RNA oligonucleotides consisting of hairpin 92 plus single-strand extensions, indicating that the YxiN RBD must exploit a different mode of RNA recognition¹⁷. Here, we report the structure of a complex between the YxiN RBD and a fragment of rRNA that reveals the mode of specific RNA recognition utilized by RBDs of proteins of the DbpA family. Even though several different schemes of RNA

recognition by RRM s have been described in recent years²⁴, the mode by which the YxiN RBD recognizes rRNA is novel.

Results

The RNA used in this work is derived from nucleotides 2508-2580 of *E. coli* 23S rRNA, which has three duplex stems designated H90, H91 and H92 (Figure 1a); stem H92 plus its five-nucleotide connecting loop is generally designated “hairpin 92”¹⁴. To produce crystals of the protein-RNA complex suitable for structure determination, a large library of RNAs differing primarily in lengths of stems H90 and H91 was screened; hairpin 92 was left invariant since it is known to be essential for RBD binding²⁵. To facilitate crystallization, a GAAA tetraloop was introduced in the loop of H91, altering a five-nucleotide segment of the native rRNA sequence. This yielded crystals that allowed us to solve the structure of the complex of the *B. subtilis* YxiN RBD and a cognate rRNA recognition fragment at 2.9 Å resolution (Figure 1c-d). As described below, the non-native tetraloop introduced in H90 is remote from the protein binding site. The *B. subtilis* rRNA sequence differs from that of *E. coli* at six positions, maintaining three base pairs in duplex regions of the RNA (shown in red in Figure 1a); these differences would not be expected to alter the structure of the RNA or specific interactions between the RBD protein and RNA. Therefore, we anticipate this complex accurately represents a complex between the *B. subtilis* RBD and its cognate 23S rRNA.

Structure of the RNA in the complex

H90 and H91 of the RNA stack to form an extended stem, while H92 protrudes via the three-way junction, giving the molecule a Y shape (Figure 1). The protein is wedged between hairpin stems H90 and H92 and the three-way junction. Hairpin 92 has five Watson-Crick (W-C) base pairs and a five-nucleotide ssRNA loop as anticipated from secondary structure prediction. In addition, the five nucleotides following H92, U2562-A2566, generally depicted in secondary structure predictions as single-stranded, form a non W-C base pair at the base of the stem (U2562:A2566) and a “uridine turn” (nucleotides U2563-A2565)²⁶ which presents the bases of two adenine nucleotides to the minor groove of two G:C base pairs of H90. It has been noted that such “A-minor” interactions are frequent determinants of RNA tertiary structure^{27, 28}.

H90 is less well-ordered in the crystal than the rest of the RNA. A five-nucleotide bulge in H90, U2571-G2576, could not be traced unambiguously in either the experimental electron density map or subsequent $F_o - F_c$ difference maps, although the maps show unequivocal positive density in the region where the nucleotides would be located (shown as green mesh contour in Figure 1c-d). Hence, these nucleotides are not included in the model. It is apparent from an $F_o - F_c$ difference map that this region of H90 interacts with the protein; this is consistent with footprinting protection data on the RNA by the *E. coli* homolog protein DbpA14 (discussed below), and with affinity measurements presented here.

Structure of the protein in the complex

The structure of the RBD alone was previously determined at 1.7 Å resolution (PDB ID 2G0C)¹⁷. In that structure, two lysine-rich loops could not be traced (residues 414-417 and 468-470), and the first half (residues 418-422) of one alpha helix displayed two alternate conformations, due to an apparent hinge motion around Gly423 in the middle of the helix. RNA binding locks these regions into specific conformations (Figure 2). The hinged helix that had two conformations, and that abuts hairpin 92 of the RNA, has a unique conformation in the complex (corresponding to “conformation 2” in PDB file 2G0C of the protein-alone structure). The polypeptide backbone can be traced unambiguously in the two

lysine-rich segments that were untraceable in the absence of RNA. The backbone of the protein in the RBD-RNA complex superimposes on that of the RBD alone with a root mean square (rms) difference in C_{α} positions of 0.84 Å, confirming the similarity of the protein tertiary fold in the presence and absence of RNA.

Specific RNA-protein interactions of the H92 stem-loop

Studies with the *E. coli* DbpA protein demonstrated sequence-specific affinity determinants within the five-nucleotide single strand loop of hairpin 92, U2552-C2556. A single U2555G mutation reduces the affinity of a 153-nucleotide RNA for DbpA substantially in a gel-shift assay¹¹. The double mutation G2553U/U2555A impairs both the RNA-dependent activation of ATPase activity¹⁰ and the RNA strand separation activity²⁵ of *E. coli* DbpA. The same double mutation increases the apparent RNA $K_{1/2}$ of RNA-dependent ATPase activity two orders of magnitude for *B. subtilis* YxiN protein and substantially reduces the RNA affinity of the RBD¹³.

In the structure of the complex, the bases of G2553-C2556 all stack on the protein surface (Figure 3) while U2552, the first nucleotide of the loop, does not interact with the protein. Notably, the base of G2553 is positioned to form three hydrogen bonds to the peptide backbone of the peptide segment Ile436-Ile439 (N.B. Since peptide backbone carbonyls cannot be discerned at 2.9 Å resolution, we are assuming the conformation of this segment of the polypeptide backbone in the complex is not dramatically different from that determined unambiguously at 1.7 Å resolution in the RBD-alone structure¹⁷). The conformation of this peptide segment appears dependent on the capability of Gly437 to adopt Ramachandran backbone angles that are unallowed for other amino acid residues. Notably, a glycine at the position of G437 is conserved throughout homologous DbpA RBD sequences (e.g. see pfam 0388029), suggesting this interaction is a conserved determinant of RNA-sequence specific recognition of the RBD.

The base of U2555 forms two apparent hydrogen bonds with the peptide backbone carbonyl of Thr466 and the backbone nitrogen of Lys468. C2556 hydrogen bonds to the side chain hydroxyl of Thr466.

There is clear electron density showing that the side chain of Arg417 “threads” the eye of the loop of hairpin 92 and stacks on the C2551-G2557 base pair (Figure 3). This arginine appears to be stabilized in position by electrostatic interactions with Asp420 and with phosphates of the RNA backbone. An arginine at this position is conserved in a subset of DbpA homologs, including the *E. coli* protein.

Electrostatic RNA-protein interactions

The RBD has one arginine and 14 lysine residues. The carboxy-terminal Lys479 could not be traced in this structure. The arginine and all but one of the remaining lysines are in positions where they could interact with one or more phosphates of RNA backbone in the crystal, either within the complex (eight lysines: 412-415, 427, 468, 470 and 473 interacting with RNA1 in Figure 4) or through crystal contacts with a neighbor RNA molecule (four lysines: 405, 460, 463 and 476 interacting with RNA2). (The exception, Lys454, forms a salt bridge with Asp435 and an apparent hydrogen bond to the carbonyl of Gly430, suggesting a role in stabilizing the polypeptide backbone conformation). Most of these lysines do not show a unique, fully-occupied conformation, suggesting the electrostatic interactions with the phosphates probably accommodate several alternative side chain conformations. These side chains have been modeled with high-probability rotamers that place the N_{ϵ} nitrogens in close proximity to phosphates of the phosphoribose backbone of the RNA and have been adjusted by crystallographic refinement. Calculation of the protein electrostatic potential

shows that the lysines form an electropositive “band” around a large fraction of the protein. The phosphoribose backbone of RNA tracks much of the band. Notably, beyond the point where H90 diverges from the protein and engages in intermolecular contacts with a neighbor molecule, a second RNA molecule converges with its phosphates tracking the electropositive band through crystal-contact interactions. This suggests the possibility that the phosphoribose backbone of other H90 variants that have been tested in biochemical studies^{10, 14} might wrap continuously around the extent of the electropositive band; whether this occurs can be explored with co-crystallization of complexes with different RNAs.

Affinities of RNA-RBD complexes

Electrophoretic mobility shift assays (EMSAs) were carried out to measure the affinities of wildtype (wt) and mutant RBD proteins for RNAs (Figure 5 and Table 1). Previous affinity measurements on YxiN and its RBD employed short oligonucleotides consisting of hairpin 92 and 5' single-strand extensions¹³. The structure of the RNA-RBD complex reveals significant interaction with the duplex and bulge of H90, including electron density in the vicinity of Tyr407. In this context, the Y407A RBD mutant that showed no difference in affinity from wt protein with short oligos was tested with the larger RNA used in the structure determination. Reciprocally, the bulge of H90 was deleted and the affinity of the RBD for an RNA with a fully-duplex H90 was measured. Gly423 in the middle of helix 1 of the RBD protein allows multiple conformations of the amino-terminal segment of the helix; a G423A mutation, which would assumably rigidify the helix and which also introduces an extraneous methyl group at the protein-RNA interface, was constructed and characterized to determine whether flexibility within the helix is essential for high-affinity RNA binding. Additionally, one RNA “sequence specificity mutant”, G2553U, was tested.

The affinity of the RNA used in crystallization, which we refer to as “wildtype”, is approximately 50 picomolar (refer to Table 1 for precise values). Mutating the glycine in alpha helix 1 that interfaces with hairpin 92 to alanine, G423A, reduces the affinity approximately 2-fold. Mutating the tyrosine that interfaces to the ambiguous electron density of the bulge in H90 to alanine, Y407A, also reduces the affinity ~2-fold. Replacing the bulge of H90 with a sequence expected to result in a fully duplex H90 reduces the affinity approximately 30-fold to 1.7 nM. A single-nucleotide mutation G2553U in the loop of hairpin 92 reduces the RNA binding affinity two orders of magnitude to ~7 nM, measured by competition with wt RNA.

Discussion

Several different schemes by which proteins with the RRM fold recognize specific RNAs have been described in recent years²⁴. To our knowledge, the mode by which the YxiN RRM (and by extrapolation, RRM of the DbpA subfamily of DEAD-box helicases) binds its target rRNA is an additional scheme that has not been described previously. There is both structural and electrostatic complementarity between the protein and the phosphoribose backbone of the 3-way junction plus stems H90 and H92 of the RNA. Additionally, RNA sequence specificity in the interaction is encoded in the single-strand loop of hairpin 92, and apparently to a lesser extent in the single-strand bulge of H90. It is notable that the guanosine base of G2553 of hairpin 92 forms three hydrogen bonds to the peptide backbone of the RBD in the same manner it hydrogen-bonds to the penultimate cytosine base of the CCA terminus of A-site tRNA in an active ribosome³⁰.

Consistency of the structure of the RNA-RBD complex with previous biochemical results

As described above, previous biochemical studies with RNAs having both a single G2553A mutation and double G2553U/U2555A mutation have shown that nucleotides in the single strand loop of hairpin 92 contribute to sequence-specific RNA recognition by DbpA10¹¹·25 and YxiN13. Consistent with these results, the bases of four nucleotides of the hairpin 92 loop interact directly with the RBD, and we have shown that a single G2553A mutation reduces the affinity of the RNA for the YxiN RBD approximately two orders of magnitude. Details of the specificity determinants of the RNA-RBD interaction merit additional structural and biochemical examination.

The structure suggests electrostatic interactions may contribute a substantial fraction of the binding energy of the complex. Previous results showing that minimal RNAs that include hairpin 92 and a ssRNA extension (minimally five nucleotides on the 5' end, or nine nucleotides on the 3' end) bind with ~10 nM affinity¹⁰. Shortening the single-strand extensions reduced binding affinity. This result can be rationalized by modeling single-strand RNA extensions on hairpin 92. Figure 6 shows hypothetical models in which poly-U extensions, 5 nucleotides in length on the 5' end or 9 nucleotides on the 3' end, have been added to hairpin 92 of the RNA-RBD complex structure. The poly-U extensions were manually adjusted such that the phosphoribose backbone tracks the positive band of electrostatic potential of the protein. These hypothetical models show that single-strand extensions of sufficient length could engage in most if not all of the electrostatic interactions between the RBD and its natural RNA binding target. The longer 3' extension is necessitated by the fact that the 3' end of hairpin 92 is distant from the RBD while the 5' end is in direct contact with the RBD. Since many of the binding studies were done with full-length DbpA or YxiN protein, it could be argued that the poly-U extensions may be interacting with the core of the helicase. However, parallel measurements on YxiN full-length protein and RBD for 5' poly-U extensions ranging from 2 to 15 nucleotides showed that the binding energies were essentially equal for the two constructs except for the shortest extensions of two or three nucleotides¹³, arguing that the binding can be attributed to the RBD.

Footprinting of the interaction of RNA with the *E. coli* DbpA protein showed that specific nucleotides of the duplex and loop of hairpin 92 (nucleotides 2549 and 2552-2554), of the single strand bulge of H90 (2572-2574), and of the 5-nucleotide segment that runs from H92 to H90 (nucleotides 2562 and 2564-2566, around the U-turn in the RNA-RBD complex; presented as single-stranded in previous work) are sites of AMPPNP-independent RNA protection (Figure 7)¹⁴. The protection of sites within hairpin 92 can be understood as due to the direct interaction with the protein. Protected sites in the bulge of H90 are not visible in the RNA-RBD model due to the ambiguity of this part of the structure, but can also be understood as due to direct protein-RNA interaction. The sites of protection in the U-turn region of the three-way junction are too distant from the protein to be due to direct steric protection within the complex. There are two alternative explanations for this result. The first is that the protection might be due to tighter folding of this part of the RNA tertiary structure when the RBD is bound. The second explanation is suggested by the intermolecular contact between the RBD and a symmetry-related RNA molecule in the crystal (denoted RNA2 in Figure 7), whose U-turn region contacts the RBD. Under the high concentrations of reagents used in the footprinting experiments, *to wit*, 50 nM RNA, 1 μ M DbpA protein¹⁴, it is possible that the protein provides steric protection of the RNA through a secondary mode of binding mimicked by this intermolecular interaction in the crystal. However, gel shift titrations of 32 μ M RNA with increasing amounts of RBD (originally carried out to characterize solution behavior of the complex under reagent concentrations near those used for crystallization) showed formation of a well-defined 1:1 RNA-protein complex (Figure S1), with formation of poorly-resolved higher molecular weight species rather than a second specific complex when protein was titrated in excess of a 1:1 molar

ratio to RNA (data not shown), arguing that such a secondary binding site would have a much weaker affinity than the primary binding site.

Comparison of the RNA structure in the RBD complex to its structure in the ribosome

The structure of this RNA fragment in the mature *E. coli* ribosome also has three stems and a three-way junction, although stem-loop 92 and stem H90 condense to interact with each other when the RBD protein is not intercalated between them³⁰. There are significant structural differences within the RNA fragment in the two different contexts; these differences correlate with specific intramolecular interactions within the ribosome subunit (Figure 8). They are most easily described by comparing the secondary structures of the RNA in the two contexts. The base pairing and uridine turn of the H92 stem-loop are essentially the same in both structures; the two adenosines of the uridine turn retain the A-minor interactions with two G-C base pairs in H91. Most notably, in H91 there is a “slippage” of three successive guanosine nucleotides, G2543-G2545. In the RNA-RBD structure, these three nucleotides base-pair with U2519, C2520 and C2521, while in the ribosome, they translocate one nucleotide to pair with C2520, C2521 and U2522. Along with the consequent “unpairing” of U2519, the base of A2518 flips out and forms specific interactions with neighboring regions of the ribosomal RNA. Additionally, on the other strand of the H91 duplex, A2542 is left unpaired, so that H91 has tandem unpaired adenosines in the ribosome, while having only a single unpaired nucleotide, A2541, in the RBD complex. In the mature ribosome, the nucleotide bases of A2572 and A2577 in stem H90 and of G2529 in H91 are flipped out and make specific intramolecular interactions within the large subunit; in the structure of the RBD-RNA complexes, these nucleotides are distant from the RBD protein.

The differences in RNA structure in the two different contexts—bound to the RBD versus folded in the mature ribosome—may be biologically relevant. That is, the RBD may bind a “pre-ribosomal” conformation of this RNA fragment during the process of ribosome maturation, and the RNA would then undergo conformational transitions when adapting to its environment in the mature ribosome. This suggests the possibility that, in addition to localizing the helicase by binding its RNA target, the RBD may serve as a folding template for the segment of RNA which it binds. However, we cannot discount the possibility that the RBD-binding RNA fragment, in the context of a larger RNA that included all the contributors of intramolecular interactions of the ribosome, would have a structure similar to that found in the mature ribosome. The RBD would then recognize a slightly different RNA structure than described here.

Localization within the ribosome large subunit and its functional implication

The DbpA family of DEAD-box RNA helicases are constituted of (a) the minimal “catalytic core” consisting of two RecA-related domains that destabilize RNA duplexes in an ATP-dependent reaction, and (b) a carboxy-terminal domain responsible for binding a specific region of 23S rRNA in the large subunit of the ribosome. Structures of the second domain of the catalytic fragment³¹ and the RBD¹⁷ of *B. subtilis* YxiN have defined the structural boundaries of these domains. Solution small-angle x-ray scattering studies have shown that these two domains are distended in solution³². Sequence alignments show the linker region between the catalytic core and the RNA recognition domain to range 5-35 amino acid residues. This information can be used to construct reasonable models of how DbpA proteins may interact with the ribosome large subunit.

We have used the structure of the catalytic core of the *Drosophila* Vasa protein with single-stranded U₁₀ bound (PDB 2DB333) as a representative model for a generic DEAD-box catalytic fragment in its RNA-binding conformation. The carboxy terminus of the Vasa

fragment has been positioned near the amino terminus of the YxiN RBD in the RBD-RNA complex (Figure 9a). The RNA of the complex has been superimposed on the equivalent RNA fragment in the large subunit of the *E. coli* ribosome³⁰. This RNA fragment, and by implication, the RBD, is in the peptidyl transferase center (Figure 9b). The orientation of the RBD reveals that a catalytic helicase fragment, connected to the RBD through a flexible peptide tether, would be in position to access a large fraction of the RNA in the extended tRNA binding cleft of the large subunit. This suggests that the function of DbpA proteins during ribosome biogenesis is to repetitively melt unstable segments of misfolded RNA in the cleft, allowing them another opportunity to re-anneal into the correct structure; DbpA would function as an “RNA chaperone” in ribosome maturation³⁴.

Materials and Methods

Reagents

The parent vector for cloning transcription sequences, pRAV12, was provided by Dr. Robert Batey, University of Colorado³⁵; the mutant hepatitis delta virus (HDV) ribozyme of pRAV12 was reverted to wildtype with a U75C mutation; this vector is referred to as pRAV12(wt). T7 RNA polymerase was prepared using standard protocols³⁶. Calf Intestinal Alkaline Phosphatase (CIP), T4 Polynucleotide Kinase (PNK), deoxynucleoside triphosphates (dNTPs), Taq DNA polymerase, T4 DNA Ligase, and restriction enzymes were purchased from New England Biolabs. Deoxyoligonucleotides used for RNA template preparation were purchased from Integrated DNA Technologies. Ribonucleoside triphosphates (NTPs) were purchased from Amersham Biosciences. G-25 Sephadex spin columns were purchased from Roche Applied Science. [γ -³²P]-Adenosine 5'-triphosphate (6000 Ci/mmol) was purchased from PerkinElmer. Iridium hexamine was provided by Dr. Robert Batey.

Protein expression and purification

YxiN RBD proteins, including native wildtype, SeMet-labeled wildtype, G423A mutant, and Y407A mutant proteins, were expressed and purified as described previously^{13, 17}.

RNA cloning, transcription, purification and labeling

RNAs were transcribed as 5' fusions with a hepatitis delta virus (HDV) self-cleaving ribozyme and purified on acrylamide gels after HDV self-cleavage at the insert-ribozyme junction.

Cloning of RNA-HDV fusions—The target sequence for *in vitro* transcription was PCR-amplified using synthetic DNA primers and template and then cloned into the pRAV12(wt) vector by restriction/ligation using EcoRI and NgoMIV restriction sites. For the sequence of RNA in the structure of the complex (Figure 1a), the template nucleotide was 5'-ATACGACTCACTATAGGCTCATCATCCTGGGGCTGGAAACGGTCCCAAGGGTATGG, and the 5' and 3' primers were 5'-GCGCGCGA**ATTCTA**ATACGACTCACTATAG and 5'-ATGGCCGGCTAGCTCGCGTACCACTTTAAATGGCGAACAGCCATACCCTTGGGACC respectively (primer-template overlaps are underlined; EcoRI and NgoMIV restriction sites of 5' and 3' primers respectively are bold; T7 promoter in 5' primer in italics). Reaction conditions for PCR were: 1 nM template, 1 μ M each primer, 0.2 mM dNTPs, 2.5 units Taq polymerase in 100 μ L. Amplification conditions were: 5 min at 95 °C followed by 25-30 cycles of 1 min. at 90 °C, 1 min. at 55 °C, 1 min. at 72 °C. The same strategy, using different oligonucleotide sequences, was used to generate an RNA with a fully-duplex H90 (Figure 1b). Accuracy of inserts ligated into pRAV12(wt) were confirmed by DNA sequencing.

In vitro RNA transcription and purification—A transcription template was PCR-amplified from the pRAV12(wt)-derived construct. RNA was transcribed by standard methods^{35, 36}; the HDV ribozyme self-cleaved during the transcription. The RNA product was purified on 8% polyacrylamide denaturing gels run with 8 M urea in TBE (89 mM Tris-Borate, 2 mM EDTA, pH 8.3), visualized by UV shadowing, excised from the gel, electroeluted from the acrylamide, and concentrated to >24 mg/ml in H₂O.

³²P-labeling of RNA—100 µg of gel-purified RNA was dephosphorylated for 3 hours at 37 °C in a 100 µL reaction volume consisting of 40 mM Na-MES (pH 6.0), 10 mM MgCl₂, 5 mM dithiothreitol (DTT), and 100 units CIP. RNA was phenol/chloroform extracted two times, precipitated with ethanol, and gel-purified. 20 pmols of dephosphorylated RNA was then radiolabeled at 37°C for 30 min. in a 20 µL reaction volume that included 70 mM Tris-HCl (pH 7.6), 10 mM MgCl₂, 5 mM DTT, 50 pmols (γ-³²P) ATP, and 20 units PNK. Unincorporated nucleotides were removed from the product by gel filtration with a G-25 Sephadex spin column.

Electrophoretic mobility shift assays (EMSA)

Direct binding EMSA experiments were performed with samples containing 4 pM ³²P-labeled RNA and protein within the range ~0.3 pM to 5 µM in an EMSA sample buffer consisting of 16 mM K-HEPES (pH 7.5), 80 mM KCl, 0.01% NP-40, 2.5 µg poly-A RNA in a volume of 100 µL. Following conditions established previously with DbpA11 and YxiN13, reactions were allowed to equilibrate on ice for 30 min before loading on native polyacrylamide gels (6%) run in TBE buffer at 4 °C. Gels were imaged on a Typhoon variable mode scanner and the signals in the gel bands corresponding to protein-bound and unbound RNA were integrated. Parameters in the following function were fit to the data for fraction of RNA bound versus protein concentration:

$$\theta = \left(\frac{[\text{protein}]}{[\text{protein}] + K_d} \right) (a - b) + b$$

where θ is fraction bound, K_d is the apparent dissociation constant, a is the upper baseline and b is the lower baseline. At least three gel shifts were performed for each sample and associated error is reported as one standard deviation from the mean (Table 1). Eight gel shifts were performed with multiple wt RBD and wt RNA samples in an effort to encompass all sources of systematic error in the measurements.

Competition experiments were effected using samples with 4 pM ³²P-labeled wt RNA and 275 pM wt RBD protein in EMSA sample buffer. Unlabeled competitor mutant RNA ranging from 1 pM to 20 µM was added to these samples, incubated on ice for 30 min., and then run on 6% polyacrylamide gels, imaged, and analyzed as described above. Parameters in the following equation were fit to these data:

$$\theta = b + \left[\frac{m - b}{1 + (IC_{50} / [C])} \right]$$

where m and b are normalization factors for the fraction bound at the upper and lower baselines respectively, θ is the fraction of labeled wt RNA bound, IC_{50} is the concentration of competitor RNA that results in 50% fraction bound for the labeled complex, and C is the concentration of competitor RNA. The K_C for the competitor was then computed using the equation³⁷:

$$K_c = \frac{2K_d \cdot IC_{50}}{2[\text{protein}] - [\text{wtRNA}] - 2K_d}$$

where the K_d is the dissociation constant for wt RNA. Competition measurements were made in triplicate and the results were averaged.

Crystallization, cryoprotection and heavy atom derivitization

Crystals were grown at 17 °C in hanging drops initially constituted of equal volumes of the RNA/YxiN-RBD complex (20 mg/ml, with RNA in 5% molar excess over the protein, in 2 mM Tris-HCl, 20 mM NaCl, pH 8.0) and precipitant (200 mM potassium acetate, 26-30% polyethylene glycol 3350 (PEG-3350)). Crystals were adapted to a cryoprotectant stabilization solution (50 mM MOPS pH 7.0, 150 mM potassium acetate, 40% PEG-3350) by gradually increasing the PEG-3350 concentration of the precipitant in 5% steps. Cryoprotectant-stabilized crystals were derivatized with iridium by transferring them to a solution of cryoprotectant plus 2 mM iridium hexamine for a minimum of 15 min. before flash-freezing in liquid nitrogen.

Data collection and structure determination

Data were collected at the Stanford Synchrotron Radiation Laboratory (SSRL) beamline 9.2 and the Advanced Light Source (ALS) beamlines 8.2.1 and 8.2.2. Diffraction images were processed and scaled with HKL200038 or Mosflm/Scala39. Data reduction and phasing statistics for multiwavelength anomalous dispersion (MAD) data collected on crystals with SeMet-labeled protein and, separately, crystals with iridium hexamine, are summarized in Table 2. Data reduction statistics for native data are summarized in Table 3.

Crystallographic computations were carried out with recent versions of phenix40, cns41, solve42, and ccp439. Model building was effected with coot43, 44 and o45. Initial phases were computed to 3.5 Å using three-wavelength MAD diffraction data on crystals grown with SeMet-labeled protein (Table 2). Location of the three selenium sites confirmed placement of the YxiN RBD in the experimental map. Automated model building of the RNA in phenix, followed by manual rebuilding, refinement, and phase combination of model and experimental phases yielded a model that included RNA stem-loops 91 and 92 (44 out of 74 nucleotides) in addition to the protein.

Experimental phasing to higher resolution was achieved with an iridium hexamine derivative46. MAD phasing statistics to 3.2 Å for the iridium hexamine data alone are given in Table 2. The improved phasing and electron density maps allowed building of the three-way junction of the RNA and most of the nucleotides of helix 90. Refinement of the model against the iridium hexamine data, while using experimental phase constraints, reduced the refinement parameters to $R_{\text{cryst}} = 0.280$, $R_{\text{free}} = 0.317$. At this point, the model was refined against native data to 2.9 Å (Table 3). Five nucleotides could not be placed unambiguously in the bulge region of helix 90 (Figure 1). The final model includes all but the carboxy-terminal residue of the RBD and 69 out of 74 nucleotides of the RNA.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by grant GM71696 from the National Institutes of Health to DBM and salary support by the Howard Hughes Medical Institute to DBM. Parts of this research were carried out at the Stanford Synchrotron Radiation Laboratory, a national user facility operated by Stanford University on behalf of the U.S. Department of Energy, Office of Basic Energy Science. The SSRL Structural Molecular Biology Program is supported by the Department of Energy and by the National Institutes of Health. Parts of this research were carried out at the Advanced Light Source of the Lawrence Berkeley Laboratory.

Abbreviations

AMPPNP	adenylyl-imidodiphosphate
CIP	calf intestinal alkaline phosphatase
EMSA	electrophoretic mobility shift assay
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
MES	2-(N-morpholino)ethanesulfonic acid
MOPS	3-(N-morpholino)propanesulfonic acid
PNK	T4 polynucleotide kinase

References

1. Rocak S, Linder P. DEAD-box proteins: the driving forces behind RNA metabolism. *Nat. Rev. Mol. Cell Biol.* 2004; 5:232–241. [PubMed: 14991003]
2. Linder P. Dead-box proteins: a family affair--active and passive players in RNP-remodeling. *Nucleic Acids Res.* 2006; 34:4168–4180. [PubMed: 16936318]
3. Caruthers JM, Johnson ER, McKay DB. Crystal structure of yeast initiation factor 4A, a DEAD-box RNA helicase. *Proc. Natl. Acad. Sci. USA.* 2000; 97:13080–13085. [PubMed: 11087862]
4. Caruthers JM, McKay DB. Helicase structure and mechanism. *Curr. Opin. Struct. Biol.* 2002; 12:123–133. [PubMed: 11839499]
5. Johnson ER, McKay DB. Crystallographic structure of the amino terminal domain of yeast initiation factor 4A, a representative DEAD-box RNA helicase. *RNA.* 1999; 5:1526–1534. [PubMed: 10606264]
6. Sharpe Elles LM, Uhlenbeck OC. Mutation of the arginine finger in the active site of *Escherichia coli* DbpA abolishes ATPase and helicase activity and confers a dominant slow growth phenotype. *Nucleic Acids Res.* 2008; 36:41–50. [PubMed: 17986459]
7. Sharpe Elles LM, Sykes MT, Williamson JR, Uhlenbeck OC. A dominant negative mutant of the *E. coli* RNA helicase DbpA blocks assembly of the 50S ribosomal subunit. *Nucleic Acids Res.* 2009; 37:6503–6514. [PubMed: 19734347]
8. Fuller-Pace FV, Nicol SM, Reid AD, Lane DP. DbpA: a DEAD box protein specifically activated by 23S rRNA. *EMBO J.* 1993; 12:3619–3626. [PubMed: 8253085]
9. Nicol SM, Fuller-Pace FV. The “DEAD box” protein DbpA interacts specifically with the peptidyltransferase center in 23S rRNA. *Proc Natl Acad Sci U S A.* 1995; 92:11681–11685. [PubMed: 8524828]
10. Tsu CA, Kossen K, Uhlenbeck OC. The *Escherichia coli* DEAD protein DbpA recognizes a small RNA hairpin in 23S rRNA. *RNA.* 2001; 7:702–709. [PubMed: 11350034]
11. Polach KJ, Uhlenbeck OC. Cooperative binding of ATP and RNA substrates to the DEAD/H protein DbpA. *Biochemistry.* 2002; 41:3693–3702. [PubMed: 11888286]
12. Kossen K, Uhlenbeck OC. Cloning and biochemical characterization of *Bacillus subtilis* YxiN, a DEAD protein specifically activated by 23S rRNA: delineation of a novel sub-family of bacterial DEAD proteins. *Nucleic Acids Res.* 1999; 27:3811–3820. [PubMed: 10481020]

13. Karginov FV, Caruthers JM, Hu Y, McKay DB, Uhlenbeck OC. YxiN is a modular protein combining a DExD/H core and a specific RNA binding domain. *J. Biol. Chem.* 2005; 280:35499–35505. [PubMed: 16118224]
14. Karginov FV, Uhlenbeck OC. Interaction of *Escherichia coli* DbpA with 23S rRNA in different functional states of the enzyme. *Nucleic Acids Res.* 2004; 32:3028–3032. [PubMed: 15173385]
15. Lorsch JR, Herschlag D. The DEAD box protein eIF4A. 2. A cycle of nucleotide and RNA-dependent conformational changes. *Biochemistry.* 1998; 37:2194–2206. [PubMed: 9485365]
16. Lorsch JR, Herschlag D. The DEAD box protein eIF4A. 1. A minimal kinetic and thermodynamic framework reveals coupled binding of RNA and nucleotide. *Biochemistry.* 1998; 37:2180–2193. [PubMed: 9485364]
17. Wang S, Hu Y, Overgaard MT, Karginov FV, Uhlenbeck OC, McKay DB. The domain of the *Bacillus subtilis* DEAD-box helicase YxiN that is responsible for specific binding of 23S rRNA has an RNA recognition motif fold. *RNA.* 2006; 12:959–967. [PubMed: 16611943]
18. Nagai K, Oubridge C, Jessen TH, Li J, Evans PR. Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. *Nature.* 1990; 348:515–520. [PubMed: 2147232]
19. Allain FH, Gubser CC, Howe PW, Nagai K, Neuhaus D, Varani G. Specificity of ribonucleoprotein interaction determined by RNA folding during complex formulation. *Nature.* 1996; 380:646–650. [PubMed: 8602269]
20. Allain FH, Howe PW, Neuhaus D, Varani G. Structural basis of the RNA-binding specificity of human U1A protein. *EMBO J.* 1997; 16:5764–5772. [PubMed: 9312034]
21. Oubridge C, Ito N, Evans PR, Teo CH, Nagai K. Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature.* 1994; 372:432–438. [PubMed: 7984237]
22. Jessen TH, Oubridge C, Teo CH, Pritchard C, Nagai K. Identification of molecular contacts between the U1 A small nuclear ribonucleoprotein and U1 RNA. *EMBO J.* 1991; 10:3447–3456. [PubMed: 1833186]
23. Nolan SJ, Shiels JC, Tuite JB, Cecere KL, Baranger AM. Recognition of an essential adenine at a protein-RNA interface: comparison of the contributions of hydrogen bonds and a stacking interaction. *J. Am. Chem. Soc.* 1999; 121:8951–8952.
24. Clery A, Blatter M, Allain FH. RNA recognition motifs: boring? Not quite. *Curr Opin Struct Biol.* 2008; 18:290–298. [PubMed: 18515081]
25. Diges CM, Uhlenbeck OC. *Escherichia coli* DbpA is an RNA helicase that requires hairpin 92 of 23S rRNA. *Embo J.* 2001; 20:5503–5512. [PubMed: 11574482]
26. Quigley GJ, Rich A. Structural domains of transfer RNA molecules. *Science.* 1976; 194:796–806. [PubMed: 790568]
27. Doherty EA, Batey RT, Masquida B, Doudna JA. A universal mode of helix packing in RNA. *Nat Struct Biol.* 2001; 8:339–343. [PubMed: 11276255]
28. Nissen P, Ippolito JA, Ban N, Moore PB, Steitz TA. RNA tertiary interactions in the large ribosomal subunit: the A-minor motif. *Proc Natl Acad Sci U S A.* 2001; 98:4899–4903. [PubMed: 11296253]
29. Bateman A, Birney E, Durbin R, Eddy SR, Howe KL, Sonnhammer EL. The Pfam protein families database. *Nucleic Acids Res.* 2000; 28:263–266. [PubMed: 10592242]
30. Voorhees RM, Weixlbaumer A, Loakes D, Kelley AC, Ramakrishnan V. Insights into substrate stabilization from snapshots of the peptidyl transferase center of the intact 70S ribosome. *Nat Struct Mol Biol.* 2009; 16:528–533. [PubMed: 19363482]
31. Caruthers JM, Hu Y, McKay DB. Structure of the second domain of the *Bacillus subtilis* DEAD-box RNA helicase YxiN. *Acta Cryst.* 2006; F62:1191–1195.
32. Wang S, Overgaard MT, Hu Y, McKay DB. The *Bacillus subtilis* RNA helicase YxiN is distended in solution. *Biophys. J.* 2008; 94:L01–L03. [PubMed: 17951299]
33. Sengoku T, Nureki O, Nakamura A, Kobayashi S, Yokoyama S. Structural basis for RNA unwinding by the DEAD-box protein *Drosophila* Vasa. *Cell.* 2006; 125:287–300. [PubMed: 16630817]
34. Russell R. RNA misfolding and the action of chaperones. *Front Biosci.* 2008; 13:1–20. [PubMed: 17981525]

35. Kieft JS, Batey RT. A general method for rapid and nondenaturing purification of RNAs. *RNA*. 2004; 10:988–995. [PubMed: 15146082]
36. Milligan JF, Uhlenbeck OC. Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol*. 1989; 180:51–62. [PubMed: 2482430]
37. Lin SY, Riggs AD. Lac repressor binding to non-operator DNA: detailed studies and a comparison of equilibrium and rate competition methods. *J Mol Biol*. 1972; 72:671–690. [PubMed: 4573844]
38. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Meth. Enzymol*. 1997; 276:307–326.
39. Bailey S. The CCP4 Suite: Programs for Protein Crystallography. *Acta Cryst*. 1994; D50:760–763.
40. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst*. 2010; D66:213–221.
41. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Cryst*. 1998; D54:905–921.
42. Terwilliger TC, Berendzen J. Automated MAD and MIR structure solution. *Acta Cryst*. 1999; D55:849–861.
43. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Cryst*. 2004; D60:2126–2132.
44. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Cryst*. 2010; D66:486–501.
45. Jones A. A graphics model building and refinement system for macromolecules. *J. Appl. Cryst*. 1978; 11:268–272.
46. Keel AY, Rambo RP, Batey RT, Kieft JS. A general strategy to solve the phase problem in RNA crystallography. *Structure*. 2007; 15:761–772. [PubMed: 17637337]
47. Schuwirth BS, Borovinskaya MA, Hau CW, Zhang W, Vila-Sanjurjo A, Holton JM, Cate JH. Structures of the bacterial ribosome at 3.5 Å resolution. *Science*. 2005; 310:827–834. [PubMed: 16272117]

Figure 1. Structure of the RBD-RNA complex

(a) Sequence and secondary structure of the RNA in the crystal of RNA-RBD complex. Sequences that differ from the native *E. coli* rRNA sequence are in shaded boxes. Sequence differences of *B. subtilis* rRNA shown in red side-by-side with nucleotides of the *E. coli* RNA sequence used in this work. High-affinity iridium hexamine binding site⁴⁶ in the native rRNA sequence is boxed. Nucleotide bulge of H90 that could not be modeled unambiguously is gray. (b) Sequence used for H90 in fully-duplex construct. Sequence of the remainder of the construct identical to (a). (c) Cartoon drawing of the RBD-RNA complex. Protein is shown as a cartoon enveloped in a semi-transparent surface representation. Color-coding for RNA: A, red; U, green; G, blue; C, orange. A F_O-F_C difference Fourier map, contoured at 3σ , is shown in green in the region where the RNA model could not be built unambiguously. (d) Same as (c), except rotated approximately 180° around a vertical axis.

Figure 2. Cartoon drawing of the RBD protein

Secondary structure elements are numbered as in the original RBD structure (PDB 2G0C)17; polypeptide segments that could not be traced in the original RBD structure are shown in red. Glycine 423 in helix 1 shown as sphere. Tyrosine 407 on $\beta 1$ is also shown. Cartoon and semitransparent surface of the RNA shown in gray. F_o-F_c difference Fourier map, contoured at 3σ , shown in green in the region where the RNA model could not be built unambiguously

Figure 3. Stereo images of the interaction of hairpin 92 with the RBD

Polypeptide backbone shown in cyan as a coil passing through C_{α} positions. Backbone and side chain atoms of selected protein residues also shown. Hydrogen bonds shown in black as dotted lines. Semitransparent surface is shown for side chain of Arg417 to emphasize its stacking on RNA duplex. Phosphoribose backbone of RNA shown in orange as a coil. Atom color scheme: carbon, green; oxygen, red; nitrogen, blue.

Figure 4. Electrostatic potential of the RBD surface

Positive electrostatic potential, blue; negative potential, red. (a-b) Two views of the RBD with electrostatic potential shown as semitransparent surface. Lysine and arginine side chains of the RBD are shown. (c-d) Interaction of RNA molecules with the RBD in the crystal. RNAs drawn as cartoons, with phosphorus atoms and nonbonded oxygen atoms of phosphates in close proximity to the protein surface drawn as spheres.

Figure 5. Electrophoretic mobility shift assay

(a) Binding of wild-type (wt) RNA by wt RBD. Protein concentrations: 0, 0.00031, 0.00094, 0.0028, 0.0085, 0.025, 0.076, 0.229, 0.690, 2.0, 6.0, 19.0, 56, 167 and 500 nM. (b) Representative binding titrations. black circles, wt RNA, wt RBD; red triangles, wt RNA, G423A RBD; green diamonds, wt RNA, Y407A RBD; blue squares, RNA with fully-duplex H90, wt RBD. (c) Representative gel showing competition displacement of labeled wt RNA by unlabeled G2553U mutant RNA. First lane is labeled wt RNA alone. Remaining lanes: RBD concentration, 0.275nM; wt RNA concentration, 0.004 nM; G2553U mutant RNA concentrations (left to right): 0.0014, 0.0042, 0.012, 0.038, 0.113, 0.339, 1.02, 3.05, 9.15, 27.4, 82.3, 247, 741, 2220, 6667 and 20000 nM. (d) Representative displacement titration showing fraction of wt RNA bound to RBD, black circles, versus concentration of unlabeled G2553U mutant RNA.

Figure 6. Models showing how single-strand extensions on hairpin 92 could interact with the RBD

(a) and (b) Sequences used in modeling of 5' and 3' extensions respectively. The poly-U extensions are shown in red. (c) and (d) Models in which the 5' and 3' extensions are adjusted to track the electropositive band of the RBD. The poly-U extensions whose conformations were modeled, corresponding to the red nucleotides in (a) and (b), shown in orange. Hairpin 92 shown: carbon, green; oxygen, red; nitrogen, blue. The 3'-UU following hairpin 92 in construct (a) is shown in yellow in panel (c). 5'=>3' directions of RNA strands indicated with arrows.

Figure 7. Sites of ATP-independent footprinting protection of the RNA by DbpA

Nucleotides which show protection shown on RNA1 as space-filling models in green; site of H90 bulge which shows protection (nucleotides 2572-2574) but is not in the model of the RBD-RNA complex is indicated with dotted arrow. Nucleotides which show protection of region 2563-2566 of symmetry-related RNA2 shown as space-filling in red. Other nucleotides shown as gray cartoons.

Figure 8. Structure of the RNA fragment in the ribosome

(a) Secondary structure of the fragment of RNA used in this work in the structure of the *E. coli* ribosome (PDB ID 2AW4 and 2AWB47). (b) Schematic drawing of the fragment of RNA in the *E. coli* ribosome; color scheme same as (Figure 1c-d).

Figure 9. Model for interaction of a DbpA protein in the large subunit of the ribosome
(a) Model in which the catalytic fragment of Vasa bound to U₁₀ (PDB 2DB333) is juxtaposed to the YxiN RBD in the RBD-RNA complex structure. Vasa fragment, orange with U₁₀ in magenta; YxiN RBD, cyan; RNA, green. Carboxy terminus of the Vasa fragment and amino terminus of the RBD are indicated with '*' connected with a dotted line. (b) Model in which the RNA of the RNA-RBD complex structure is superimposed on the equivalent RNA (shown in red) in the *E. coli* large ribosome subunit (PDB 2WDI30). Ribosome surface is shown: gray, RNA; purple, protein.

Table 1

Weighted mean values and standard deviations of dissociation constants (K_d s) for RNA-RBD complex computed from EMSA data

RBD construct	RNA construct	Number of measurements	K_d (nM)	sigma (nM)
wt	wt	8	0.054	0.021
Y407A	wt	3	0.128	0.013
G423A	wt	4	0.113	0.015
wt	duplex H90	4	1.70	0.23
wt *	G2553U *	3	7.1	2.0

* Measured with competition assay.

Table 2
Data collection and MAD phasing statistics for RBD-RNA complex

Crystal spacegroup P2 ₁ 2 ₁ 2 ₁ SeMet-labeled RBD, a = 45.04 Å, b = 91.65 Å, c = 116.91 Å.						
Data Statistics						
Wavelength (Å)	Resolution (Å) (highest shell)	Completeness (%)	R _{sym} ^a	b _F ^c	b _F ^c	b _F ^c
λ1 = 0.97891 (peak)	50.0-3.50(3.63-3.50)	0.974 (0.875)	0.075 (0.177)	-7.0	3.7	3.7
λ2 = 0.97916 (edge)	50.0-3.50(3.63-3.50)	0.917 (0.810)	0.060 (0.287)	-8.4	4.1	4.1
λ3 = 0.95572 (remote)	50.0-3.50(3.63-3.50)	0.925 (0.779)	0.065 (0.403)	-3.8	2.0	2.0
Diffraction Ratios and Phasing Statistics						
Anomalous Diffraction Ratios ^c			Anomalous Phasing Power ^d			
Wavelength (Å)	λ1	λ2	λ3			
λ1 (peak)	0.079	0.073	0.090	0.89	1.54	
λ2 (edge)		0.076	0.077	1.53	2.32	
λ3 (remote)			0.085	reference	1.37	
Figure of Merit <m>						
Resolution (Å)	48.8-6.97	6.97-5.54	5.54-4.84	4.84-4.40	4.40-4.08	4.08-3.84
<m>	0.884	0.852	0.771	0.709	0.638	0.563
						0.506
						0.432
						0.694
Native RBD with iridium hexamine, a = 45.47 Å, b = 91.74 Å, c = 114.03 Å.						
Data Statistics						
Wavelength (Å)	Resolution (Å) (highest shell)	Completeness (%)	R _{sym} ^a	b _F ^c	b _F ^c	b _F ^c
λ1 = 1.10604 (peak)	50.0-3.20(3.36-3.20)	0.940 (0.885)	0.076 (0.340)	-41.4	15.3	15.3
λ2 = 1.10629 (edge)	50.0-3.20(3.36-3.20)	0.937 (0.908)	0.073 (0.294)	-44.4	8.2	8.2
λ3 = 1.05970 (remote)	50.0-3.20(3.36-3.20)	0.965 (0.960)	0.105 (0.320)	-10.3	15.9	15.9
Diffraction Ratios and Phasing Statistics						
Anomalous Diffraction Ratios ^c			Anomalous Phasing Power ^d			
Wavelength (Å)	λ1	λ2	λ3			
λ1 (peak)	0.105	0.110	0.113	0.88	1.16	

Diffraction Ratios and Phasing Statistics				
Wavelength (Å)	Anomalous Diffraction Ratios ^c		Anomalous Phasing Power ^d	
	λ_1	λ_2	λ_3	
		0.079	0.120	1.15
			0.078	reference

Figure of Merit <m>									
Resolution (Å)	50.0-6.37	6.37-5.06	5.06-4.42	4.42-4.02	4.02-3.73	3.73-3.51	3.51-3.34	3.34-3.20	Overall
<m>	0.613	0.631	0.587	0.535	0.505	0.443	0.368	0.323	0.504

^a $R_{\text{sym}} = \sum |I_{\text{hkl}}| - \langle I_{\text{hkl}} \rangle / \sum \langle I_{\text{hkl}} \rangle$ where I_{hkl} = single value of measured intensity of hkl reflection, and $\langle I_{\text{hkl}} \rangle$ = mean of all measured value intensity of hkl reflection. Bijvoet measurements were treated as independent reflections for the MAD phasing data sets.

^b Values of f' and f'' were initially estimated from an EXAFS scan and refined in SOLVE42 or CNS41.

^c Anomalous diffraction ratio values are $= \langle (\Delta F)^2 \rangle^{1/2} / \langle |F|^2 \rangle^{1/2}$, where ΔF is the dispersive (off-diagonal element), or Bijvoet (diagonal element) difference, computed in CNS41.

^d Phasing Power = $\langle |F_H| \rangle / E$, where $\langle |F_H| \rangle$ is the rms structure factor amplitude for anomalous scatterers and E is the estimated lack of closure error; computed in CNS41 using all data between 50.0 Å and highest resolution.

Table 3
Crystallographic data collection and refinement statistics for native data

Spacegroup $P2_12_12_1$, $a = 45.01 \text{ \AA}$ $b = 91.63 \text{ \AA}$ $c = 115.28 \text{ \AA}$

	Cryst x
Data Collection	
Wavelength ($\text{\AA}$)	1.000
Resolution range (last shell) ($\text{\AA}$)	50.0-2.9 (3.0-2.9)
Observations (total/unique)	44210/10490
Redundancy	4.2 (3.1)
Completeness (%)	94.3 (71.6)
$R_{\text{sym}}^{\pm}$	0.054 (0.296)
Refinement	
Resolution range (last shell) ($\text{\AA}$)	20.0-2.90 (3.05-2.90)
$R_{\text{cryst}}^{\S}$	0.209 (0.266)
R_{free}	0.275 (0.322)
Number of reflections (working set)	8956
Number of reflections (test set)	1039
Number of protein/RNA atoms	2044
Average B value, all atoms ($\text{\AA}^2$)	134
Rmsd [#] bond length ($\text{\AA}$)	0.005
Rmsd angles ($^{\circ}$)	1.060

* $R_{\text{sym}} = \sum |I_{hkl} - \langle I_{hkl} \rangle| / \sum \langle I_{hkl} \rangle$ where I_{hkl} = single value of measured intensity of hkl reflection, and $\langle I \rangle$ = mean of all measured value intensity of hkl reflection.

$R_{\text{cryst}}^{\S} = \sum |F_{\text{obs}} - F_{\text{calc}}| / \sum F_{\text{obs}}$ where F_{obs} = observed structure factor amplitude and F_{calc} = structure factor calculated from model. R_{free} is computed in the same manner as R_{cryst} , using the test set of reflections.

[#] Rmsd is root mean square deviation from ideal value.