Published in final edited form as:

Cell Rep. 2019 December 17; 29(12): 4024–4035.e5. doi:10.1016/j.celrep.2019.11.059.

RNA Specificity and Autoregulation of DDX17, a Modulator of MicroRNA Biogenesis

Tri D. Ngo[1,2,3], **Alexander C. Partin**[1,2,3], **Yunsun Nam**[1,2,3,4,*]

[1]Cecil H. and Ida Green Center for Reproductive Biology Sciences and Division of Basic Reproductive Biology Research, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

[2]Department of Biophysics, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

[3]Department of Obstetrics and Gynecology, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

[4]Lead Contact

SUMMARY

DDX17, a DEAD-box ATPase, is a multifunctional helicase important for various RNA functions, including microRNA maturation. Key questions for DDX17 include how it recognizes target RNAs and influences their structures, as well as how its ATPase activity may be regulated. Through crystal structures and biochemical assays, we show the ability of the core catalytic domains of DDX17 to recognize specific sequences in target RNAs. The RNA sequence preference of the catalytic core is critical for DDX17 to directly bind and remodel a specific region of primary microRNAs 3′ to the mature sequence, and consequently enhance processing by Drosha. Furthermore, we identify an intramolecular interaction between the N-terminal tail and the DEAD domain of DDX17 to have an autoregulatory role in controlling the ATPase activity. Thus, we provide the molecular basis for how cognate RNA recognition and functional outcomes are linked for DDX17.

Graphical Abstract

*Correspondence: yunsun.nam@utsouthwestern.edu.

AUTHOR CONTRIBUTIONS

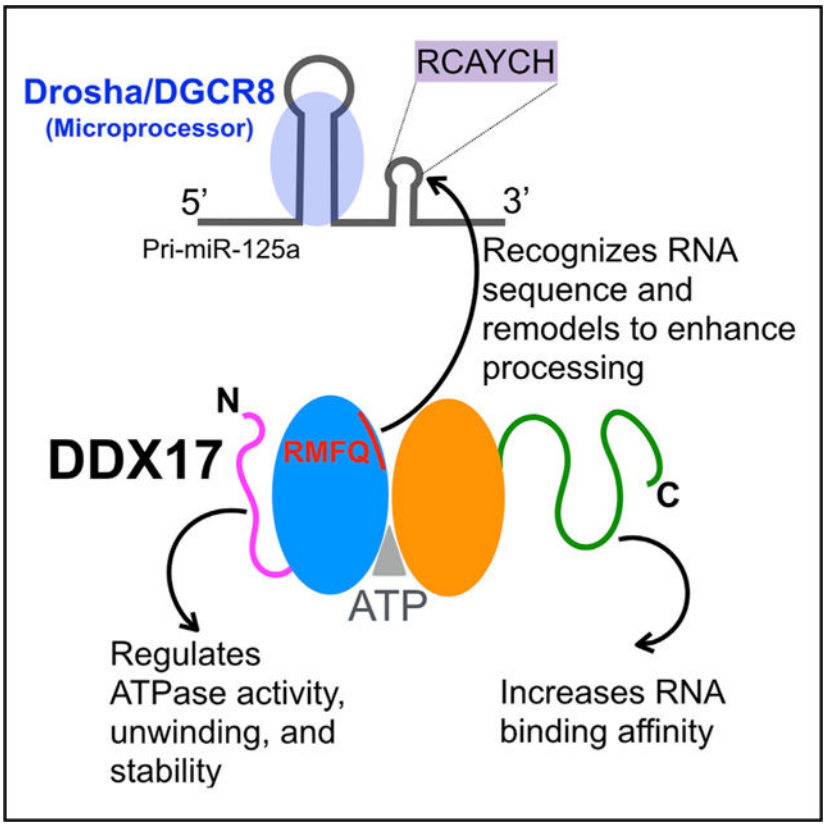
T.D.N. and Y.N. designed experiments. T.D.N. carried out most of the experiments, including determining and refining the crystal structures. A.C.P. performed qPCR experiments and assisted with biochemical assays. T.D.N., A.C.P., and Y.N. wrote the manuscript.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.celrep.2019.11.059.

DECLARATION OF INTERESTS

The authors declare no competing interests.



In Brief

Ngo et al. reveal crystal structures of DEAD-box 17 (DDX17) and show that the core catalytic domains recognize RNA sequence motifs in primary transcripts of microRNAs, to regulate processing by Drosha. DDX17 also has a unique N-terminal tail that can attenuate the ATPase activity.

INTRODUCTION

DDX17 is a member of the large family of DEAD-box RNA helicases found within the superfamily 2 ATPases (Giraud et al., 2018; Linder and Jankowsky, 2011; Xing et al., 2019). DEAD-box proteins are necessary for proper function of RNA in many cellular processes (Cordin et al., 2006; Fuller-Pace, 2013b; Gustafson and Wessel, 2010; Janknecht, 2010; Jarmoskaite and Russell, 2011). DDX17 and its close homolog DDX5 play an important role in various contexts, including processing of primary microRNA transcripts (pri-miRs) in the nucleus (Kao et al., 2019; Li et al., 2017; Mori et al., 2014), pre-mRNA alternative splicing (Dardenne et al., 2014; Hönig et al., 2002), ribosome biogenesis (Jalal et al., 2007), mRNA export (Montpetit et al., 2011), and coregulation of transcription (Dardenne et al., 2014; Fuller-Pace, 2013a; Lambert et al., 2018; Samaan et al., 2014). DDX17 has also been implicated in immunity by affecting viral infectivity (Lorgeoux et al., 2013; Moy et al., 2014; Sithole et al., 2018). Dysregulated expression of DDX17 has been associated with many cancers, including those in colon, prostate, breast, brain, lung, stomach, and blood

(Cai et al., 2017; Fuller-Pace and Moore, 2011). While the long list of significant biological functions highlights the importance of DDX17, the multi-tasking nature of DDX17 has also made its mechanisms perplexing.

Similar to other DEAD-box proteins, DDX17 is an RNA-dependent ATPase. For its helicase function, DDX17 is thought to remodel RNA in a non-processive manner (Huang and Liu, 2002; Pyle, 2008; Rössler et al., 2001; Xing et al., 2017, 2019). DDX17 has been reported to bind various RNA sequences and structures, including stem-loop structures (Mori et al., 2014; Moy et al., 2014; Remenyi et al., 2016), G-quadruplexes (Herdy et al., 2018), and R-loops (Wang et al., 2018). However, how DDX17 operates on specific RNA targets is unclear. The structural requirement for a stem structure (Moy et al., 2014) differs from the genomic studies that identified enhancer elements that also attract DDX17 binding (Hönig et al., 2002). Many DEAD-box proteins bind RNA in a sequence-independent manner (Gilman et al., 2017; Sengoku et al., 2006), and how DDX17 function might depend on RNA sequence is unclear. Multiple C-rich motifs have been associated with DDX17 function, but the patterns are different and whether the sequences reflect direct binding sites of DDX17 is also unknown (Hönig et al., 2002; Mori et al., 2014; Moy et al., 2014).

Several regulatory mechanisms have been proposed for DDX17. DDX17 and its close homolog DDX5 share high sequence similarity and may have overlapping roles (Fuller-Pace, 2013a; Jalal et al., 2007; Janknecht, 2010). DDX17 seems to self-associate and bind DDX5 from coimmunoprecipitation studies, but the molecular details of homotypic interactions are yet unclear (Ogilvie et al., 2003). DDX17 can also be modulated via subcellular localization. Interactions of DDX17 with YAP downstream of Hippo signaling can alter nuclear access—and thus pri-miR processing—at different cell densities (Mori et al., 2014). Finally, the N-terminal extension of DDX17 can be modified with SUMO, which can regulate protein stability and interactions with other transcription factors (Mooney et al., 2010).

To gain insight into the molecular mechanisms underlying substrate specificity, remodeling activity, and regulation of DDX17, we studied the biochemical activity of DDX17 in the context of its tertiary structure. Here, we present high-resolution crystal structures of the helicase core domains of DDX17 in free form and bound with RNA and an ATP analog, ADP-BeF_3^-. We identify several mutations of DDX17 that affect ATP hydrolysis or RNA binding, some of which have been associated with cancer. We show that RNA binding but not ATPase activity is critical for DDX17 to enhance the processing of pri-miRs *in vitro*, by remodeling the 3′ flanking region. In addition to mapping the DDX17 binding site on a pri-miR, we use multiple crystal structures to show that the core catalytic domains of DDX17 prefer to bind an RNA containing the sequence “RCAYCH.” Furthermore, we show that an N-terminal region of DDX17 serves to autoregulate the ATPase activity and domain stability, raising the possibility of modulating DDX17 through post-translational mechanisms.

RESULTS

Crystal Structure of DDX17 Catalytic Core

DDX17 contains RecA-like DEAD and C-terminal helicase (cHelicase) domains, with many conserved motifs including the DEAD motif (Figure 1A). Among DDX5 and DDX17 homologs, the N-terminal region also shares significant sequence similarity (Figure S1A). As a starting point for biochemical and structural studies of DDX17, we purified recombinant full-length human DDX17 from *E. coli* and insect cells. Irrespective of the expression host, the purified polypeptides exhibit similar ATPase activities (Figure S1B). Therefore, for most studies involving recombinant proteins, we relied on those expressed in *E. coli.* Recombinant DDX17 polypeptides have the ability to unwind an asymmetric RNA duplex (Jankowsky and Fairman, 2008), regardless of whether the duplex contains a 3′ or 5′ overhang (Figures S1C and S1D). A double-stranded RNA (dsRNA) lacking any overhang is resistant to unwinding, indicating that the single-stranded RNA (ssRNA) region is necessary for the unwinding activity of DDX17 (Figure S1E) as seen for certain DEAD-box helicases (Yang and Jankowsky, 2006). To determine whether purified DDX17 can self-associate *in vitro*, we used size exclusion chromatography combined with multi-angle light scattering (SEC-MALS) (Figure 1B). The measured solution molecular weight of full-length DDX17 corresponds to the theoretical size of the monomer, indicating that DDX17 does not dimerize in the absence of other factors.

Despite numerous attempts, we were unable to crystallize full-length DDX17. A shorter construct that contains the two predicted RecA-like domains and some of the N-terminal region (which we refer to as “Core”; Figure 1A) is also monodisperse and monomeric in solution (Figure 1C). We were able to crystallize the Core construct both in apo-state, and bound to an RNA oligonucleotide (uridine decamer, rU10) and an ATP analog, ADP-BeF_3^-. The relative orientation of the catalytic domains changes dramatically upon binding RNA and nucleotide substrates (Figures 1D and 1E). Comparing the free DDX17 Core structure to other DEAD-box proteins crystallized without substrates shows that the relative orientation between DEAD and cHelicase domains vary widely, likely due to the flexible inter-domain linker (Figure 1F).

A high-resolution structure of DDX17 core in complex with RNA and ADP-BeF_3^- allows us to examine the contribution of each residue lining the binding sites to substrate specificity. While the Q motif makes important hydrogen bonds with the adenine nucleobase, the motifs I, Va, and VI help to recognize the ribose sugar geometry and create a pocket for the triphosphate (Figure 2A). The charged phosphates are balanced with basic side chains as well as a magnesium ion coordinated by the DEAD motif. Less elaborate hydrogen bonding networks are observed for the RNA binding pocket. The uridine polymer mostly forms favorable electrostatic contacts through the phosphate backbone, and only six uridines out of the 10 included in the experiment are ordered enough to yield discernable electron density (Figure 2B). Therefore, we have numbered the uridine positions (1 through 6) arbitrarily starting from the first visible nucleotide in the RNA oligonucleotide.

Several point mutations in DDX17 have been found in cancers, such as those in the large intestine, bladder, stomach, brain, and skin (Cai et al., 2017; Fuller-Pace, 2013b; Robert and

Pelletier, 2013) (Figure S2A). Some of the mutations involve residues that make direct contact with ATP, such as R432C in motif VI (colorectal carcinoma samples: COSMIC Project COSP44021; and skin cutaneous melanoma: TCGA; US; study ID: COSU540) and T143A in motif I (bladder cancer samples; study ID: COSU581) (Guo et al., 2013). Introducing the cancer-derived mutations to recombinant DDX17 causes a dramatic decrease in its ability to hydrolyze ATP (Figure 2C). A E247Q/K142A double mutation in DDX17 completely abolishes its ATPase activity. Another set of cancer-associated mutations lie in the RNA-binding groove. Introducing single-point mutations at the amino acids directly interacting with RNA—R250G, R175G, G371R, or R378T (all patient mutations)—decreases RNA-dependent ATPase activity of full-length DDX17, although not as dramatically as mutations that interfere with ATP binding (Figure 2C).

To test the direct impact on RNA affinity, we performed electrophoretic mobility shift assays (EMSAs) using mutant DDX17 polypeptides. When an RNA-binding groove mutation is introduced into the DDX17 Core, the affinity for RNA decreases dramatically (Figure S2B). In contrast, the dissociation constant (K_D) for RNA is not significantly affected when the same amino acid substitutions are introduced into the full-length DDX17 constructs (Figure S2C). To determine how different parts of DDX17 contribute to RNA binding, we extended the DDX17 Core construct to include the N- or C-terminal regions (N-Core and Core-C, respectively; Figure S3A). Full-length DDX17 binds an overhang-containing RNA duplex with a K_D of ~310 nM (Figure S3B). Core-C exhibits similar affinity for the same RNA, but the difference in gel-shift patterns suggests that the binding modes may be different. The N-terminal region may not contribute significantly to RNA binding, because there is little difference between Core and N-Core. As evident from the binding assays, the Core in isolation has weak affinity for RNA in comparison to the C-terminal region. For a dsRNA with no overhang, none of the DDX17 constructs shows detectable binding affinity (Figure S3C). However, ssRNA binds DDX17 with high affinity, and the tight affinity requires the presence of the C-terminal tail of DDX17 (Figure S3D). Considering the smeary RNA-protein complex bands in EMSAs and the low complexity of the polypeptide sequence, we propose that the C-terminal region has a high affinity for ssRNA with little discretion for nucleotide sequence.

We then tested how different types of DDX17 mutations affect the unwinding activity (Figure 2D). In contrast to the robust unwinding activity of wild-type DDX17, mutant full-length DDX17 constructs that show diminished ATPase activity also manifest loss of unwinding activity, regardless of whether the mutations are in the ATP- or RNA-binding regions. Therefore, for both ATPase activity and unwinding activity, DDX17 requires proper interactions with both RNA and ATP as visualized in the crystal structures.

DDX17 Regulates pri-miR Processing by Remodeling the 3′ Flanking Region

Multiple groups have shown that DDX17 associates with Drosha and DGCR8 to regulate pri-miR processing, which is an obligatory maturation step for most microRNAs (Dardenne et al., 2014; Gregory et al., 2004; Mori et al., 2014; Remenyi et al., 2016). To gain insight into the mechanism through which DDX17 affects pri-miR processing, we tested how various DDX17 mutations affect Drosha cleavage of pri-miR-125a. Using a previously

established reconstituted pri-miR processing assay with recombinant Drosha and DGCR8 (Partin et al., 2017), we can recapitulate the enhancing effect of wild-type ADDX17 on the processing of pri-miR-125a (Figures 3A and S4A). When the E247Q/K142A double mutation is introduced, thus abrogating ATPase activity, DDX17 retains most of its ability to enhance Drosha activity. Conversely, mutations that impair RNA binding dramatically diminish DDX17-mediated enhancement of pri-miR processing. Pri-miR-125a contains a CNNC motif, similar to the ones associated with DDX17-regulated miRNAs (Mori et al., 2014) and efficient processing (Auyeung et al., 2013). To test how DDX17 affects other pri-miRs, we tested another pri-miR that contains a CNNC sequence (pri-miR-18a), as well as another pri-miR lacking the motif (pri-miR-335) (Figures 3B, 3C, S4B, and S4C). Similar to pri-miR-125a, wild-type and catalytically inactive DDX17 can both enhance processing of pri-miR-18a. However, we do not observe a similar degree of processing enhancement for pri-miR-335 (Figure 3D). Therefore, DDX17 enhances processing of a subset of pri-miRs con taining a CNNC motif, and while ATPase activity is dispensable for this enhancement, proper RNA binding is critical.

To further investigate the effector function of DDX17 on pri-miR processing, we studied the interaction of the DDX17 Core with pri-miR-125a in more detail. We focused on the Core, because the C-terminal tail binds nonspecifically to RNA with high affinity (Figure S3). First, as a way to map DDX17 Core binding site on pri-miR-125a, we truncated the RNA. Pri-miRs with longer arms (the 5′ and 3′ extensions that are cropped by Drosha) can bind DDX17 Core, but truncating the 5′ and 3′ arms down to 10 and 8 nucleotides, respectively, abrogates detectable binding (Figures 3E–3G). To identify the DDX17 Core binding site in the larger RNA context, we performed a footprinting assay using hydroxyl radical protection (Figures 3H, S5A, and S5C). We observe a protected region on the 3′ flanking arm of pri-miR-125a with wild-type Core domains, and the protection is lost when the RNA-binding mutation G371R is incorporated into the DDX17 Core construct. The 3′ region of pri-miR-125a is relatively rich in cytidines, and multiple “CNNC” motifs can be found (Auyeung et al., 2013). Remarkably, the most protected region contains “ACACCU,” which resembles the motif enriched among the DDX17-dependent miRNAs in another study, “VCAUCH” (Mori et al., 2014).

To investigate the effect of DDX17 on pri-miR-125a structure, we performed selective hydroxyl acylation followed by primer extension (SHAPE) analysis (Wilkinson et al., 2006) in the presence and absence of full-length DDX17 (Figures 3I, S5B, and S5D). SHAPE reactivity in the 3′ flanking region increases upon binding DDX17, suggesting that the helicase induces a more single-stranded, flexible RNA conformation. Similar to the footprinting assay, introducing the G371R mutation into DDX17 results in no significant change in SHAPE reactivity. Therefore, an RNA-binding mutation, G371R, interferes with DDX17 localizing to a particular site in the 3′ region where it usually remodels the RNA to become more single-stranded. The same mutation also abolishes the ability of DDX17 to enhance pri-miR processing (Figure 3D), suggesting that the RNA-binding and remodeling activities of the helicase are directly linked to its impact on Microprocessor efficiency.

RMFQ Groove of DDX17 Core Binds RNA Sequences Containing RCAYCH

The specific binding of pri-miR-125a by the DDX17 Core (Figure 3) led us to postulate that the catalytic domains may recognize a particular RNA sequence. As discussed above, the C-terminal region of DDX17 binds tightly to ssRNAs with no apparent sequence specificity. To understand how the C-terminal region and the Core together recognize target RNAs, it is important to determine the intrinsic specificity of the Core in isolation. Therefore, we determined multiple structures of the Core in complex with various RNA oligonucleotides (Figure 4A).

When we co-crystallize DDX17 Core with 125a-oligo1, a fragment of pri-miR-125a containing the DDX17 footprint—"ACACCU"—we observe additional protein-RNA contacts in comparison to the complex containing rU10, particularly through a groove formed by the four residues: R250, M254, F256, and Q259 (RMFQ groove) (Figure 4B). The RMFQ groove binds the "ACAC" portion of 125a-oligo1 more tightly than rU10, mostly due to shape complementarity. The closer contact between RNA and protein is especially apparent when comparing the crevice that narrows between the two molecules with a change in RNA sequence (Figure 4C). When we compare the RMFQ region between the free and substrate-bound states, a significant conformational change is observed, suggesting an induced folding event as the protein-RNA complex forms (Figure S6A).

It is possible that the "ACAC" motif from 125a-oligo1 coincidentally interacts with the RMFQ groove, due to where it lies with respect to the ends of the oligonucleotides, potentially due to crystal packing. To test the ability of the DDX17 Core to recognize the "ACAC" motif, we shifted its position by one nucleotide to generate 125a-oligo2 (Figure 4A). Remarkably, the location of the oligonucleotide with respect to the domain also shifts, to maintain the same favorable interactions between "ACAC" and RMFQ (Figure 4D). Thus, the preferred register of the RNA sequence in the protein groove is visible by comparing the crystal structures. We also co-crystallized a fragment from another DDX17-sensitive miR, pri-miR-18a (Figure 4E). We observe a similar motif, "UCAU," to align at the location where "ACAC" binds. In summary, through multiple crystal structures, we can observe the intrinsic sequence preferences of the DDX17 Core.

To test the importance of the specific RNA interactions, we used EMSAs to measure the affinity of the DDX17 Core for different RNA oligonucleotides (Figure 4F). Although we were able to crystallize rU10 with DDX17 core, it is difficult to detect the binding affinity of DDX17 for rU10 or rA10 *in vitro*. However, when a more preferred sequence is included, there is a dramatic increase in binding affinity. When we compare the structures with 125a-oligo1 and 18a-oligo1, having a smaller nucleobase (uracil) in position 1 prevents favorable interactions with M254 (Figure 4E). Loss of a favorable contact may explain the slightly higher K_D for 18a-oligo1. Taken together, the biochemical and structural data help us redefine the preferred binding site of DDX17 from "VCAUCH" (Mori et al., 2014) to "RCAYCH."

Since the RMFQ groove directly binds RCAYCH, we also tested the effect of protein mutations on the binary interaction, using 125a-oligo1 (Figure 4G). Mutating the RMFQ motif results in a significant loss of binding affinity, supporting our hypothesis that the

RMFQ motif is critical to recognize RCAYCH. Notably, the R250G mutation has been reported in low-grade glioma patients (study ID: COSU545). The RMFQ residues are conserved in the DDX5/DDX17 homologs (Figure 4H) but not in most of the other DEAD-box proteins (Figure S6B), suggesting that the RNA sequence preference is likely to change to match the distinct biological roles they play. In summary, we conclude that DDX17 Core participates in direct binding and sequence recognition of the target RNA, RCAYCH, through the RMFQ groove.

DDX17 N-Terminal Extension Regulates ATPase Activity

Our DDX17 Core structure includes an N-terminal extension with an unusual conformation in comparison to other DEAD-box helicase structures. When the DEAD domains of these structures are superimposed onto each other, the N-terminal extensions protrude in various directions. DDX17, its close relative DDX5, and Prp5 appear to have similarly oriented N-terminal extensions, but in DDX17 we observe that it traverses the entire length of the DEAD domain, reaching near the RNA substrate binding area (Figure 5A). The N-terminal tail is mostly in an extended conformation and makes favorable interactions with the surface of the DEAD domain on the opposite side of the cHelicase-binding face. Though mostly lacking in secondary structure, the conformation of the bound N-terminal tail is similar in both the open and closed structures of DDX17 Core, suggesting that the intramolecular contacts are sufficient to stabilize the specific peptide conformation observed in the crystals (Figure 5B). Favorable hydro-phobic interactions are observed, and larger aromatic residues such as F55 and Y56 become buried. Moreover, Y56 is stabilized by the network of hydrogen bonds formed with two nearby waters. The binding site for the N-terminal peptide on the DEAD domain is well conserved in the DDX5/17 family (Figures 5C and 5D), suggesting that the intramolecular interaction may have a physiological role.

To determine how the interaction between the N-terminal tail and the DEAD domain affects DDX17 function, we introduced point mutations to disrupt key residues at the interface. When we purify DDX17 from insect cells, the N-terminal tail can undergo phosphorylation at Y56, as shown by tandem mass spectrometry (Figure 6A). Although its role is yet unclear, phosphorylation of the same residue has also been found in human lung cancer cells in at least one reported study (Tsai et al., 2015). A phosphorylated tyrosine is difficult to mimic with a natural amino acid, but we generated Y56E and Y56A substitutions to study the role of Y56. Similar mutations were also introduced to other residues in the N-terminal tail that make important contacts with the rest of DDX17. To determine how the mutations affect protein stability, we used differential scanning fluorimetry. Most mutations lead to a significant decrease in protein stability, as indicated by lower melting temperatures upon thermal denaturation (Figure 6B). Using the most destabilizing mutation, Y56E, we also tested how disrupting the intramolecular binding interface affects the ATPase activity. Introducing the Y56E mutation causes a slight increase in enzymatic efficiency, mostly because of an increase in ATP turnover (Figures 6C and 6D). Accordingly, the unwinding activity of DDX17 is also enhanced with the Y56E mutation (Figure 6E). Neither of the changes in enzymatic activity is due to a change in RNA affinity because EMSAs show similar binding constants for wild-type and mutant proteins (Figure S7A). In summary,

biochemical studies suggest that the N-terminal tail can modulate the ATPase activity of DDX17.

We also investigated how the N-terminal tail affects pri-miR processing. When Y56 is mutated to Glu or Ala, the mutant DDX17 polypeptides can enhance Drosha activity as much as the wild type (Figures 6F, S7B, and S7C). This is consistent with our results that show that the ATPase activity of DDX17 is not important for the ability of DDX17 to enhance pri-miR processing *in vitro* (Figure 3). Furthermore, we tested how the DDX17 mutations affect microRNA biogenesis in cells. We co-transfected expression constructs of pri-miR-125a along with wild-type or mutant DDX17 constructs into HEK293T cells. Representative mutations for DDX17 were chosen to target the ATPase (E247Q/K142A), N-terminal tail (Y56A), and RNA-binding (G371R) regions. Unlike *in vitro* processing assays, we observe that all three types of mutations can affect the levels of mature miRs and/or pri-miRs (Figures S7D–S7G). Thus, consistent with previous reports, DDX17 is likely to have roles in addition to affecting pri-miR processing. Our in-depth evaluation and categorization of various mutations *in vitro* will help future studies to dissect how RNA-binding and ATPase activities of DDX17 affect its multiple cellular functions.

DISCUSSION

DDX17, like many other DEAD-box RNA helicases, is not only known for its role in many aspects of RNA biology, but also for its association with multiple cancers. Using structure-guided functional assays, we show how various cancer-derived mutations can affect different aspects of DDX17 function, such as RNA-binding, ATP hydrolysis, and unwinding activity. Although many of the DEAD-box proteins have dedicated functions with a specific set of RNAs they act upon, little is known about how DEAD-box proteins recognize their target RNAs. Here, we show that DDX17 has different modules with dedicated RNA specificity: the C-terminal region of DDX17 has a high affinity for single-stranded RNA with no apparent sequence preference, while the ATPase Core shows a preference for sequences containing RCAYCH (Figure 7). Moreover, we show that the ability of DDX17 Core to recognize a specific sequence within pri-miRs is important for its ability to remodel the 3′ flanking region, thus enhancing Drosha processing. Although ATP hydrolysis does not seem to be required for improving pri-miR processing, DDX17 has a robust RNA-activated ATPase activity. We show that an unusual N-terminal extension of DDX17 visible in crystal structures can modulate the ATPase activity and unwinding activity. Together, our study provides a mechanistic model for substrate specificity and regulation of DDX17, while also providing a molecular explanation for how DDX17 can enhance processing of a subset of pri-miRs.

Unlike most DEAD-box proteins, DDX17 function depends on the ability of its Core to recognize a preferred sequence motif, RCAYCH. The intrinsic binding affinity of the Core for RNA is rather weak. Therefore, the C-terminal tail is likely to help localize DDX17 in the vicinity of the target, so that the DEAD and cHelicase domains can act on a specific site that requires remodeling. The DDX17 Core can crystallize with a poly-U sequence, similar to several other DEAD-box helicases, although poly-U is not a preferred binding site. However, with a cognate RNA sequence, we observe more favorable contacts in the

structures and higher binding affinity biochemically. Such cryptic specificity may also play a part in other DEAD-box helicases.

The role of DDX17 as an enhancer of pri-miR processing has been established in various cellular contexts (Kao et al., 2019; Mori et al., 2014; Remenyi et al., 2016). Using a reconstituted system, we provide a mechanistic model for how DDX17 enhances the maturation of certain miRs. DDX17 directly and specifically binds the pri-miR 3′ region containing "RCAYCH" via the RMFQ groove. The impact on the pri-miR is an increase in SHAPE reactivity—and thus likely also RNA flexibility—in the 3′ flanking region, which renders the RNaseIII cleavage of the pri-miR more efficient. Because DDX17-dependent Microprocessor activity is not affected by mutations that obliterate ATPase activity, DDX17 may affect pri-miRs through clamping a specific RNA segment. Our findings on DDX17 specificity in the context of pri-miRs *in vitro* are similar to the previously proposed consensus sequence among the miRs that depend on DDX17 for maturation *in vivo*, which allows us to link our *in vitro* binding and processing results to the extensive work that has already been done in intact cells. "RCAYCH" in the 3′ flanking region overlaps with the "CNNC" motif in 3′ ends of pri-miRs, which was suggested to recruit SRSF3 in a separate study (Änkö et al., 2012; Auyeung et al., 2013; Kim et al., 2018). DDX17 may mediate the interaction between SRSF3 and pri-miRs, or DDX17 and SRSF3 may act on a similar RNA region independently. Our reconstitution studies show that recombinant DDX17 is sufficient to affect processing of certain pri-miRs by Drosha, in the absence of SRSF3. How SRSF3 affects the described function of DDX17 is yet unclear.

Various DEAD-box proteins can be regulated via additional domains or regions tethered to the core catalytic domains, usually at the N- or C-terminal ends (Rudolph and Klostermeier, 2015). We present evidence that the N-terminal tail of DDX17 has specific, extensive interactions with the DEAD domain of DDX17, and interfering with it can modulate the ATPase activity. Even though the ATPase activity is not important for enhancing microRNA biogenesis *in vitro*, it is critical for unwinding activity and thus likely important for some other cellular roles of DDX17. Our studies of various specific mutations that only affect one aspect of DDX17 activity will be useful to dissect various molecular mechanisms of DDX17 on other target RNAs. The N-terminal tail of DDX17 is reminiscent of that of DDX5 and Prp5, although we observe more extensive interactions in DDX17. Although the N-terminal tail can affect the catalytic activity and/or interactions with other factors in DDX5 and Prp5 (Dutta et al., 2012; Zhang et al., 2013), further study will be required to compare how the N-terminal regions affect each helicase. DDX17 is also known to be sumoylated in the N-terminal region (Mooney et al., 2010). Thus, we propose that post-translational mechanisms may be utilized to regulate DDX17 ATPase activity, by changing how the N-terminal peptide interacts with the catalytic domains.

STAR★METHODS

LEAD CONTACT AND MATERIALS AVAILABILITY

The Lead Contact is Yunsun Nam (yunsun.nam@utsouthwestern.edu). Any questions should be directed to the Lead Contact. Any reagents generated in this study are available from the Lead Contact with a completed Materials Transfer Agreement.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Bacterial strains: Expression of recombinant human DDX17 was performed using Rosetta 2 (BL21-DE3)(Novagen) cells, genotype: F^{-}ompT $hsdS_B(r_B^{-}\ m_B^{-})$ gal dcm pRARE2 (Cam^R).

Cell Lines. Sf21 cells (Expression Systems) (cultured in Grace's Supplemented Media (Expression Systems) with 10% v/v fetal bovine serum (Sigma) at 27°C for expression of Baculovirus. HighFive cells (ThermoFisher) were cultured in ESF 921 insect cell culture medium (Expression Systems) at 27°C for expression of DDX17 protein using the baculovirus system.

HEK293T cells (ATCC) were cultured in Dulbecco's modified eagle medium (DMEM) supplemented with 10% (v/v) FBS (Sigma).

METHOD DETAILS

Materials—Oligos rU10, rA10, 16-mer-3OV, 41-mer-3OV, 16-mer-5OV, 41-mer-5OV, 125a-oligo1, 125a-oligo2, and 18a-oligo1 were custom synthesized (IDT). ds41–16nt-3OV and ds41–16nt-5OV were prepared by annealing 16-mer-3OV with 41-mer-3OV and 16-mer-5OV with 41-mer-5OV, respectively, at a ratio of 1.2. The sequence of the RNA substrate is from Jankowsky and Fairman (2008)

16-mer-3OV: 5′-AGCACCGUAAAGACGC-3′

41-mer-3OV: 5′-GCGUCUUUACGGUGCUUAAAACAAAACAAAACAAAACAAAA-3′
16-mer-5OV: 5′-AGCACCGUAAAGACGC-3′

41-mer-5OV: 5′-AAAACAAAACAAAACAAAACAAAAUGCGUCUUUACGGUGCU-3′

Adenylyl-imidodiphosphate (AMP-PNP), adenosine 5′-diphosphate (ADP), adenosine 5′-triphosphate (ATP), beryllium chloride ($BeCl_2$), and sodium fluoride (NaF) were purchased from Sigma.

***in vitro* RNA transcription and purification—**DNA templates for pri-miR-125a, pri-miR-18a and pri-miR-335 constructs were cloned into the pRZ vector containing the self-cleaving ribozymes hammerhead (HH) and hepatitis delta virus (HDV) on the 5′ and 3′ ends, respectively, to produce homogeneous ends (Walker et al., 2003). *In vitro* transcription was carried out at 37°C under the following conditions: 4 mM rNTPs, 35 mM $MgCl_2$, 20 mM Tris pH 8.0, 1 mM spermidine, 0.01% Triton X-100, 10 mM DTT and 100 U RNase inhibitor (Thermo) with 2 μg of DNA template in the final volume of 500 μL reaction. Target RNA was purified by denaturing PAGE.

Protein purification—DDX17 Core and WT were expressed in Rosetta 2 (BL21-DE3) cells (Novagen) using autoinduction media (Studier, 2014) at 20°C. DDX17 Core was purified through nickel affinity, cation-exchange and size-exclusion chromatography steps. The final buffer contained 20 mM Tris pH 8.0, 500 mM NaCl, 5 mM DTT and 2 mM $MgCl_2$. Full-length DDX17 was purified through nickel affinity, cation exchange and size-exclusion chromatography (in a final buffer containing 20 mM Tris pH 8.0, 1M NaCl, 5 mM

DTT, 2 mM $MgCl_2$, 200 mM Arginine pH 8.5). All variants were purified using the same protocol as their wild-type counterparts. Full-length DDX17 was also expressed using the Bac-to-Bac system (ThermoFisher) in HighFive cells. Protein from insect cells was purified using the same protocol as their bacterial equivalents.

Microprocessor purification was performed as described previously (Partin et al., 2017). Briefly, Drosha/DGCR8 complexes were co-expressed using Bac-to-Bac system in HighFive cells. Proteins were purified using Ni-NTA affinity, cation-exchange and size-exclusion chromatography.

RNA Unwinding Assay—Unlabeled single-stranded 16nt RNA was mixed with ^{32}P-5′ end labeled single-stranded 41nt RNA with a molar ratio of 1:1.2. The reaction was incubated at 95°C for two minutes and cooled to room temperature for two hours. The radiolabeled duplex ds41–16nt was then purified by native PAGE. ^{32}P-5′ end labeled RNA duplex ds41–16nt was incubated with proteins in a buffer containing 20 mM Tris pH 7.5, 10 mM $MgCl_2$, 2 mM DTT, 50 mM KCl and 5 mM ATP for 30 minutes. The reaction was quenched by adding one volume of stop buffer (15 μl) containing 50 mM EDTA and 1% SDS, followed by 4 μL of Proteinase K (3 mg/ml). The samples were then analyzed by native PAGE.

Size-exclusion chromatography–Multiangle light scattering—Purified full-length DDX17 protein was loaded onto a Superdex 200 Increase 10/300 column (GE) on Agilent 1200 Infinity series HPLC machine in the same buffers described for purification above. DAWN HELIOS II and tREX detectors (Wyatt) were used to monitor scattering and differential refractometry. Astra 6.1 (Wyatt) was used to process the data. The results were fitted using the Zimm model to obtain molar mass.

Electrophoretic Mobility Shift Assay—^{32}P-5′ end labeled RNA duplex ds41–16nt (same as unwinding assay) and pri-miR-125a were mixed with varying concentrations of proteins in the binding buffer containing 20 mM Tris pH 7.5, 2 mM DTT, 12 mM $MgCl_2$, 100 mM NaCl, 2 mM AMP-PNP, 50 ng/μl yeast tRNA and 10% glycerol. The complexes were separated by native PAGE and visualized using a phosphorimager.

***in vitro* ATP hydrolysis Assay—**ATPase assays were performed using the ADP-Glo Max Assay kit (Promega). All reactions were performed in triplicates, in a buffer containing 20 mM Tris pH 7.5, 12 mM $MgCl_2$, 1 mM DTT and 40 mM NaCl for 30 minutes in the presence of 1.7 μM of dsRNA 41–16nt and 5 mM ATP. For steady-state kinetic analysis, the proteins were incubated with different ATP concentrations ranging from 2.5 mM to 156.25 μM, at saturating substrate concentration (2.4 μM dsRNA 41–16nt) for 5–20 minutes at room temperature. Data from ATP-dependent experiments (K_m and K_{cat}) were calculated using the nonlinear regression method, using PRISM.

Differential Scanning Fluorimetry—Proteins (final concentration 1 μM) were incubated with 6X final SYPRO Orange concentration (Thermo Fisher) in a buffer containing 20 mM Tris pH 8.5, 1M NaCl, 10 mM $MgCl_2$, 5 mM DTT, 5% Glycerol. Reactions were performed using the LightCycler 480 System (Roche) with the excitation

wavelength set to 470 nm and detecting emission at 557 nm. The fluorescence intensity was plotted as a function of temperature.

Crystallization and Data Collection—All proteins were crystallized via the hanging drop vapor-diffusion method by mixing equal volumes of protein solution and crystallization buffer. Apo Core at a final concentration of 1 mM with 2 mM AMP-PNP formed crystals after 8 days from the condition 24% PEG, 0.26M NaSCN. To obtain crystals of complex Core·rU10·ADP-BeF_3^-, 0.62 mM of Core was incubated with 1.2 mM rU10 and 5 mM ADP-BeF_3^- at room temperature for 30 minutes before setting up crystallization trials. Crystals grown from the condition 0.1M sodium citrate pH 6.0, 2M NaCl were used for data collection. Crystals of Core with 125a-oligo1 and 125a-oligo2 were grown from condition 0.1 M HEPES pH 7.5 and 40% PEG200 after several days. Crystals of Core with 18a-oligo1 were obtained from the condition 0.1M sodium acetate pH 4.6, 0.1M sodium chloride, 14% MPD, 10 mM $CaCl_2$. 30% glycerol was used as cryoprotectant. Diffraction data were collected at 100K on beamline APS ID-19 at the energy of Se-K edge (λ = 0.9792 Å) and processed with HKL3000. Data collection statistics are summarized in Table S1.

Structure determination and refinement—The crystal structure of the apo Core was solved by molecular replacement using Phaser (McCoy et al., 2007) with individual recA domains from the *Drosophila* Vasa structure (PDB: 2DB3) (Sengoku et al., 2006) as the search model using Phaser implement in Phenix suite. The crystal structure of Core complexed with ssRNA rU10 and ADP-BeF_3^- was solved by molecular replacement, using the individual domains from the apo DDX17 Core structure as separate search models. To build the model, Autobuild was used initially and followed by manual building by Coot (Emsley et al., 2010). Core·U10·ADP-BeF_3^- was used as a search model to solve other Core·RNA·ADP-BeF_3^- complexes. Models were refined using Phenix.refine (Afonine et al., 2012) and validated using MolProbity (Chen et al., 2010). Detailed refinement statistics are shown in Table S1.

***in vitro* pri-miR processing assays—**5′ end-labeling was carried out using T4 Polynucleotide Kinase (NEB) and γ–^{32}P-ATP. Pri-miRNA Microprocessor cleavage was performed in 15 μL reactions containing 30 mM Tris pH 7.5, 50 mM NaCl, 5% glycerol, 10 mM $MgCl_2$, 5 mM DTT, 2 mM ATP, 8 U RNase inhibitor (Thermo) and 0.4 μg yeast tRNA. Final concentration of Microprocessor and DDX17 were 4 nM and 130 nM, respectively.

Hydroxyl radical footprinting—40 nM pri-miR-125a and proteins ($Core^{WT}$ and $Core^{G371R}$) at varying concentrations were incubated in a buffer containing 10 mM HEPES pH 7.5, 40 mM NaCl, 5 mM $MgCl_2$, 1 mM ADP-BeF_3^-, 8 U RNase inhibitor and 32 ng of yeast tRNA for 30 min at 30°C. The hydroxyl radical mixture was prepared by quickly mixing 5.3 mM Fe(II) sulfate hexahydrate with 16 mM EDTA, 2% H_2O_2 and 63 mM ascorbic acid. After a one-minute incubation at room temperature, the reaction was quenched with 200 μL of 300mM sodium acetate pH5.0, 20 mM EDTA and RNA was extracted with acid phenol-chloroform. Total RNA was then reverse transcribed using a ^{32}P-labeled DNA primer specific for pri-miR-125a. The radioactive products were visualized using the sequencing gel electrophoresis. For quantitation, band intensities were measured

using Image Lab (Bio-Rad). H_{Sample} is defined as the extent of hydroxyl radical protection of each nucleotide in each sample, calculated using the formula: $H_{Sample} = (I_{Sample} - I_{Buffer}) / I_{Buffer}$. I_{Sample} and I_{Buffer} are the band intensities of the particular nucleotide, in the sample lane with protein or the buffer-only lane, respectively.

Selective hydroxyl acylation followed by primer extension (SHAPE)—40 nM pri-miR-125a and proteins (625 nM of full-length DDX17 and the full-length mutant G371R, 5 μM of Core) were incubated in a buffer containing 20 mM Tris pH 7.5. 66 mM NaCl, 0.5 mM DTT, 10 mM $MgCl_2$, 8U RNase inhibitor, 4 mM ATP and 25 ng of yeast tRNA for 30 min at 30°C. RNA was modified by incubating the pre-assembled protein-RNA complex with 100 mM NAI (EMD Millipore) at 37°C for five minutes. The reaction was quenched with 200 μL of 300mM sodium acetate pH5.0 and 20 mM EDTA and total RNA was extracted using acid phenol-chloroform (Thermo). SuperScript III reverse transcriptase (Thermo) was used for reverse transcription using ^{32}P-labeled primers for pri-miR-125a and the product DNA was subjected to sequencing gel electrophoresis. Dideoxynucleotides (ddATP, ddCTP, ddGTP, or ddTTP) were used to generate the sequencing ladder. The radioactive products were visualized using a phosphorimager. For quantitation, band intensities were measured using Image Lab (Bio-Rad). SHAPE reactivity of each sample was calculated from the differences in absolute intensities between NAI-treated and DMSO-control. S_{Sample} is the SHAPE reactivity of each nucleotide in each sample, calculated by the formula: $S_{Sample} = (I_{NAI} - I_{DMSO}) / I_{DMSO}$. I_{NAI} and I_{DMSO} are the band-intensities of the particular nucleotide in the sequence for the NAI-treated sample or the DMSO-treated sample, respectively. The difference in SHAPE reactivity, calculated by subtracting the solvent signal (S_{Sample} - S_{Buffer}) was plotted for each nucleotide.

Cell Culture and quantitative real-time PCR (qPCR)—HEK293T cells were cultured using Dulbecco's modified eagle medium (DMEM) containing 10% FBS (Sigma). Transient co-transfections (1 μg of plasmid expressing DDX17 and 1.5 μg of plasmid expressing pri-miR-125a) were performed using Lipofectamine3000 (Thermo) in 60mm dishes according to the instructions provided by the manufacturer. 48 hours after transfection, total RNA was extracted using Trizol (Thermo). RNA samples were DNase-treated, and reverse transcription was performed in 20 uL reactions containing 1x First Strand buffer (Thermo), 10 mM DTT, 0.5 mM dNTPs, 8U Ribolock (Thermo), 50U SuperScript II (Thermo), 100 ng DNase-treated total RNA, and 150 nM reverse transcription primer. qPCR was performed in 10 μL reactions on a LightCyclerII (Roche) containing 1x HF buffer (Thermo), 200 nM dNTPs (NEB), 0.5x SYBR Green I (ThermoFisher), 100 nM forward and reverse primers, 2% DMSO, and 0.4 U Phusion DNA polymerase (NEB). Cycle times (Ct) for pri-miR-125a were normalized to U6 expression levels, and error bars were calculated using propagation of uncertainty analysis. Statistics were carried out with two biological replicates (three technical replicates each), using paired Student's t test (two tailed).

Splinted ligation small RNA detection assays—1 μg of total RNA was annealed with 1 pmol of ^{32}P-end-labeled ligation oligonucleotide and 1 pmol of a bridge nucleotide specific to the target miR-125a, in a reaction containing 75 mM KCl and 20 mM Tris pH 7.5. The sample was then incubated in a ligation reaction containing 30 mM KCl, 75 mM

Tris pH 7.6, 10 mM MgCl2, 1 mM DTT, 0.5 mM ATP, 7.5% w/v PEG 6000, and 2 U T4 DNA ligase (NEB). After incubation for 1 hour at room temperature, the reactions were stopped by adding 10 units of alkaline phosphatase and incubated at 37 °C for 15 min. The ligated samples separated by denaturing PAGE and visualized by autoradiography followed by quantitation using ImageLab. Statistics were carried out with two biological replicates (two technical replicates each), using paired Student's t test (two tailed).

Northern blotting—Total RNA (5 μg) mixed with equal volume of 2x formamide RNA loading dye was resolved using denaturing polyacrylamide-urea gel. RNA was transferred to Hybond-N membrane (GE) in TBE buffer using a Trans-Blot cell (Bio-Rad) followed by UV-crosslinking. Hybridization was performed in a buffer containing 0.5 M sodium phosphate pH 7.2, 7% SDS, and 1 mM EDTA with 10 ng of radiolabeled probe at 42 °C overnight. The membrane was washed in washing buffer containing 60 mM sodium phosphate, 1 M sodium chloride, and 6 mM EDTA at 42 °C for 30 min before being exposed to a phosphor screen. The probe for detecting U6 was comprised of the sequence GCAGGGGCCATGCTAATCTTCTCTGTATCG.

QUANTIFICATION AND STATISTICAL ANALYSIS

ATPase activity assays were performed in triplicates, and the error bars represent ± standard error. Differential scanning fluorimetry assays were performed in triplicates, and the errors represent ± standard deviation. Two replicates were performed for the RNA unwinding assays. EMSAs were performed in three replicates. Quantified *in vitro* pri-miR processing assays were performed in triplicates, and the error bars represent ± standard error between the biological replicates. Quantified *in vivo* pri-miR processing assays were performed in triplicates for each of the two biological replicates, and the error bars represent ± standard error between the biological replicates. Splinted ligation were performed in duplicates, for each of the two biological replicates replicates, and the error bars represent ± standard deviation. For both *in vivo* pri-miR processing assay and splinted ligation assay, Student's t test (two-sided) was performed between the biological replicates.

DATA AND CODE AVAILABILITY

The structures of Core-Apo, Core·ADP-BeF3−·rU10, Core·ADP-BeF3−·125a-oligo1, Core·ADP-BeF3−·125a-oligo2, and Core·ADP-BeF3−·18a-oligo1, have been deposited under the following PDB codes, respectively: 6UV0, 6UV1, 6UV2, 6UV3 and 6UV4.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank the support from the Cecil H. and Ida Green Center Training Program in Reproductive Biology Sciences Research and also members of the Structural Biology Laboratory at UT Southwestern for help with synchrotron data collection. Y.N. is a Southwestern Medical Foundation Scholar in Biomedical Research (Endowed Scholar Program at UT Southwestern), a Pew Scholar in the Biomedical Sciences (27339), and a Packard Fellow (2013–39275). This study was supported by grants from the NIH NIGMS (R01GM122960), the Welch Foundation (I-1851–20170325), and the Cancer Prevention Research Institute of Texas (R1221). The use of SBC 19ID

beamline at the Advanced Photon Source is supported by the United States Department of Energy contract DE-AC02–06CH11357.

REFERENCES

Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH, and Adams PD (2012). Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr. D Biol. Crystallogr 68, 352–367. [PubMed: 22505256]

Änkö ML, Müller-McNicoll M, Brandl H, Curk T, Gorup C, Henry I, Ule J, and Neugebauer KM (2012). The RNA-binding landscapes of two SR proteins reveal unique functions and binding to diverse RNA classes. Genome Biol. 13, R17. [PubMed: 22436691]

Auyeung VC, Ulitsky I, McGeary SE, and Bartel DP (2013). Beyond secondary structure: primary-sequence determinants license pri-miRNA hairpins for processing. Cell 152, 844–858. [PubMed: 23415231]

Cai W, Xiong Chen Z, Rane G, Satendra Singh S, Choo Z, Wang C, Yuan Y, Zea Tan T, Arfuso F, Yap CT, et al. (2017). Wanted DEAD/H or alive: helicases winding up in cancers. J. Natl. Cancer Inst 109, djw278.

Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, and Richardson DC (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr 66, 12–21. [PubMed: 20057044]

Cordin O, Banroques J, Tanner NK, and Linder P (2006). The DEAD-box protein family of RNA helicases. Gene 367, 17–37. [PubMed: 16337753]

Dardenne E, Polay Espinoza M, Fattet L, Germann S, Lambert MP, Neil H, Zonta E, Mortada H, Gratadou L, Deygas M, et al. (2014). RNA helicases DDX5 and DDX17 dynamically orchestrate transcription, miRNA, and splicing programs in cell differentiation. Cell Rep. 7, 1900–1913. [PubMed: 24910439]

Dutta S, Gupta G, Choi Y-W, Kotaka M, Fielding BC, Song J, and Tan Y-J (2012). The variable N-terminal region of DDX5 contains structural elements and auto-inhibits its interaction with NS5B of hepatitis C virus. Biochem.J 446, 37–46. [PubMed: 22640416]

Emsley P, Lohkamp B, Scott WG, and Cowtan K (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr 66, 486–501. [PubMed: 20383002]

Fuller-Pace FV (2013a). The DEAD box proteins DDX5 (p68) and DDX17 (p72): multi-tasking transcriptional regulators. Biochim. Biophys. Acta 1829, 756–763. [PubMed: 23523990]

Fuller-Pace FV (2013b). DEAD box RNA helicase functions in cancer. RNA Biol. 10, 121–132. [PubMed: 23353573]

Fuller-Pace FV, and Moore HC (2011). RNA helicases p68 and p72: multifunctional proteins with important implications for cancer development. Future Oncol. 7, 239–251. [PubMed: 21345143]

Gilman B, Tijerina P, and Russell R (2017). Distinct RNA-unwinding mechanisms of DEAD-box and DEAH-box RNA helicase proteins in remodeling structured RNAs and RNPs. Biochem. Soc. Trans 45, 1313–1321. [PubMed: 29150525]

Giraud G, Terrone S, and Bourgeois CF (2018). Functions of DEAD box RNA helicases DDX5 and DDX17 in chromatin organization and transcriptional regulation. BMB Rep. 51, 613–622. [PubMed: 30293550]

Gregory RI, Yan K-P, Amuthan G, Chendrimada T, Doratotaj B, Cooch N, and Shiekhattar R (2004). The Microprocessor complex mediates the genesis of microRNAs. Nature 432, 235–240. [PubMed: 15531877]

Guo G, Sun X, Chen C, Wu S, Huang P, Li Z, Dean M, Huang Y, Jia W, Zhou Q, et al. (2013). Whole-genome and whole-exome sequencing of bladder cancer identifies frequent alterations in genes involved in sister chromatid cohesion and segregation. Nat. Genet 45, 1459–1463. [PubMed: 24121792]

Gustafson EA, and Wessel GM (2010). DEAD-box helicases: posttranslational regulation and function. Biochem. Biophys. Res. Commun 395, 1–6. [PubMed: 20206133]

Herdy B, Mayer C, Varshney D, Marsico G, Murat P, Taylor C, D'Santos C, Tannahill D, and Balasubramanian S (2018). Analysis of NRAS RNA G-quadruplex binding proteins reveals DDX3X as a novel interactor of cellular G-quadruplex containing transcripts. Nucleic Acids Res. 46, 11592–11604. [PubMed: 30256975]

Hönig A, Auboeuf D, Parker MM, O'Malley BW, and Berget SM (2002). Regulation of alternative splicing by the ATP-dependent DEAD-box RNA helicase p72. Mol. Cell. Biol 22, 5698–5707. [PubMed: 12138182]

Huang Y, and Liu ZR (2002). The ATPase, RNA unwinding, and RNA binding activities of recombinant p68 RNA helicase. J. Biol. Chem 277, 12810–12815. [PubMed: 11823473]

Jalal C, Uhlmann-Schiffler H, and Stahl H (2007). Redundant role of DEAD box proteins p68 (Ddx5) and p72/p82 (Ddx17) in ribosome biogenesis and cell proliferation. Nucleic Acids Res. 35, 3590–3601. [PubMed: 17485482]

Janknecht R (2010). Multi-talented DEAD-box proteins and potential tumor promoters: p68 RNA helicase (DDX5) and its paralog, p72 RNA helicase (DDX17). Am. J. Transl. Res 2, 223–234. [PubMed: 20589163]

Jankowsky E, and Fairman ME (2008). Duplex unwinding and RNP remodeling with RNA helicases. Methods Mol. Biol 488, 343–355. [PubMed: 18982301]

Jarmoskaite I, and Russell R (2011). DEAD-box proteins as RNA helicases and chaperones. Wiley Interdiscip. Rev. RNA 2, 135–152. [PubMed: 21297876]

Kao SH, Cheng WC, Wang YT, Wu HT, Yeh HY, Chen YJ, Tsai MH, and Wu KJ (2019). Regulation of miRNA biogenesis and histone modification by K63-polyubiquitinated DDX17 controls cancer stem-like features. Cancer Res. 79, 2549–2563. [PubMed: 30877109]

Kim K, Nguyen TD, Li S, and Nguyen TA (2018). SRSF3 recruits DROSHA to the basal junction of primary microRNAs. RNA 24, 892–898. [PubMed: 29615481]

Lambert MP, Terrone S, Giraud G, Benoit-Pilven C, Cluet D, Combaret V, Mortreux F, Auboeuf D, and Bourgeois CF (2018). The RNA helicase DDX17 controls the transcriptional activity of REST and the expression of pro-neural microRNAs in neuronal differentiation. Nucleic Acids Res. 46, 7686–7700. [PubMed: 29931089]

Li H, Lai P, Jia J, Song Y, Xia Q, Huang K, He N, Ping W, Chen J, Yang Z, et al. (2017). RNA helicase DDX5 inhibits reprogramming to pluripotency by miRNA-based repression of RYBP and its PRC1-dependent and -independent functions. Cell Stem Cell 20, 571.

Linder P, and Jankowsky E (2011). From unwinding to clamping—the DEAD box RNA helicase family. Nat. Rev. Mol. Cell Biol 12, 505–516. [PubMed: 21779027]

Lorgeoux RP, Pan Q, Le Duff Y, and Liang C (2013). DDX17 promotes the production of infectious HIV-1 particles through modulating viral RNA packaging and translation frameshift. Virology 443, 384–392. [PubMed: 23769241]

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, and Read RJ (2007). Phaser crystallographic software. J. Appl. Cryst 40, 658–674. [PubMed: 19461840]

Montpetit B, Thomsen ND, Helmke KJ, Seeliger MA, Berger JM, and Weis K (2011). A conserved mechanism of DEAD-box ATPase activation by nucleoporins and InsP6 in mRNA export. Nature 472, 238–242. [PubMed: 21441902]

Mooney SM, Grande JP, Salisbury JL, and Janknecht R (2010). Sumoylation of p68 and p72 RNA helicases affects protein stability and transactivation potential. Biochemistry 49, 1–10. [PubMed: 19995069]

Mori M, Triboulet R, Mohseni M, Schlegelmilch K, Shrestha K, Camargo FD, and Gregory RI (2014). Hippo signaling regulates microprocessor and links cell-density-dependent miRNA biogenesis to cancer. Cell 156, 893–906. [PubMed: 24581491]

Moy RH, Cole BS, Yasunaga A, Gold B, Shankarling G, Varble A, Molleston JM, tenOever BR, Lynch KW, and Cherry S (2014). Stem-loop recognition by DDX17 facilitates miRNA processing and antiviral defense. Cell 158, 764–777. [PubMed: 25126784]

Ogilvie VC, Wilson BJ, Nicol SM, Morrice NA, Saunders LR, Barber GN, and Fuller-Pace FV (2003). The highly related DEAD box RNA helicases p68 and p72 exist as heterodimers in cells. Nucleic Acids Res. 31, 1470–1480. [PubMed: 12595555]

Partin AC, Ngo TD, Herrell E, Jeong BC, Hon G, and Nam Y (2017). Heme enables proper positioning of Drosha and DGCR8 on primary microRNAs. Nat. Commun 8, 1737. [PubMed: 29170488]

Pyle AM (2008). Translocation and unwinding mechanisms of RNA and DNA helicases. Annu. Rev. Biophys 37, 317–336. [PubMed: 18573084]

Remenyi J, Bajan S, Fuller-Pace FV, Arthur JSC, and Hutvagner G (2016). The loop structure and the RNA helicase p72/DDX17 influence the processing efficiency of the mice miR-132. Sci. Rep 6, 22848. [PubMed: 26947125]

Robert F, and Pelletier J (2013). Perturbations of RNA helicases in cancer. Wiley Interdiscip. Rev. RNA 4, 333–349. [PubMed: 23658027]

Rössler OG, Straka A, and Stahl H (2001). Rearrangement of structured RNA via branch migration structures catalysed by the highly related DEAD-box proteins p68 and p72. Nucleic Acids Res. 29, 2088–2096. [PubMed: 11353078]

Rudolph MG, and Klostermeier D (2015). When core competence is not enough: functional interplay of the DEAD-box helicase core with ancillary domains and auxiliary factors in RNA binding and unwinding. Biol. Chem 396, 849–865. [PubMed: 25720120]

Samaan S, Tranchevent LC, Dardenne E, Polay Espinoza M, Zonta E, Germann S, Gratadou L, Dutertre M, and Auboeuf D (2014). The Ddx5 and Ddx17 RNA helicases are cornerstones in the complex regulatory array of steroid hormone-signaling pathways. Nucleic Acids Res. 42, 2197–2207. [PubMed: 24275493]

Sengoku T, Nureki O, Nakamura A, Kobayashi S, and Yokoyama S (2006). Structural basis for RNA unwinding by the DEAD-box protein *Drosophila* Vasa. Cell 125, 287–300. [PubMed: 16630817]

Sithole N, Williams CA, Vaughan AM, Kenyon JC, and Lever AML (2018). DDX17 specifically, and independently of DDX5, controls use of the HIV A4/5 splice acceptor cluster and is essential for efficient replication of HIV. J. Mol. Biol 430 (18 Pt B), 3111–3128. [PubMed: 30131116]

Studier FW (2014). Stable expression clones and auto-induction for protein production in E. coli. Methods Mol. Biol 1091, 17–32. [PubMed: 24203322]

Tsai CF, Wang YT, Yen HY, Tsou CC, Ku WC, Lin PY, Chen HY, Nesvizhskii AI, Ishihama Y, and Chen YJ (2015). Large-scale determination of absolute phosphorylation stoichiometries in human cells by motif-targeting quantitative proteomics. Nat. Commun 6, 6622. [PubMed: 25814448]

Walker SC, Avis JM, and Conn GL (2003). General plasmids for producing RNA in vitro transcripts with homogeneous ends. Nucleic Acids Res. 31, e82. [PubMed: 12888534]

Wang IX, Grunseich C, Fox J, Burdick J, Zhu Z, Ravazian N, Hafner M, and Cheung VG (2018). Human proteins that interact with RNA/DNA hybrids. Genome Res. 28, 1405–1414. [PubMed: 30108179]

Wilkinson KA, Merino EJ, and Weeks KM (2006). Selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution. Nat. Protoc 1, 1610–1616. [PubMed: 17406453]

Xing Z, Wang S, and Tran EJ (2017). Characterization of the mammalian DEAD-box protein DDX5 reveals functional conservation with *S. cerevisiae* ortholog Dbp2 in transcriptional control and glucose metabolism. RNA 23, 1125–1138. [PubMed: 28411202]

Xing Z, Ma WK, and Tran EJ (2019). The DDX5/Dbp2 subfamily of DEAD-box RNA helicases. Wiley Interdiscip. Rev. RNA 10, e1519. [PubMed: 30506978]

Yang Q, and Jankowsky E (2006). The DEAD-box protein Ded1 unwinds RNA duplexes by a mode distinct from translocating helicases. Nat. Struct. Mol. Biol 13, 981–986. [PubMed: 17072313]

Zhang ZM, Yang F, Zhang J, Tang Q, Li J, Gu J, Zhou J, and Xu YZ (2013). Crystal structure of Prp5p reveals interdomain interactions that impact spliceosome assembly. Cell Rep. 5, 1269–1278. [PubMed: 24290758]

Highlights

- Crystal structures of human DDX17 core domains in multiple states are determined
- DDX17 core domains exhibit RNA sequence preference via the RMFQ motif
- DDX17 can enhance processing of pri-miRs by remodeling the 3′ flanking region
- N-terminal tail of DDX17 can regulate the ATPase activity

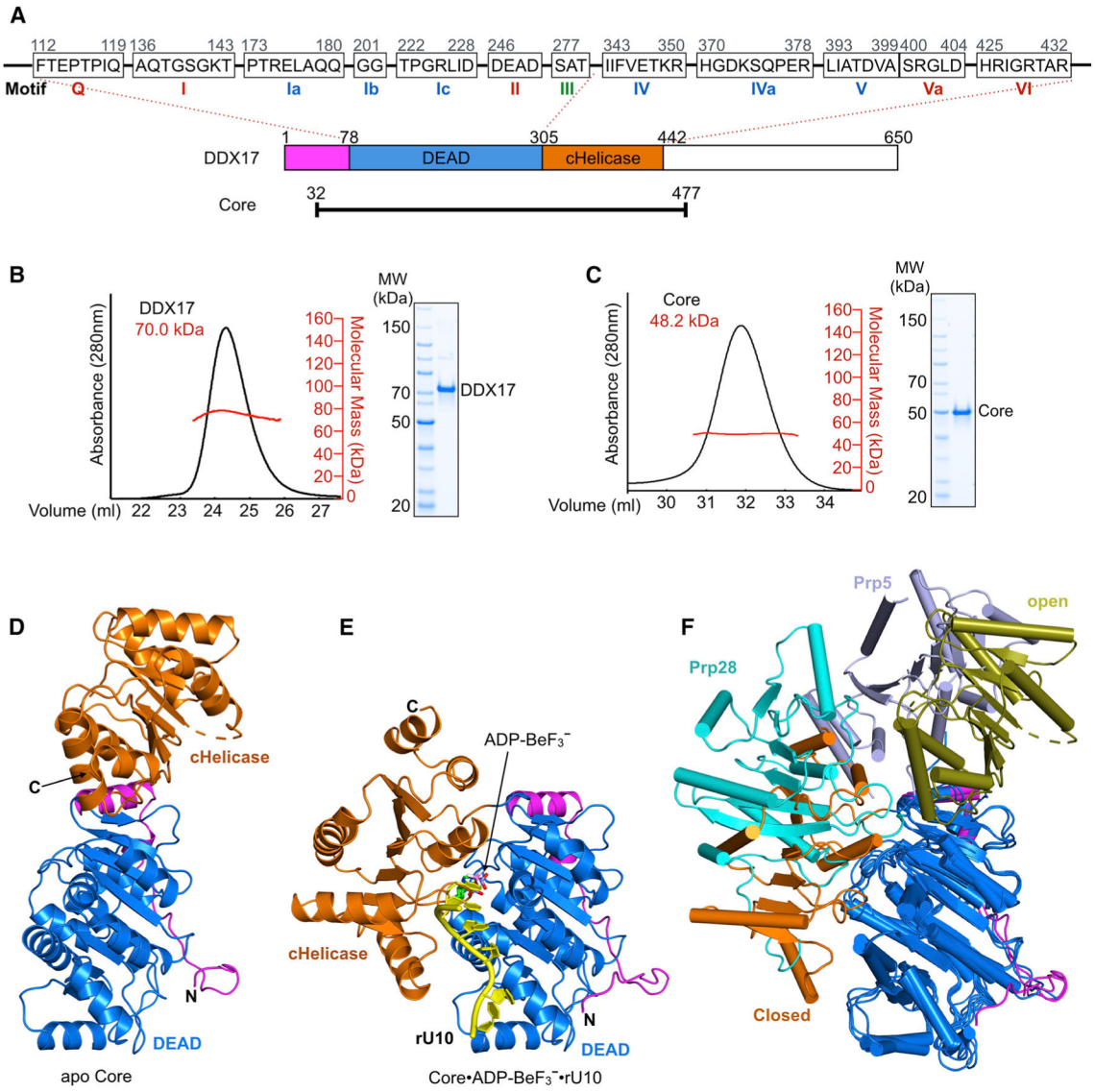


Figure 1. Overall Structures of the Apo and RNA-Bound DDX17 Core Reveal a Large Conformational Change

(A) Domain organization and conserved motifs of human DDX17 and the DDX17 Core construct used in this study. Motifs associated with ATP binding, RNA binding, and coupling of ATPase-unwinding are colored in red, blue, and green, respectively. (B) SEC-MALS analysis of full-length DDX17 reports a molecular mass of 70.0 kDa (theoretical, 72.4 kDa). SDS-PAGE of the purified protein is shown on the right. (C) SEC-MALS analysis of the DDX17 Core reports a molecular mass of 48.2 kDa (theoretical, 51.1 kDa). SDS-PAGE of the purified protein is shown on the right. (D) Cartoon representation of DDX17 Core in open conformation. The N-terminal tail, DEAD domain, and cHelicase domain are colored magenta, blue, and orange, respectively. (E) Cartoon representation of the DDX17 Core in closed conformation with rU10 and ADP-BeF_3^-. The N-terminal tail, DEAD domain, and cHelicase domain are colored magenta, blue, and orange, respectively. RNA rU10 is shown in yellow, and ADP-BeF_3^- is shown in green. (F) Superimposition of

the DEAD domains (blue) from DDX17 Core structures (orange for both conformations) and other DEAD-box helicases, Prp28 (cyan) and Prp5 (gray).

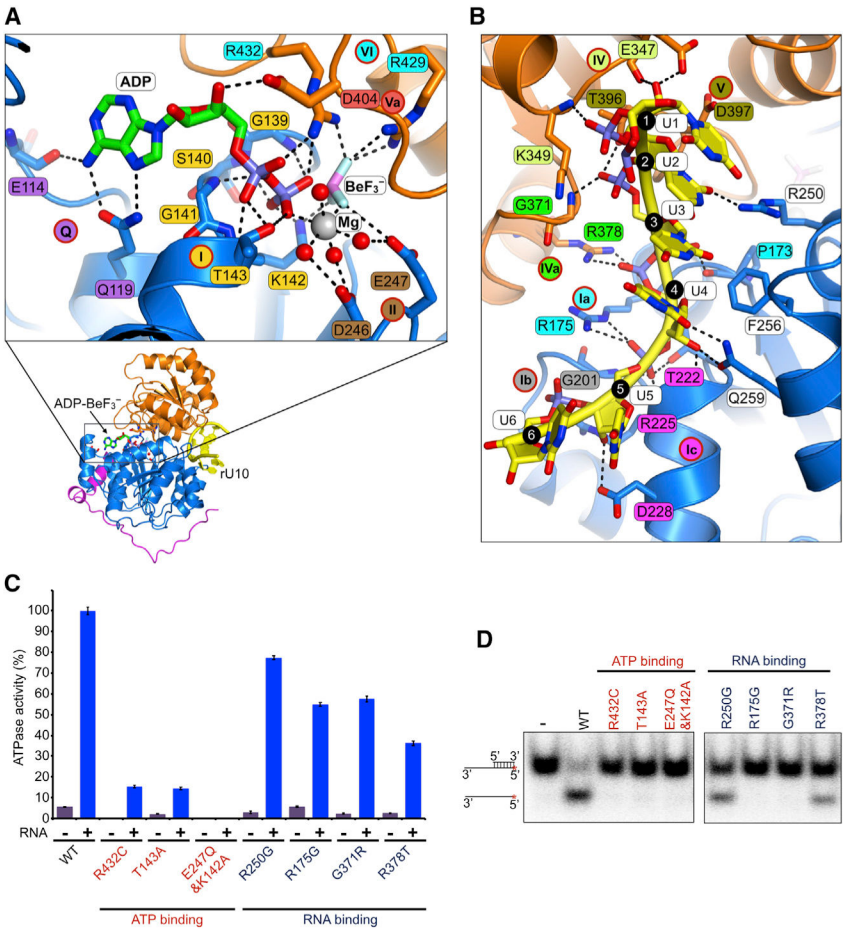


Figure 2. Cancer Mutations Found in DDX17 ATP- and RNA-Binding Sites Are Deficient in ATP Hydrolysis and Unwinding Activity

(A) Detailed view of the ATP-binding site of DDX17 Core in closed conformation. Contact residues are shown as sticks, and the labels are colored differently for each helicase motif. Hydrogen bonds are presented as black dashes. (B) Detailed view of RNA-protein contacts involved in the DDX17 Core/rU10 complex. Arbitrary nucleotide positions are marked with black circles with white numbers. (C) ATPase activities of wild-type (WT) and mutant full-length DDX17. ATPase assays were performed with or without 1.7 μM dsRNA with 3′ overhang (Figure S1B). (D) RNA unwinding activities WT and mutant full-length DDX17, using dsRNA with a 3′ overhang.

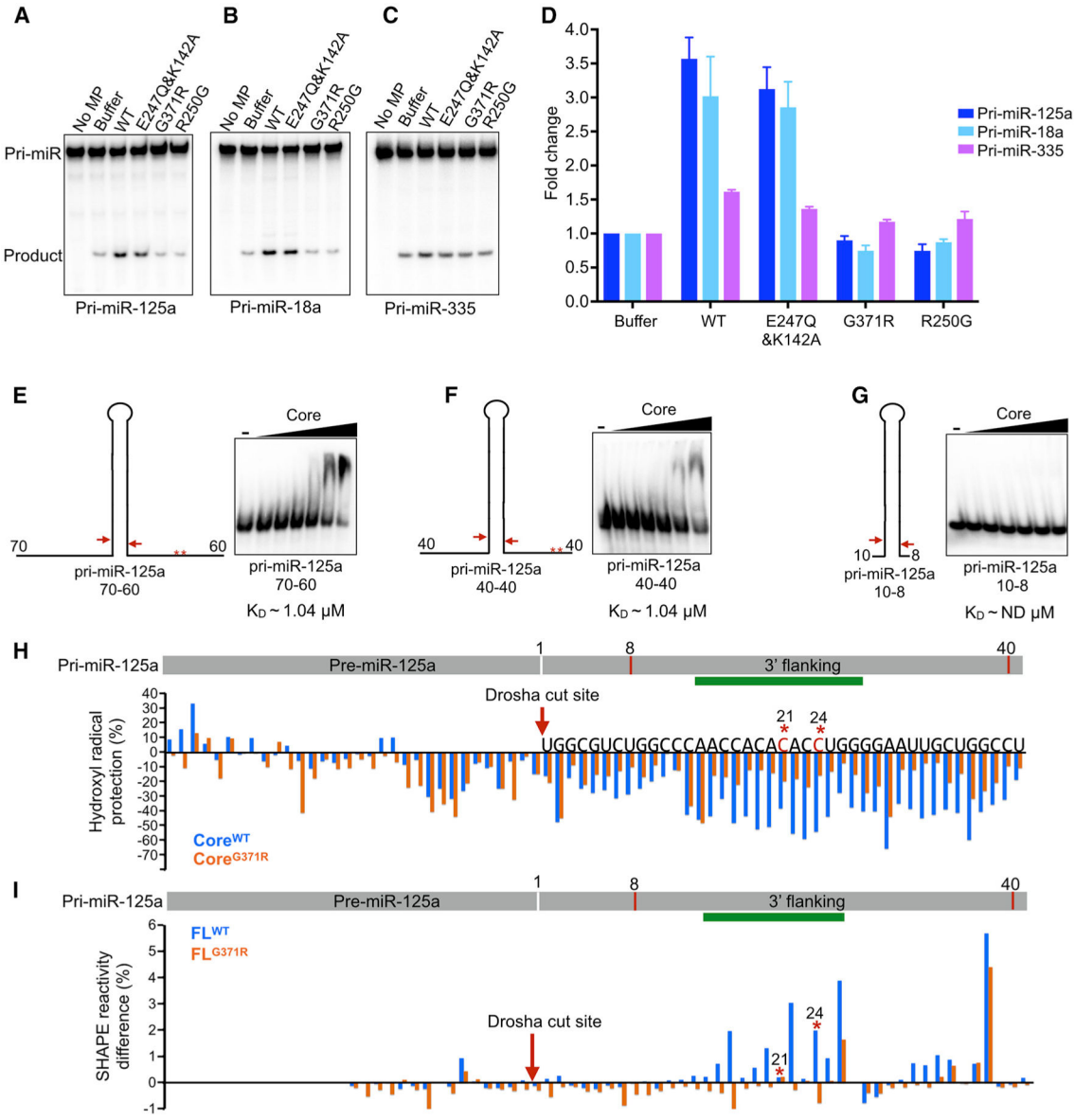


Figure 3. DDX17 Enhances Microprocessor Activity by Remodeling the 3′ Flanking Extension of Pri-miR-125a

(A–C) *In vitro* pri-miRNA processing assays using pri-miR-125a (A), pri-miR-18a (B), and pri-miR-335 (C) in the presence of wild-type or mutant full-length DDX17. (D) Quantification of *in vitro* pri-miRNA processing assays shown in (A)–(C). Data are shown as mean fold change ± SE from three replicates. (E–G) EMSAs of DDX17 Core with pri-miR-125a probes with varying arm lengths (E, 70/60; F, 40/40; G, 10/8). The arm lengths are measured from Drosha cut sites. Protein concentrations from left to right are as follows: 0.065, 0.13, 0.26, 0.52, 1.04, and 2.08 μM. (H) Quantitation of hydroxyl radical footprinting experiments for pri-miR-125a in the presence of 625 nM DDX17 Core^{WT} and Core^{G371R}. Left to right, the bars represent nucleotides in the 5′-to-3′ direction. Cytosines in the CNNC motif are indicated with red asterisks. The 3′ Drosha cut site is marked with a red arrow. The green bar indicates the region undergoing the highest degree of protection. (I) SHAPE

analysis of pri-miR-125a in the presence of 625 nM full-length $DDX17^{WT}$ or $DDX17^{G371R}$. Left to right, the bars represent nucleotides in the 5′-to-3′ direction. Cytosines in the CNNC motif are indicated with red asterisks. The 3′ Drosha cut site is marked with a red arrow. The green bar indicates the region with the highest SHAPE reactivity difference.

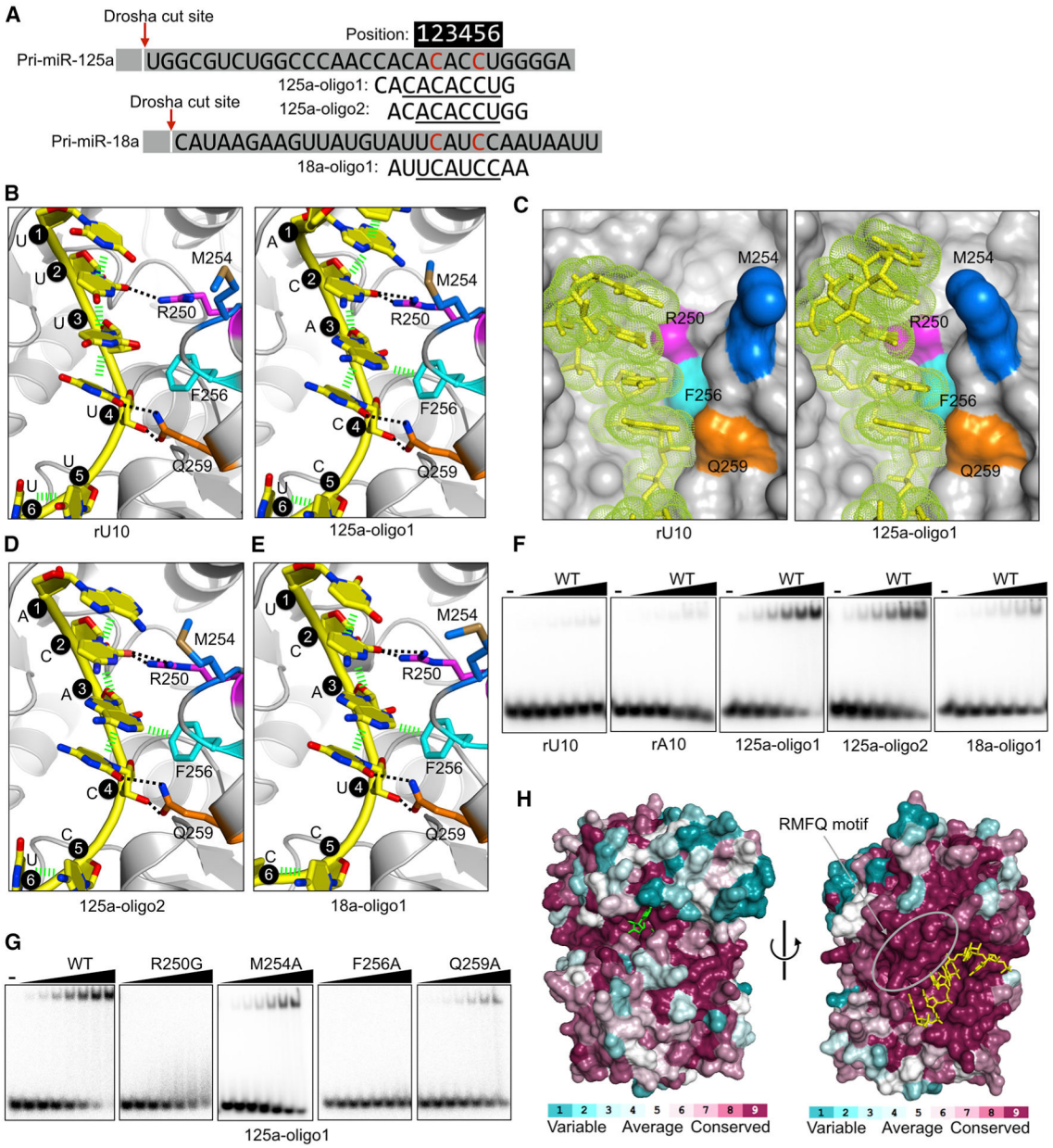


Figure 4. DDX17 Core Exhibits Sequence Preference via RMFQ Motif

(A) Sequences of oligonucleotides derived from the 3′ arms of pri-miR-125a and pri-miR-18a used for crystallization. (B) Detailed view of rU10 and 125a-oligo1 bound to the RMFQ groove. Ring stacking interactions are marked with fluorescent green dashes, and hydrogen bonds are marked with black dotted lines. (C) Surface representation of the RMFQ groove with RMFQ colored in magenta, blue, cyan, and orange, respectively. RNA is shown in stick representation with dotted surface. (D and E) Detailed view of 125a-oligo2 (D) and 18a-oligo1 (E) bound to the RMFQ groove. Ring stacking interactions are marked with fluorescent green dashes, and hydrogen bonds are marked with black dotted lines. (F) EMSAs of DDX17 Core binding 10-mer ssRNA oligonucleotides: rU10, rA10, 125a-oligo1, 125a-oligo2, and 18a-oligo1, in the presence of 2 mM adenylylimidodiphosphate (AMP-

PNP). Protein concentrations (left to right) are as follows: 0.52, 1.04, 2.08, 4.17, 8.33, and 16.68 μM. (G) EMSAs of the wild-type or mutant DDX17 Core constructs binding radiolabeled 125a-oligo1 in the presence of 2 mM AMP-PNP. Protein concentrations (left to right) are as follows: 0.26, 0.52, 1.04, 2.08, 4.17, 8.33, and 16.68 μM. (H) Sequence conservation among DDX17 orthologs, shown as surface colored by conservation score. Binding sites for ATP and RNA, including the RMFQ motif (gray circle), are more conserved than the rest of the protein.

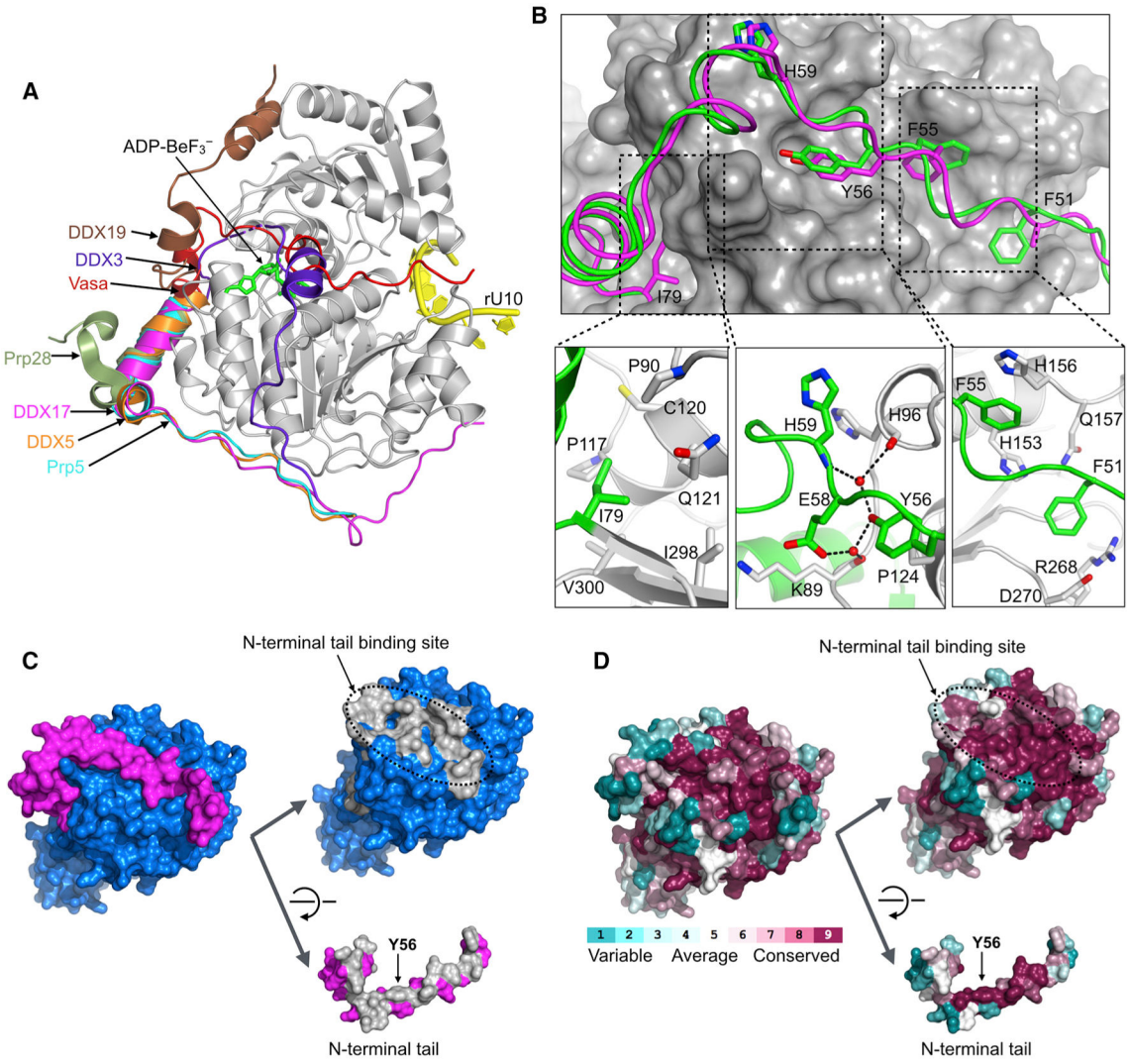
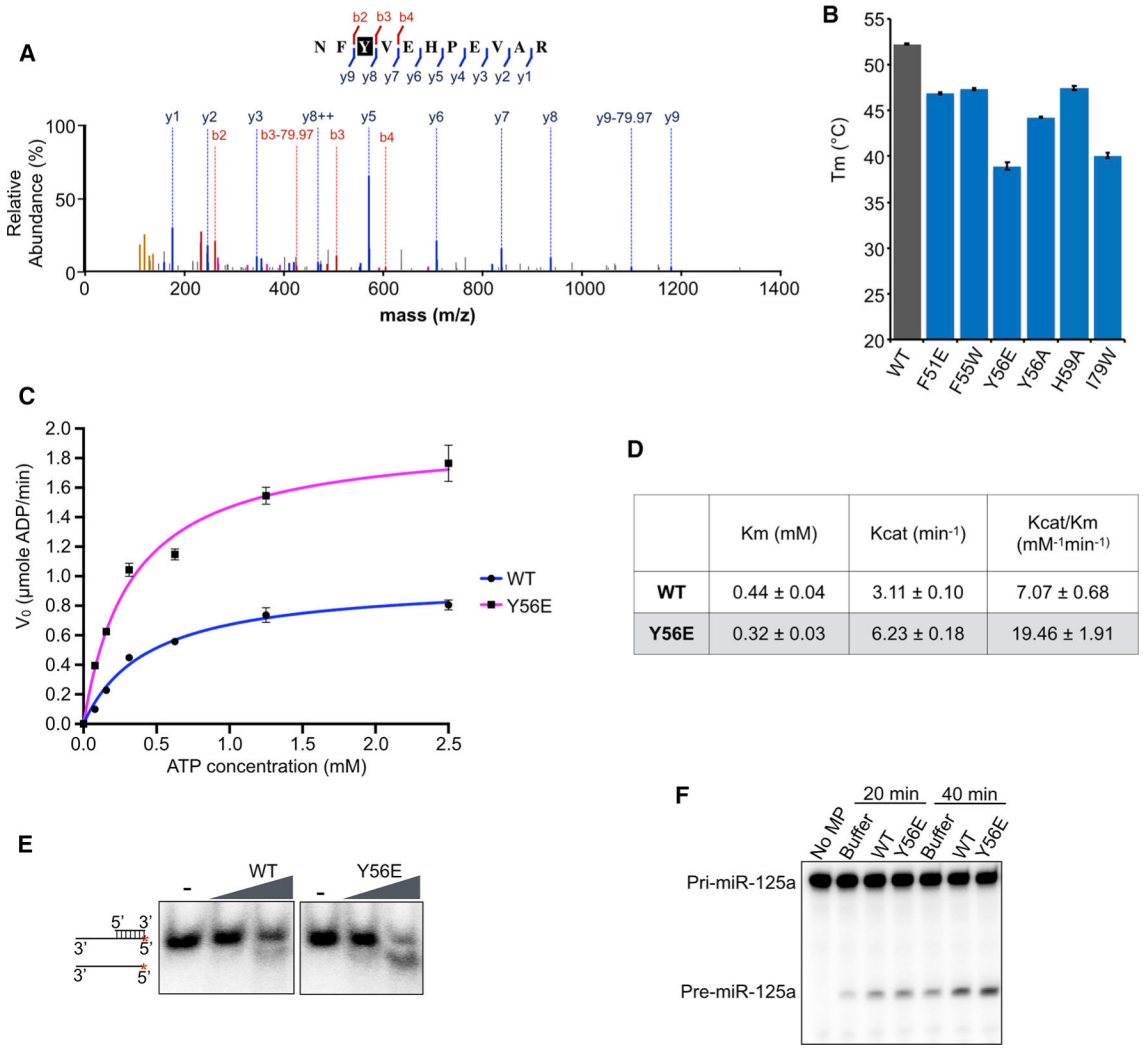


Figure 5. N-Terminal Tail of DDX17 Specifically Binds the DEAD Domain

(A) Superimposition of DDX17 with other DEAD-box helicases, DDX19 (PDB: 3G0H), Vasa (PDB: 2DB3), Prp5 (PDB: 4LJY), DDX3 (PDB: 6CZ5), prp28 (PDB: 4W7S), and DDX5 (PDB: 4A4D), by aligning only the DEAD domain. DEAD and cHelicase domains of DDX17 are shown as gray cartoon. For other helicases, only the N-terminal extensions are displayed. (B) Detailed view of the intramolecular contacts involved in maintaining the orientation of the N-terminal tail of DDX17. N-terminal extension from the closed structure (Core·ADP-BeF_3^-·$rU10) is shown as magenta cartoon, while that of the free DDX17 structure is shown as green cartoon. DEAD and cHelicase domains are shown in gray. Key residues with substantial contact with the DEAD domain are shown as sticks. (C) Surface representation of free DDX17 Core, viewing from the N-terminal tail (magenta) side, binding the rest of the protein (blue). The complex (left) is peeled off to show the interacting surfaces in gray (right). (D) Similar views as in (C), colored by sequence conservation score.



	Km (mM)	Kcat (min^{-1})	Kcat/Km (mM^{-1}min^{-1})
WT	0.44 ± 0.04	3.11 ± 0.10	7.07 ± 0.68
Y56E	0.32 ± 0.03	6.23 ± 0.18	19.46 ± 1.91

Figure 6. Destabilization of DDX17 in N-Terminal Extension Increases ATP Turnover in the RNA-Dependent ATPase Activity but Has No Effect on Microprocessor Activity

(A) Phosphorylated Y56 in insect cell-produced DDX17 detected using liquid chromatography-tandem mass spectrometry (LC-MS/MS). (B) Differential scanning fluorimetry shows different melting temperatures (Tms) for various full-length DDX17 constructs. (C) Michaelis-Menten plot for the ATPase activity of full-length DDX17 constructs. Data are shown as mean initial velocity ±SE from three replicates. (D) Kinetics data derived from ATPase activity assays performed at saturating (2.4 μM) RNA concentrations. (E) RNA unwinding assays for WT or mutant full-length DDX17 constructs. (F) Microprocessor assay of pri-miR-125a in the presence of WT or mutant full-length DDX17.

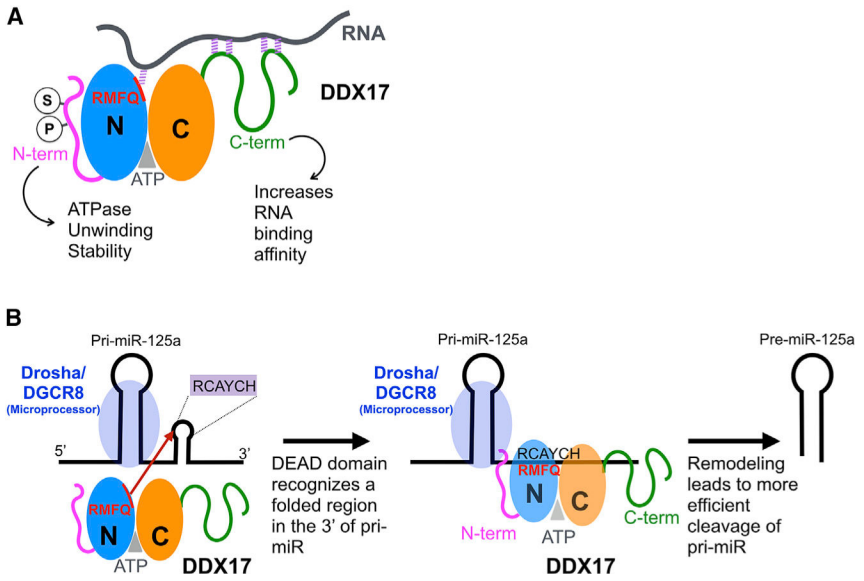


Figure 7. Mechanistic Model for DDX17 Function

(A) DDX17 contains multiple modules with different roles. The N-terminal tail (magenta) is involved in modulating the ATPase activity, and post-translational modifications; e.g., phosphorylation (p) or sumoylation (s) may be used to switch the N-terminal extension conformation. The catalytic domains DEAD (N) and cHelicase (C) bind ATP and RNA, and RMFQ groove can dictate sequence preference despite weak affinity for ssRNA. The C-terminal region of DDX17 has high affinity for ssRNA. (B) DDX17 binds to the 3′ flanking region of certain pri-miRs, partly by recognizing the RCAYCH motif via the RMFQ groove. DDX17 induces remodeling of the pri-miR in the 3′ region, which leads to more efficient cleavage without affecting affinity for Drosha/DGCR8 (Microprocessor).

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse anti-Actin	Abcam	Cat#ab6276; RRID:AB_2223210
Rabbit anti-DDX17	ThermoFisher	PIPA5–26330; RRID:AB_2543830
Goat anti-Rabbit HRP	Invitrogen	31460; RRID:AB_228341
Rabbit anti-DGCR8	ThermoFisher	PA5–40122; RRID:AB_2606318
Bacterial and Virus Strains		
Rosetta 2 (BL21-DE3)	Novagen	Cat# 70956
Chemicals, Peptides, and Recombinant Proteins		
AMP-PNP	Sigma	10102547001
ADP	Sigma	A2754
ATP	Sigma	A2383
$BeCl_2$	Santa Cruz	sc-257152
NaF	Sigma	201154
gamma-^{32}P-ATP	Perkin Elmer	NEG035C001MC
NAI	EMD Millipore	03–310
SYPRO Orange	Sigma	S5692–500UL
SuperScript III Reverse Transcriptase	ThermoFisher	18080044
DDX17 FL WT	This paper	N/A
DDX17 FL E247Q&K142A	This paper	N/A
DDX17 FL R432C	This paper	N/A
DDX17 FL T143A	This paper	N/A
DDX17 FL R250G	This paper	N/A
DDX17 FLR175G	This paper	N/A
DDX17 FL G371R	This paper	N/A
DDX17 FL R378T	This paper	N/A
DDX17 FL F51W	This paper	N/A
DDX17 FL F55W	This paper	N/A
DDX17 FL Y56E	This paper	N/A
DDX17 FL Y56A	This paper	N/A
DDX17 FL H59A	This paper	N/A
DDX17 FL I79W	This paper	N/A
DDX17 Core WT (32–477)	This paper	N/A
DDX17 N-Core (1–477)	This paper	N/A
DDX17 Core-C (32–650)	This paper	N/A
DDX17 Core R250G	This paper	N/A
DDX17 Core R175G	This paper	N/A
DDX17 Core G371R	This paper	N/A

REAGENT or RESOURCE	SOURCE	IDENTIFIER
DDX17 Core R378T	This paper	N/A
DDX17 Core Y56E	This paper	N/A
Critical Commercial Assays		
ADP-Glo Max Assay kit	Promega	V7001
Deposited Data		
Core-apo	This study	PDB entry: 6UV0
Core·ADP-BeF_3^-·rU10	This study	PDB entry: 6UV1
Core· ADP-BeF_3^-· 125a-oligo1	This study	PDB entry: 6UV2
Core· ADP-BeF_3^-· 125a-oligo2	This study	PDB entry: 6UV3
Core· ADP-BeF_3^-· 18a-oligo1	This study	PDB entry: 6UV4
Experimental Models: Cell Lines		
High Five insect cells	ThermoFisher	Cat# B85502
Sf21 insect cells	Expression Systems	94–003F
HEK293T cells (female)	ATCC	ATCC® CRL-11268
Oligonucleotides		
rU10: rUrUrUrUrUrUrUrUrU	IDT	N/A
rA10: rArArArArArArArArArA	IDT	N/A
16-mer-3OV: rArGrCrArCrCrGrUrArArArGrArCrGrC	IDT	N/A
41mper-3OV: rGrCrGrUrCrUrUrUrArCrGrGrUr GrCrUrUrArArArArCrArArArArArCrArArArArArCrAr ArArArCrArArArA	IDT	N/A
16-mer-5OV: rArGrCrArCrCrGrUrArArArGrArCrGrC	IDT	N/A
41-mer-5OV: 5'-rArArArArCrArArArArCrArArArArCrArArArArCrArArArArUrGrCrGrUrCrUrUrUrArCrGrGrUrGrCrU-3'	IDT	N/A
125a-oligo1: rCrArCrArCrArCrCrUrG	IDT	N/A
125a-oligo2: rArCrArCrArCrCrUrGrG	IDT	N/A
18a-oligo1: rArUrUrCrArUrCrCrArA	IDT	N/A
Recombinant DNA		
pET21 constructs for expression of recombinant proteins	This study	N/A
pET21	Novagen (EMD Millipore)	Cat# 69770–3
pRZ constructs for *in vitro* RNA transcription templates	This study	N/A
pRZ	Walker et al., 2003	N/A
pcDNA3.1 constructs for overexpression of RNA and DDX17 in mammalian cells	This study	N/A
pcDNA3.1	ThermoFisher	Cat# V79020
pFastBacDual construct for expression of full-length DDX17 in insect cells	This study	N/A
pFastBacDual	ThermoFisher	Cat# 0712024
Software and Algorithms		
Astra 6.1 software	Wyatt	N/A
PRISM 8.0 software	GraphPad Software	N/A
Phaser/Phenix	McCoy et al., 2007	N/A
Coot	Emsley et al., 2010	N/A
MolProbity	http://molprobity.biochem.duke.edu/	N/A

REAGENT or RESOURCE	SOURCE	IDENTIFIER
PyMOL	https://pymol.org/2/	N/A
Image Lab	Bio-Rad	N/A
Other		
HPLC (1200 Ininity)	Agilent	N/A
DAWN HELIOS II detector	Wyatt	N/A
tREX detector	Wyatt	N/A
Superdex 200 16/60	GE	N/A