Title

Coordinated Activities of Human Dicer Domains in Regulatory RNA Processing

Journal

Journal of Molecular Biology, 422(4)

ISSN

0022-2836

Authors

Ma, Enbo

Zhou, Kaihong

Kidwell, Mary Anne

et al.

Publication Date

2012-09-01

DOI

10.1016/j.jmb.2012.06.009

Published in final edited form as:
J Mol Biol. 2012 September 28; 422(4): 466–476. doi:10.1016/j.jmb.2012.06.009.

Coordinated Activities of Human Dicer Domains in Regulatory RNA Processing

Enbo Ma[1], **Kaihong Zhou**[3], **Mary Anne Kidwell**[1], and **Jennifer A. Doudna**[1,2,3,4,*]

[1]Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720

[2]Department of Chemistry, University of California, Berkeley, CA 94720

[3]Howard Hughes Medical Institute, Lawrence Berkeley National Laboratory, Berkeley, CA 94720

[4]Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720

Summary

The conserved ribonuclease Dicer generates microRNAs and short interfering RNAs that guide gene silencing in eukaryotes. The specific contributions of human Dicer's structural domains to RNA product length and substrate preference are incompletely understood, due in part to the difficulties of Dicer purification. Here we show that active forms of human Dicer can be assembled from recombinant polypeptides expressed in bacteria. Using this system, we find that three distinct modes of RNA recognition give rise to Dicer's fidelity and product length specificity. The first involves anchoring one end of a dsRNA helix within the PAZ domain, which can assemble *in trans* with Dicer's catalytic domains to reconstitute an accurate but non-substrate-selective dicing activity. The second entails non-specific RNA binding by the double-stranded RNA binding domain (dsRBD), an interaction that is essential for substrate recruitment in the absence of the PAZ domain. The third mode of recognition involves hairpin RNA loop recognition by the helicase domain, which ensures efficient processing of specific substrates. These results reveal distinct interactions of each Dicer domain with different RNA structural features, and provide a facile system for investigating the molecular mechanisms of human miRNA biogenesis.

Keywords

Ribonuclease; Dicer; RNAi

Introduction

Dicer plays a central role in RNA silencing by generating ~22 nt short interfering RNAs (siRNAs) or microRNAs (miRNAs) from long dsRNAs or hairpin RNAs, respectively. These small regulatory RNAs assemble with Argonaute-family proteins such as Ago2 to

*Correspondence: doudna@berkeley.edu.

AUTHOR CONTRIBUTIONS

E.M. planned the experiments, generated hDcr constructs, purified proteins, performed *in vitro* RNA binding and cleavage experiments, analyzed and interpreted the data, and wrote the manuscript. K.H.Z. purified proteins and prepared the RNA substrates. M.A.K wrote the manuscript. J.A.D planned the experiments, interpreted the data and wrote the manuscript.

direct the degradation or translational repression of mRNAs bearing complementary sequences [1; 2]. The multiple domains of human Dicer (hDcr) contribute to double-stranded RNA recognition and processing by mechanisms that have not been fully elucidated. Determining how Dicer's RNA interacting domains ensure accurate cleavage of different RNA substrates will extend our understanding of the biogenesis of small RNAs under both physiological and pathological conditions.

Structural analysis of *Giardia* Dicer [3] and biochemical studies of hDcr [4] showed that two tandem RNase III domains in the enzyme form an intramolecular dimer that cleaves dsRNA by a mechanism reminiscent of the homodimeric bacterial ribonuclease III [5]. Recently it was shown that the PAZ domain of hDcr anchors the 3' and 5' termini of a dsRNA substrate to determine the product size of hDcr cleavage [6]. In addition to the RNase III and PAZ domains, hDcr also contains an N-terminal ATPase/helicase domain (ATPase/hel), a domain of unknown function (DUF283) and a C-terminal dsRNA-binding domain (dsRBD). Although these domains are highly conserved, their roles in small RNA biogenesis may differ in different organisms [7; 8; 9; 10; 11; 12].

Clues to the possible functions of hDcr's non-catalytic domains have emerged from previous studies using mutated forms of the enzyme both *in vitro* and *in vivo*. Kinetic experiments revealed large differences in cleavage rates for long dsRNA versus hairpin substrates [13; 14]. Dicer processes pre-miRNA substrates much faster than pre-siRNA substrates under both single and multiple turnover conditions. Deletion of the ATPase/hel domain from hDcr enhances cleavage of pre-siRNA substrates without affecting pre-miRNA processing rates [13]. In addition, mutation of the helicase domain leads to changes in dsRNA processing *in vivo* [15].

A major challenge to investigating hDcr function and structure has been the difficulty of producing large amounts of active enzyme. Although recombinant Dicer can be expressed and purified from insect cell culture [13; 16; 17], the low yield and expense of protein production have inhibited progress. Here we show that active forms of hDcr can be reconstituted from separate polypeptides expressed in *E. coli*. Using this system, we find that hDcr employs its component RNA-interacting domains for distinct aspects of the dicing reaction, including determination of product length and recognition of different substrate classes. The combined footprint of the PAZ and RNase III domains on dsRNA determines the unique and accuracy of dsRNA product length generated by hDcr. The C-terminal double-stranded RNA binding domain (dsRBD), which does not influence product length or PAZ domain association, is required for substrate binding in the absence of the PAZ domain. Finally, the ATPase/hel domain enables hDcr to distinguish between pre-miRNA and long dsRNA substrates by interaction with the single-stranded hairpin loop. Together, the coordinated interactions among these RNA recognition domains ensure accurate processing of different RNA substrates by hDcr.These results provide a powerful system for investigating the molecular mechanisms of human miRNA biogenesis, and also offer practical insights relevant to the production of small hairpin RNAs.

Results

Active hDcr can be reconstituted from trans-expressed fragments

The large size and multi-domain composition of hDcr have presented challenges to its expression, purification and analysis in recombinant form. Previous studies have relied on the presence of endogenous hDcr in cell extracts or purified hDcr obtained by over-expression in baculovirus-infected insect cells [4; 13; 14; 18; 19]. The limitations of these approaches have hindered biochemical and structural analyses of hDcr domain functions. Although prior attempts to express hDcr in *E. coli* were unsuccessful, we reasoned that it

might be possible to break the protein into smaller fragments that could be expressed individually in bacteria. Using full-length active recombinant hDcr, we performed limited proteolysis using endoproteinase Glu-C to identify stable fragments. This treatment produced two polypeptides whose sizes together equaled ~95% of the full-length hDcr sequence (Fig. 1A). Results from both mass spectrometry and Edman degradation sequencing showed that one fragment contains the ATPase/hel, DUF283, and PAZ domains (N-terminal fragment, hDcr-N) and the other contains the two tandem RNase III domains and the C-terminal dsRBD (C-terminal fragment, hDcr-C) (Fig. 1A).

We next prepared recombinant baculovirus constructs encoding these polypeptides and tested their expression in insect (Sf9) cells. Although the two fragments could not be expressed individually in this system, co-expression led to production of a stable complex that could not be disrupted by either 1 M sodium chloride or 4 M urea (Fig. 1B; Supplementary Fig. S1). To check whether the co-expressed complex was correctly folded and functional, we performed cleavage assays using a 35-base pair substrate (37ab) comprised of two complementary RNA oligos (37a and 37b) with 2-nt 3' single-stranded overhangs. These single turnover dicing assays showed that the hDcr-N/C complex has an activity similar to that of wild-type hDcr, producing 22-bp products (Fig. 1C). To quantitate the dicing activity of the hDcr-N/C complex, we performed multiple turnover assays. These studies showed that the complex has a similar K_M as that of intact wild-type hDcr, but its k_{cat} is 3.1×10^{-4} s^{-1},4-fold of wild-type hDcr's (7.0×10^{-5} s^{-1}) (Supplementary Table S1) [13].

Physical interaction of the PAZ and RNase III domains generates length–specific products

The successful expression of hDcr fragments *in trans* in the baculovirus system led us to test whether these polypeptides could be further truncated and expressed in *E. coli*. We focused initially on the catalytic region of the enzyme corresponding to the version of Dicer found in *Giardia* [3; 20]. Based on published sequence alignments [4; 18], we cloned and over-expressed this region as two separate polypeptides that included the DUF and PAZ domains (DP), and the RNase III and dsRBD domains (hDcr-C), respectively (Fig. 2A). RNA cleavage assays showed that the purified hDcr-C protein is catalytically active but yields a product set in which a 15 bp fragment is prominent rather than the characteristic 22 bp product generated by full-length hDcr (Fig. 2B). For comparison, the main products generated by *E. coli* RNase III, a structural homolog of the RNase III domains of hDcr, are 12 bp in length (Fig. 2B, left panel, lane 4). We also found that *E. coli* RNase III could cleave a 19 bp substrate, whereas the hDcr-C showed no cleavage activity on this substrate (Supplementary Fig. S2). This indicated that the ~21-23 bp RNAs produced by hDcr-C cleavage (Fig. 2B, left and middle panels, marked with short vertical lines) could not be further processed by this polypeptide.

Additional cleavage assays showed that the hDcr-C protein can also cleave a hairpin RNA (pre-miR-20a) in a similar manner, generating 15 bp products (Fig. 2B, right panel). However, the cleavage rate for the pre-miRNA was slower than that for the pre-siRNA substrate (Fig. 2B, compare middle and right panels). To control for the possibility that this cleavage activity arises from RNase contamination during protein preparation, we expressed an hDcr-C protein variant containing point mutations in the two RNaseIII active sites (Glu1316Ala and Glu1705Ala). These mutations abolished cleavage activity (Fig. 2B, left panel, lane 5).

Addition of the DP polypeptide to hDcr-C-catalyzed dsRNA cleavage reactions restored the production of 22-bp RNA products similar to those produced by full-length hDcr (Fig. 2C, lanes 6-7). This observation indicates that DP and hDcr-C are correctly folded and interact with each other, either directly or through their mutual binding to dsRNA. To test their physical interaction, DP and hDcr-C polypeptides were incubated together in the absence of

RNA and then analyzed by size exclusion chromatography. The elution profile suggests that DP and hDcr-C form a stable complex in the absence of RNA (Supplementary Fig. S3). We note that the 7-bp RNA footprint conferred by the PAZ domain within the DP polypeptide [21; 22] and the hDcr-C-generated 15-bp product suggests that the size of hDcr products (22 bp) is determined by the combined footprints of the PAZ and RNase III domains on the RNA.

The C-terminal dsRBD is required for RNA substrate binding and cleavage activities of hDcr-C

It has been reported that the dsRBD of *E. coli* RNase III is not required for substrate cleavage [23], whereas this domain is necessary for the activity of human *Drosha*, another RNase III family enzyme in the microRNA pathway [24]. In addition, deletion of the dsRBD from hDcr may contribute to reduced dsRNA processing efficiency [4]. To assess the importance of the C-terminal dsRBD in the hDcr-C construct, we expressed a shortened version of hDcr-C lacking the dsRBD (hDcr-CΔRBD, Fig. 3A). Our analysis showed that the hDcr-CΔRBD protein alone had no cleavage activity (Fig. 3B, lanes 1-2), indicating that the terminal dsRBD is necessary for hDcr-C to bind and/or cleave dsRNA. To test whether the bacterially expressed hDcr-CΔRBD retains its native fold and catalytic capability, we tested dsRNA cleavage assays in the presence of the DP polypeptide. These assays show that the presence of DP restored the dicing pattern of hDcr (Fig. 3B, lanes 3-4). We also found that deletion of the dsRBD from hDcr-C did not affect the ability of the catalytic, tandem RNase III domains to form a complex with DP (Supplementary Fig. S4). Therefore, the terminal dsRBD facilitates hDcr-C substrate cleavage, but does not affect the folding or catalytic function of the RNaseIII domains.

To establish the relationship between cleavage activity and substrate binding, we performed nitrocellulose filter-binding assays with three kinds of RNAs under conditions in which divalent ions are chelated to prevent dsRNA cleavage: substrate dsRNA (37ab), Dicer product-mimic dsRNA (21ab) and a pre-miRNA (pre-hlet-7a-1). The DP fragment had a higher affinity for perfectly matched dsRNAs (either substrate or product RNAs) than for the hairpin pre-miRNA (K_d ~200 nM versus ~1 μM, Table 1). By contrast, the hDcr-C fragment bound with measurable affinity only to the substrate dsRNA substrate (K_d ~300 nM) and displayed almost no detectable binding to either the hairpin or product RNAs (Table 1). These RNA binding data are consistent with the above cleavage results, showing that the hDcr-C protein is more active towards long, perfectly matched dsRNA substrates relative to pre-miRNAs. Removal of the terminal dsRBD domain from hDcr-C abolishes its RNA binding ability, indicating that this domain is required for the binding activity of hDcr-C to dsRNA in the absence of the PAZ domain (Table 1).

The hDcr ATPase/hel domain is important for pre-miRNA substrate selectivity

Based on assays with specific types of RNA substrates, the C-terminal hDcr fragment appears to bind and cleave perfect duplexes more efficiently than hairpin RNAs (Fig. 2B). However, wild-type hDcr has the opposite preference, cleaving hairpin RNA ~100-fold faster than a perfect dsRNA duplex under k_{cat}/K_m conditions [13; 14]. We hypothesized that the hDcr-N polypeptide, which includes the ATPase/hel, DUF and PAZ domains, might play a regulatory role in selective pre-miRNA processing. Although this fragment could not be expressed on its own, either in insect cells or in *E. coli*, we generated a construct containing the complete ATPase/helicase domain of hDcr fused with maltose-binding protein (MBP) that could be produced in *E. coli* (Fig. 4A).

Since hDcr interacts with human TAR-RNA binding protein (hTRBP) via its helicase domain [16; 25; 26; 27], we tested whether the MBP-ATPase/hel fusion retains the ability to

bind to recombinant hTRBP. Both size exclusion chromatography and co-immunoprecipitation assays showed that the helicase domain interacts with hTRBP, indicating that the purified MBP-ATPase/hel protein is likely to be correctly folded (Fig. 4B, C).

We previously demonstrated that wild-type hDcr prefers to cleave the pre-hlet-7a-1 RNA relative to a perfectly matched duplex RNA substrate [13; 14]. Furthermore, it has also been reported that the ATPase/hel domain is involved in the production of siRNAs from long dsRNA substrates [9; 15]. To further understand the role of the helicase domain in the processing of RNA substrates, we first studied the substrate binding properties of the MBP-ATPase/hel protein using nitrocellulose filter binding assays. The helicase domain prefers to bind to the pre-hlet-7a-1 substrate with a K_d of ~100 nM for the hairpin RNA. In contrast, the helicase domain bound the 37ab RNA with a K_d of ~500 nM, while it did not bind appreciably to a 21-bp RNA (Table 1). Although the helicase domain is involved in RNA substrate binding, it plays no role in determination of the length of hDcr's products (Figs. 2 and 3).

The preferred binding of the helicase domain to pre-hlet-7a-1 may reflect the existence of an interaction between the helicase domain and the terminal loop, which could play an important role in the selection of this type of RNA substrate by hDcr. To test this possibility, we designed a hairpin RNA (37ab-loop, Fig. 5A) containing the perfectly matched stem derived from the 37ab RNA substrate (a slowly-cleaved RNA) and the terminal loop from pre-hlet-7a-1 (a rapidly-cleaved RNA). Dicing assays showed that hDcr cleaves the 37ab-loop substrate with a rate similar to that observed for the wild-type pre-hlet-7a-1 RNA (Fig. 5B, C). Specifically, under single-turnover conditions, the time required to cleave 50% of the labeled substrate ($t_{1/2}$) was approximately 1 min, 3 min, and 65 min for pre-hlet-7a-1, 37ab-loop, and 37ab, respectively (Fig. 5C, left panel). Furthermore, a substrate derived from pre-hlet-7a-1 but containing additional base pairs in place of the loop (hlet7-stem) became an unfavorable substrate, with a cleavage pattern similar to that of the 37ab RNA (Fig. 5C, left panel). Notably, hDcr lacking the helicase domain was capable of cleaving all of the substrates (perfectly matched or bulged dsRNA, or pre-miRNA) in a similar manner (Fig. 5C, right panel). Furthermore, filter binding assays showed that the helicase domain has detectable affinity for single-stranded RNA, with K_d of ~250 nM (Supplementary Fig. S5). Taken together with the above binding data, these results suggest that the ATPase/hel domain plays the role of a “gate-keeper” in order to screen RNA substrates. Its interaction with the single-stranded terminal loop of an RNA hairpin, rather than the dsRNA helix, regulates the dicing activity of hDcr on pre-hlet-7a-1, and perhaps other pre-miRNAs as well.

Discussion

Human Dicer is a 220 kDa RNase III-family endonuclease containing multiple functional domains. How these domains contribute to RNA substrate selection and product size determination remains elusive, mainly due to the technical challenges of preparing recombinant hDcr protein. Here we used limited proteolysis and sequence alignment data to construct hDcr fragments that could be expressed in *E. coli*. Using these hDcr fragment proteins, we determined how the hDcr enzyme selects substrates and produces unique products. Such a “divide-and-conquer” strategy for expressing large proteins in bacteria has worked occasionally for other systems, such as ubiquitin ligases [28]. The over-expression of hDcr in bacteria now provides a powerful tool for further study of the structural and functional relationships of hDcr domains.

Uncovering how Dicer recognizes and processes RNA substrates is essential to understanding the regulatory roles of small RNAs under both physiological and pathological conditions. Biochemical studies of hDcr point mutants, as well as structural studies of *Giardia* Dicer, suggested that the size of Dicer products is determined by the distance between the PAZ domain and the active sites of the RNase III domains [3; 4]. Using separately expressed hDcr fragments, we found that hDcr-C (containing both RNase IIIa/b and the C terminal dsRBD domains) produced 15-bp products, whereas addition of the PAZ domain fragment to these cleavage reactions led to the generation of 22-bp products, the same size product produced by full-length hDcr [4; 13]. These results show that the PAZ domain is a key dsRNA-end anchor domain that establishes the size of Dicer products. This conclusion is consistent with results using a PAZ domain-deleted version of *Giardia* Dicer [29] and point mutations in the PAZ domain in hDcr [6]. Structural studies of the Argonaute PAZ domain revealed a conserved pocket that binds a 7 bp segment of RNA duplex [21; 22]. This finding together with our observation that hDcr-C (lacking the PAZ domain) produces 15-bp products supports the conclusion that the combined footprint of the PAZ and tandem RNase III domains specifies the 22-bp product size generated by hDcr.

The importance of the terminal dsRBD for RNase III cleavage has been unclear. It has been reported that this domain is not required for cleavage by bacterial RNase III [23], while it is essential for *Drosha*'s ribonuclease activity [24]. Here we find that the dsRBD domain in hDcr is required for dsRNA binding only in the absence of the PAZ domain. The dsRBD does not play a role in determining product length, as the hDcr-C protein lacking the dsRBD can generate full-length products provided that the PAZ domain is present. Therefore, the dsRBD domain in full-length hDcr likely plays an auxiliary role in dsRNA binding and cleavage. It could also be involved in protein-protein interactions *in vivo*.

Previously we showed that the helicase domain inhibits the cleavage of pre-siRNA substrates, and that wild-type hDcr cleaves pre-miRNAs at a faster rate [13; 14]. The emerging importance of the ATPase/helicase domain of Dicer proteins in small RNA substrate selection is supported by studies in other eukaryotes. The helicase domain has been shown to discriminate different RNA ends in *C. elegans* and *D. melanogaster* [7] and additional studies of hDcr demonstrate changes in processing long, thermodynamically stable hairpins when the helicase domain is disrupted by a 43 amino acid insertion [15]. By purifying the hDcr helicase domain alone, we could begin to probe the specific interactions that are made with RNA substrates that may explain these differences in RNA processing. We found that the terminal loop in the pre-hlet-7a-1 RNA is a key structure recognized by the hDcr helicase domain. These biochemical data correspond well to an updated model for hDcr as well as recent evidence for pre-miRNA loop interactions with the hDcr helicase domain[11; 30].

Our data showing that deletion of the ATPase/hel domain makes the protein equally active on both the pre-miRNA and dsRNA substrates may indicate that the ATPase/hel domain is a “gate-keeper” that discriminates RNA substrates by interacting with their terminal loops. Several lines of evidence suggest that the terminal loops of pri- or pre-miRNAs play an important role in the biogenesis of miRNAs by interacting with RNA-binding proteins. For example, the RNA-binding protein KSRP promotes the biogenesis of some microRNAs by a direct interaction with their terminal loops [31]. In contrast, Lin-28, a pluripotency factor, and hnRNP A1, a heteronuclear ribonucleoprotein, inhibit the processing of pre-hlet-7a-1 via direct binding to its terminal loop [32; 33]. Our data demonstrate that the putative interaction of the ATPase/hel domain with the terminal loop of pre-hlet-7a-1 adds another level of regulation to miRNA biogenesis. Using the bacterial expression system established in this study, further dissection of the interaction between the ATPase/hel domain and pre-miRNAs will now be possible.

These findings underscore how the PAZ domain cooperates with the RNase III domains to specify the lengths of hDcr products. In addition to this catalytic core, the accessory RNA binding domains ensure accurate processing and selectively. Future investigation will focus on these regulatory mechanisms, particularly how the ATPase/helicase domain of hDcr influences pre-miRNA processing. These results lay a foundation for determining the structural relationships among these hDcr domains, advancing our understanding of RNA biogenesis regulation.

EXPERIMENTAL PROCEDURES

RNA substrates

All of the RNA oligos with exception of human pre-let-7a-1 hairpin RNA (pre-hlet-7a-1) listed below were synthesized by IDT (Integrated DNA Technologies, Inc, Coralville, IA). All RNA oligos were purified by 16% urea-PAGE before use. Pre-hlet-7a-1 (73nt) was transcribed *in vitro* by T7 RNA polymerase from a construct containing a double ribozyme system to ensure homogeneous 5’ and 3’ ends [34]. An artificial hairpin RNA (37ab-loop, 79nt) was made by the splinted ligation [35] of 37a-loop and 5’-phosphorylated 37b-loop (see below) with T4 RNA ligase (New England BioLabs, Inc, Beverly, MA). The 37ab-loop RNA was constructed such that it contains a perfectly matched stem from 37ab (see below) and a terminal loop from pre-hlet-7a-1. The 37ab perfectly matched duplex was formed by annealing the RNA oligos 37a and 37b. The siRNA-mimic dsRNA (21ab) was made by annealing 21a and 21b RNA oligos. The hairpin-stem dsRNA (hlet7-stem) was generated by annealing RNA oligos hlet7-stem-a and hlet7-stem-b. For both filter binding and dicing assays, the purified RNA substrates were 5’-end labeled with ^{32}P using T4 polynucleotide kinase (New England Biolabs, Inc. Beverly, MA), gel-purified and annealed before use. The sequences of all of RNA substrates or oligos used in this study are:

pre-hlet-7a-1: 5’-UGAGGUAGUAGGUUGUAUAGUUUUAGGGUCACACCCACCACUGGGAGAUAACUAUACAAUCUACUGUCUUACC-3’;

hlet7-stem-a: 5’-UGAGGUAGUAGGUUGUAUAGUUUGAAAGUUCACGAUU-3’

hlet7-stem-b: 5’-AAUCGUGAACUUUCAAACUAUACAAUCUACUGUCUUACC-3’

37a-loop: 5’-UGAGGUAGUAGGUUGUAUAGUUUGAUUAGGGUCACACCCACC-3’

37b-loop: 5’-P-ACUGGGAGAUUCAAACUAUACAACCUACUACCUCAUU-3’

37a: 5’-UGAGGUAGUAGGUUGUAUAGUUUGAAAGUUCACGAUU-3’;

37b: 5’-UCGUGAACUUUCAAACUAUACAACCUACUACCUCAUU-3’;

pre-miR20a: 5’UAAAGUGCUUAUAGUGCAGGUAGUGUGUAGCCAUCUACUGCAUUACGAGCACUUAAAG-3’

21a: 5’-UAUACAAUGUGCUAGCUUUCU-3’

21b: 5’-AAAGCUAGCACAUUGUAUAGU-3’

38a: : 5’-UGAGGUAGUAGGUUGUAUAGUUUGAAAGUUCACGAUUAUU

38b: 3’-UUACUCCAUCAUCCAACAUAUCAAACUUUCAAGUGCUAAUAA

Human Dicer and TRBP constructs for Sf9 and bacterial expression

To analyze hDcr, limited proteolysis was performed with endoproteinase Glu-C (Sigma-Aldrich, St. Louis, MO). Specifically, 60 ng of Glu-C was incubated with 30 μg of hDcr on ice for 60 min. The proteolytic fragments were separated on a 10% SDS-PAGE and were then either stained with Coomassie Brilliant Blue and excised for mass spectrometric analysis or transferred onto a PVDF membrane (Millipore, Billerica, MA) for Edman degradation-sequencing.

The N-terminal (hDcr-N: 1-1068) and C-terminal (hDcr-C: 1235-1922) fragments were co-expressed in Sf9 cells transfected with baculovirus expression constructs as described previously [13]. The bacterial expression constructs were designed based on the alignment of published Dicer sequences [4; 18] (ATPase/helicase (ATPase/Hel): 1-604; DUF283-PAZ (DP): 605-1068; hDcr-C: 1235-1922; and hDcr-CΔRBD: 1235-1844). The corresponding DNA fragments were generated by PCR and then cloned into pENTR/TEV/D-TOPO vector (Invitrogen). After being confirmed by sequencing, the correct inserts were sub-cloned into the destination vector pHMGWA-His6-MBP by LR Clonase™ II enzyme mix (Invitrogen). The pHMGWA-His6-MBP vector was kindly provided by Dr. Busso, CNRS/INSERM/ Université Louis Pasteur, France [36].

In this study, human TRBP (hTRBP) was the larger isoform TRBP2, which was over-expressed and purified from bacteria [16].

Filter binding assays

Filter binding assays of hDcr and its different hDcr fragments were performed in the same way as described previously [13]. Briefly, serial dilutions of hDcr protein were incubated in a buffer containing 20 mM Tris-HCl (pH 7.5), 25 mM NaCl, 5 mM EDTA, 1 mM dithiothreitol (DTT), 1% glycerol and ~0.5-1 nM (1500 CPM) of 5'-end ^{32}P-labeled duplex RNA substrate (one strand was labeled) at room temperature for 60 min in a 30 μl total volume. Following incubation, a 25 μl aliquot of each reaction was applied to a dot-blot apparatus equipped with three membranes: Tuffryn, Protran and Nytran (from top to bottom). After drying, the bound (on Protran) or free (on Nytran) RNAs were quantified using a Phosphorimager (GE Healthcare). Percent bound RNA, calculated as the ratio of radioactivity detected on the Protran membrane over the total input radioactivity, was plotted as a function of protein concentration. K_d was determined by global fitting to the equation: $k_{obsd} = (k_{max} \times [Dicer]) (K_d + [Dicer])\text{-}1$, where k_{obsd} is the observed rate constant at a given protein concentration, k_{max} is the maximal rate constant with saturating protein, and K_d is the protein concentration that provides half the maximal rate. Curve fitting was conducted with KaleidaGraph (Synergy Software, Reading, PA).

Dicing assays

Cleavage assays with hDcr were carried out similarly as described previously [13]. Briefly, dsRNA substrates were 5'-end labeled with γ-^{32}P-ATP, annealed and incubated with 30 nM of hDcr (unless otherwise stated in figure legends) at 37°C for the specified time in a 10 μl volume (unless otherwise indicated) containing 20 mM Tris-HCl (pH 6.5), 1.5 mM MgCl2, 25 mM NaCl, 1 mM DTT and 1% glycerol. Reactions were stopped by addition of 1.2 volumes of loading buffer (95% formamide, 18 mM EDTA, 0.025% SDS, 0.1% xylene cyanole FF and 0.1% bromophenol blue). After heating at 70 °C for 10 min, the samples were analyzed by electrophoresis through a 15% polyacrylamide-7M urea gel run in TBE buffer and analyzed using a Phosphorimager, and data quantification was conducted using ImageQuant TL.

Immunoprecipitation assays

15 picomoles of hTRBP and and 5 picomoles of mbp-ATPase/hel-HA proteins were mixed with 15 μl of anti-HA agarose beads in 1XPBS buffer (Sigma, Saint Louis, MO) and incubated and rocked in the cold room for 60 min. The mixture was pelleted by a 30-sec spin at 10,000 × *g* and then washed once with 1xPBS and followed by 5 times with the washing buffer (20 mM HEPES (pH 7.5), 250 mM NaCl, 1% glycerol and 0.1% Triton X-100). After the last wash, the pellet was boiled for 3 min in 1.2X SDS protein loading buffer. As a control, hTRBP alone was processed in the same way.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the members of the Doudna lab for helpful discussions. This work was supported by a grant from the US National Institutes of Health (J.A.D.). M.A.K. was supported by a Berkeley Chancellor's Fellowship. J.A.D. is a Howard Hughes Medical Institute Investigator. The authors declare that they have no competing financial interests.

REFERENCES

1. Siomi H, Siomi MC. RISC hitches onto endosome trafficking. Nat Cell Biol. 2009; 11:1049–51. [PubMed: 19724258]
2. Jinek M, Doudna JA. A three-dimensional view of the molecular machinery of RNA interference. Nature. 2009; 457:405–12. [PubMed: 19158786]
3. Macrae IJ, Zhou K, Li F, Repic A, Brooks AN, Cande WZ, Adams PD, Doudna JA. Structural basis for double-stranded RNA processing by Dicer. Science (New York, N.Y. 2006; 311:195–8.
4. Zhang H, Kolb FA, Jaskiewicz L, Westhof E, Filipowicz W. Single processing center models for human Dicer and bacterial RNase III. Cell. 2004; 118:57–68. [PubMed: 15242644]
5. Nicholson RH, Nicholson AW. Molecular characterization of a mouse cDNA encoding Dicer, a ribonuclease III ortholog involved in RNA interference. Mamm Genome. 2002; 13:67–73. [PubMed: 11889553]
6. Park JE, Heo I, Tian Y, Simanshu DK, Chang H, Jee D, Patel DJ, Kim VN. Dicer recognizes the 5' end of RNA for efficient and accurate processing. Nature. 2011; 475:201–205. [PubMed: 21753850]
7. Welker NC, Maity TS, Ye X, Aruscavage PJ, Krauchuk AA, Liu Q, Bass BL. Dicer's helicase domain discriminates dsRNA termini to promote an altered reaction mode. Mol Cell. 2011; 41:589–99. [PubMed: 21362554]
8. Cenik ES, Fukunaga R, Lu G, Dutcher R, Wang Y, Tanaka Hall TM, Zamore PD. Phosphate and R2D2 restrict the substrate specificity of Dicer-2, an ATP-driven ribonuclease. Mol Cell. 2011; 42:172–84. [PubMed: 21419681]
9. Lee YS, Nakahara K, Pham JW, Kim K, He Z, Sontheimer EJ, Carthew RW. Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. Cell. 2004; 117:69–81. [PubMed: 15066283]
10. Welker NC, Pavelec DM, Nix DA, Duchaine TF, Kennedy S, Bass BL. Dicer's helicase domain is required for accumulation of some, but not all, C. elegans endogenous siRNAs. RNA. 2010; 16:893–903. [PubMed: 20354150]
11. Tsutsumi A, Kawamata T, Izumi N, Seitz H, Tomari Y. Recognition of the pre-miRNA structure by Drosophila Dicer-1. Nat Struct Mol Biol. 2011; 18:1153–8. [PubMed: 21926993]
12. Lim do H, Kim J, Kim S, Carthew RW, Lee YS. Functional analysis of dicer-2 missense mutations in the siRNA pathway of Drosophila. Biochem Biophys Res Commun. 2008; 371:525–30. [PubMed: 18454937]
13. Ma E, MacRae IJ, Kirsch JF, Doudna JA. Autoinhibition of human dicer by its internal helicase domain. Journal of molecular biology. 2008; 380:237–43. [PubMed: 18508075]

14. Chakravarthy S, Sternberg SH, Kellenberger CA, Doudna JA. Substrate-specific kinetics of Dicer-catalyzed RNA processing. J Mol Biol. 2010; 404:392–402. [PubMed: 20932845]
15. Soifer HS, Sano M, Sakurai K, Chomchan P, Saetrom P, Sherman MA, Collingwood MA, Behlke MA, Rossi JJ. A role for the Dicer helicase domain in the processing of thermodynamically unstable hairpin RNAs. Nucleic acids research. 2008; 36:6511–22. [PubMed: 18927112]
16. MacRae IJ, Ma E, Zhou M, Robinson CV, Doudna JA. In vitro reconstitution of the human RISC-loading complex. Proc Natl Acad Sci U S A. 2008; 105:512–7. [PubMed: 18178619]
17. Myers JW, Ferrell JE. Silencing gene expression with Dicer-generated siRNA pools. Methods Mol Biol. 2005; 309:93–196. [PubMed: 15990400]
18. Provost P, Dishart D, Doucet J, Frendewey D, Samuelsson B, Radmark O. Ribonuclease activity and RNA binding of recombinant human Dicer. EMBO J. 2002; 21:5864–74. [PubMed: 12411504]
19. Myers JW, Jones JT, Meyer T, Ferrell JE Jr. Recombinant Dicer efficiently converts large dsRNAs into siRNAs suitable for gene silencing. Nat Biotechnol. 2003; 21:324–8. [PubMed: 12592410]
20. Prucca CG, Slavin I, Quiroga R, Elias EV, Rivero FD, Saura A, Carranza PG, Lujan HD. Antigenic variation in Giardia lamblia is regulated by RNA interference. Nature. 2008; 456:750–4. [PubMed: 19079052]
21. Yan KS, Yan S, Farooq A, Han A, Zeng L, Zhou MM. Structure and conserved RNA binding of the PAZ domain. Nature. 2003; 426:468–74. [PubMed: 14615802]
22. Ma JB, Ye K, Patel DJ. Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. Nature. 2004; 429:318–22. [PubMed: 15152257]
23. Sun W, Jun E, Nicholson AW. Intrinsic double-stranded-RNA processing activity of Escherichia coli ribonuclease III lacking the dsRNA-binding domain. Biochemistry. 2001; 40:14976–84. [PubMed: 11732918]
24. Han J, Lee Y, Yeom KH, Kim YK, Jin H, Kim VN. The Drosha-DGCR8 complex in primary microRNA processing. Genes Dev. 2004; 18:3016–27. [PubMed: 15574589]
25. Daniels SM, Melendez-Pena CE, Scarborough RJ, Daher A, Christensen HS, El Far M, Purcell DF, Laine S, Gatignol A. Characterization of the TRBP domain required for dicer interaction and function in RNA interference. BMC Mol Biol. 2009; 10:38. [PubMed: 19422693]
26. Kok KH, Ng MH, Ching YP, Jin DY. Human TRBP and PACT directly interact with each other and associate with dicer to facilitate the production of small interfering RNA. J Biol Chem. 2007; 282:17649–57. [PubMed: 17452327]
27. Chendrimada TP, Gregory RI, Kumaraswamy E, Norman J, Cooch N, Nishikura K, Shiekhattar R. TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. Nature. 2005; 436:740–4. [PubMed: 15973356]
28. Li T, Pavletich NP, Schulman BA, Zheng N. High-level expression and purification of recombinant SCF ubiquitin ligases. Methods Enzymol. 2005; 398:125–42. [PubMed: 16275325]
29. MacRae IJ, Zhou K, Doudna JA. Structural determinants of RNA recognition and cleavage by Dicer. Nature structural & molecular biology. 2007; 14:934–40.
30. Lau PW, Guiley KZ, De N, Potter CS, Carragher B, MacRae IJ. The molecular architecture of human Dicer. Nat Struct Mol Biol. 2012; 19:436–40. [PubMed: 22426548]
31. Trabucchi M, Briata P, Garcia-Mayoral M, Haase AD, Filipowicz W, Ramos A, Gherzi R, Rosenfeld MG. The RNA-binding protein KSRP promotes the biogenesis of a subset of microRNAs. Nature. 2009; 459:1010–4. [PubMed: 19458619]
32. Rybak A, Fuchs H, Smirnova L, Brandt C, Pohl EE, Nitsch R, Wulczyn FG. A feedback loop comprising lin-28 and let-7 controls pre-let-7 maturation during neural stem-cell commitment. Nat Cell Biol. 2008; 10:987–93. [PubMed: 18604195]
33. Michlewski G, Caceres JF. Antagonistic role of hnRNP A1 and KSRP in the regulation of let-7a biogenesis. Nat Struct Mol Biol. 2010; 17:1011–8. [PubMed: 20639884]
34. Ferre-D'Amare AR, Scott WG. Small self-cleaving ribozymes. Cold Spring Harb Perspect Biol. 2010; 2:a003574. [PubMed: 20843979]
35. Moore MJ, Query CC. Joining of RNAs by splinted ligation. Methods Enzymol. 2000; 317:109–23. [PubMed: 10829275]

36. Busso D, Delagoutte-Busso B, Moras D. Construction of a set Gateway-based destination vectors for high-throughput cloning and expression screening in Escherichia coli. Anal Biochem. 2005; 343:313–21. [PubMed: 15993367]

Highlights

- Active human Dicer can be reconstituted from polypeptides expressed in bacteria
- Dicer's PAZ and catalytic domains contribute to accurate but non-substrate-selective dicing activity
- Dicer helicase domain interacts with RNA hairpin loops to enhance processing of specific substrates
- This system provides a facile means of investigating the molecular mechanisms of human miRNA biogenesis

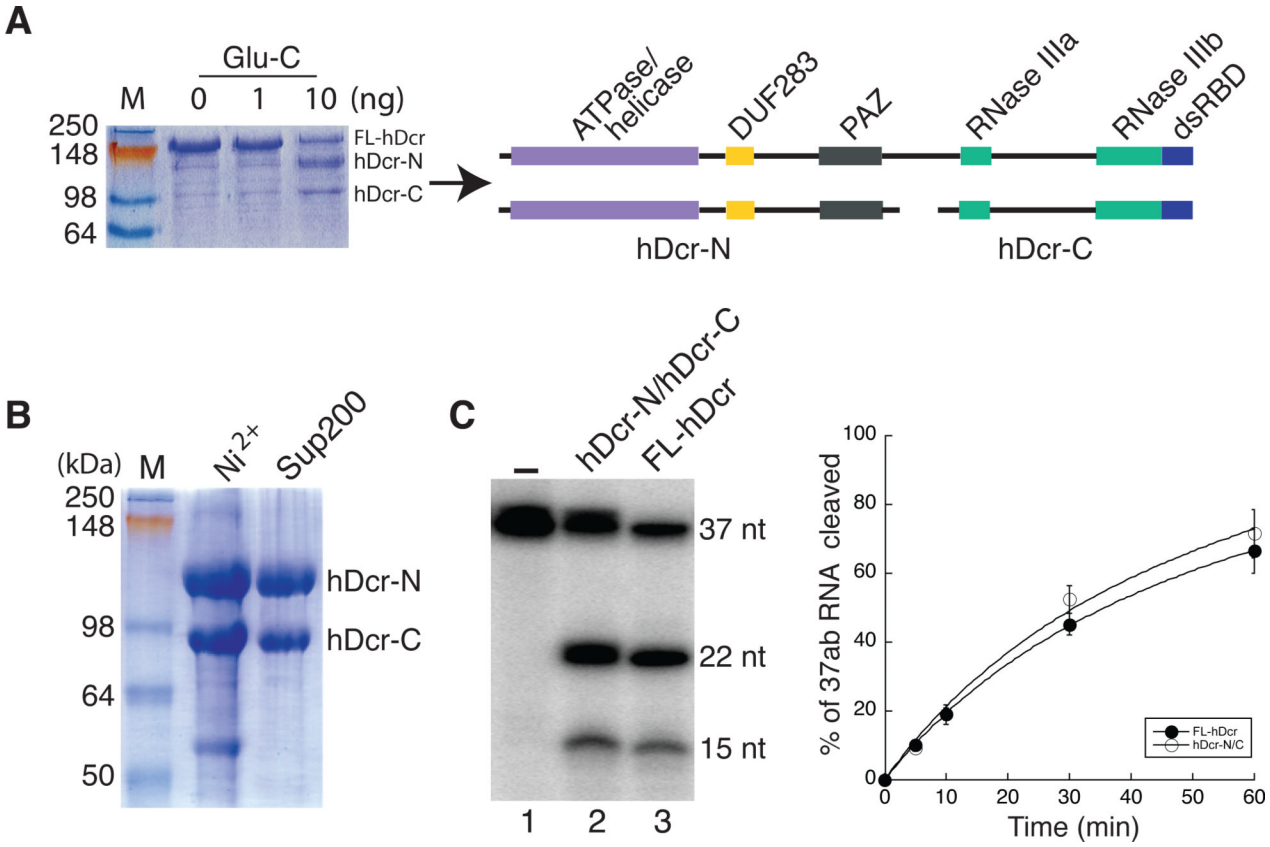

Figure 1. Human Dicer can be separated into functional fragments that interact *in trans*

A. Proteolysis of full-length recombinant hDicer (FL-hDcr) protein. Dose-dependent proteolysis of FL-hDcr protein (10 μg for each reaction) with endoproteinase Glu-C was used to screen for optimal proteolytic conditions (left panel). The two identified globular protein fragments marked with hDcr-N and hDcr-C were isolated for mass spectrometry and Edman degradation sequencing. The isolated fragments of hDcr-N and hDcr-C from the limited partial proteolysis are represented in relation to wild-type FL-hDcr (right panel). **B.** Co-expression of the hDcr fragments in Sf9 cells. An SDS-PAGE gel shows the two protein fragments either from a Ni^{2+}-column (Ni^{2+}) after TEV protease cleavage or from a Superdex 200 size-exclusion column (Sup200). The ~55 kDa band in Ni^{2+} is a non-specific contaminant that was removed through size-exclusion chromatography. M is prestained protein ladder, SeeBlue Plus2 (Invitrogen). **C.** The complex (hDcr-N/hDcr-C) displays cleavage activity similar to that of FL-hDcr. In the cleavage assay, the hDcr-N/hDcr-C complex (lane 2) or FL-hDcr (lane 3) was incubated with the 37ab RNA substrate, of which 37a was ^{32}P-labeled. From this substrate, hDcr generates two products of 22 nt and 15 nt. The single turnover dicing assays with the 37ab RNA substrate show no significant difference between co-expressed hDcr-N/C complex and wild-type hDcr (right panel). Note that single-turnover assays (enzyme in excess over substrate) were used in this study to provide data for comparison to previous published results for hDcr.

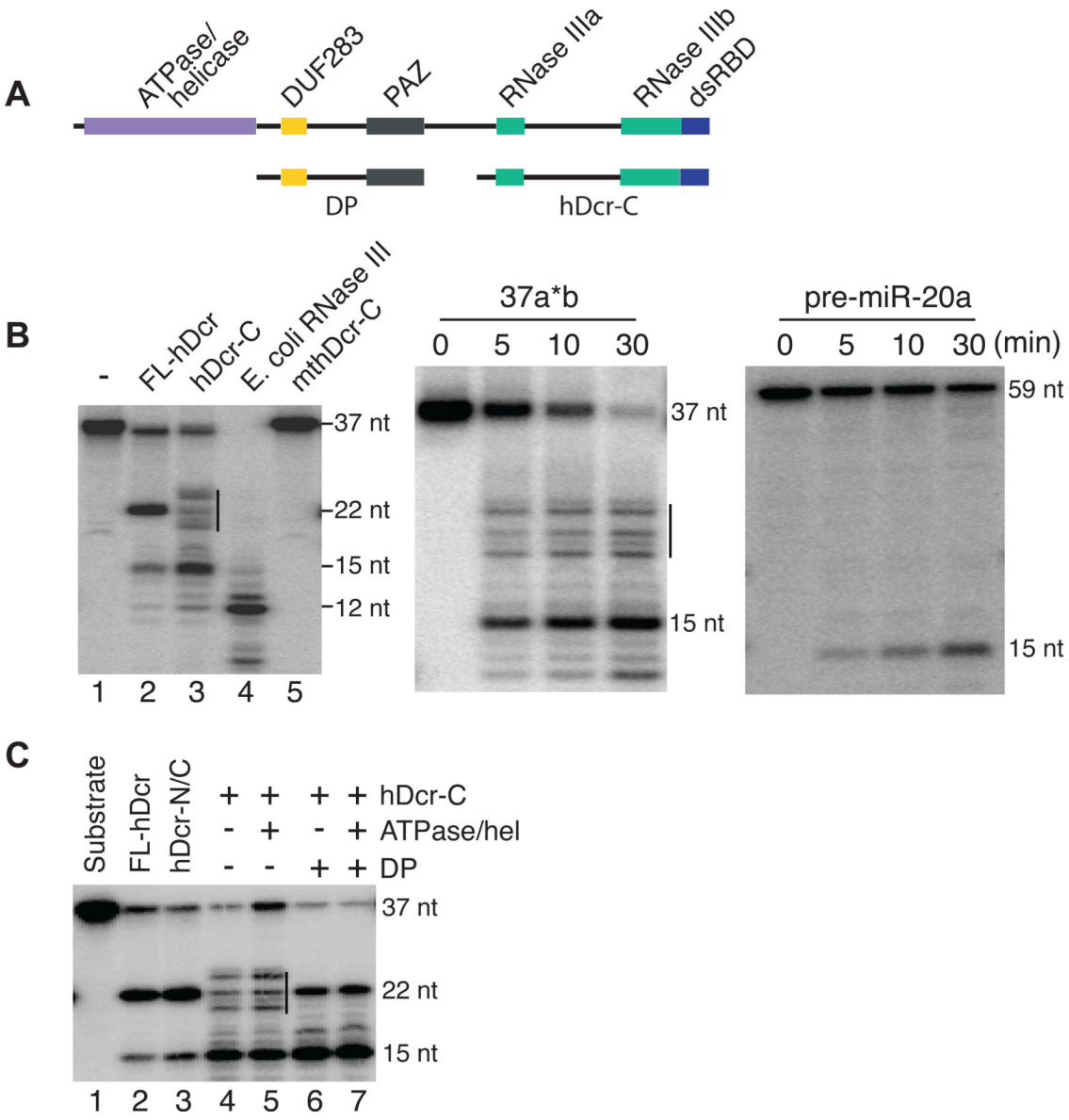

Figure 2. Cooperative action between the PAZ and RNase III domains determines the size of hDcr products

A. Schematic representation of the bacterially expressed tandem DUF283 and PAZ domains (DP) and hDcr-C. **B.** Cleavage assays with hDcr-C. hDcr-C mainly generates 15 nt products from a dsRNA (lane 3), while *E. coli* RNase III gives 12 nt products (lane 4) when a 35bp dsRNA was used. As a negative control, hDcr-C with mutations (mthDcr-C) in the active site glutamines (1316(E/A) and 1705(E/A)) in the RNase III domains displayed no cleavage activity (lane 5). Middle and right panels are the cleavage assays of hDcr-C on a dsRNA (37ab) and a pre-microRNA (pre-miR-20a). In both cases, hDcr-C mainly generates a 15 nt product. **C.** PAZ and RNase III domains together determine the size of hDcr product. Addition of the middle domains of hDcr (DP) to the cleavage reaction (lane 6-7) restored dicing patterns displayed by FL-hDcr (compare lanes 2-3 to lanes 6-7). ATPase/hel domain played no role in determination of the length of cleavage products (compare lane 4 to lane 5, or lane 6 to lane 7). FL-hDcr (lane 2) and hDcr-N/hDcr-C complex (lane 3) were used as positive controls, which generate the 22 nt and 15 nt products. The vertical lines indicate the

intermediate ~21-23 nt products that cannot be cleaved by hDcr-C. The RNA substrate used in these assays was 37ab RNA, of which 37a was 5'-^{32}P-labeled.

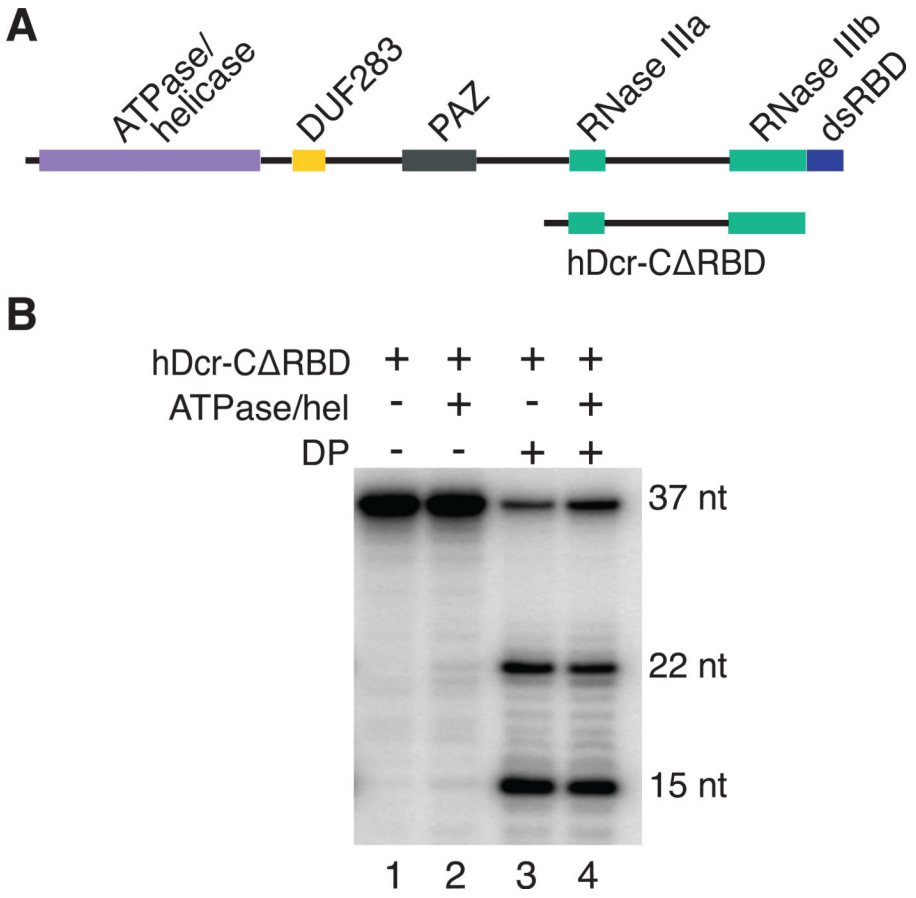

Figure 3. The C-terminal dsRBD is required for RNA binding and cleavage in the absence of the PAZ domain

A. Schematic representation of bacterially expressed hDcr-C without the C-terminal dsRBD (hDcr-CΔRBD). **B.** Requirement of dsRBD for the cleavage activity of hDcr-C. Deletion of dsRBD from hDcr-C fragment abolishes its substrate cleavage activity (lane 1-2). Addition of the middle domains of hDcr (DP) into the cleavage reactions restored FL-hDcr cleavage pattern (lanes 3-4). The ATPase/helicase domain played no role in the cleavage activity (compare lane 1 to lane 2, or lane 3 to lane 4).

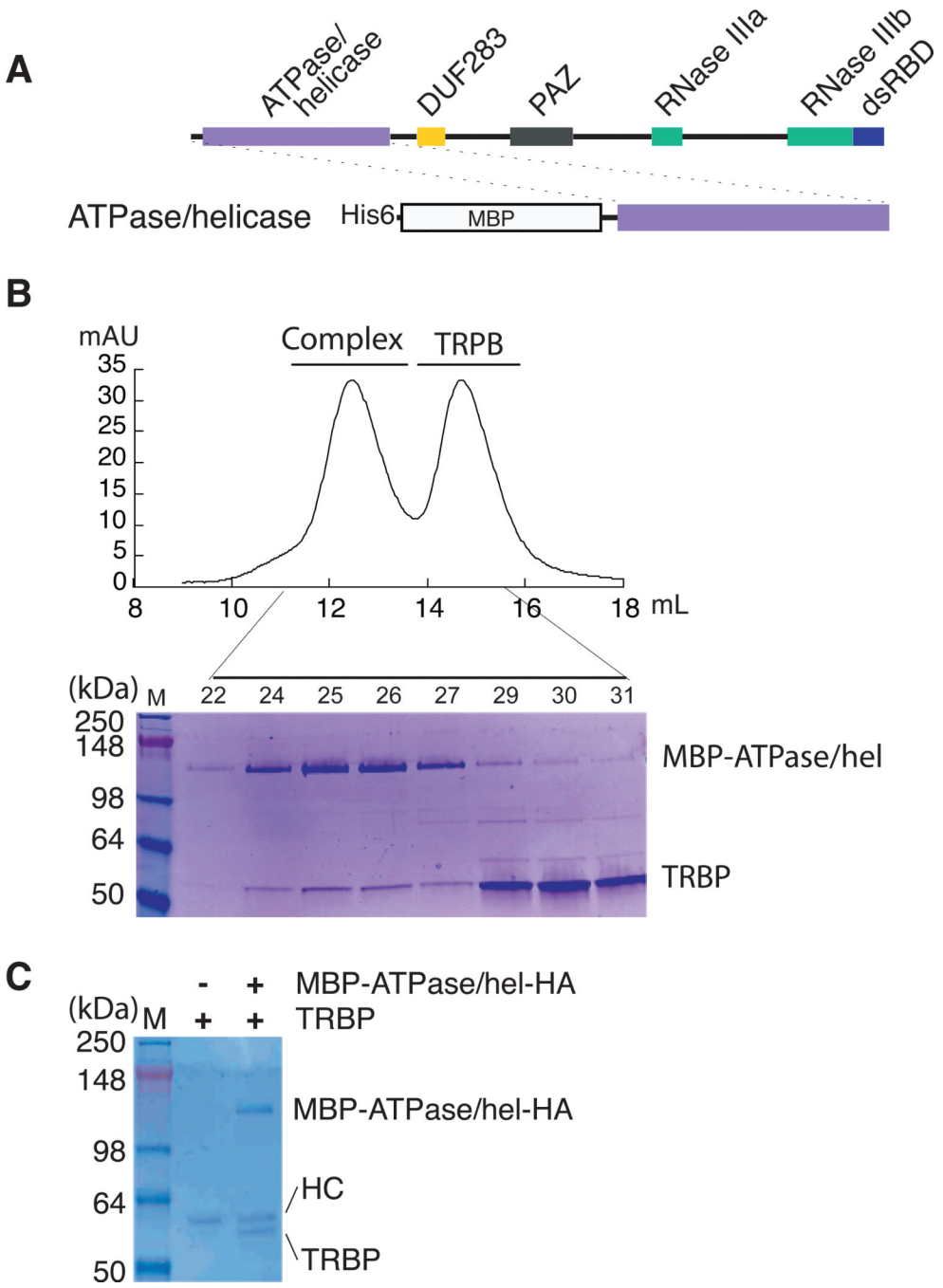

Figure 4. ATPase/Helicase domain of hDcr interacts with hTRBP

A. Schematic representation of bacterially expressed ATPase/hel domain tagged with MBP. **B.** The interaction of MBP-ATPase/hel fragment with hTRBP. A pre-incubated mixture of the MBP-ATPase/hel fragment with 3-fold molar excess of hTRBP was fractionated with a Superdex 200 size-exclusion column (upper panel, elution profile). SDS/PAGE gel analysis of the Superdex 200 fractions indicates that MBP-ATPase/hel and hTRBP interact as shown in the first peak (lower panel). The excess hTRBP elutes in the second peak. **C.** MBP-ATPase/hel can pull-down hTRBP. The MBP-ATPase/hel-domain was purified with a C-terminal hemagglutinin (HA) epitope tag. The two purified proteins (60 pmol of MBP-ATPase/hel and 130 pmol of hTRBP) were incubated on ice with anti-HA antibody agarose

beads (Sigma-Aldrich) for 60 min prior to several washes. The bound proteins are eluted via boiling with 1.2X SDS buffer. HC represents the antibody heavy chain, while the light chain was run out. M is prestained protein ladder, SeeBlue Plus2 (Invitrogen).

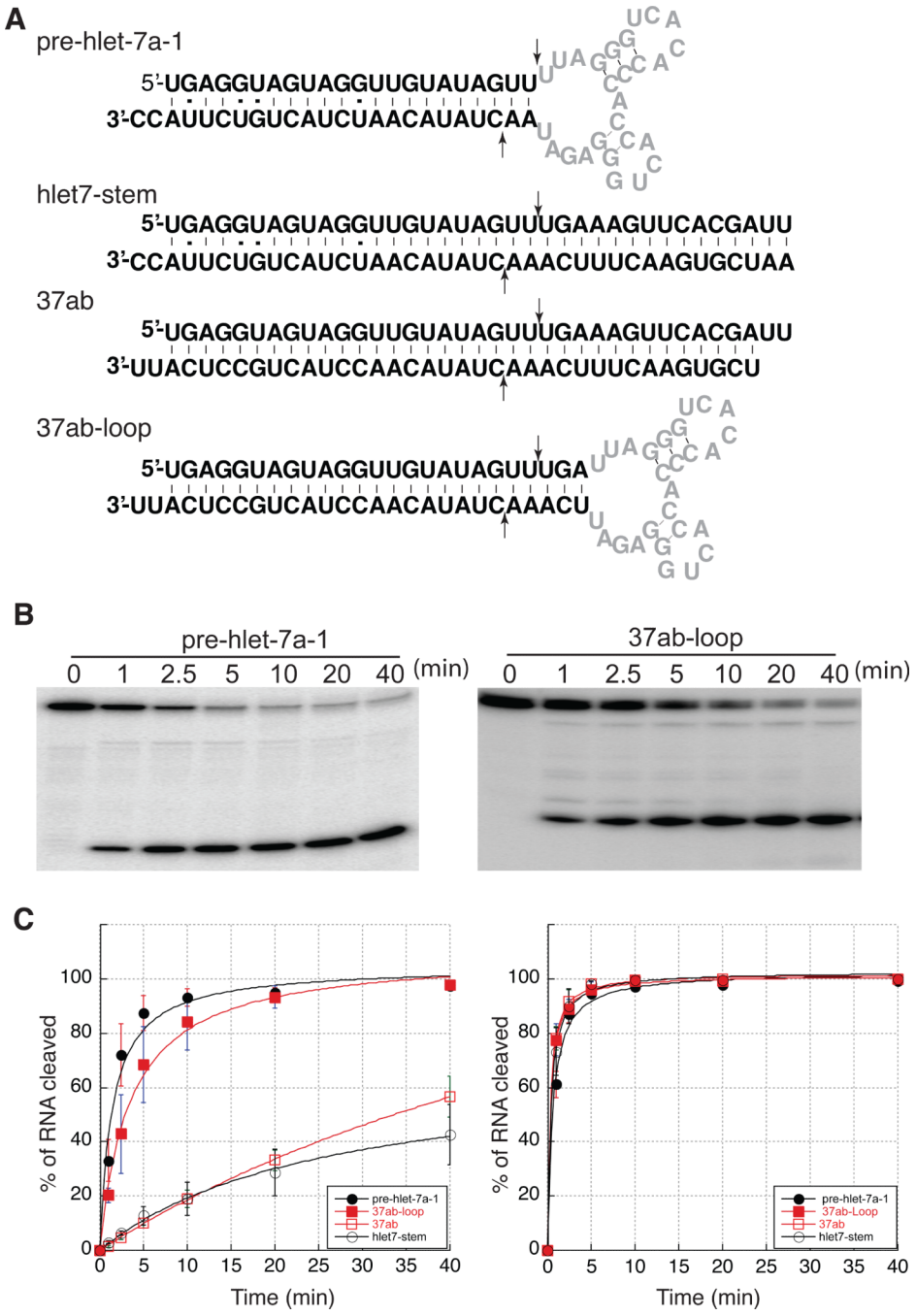

Figure 5. Terminal loop of pre-hlet-7a-1 determines the substrate selection by interacting with the ATPase/helicase domain

A. Schematic representation of four RNA substrates: pre-hlet-7a-1 is abbreviated from human pre-let-7a-1; hlet7-stem is constructed from pre-hlet-7a-1 stem plus an additional 15 bps; 37ab represents a pre-siRNA; and 37ab-loop is an artificial hairpin RNA made of the 37ab stem and the terminal loop from pre-hlet-7a-1. The perfectly-matched base pairs are depicted with vertical lines in the cartoons, while G-U wobbles are marked with dots. Cleavage sites are indicated by arrowheads. The terminal loop structure is predicted from MFOLD and marked with grey color. **B.** Actual cleavage images of a natural hairpin RNA (pre-hlet-7a-1) and an artificial hairpin RNA (37ab-loop). **C.** Interaction of the terminal loop

with ATPase/helicase domain determines processing activity of hDcr. Each graph shows a quantification of dicing assays from FL-hDcr (left panel) and hDcr without ATPase/hel domain (right panel) on the four RNA substrates shown in **A**.

Table 1

K_d values (nM) for human Dicer proteins*

RNA substrate	pre-hlet-7a-1	37ab	21ab
FL-hDicer	39±5	53±8	144±23
MBP-ATPase/hel	96±10	476±30	n.d.
DP	~1000	200±34	220±40
hDcr-C	n.d.	300±40	n.d.
hDcr-CΔRBD	n.d.	n.d.	n.d.

* n.d.=out of the detectable limit