

Structure of the human DICER–pre-miRNA complex in a dicing state

<https://doi.org/10.1038/s41586-023-05723-3>

Young-Yoon Lee^{1,2,4}, Hansol Lee^{2,3,4}, Haedong Kim^{1,2,4}, V. Narry Kim^{1,2,✉} & Soung-Hun Roh^{2,3,✉}

Received: 30 May 2022

Accepted: 14 December 2022

Published online: 22 February 2023

 Check for updates

Dicer has a key role in small RNA biogenesis, processing double-stranded RNAs (dsRNAs)^{1,2}. Human DICER (hDICER, also known as DICER1) is specialized for cleaving small hairpin structures such as precursor microRNAs (pre-miRNAs) and has limited activity towards long dsRNAs—unlike its homologues in lower eukaryotes and plants, which cleave long dsRNAs. Although the mechanism by which long dsRNAs are cleaved has been well documented, our understanding of pre-miRNA processing is incomplete because structures of hDICER in a catalytic state are lacking. Here we report the cryo-electron microscopy structure of hDICER bound to pre-miRNA in a dicing state and uncover the structural basis of pre-miRNA processing. hDICER undergoes large conformational changes to attain the active state. The helicase domain becomes flexible, which allows the binding of pre-miRNA to the catalytic valley. The double-stranded RNA-binding domain relocates and anchors pre-miRNA in a specific position through both sequence-independent and sequence-specific recognition of the newly identified ‘GYM motif’³. The DICER-specific PAZ helix is also reoriented to accommodate the RNA. Furthermore, our structure identifies a configuration of the 5' end of pre-miRNA inserted into a basic pocket. In this pocket, a group of arginine residues recognize the 5' terminal base (disfavouring guanine) and terminal monophosphate; this explains the specificity of hDICER and how it determines the cleavage site. We identify cancer-associated mutations in the 5' pocket residues that impair miRNA biogenesis. Our study reveals how hDICER recognizes pre-miRNAs with stringent specificity and enables a mechanistic understanding of hDICER-related diseases.

Small regulatory RNAs serve as guide molecules in RNA interference (RNAi) by inducing translational repression and destabilization of the cognate mRNAs^{4–6}. Central to the pathway is the ribonuclease (RNase) III enzyme Dicer, which cleaves long dsRNAs or short hairpin RNAs to generate small RNAs of 21–25 nucleotides (nt) in length^{1,2}. Dicer homologues are found throughout eukaryotes and show substantial diversity in their substrate specificity and mechanism of action. Some Dicer proteins are specific to long dsRNAs, as observed in structural and biochemical studies on *Giardia* Dicer, fly Dicer-2 (Dcr-2) and plant Dicer-like proteins^{7–12}. By contrast, other homologues, such as fly Dicer-1 (Dcr-1), are highly selective to hairpin-shaped pre-miRNAs¹³. hDICER can cleave both types of substrate, with a clear preference for short hairpins over long dsRNAs^{14,15}.

hDICER recognizes several features of its substrates: a dsRNA stem of approximately 22 bp; a 2-nt 3' overhang; and a flexible loop next to the cleavage site^{7,12,13,16–19}. The flexible loop is known to be sensed by the helicase domain^{13,14}, whereas the 5' phosphorylated end and the 3' overhang are recognized by basic pockets in the platform and the PAZ (Piwi–Argonaute–Zwille) domains, respectively^{16,17,20}. By anchoring the termini, hDICER can act as a ‘molecular ruler’ to measure around 22 nt

away from the 5' end (‘5' counting rule’) and 3' end (‘3' counting rule’) of the substrate^{7,12,16,17}. In addition, the GYM motif at the cleavage site enables the cleavage site to be precisely determined (see the partner paper to this one³). However, the structural basis of the substrate specificity of hDICER remains largely unknown, owing to the lack of an active-state structure.

Early electron microscopy (EM) analyses of hDICER revealed its overall L-shape^{21–23}. Crystal structures of a partial fragment containing the platform–PAZ domain showed the 5' and 3' pockets, which recognize the respective ends of a small RNA duplex²⁰. A more recent cryo-electron microscopy (cryo-EM) study showed the overall topology of full-length hDICER in the apo and RNA-bound states²⁴. However, in this structure, the pre-miRNA is situated distant from the catalytic valley, probably representing a ‘pre-dicing’ state. Thus, we still lack a structural understanding of how hDICER recognizes pre-miRNAs in an active state.

Here we aimed to determine the cryo-EM structure of hDICER with pre-miRNA in a cleavage-competent state. The structure reveals dynamic spatial rearrangements of several domains of hDICER during the transition to a catalytic state, and explains how hDICER selects its substrates with specificity.

¹Center for RNA Research, Institute for Basic Science (IBS), Seoul, Republic of Korea. ²School of Biological Sciences, Seoul National University, Seoul, Republic of Korea. ³Institute of Molecular Biology and Genetics, Seoul National University, Seoul, Republic of Korea. ⁴These authors contributed equally: Young-Yoon Lee, Hansol Lee, Haedong Kim. [✉]e-mail: narrykim@snu.ac.kr; shroh@snu.ac.kr

Structural determination

To reconstitute the enzyme–substrate complex, we used mono-uridylated pre-let-7a-1, which has an optimal 2-nt 3' overhang, stem length and terminal loop. Meanwhile, in our concurrent study³, we performed massively parallel assays using more than one million pre-miRNAs with random sequences near the cleavage sites. An analysis of the top variants revealed that nucleotide sequences at positions −1, 0 and 1 relative to the 3p DICER cleavage site were enriched with a paired guanine (G), a paired pyrimidine (Y) and a mismatched cytosine or adenine (M), respectively. We therefore termed this motif the GYM motif. The GYM motif robustly enhances dsRNA processing at a specific site, suggesting that it has a role in the catalytic step. The mismatched ‘M’ is particularly important for the decision of cleavage site. To design an optimal substrate for structural determination, as well as to obtain mechanistic insights into GYM-motif-mediated processing, we incorporated the highest-scoring GYM motif (5'-CGC/GCC-3') into pre-let-7a-1 (referred to as pre-let-7a-1^{GYM}) (Extended Data Fig. 1a), which substantially increased the processing rate compared to the wild-type sequence.

We overexpressed and purified the full-length hDICER protein, which cleaved the pre-let-7a-1^{GYM} precisely, yielding the expected 22-nt fragment (Extended Data Fig. 1b–d). Previous cryo-EM structures of hDICER were determined with its accessory protein, TRBP, which exhibits flexible and heterogenous conformations²⁴. As the vast majority of human miRNAs, including let-7, do not require TRBP both in vitro and in cells^{16,25,26}, we set out to determine the cryo-EM structure of hDICER alone so as to reduce the structural heterogeneity.

We first validated the architecture of the purified apo-hDICER at 4.0 Å resolution by cryo-EM (Extended Data Fig. 2a–f and Extended Data Table 1). This structure exhibits a compact hatchet-like (or L-shaped) architecture^{21–23} with mostly globular domains, as observed in the previous cryo-EM structure of the hDICER–TRBP complex²⁴. Our map clearly shows overall domains with defined helical features (Extended Data Fig. 2g,h). The helicase domain shows lower local resolution, suggesting its intrinsically flexible behaviour (Extended Data Fig. 2f). Next, to build a model of apo-hDICER, we introduced the cryo-EM model of hDICER from hDICER–TRBP into our map and refined it by flexible fitting²⁴. Our model covers the three-dimensional (3D) position of most folded domains and partial interdomain loops, including DUF283 linker (D-linker), which connects DUF283 to the platform domain (Extended Data Fig. 2g,h). Notably, D-linker, which was mostly undetermined in the previous hDICER–TRBP structure, forms a short β-sheet with the N-terminal region of RNase III domain a (RIIIDa).

Next, we reconstituted the hDICER–pre-miRNA complex (Extended Data Fig. 1e–g) and were able to reconstruct the map of the complex at 3.0-Å resolution using a cryo-EM dataset of around 1,210,800 particle images (Extended Data Fig. 3a–f and Extended Data Table 1). In this map, hDICER embraces the helix of pre-miRNA within the catalytic centre of the enzyme through an extensive surface contact (Fig. 1a–c). Our map shows the side-chain details of the platform, PAZ, RIIIDa and RIIIDb domains and the secondary structural features for the double-stranded RNA-binding domain (dsRBD) and interdomain linkers. We could not identify the featured densities from helicase (residues 1–564), DUF283 (residues 590–714) and several loops in RIIID domains (residues 1392–1545 and 1595–1684) (Fig. 1a,b). We referenced the model of apo-hDICER and pre-let-7a-1 structures and built an atomic model of hDICER–pre-let-7a-1^{GYM} (Extended Data Fig. 3g–i).

Overall structure in a dicing state

The structure of the hDICER–pre-let-7a-1^{GYM} complex shows that the pre-miRNA is fully docked and poised for cleavage within the catalytic centre that is formed by intramolecular dimerization of RIIIDa and RIIIDb (Fig. 1d). The catalytic centre has two clusters of conserved acidic

amino acid residues in the RIIIDa (E1316, D1320, D1561 and E1564) and RIIIDb (E1705, D1709, D1810 and E1813) (Fig. 1e,f and Extended Data Fig. 4a–c). Our map also shows extra densities in the catalytic core, which are attributed to calcium ions used to substitute magnesium ions to prevent hydrolysis²⁷. The calcium ions are situated near the oxygen atoms of the scissile phosphodiester bonds in the 5p strand (between 22U and 23U) and the 3p strand (between 51G and 52C) (Fig. 1e,f), the position of which coincides with the actual cleavage sites of pre-let-7a-1. This spatial arrangement is highly homologous to that of other RNase III type enzymes, including human DROSHA and *Aquifex aeolicus* RNase III (Aa RNase III)^{28–30} (Extended Data Fig. 4d,e).

We could build the 3D model for most of the pre-miRNA, including the stem region and the additional 4 nt and 6 nt beyond the cleavage sites in the 5p and 3p strands, respectively (Fig. 1c,d). The rest of the terminal loop could not be modelled, affirming the flexible nature of the terminal loop. Our structure shows considerable contacts between hDICER and pre-miRNA, with a total buried surface area of 10,290.8 Å² (Extended Data Fig. 4f). The root-mean-square deviation (RMSD) between the apo and dicing states was 2.9 Å, mostly accounting for differences in the dsRBD and PAZ domain, with RMSDs of 17.7 Å and 13.6 Å, respectively (Extended Data Fig. 4g). Compared to the pre-dicing state²⁴, the dicing-state structure shows large differences both in protein domain organization and RNA interaction. Note that in the pre-dicing state, there is only limited interaction with pre-miRNA, mainly through its termini and loop (with a total buried surface area of 3,206.9 Å²) (Extended Data Fig. 4f).

One of the prominent changes observed during the transition between apo and dicing states was the fade-out of the density of helicase domain. In the apo state structure, there appear to be interdomain interactions among the helicase domain, dsRBD and RIIIDb (Extended Data Fig. 5a). These interactions are likely to support the overall architecture of these domains in a stable ‘closed’ conformation. However, superposition of the apo-hDICER structure, into which the helicase domain could be modelled, shows a steric clash between the helicase domain and the pre-miRNA loop (Extended Data Fig. 5b). Consistently, the N-terminal helicase and DUF283 domains exhibit substantial flexibility in a dicing state in our analysis (Fig. 1b). We observed the same result with pre-miR-3121^{GYM} that has a small loop (11 nt), suggesting that the helicase domain becomes flexible generally during the dicing step regardless of the terminal loop size (Extended Data Fig. 5c). Further in-depth particle classification of around 4,000 particles revealed the extended map of the helicase domain, dislocated from other domains (Extended Data Fig. 5d). To test whether this missing density is due to chemical integrity, we performed another cryo-EM imaging of the same specimens, after incubating with 2 mM MgCl₂, which allowed the cleavage reaction (Extended Data Fig. 5e). We obtained particles at multiple structural states including the dicing state (25%), some intermediates (51%) and apo-like structures (34%) (Extended Data Fig. 5f). This structural analysis indicates that the helicase domain is chemically intact and that the structural heterogeneity of the helicase domain is induced transiently in the dicing state. After the cleavage reaction, the protein returns to the original apo state conformation. Collectively, these analyses suggest that the conformational change in the helicase domain is required to allow productive interaction with the substrate. These observations contrast with structures of fly Dcr-2—specialized in the short interfering RNA (siRNA) pathway—that showed a fixed orientation of the helicase domain, which is engaged in ATP-dependent translocation of the long dsRNA^{10,11}.

Stem recognition by the dsRBD and RIIID

Near the catalytic sites in the upper stem of pre-miRNA, we observed a pronounced movement of the C-terminal dsRBD (Fig. 2a), mediated by the flexible linker that connects to the RIIIDb. This RNA-induced conformational switching orients the dsRBD away from the inner core

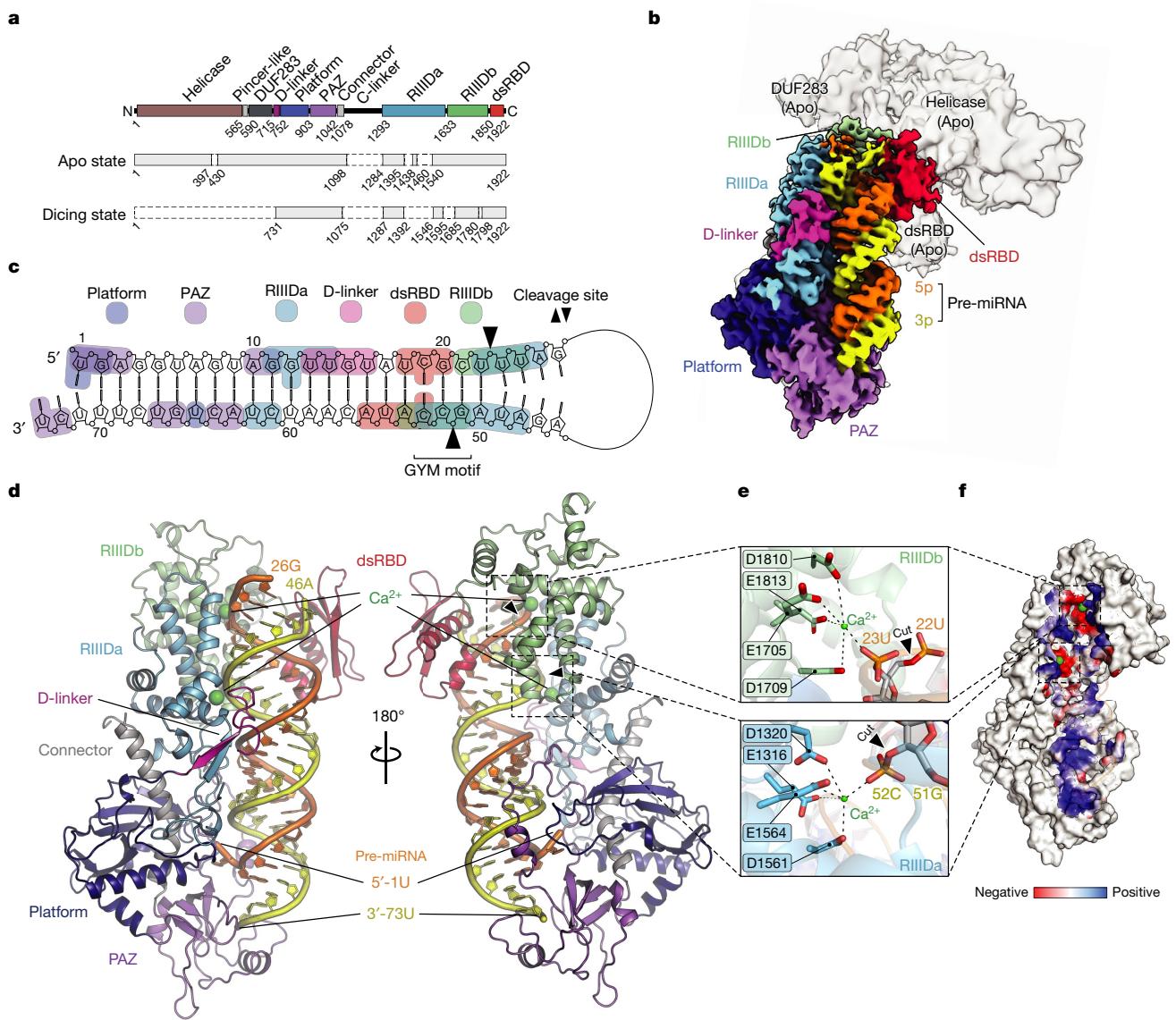


Fig. 1 | Cryo-EM structure of hDICER in complex with a pre-miRNA in a dicing state. **a**, Domain organization of hDICER. Schematics for the apo and dicing states indicate amino acid residues that are included (solid lines) or not included (dashed lines) in the model. C-linker, connector linker; DUF, domain of unknown function. **b**, Cryo-EM map of hDICER in a dicing state (grey, 3.0 Å) overlaying that of hDICER in an apo state (grey, 4.0 Å). **c**, Protein–RNA

interactions in the dicing state at the domain level. Sequences of pre-let-7a-1^{GYM} that are not included in the model are not shown. **d**, Overall structure of the hDICER with a pre-miRNA in a cleavage-competent state. Black arrowheads point to the DICER cleavage sites. **e**, Magnified views of the catalytic sites in the RIIIDa and RIIIDb domains. **f**, Electrostatic potential surface model of the catalytic valley along the protein–RNA interface.

of the catalytic valley and relieves the steric clash between the dsRBD and dsRNA (Extended Data Fig. 6a) so as to permit dsRNA recognition. With respect to its original position in the apo and pre-dicing states, the dsRBD swings about 12.6 Å and 16.5 Å, respectively (Fig. 2a and Extended Data Fig. 6a).

Of note, close to the dsRBD–dsRNA interface, we observed a large change in the helical structure of dsRNA, which deviates from the ideal A-form dsRNA structure (Fig. 2b). This conformational distortion expands the width of the major groove of pre-let-7a-1^{GYM} to 15.6 Å, compared with 8.0 Å in the ideal A-form RNA. The successive major and minor grooves across the region are sandwiched between dsRBD and RIIIDa, and form extensive interactions with both domains (Fig. 2c), suggesting a possible basis for the local distortion in the dsRNA structure. Similar observations were made in the high-resolution cryo-EM structure of *Arabidopsis* DCL-3 (AtDCL3), in complex with a pre-siRNA⁹ (Extended Data Fig. 6b), implying that the conformational distortion in the dsRNA helix is not specific to the pre-let-7a-1 sequence, but induced

by protein–RNA interactions that are unique to a certain group of Dicer homologues.

In addition to the dsRBD, the RIIIDa and RIIIDb domains wrap around the RNA, forming extensive electrostatic interactions with the upper stem region of the pre-miRNA. As well as the contacts at the catalytic core of the DICER cleavage sites, the RIIIDa, situated on the opposite face to the dsRBD-binding site, makes tight interactions with the RNA in the minor groove (Fig. 2c). We observed that α-helices 2 and 3 of RIIIDa potentially interact with the ribose sugars and internucleotide phosphate groups (Fig. 2d). The symmetrically located α-helices 2 and 3 in the RIIIDb may also participate in dsRNA recognition (Fig. 2e).

Interacting with the distorted dsRNA, the dsRBD adopts a canonical $\alpha\beta\beta\alpha$ topology to cover the dsRNA across the minor and major grooves (Fig. 2c). Basically, the reoriented dsRBD interacts with the RNA backbone through its mostly basic patch on the surface (Extended Data Fig. 6c). For instance, in the major groove, the positively charged residues contact the RNA backbone near 19C (Fig. 2f and Extended Data

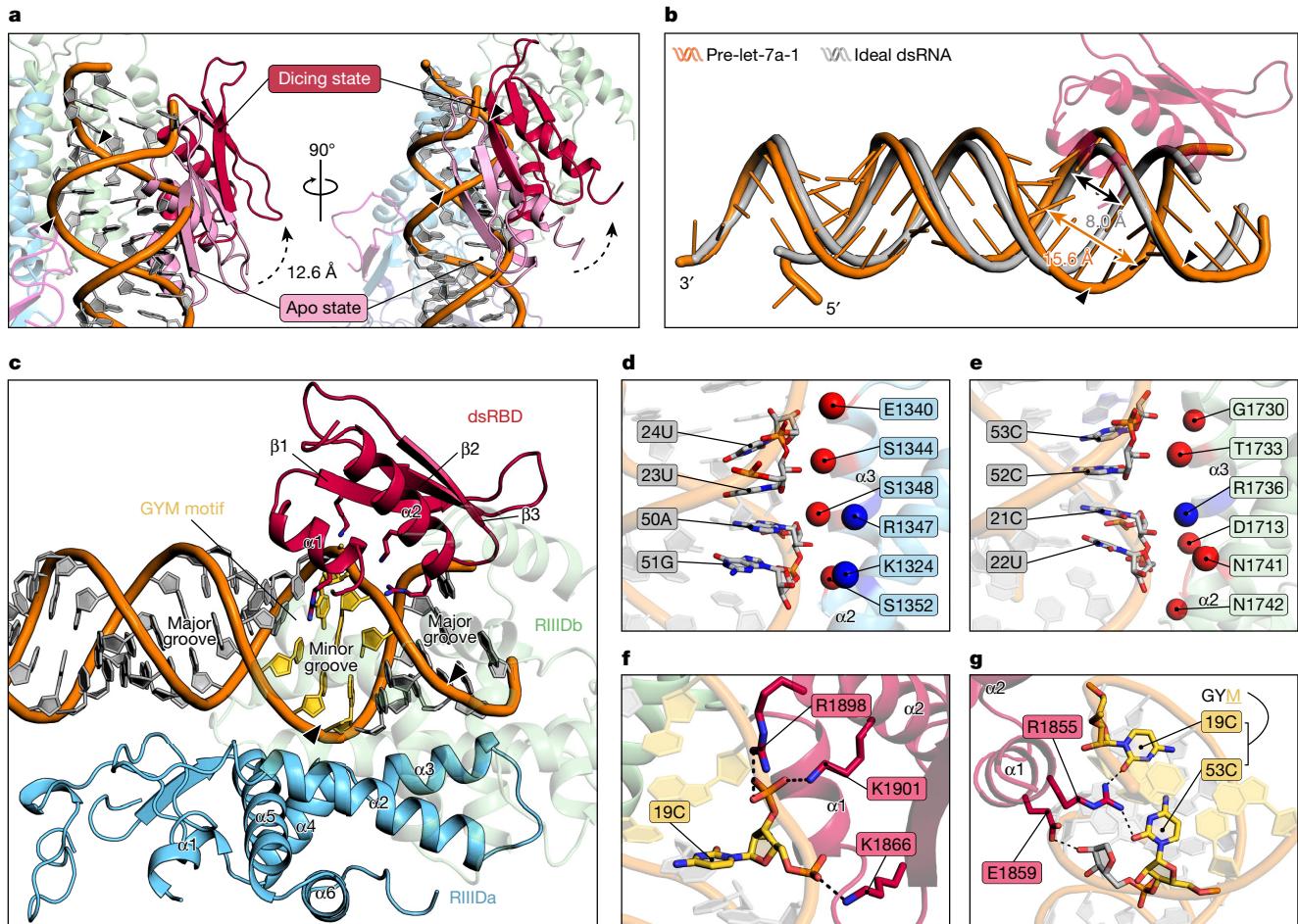


Fig. 2 | Sequence-specific and non-specific binding to RNA by the dsRBD and RIIID domains. **a**, Conformational change in the dsRBD during the transition from an apo state to a dicing state. Black arrowheads near RNA backbones indicate cleavage sites. **b**, Comparison between the structures of ideal A-form dsRNA helix and pre-let-7a-1^{GYM}. **c**, Protein–RNA interactions near the cleavage sites in the minor groove. **d**, Protein–RNA interactions in the interface between RIIIDa and the RNA backbone. Residues that may engage in

electrostatic interactions with the RNA backbone are displayed as balls. **e**, Protein–RNA interactions in the interface between RIIIDb and the RNA backbone. Residues that may engage in electrostatic interactions with the RNA backbone are displayed as balls. **f**, Non-sequence-specific interactions between the dsRBD and the RNA phosphate backbone. **g**, Sequence-specific interactions between the dsRBD and the C–C mismatch of the GYM motif.

Fig. 6d). In the minor groove, however, we observed a sequence-specific interaction between the dsRBD and the RNA. The α -helix 1 of the dsRBD is situated in the vicinity of the GYM motif (Fig. 2c,g), the *cis*-acting element that markedly improves the fidelity of processing (see the partner paper³). An arginine residue (R1855) in α -helix 1 protrudes into the minor groove and may form hydrogen bonds with the C–C mismatch (Fig. 2g and Supplementary Video 1). This is consistent with the observation that mutating this arginine residue abolishes the effect exerted by the mismatch³. Thus, in contrast to previous papers that suggest an auxiliary function of the hDICER dsRBD^{31,32}, our structure provides the structural basis for its predominant role in the selection of the cleavage site, which can even override the effects of 5' and 3' counting mechanisms³. Together, our data indicate that the dsRBD and RIIID domains anchor the upper region of the pre-miRNA and induce local distortion of the RNA, which facilitates the recognition of not only the RNA backbone but also the GYM motif.

Role of the PAZ helix in a dicing state

A previous structural study on a hDICER fragment containing the platform–PAZ domain showed a knob-like protrusion with a small α -helical segment within the PAZ domain, which is specifically found in Dicer²⁰.

This ‘PAZ helix’ (also known as hDICER-specific helix) separates the 5' and 3' pockets and orients the bound RNA away from the surface of hDICER, which is thought to occur in a product-release state and/or pre-dicing state²⁰ (Extended Data Fig. 7a, middle). This helix, however, may be dynamic given that an additional ‘melted’ conformation of the PAZ helix was observed in the platform–PAZ–small RNA duplex complex²⁰ (Extended Data Fig. 7a, right). Indeed, we found a large conformational change in the PAZ helix resulting in a tilt angle of around 54° from its position in the pre-dicing state (Fig. 3a). The PAZ helix is consequently located near the lower stem region of pre-miRNA (Fig. 3a and Supplementary Video 2). Together, our findings show that the spatial rearrangement of the PAZ helix is necessary to allow the pre-miRNA to be aligned parallel to the catalytic valley for subsequent cleavage (Extended Data Fig. 7b).

In addition, this conformational change puts the short stretch of positively charged amino acids ($\text{K}_{109}\text{R}\text{K}\text{K}\text{A}\text{K}_{1023}$) in the vicinity of the negatively charged backbone of the 3p strand (Fig. 3b). To assess the importance of the observed interaction between the PAZ helix and the dsRNA, we introduced mutations by replacing the positively charged residues with five glutamate (E5) or alanine (A5) residues, or by deleting the helix (ΔPAZh) (Fig. 3c). The PAZ-helix-mutant proteins were purified (Extended Data Fig. 1b,c) and incubated with pre-let-7a-1 to

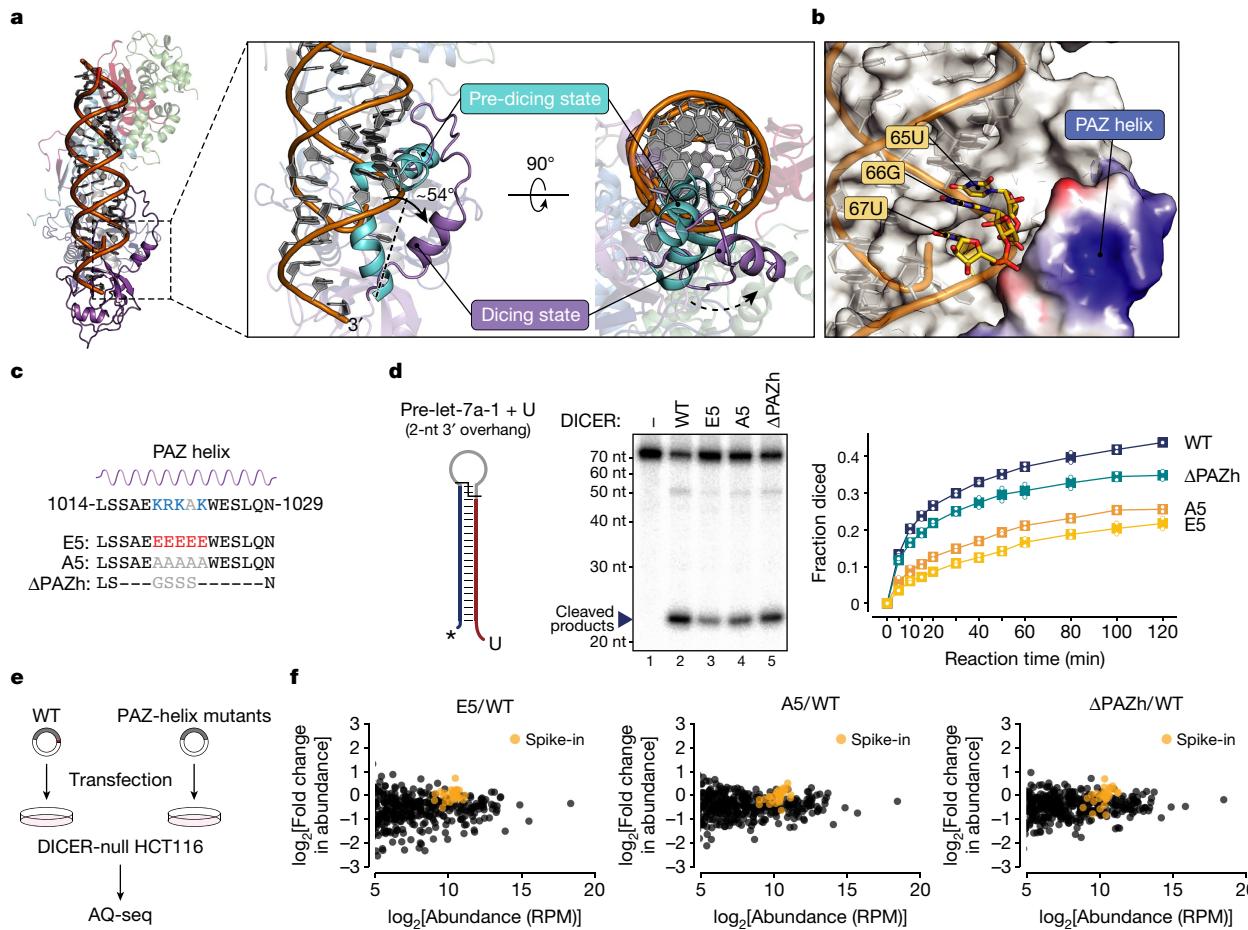


Fig. 3 | The PAZ helix reorganizes to accommodate pre-miRNA in a dicing state. **a**, Conformational change of the PAZ helix between a dicing state (this study) and a pre-dicing state²⁴. **b**, Electrostatic interactions between the positively charged PAZ helix and the negatively charged RNA phosphate backbone. **c**, Sequences of DICER mutants. **d**, In vitro processing of mono-uridylated pre-let-7a-1. Relative cleavage was calculated by quantifying the

band intensity (1 – uncleaved/input). Squares indicate mean ($n = 2$, independent experiments). Asterisk indicates radiolabelled 5' phosphate. For gel source data, see Supplementary Fig. 1. **e**, Schematic outline of the rescue experiment ($n = 2$, biological replicates). **f**, Comparison of miRNA abundance. Spike-ins were used for normalization. RPM, reads per million.

quantitatively measure their effects on the cleavage efficiency in vitro. The mutant proteins showed reduced cleavage efficiencies, regardless of the length of the 3' overhang (Fig. 3d and Extended Data Fig. 7c,d). PAZ helix^{E5} showed a more severe effect than PAZ helix^{A5}, presumably owing to the electrostatic repulsive forces created between the PAZ helix and the RNA backbone. Notably, the deletion of the PAZ helix led to a modest but consistent reduction in cleavage efficiency. This result, together with the structural observations, implies that the PAZ helix has an autoinhibitory effect on the transition to a dicing state besides its contribution to RNA-binding affinity, once the dicing state is achieved.

We next sought to investigate the role of the PAZ helix in miRNA biogenesis by transiently expressing the mutant DICER in *DICER*-knockout HCT116 cells, and then performed Accurate Quantification by Sequencing (AQ-seq), which reliably profiles cellular miRNAs³³ (Fig. 3e). The mutations resulted in a global reduction in the abundance of miRNAs (Fig. 3f), corroborating the in vitro results. Our data collectively suggest that the conformational change in the PAZ helix and its subsequent interaction with the RNA backbone are important for pre-miRNA processing.

Architecture of the 5' and 3' pockets

Consistent with the idea that hDICER recognizes the pre-miRNA termini for accurate processing, our structure shows both 5' and 3' ends stably anchored within the platform and PAZ domains, respectively (Fig. 4a–c

and Extended Data Fig. 8a,b). In the 3' pocket, the last phosphodiester linkage makes close interactions with a cluster of four conserved tyrosine residues (Y936, Y971, Y972 and Y976) and an arginine residue (R937) through potential hydrogen bonding, which is in line with previous reports^{20,24} (Fig. 4c).

The 5' end of pre-miRNA is in a unique kinked conformation (Fig. 4a), which is very different from the previous structures of hDICER with a small RNA duplex or a pre-miRNA in the pre-dicing state and other Dicer homologues in the dicing state^{9,20,24} (Extended Data Fig. 8c–g). The conformation in our structure allows the 5' monophosphate to be inserted into the 5' pocket and possibly interact through hydrogen bonds with the main-chain amide and the amine group of two arginine residues, R996 and R1003, respectively (Fig. 4b and Extended Data Fig. 8a). In addition, the 5' base unexpectedly flips out to interact with a cluster of three arginine residues—R788, R790 and R821—that make hydrogen bonds with the base (Fig. 4b). These results collectively suggest that the 3' pocket is conserved whereas the 5' pocket may have emerged more recently to meet the needs of individual Dicer homologues with different substrate types.

Cancer-associated 5' pocket mutations

The 5' counting mechanism is important for miRNA biogenesis because it ensures accurate and efficient processing regardless of the frequent

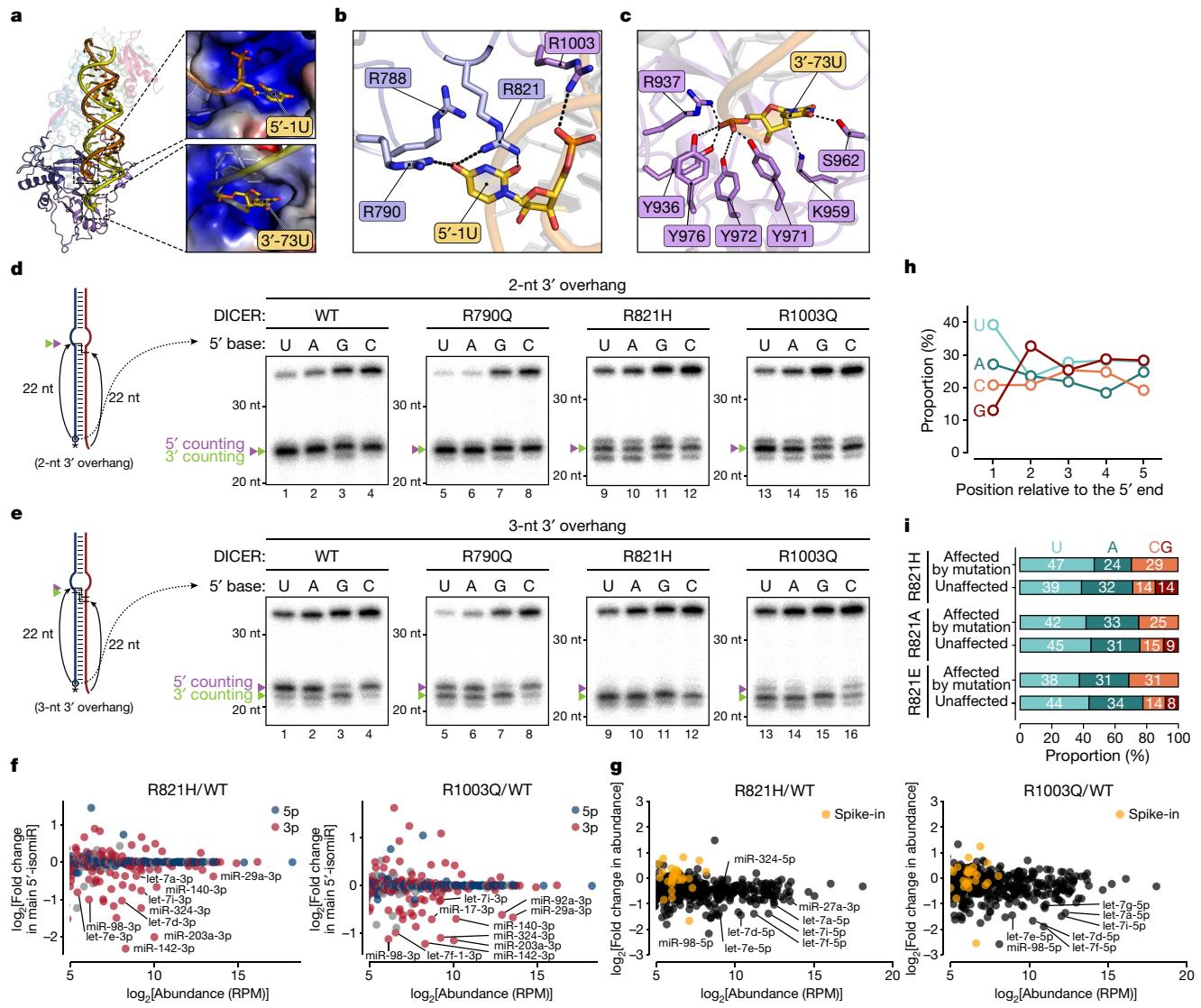


Fig. 4 | Cancer-associated mutations of the 5' pocket affect miRNA biogenesis. **a**, The 5' and 3' ends of pre-miRNA anchored in the basic 5' and 3' pockets in the platform and PAZ domains, respectively. **b**, The 5'-end recognition by the 5' pocket through hydrogen bonding. **c**, The 3'-end recognition by the 3' pocket through hydrogen bonding. **d,e**, Cancer-associated mutations in the 5' pocket impair 5' counting. In vitro processing of duplex RNAs with either a 2-nt (**d**) or a 3-nt (**e**) 3' overhang and a varying sequence at the 5' end. The nucleotide in the 3p strand opposite to the varying sequence is A. Asterisk indicates radiolabelled 5' phosphate. For gel source data, see Supplementary Fig. 1. **f**, Changes in cleavage accuracy, estimated with the fold change of the proportion of the major 5'-isomiR. For a given miRNA, the most abundant

5'-isomiR was identified in the wild-type (WT) sample. Grey, unannotated strand. **g**, Change in abundance of endogenous miRNAs. Spike-ins were used for normalization. **h**, Nucleotide composition of human pre-miRNAs. The proportion of each base at the indicated position relative to the 5' end of pre-miRNAs is shown. miRNAs whose 5p is registered in MirGeneDB (<https://mirgenedb.org/>) ($n = 502$) were included in this analysis. **i**, The 5' terminal base identity of pre-miRNAs. Compare the miRNA groups whose major 5'-isomiRs are affected by the R821 mutation with those that are unaffected. Affected miRNAs ($P < 0.05$, \log_2 -transformed fold change > 0.6 or < -0.6); unaffected miRNAs ($P > 0.05$); P value by two-sided Student's *t*-test compared to wild type.

3' end trimming and tailing in cells¹⁶. Of note, three amino acids in the 5' pocket are mutated in diverse cancer types, according to cancer databases including The Cancer Genome Atlas: R790Q (in colon adenocarcinoma); R821H (in colorectal adenocarcinoma); and R1003Q (in uterine endometrioid carcinoma and rectal adenocarcinoma)^{34,35}.

In vitro, the R821H and R1003Q mutants show defects in cleavage-site selection; they produced multiple products (21–23 nt) from dsRNA with a canonical 2-nt 3' overhang (Fig. 4d, lanes 9–16). Note that wild-type hDICER cleaves this substrate homogeneously because the 5' and 3' counting mechanisms corroborate to ensure the generation of a 22-nt product (Fig. 4d, lanes 1–4). To differentiate the 5' counting and 3' counting mechanisms more clearly, we next used dsRNA with a 3-nt 3' overhang (Fig. 4e). In this assay, all three mutants showed a considerable

reduction in the 5' counting capability, validating our structural observation that these residues are in close contact with the 5' end of RNA. The R821H mutation as well as two other mutations of R821 (R821E and R821A) resulted in a complete loss of 5' counting (Fig. 4e and Extended Data Fig. 8h).

To interrogate the effects on miRNA biogenesis, we performed rescue experiments with wild-type and mutant hDICER (Extended Data Fig. 8i). We observed marked alterations in the DICER cleavage sites of many miRNAs (Fig. 4f and Extended Data Fig. 8j), as indicated by the 5' end of 3p miRNAs, which is determined at the DICER processing step (Extended Data Fig. 8k). The DROSHA processing sites (5' ends of 5p mature miRNAs) were largely unaffected, as expected. The most notable examples are let-7-3p and miR-324-3p, which are dependent on the 5'

pocket^{16,36} (Fig. 4f and Extended Data Fig. 8j). Other 3p miRNAs (such as miR-29a, miR-92a, miR-140, miR-142 and miR-203a) also showed substantial changes in the processing sites, which alter the miRNA isoform landscape and their target repertoire.

Moreover, the abundance of many miRNAs was decreased by the mutations (Fig. 4g). Notably, the mutations led to a significant reduction in the levels of tumour-suppressive let-7 family members—let-7a, let-7d, let-7e, let-7f, let-7i and miR-98—which were shown to be down-regulated in various cancers^{37,38}. Our results indicate the key role of the 5'-end recognition and suggest the oncogenic potential of the 5' pocket mutations.

Role of the 5' base

In the 5' pocket, R821 potentially forms hydrogen bonds with the O2 carbonyl of the 5' terminal uracil (Fig. 4b), hinting at a previously unknown base specificity. To examine the base specificity, we modelled other bases—cytosine, adenine and guanine—into the corresponding position within the 5' pocket structure (Extended Data Fig. 9a). The guanine, unlike other bases, has a 2-amino group that sterically overlaps with the guanidino group of R821, prompting us to experimentally test the effect of the 5' terminal sequence on the 5' counting rule.

In vitro, dsRNAs with 5'-U, 5'-C and 5'-A bases are cleaved predominantly by the 5' counting mechanism, whereas the substrate with 5'-G is cleaved mainly by the 3' counting mechanism (Fig. 4e, lanes 1–4). Even in the presence of a 2-nt 3' overhang, the dsRNA substrate with 5'-G was imprecisely cleaved, yielding 21–22-nt products (Fig. 4d, lanes 1–4), which is consistent with a previous observation that defects in 5' counting give rise to multiple products¹⁶. These observations are not explained by the previous notion that pre-miRNAs with high thermodynamic stability at the 5' terminus do not follow the 5' counting rule¹⁶, because among our substrates, only the dsRNA starting with 5'-U forms a base pair at the terminus. Together, these data suggest that 5'-G may be inefficiently inserted to the 5' pocket, compromising the 5'-end recognition.

To investigate whether R821 is involved in the recognition of the 5' end, we performed the same *in vitro* processing assays with R821 mutants (Extended Data Fig. 1b,c). The R821 mutants completely lost the 5' counting capability, and were no longer influenced by the 5' terminal base regardless of the length of the 3' overhang (Fig. 4d,e, lanes 9–12 and Extended Data Fig. 8h). Thus, R821 has a key role in recognizing the 5' terminal sequence, disfavouring 5'-G.

To further examine the 5' terminal base identity, we looked at the sequences of human pre-miRNAs. We found a substantial enrichment of uridine and adenine—bases that are known to facilitate the loading of the Argonaute (AGO) protein³⁹ (Fig. 4h). Furthermore, we found a marked depletion of guanine compared to cytosine at the 5' end relative to other neighbouring positions, despite the fact that both bases are equally disfavoured by the AGO protein. This suggests an evolutionary pressure against 5'-G in natural pre-miRNAs.

Finally, by examining miRNAs that are affected by the R821 mutations (Fig. 4f and Extended Data Figs. 8j and 9b), we found that the G base is depleted at the 5' end of affected miRNAs, in contrast to those that are unaffected (Fig. 4i). This is in line with the structural prediction (Extended Data Fig. 9a) that U, C or A can be inserted into the 5' pocket, whereas G is incompatible. Thus, owing to the structural restriction posed by R821, hDICER has a nucleotide preference for H (any base but G) at the 5' terminal position for the recognition of the 5' end.

Discussion

We have here captured the cryo-EM structure of hDICER–pre-let-7a-1^{GYM} in a cleavage-competent state. Our structure reveals how hDICER can achieve its substrate specificity through extensive contacts (Fig. 5). Most notably, we discover that a frayed 5' terminal nucleotide other than G ('5'-H') is inserted into the 5' pocket in the platform domain.

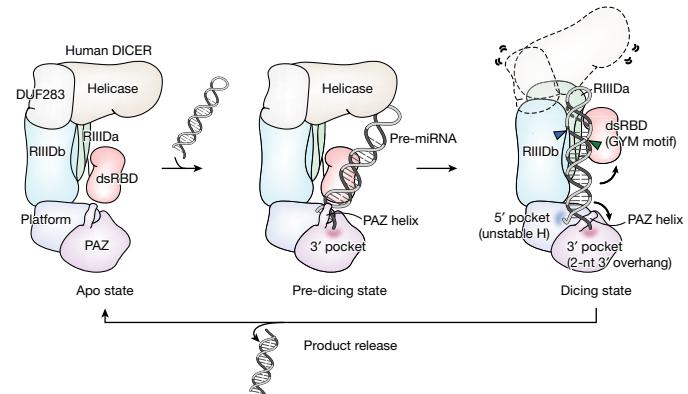


Fig. 5 | Model of the structural transition and substrate recognition of hDICER during the pre-miRNA processing cycle. In the apo state (this study, Protein Data Bank (PDB): SZAK)²⁴, the N-terminal helicase domain, the PAZ helix and the C-terminal dsRBD block the entrance of the pre-miRNA substrate. In the pre-dicing state (PDB: SZAL)²⁴, the 3' end of the pre-miRNA is stably inserted into the 3' pocket and the terminal loop contacts the helicase domain while no major conformational changes occur, limiting the access of RNA to the catalytic centre. In transition to the dicing state, three major conformational changes occur, with the helicase domain in highly flexible conformations, the PAZ helix reorienting to allow a simultaneous recognition of the 5' and 3' termini, and the dsRBD swinging out to accommodate pre-miRNA in the catalytic centre. After cleavage, the product (miRNA duplex) is released and the enzyme returns to the original closed conformation. Note that the *cis*-elements are recognized by the specific domains of hDICER to allow RNA binding and dictate cleavage sites: the GYM motif by the dsRBD and RIIIDA; the thermodynamically unstable 5'-H by the 5' pocket; and the 2-nt 3' overhang by the 3' pocket.

Thermodynamically unstable 5'-H contributes to efficient and accurate processing. We also find that the dsRBD not only interacts with the RNA backbone but also actively engages in a sequence-specific interaction by recognizing the mismatch of the GYM motif. The dsRBD, through the flexible linker, may scan the upper stem of dsRNA, locate the motif and facilitate the cleavage two nucleotides away from the mismatch.

Combined with the apo and pre-dicing states²⁴, our dicing-state structure completes the structural landscape of the pre-miRNA processing cycle (see Fig. 5 for a cartoon model and Supplementary Video 3). In the apo state, hDICER exhibits a compact architecture in which the helicase domain, the PAZ helix and the dsRBD limit the access of RNA to the catalytic valley. During the transition to the pre-dicing state, the protein conformation remains largely unchanged, creating only limited protein–RNA contacts through the 3' pocket, the outer surface of the dsRBD and the helicase domain²⁴. Next, the transition from the pre-dicing to the dicing state induces major changes in the helicase domain, the PAZ helix and the dsRBD to fully dock and poise pre-miRNA for cleavage within the catalytic centre. After cleavage, the product is subsequently released from DICER, and transferred to AGO. Further structural investigation will be required to improve our understanding of the AGO-loading step and the role of TRBP in the process.

Comparisons between Dicer homologues provide valuable insights into the mechanism and evolution of eukaryotic small RNA pathways. Some primitive homologues, such as *Giardia* Dicer, do not have the helicase, the PAZ helix or the dsRBD, implying that these domains may have emerged relatively recently for regulatory purposes. The helicase domains of Dicer homologues seem to be particularly diverse even among higher eukaryotes. Our structure shows that the helicase domain of hDICER becomes largely flexible in a dicing state, potentially serving as the regulatory barrier that contributes to substrate specificity. The helicase domain is also known to interact with TRBP, but the mutations or deletions of the helicase domain do not impair or even enhance the dicing activity, suggesting an autoinhibitory

function¹⁴. Unlike Dicer homologues that processively cleave long dsRNAs using an ATP-dependent translocation by the helicase domain, such as *Drosophila* Dcr-2 and plant Dicer-like proteins^{40,41}, animal Dicer proteins, the main substrates of which are short hairpin-shaped pre-miRNAs, do not require ATP and cannot perform processive dicing on long dsRNAs^{17,19} (Extended Data Fig. 10a). On the basis of sequence similarity, vertebrate Dicer proteins form a separate group from the other Dicer homologues in the miRNA pathway (such as fly Dcr-1), as well as those in the long dsRNA pathway (including fly Dcr-2) (Extended Data Fig. 10b). hDICER prefers hairpins with a flexible terminal loop, but it can also process long dsRNAs—unlike fly Dcr-1, which is highly specialized in pre-miRNA processing^{13,18,42,43}. Together with these observations, our current structure suggests that the helicase domain of hDICER may contribute to processing by interacting with a single-stranded RNA region at the initial binding step, but it is not required for the catalytic step and needs to move flexibly to transit to an open conformation. Also supporting this notion, the helicase and DUF283 domains of an oocyte-specific mouse ΔHEL1 Dicer isoform were difficult to model owing to flexibility⁴⁴. Previous structures of *Drosophila* Dcr-1 and its partner Loqs-PB show an overall similar configuration to that of hDICER, but the helicase part seems to be more rigid⁴⁵. Together, these structures suggest conserved yet unique features of vertebrate Dicer proteins that have evolved to act mainly in the miRNA pathway while still maintaining a role in the siRNA pathway.

Collectively, our work on hDICER will provide practical benefits for the rational design of short hairpin RNAs (shRNAs) and Dicer-substrate siRNAs (DsRNAs) to improve RNAi^{46–48}. Moreover, our structure offers a molecular framework for understanding DICER-related diseases in humans. On the basis of the structure, we identified hDICER mutations in the 5' pocket, which had not been characterized previously. These mutations—R821H and R1003Q—reduce well-known tumour-suppressive miRNAs (for example, let-7 members, miR-324 and miR-142). Given that various DICER mutations have been described in cancers and in a genetic disorder called DICER1 syndrome⁴⁹, functional analyses based on the structure will help us to understand how mutations in various parts of hDICER differentially affect the miRNA population and shape the transcriptome in favour of tumorigenesis.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-023-05723-3>.

- Elbashir, S. M., Lendeckel, W. & Tuschl, T. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev.* **15**, 188–200 (2001).
- Hutvagner, G. et al. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* **293**, 834–838 (2001).
- Lee, Y.-Y., Kim, H. & Kim, V. N. Sequence determinant of small RNA production by DICER. *Nature* <https://doi.org/10.1038/s41586-023-05722-4> (2023).
- Hutvagner, G. & Zamore, P. D. A microRNA in a multiple-turnover RNAi enzyme complex. *Science* **297**, 2056–2060 (2002).
- Mourelatos, Z. et al. miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. *Genes Dev.* **16**, 720–728 (2002).
- Liu, J. et al. Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **305**, 1437–1441 (2004).
- Macrae, I. J., Li, F., Zhou, K., Cande, W. Z. & Doudna, J. A. Structure of Dicer and mechanistic implications for RNAi. *Cold Spring Harb. Symp. Quant. Biol.* **71**, 73–80 (2006).
- MacRae, I. J., Zhou, K. & Doudna, J. A. Structural determinants of RNA recognition and cleavage by Dicer. *Nat. Struct. Mol. Biol.* **14**, 934–940 (2007).
- Wang, Q. et al. Mechanism of siRNA production by a plant Dicer–RNA complex in dicing-competent conformation. *Science* **374**, 1152–1157 (2021).
- Sinha, N. K., Iwasa, J., Shen, P. S. & Bass, B. L. Dicer uses distinct modules for recognizing dsRNA termini. *Science* **359**, 329–334 (2018).
- Su, S. et al. Structural insights into dsRNA processing by *Drosophila* Dicer-2–Loqs-PD. *Nature* **607**, 399–406 (2022).
- Macrae, I. J. et al. Structural basis for double-stranded RNA processing by Dicer. *Science* **311**, 195–198 (2006).

- Tsutsumi, A., Kawamata, T., Izumi, N., Seitz, H. & Tomari, Y. Recognition of the pre-miRNA structure by *Drosophila* Dicer-1. *Nat. Struct. Mol. Biol.* **18**, 1153–1158 (2011).
- Ma, E., MacRae, I. J., Kirsch, J. F. & Doudna, J. A. Autoinhibition of human Dicer by its internal helicase domain. *J. Mol. Biol.* **380**, 237–243 (2008).
- Chakravarthy, S., Sternberg, S. H., Kellenberger, C. A. & Doudna, J. A. Substrate-specific kinetics of Dicer-catalyzed RNA processing. *J. Mol. Biol.* **404**, 392–402 (2010).
- Park, J. E. et al. Dicer recognizes the 5' end of RNA for efficient and accurate processing. *Nature* **475**, 201–205 (2011).
- Zhang, H., Kolb, F. A., Brondani, V., Billy, E. & Filipowicz, W. Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. *EMBO J.* **21**, 5875–5885 (2002).
- Gu, S. et al. The loop position of shRNAs and pre-miRNAs is critical for the accuracy of dicer processing in vivo. *Cell* **151**, 900–911 (2012).
- Feng, Y., Zhang, X., Graves, P. & Zeng, Y. A comprehensive analysis of precursor microRNA cleavage by human Dicer. *RNA* **18**, 2083–2092 (2012).
- Tian, Y. et al. A phosphate-binding pocket within the platform–PAZ–connector helix cassette of human Dicer. *Mol. Cell* **53**, 606–616 (2014).
- Lau, P. W., Potter, C. S., Carragher, B. & MacRae, I. J. Structure of the human Dicer–TRBP complex by electron microscopy. *Structure* **17**, 1326–1332 (2009).
- Wang, H. W. et al. Structural insights into RNA processing by the human RISC-loading complex. *Nat. Struct. Mol. Biol.* **16**, 1148–1153 (2009).
- Lau, P. W. et al. The molecular architecture of human Dicer. *Nat. Struct. Mol. Biol.* **19**, 436–440 (2012).
- Liu, Z. et al. Cryo-EM structure of human Dicer and its complexes with a pre-miRNA substrate. *Cell* **173**, 1549–1550 (2018).
- Kim, Y. et al. Deletion of human *tarbp2* reveals cellular microRNA targets and cell-cycle function of TRBP. *Cell Rep.* **9**, 1061–1074 (2014).
- Wilson, R. C. et al. Dicer–TRBP complex formation ensures accurate mammalian microRNA biogenesis. *Mol. Cell* **57**, 397–407 (2015).
- Provost, P. et al. Ribonuclease activity and RNA binding of recombinant human Dicer. *EMBO J.* **21**, 5864–5874 (2002).
- Partin, A. C. et al. Cryo-EM structures of human Drosha and DGCR8 in complex with primary microRNA. *Mol. Cell* **78**, 411–422 (2020).
- Jin, W., Wang, J., Liu, C. P., Wang, H. W. & Xu, R. M. Structural basis for pri-miRNA recognition by Drosha. *Mol. Cell* **78**, 423–433 (2020).
- Gan, J. et al. Structural insight into the mechanism of double-stranded RNA processing by ribonuclease III. *Cell* **124**, 355–366 (2006).
- Zhang, H., Kolb, F. A., Jaskiewicz, L., Westhof, E. & Filipowicz, W. Single processing center models for human Dicer and bacterial RNase III. *Cell* **118**, 57–68 (2004).
- Ma, E., Zhou, K., Kidwell, M. A. & Doudna, J. A. Coordinated activities of human dicer domains in regulatory RNA processing. *J. Mol. Biol.* **422**, 466–476 (2012).
- Kim, H. et al. Bias-minimized quantification of microRNA reveals widespread alternative processing and 3' end modification. *Nucleic Acids Res.* **47**, 2630–2640 (2019).
- Cerami, E. et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* **2**, 401–404 (2012).
- Gao, J. et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* **6**, pl1 (2013).
- Kim, H. et al. A mechanism for microRNA arm switching regulated by uridylation. *Mol. Cell* **78**, 1224–1236 (2020).
- Takamizawa, J. et al. Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. *Cancer Res.* **64**, 3753–3756 (2004).
- Johnson, S. M. et al. RAS is regulated by the let-7 microRNA family. *Cell* **120**, 635–647 (2005).
- Frank, F., Sonenberg, N. & Nagar, B. Structural basis for 5'-nucleotide base-specific recognition of guide RNA by human AGO2. *Nature* **465**, 818–822 (2010).
- Zamore, P. D., Tuschl, T., Sharp, P. A. & Bartel, D. P. RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* **101**, 25–33 (2000).
- Nykanen, A., Haley, B. & Zamore, P. D. ATP requirements and small interfering RNA structure in the RNA interference pathway. *Cell* **107**, 309–321 (2001).
- Liu, Z., Wang, J., Li, G. & Wang, H. W. Structure of precursor microRNA's terminal loop regulates human Dicer's dicing activity by switching DExH/D domain. *Protein Cell* **6**, 185–193 (2015).
- Nguyen, T. D., Trinh, T. A., Bao, S. & Nguyen, T. A. Secondary structure RNA elements control the cleavage activity of DICER. *Nat. Commun.* **13**, 2138 (2022).
- Zapletal, D. et al. Structural and functional basis of mammalian microRNA biogenesis by Dicer. *Mol. Cell* **82**, 4064–4079 (2022).
- Jouravleva, K. et al. Structural basis of microRNA biogenesis by Dicer-1 and its partner protein Loqs-PB. *Mol. Cell* **82**, 4049–4063 (2022).
- Paddison, P. J., Caudy, A. A., Bernstein, E., Hannon, G. J. & Conklin, D. S. Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev.* **16**, 948–958 (2002).
- Amarzguioui, M. et al. Rational design and in vitro and in vivo delivery of Dicer substrate siRNA. *Nat. Protoc.* **1**, 508–517 (2006).
- Kim, D. H. et al. Synthetic dsRNA Dicer substrates enhance RNAi potency and efficacy. *Nat. Biotechnol.* **23**, 222–226 (2005).
- Hill, D. A. et al. DICER1 mutations in familial pleuropulmonary blastoma. *Science* **325**, 965 (2009).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2023

Methods

Plasmid construction

The DNA encoding hDICER (RefSeq NM_030621) was PCR-amplified by a human cDNA library and subcloned into either a pX vector (for protein purification) containing an N-terminal His₁₀-eYFP-SUMOstar-Strep or a pCK vector (for ectopic expression in *DICER*-knockout cells) with its original CMV promoter replaced with a PGK promoter. Site-directed mutagenesis was performed to introduce mutations for functional studies.

Cell culture and protein expression and purification

Suspension HEK293E cells were subcultured in a 37 °C shaking incubator with a humidified atmosphere of 8% CO₂, in Dulbecco's modified Eagle's medium (DMEM, WELGENE) supplemented with 5% fetal bovine serum (FBS, WELGENE). For a half-litre culture, 0.15 mg of the plasmid encoding the full-length hDICER, N-terminally tagged with His₁₀-eYFP-SUMOstar-Strep tag, was transiently transfected using 1.5 mg of linear polyethylenimine (PEI) and 1% DMSO. After transfection, the cells were incubated at 33 °C.

The entire protein purification steps were performed at 4 °C. The cells were collected after 72 h, washed with cold PBS and resuspended in buffer A (100 mM Tris-HCl (pH 8.0), 150 mM NaCl, 10% glycerol, 2 mM β-mercaptoethanol and 0.1 mM phenylmethylsulfonyl fluoride) supplemented with EDTA-free Pierce Protease Inhibitor Mini Tablets (Thermo Fisher Scientific), 20 µg ml⁻¹ micrococcal nuclease and 5 mM CaCl₂. Cells were subsequently sonicated and the lysate was clarified by centrifugation at 35,000g for an hour. Clarified lysate was loaded onto a column packed with Ni-NTA Superflow resin (Qiagen) that was pre-equilibrated with buffer A. The resin with the bound protein was washed with 5 column volumes of buffer A supplemented with 40 mM imidazole. The protein was eluted with buffer A supplemented with 200 mM imidazole and incubated with SUMOstar protease (LifeSensors) at 4 °C to cleave the N-terminal tag. The cleavage was confirmed by SDS-PAGE. The protein was loaded onto a column packed with Strep-Tactin Superflow (IBA Lifesciences) resin that was pre-equilibrated with buffer A. The resin was then washed with 5 column volumes of buffer A with 2 mM EDTA and subsequently with 5 column volumes of buffer A without EDTA. The protein was eluted with buffer A with 50 mM biotin. To remove uncleaved fusion protein, the eluate was loaded onto Ni-NTA Superflow resin (Qiagen). The unbound protein was concentrated in a 100-kDa molecular weight cut off Amicon Ultra-15 Centrifugal Filter Unit (Merck). The concentrated protein was subjected to size-exclusion chromatography on a Superose 6 Increase 5/150 GL (GE Healthcare) pre-equilibrated with 50 mM Tris (pH 8.0), 100 mM NaCl and 0.5 mM TCEP. For the apo structure, the protein was concentrated to around 1 mg ml⁻¹, snap-frozen in liquid nitrogen and stored at -80 °C. The same purification method was used for mutant DICER proteins.

In vitro reconstitution of the DICER and pre-miRNA complex

Synthetic pre-let-7a-1^{GYM} (Integrated DNA Technologies) (5'-UGAG GUAGUAGGUUGUAUCGGCUUUAGGGUCACACCCACCACUGGGAGAU AGCCAUACAAUCUACUGUCUUUCU-3') was resuspended in 20 mM Tris (pH 7.5), 80 mM NaCl and 1 mM EDTA. The RNA was heated at 80 °C for 5 min and annealed by slowly decreasing the temperature to 4 °C at a rate of -1 °C min⁻¹. Complex formation was performed by mixing DICER and RNA at final concentrations of 15 µM and 45 µM, respectively, in a buffer containing 50 mM Tris-HCl (pH 8.0), 100 mM NaCl, 0.5 mM TCEP and 3 mM CaCl₂. The mixture was incubated on ice for 30 min and loaded onto a Superose 6 Increase 5/150 GL (GE Healthcare) pre-equilibrated with 50 mM Tris (pH 8.0), 100 mM NaCl, 0.5 mM TCEP and 2 mM CaCl₂. The fractions containing the complex were pooled, snap-frozen in liquid nitrogen and stored at -80 °C. The concentration of the complex was evaluated on an SDS-PAGE gel and 15% urea-polyacrylamide gel.

Cryo-EM specimen preparation and data collection

Apo-hDICER. An aliquot (2.5 µl) of around 1.6 µM apo-hDICER sample was applied to glow-discharged 300-mesh UltraAuFoil R1.2/1.3 holey-gold grids (Quantifoil), blotted for 2 s with blot force 5 at 15 °C in 100% humidity and vitrified using Vitrobot Mark IV (Thermo Fisher Scientific) at the Center for Macromolecular and Cell Imaging (CMCI) at Seoul National University (SNU). A total of 5,416 movies (1,390 movies at 1.1 Å per pixel and 4,026 movies at 0.87 Å per pixel) were collected in electron-event representation (EER) mode with a defocus range of -1.0 to -2.75 µm on a 200-kV Glacios (Thermo Fisher Scientific, CMCI at SNU) equipped with a Falcon 4 direct electron camera using EPU software (Thermo Fisher Scientific). The movies at 0.87 Å per pixel were then resampled into 1.1 Å per pixel and combined for further processing. Detailed information on data collection is provided in Extended Data Table 1.

hDICER-pre-let-7a-1^{GYM}. An aliquot (2.5 µl) of around 1.0 µM DICER-pre-let-7a-1^{GYM} complex sample was applied to glow-discharged 300-mesh UltraAuFoil R1.2/1.3 holey-gold grids, blotted for 2 s with blot force 5 at 15 °C in 100% humidity and vitrified using Vitrobot Mark IV in SNU CMCI. A total of 5,964 micrographs were collected at a magnification of 105,000× (corresponding to a calibrated pixel size of 0.849 Å) and a defocus range of -0.9 to -2.2 µm on a 300-kV Titan Krios (Thermo Fisher Scientific, Institute of Basic Science, Daejeon, Korea) equipped with a Gatan K3 BioQuantum (Gatan) detector using EPU software (Thermo Fisher Scientific). Detailed information on data collection is provided in Extended Data Table 1.

hDICER-pre-let-7a-1^{GYM} + Mg²⁺. Around 1.0 µM DICER-pre-let-7a-1^{GYM} complex was incubated with MgCl₂ added at a final concentration of 2 mM for more than 10 min at room temperature. The cryo grid was prepared as described above for hDICER-pre-let-7a-1^{GYM}. A total of 2,882 movies at 1.1 Å per pixel were collected in EER mode with a defocus range of -1.1 to -2.3 µm on a 200-kV Glacios (Thermo Fisher Scientific, CMCI at SNU) equipped with a Falcon 4 direct electron camera using EPU software (Thermo Fisher Scientific).

Data processing and 3D refinement

All image processing was done in cryoSPARC v.3.2 (ref.⁵⁰). Computing resources were used in the CMCI at SNU.

Apo-hDICER. Movies were aligned in 5 × 5 patches in MotionCor2 (ref.⁵¹), and CTF parameters were estimated with GCTF (ref.⁵²). Using template-based autopicking in cryoSPARC, raw particles were initially picked and extracted. After two-dimensional (2D) classification and removing bad particles, 128,635 particles were subjected to 3D heterogeneous refinement using an ab initio model in cryoSPARC. After further 3D classification, global CTF refinements and non-uniform refinement were performed using 128,635 particles, yielding a 4.04-Å map of apo-hDICER based on the gold-standard Fourier shell correlation (FSC) at 0.143. This analysis workflow is illustrated in Extended Data Fig. 2. Detailed information on data processing and 3D refinement is provided in Extended Data Table 1.

hDICER-pre-let-7a-1^{GYM}. Motion correction and CTF estimation was performed in the same process of apo-hDICER. For further automated particle-picking process, an initial particle template was roughly generated based on manually picked 189 particles. Using 3,762,043 particles from an initial template-based picking method, 2D classification and 2D selection was performed. After excluding bad particles and a heterogeneous portion especially containing partial flexible helicase domains, 1,386,301 particles were used for ab initio modelling and 3D classification yielding a homogeneous particle population. Using non-uniform refinement and local refinement, hDICER-pre-let-7a-1 was

Article

reconstructed at a resolution of 3.04 Å map based on the gold-standard FSC at 0.143. This analysis workflow is illustrated in Extended Data Fig. 3. Detailed information on data processing and 3D refinement is summarized in Extended Data Table 1.

hDICER-pre-let-7a-1^{GYM} + Mg²⁺. Motion correction and CTF estimation was performed using the same process as that for hDICER-pre-let-7a-1^{GYM}. An initial particle template was roughly generated based on manually picked 87 particles, and 2,857,379 particles were automatically picked by Blob Tuner with a blob diameter of 30 Å to 300 Å. For further automated particle-picking processes, an initial particle template was roughly generated based on the manually picked 189 particles. After excluding bad particles and a heterogeneous portion especially containing partial flexible helicase domains through 2D classification and 2D selection, 279,763 particles were used for ab initio modelling and 3D classification, yielding a homogeneous particle population. Further subclassification was performed by hetero refinement based on 3D templates generated with structural models of apo-like and dicing-state-like DICER in this study.

Model building

Apo-hDICER. Model building started from an initial protein model from the previously reported hDICER–TRBP structure (PDB: 5ZAK)²⁴. The initial model was fitted into the density map by Dock in map tool from Phenix v.1.18.1 and manually refined in Coot^{53,54}. Then, the model was refined on the Namdinator server using MDFF and Phenix real space refinement default options⁵⁵. After a few rounds, the model was further corrected using Phenix and Coot⁵⁶.

hDICER-pre-let-7a-1^{GYM}. An initial protein model from our apo-hDICER structure and an initial pre-let-7a-1 model from a previously reported hDICER structure (PDB: 5ZAL)²⁴ were rigidly fitted into the density by rigid-body fitting with the Fit in Map tool from Chimera v.1.14. This fitted model was inspected and manually adjusted in Coot, and further refined with phenix.real_space_refine in Phenix, ISOLDE v.1.1.0.

All models were validated by phenix.validation_cryoem⁵⁶. All figures in the manuscript were illustrated by ChimeraX⁵⁷ and PyMol (Schrödinger). For the residual validation analysis, Q-scores for each residue were derived from MapQ of Segger tool⁵⁸ plugged in Chimera v.1.15. B-factor values were derived from real space refinement in Phenix ISOLDE v.1.1.0.

In vitro DICER processing assay

All of the RNA oligos described below were chemically synthesized (Integrated DNA Technologies) and gel-purified before use. To prepare pre-miRNA substrates with 1-nt, 2-nt and 3-nt 3' overhang lengths, RNA oligos (5'-UGAGGUAGUAGGUUGUAUAGUUUUAGGGUCACA CCCACCACUGGGAGAUACUAUACAAUCUACUGUCUUUC(U)(U)-3') were radiolabelled at their 5' ends with [γ -³²P]ATP by T4 polynucleotide kinase (Takara). The RNA oligos were then purified using Oligo Clean & Concentrator (Zymo Research) according to the manufacturer's instructions. The eluted RNAs were annealed in a buffer containing 20 mM Tris (pH 7.5), 80 mM NaCl and 1 mM EDTA, by slowly decreasing the temperature to 4 °C at a rate of -1 °C min⁻¹. To prepare pre-miRNA-like duplex RNA substrates, RNA oligos of the 5'-arm with varying 5' terminal base (5'-(N)GAGGUAGUAGGUUGUAUAGAAAGGACAAAGAG-3') were radiolabelled, purified and annealed to the complementary RNA oligos of the 3'-arm (5'-CUCUUUGUCCAAACUACUUACAACCUACUACCUUAU U(U)-3' as described above.

The RNA substrates were cleaved with the purified DICER proteins added at a final concentration of 0.5 μM, in a buffer containing 50 mM Tris (pH 7.5), 100 mM NaCl, 0.5 mM TCEP, 2 mM MgCl₂ and 5 μg ml⁻¹ yeast RNA. After incubation at 37 °C, the reaction was halted by addition of an equivalent volume of 2× RNA Loading Dye (NEB) supplemented with 0.5 mg ml⁻¹ proteinase K (Roche). The RNA was then resolved on a

15% urea-polyacrylamide gel, along with 5'-end-radiolabelled synthetic miRNAs and Decade Markers System (Ambion). The product was visualized by phosphorimaging with Typhoon FLA 7000 (GE Healthcare).

DICER rescue experiments and data analysis

DICER-knockout HCT116 cells were cultured in McCoy's 5A medium (WELGENE), supplemented with 10% FBS (WELGENE)⁵⁹. Cells were authenticated by short tandem repeat profiling (ATCC) and confirmed to be mycoplasma-negative. To ectopically express DICER proteins, the cells were transiently transfected with DICER-expressing vectors using FuGENE HD (Promega). The cells were treated with TRIzol (Thermo Fisher Scientific) 24 hr 48 h after transfection to extract total RNAs. For library construction, AQ-seq was performed following the protocol described previously³³, except that 100 μg of total RNAs was used and that NEBNext Ultra II Q5 Master Mix (NEB) was used to amplify the cDNAs. The libraries were sequenced on the NovaSeq 6000 platform.

Data were preprocessed as previously described³³. In brief, the 3' adapter with 5' 4-nt degenerate sequences was removed using cutadapt⁶⁰ and FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). Next, short, low-quality and artefact reads were filtered out using FASTX-Toolkit. The output reads were mapped to the spike-in reference first and then the unmapped reads were aligned to the human genome (hg38) using BWA⁶¹. miRNA annotations corresponding to each alignment were retrieved with miRBase release 21 using the intersect tool in BEDTools^{62,63}. Results from this data are summarized in Supplementary Table 1.

Statistics and reproducibility

All in vitro assays were repeated at least twice (Figs. 3d and 4d,e and Extended Data Figs. 1b–g, 5e, 7c,d and 8h). The particle distribution pattern of the representative micrograph of apo-DICER (Extended Data Fig. 2b) repetitively appeared in 3,997 macrographs selected from 5,416 micrographs (Extended Data Fig. 2a). The particle distribution pattern of the representative micrograph of DICER-pre-let-7a-1^{GYM} (Extended Data Fig. 3b) repetitively appeared in 4,817 macrographs selected from 5,974 micrographs (Extended Data Fig. 3a).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The structural models and density maps have been deposited in the PDB under the accession codes 7WX3 (apo-hDICER) and 7WX2 (hDICER-pre-let-7a-1^{GYM}), as listed in Extended Data Table 1. The raw images have been deposited in the Electron Microscopy Data Bank (EMDB) under the accession codes EMD-33490 (apo-hDICER) and EMD-33489 (hDICER-pre-let-7a-1^{GYM}), as listed in Extended Data Table 1. Other structural models cited in this study for analysis (5ZAL, 5ZAK, 2EZ6, 7VG2, 7VG3, 4NGD, 4NHA and 4NH6) are also accessible through the PDB. The rescue data were deposited to the Gene Expression Omnibus (GSE215867).

Code availability

Custom analysis codes are available at https://github.com/haedong-kim615/dicer_dicing_state_structure.

50. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
51. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
52. Zhang, K. Gtcf: real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
53. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D* **75**, 861–877 (2019).

54. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
55. Kidmose, R. T. et al. Namdinator—automatic molecular dynamics flexible fitting of structural models into cryo-EM and crystallography experimental maps. *IUCrJ* **6**, 526–531 (2019).
56. Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D* **74**, 531–544 (2018).
57. Goddard, T. D. et al. UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).
58. Pintilie, G. et al. Measurement of atom resolvability in cryo-EM maps with Q-scores. *Nat. Methods* **17**, 328–334 (2020).
59. Kim, Y. K., Kim, B. & Kim, V. N. Re-evaluation of the roles of DROSHA, Exportin 5, and DICER in microRNA biogenesis. *Proc. Natl Acad. Sci. USA* **113**, E1881–E1889 (2016).
60. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.j.* **17**, 10–12 (2011).
61. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
62. Kozomara, A. & Griffiths-Jones, S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.* **42**, D68–73 (2014).
63. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

Acknowledgements We thank J.-S. Woo for the mammalian cell transfection protocol and Y.-G. Choi, S.-M. Ji, J. Yang, D.-E. Choi, S. Bang and E. Kim for technical assistance. This research was supported by the Institute for Basic Science from the Ministry of Science and ICT of

Korea (IBS-R008-D1 to Y.-Y.L., H.K. and V.N.K.); BK21 research fellowships from the Ministry of Education of Korea (to Y.-Y.L. and H.K.); and the National Research Foundation of Korea (NRF-2018-Global Ph.D. Fellowship Program to Y.-Y.L. and NRF-2015-Global Ph.D. Fellowship Program to H.K.). S.-H.R. acknowledges financial support from the Creative-Pioneering Researchers Program through Seoul National University, NRF grants (2019M3E5D6063871, 2019RIC1C1004598, 2020R1A5A1018081 and 2021M3A914021220) and the SUHF Foundation. Computing resources were used in the CMCI at SNU and the Global Science Experimental Data Hub Center (GSDC) at Korea Institute of Science and Technology Information (KISTI).

Author contributions All of the authors conceived the project. V.N.K. and S.-H.R. collected financial support. Y.-Y.L. performed protein purification and cryo-EM sample preparation. Y.-Y.L. and H.K. performed biochemical and cellular experiments. H.L. and S.-H.R. performed structural studies. H.K. performed bioinformatic analyses.

Competing interests The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41586-023-05723-3>.

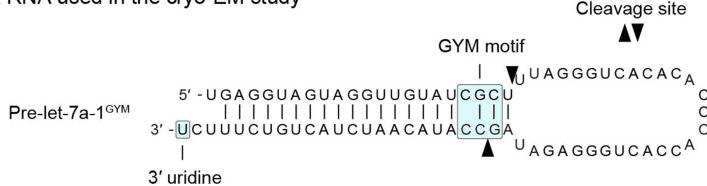
Correspondence and requests for materials should be addressed to V. Narry Kim or Soung-Hun Roh.

Peer review information *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

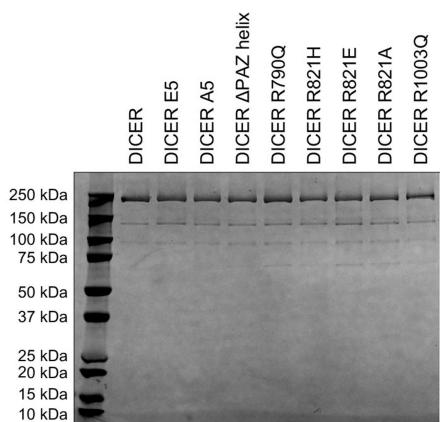
Reprints and permissions information is available at <http://www.nature.com/reprints>.

Article

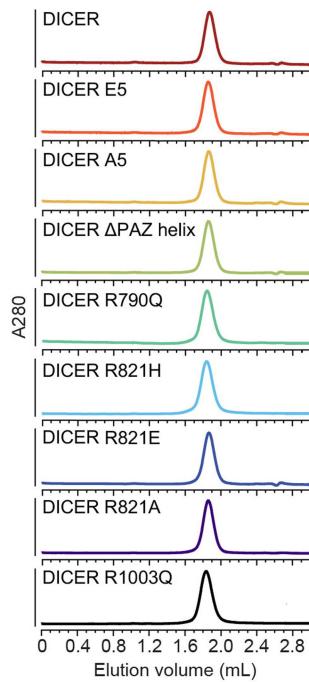
a RNA used in the cryo-EM study



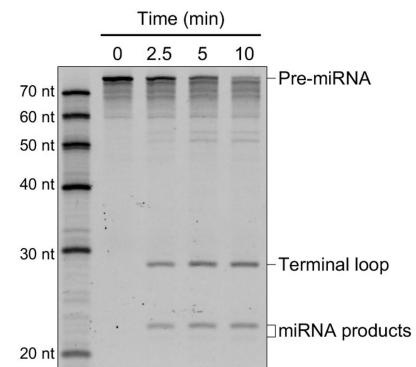
b SDS-PAGE



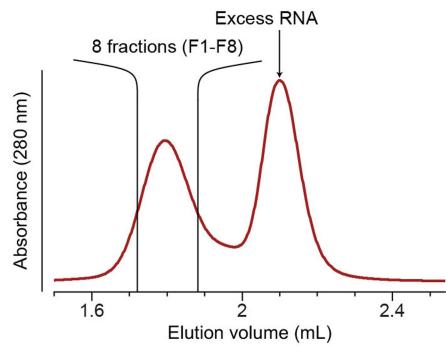
c Size-exclusion chromatography



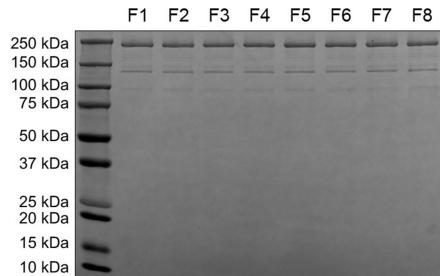
d Cleavage assay



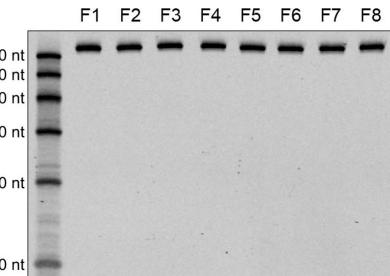
e Complex formation



f SDS-PAGE of protein-RNA complex



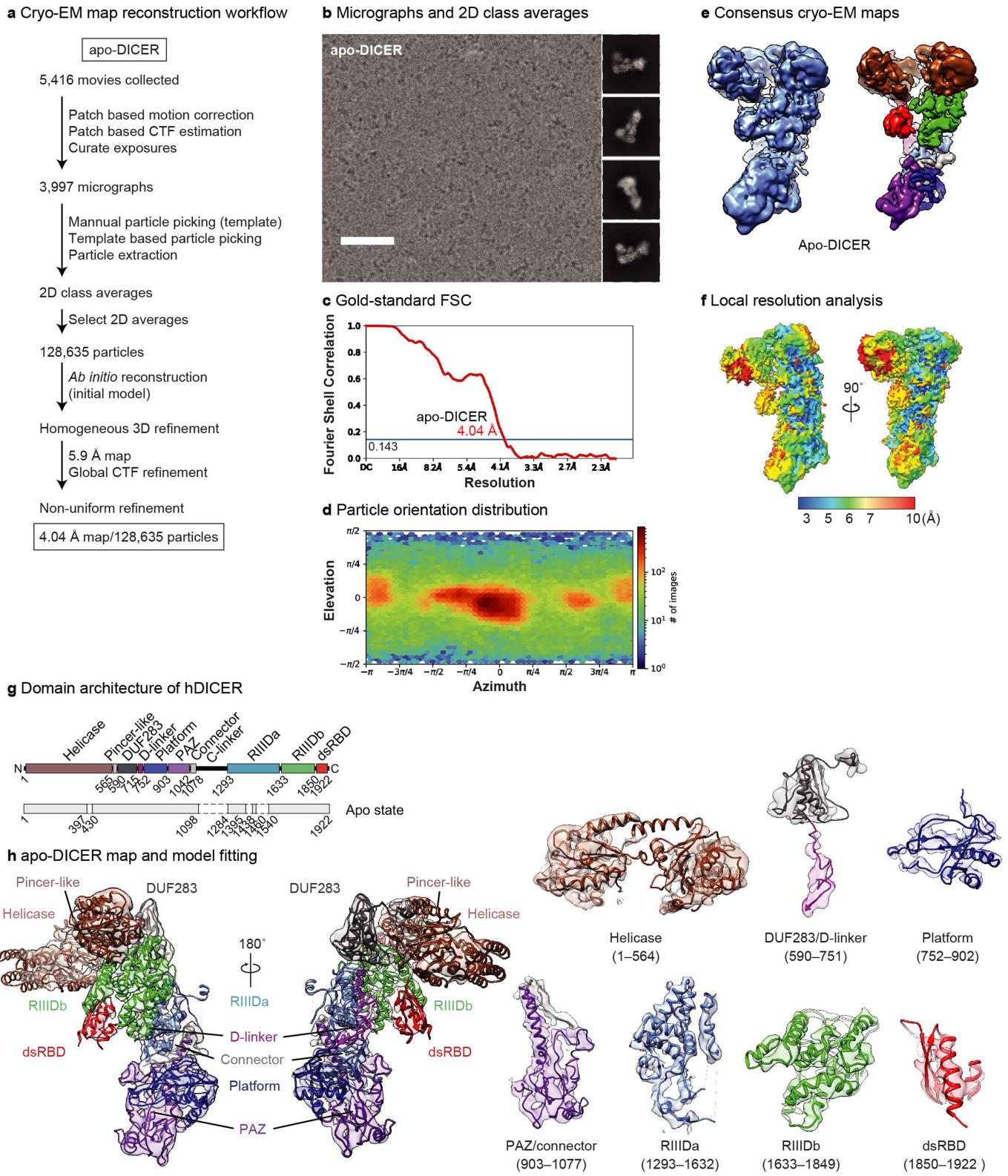
g Urea-PAGE of protein-RNA complex



Extended Data Fig. 1 | Purification of hDICER and the hDICER–RNA complex.

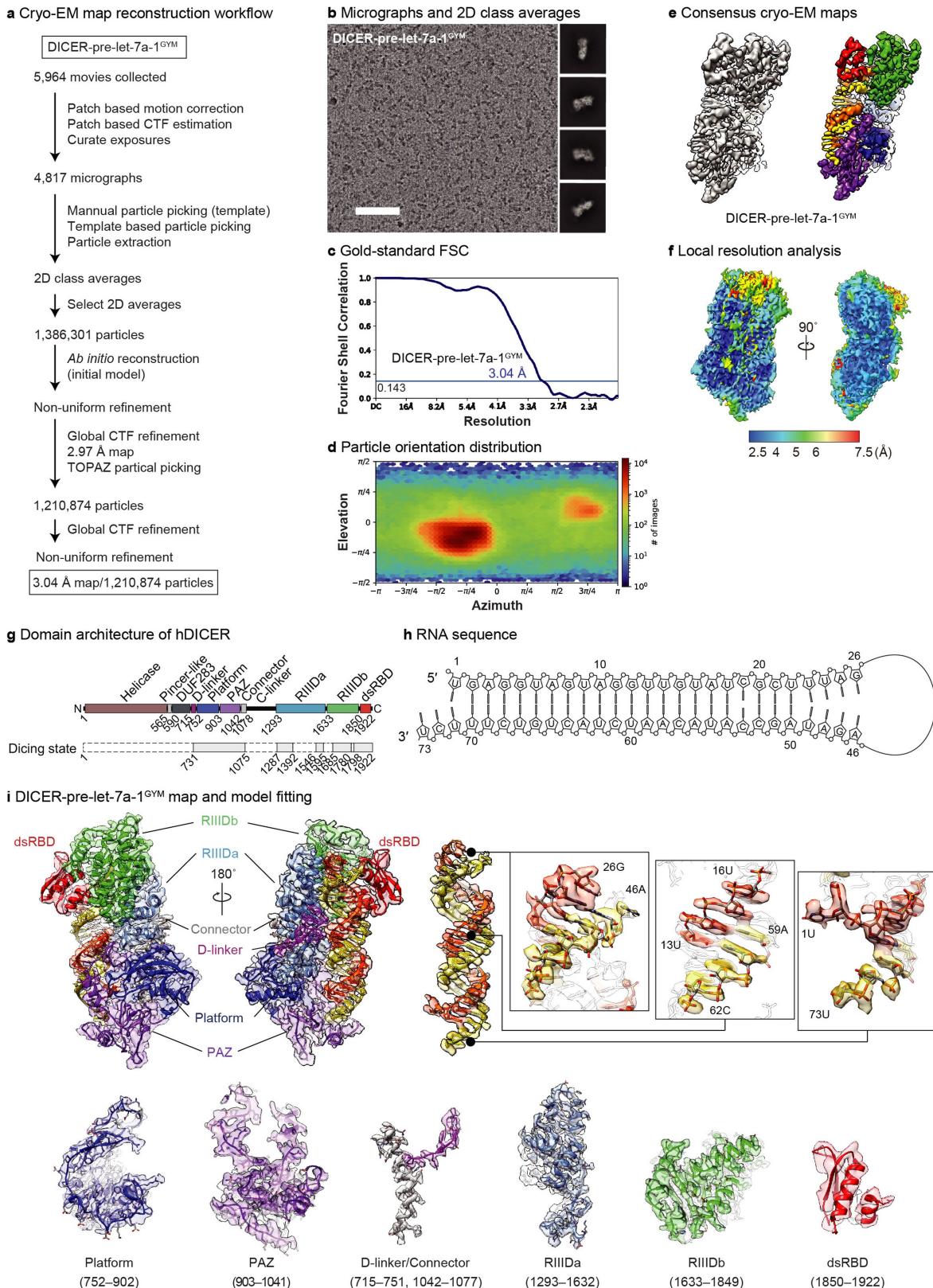
a, The sequence of pre-let-7a-1^{GYM} used for structural determination. **b**, SDS-PAGE of wild-type and mutant hDICER proteins. **c**, Size-exclusion chromatography of purified proteins. **d**, In vitro processing of pre-let-7a-1 by purified hDICER. **e**, Size-exclusion chromatography of the hDICER–pre-let-7a-1^{GYM} complex. **f**, SDS-PAGE of the hDICER–pre-let-7a-1^{GYM} complex visualized by

Coomassie blue staining. Protein concentration for each fraction was estimated by Bradford protein assay, and the same amount of protein was loaded for each fraction. **g**, Urea-PAGE of the hDICER–pre-let-7a-1^{GYM} complex visualized by SYBR gold staining. RNA concentration for each fraction was estimated by absorbance at 260 nm, and the same amount of RNA was loaded for each fraction. For gel source data, see Supplementary Fig. 1.



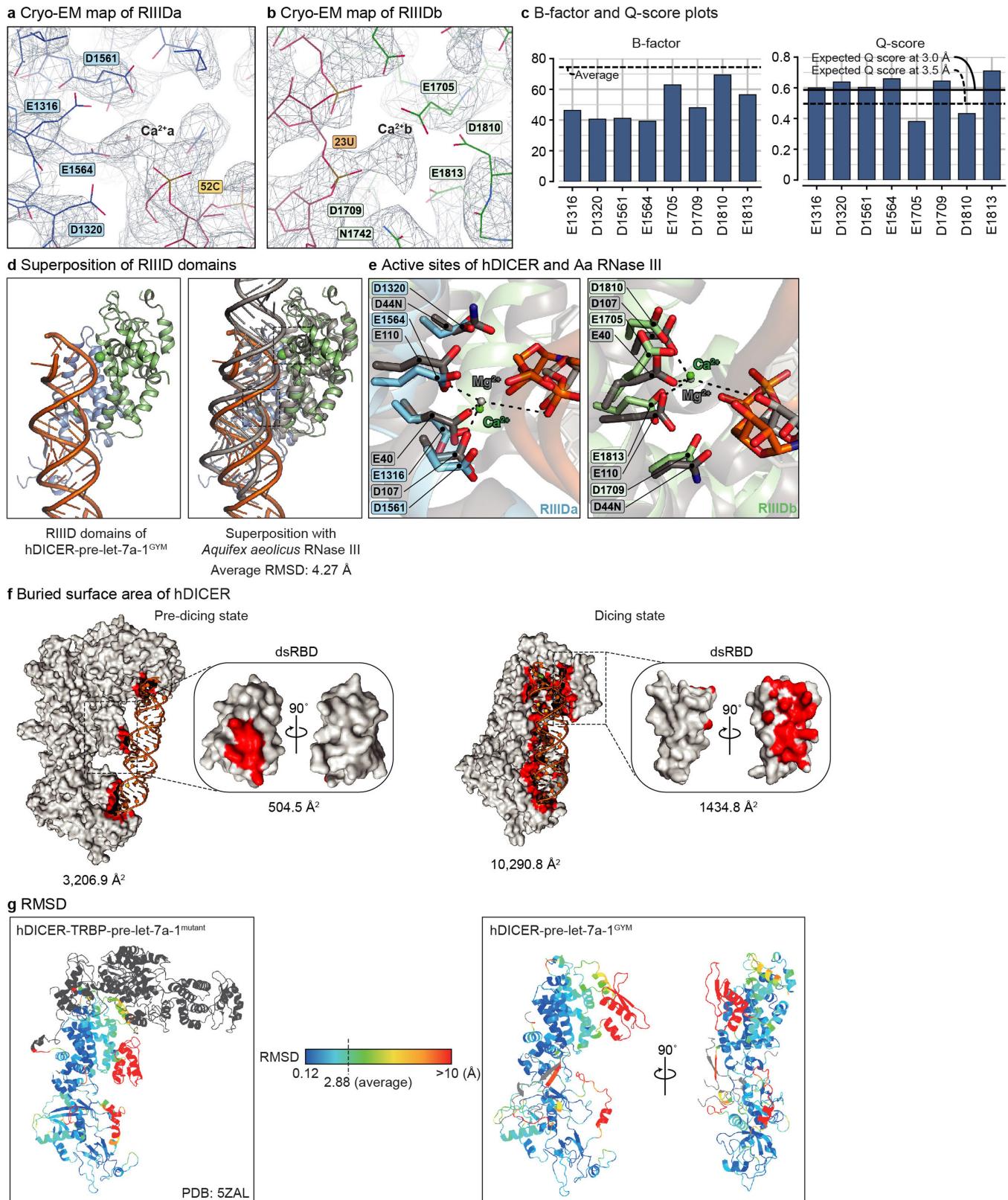
Extended Data Fig. 2 | Cryo-EM image processing procedure for apo-hDICER. **a**, Overview of the image processing procedure (see Methods). **b**, Representative micrograph and 2D class averages of the apo-hDICER (scale bar, 50 nm). **c**, Gold-standard FSC at 0.143 of the apo-hDICER. **d**, Angular particle distribution heat map. **e**, Consensus map of apo-hDICER. Each domain is indicated in a different colour. **f**, Local-resolution analysis shown in rainbow scale. **g**, Domain organization of hDICER with colour code for each domain. Schematics for the apo state shows amino acid residues included (solid lines) or not included (dashed lines) in the model. **h**, Atomic model fitting to the map of apo-hDICER.

Article



Extended Data Fig. 3 | Cryo-EM image processing procedure for hDICER-pre-let-7a-1^{GYM}. **a**, Overview of the image processing procedure (see Methods). **b**, Representative micrograph and 2D class averages of hDICER-pre-let-7a-1^{GYM} (scale bar, 50 nm). **c**, Gold-standard FSC at 0.143 of the hDICER-pre-let-7a-1^{GYM}. **d**, Angular particle distribution heat map. **e**, Consensus map of hDICER-pre-let-7a-1^{GYM}. Each domain is indicated in a different colour. **f**, Local-resolution

g, Domain organization of hDICER with colour code for each domain. Schematics for the dicing state shows amino acid residues included (solid lines) or not included (dashed lines) in the model. **h**, Sequence of pre-let-7a-1^{GYM} in the model. **i**, Atomic model fitting to the map of hDICER-pre-let-7a-1^{GYM}.

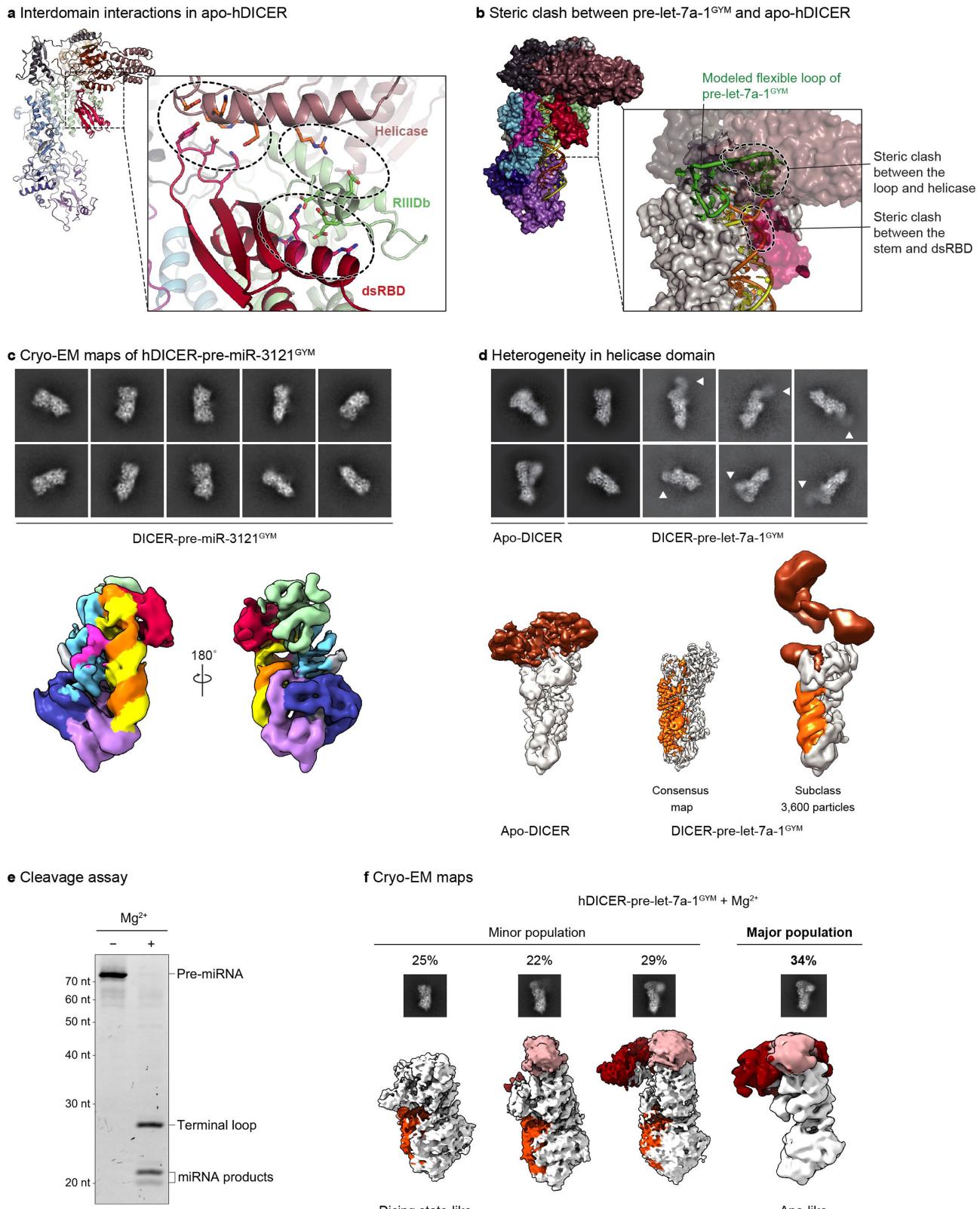


Extended Data Fig. 4 | Overall structure of hDICER in a dicing state.

a, Cryo-EM map of the catalytic site created by RIIIDa. **b**, Cryo-EM map of the catalytic site created by RIIIDb. **c**, B-factor and Q-score plots for active site residues in the hDICER-let-7a-1^{GYM} complex structure. Q-scores for each residue were derived from MapQ of Segger tool plugged in Chimera v.1.15. B-factor values were derived from real space refinement in Phenix ISOLDE v.1.1.0.

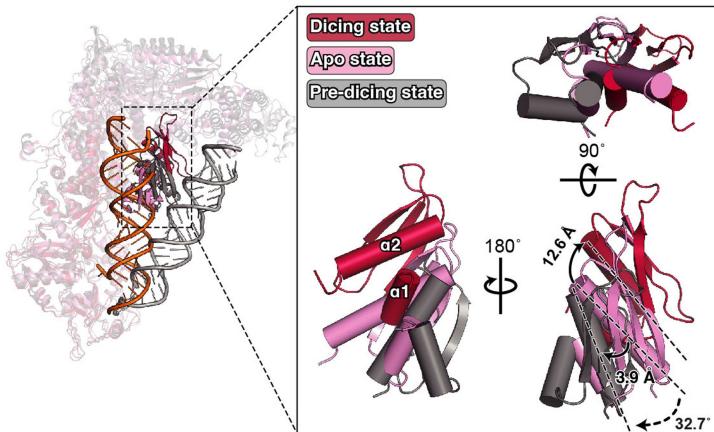
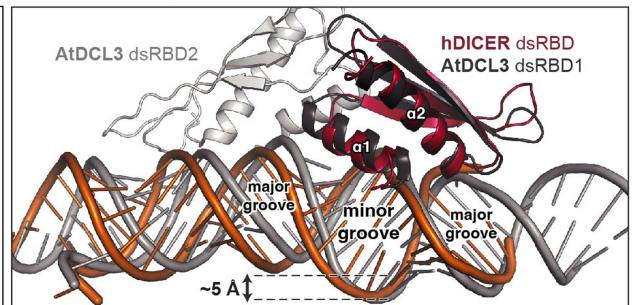
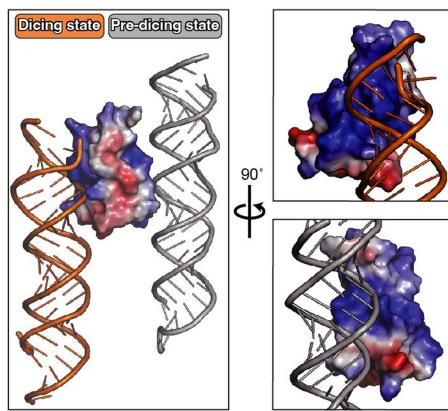
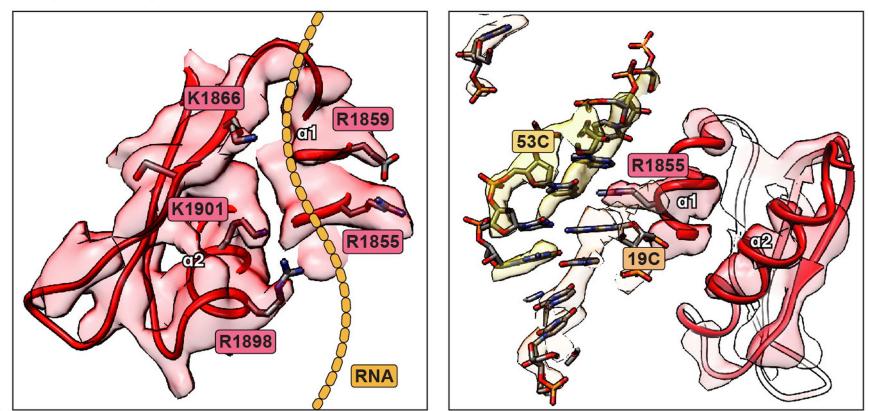
d, Superposition of RIIID domains of hDICER (this study) and Aa RNase III (PDB: 2EZ6, grey)³⁰. **e**, Active sites of hDICER and Aa RNase III (PDB: 2EZ6, grey)³⁰. **f**, Buried surface area of hDICER in a pre-dicing state (PDB: 5ZAL)²⁴ and a dicing state. **g**, RMSD of hDICER-pre-let-7a-1^{GYM} (this study) compared to hDICER-TRBP-pre-let-7a-1^{mutant} (PDB: 5ZAL)²⁴. Residues not resolved in the dicing state are coloured in grey.

Article



Extended Data Fig. 5 | The structure of the helicase domain. **a**, Interdomain interactions in apo-hDICER. **b**, Steric clash between pre-let-7a-1^{GYM} and apo-hDICER. **c**, Cryo-EM map of the hDICER-pre-miR-3121^{GYM} complex in a dicing state. **d**, Selected 2D class averages and 3D maps showing heterogeneity in the helicase domain. White arrowhead indicates the location of the helicase

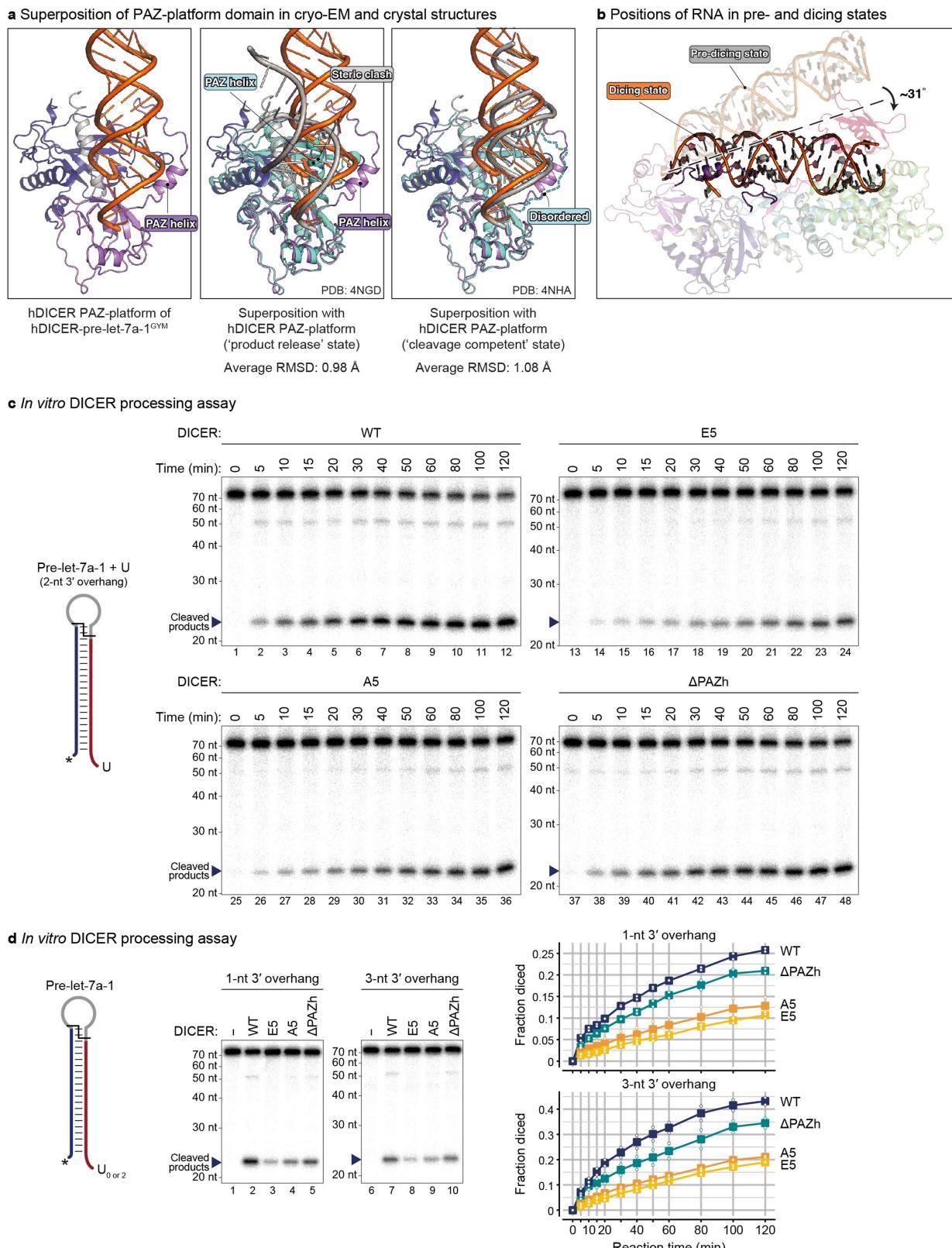
domain in 2D averages. Bound RNA density is indicated in orange. **e**, Urea-PAGE of hDICER-pre-let-7a-1^{GYM} complex incubated with or without MgCl₂ for 10 min at room temperature, visualized by SYBR gold staining. For gel source data, see Supplementary Fig. 1. **f**, Selected 2D class averages and 3D maps showing heterogeneity of the helicase domain of hDICER-pre-let-7a-1^{GYM} complex.

a Translational & rotational movement of dsRBD**b** Structural comparison of hDICER dsRBD with AtDCL3**c** Surface charge of dsRBD**d** Cryo-EM map and model of dsRBD with RNA

Extended Data Fig. 6 | The structure of dsRBD in different states. **a**, Conformational changes of the dsRBD in the apo (this study), dicing (this study) and pre-dicing states (PDB: 5ZAL)²⁴. **b**, Superposition of the dsRBDs of hDICER

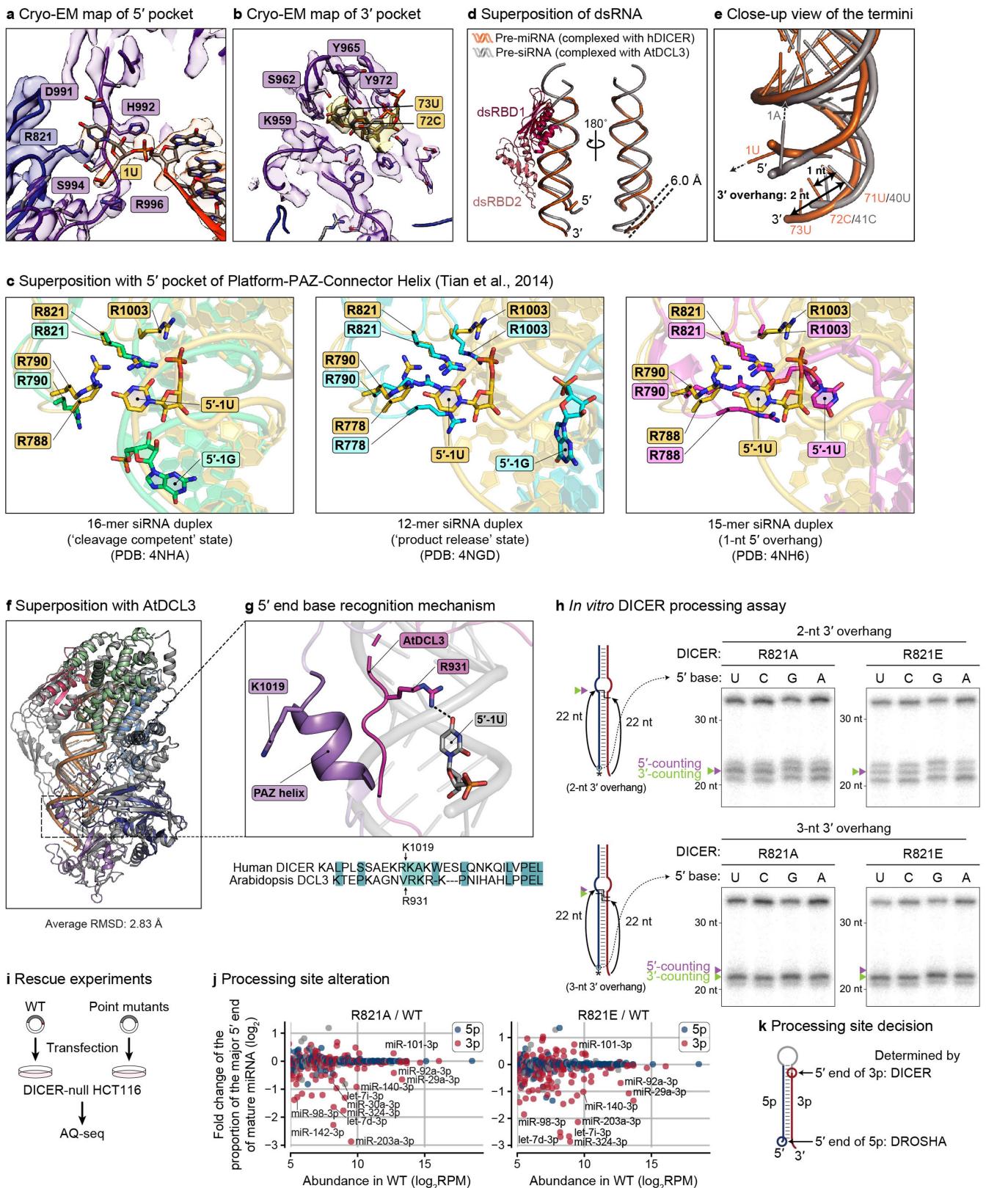
and AtDCL3 (PDB: 7VG2)⁹. **c**, Surface charge of the dsRBD, with the dsRNA–dsRBD interface in dicing and pre-dicing states (PDB: 5ZAL)²⁴. **d**, Cryo-EM map and model of the hDICER dsRBD with dsRNA.

Article



Extended Data Fig. 7 | The PAZ helix rearranges to interact with pre-miRNA in a cleavage-competent position. **a**, Superposition of hDICER PAZ-platform domain in the cryo-EM structure (this study) and in the crystal structure (PDB: 4NHA, grey)²⁰. **b**, Changes in the position of the pre-miRNA in a dicing state (this study) and a pre-dicing state (PDB: 5ZAL)²⁴. **c**, In vitro processing of

pre-let-7a-1 with a 2-nt 3' overhang. *, radiolabelled 5' phosphate. **d**, In vitro processing of pre-let-7a-1 with a 1-nt 3' overhang (lanes 1–5) or a 3-nt 3' overhang (lanes 6–10). Relative cleavage was calculated by quantifying the band intensity (1 – uncleaved/input). Squares indicate mean ($n = 2$, independent experiments). *, radiolabelled 5' phosphate. For gel source data, see Supplementary Fig. 1.

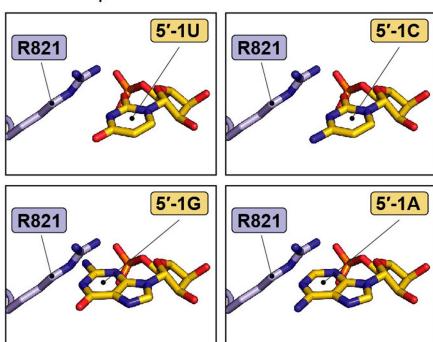


Extended Data Fig. 8 | End recognition mechanism of hDICER. **a**, Cryo-EM map of the 5' pocket. **b**, Cryo-EM map of the 3' pocket. **c**, Superposition of hDICER-pre-let-7a-1^{GMY} and Platform-PAZ-Connector Helix (PDB: 4NHA, 4NGD and 4NH6)²⁰. **d**, Superposition of dsRNAs complexed with hDICER and AtDCL3 (PDB: 7VG2)². **e**, Close-up view of the ends of the dsRNAs complexed with hDICER and AtDCL3 (PDB: 7VG2, grey)². **f**, Superposition of hDICER in a dicing state and AtDCL3-pre-siRNA complex (PDB: 7VG3, magenta)². **g**, 5' terminal base recognition by AtDCL3 via PAZ region corresponding to the PAZ helix of

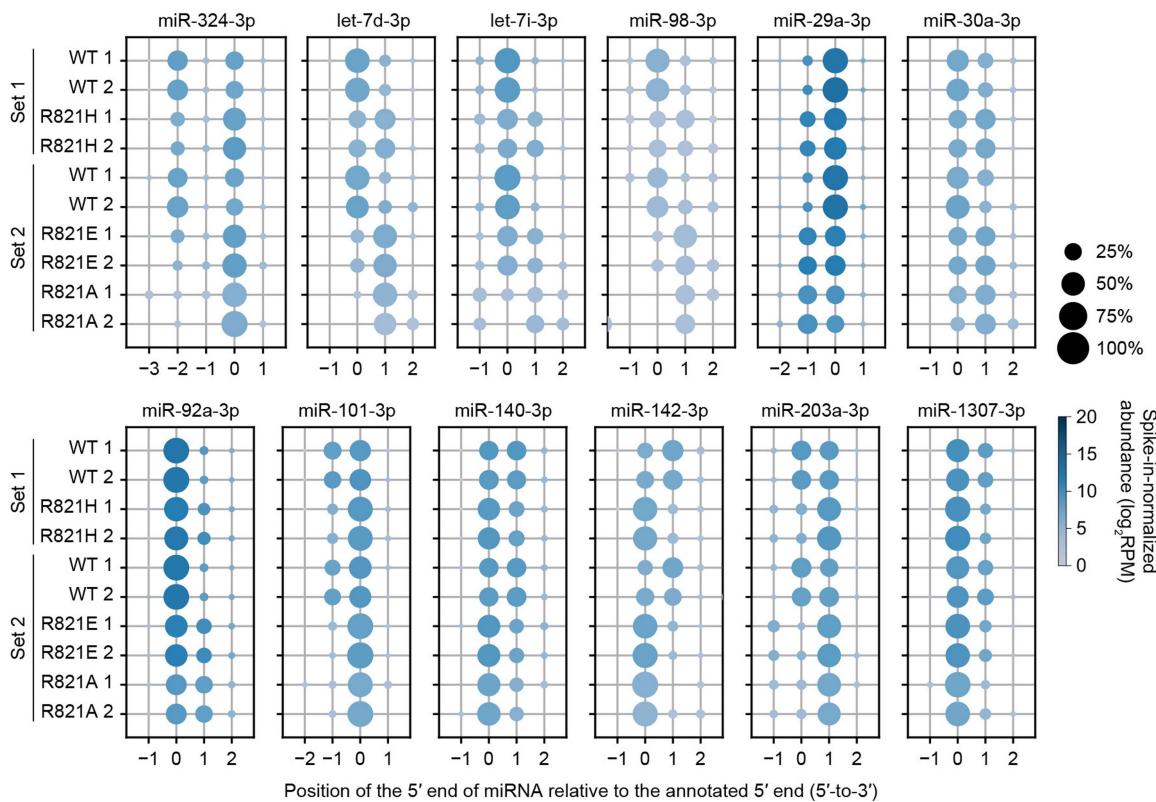
hDICER. **h**, In vitro DICER processing of pre-miRNA-like duplex with a 2-nt or 3-nt 3' overhang. The base opposite to the varying sequence is A on the 3p strand. For gel source data, see Supplementary Fig. 1. *, radiolabelled 5' phosphate. **i**, Schematic outline of the rescue experiment ($n = 2$, biological replicates). **j**, Changes in cleavage accuracy, estimated with the fold change of the proportion of the major 5'-isomiR. For a given miRNA, the most abundant 5'-isomiR was identified in the wild-type sample. Grey, unannotated strand. **k**, DROSHA/DICER cleavage sites dictated by 5' ends of mature miRNAs.

Article

a 5' base prediction models



b Proportion of 5' ends of miRNAs in the rescue experiment

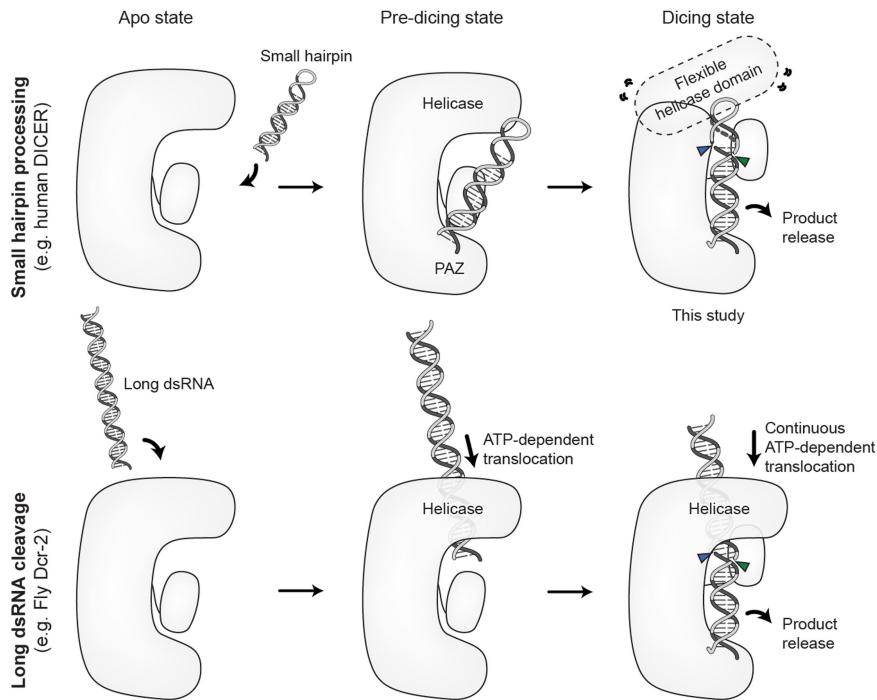


Extended Data Fig. 9 | Rescue experiments with DICER 5' pocket mutants.

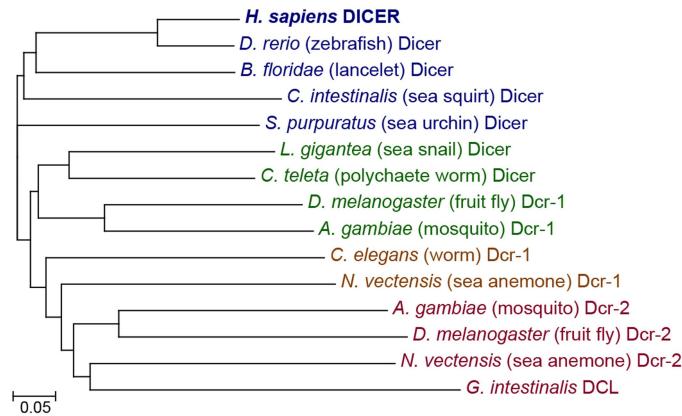
a, Predicted structural effect of the 5' end base substitutions on the interaction with the 5' pocket. **b**, Examples of altered processing sites observed in the rescue experiments. Note that the DICER cleavage sites can be inferred from

the 5' end of 3p miRNAs. miRNA isoforms beginning at the indicated position are plotted with circles, with the size of the circle reflecting the proportion of the cleavage-site usage at the given position.

a Two distinct pathways of DICER processing



b Phylogenetic tree



Extended Data Fig. 10 | Distinct functions and evolution of Dicer proteins in two small RNA pathways. **a.**, Comparison of the substrate RNA movement during DICER processing between two small RNA pathways. In the miRNA pathway, a hairpin-shaped small RNA (pre-miRNA) is bound to DICER by the helicase and PAZ domains. For cleavage, the helicase domain becomes flexible to accommodate the pre-miRNA into the catalytic centre. By contrast, in the

siRNA pathway, a long dsRNA comes into DICER by passing through the helicase domain. The ATP-dependent translocation by the helicase domain leads to processive cleavage of long dsRNAs. **b.**, A phylogenetic tree of Dicer homologues. The scale bar indicates the length for the indicated frequency of amino acid variation.

Article

Extended Data Table 1 | Cryo-EM data collection, refinement and validation statistics

	#1 apo-hDICER (EMDB-33490) (PDB 7XW3)	#2 hDICER-let7a-1 ^{GYM} (EMDB-33489) (PDB 7XW2)
Data collection and processing		
Magnification	120,000	105,000
Voltage (kV)	200	300
Electron exposure (e-/Å ²)	44.163	40
Defocus range (μm)	-1.0 ~ -2.75	-0.9 ~ -2.2
Pixel size (Å)	1.102	0.849
Symmetry imposed	C1	C1
Initial particle images (no.)	189,908	1,386,301
Final particle images (no.)	128,635	1,210,874
Map resolution (Å)	4.04	3.04
FSC threshold	0.143	0.143
Map resolution range (Å)	3.0-10.0	2.5-7.5
Refinement		
Initial model used (PDB code)	5ZAK	7XW3
Model resolution (Å)	4.04	4.03
FSC threshold	0.143	0.143
Model resolution range (Å)	3.0-10.0	2.5-7.5
Map sharpening <i>B</i> factor (Å ²)	-120.8	-157.2
Model composition		
Non-hydrogen atoms	12087	6997
Protein residues	1502	725
Nucleotides	-	54 (RNA)
Ligands	--	2 (Ca)
<i>B</i> factors (Å ²)		
Protein	98.43	74.38
Nucleotides	--	83.33 (RNA)
Ligands	--	30.00 (Ca)
R.m.s. deviations		
Bond lengths (Å)	0.009	0.007
Bond angles (°)	1.238	1.058
Validation		
MolProbity score	2.49	1.77
Clashscore	15.50	7.50
Poor rotamers (%)	2.44	0.92
Ramachandran plot		
Favored (%)	91.54	95.10
Allowed (%)	8.39	4.90
Disallowed (%)	0.07	0.00

The top part of the table provides conditions and parameters for the data collection process. The bottom part contains statistic values related to the refinement of density maps and molecular models.

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection	Cryo-EM data were collected using EPU software (version 2.10.0.5REL). Preprocessing for the sequencing data was done with cutadapt 1.8.3, FASTX Toolkit 0.0.13.2, bwa 0.7.10-r789, bedtools v2.25.0.
Data analysis	Cryo-EM data analysis was conducted using cryoSPARC (version 3.2), MotionCor2 (version 1.3.0) and GCTF (version 1.06). Initial model fitting was carried out using chimera (version 1.14) and Molecular Dynamics Flexible Fitting (MDFF) provided by Namdinator server (www.namdinator.au.dk). Further model building and refinement process for cryoEM structures were performed using COOT (Version 0.9) and ISOLDE (version 1.1.0) included in PHENIX (version 1.18.1). Visualization of structural models were performed using Chimera X (version 1.3rc202111060828) and PyMol (version 2.4.1). Custom analysis codes are available at https://github.com/haedongkim615/dicer_dicing_state_structure .

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The cryo-EM structures of the apo-hDICER and hDICER-pre-let7a-1-GYM have been deposited to Protein Data Bank (www.pdb.org) with PDB access codes of 7XW3 and 7XW2, respectively. The cryo-EM density maps of apo-hDICER and hDICER-pre-let7a-1-GYM have been deposited to EMDB under the access codes of EMD-33490 and EMD-33489, respectively. There are no restrictions on data availability. For material requests, please contact the corresponding authors. Other structural models cited in this study for analysis (5ZAL, 5ZAK, 2EZ6, 7VG2, 7VG3, 4NGD, 4NHA, and 4NH6) are also accessible on PDB. The rescue data were deposited to the GEO repository (accession number: GSE215867).

Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

Reporting on sex and gender	Not applicable
Population characteristics	Not applicable
Recruitment	Not applicable
Ethics oversight	Not applicable

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Described in the corresponding figure legends. No sample-size calculation was performed. For in vitro processing assays, we have performed independent experiments, which is sufficient to estimate the variation.
Data exclusions	No data was excluded.
Replication	Experiments were done with replicates as indicated in the figure legends. Independent methods, such as in vitro DICER processing, were in agreement.
Randomization	This project does not involve allocation into experimental groups.
Blinding	Blinding was not relevant to this study. The results from in vitro processing assays are visualized by urea-PAGE as individual bands that are quantified using software (i.e., ImageJ), requiring no subject interpretation or judgement.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	Antibodies
<input type="checkbox"/>	Eukaryotic cell lines
<input checked="" type="checkbox"/>	Palaeontology and archaeology
<input checked="" type="checkbox"/>	Animals and other organisms
<input checked="" type="checkbox"/>	Clinical data
<input checked="" type="checkbox"/>	Dual use research of concern

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	ChIP-seq
<input checked="" type="checkbox"/>	Flow cytometry
<input checked="" type="checkbox"/>	MRI-based neuroimaging

Eukaryotic cell linesPolicy information about [cell lines and Sex and Gender in Research](#)

Cell line source(s)

HCT116 cells

Authentication

Authenticated using the ATCC short tandem repeat profiling. Last performed on Mar 25th, 2022.

Mycoplasma contamination

Cell lines were confirmed mycoplasma negative prior to use.

Commonly misidentified lines
(See [ICLAC](#) register)

No commonly misidentified cell lines were used in this study.