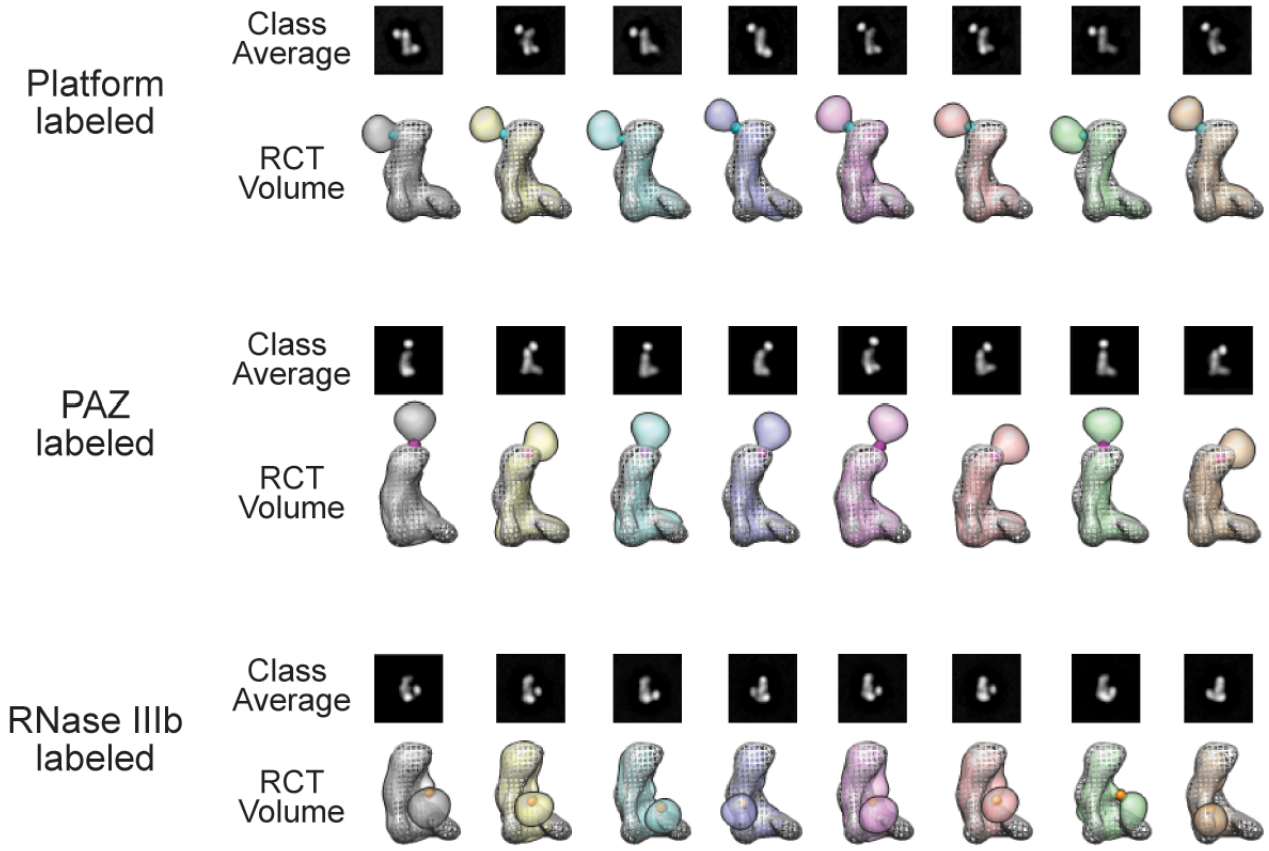
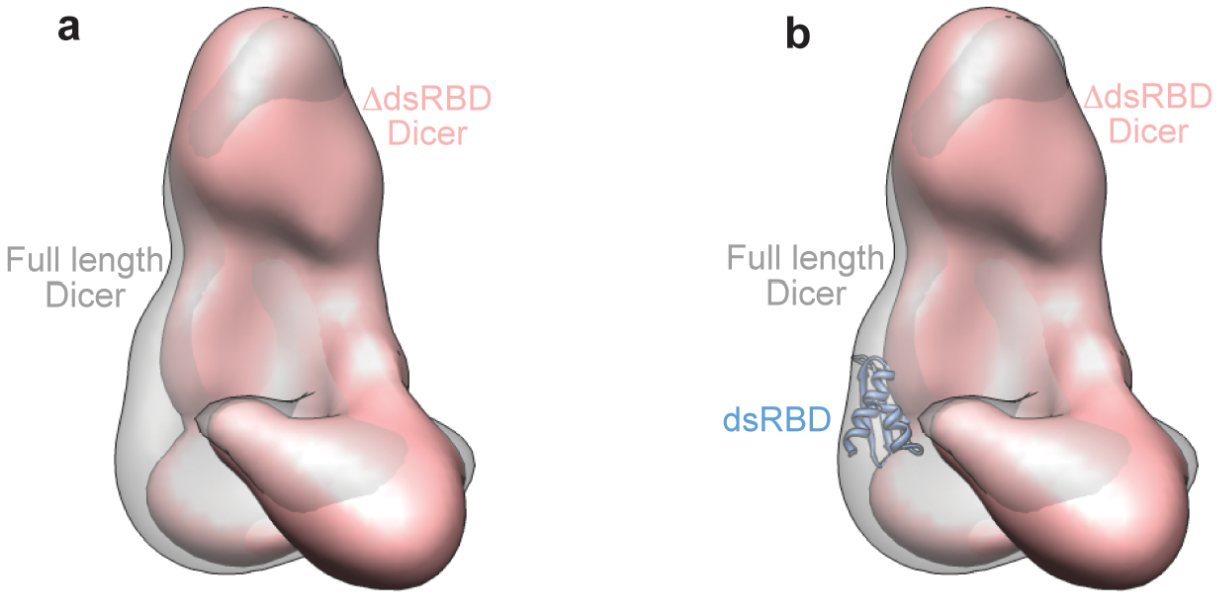
Supplementary Information

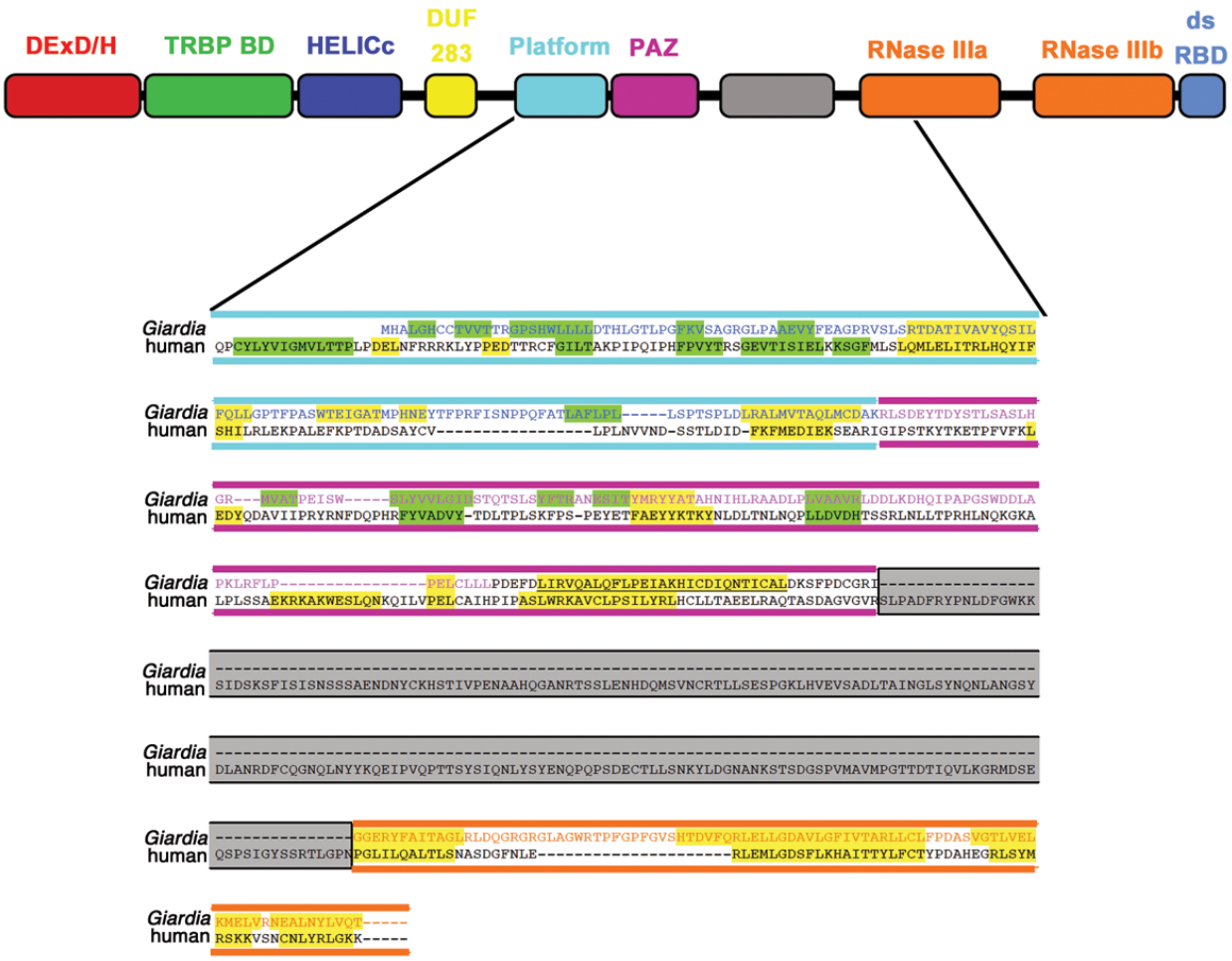
Supplementary Figures:



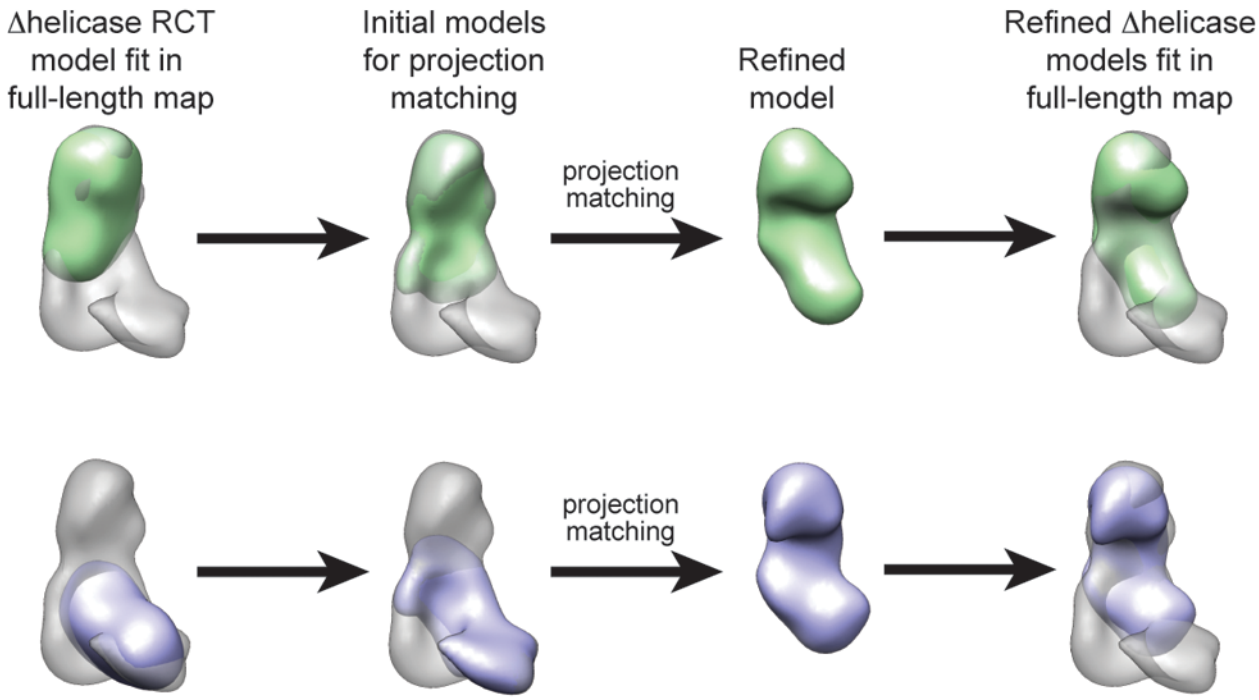
Supplementary Figure 1: RCT reconstructions of Platform, PAZ and RNase III labeled Dicer.
For each streptavidin labeled sample, the eight RCT reconstructions with the highest correlation to the refined, unlabeled Dicer map are shown. The refined, unlabeled Dicer map is shown (in wire mesh) overlaid on each RCT reconstruction. The 2D class averages that resulted in the 3D reconstructions are shown above each map. The estimated attachment sites of the streptavidin are indicated with spheres (cyan: Platform; magenta: PAZ; orange: RNase III).



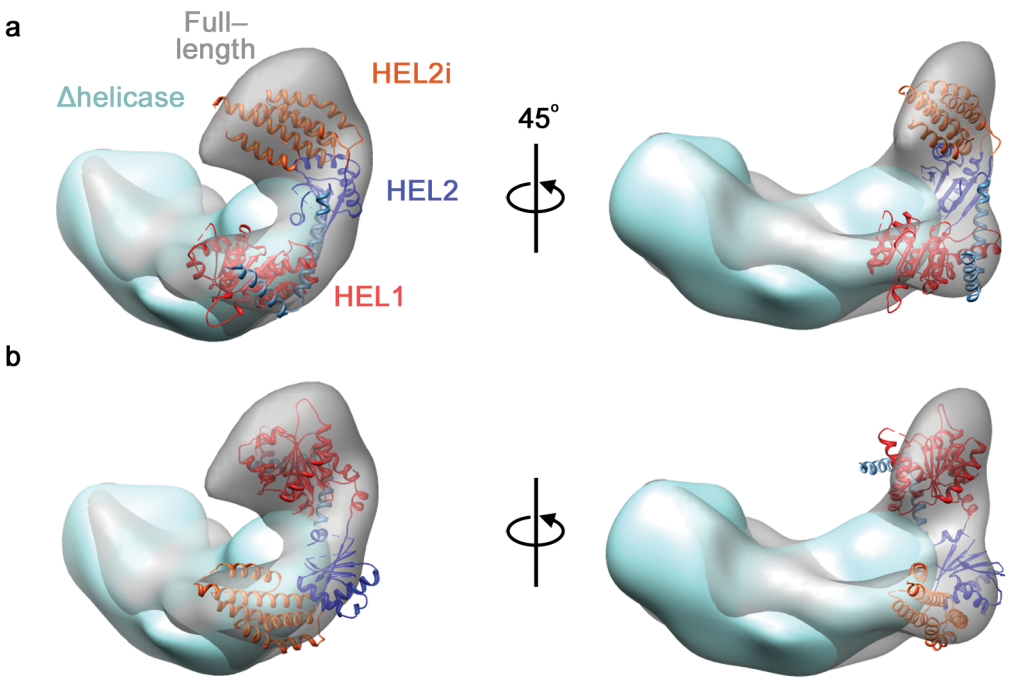
Supplementary Figure 2: Reconstruction of ΔdsRBD Dicer
Truncated human Dicer particles lacking the C-terminal dsRBD (ΔdsRBD Dicer) were subjected to projection matching refinement using the full-length protein as an initial model. **a**, The refined ΔdsRBD reconstruction (salmon) had a missing density on the right side of the body compared to the full-length model (gray). **b**, The crystal structure of the dsRBD from mouse Dicer (PDB 3C4T) docked into the difference density.



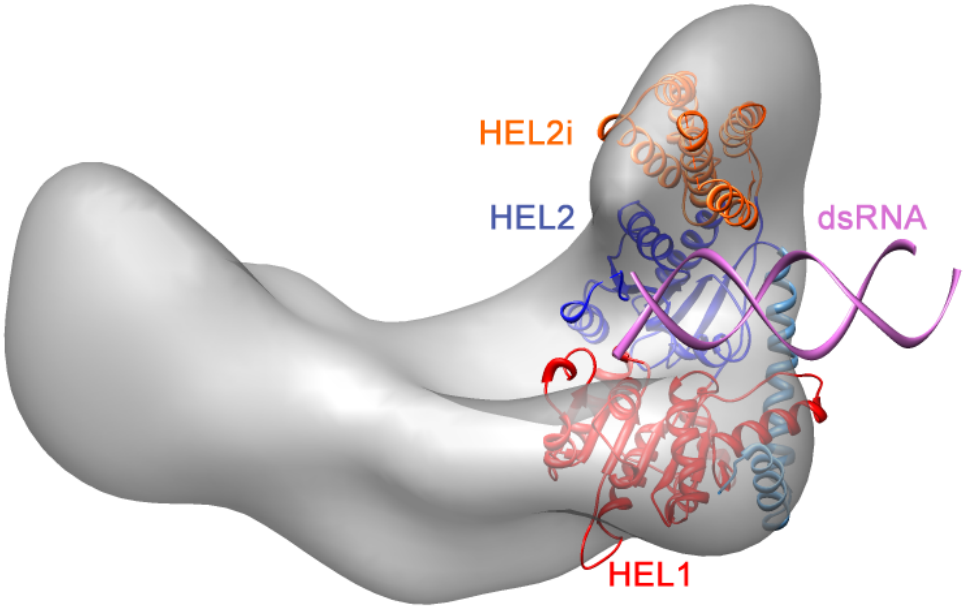
Supplementary Figure 3: Possible sequence of the "Ruler Domain".
Alignment of the amino acid sequences of human and *Giardia* Dicer between the PAZ and RNase IIIa domains. Alignments were made by sequence and structural homology. The grey region is not present in *Giardia* Dicer and, taken with its position in the primary sequence, is therefore our favorite candidate for making up the ruler domain observed in the EM map of human Dicer EM. The human RNase IIIa domain also has sequences not present in *Giardia*, and thus may also contribute to the ruler domain (not shown). It is also possible that DUF283, which is also absent in *Giardia* Dicer, may contribute to the structure of this region of the protein as well. β-strands and α-helices are highlighted green and yellow, respectively.



Supplementary Figure 4: Reconstruction and validation of the Δhelicase Dicer structure.
The RCT model of Δhelicase Dicer was docked into the two positions it best fit in full length Dicer map: head and body (top panel), and body and arm (bottom panel). Based on these dockings, two synthetic initial models were generated from the refined map. Single particle projection matching refinements were performed with each initial model. The resulting refined structures were very similar despite starting from two different initial models. Both models most closely resemble the head and body portion of Dicer.



Supplementary Figure 5: Two possible orientations for the RIG-I Helicase in the base of the L. Because the RIG-I helicase (PDB code: 4A2P) has a "C" shape with psudeo 2-fold symmetry (when viewed at moderate resolution) there were two possible ways in which it could be docked into the base of the Dicer L. **a**, Orientation A, in which the HEL1 domain is in the corner of the L. **b**, Orientation B, in which the HEL1 domain is in the arm of the L.



Supplementary Figure 6: Docking the RNA-bound RIG-I helicase crystal structure. The crystal structure of the RIG-I bound to dsRNA (PDB: 4A36) was docked into the base portion of the Dicer reconstruction. The docked dsRNA passes through the base of the "L" and leads into the nuclease core. We suggest that Dicer interacts with dsRNA substrates a similar manner.

Supplementary Materials and Methods

Generation of Avitag Dicer constructs

The Avitag sequence was cloned into specific positions encoding surface loops in a cDNA clone of human Dicer (NM_030621) in the plasmid pFastBac HT A (Invitrogen). Surface accessible loops were identified by inspection of crystal structures of the isolated human Dicer RNase IIIb domain (PDB 2EB1) and the human Dicer PAZ-platform module (generously provided by Dinshaw Patel). Restriction sites were first engineered into each loop using QuikChange PCR (Agilent Technologies). DNA oligonucleotides encoding the Avitag sequence (LNDILEAQ**K**IEWHEG) with the appropriate overhangs were then ligated into the engineered restriction sites.

For the Avitag-PAZ Dicer, the Avitag sequence was inserted between amino-acid residues K916 and E917. The primer set PAZ-QC-F/PAZ-QC-R was used to introduce *Bam*HI and *Not*I restriction sites. The oligonucleotide pair Avitag-BamHI-Not1-F/Avitag-BamHI-Not1-R was then ligated into the introduced restriction sites. For the Platform domain, *Bam*HI and *Not*I restriction sites were introduced between residues D886 and S887 using Plat-QC-F/Plat-QC-R. The Avitag-BamHI-Not1-F/Avitag-BamHI-Not1-R oligonucleotides were inserted into the restriction sites. For the Avitag-RNase IIIb Dicer, a loop composed of residues K1779–E1800 was replaced with the Avitag sequence by first replacing the encoding region with *Sfo*I and *Spe*I restriction sites using the primer pair RNaseIIIb-QC-F/RNaseIIIb-QC-R. The oligo pair Avitag-SfoI-SpeI-F/Avitag-SfoI-SpeI-R was then inserted into the engineered position.

The DNA oligonucleotides are shown below (BamHI, NotI, SfoI and SpeI sites are underlined):

PAZ-QC-F: 5′-ATCCGGTGGGGCGGCCGCTGAAACACCCTTTGTTTTTAAATTAG-3′

PAZ-QC-R: 5′-GGCCGCCACCACCGGATCCTTTTGTATACTTTGTACTGGGAA-3′

Plat-QC-F： 5′-ATCCGGTGGGGCGGCCGCTTCCAGCACTTTGGATATTGACTTT-3′

Plat-QC-R: 5′-GGCCGCCACCACCGGATCCGTCATTAACAACATTAAGAGGTAGAAC-3′

Avitag-BamHI-NotI-F/Avitag-BamHI-NotI-R oligo pair：

```
5' GATCCGGAGGTCTGAACGATATCTTGGAAGCGCAGAAGATTGAATGGCATGAAGGCGG
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    3' GCCTCCAGACTTGCTATAGAACCTTCGCGTCTTCTAACTTACCGTACTTCCGCCCCGG
```

RNaseIIIb-QC-F: 5′-GGCGCCGGAGGAACTAGTGAGGATATTGAAGTTCCAAAGGCCA-3′

RNaseIIIb-QC-R: 5′-ACTAGTTCCTCCGGCGCCCTTCTCAAGCTGAAACTGCACAAAGTCAT-3′

Avitag-SfoI-SpeI-F/Avitag-SfoI-SpeI-R oligo pair:

```
5' GGAGGTCTGAACGATATCTTGGAAGCGCAGAAGATTGAATGGCATGAAGGCA 3′
   ||||||||||||||||||||||||||||||||||||||||||||||||||||
3' CCTCCAGACTTGCTATAGAACCTTCGCGTCTTCTAACTTACCGTACTTCCGTGATC 3′
```

Generation of truncated Dicer constructs

Deletion mutants of human Dicer were amplified by PCR using the full-legth cDNA clone (NM_030621) as a template. The ΔHEL1 Dicer and ΔHelicase constructs were amplified with the primer sets HEL1-F/hDcr-R and Helicase-F/hDcr-R, respectively. PCR products were cloned as *Sfo*I-*Xho*I fragments into pFastBac HT A. PCR primers are shown below (SfoI and XhoI sites are underlined):

hDcr-R: 5′-CCCCTCGAGTCAGCTATTGGGAACCTGAGGTTGATTAGC-3′

HEL1-F: 5′-GGGGGCGCCGGGAAATGTGATCCAGA-3′

Helicase-F: 5′-GGGGGCGCCATGGATGATGATGACGT-3′

Dicer expression and purification

All the Dicer constructs were produced and purified from Sf9 cells using the Bac-to-Bac system (Invitrogen) as described[1]. The Drosophila Dicer-2 expression plasmid was a generous gift from Qinghua Liu[2]. Briefly, after the infected Sf9 cells were homogenized, insoluble material was removed by centrifugation. The supernatant, which contained the Dicer protein, was bound to His Select resin (Sigma). After 3 – 5 rounds of washing, protein was eluted from the resin using high imidazole buffer. Dicer proteins were then dialyzed against a low imidazole buffer in the presence of the tobacco etch virus (TEV) protease to remove the N-terminal His_6 tag from Dicer. After dialysis the solution was passed through a 5 ml His-Trap FF column (GE Healthcare Life Sciences) to remove TEV and Ni-binding contaminants (the flow-through solution contained untagged Dicer protein). AviTag Dicer proteins were biotinylated and labeled with streptavadin at this point in the preparation (see below). Finally, samples were concentrated and run on a HiLoad 16/60 Superdex 200 column (GE Healthcare Life Sciences) equilibrated in 0.1 M KCl, 0.5 mM TCEP, 20 mM Tris, pH 8. Additionally, in some cases, a glycerol gradient purification step was included as a final step to increase the purity of sample, as described previously[3].

Negative staining

Specimens were negatively stained using either the carbon sandwich[4] or deep staining[5] method. Approximately 4 µL of protein sample was deposited on plasma-cleaned, holey C-flat grids covered with an additional layer of thin carbon, and allowed to absorb for 30 to 60 seconds. Excess sample was blotted from the side of the grid and replaced with 2% uranyl acetate (UA) or uranyl formate (UF) solution. The stain was blotted off and replaced with fresh stain, repeated 3 to 5 times. During the last iteration of staining, the staining solution was retained on the grid for an additional 30 to 60 sec before proceeding to the following steps. For the carbon sandwich method, a thin carbon was floated over a well of stain. The grid was then used to pick up the piece of thin carbon. With the top layer of carbon, the grid was blotted from the bottom to remove excess stain and allowed to dry. For deep staining, the grid was blotted from the side using a single layer of Kimwipe such that a very thin layer of stain is left on the grid and allowed to dry.

Electron microscopy

Data were acquired using a Tecnai F20 Twin transmission electron microscope operating at 120 keV, using a dose of ~20 e-/Å^2 and a nominal defocus range of -1 to -3 µm; Images were automatically collected during multiple sessions at a nominal magnification of between 50,000x and 62,000x at a pixel size at the specimen level of 0.151 nm and 0.131 nm respectively. Images were recorded using either a Tietz F415 4 × 4 K pixel CCD camera (15 µm pixel) or Gatan 4 x 4 K pixel CCD camera utilizing the

Leginon data collection software[6]. Random Conical Tilt (RCT) experiments were carried out using the RCT node of Leginon[7], with image pairs taken at 0 and (-) 50 degrees.

Image processing and model reconstruction

All experimental data were processed using the Appion software package, which interfaces with the Leginon database infrastructure[8]. For the RCT dataset, particle pairs were picked using tiltpicker[9]. Contrast transfer function (CTF) corrections for the untilted images were estimated using Automated CTF Estimation (ACE)[10]. Only images whose CTF estimation had an ACE confidence of 0.8 or better were extracted.

For the PAZ-labeled Dicer, a total of 653 image pairs resulting in 36,809 untilted particles were subjected to alignment and classification. 37 3D RCT reconstructions were generated. For the platform-labeled Dicer, a total of 340 image pairs resulting in 12,639 untilted particles were analyzed. From these particles, 27 RCT reconstructions were generated. For the RNase IIIb-labeled Dicer dataset, 880 image pairs were collected. From these micrographs, 27,607 untilted particles were extracted, resulting in 40 RCT reconstructions.

Each reconstruction was visually inspected for evident Dicer features such as the correct dimensions and defined “head” and “arm” regions. For each labeled from of Dicer, approximately 20 visually acceptable reconstructions were retained and aligned against the full length Dicer map using the “fit in map” protocol within Chimera[11] with the density of streptavidin segmented out using Segger[12]. Cross-correlation values of the aligned RCT maps against the full length Dicer model were calculated using Spider’s

CC C routine[13]. For each labeled form, the top eight RCT reconstructions were used for estimating the attachment site of streptavidin (Figure S2).

For the ΔHelicase Dicer, a total of 150 RCT image pairs were analyzed. 12,232 particles were picked, aligned and classified using a combination of Spider[14] and Xmipp[15] protocols. 639 out of 5,703 particles were used in the RCT reconstruction. The generated RCT model at 50-Å resolution was fit into the full length Dicer map. Two positions were chosen to be a potential fit of the RCT model (Figure S4). Based on the two fits, two synthetic models of ΔHelicase Dicer were generated by manually erasing the non-overlapping density from the refined full length map using Chimera. For the projection matching refinement, the two synthetic models were used as initial models against a dataset of 63,457 particles originating from 971 micrographs. Both refinements using the EMAN package[16] converged to similar structures (Figure S4) with the resolution assessed by $FSC_{0.5}$ to be 25 Å.

For the ΔHEL1 Dicer, 64,185 particles were extracted from 1,419 micrographs at a boxsize of 120 pixels at 3.1 Å/pixel. The final 3D reconstruction was performed using a combination of SPIDER and EMAN reconstruction packages[14,16]. The starting model was obtained from full length Dicer-TRBP previously reconstructed[3]. The resolution of the final volume assessed by $FSC_{0.5}$ was 17 Å.

For Drosophila Dicer2 reconstructions, 135,377 particles were extracted from 885 micrographs at a box size of 96 pixels at 3.38. Å/pixel. The final 3D reconstruction was performed using a combination of SPIDER and EMAN reconstruction packages[14,16]. The starting model was obtained from full length Dicer-TRBP previously reconstructed[3]. The resolution of the final volume assessed by $FSC_{0.5}$ was 15 Å.

For the RCT reconstructions of Dicer bound to RNA, 34,743 untilted and 36,305 tilted particles were extracted from 637 image pairs. The two RCT reconstructions were generated from 1,643 and 1,322 particles. The resolution of each model is estimated to be 35 Å by $FSC_{0.5}$ criterion. To visualize the conformational flexibility of the helicase of Dicer, the two RCT models were aligned using the "fit in map" protocol within Chimera[11] with manual adjustment to ensure a good alignment within the body of the L.

For the subsequent modeling of RIG-I helicase crystal structures into the two conformations of Dicer, the two RCT models were used as initial models for a further multi-model refinement procedure using the EMAN package[16] with an untilted stack of 56,696 particles. The resolution of the final volumes assessed by $FSC_{0.5}$ was 19 Å.

Model fitting

Full length Dicer map was segmented using Segger[12] within the Chimera visualization software[11]. Using one smoothing step with a 1 voxel step size, the map was segmented into seven distinct regions. Fitting of the crystal structures of components of Dicer was carried out using both Chimera's fit in map routine as well as manual fitting to allow for the correspondence to streptavidin tagging data.

For the analysis using the RIG-I helicase crystal structures, the two refined Dicer maps obtained in the presence of RNA were segmented using Segger[12]. The helicase portion at the base of the L were extracted and aligned using the "fit in map" protocol within Chimera[11] with manual adjustment to ensure a good alignment at the density corresponding to the HEL1 domain. Crystal structures of the RIG-I helicase alone (PDB code: 4A2P) and bound to RNA substrate (PDB code: 4A36) were modeled into each of

the densities using also a combination of the "fit in map" protocol within Chimera[11] and manual adjustment.

REFERENCES

1 MacRae, I. J. et al., In vitro reconstitution of the human RISC-loading complex. *Proc Natl Acad Sci U S A* **105** (2), 512 (2008).

2 Liu, Q. et al., R2D2, a bridge between the initiation and effector steps of the Drosophila RNAi pathway. *Science* **301** (5641), 1921 (2003).

3 Lau, P. W., Potter, C. S., Carragher, B., and MacRae, I. J., Structure of the human Dicer-TRBP complex by electron microscopy. *Structure* **17** (10), 1326 (2009).

4 Ohi, M., Li, Y., Cheng, Y., and Walz, T., Negative Staining and Image Classification - Powerful Tools in Modern Electron Microscopy. *Biol Proced Online* **6**, 23 (2004).

5 Ruiz, T. and Radermacher, M., Three-dimensional analysis of single particles by electron microscopy: sample preparation and data acquisition. *Methods Mol Biol* **319**, 403 (2006).

6 Suloway, C. et al., Automated molecular microscopy: the new Leginon system. *J Struct Biol* **151** (1), 41 (2005).

7 Yoshioka, C. et al., Automation of random conical tilt and orthogonal tilt data collection using feature-based correlation. *J Struct Biol* **159** (3), 335 (2007).

8 Lander, G. C. et al., Appion: an integrated, database-driven pipeline to facilitate EM image processing. *J Struct Biol* **166** (1), 95 (2009).

9 Voss, N. R. et al., DoG Picker and TiltPicker: software tools to facilitate particle selection in single particle electron microscopy. *J Struct Biol* **166** (2), 205 (2009).

10 Mallick, S. P., Carragher, B., Potter, C. S., and Kriegman, D. J., ACE: automated CTF estimation. *Ultramicroscopy* **104** (1), 8 (2005).

11 Pettersen, E. F. et al., UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25** (13), 1605 (2004).

12 Pintilie, G. D. et al., Quantitative analysis of cryo-EM density map segmentation by watershed and scale-space filtering, and fitting of structures by alignment to regions. *J Struct Biol* **170** (3), 427 (2010).

13 Baxter, W. T., Leith, A., and Frank, J., SPIRE: the SPIDER reconstruction engine. *J Struct Biol* **157** (1), 56 (2007).

14 Frank, J. et al., SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J Struct Biol* **116** (1), 190 (1996).

15 Scheres, S. H. et al., Maximum-likelihood multi-reference refinement for electron microscopy images. *J Mol Biol* **348** (1), 139 (2005).

16 Ludtke, S. J., Baldwin, P. R., and Chiu, W., EMAN: semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol* **128** (1), 82 (1999).