npg

RESEARCH HIGHLIGHT

Drosha and Dicer: Slicers cut from the same cloth

DROSHA and its partner DGCR8 form a heterotrimeric complex named Microprocessor, which is essential for microRNA biogenesis. A recent study by Kwon *et al.* in *Cell* reveals the structure of a DROSHA construct in complex with the C-terminal region of DGCR8, thereby unveiling the topology and interactions between components of the Microprocessor and insights into its 'ruler'-based cleavage activity and function.

MicroRNAs (miRNAs) are a class of ~22-nt endogenous non-coding RNAs with important roles in the post-transcriptional regulation of genes in eukaryotes. The biogenesis of miRNAs is under strict control and their dysregulation leads to the onset of diverse disease states. In humans, miRNA maturation involves two consecutive cleavage steps mediated by RNaseIII enzymes DROSHA and DICER. In the nucleus, DROSHA together with its partner DiGeorge syndrome critical region 8 (DGCR8) site-specifically crops a long primary transcript (pri-miRNA) to release a hairpin RNA (pre-miRNA) of ~65-nt in length. Pre-miRNA is subsequently exported into the cytoplasm and undergoes a second cleavage by DICER, liberating a ~22-nt mature miRNA duplex. This duplex is subsequently loaded onto Argonaute, with one strand termed the guide retained, to form a RNA-induced silencing complex (RISC).

DROSHA belongs to the RNaseIII endonuclease family with the domain architecture shown in Figure 1A (top panel). The critical domains of DROSHA include a highly conserved central domain (CED), which is essential for its cleavage activity, and a C-terminal segment containing two tandem RNaseIII

domains (RIIIDa and RIIIDb) and one dsRNA-binding domain (dsRBD). The pair of RIIIDs dimerize intramolecularly to form a composite processing center with the ability to cut the 3′ and 5′ strands of the pri-miRNA stem, thereby producing characteristic staggered ends containing 2-nt 3′-overhangs. The dsRBDs of DROSHA have weak RNA-binding capacity, which is augmented by DGCR8 (domain architecture in Figure 1A, bottom panel), such that DROSHA and DGCR8 together form a complex termed Microprocessor [1, 2]. The binding stoichiometry of Microprocessor has been elucidated from biochemical assays to consist of a heterotrimer composed of one DROSHA and two DGCR8 molecules [3]. The central RNA-binding heme domain (Rhed) of DGCR8 contributes to dimerization and Microprocessor processing fidelity [4], while the dsRBDs of DGCR8 have higher binding affinity to RNA and the C-terminal tail region (CTT) is known to stabilize DROSHA.

Pri-miRNAs, which are transcribed by RNA Polymerase II (Pol II), contain a ~33-35-bp stem containing helical imperfections, a hairpin loop at the apical junction and single-strand RNA (ssRNA) overhangs at the basal junction (Figure 1B). Based on previous biochemical data, it has been proposed that Microprocessor cleaves at a distance of ~11-bp from the basal junction and at a distance of ~22-bp from the apical junction. Recent studies have identified four primary sequence determinants located in the pri-miRNA scaffold (UGUG element in apical loop, GHG (H = A, U or C) elements in the stem, and UG and CNNC elements in the basal ssRNA overhangs (Figure 1B)) that contribute

to cleavage specificity and efficiency [5], providing additional constraints on Microprocessor recognition.

During the last decade, structural information on key miRNA processing enzymes DICER [6] and Argonaute [7, 8] have added to our understanding of the later steps of miRNA biogenesis. Although DROSHA was only discovered over a decade ago [9], it was recently that Kwon *et al*. [10] have reported the 3.2 Å crystal structure of a DROSHA construct (aa 390-1 365) in complex with the C-terminal helix (aa 728-750) of DGCR8 (Figure 1C). The overall structure adopts an elongated shape with the two RIIIDs located on one side and the N-terminal segment of the CED domain located on the other side. One small 23-aa DGCR8 helix (named as G1) interacts with each DROSHA RIIID, with the interaction between G1s and RIIIDs playing an important role in stabilizing DROSHA. The superposition of RNaseIII domains of Drosha with *Aquifex aeolicus* RNaseIII (AaRNaseIII) [11] indicates that DROSHA RIIIDs utilize a canonical catalytic mechanism involving two metal ions at the catalytic sites, thereby confirming previous predictions that a pair of RIIIDs generate a composite processing center. However, there is an unusual long conserved insertion (aa 898-964) positioned within the DROSHA RIIIDa predicted to involve two helical segments. One of these, designated as the Bump helix, is buried into the CED region, while the other one is disordered in the structure due to flexibility. RIIIDa is linked to the bottom of the CED by a long helix (named Connector) that is surrounded by the N-terminal part of CED (named Platform) and the long
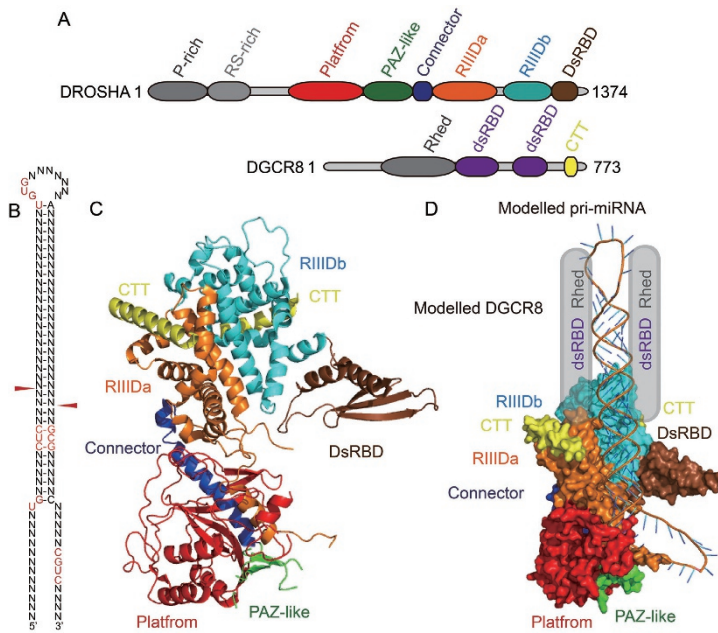
Figure 1 The structure of DROSHA-DGCR8 complex. **(A)** Schematic representation of the domain architecture of human DROSHA and DGCR8. PAZ, Piwi-Argonaute-Zwille; RIIID, RNase III domain; dsRBD, double-stranded RNA-binding domain; Rhed, RNA-binding heme domain; CTT, C-terminal tail. **(B)** A schematic representation of pri-miRNA sequence and DROSHA cleavage site. The key invariant nucleotides are highlighted in red and the cleavage sites are marked with red arrows. **(C)** The crystal structure of human DROSHA in complex with DGCR8 (PDB code: 5B16). The Platform, PAZ-like, Connector, RIIIDa, RIIIDb, dsRBD, and CTT domains are colored in red, green, blue, orange, cyan, brown, and yellow, respectively. **(D)** Modeling of pri-miRNA onto the DROSHA complex with full-length DGCR8. Note: in **B** and **C**, the DROSHA-DGCR8 complex is aligned in the same orientation. We thank Drs V Narry Kim and Jae-Sung Woo for providing the coordinates of DROSHA-DGCR8 complex and DROSHA-DGCR8-pri-miRNA complex.

insertion of RIIIDa. Platform has two unanticipated zinc finger (ZnF) motifs, with the ZnF1 motif stabilizing a protruding loop from the Platform that interacts with RIIIDa. The DROSHA dsRBD domain is positioned adjacent to the RIIIDb site and projects away from the main scaffold in the absence of bound pri-miRNA.

An unexpected and insightful observation is that the overall folding of DROSHA in the human DROSHA-DGCR8 complex resembles that observed for *Giardia* Dicer [6], although there is no sequence homology between the two proteins, except that they contain a pair of RIIIDs. Moreover, they both contain a Connector helix of similar size and a Platform of different size surrounding the Connector that adopts the same folding topology, implying that DROSHA and DICER most likely evolved from a common ancestor. There is a partially disordered region between Platform and Connector tentatively named PAZ-like region in DROSHA.

In order to illustrate how Microprocessor could potentially recognize substrate pri-miRNAs, Kwon *et al.* built a model of the pri-miRNA bound to DROSHA-DGCR8 (Figure 1D),

with the model guided by the structure of dsRNA-bound AaRNase III [11]. Given the directionality of the DGCR8 CTT domain relative to the DROSHA RIIIDs, it is conceivable that full-length DGCR8 can form a symmetrically elongated complex, such that dimerized Rhed and dsRBD domains are positioned to interact with the upper stem and apical loop of bound pri-miRNAs. The long insertion loop of RIIIDa (Bump and Mobile helices) together with Platform and PAZ-like domains could collaborate so as to recognize and hold the basal junction of pri-miRNA, thereby allowing the RIIIDa catalytic site to measure ~11-bp away from the basal junction.

This landmark contribution reveals the structure of the catalytic core of DROSHA, clarifies how DROSHA assembles with an interacting segment of DGCR8 and confirms a 1:2 binding stoichiometry of DROSHA and DGCR8. Interestingly, the similarity between DROSHA and DICER implicates that class II RNaseIII nucleases (DORSHA) may have evolved from class III RNaseIII nucleases (DICER) early in evolution. A future challenge remains the structural characterization

of the complex of DROSHA, full-length DGCR8 and bound pri-miRNA, so as to more fully decipher the rules governing pri-miRNA recognition, thereby providing a molecular understanding underlying accurate cleavage-based processing.

Sisi Li[1], Dinshaw J Patel[1,2]

[1]*Department of Biology, South University of Science and Technology of China, Shenzhen, Guangdong 518055, China;* [2]*Structural Biology Program, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA*
Correspondence: Dinshaw J Patel
E-mail: pateld@mskcc.org

References

1   Denli AM, Tops BB, Plasterk RH, *et al. Nature* 2004; **432**:231-235.
2   Gregory RI, Yan KP, Amuthan G, *et al. Nature* 2004; **432**:235-240.
3   Nguyen TA, Jo MH, Choi YG, *et al. Cell* 2015; **161**:1374-1387.
4   Barr I, Smith AT, Chen Y, *et al. Proc Natl Acad Scis USA* 2012; **109**:1919-1924.
5   Fang W, Bartel DP. *Mol Cell* 2015; **60**:131-145.
6   Macrae IJ, Zhou K, Li F, *et al. Science* 2006; **311**:195-198.
7   Wang Y, Juranek S, Li H, *et al. Nature* 2009; **461**:754-761.
8   Schirle NT, Sheu-Gruttadauria J, MacRae IJ. *Science* 2014; **346**:608-613.
9   Lee Y, Ahn C, Han J, *et al. Nature* 2003; **425**:415-419.
10 Kwon SC, Nguyen TA, Choi YG, *et al. Cell* 2016; **164**:81-90.
11 Gan J, Tropea JE, Austin BP, *et al. Cell* 2006; **124**:355-366.