# How RNA Guides Pseudouridine Formation

Pseudouridine is the most frequent type of RNA modification and has been called as the fifth nucleoside. Conversion of uridine to pseudouridine occurs only at the particular sites of substrate RNAs, so pseudouridine synthases must know where to act before catalysis. One type of enzymes comprising single polypeptide relies on specific protein-RNA interaction for substrate selection. In contrast, H/ACA RNA protein complexes (RNPs) are a family of pseudouridine synthases that use guide RNA to recruit suitable substrate RNAs [1]. H/ACA RNPs have a complex structure comprising the pseudouridine catalyst Cbf5/dyskerin and accessory proteins Nop10, Gar1 and L7ae and a distinct H/ACA guide RNA.

H/ACA guide RNAs adopts a consensus structure that minimally includes a single hairpin with a large internal loop, a conserved ACA motif in the 3' tail (Fig. 1(a)) and a K-turn motif in archaeal RNAs. Sequences in the internal loop would form two short duplexes with complementary substrate RNA and thereby specify a central unpaired uridine for modification. There are about 100 human H/ACA RNAs directing pseudouridine formation in ribosomal and splieosomal RNAs. In addition, certain H/ACA RNPs have roles other than modification, like in rRNA processing and telomerase function. Genetic analysis has shown that mutations in dyskerin, a protein component in the complex, could lead to a rare inheritable disease called dyskeratosis congenita.

To understand how Cbf5 protein cooperates with guide RNA and accessory proteins to achieve RNA-guided pseudouridylation, we have determined the crystal structure of an entire H/ACA RNP from archaeal organism *Pyrococcus furiosus* at 2.3 Å resolution by using X-ray light at beamline **BL41XU** [2].

The crystal structure reveals a triangle shaped complex with the catalytic domain of Cbf5 located in the center (Figs. 1(b)). Proteins L7Ae together with Nop10, Gar1 and the PUA domain of Cbf5 extends the catalytic core at three corners, respectively. The RNA hairpin adopts an extended conformation and binds at one side of the active site cleft that divides the catalytic domain into two roughly even parts D1 and D2. The RNA is primarily bound at the upper stem region by L7Ae, Nop10 and Cbf5 and at the lower stem region by the Cbf5 PUA domain. The guide sequences franked by the upper and lower stems are therefore positioned just on the vicinity of the active site cleft, a reasonable place for substrate binding.

The RNP structure shows that accessory proteins L7Ae, Nop10 and the peripheral PUA domain critically coordinate the position of guide RNA in the complex.

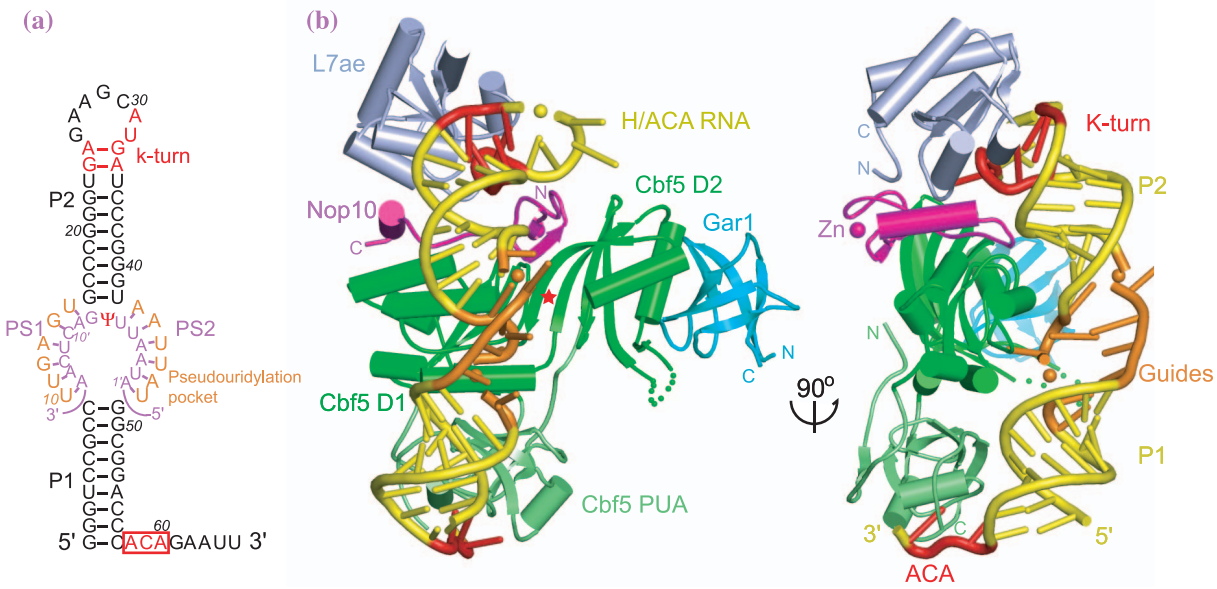


Fig. 1. Structure of H/ACA RNP. **(a)** The H/ACA RNA along with a purple cognate substrate RNA bound in the pseudouridylation pocket. **(b)** Crystal structure of the H/ACA RNP shown in two orthogonal orientations. Dots represent disordered chains. Star denotes the active site of Cbf5. [2]

The binding of the guide RNA depends on its conserved secondary structures and sequence motifs, which allows guide RNA of different sequences to be incorporated into the same protein complex.

Using a previously determined structure of pseudouridine synthase TruB in complex with a substrate hairpin RNA [3], we have built a structural model showing how guide RNA recruits substrate and orients the uridine target in the catalytic site (Fig. 2(a)).

The function of Gar1 is mysterious since it is the only protein that does not bind the guide RNA in the complex. Important clues about its function came from the observation that Gar1 interacts with a partially disordered loop of Cbf5 (Fig. 2(b)). The corresponding loop in stand-alone pseudouridine synthases, which have already had the RNA complex structure determined, interacts extensively with the loaded substrate RNA. The loop of Cbf5 likely assumes a similar role in substrate locking. However, the loop in the substrate-free complex docks against Gar1 and appears too far away from the substrate. Gar1 might control the loop in such an open conformation to aid substrate loading and release.

Guide RNA-mediated substrate recognition allows H/ACA RNP to become a programmable machine: the enzyme is able to recognize different target sites once associated with different guide RNA. Our structure has generated a framework for understanding the mechanism by which H/ACA guide RNA directs pseudouridylation, as well as other functions of H/ACA RNP.

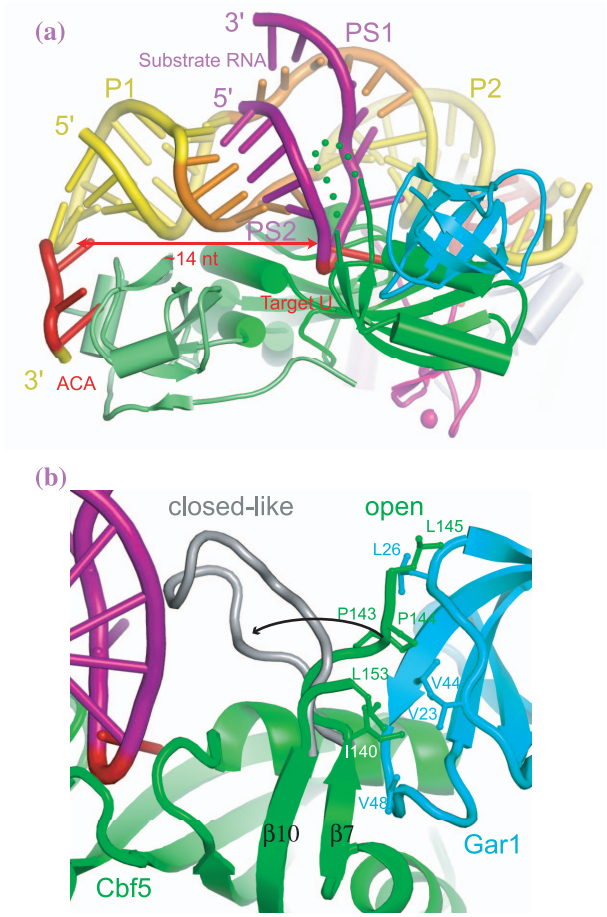


Fig. 2. Proposed mechanism of RNA-guided pseudouridylation **(a)** Structure model of the substrate complex. **(b)** Two states of the putative substrate-binding loop. The possible substrate bound state of the loop is shown in grey. The loop in our structure adopts an open conformation, in which it docks against Gar1 and appears to be unable to bind substrate. [2]

Ling Li and Keqiong Ye*

National Institute of Biological Sciences, China

*E-mail: yekeqiong@nibs.ac.cn

**References**

[1] U.T. Meier: Chromasoma **114** (2005) 1.
[2] *L. Li and K. Ye: Nature **443** (2006) 302.*
[3] C. Hoang and A.R. Ferre-D'Amare: Cell **107** (2001) 929.