nature structural & molecular biology

Structure of a functional ribonucleoprotein pseudouridine synthase bound to a substrate RNA

Bo Liang[1], Jing Zhou[2], Elliot Kahen[2], Rebecca M Terns[3], Michael P Terns[3] & Hong Li[1,2]

Box H/ACA small nucleolar and Cajal body ribonucleoprotein particles comprise the most complex pseudouridine synthases and are essential for ribosome and spliceosome maturation. The multistep and multicomponent-mediated enzyme mechanism remains only partially understood. Here we report a crystal structure at 2.35 Å of a substrate-bound functional archaeal enzyme containing three of the four proteins, Cbf5, Nop10 and L7Ae, and a box H/ACA RNA that reveals detailed information about the protein-only active site. The substrate RNA, containing 5-fluorouridine at the modification position, is fully docked and catalytically rearranged by the enzyme in a manner similar to that seen in two stand-alone pseudouridine synthases. Structural analysis provides a mechanism for plasticity in the diversity of guide RNA sequences used and identifies a substrate-anchoring loop of Cbf5 that also interacts with Gar1 in unliganded structures. Activity analyses of mutated proteins and RNAs support the structural findings and further suggest a role of the Cbf5 loop in regulation of enzyme activity.

Pseudouridine (ψ), a rotational isomer of uridine, is the most abundant modified nucleotide and is found in virtually all transfer RNA (tRNA), ribosomal RNA (rRNA) and small nuclear RNA (snRNA)[1–4]. In bacteria, five families of pseudouridine synthases catalyze site-specific isomerization in tRNA and rRNA[5–7]. In eukaryotes and archaea, a unique class of pseudouridine synthases that depend on noncoding RNAs is responsible for site-specific isomerization of rRNA and snRNA[8–11]. The box H/ACA small nucleolar or Cajal body ribonucleoprotein particles (sno/scaRNPs) comprise a minimal set of four proteins and a box H/ACA RNA. The H/ACA RNA secures the substrate RNA by base-pairing, whereas the proteins are believed to catalyze the chemical reaction. The four proteins are Cbf5 (NAF57 in mammals and dyskerin in human), Nop10, NH2P (L7Ae in archaea) and Gar1 (refs. 11–16). Cbf5 has been identified as the putative catalytic subunit based on its sequence similarities with the TruB family of bacterial pseudouridine synthases. In addition to the nucleotide-isomerization function, some members of H/ACA snoRNPs are responsible for rRNA processing[17,18], and one vertebrate scaH/ACA RNP is required for telomere maintenance[19,20]. Vertebrate telomerase is known to harbor a box H/ACA RNP subdomain that is crucial to its biogenesis and stability in small Cajal bodies[19–21]. Significantly, mutations in human Cbf5 or dyskerin[22], NOP10 (ref. 23) and NH2P[24] have all been linked to the rare genetic disorder dyskeratosis congenita[24].

The process of uridine isomerization by Cbf5-like pseudouridine synthases is only partially understood, and what is known has come largely from studies of the stand-alone pseudouridine synthases in bacteria[25–27]. Cbf5 resembles the best-characterized bacterial pseudouridine synthase TruB both in sequence and structure and, in limited cases, can act as a stand-alone pseudouridine synthase on tRNA substrates[28,29]. It is thus believed that Cbf5 shares the same catalytic mechanism as TruB and differs only in its substrate-recognition mechanism.

The sno/scaRNA-guided processes face unique challenges in substrate binding and release. The snoRNAs mediate modification of core regions of the ribosomal RNA[29,30]. Thus, in order for rRNA to be processed and modified in a timely manner, snoRNPs must gain access to specific nucleotides within the large and complex rRNA, carry out processing or modification reactions, and then release the mature rRNA to allow folding and assembly with ribosomal proteins. This process requires all box H/ACA complex proteins and the guide RNA via a complicated mechanism for which the molecular basis remains unknown[31–33]. Recent advances in structural studies of archaeal box H/ACA RNPs have provided glimpses into the architecture of the fully assembled enzyme and suggested intriguing roles for the noncatalytic subunits in substrate placement[34–40]. Regardless of their positions and interactions in the assembled RNP, individual protein subunits affect the placement process and the final conformation of the substrate RNA[36,41].

Here we describe a 2.35-Å crystal structure of a functional archaeal H/ACA RNP containing *Pyrococcus furiosus* (Pf) Cbf5, Nop10, L7Ae and a guide RNA with a substrate RNA bound at the active site. The guide RNA used in crystallization is a composite RNA based on the previously characterized Pf9 RNA[31] and another computationally identified Pf H/ACA RNA, Pf6 (**Fig. 1a**)[42,43]. The substrate RNA is a 21-nt oligomer containing 5-fluorouridine (f^5U) at the target uridine position (position 10). We also report a structure of a subcomplex containing Cbf5, Nop10 and a model guide RNA[36] with a bound but undocked substrate RNA at 3.65 Å.

[1]Institute of Molecular Biophysics, [2]Department of Chemistry and Biochemistry, Florida State University, Tallahassee, Florida, USA. [3]Department of Biochemistry and Molecular Biology, University of Georgia at Athens, Athens, Georgia, USA. Correspondence and requests for materials should be addressed to H.L. (hong.li@fsu.edu).

Received 21 May; accepted 22 May; published online 28 May 2009; doi:10.1038/nsmb.1624

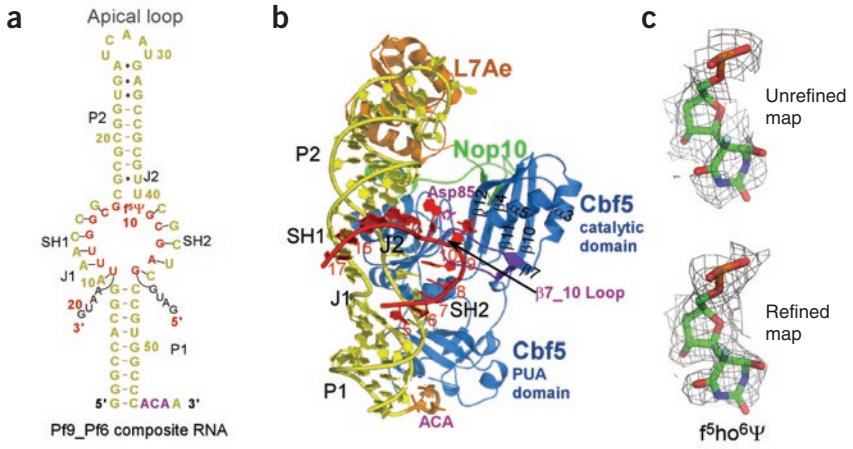

Figure 1 Overview of the structure. (**a**) The sequence of the composite guide RNA Pf9_Pf6 is shown in yellow and that of the substrate RNA is in red (modeled) and black (not modeled). RNA secondary-structure elements are named as in ref. 36. (**b**) Overview of the RNP structure. Cbf5 is in blue, Nop10 is in green and L7Ae is in orange. The RNA molecules are colored as in **a**. The substrate RNA anchoring loop of Cbf5, β7_10, is shown in purple. Nomenclature of the protein domains is as in ref. 37. (**c**) Electron density maps around the modified nucleotide. Above, cross-validated and σ_A-weighted $3F_o - 2F_c$ map computed before modeling of the RNA. Below, composite omit $3F_o - 2F_c$ map of the final refined model.

RESULTS

Overall structure

The overall guide RNA structure resembles that of the fully assembled RNP in the absence of the substrate RNA[34] in which its lower stem is anchored by the hallmark trinucleotide ACA on the PUA domain of Cbf5 and its upper stem is anchored by the kink-turn motif bound to L7Ae (**Fig. 1b** and **Table 1**). The structure of the bound guide-substrate RNA complex in the active enzyme differs from that in the L7Ae-minus enzyme (PDB 2RFK)[36] and more so from the protein-free complexes (PDB 2P89 and 2PCW)[44,45], suggesting an important role of proteins in shaping the flexible RNA. Although crystallographic contacts are observed at the end of the lower stem of the guide RNA, the bound substrate RNA is completely free of crystallographic contacts. The proteins, Cbf5, Nop10 and L7Ae, form an elongated platform on which the guide-substrate RNA complex lies longitudinally. The RNA-protein interface buries an extensive solvent-accessible surface (6,694 Å^2), of which ~30% is attributable to the substrate. The protein-RNA interface is greater in this structure than that in the L7Ae-minus RNP structure (PDB 2RFK, 5,332 Å^2)[36] and that of the substrate-minus RNP structure (PDB 2HVY, 4,488 Å^2)[34]. This finding suggests that binding of the substrate RNA enhances the interaction between the protein complex and the RNA.

Table 1 Data collection and refinement statistics

	CNL+g+s crystals	CN+g+s crystals
Data collection		
Space group	$P2_12_12$	$P6_222$
Cell dimensions		
a, *b*, *c* (Å)	186.013, 63.026, 85.447	238.341, 238.341, 127.369
Resolution (Å)	46.52–2.35 (2.41–2.3)*	50.41–3.65 (3.74–3.60)
R_{sym}	0.079 (0.550)	0.110 (0.590)
$I / \sigma I$	20.9 (2.2)	20.2 (2.7)
Completeness (%)	98.7 (92.6)	97.8 (83.5)
Redundancy	6.1 (4.1)	10.6 (9.4)
Refinement		
Resolution (Å)	46.52–2.35 (2.41–2.35)	50.0–3.65 (3.74–3.65)
No. reflections	40,393	22,492
R_{work} / R_{free}	21.7/24.8	27.5/30.6
No. atoms		
Protein	3,964	3,049
RNA	1,516	1,272
Water/ions	99/2	—
B-factors		
Protein	34.9	112.5
RNA	42.1	181.3
Water	22.4	—
r.m.s. deviations		
Bond lengths (Å)	0.010	0.013
Bond angles (°)	1.397	1.505

'CNL+g+s' denotes the Cbf5–Nop10–L7Ae complex bound with a guide and substrate RNA. 'CN+g+s' denotes Cbf5–Nop10 bound with a guide and substrate RNA.

*Values in parentheses refer to the highest-resolution shell.

Active site structure

Unbiased electron density maps calculated before modeling the substrate RNA indicate that the nucleotide targeted for modification is fully docked into the active site of Cbf5 (**Fig. 1c**). The f^{5}U is rearranged and also seems to be *cis* hydrated at position 6 to (5S,6R)-5-fluoro-6-hydroxy-pseudouridine (f^{5}ho$^6\Psi$), as was previously observed in the TruB-substrate RNA complex[46]. However, the features of the electron density do not completely exclude a hydrated *trans* isomer of f^{5}ho$^6\Psi$ or a nonhydrated f^{5}U, as was suggested by MS analysis of the reaction products of a f^{5}U-containing substrate by TruB[47]. For simplicity, we modeled the nucleotide as the rearranged and *cis*-hydrated product, f^{5}ho$^6\Psi$. We superimposed the f^{5}ho$^6\Psi$ in the structure of this complex with those found in the TruB-RNA[46] and RluA-RNA[48] complex structures and compared its surrounding amino acids (**Fig. 2**). The three active sites share strong similarities. In addition to the strictly conserved aspartate, a tyrosine, two glycines, an arginine, a lysine and a hydrophobic residue form a nearly identical pocket that accommodates f^{5}ho$^6\Psi$ (**Fig. 2**). This result establishes that Cbf5 is the catalytic subunit of the H/ACA RNP pseudouridine synthase and that Cbf5 probably shares a similar catalytic mechanism with TruB and possibly RluA. Our finding that the nucleotide to be modified is fully docked into the active site and rearranged, as previously observed for the TruB and RluAcomplexes[46,48], demonstrates that the RNP lacking Gar1 is a functional pseudouridine synthase, although its efficiency is significantly compromised in the absence of Gar1 (refs. 31,33).

The active site of Cbf5 is accessible to both ordered and bulk solvent. Cross-validated and σ_A-weighted difference density maps reveal bound solvent or ion molecules. A cluster of peaks is found near the sugar phosphate moiety of f^{5}ho$^6\Psi$ and is assigned to a hydrated potassium ion based on coordination geometries (**Fig. 3a**).

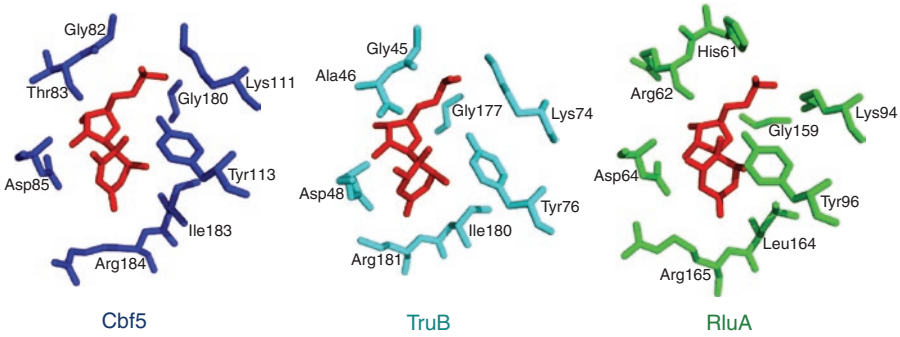

Figure 2 Comparison of active site structure. The modified nucleotide $f^5ho^6\Psi$ was superimposed and oriented in the same way for all three structures. $f^5ho^6\Psi$ is shown in red and the surrounding residues are in blue (Cbf5), cyan (TruB) and green (RluA). The structures of TruB–RNA and RluA–RNA complexes are from PDB 1K8W (ref. 46) and 2I82 (ref. 48), respectively.

The potassium ion is within coordination distance (~2.4–2.9 Å) of two water molecules, the carbonyl oxygen of Thr181 and the O5′ atom of $f^5ho^6\Psi$. It is also within 3.3 Å to O4′ and O6 of $f^5ho^6\Psi$ and the hydroxyl group of Tyr113. In both the TruB and RluA complex structures, this site also binds a solvent molecule[46,48]. This solvent molecule(s) can potentially hydrate the rearranged f^5U and may therefore have a role in the isomerization reaction. A second solvent site is 2.7 Å away from N1, which mediates recognition of the nucleobase of $f^5ho^6\Psi$ by the N terminus of the α5 helix (**Fig. 3a**). Unexpectedly, the nucleobase is also accessible to bulk solvent, with a notable solvent-accessible area (7.7 $Å^2$). In contrast, the nucleobase bound to TruB or RluA has a negligibly small solvent-accessible area (<0.8 $Å^2$)[46,48]. The difference in solvent accessibility was found not to be due to the insertion regions in TruB (resides 83–101) and RluA (resides 175–195) that block the back entrance to the active site[46,48]. Rather, the difference is attributable to the less compact active site of Cbf5. Compression of the active site, through Gar1 binding for example, could prevent escape of the uracilate intermediate and, therefore, enhance the activity of Cbf5.

Substrate RNA binding to Cbf5

The substrate RNA interacts exclusively with Cbf5 residues (**Fig. 3**) at the tip of the V-shape that ends with $f^5ho^6\Psi$ (G7–C12). The rearranged nucleotide $f^5ho^6\Psi$ establishes the most extensive interactions with a polar pocket of Cbf5 (**Fig. 3**). The carboxylate group of the putative catalytic residue Asp85 establishes two hydrogen bonds with $f^5ho^6\Psi$: Oδ1 with N3 (2.9 Å) and Oδ2 with 2′-OH (2.4 Å). Two amide nitrogen atoms (Ile183 and Arg184) contact N1 and O6, respectively (**Fig. 3a**). The amide nitrogen of Gly180 and the guanidinium group of Arg205 further enhance the interaction by contacting two nonbridging oxygen atoms (**Fig. 3a**).

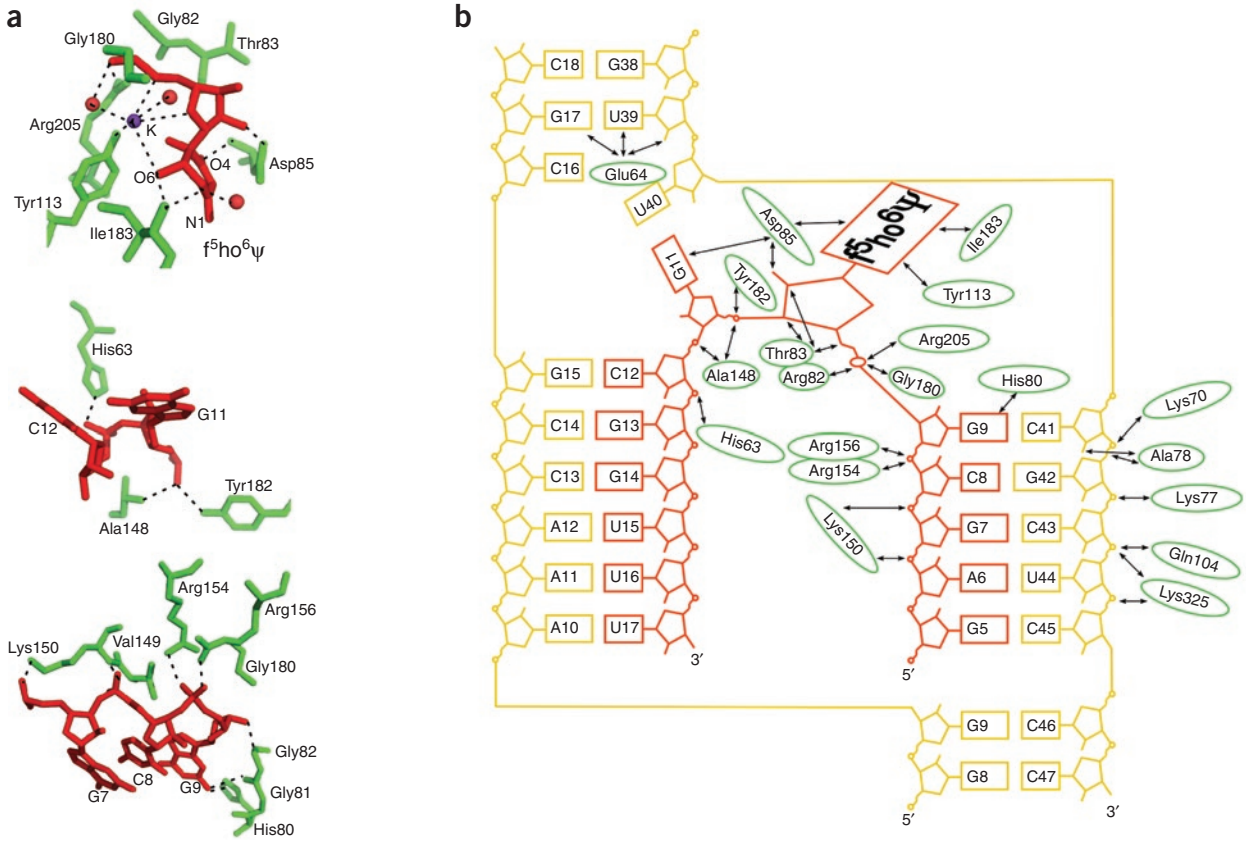

Figure 3 Interactions of the RNA with Cbf5. (**a**) Interactions between the substrate RNA (red) and Cbf5 (green). Dashed lines indicate polar atom contacts within 3.4 Å. Red spheres indicate solvent molecules, and the purple sphere indicates bound potassium. (**b**) Schematic interactions between the guide (yellow) and substrate RNA (red) with Cbf5 (green).

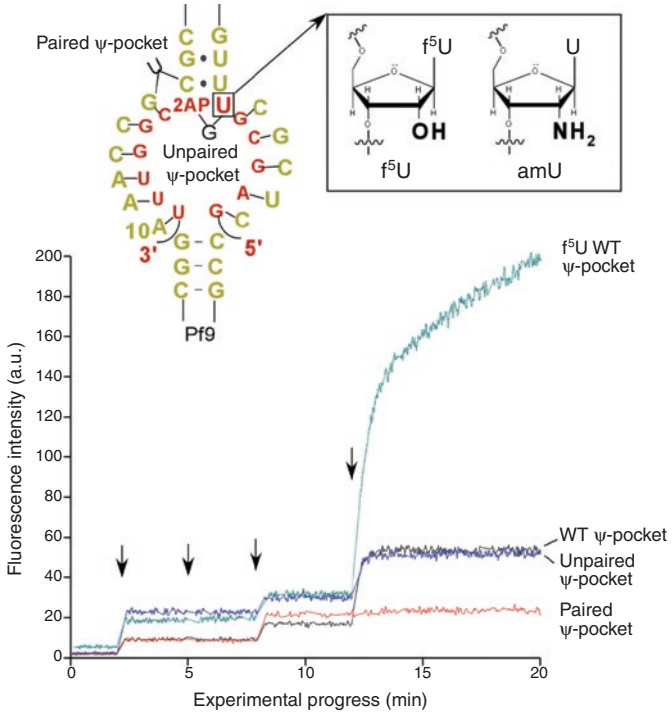

Figure 4 Structural features of the wild-type ψ-pocket important for substrate docking examined by fluorescence analysis. Color traces are fluorescence intensities in arbitrary unit (a.u.) as a function of experiment progress (in minutes). Arrows indicate the time points at which each specific component was added, where 's' denotes 2AP-labeled substrate RNA, 'g' denotes the Pf9_Pf6 composite RNA, 'CGN' denotes the Cbf5–Gar1–Nop10 complex and 'L' denotes L7Ae. Mutations in either guide or substrate RNA leading to a fully paired or an additionally unpaired ψ-pocket are shown in the RNA secondary-structure diagram, and their docking fluorescence traces are labeled accordingly. All proteins are in molar excess of RNA to ensure full binding. The nomenclature for guide and substrate RNA mutations and modifications is shown above.

Two Cbf5 regions interact with nucleotides flanking $f^5ho^6\Psi$: an amino acid cluster comprising His63, His80, Gly81, Gly82 and Thr83 (His-Gly-Thr cluster) and the loop connecting the β7 and β10 strands (β7_10 loop) (**Fig. 3** and **Supplementary Fig. 1**). The His-Gly-Thr cluster interacts with the major groove, whereas the β7_10 loop interacts with the minor groove side of the substrate RNA. These interactions are largely nonspecific, consistent with the flexibility observed in the identity of substrate RNA nucleotides in this region, and they involve hydrophobic and electrostatic contacts to the backbone of the RNA. The sugar phosphate backbone of substrate RNA nucleotides G7–G9 directs the approach of the target uridine to the active site and is stabilized by interactions with protein backbone atoms and positively charged arginine residues of the β7_10 loop (Lys150, Arg154 and Arg156). On the strand leaving the active site, the G11 nucleobase stacks with the aliphatic chain of Arg146. In Cbf5 proteins from other organisms, this position is often substituted by a hydrophobic residue (**Supplementary Fig. 1**). Thr83 interacts with both G9 and C12 on either side of $f^5ho^6\Psi$ by forming a hydrogen bond with its amide nitrogen to the 2′-hydroxyl oxygen of G9 and a water-mediated interaction with its hydroxyl oxygen to the nonbridging oxygen atoms of C12. This interaction is important for maintaining the V-shape of the substrate RNA that places the substrate uridine in the active site.

The impact of observed structural features on substrate binding

We further examined the importance of certain structural features of the composite Pf9-Pf6 RNA in substrate binding using a previously developed fluorescence assay[41]. In this assay, the nucleotide immediately downstream of the target uridine is substituted with fluorescent 2-aminopurine (2AP), which produces a high fluorescence intensity as a result of substrate docking. Furthermore, titrating the substrate RNA with a protein (or a protein complex) yields an apparent K_d that measures the free energy for the last-assembled protein (or protein complex) to place the substrate at its fully docked position. We used the K_d of docking a 2′-amino substituted substrate (amU, **Fig. 4**) as an upper limit for the wild-type K_d, which is experimentally difficult to obtain owing to slow isomerization of the substrate or release of the product.

Fluorescence evidence indicates the importance of the structure of the pseudouridine pocket (Ψ-pocket) in substrate binding. The nucleotide immediately downstream of the target uridine is unpaired in the overwhelming majority of known substrate RNAs. This unpaired nucleotide (G11) is extruded and is free of base-specific interactions (**Fig. 3**). We inserted a uridine between G15 and C16 in the Pf9-Pf6 guide RNA that can potentially form a base pair with the 2AP-substituted G11 in the fluorescence substrate RNA (paired Ψ-pocket). The substrate no longer docks into the active site (**Fig. 4**). We further asked whether the substrate RNA can accommodate additional unpaired nucleotides downstream of the target uridine, and we tested this by inserting a guanine between 2AP and the target uridine (unpaired Ψ-pocket). We observed that the substrate RNA with the insertion was docked to the active site of the RNP enzyme, although with an increased energetic cost (**Fig. 4** and **Supplementary Table 1**). These results are consistent with the conservation of one unpaired 3′ adjacent nucleotide while accommodating some occurrences of two unpaired 3′ adjacent nucleotides[48].

We further used mutational analysis to assess the impact of the β7_10 loop. We deleted three highly conserved residues within this loop: Ala148, Val149 and Lys150 ($C_{\Delta loop}$) (**Supplementary Fig. 1**). This deletion led to a substantial increase in the K_d of the complex (**Supplementary Table 1**), suggesting a defect in substrate docking. Notably, omission of Gar1 from the complex ($C_{\Delta loop}N$) restored binding efficiency (**Supplementary Table 1**). We interpret this result to indicate that Gar1 functions through the β7_10 loop of Cbf5 and that both an intact β7_10 loop and Gar1 are required for correct placement of the substrate. The structural basis for the cooperativity of Gar1 and the β7_10 loop is discussed in a later section.

Placement of substrate RNA by L7Ae

The effect of L7Ae on substrate docking has been well discussed[35,36,41,49]. Our structure provides unambiguous support for the proposed role of L7Ae in substrate placement. The substrate RNA binds in the absence of L7Ae but is far from the active site. Delivery to the active site upon L7Ae binding is almost entirely based on a rotation of the guide-substrate helix, SH1, as a result of anchoring the upper stem, P2, by L7Ae (**Fig. 5a**,**b**). More unexpectedly, comparison of the new substrate-bound structure with a previously determined structure lacking the substrate RNA (this work and ref. 38) suggests plasticity in the complex that could be important in the interaction of the proteins with the diverse family of H/ACA RNAs and/or with their substrates (**Fig. 5a**). Whereas the Cbf5 structures in complexes containing or lacking the substrate RNA are closely superimposable, in the presence of the substrate, L7Ae, Nop10 and the upper stem of the guide RNA are positioned ~6 Å closer to the back face of the catalytic domain of Cbf5, leading to further anchoring of the guide RNA (**Fig. 5a**). The movement of L7Ae and Nop10 could be a result of differences in the secondary structures of the distinct guide RNAs used as well as of substrate binding. The large shift suggests a

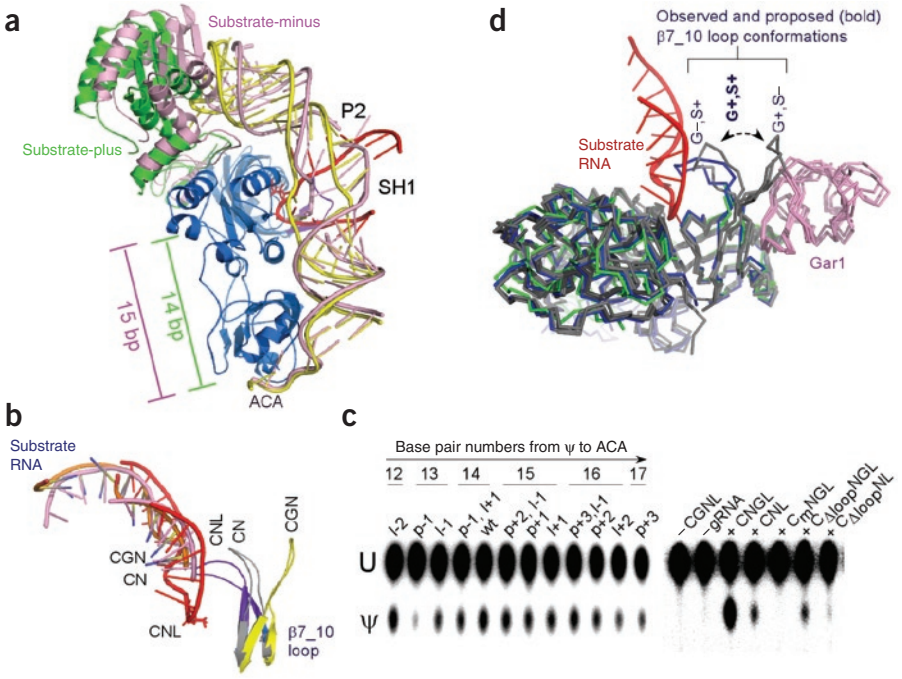

Figure 5 Structure and activity analysis of variously assembled wild-type and mutant RNPs reveal mechanisms for plasticity and the sensitivity of β7_β10 loop to substrate and Gar1 binding. (**a**) The RNP structures of the substrate-plus (this work) and substrate-minus (PDB 2HVY)[38] complexes are superimposed. A large downward shift of L7Ae, Nop10 and the guide RNA is observed upon binding of substrate RNA. L7Ae and Nop10 of the substrate-plus complex are labeled in green and those of the substrate-minus complex are in pink. Note that the substrate-minus complex would have 15 base pairs between the target uridine and the ACA trinucleotide if substrate is bound, whereas the current substrate-plus structure has 14 base pairs. (**b**) Cbf5-superimposed structures containing a bound substrate RNA. For clarity, only the substrate RNA and the β7_10 loop are shown. Each structure is labeled, where CNL denotes the structure of the Cbf5–Nop10–L7Ae complex bound with a guide and substrate RNA (this work), CN denotes that of the Cbf5–Nop10 complex bound with a guide and substrate RNA (this work) and CGN denotes that of Cbf5–Gar1–Nop10 complex bound with a guide and substrate RNA (PDB 2RFK)[36]. (**c**) Thin-layer chromatography radiograms of reacted and nuclease-digested substrate RNA with various RNP complexes. 'gRNA' denotes the Pf9_Pf6 composite RNA, 'C', 'G', 'N', 'L' denote Cbf5, Gar1, Nop10 and L7Ae respectively. 'C_m' denotes Cbf5 containing D85A, and '$C_{\Delta loop}$' denotes the β7_10 loop lacking Ala148, Val149 and Lys150. Letters 'I' and 'p' refer to 'lower stem' and 'psi-pocket' respectively. Thus, 'I-2' refers to the Pf9_Pf6 mutant containing a 2-base-pair deletion in its lower stem, and so on. Structures of mutant Pf9_Pf6 RNA are shown in **Supplementary Figure 2**. (**d**) Superimposition of Cbf5 proteins present in various box H/ACA complex structures (this work and PDB 2APO, 2HVY, 2EY4 and 2RFK)[36–39] reveals two distinct conformations of the β7_10 loop that depend on Gar1 binding and the presence of the substrate RNA. Cbf5 structures observed in the presence of Gar1 (pink) and absence of substrate are shown in gray. The loop is in the Gar1-plus, substrate-minus conformation (G+,S–) with one exception (PDB 2EY4) that is influenced by crystal packing. In the absence of Gar1 and presence of substrate RNA, Cbf5 is shown in blue (this work). The loop is in Gar1-minus and substrate-plus conformation (G–,S+). The structure in the absence of both Gar1 and substrate is shown in green (PDB 2APO), and the loop (or β7 and β10) adopts the Gar1-minus, substrate-minus conformation (G–,S–). The proposed third conformation of the loop in the presence of both Gar1 and substrate RNA (G+,S+) is indicated by a double arrow.

molecular basis for the ability of box H/ACA proteins to accommodate structurally diverse guide RNAs and/or substrate binding. Notably, the residues that form the Cbf5, L7Ae and Nop10 interfaces remain nearly unchanged in the two structures (data not shown), suggesting that a minor rearrangement of the protein interface can accommodate relatively large differences in RNA interactions.

To assess the range of structural plasticity in the guide RNA, we tested the pseudouridylation activities of the Pf9_Pf6 RNAs that form 12–17 base pairs between the target uridine and the ACA trinucleotide (**Supplementary Fig. 2**). Notably, all mutant Pf9_Pf6 RNAs were able to guide pseudouridylation of the substrate RNA with exception of the guide RNA that forms a ψ-pocket with one less base pair (p–1) (**Fig. 5c**), suggesting an impressive range of flexibility in the guide RNA. The structural mechanism identified here offers one possible solution to binding structurally varied guide RNAs.

Gar1 affects substrate RNA conformation via a Cbf5 loop

Gar1 binds to a peripheral portion of the catalytic subunit and is unable to contact either the guide or the substrate RNA[34,36]. Yet both *in vitro* pseudouridylation activity assays[31,50] and a substrate-docking fluorescence assay[41] provide clear evidence for its impact on the rate of pseudouridylation and substrate placement. To understand the functional role of Gar1 and its cooperativity with the β7_10 loop of Cbf5, we determined a crystal structure of the substrate-bound complex containing the guide RNA, Cbf5 and Nop10 (**Table 1**) to compare with a previously determined structure of this complex in the presence of Gar1. In the absence of both L7Ae and Gar1, the substrate is undocked. However, the substrate is shifted closer to the active site than that in the presence of Gar1 (**Fig. 5b**)[36], supporting the idea that Gar1 may restrict access to the active site in the absence of L7Ae. This structural observation is consistent with a decrease in the K_d of substrate docking without Gar1 (**Supplementary Table 1**).

Superimposition of the currently available Cbf5 structures[34–40] shows that the β7_10 hairpin is found in two distinct conformations that depend on the presence of Gar1 and substrate RNA (**Fig. 5b**,**d**). In the absence of Gar1 and the presence of a fully docked substrate RNA, the β7_10 loop is oriented to interact with the substrate RNA (Gar1-minus, substrate-plus, or 'G–,S+', conformation). With one exception (in which the β7_10 loop is involved in a crystal packing interaction[37]), in the presence of Gar1 and the absence of substrate RNA, the β7_10 loop is moved toward Gar1 and away from the substrate-docking position (Gar1-plus, substrate-minus, or 'G+,S–', conformation) (**Fig. 5d**). Thus, the β7_10 loop conformation is sensitive to binding of both substrate RNA and Gar1. However, removal of three β7_10 loop residues ($C_{\Delta loop}$NGL) did not completely abolish pseudouridylation activity under the condition of excess enzyme over substrate (**Fig. 5c**), suggesting that the β7_10 loop has a role in the regulation of enzyme activity. We predict that in the presence of both substrate RNA and Gar1, the Cbf5 β7_10 loop experiences an energetic 'tug-of-war', which may result in positioning of the loop in a third, intermediate

conformation. The third (Gar1-plus, substrate-plus, or 'G+,S+') loop conformation may optimally position the substrate in the active site, or close the active site, or mediate substrate binding and release (**Fig. 5d**). This interpretation is completely consistent with the conformational behavior of the substrate RNA provoked by Gar1 in either the wild-type or a β7_10 loop mutant RNP[41] and with the demonstrated role of the β7_10 loop in the thermodynamics of substrate RNA binding (**Supplementary Table 1**).

DISCUSSION

Although box H/ACA snoRNPs have long been known to be responsible for site-specific pseudouridylation of rRNA and snRNA, and box H/ACA RNPs that efficiently modify RNA have been reconstituted, the enzymatic mechanism of this unique class of pseudouridine synthases has remained elusive. Box H/ACA RNP enzymes have a complex mechanism for recruiting substrate RNA and require a number of proteins in addition to the catalytic subunit for activity. The structure of an archaeal box H/ACA RNP interacting with its substrate RNA provides important insight into the processes of substrate binding and modification.

The arrangement of active site residues around $f^5ho^6\Psi$ is similar to that found in the previously studied bacterial pseudouridine synthases. This result directly links the catalytic mechanism of the RNP pseudouridine synthase to that of stand-alone pseudouridine synthases. More notably, the process of substrate binding to the RNP pseudouridine synthase is extraordinarily complex and requires all protein and RNA components. Systematic disruption of structural elements that are directly involved in the substrate-enzyme interaction, including the conserved pseudouridine pocket structure and the β7_10 loop of Cbf5, either causes misplacement of the substrate RNA or reduces binding efficiency. These findings suggest a structural plasticity of the box H/ACA RNP assembly that is optimized for the RNA-modification process.

The architecture of the archaeal and eukaryotic H/ACA RNPs is probably similar. Generally, the eukaryotic H/ACA RNP proteins contain additional sequences; however, the elements responsible for particle assembly are conserved between the two domains. The finding that mammalian Gar1 is assembled with the other H/ACA RNP components separately and at the last stage of initial biogenesis of the complex is consistent with its subtle conformational role in substrate docking. In mammalian cells, Gar1 is the last of the four core proteins to assemble on the RNP, competing off the early-binding assembly factor NAF1 in the nucleolus[51]. It is possible that, through this conformational role, Gar1 controls the onset of the pseudouridylation activity or the release of modified RNA substrate.

Key questions remain with regard to the catalytic mechanism of Cbf5 and all pseudouridine synthases. In the absence of cofactors or external sources of energy, pseudouridine synthases break one chemical bond, rotate the uracilate ring and form a different chemical bond. With new knowledge of the active site arrangement, understanding this remarkable enzymatic acrobat is now within reach.

Accession codes. Protein Data Bank: coordinates for the Cbf5–Nop1–L7Ae and Cbf5–Nop10 bound to guide and substrate RNA have been deposited with accession codes 3HJW and 3HJY, respectively.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/nsmb/.

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

ACKNOWLEDGMENT

This work was supported in part by the US National Institutes of Health grants R01 GM66958-01 (H.L.) and R01 GM54682 (M.T. and R.T.). B.L. (0815118E) and J.Z. (0815267E) are supported by the American Heart Association, Greater Southeast Affiliate. X-ray diffraction data were collected by the Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beamline at the Advanced Photon Source (APS), Argonne National Laboratory. Supporting institutions for APS beamlines may be found at http://necat.chem.cornell.edu and http://www.ser-cat.org/members.html. Use of the APS was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-Eng-38.

AUTHOR CONTRIBUTIONS

B.L. determined the crystal structure; J.Z. prepared samples; E.K. performed fluorescence assays; M.P.T. and R.M.T. prepared the manuscript; H.L. supervised the project and prepared the manuscript.

Published online at http://www.nature.com/nsmb
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Maden, B.E. The numerous modified nucleotides in eukaryotic ribosomal RNA. *Prog. Nucleic Acid Res. Mol. Biol.* **39**, 241–303 (1990).
2. Limbach, P.A., Crain, P.F. & McCloskey, J.A. Summary: the modified nucleosides of RNA. *Nucleic Acids Res.* **22**, 2183–2196 (1994).
3. Grosjean, H. & Benne, R. Modification and editing of RNA (ASM, Washington, DC, 1998).
4. Charette, M. & Gray, M.W. Pseudouridine in RNA: what, where, how, and why. *IUBMB Life* **49**, 341–351 (2000).
5. Ofengand, J. Ribosomal RNA pseudouridines and pseudouridine synthases. *FEBS Lett.* **514**, 17–25 (2002).
6. Koonin, E.V. Pseudouridine synthases: four families of enzymes containing a putative uridine-binding motif also conserved in dUTPases and dCTP deaminases. *Nucleic Acids Res.* **24**, 2411–2415 (1996).
7. Kaya, Y. & Ofengand, J. A novel unanticipated type of pseudouridine synthase with homologs in bacteria, archaea, and eukarya. *RNA* **9**, 711–721 (2003).
8. Decatur, W.A. & Fournier, M.J. RNA-guided nucleotide modification of ribosomal and other RNAs. *J. Biol. Chem.* **278**, 695–698 (2003).
9. Ganot, P., Bortolin, M.L. & Kiss, T. Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs. *Cell* **89**, 799–809 (1997).
10. Meier, U.T. The many facets of H/ACA ribonucleoproteins. *Chromosoma* **114**, 1–14 (2005).
11. Yu, Y.T., Terns, R.M. & Terns, M.P. Fine-tuning of RNA functions by modification and editing. in *Topics in Current Genetics* (ed. Grosjean, H.) 223–262 (Springer, New York, USA, 2005).
12. Balakin, A.G., Smith, L. & Fournier, M.J. The RNA world of the nucleolus: two major families of small RNAs defined by different box elements with related functions. *Cell* **86**, 823–834 (1996).
13. Ganot, P., Caizergues-Ferrer, M. & Kiss, T. The family of box ACA small nucleolar RNAs is defined by an evolutionarily conserved secondary structure and ubiquitous sequence elements essential for RNA accumulation. *Genes Dev.* **11**, 941–956 (1997).
14. Watkins, N.J. *et al.* Cbf5p, a potential pseudouridine synthase, and Nhp2p, a putative RNA-binding protein, are present together with Gar1p in all H BOX/ACA-motif snoRNPs and constitute a common bipartite structure. *RNA* **4**, 1549–1568 (1998).
15. Henras, A. *et al.* Nhp2p and Nop10p are essential for the function of H/ACA snoRNPs. *EMBO J.* **17**, 7078–7090 (1998).
16. Lafontaine, D.L., Bousquet, A.C., Henry,Y., Caizergues, F.M. & Tollervey, D. The box H^+ ACA snoRNAs carry Cbf5p, the putative rRNA pseudouridine synthase. *Genes Dev.* **12**, 527–537 (1998).
17. Morrissey, J.P. & Tollervey, D. Yeast snR30 is a small nucleolar RNA required for 18S rRNA synthesis. *Mol. Cell. Biol.* **13**, 2469–2477 (1993).
18. Eliceiri, G.L. The vertebrate E1/U17 small nucleolar ribonucleoprotein particle. *J. Cell. Biochem.* **98**, 486–495 (2006).
19. Mitchell, J.R., Cheng, J. & Collins, K. A box H/ACA small nucleolar RNA-like domain at the human telomerase RNA 3′ end. *Mol. Cell. Biol.* **19**, 567–576 (1999).
20. Collins, K. The biogenesis and regulation of telomerase holoenzymes. *Nat. Rev. Mol. Cell Biol.* **7**, 484–494 (2006).
21. Lukowiak, A.A., Narayanan, A., Li, Z.H., Terns, R.M. & Terns, M.P. The snoRNA domain of vertebrate telomerase RNA functions to localize the RNA within the nucleus. *RNA* **7**, 1833–1844 (2001).
22. Heiss, N.S. *et al.* X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions. *Nat. Genet.* **19**, 32–38 (1998).
23. Walne, A.J. *et al.* Genetic heterogeneity in autosomal recessive dyskeratosis congenita with one subtype due to mutations in the telomerase-associated protein NOP10. *Hum. Mol. Genet.* **16**, 1619–1629 (2007).
24. Vulliamy, T. *et al.* Mutations in the telomerase component NHP2 cause the premature ageing syndrome dyskeratosis congenita. *Proc. Natl. Acad. Sci. USA* **105**, 8073–8078 (2008).
25. Spedaliere, C.J., Ginter, J.M., Johnston, M.V. & Mueller, E.G. The pseudouridine synthases: revisiting a mechanism that seemed settled. *J. Am. Chem. Soc.* **126**, 12758–12759 (2004).

26. Hamma, T. & Ferre-D'Amare, A.R. Pseudouridine synthases. *Chem. Biol.* **13**, 1125–1135 (2006).
27. Phannachet, K., Elias, Y. & Huang, R.H. Dissecting the roles of a strictly conserved tyrosine in substrate recognition and catalysis by pseudouridine 55 synthase. *Biochemistry* **44**, 15488–15494 (2005).
28. Roovers, M. *et al.* Formation of the conserved pseudouridine at position 55 in archaeal tRNA. *Nucleic Acids Res.* **34**, 4293–4301 (2006).
29. Schattner, P. *et al.* Genome-wide searching for pseudouridylation guide snoRNAs: analysis of the *Saccharomyces cerevisiae* genome. *Nucleic Acids Res.* **32**, 4281–4296 (2004).
30. Decatur, W.A. & Fournier, M.J. rRNA modifications and ribosome function. *Trends Biochem. Sci.* **27**, 344–351 (2002).
31. Baker, D.L. *et al.* RNA-guided RNA modification: functional organization of the archaeal H/ACA RNP. *Genes Dev.* **19**, 1238–1248 (2005).
32. Wang, C. & Meier, U.T. Architecture and assembly of mammalian H/ACA small nucleolar and telomerase ribonucleoproteins. *EMBO J.* **23**, 1857–1867 (2004).
33. Charpentier, B., Muller, S. & Branlant, C. Reconstitution of archaeal H/ACA small ribonucleoprotein complexes active in pseudouridylation. *Nucleic Acids Res.* **33**, 3133–3144 (2005).
34. Li, L. & Ye, K. Crystal structure of an H/ACA box ribonucleoprotein particle. *Nature* **443**, 302–307 (2006).
35. Li, H. Unveiling substrate RNA binding to H/ACA RNPs: one side fits all. *Curr. Opin. Struct. Biol.* **18**, 78–85 (2008).
36. Liang, B., Xue, S., Terns, R.M., Terns, M.P. & Li, H. Substrate RNA positioning in the archaeal H/ACA ribonucleoprotein complex. *Nat. Struct. Mol. Biol.* **14**, 1189–1195 (2007).
37. Rashid, R. *et al.* Crystal structure of a Cbf5-Nop10-Gar1 complex and implications in RNA-guided pseudouridylation and dyskeratosis congenita. *Mol. Cell* **21**, 249–260 (2006).
38. Ye, K. H/ACA guide RNAs, proteins and complexes. *Curr. Opin. Struct. Biol.* **17**, 287–292 (2007).
39. Hamma, T., Reichow, S.L., Varani, G. & Ferre-D'Amare, A.R. The Cbf5-Nop10 complex is a molecular bracket that organizes box H/ACA RNPs. *Nat. Struct. Mol. Biol.* **12**, 1101–1107 (2005).
40. Manival, X. *et al.* Crystal structure determination and site-directed mutagenesis of the *Pyrococcus abyssi* aCBF5-aNOP10 complex reveal crucial roles of the C-terminal domains of both proteins in H/ACA sRNP activity. *Nucleic Acids Res.* **34**, 826–839 (2006).
41. Liang, B. *et al.* Long-distance placement of substrate RNA by H/ACA proteins. *RNA* **14**, 2086–2094 (2008).
42. Klein, R.J., Misulovin, Z. & Eddy, S.R. Noncoding RNA genes identified in AT-rich hyperthermophiles. *Proc. Natl. Acad. Sci. USA* **99**, 7542–7547 (2002).
43. Rozhdestvensky, T.S. *et al.* Binding of L7Ae protein to the K-turn of archaeal snoRNAs: a shared RNA binding motif for C/D and H/ACA box snoRNAs in Archaea. *Nucleic Acids Res.* **31**, 869–877 (2003).
44. Wu, H. & Feigon, J. H/ACA small nucleolar RNA pseudouridylation pockets bind substrate RNA to form three-way junctions that position the target U for modification. *Proc. Natl. Acad. Sci. USA* **104**, 6655–6660 (2007).
45. Jin, H., Loria, J.P. & Moore, P.B. Solution structure of an rRNA substrate bound to the pseudouridylation pocket of a box H/ACA snoRNA. *Mol. Cell* **26**, 205–215 (2007).
46. Hoang, C. & Ferre-D'Amare, A.R. Cocrystal structure of a tRNA Psi55 pseudouridine synthase: nucleotide flipping by an RNA-modifying enzyme. *Cell* **107**, 929–939 (2001).
47. Spedaliere, C.J. & Mueller, E.G. Not all pseudouridine synthases are potently inhibited by RNA containing 5-fluorouridine. *RNA* **10**, 192–199 (2004).
48. Hoang, C. *et al.* Crystal structure of pseudouridine synthase RluA: indirect sequence readout through protein-induced RNA structure. *Mol. Cell* **24**, 535–545 (2006).
49. Baker, D.L. *et al.* Determination of protein-rna interaction sites in the Cbf5-H/ACA Guide RNA complex by mass spectrometric protein footprinting. *Biochemistry* **47**, 1500–1510 (2008).
50. Muller, S., Fourmann, J.B., Loegler, C., Charpentier, B. & Branlant, C. Identification of determinants in the protein partners aCBF5 and aNOP10 necessary for the tRNA:{Psi}55-synthase and RNA-guided RNA:{Psi}-synthase activities. *Nucleic Acids Res.* **35**, 5610–5624 (2007).
51. Darzacq, X. *et al.* Stepwise RNP assembly at the site of H/ACA RNA transcription in human cells. *J. Cell Biol.* **173**, 207–218 (2006).

ONLINE METHODS

Protein and RNA preparation. We purified Pf H/ACA proteins as described[37] with slight modification. Briefly, we purified Pf Nop10 and Pf Cbf5 as a binary complex by a nickel–nitrilotriacetic acid (Ni-NTA) affinity and a gel filtration procedure. We purified Pf L7Ae separately by treating cell supernatant with polyethyleneimine and precipitating with ammonium sulfate, followed by a Ni-NTA affinity purification and a gel filtration procedure. The proteins were concentrated and stored at –80 °C before crystallization. The full-length Pf9_Pf6 composite guide RNA was transcribed *in vitro* using T7 RNA polymerase and purified as described[52]. The substrate RNA with sequence 5′-GAUGGAGCG(f^5U)GCGGUUUAAUG-3′ was purchased from Dharmacon (Thermo Fisher Scientific) and purified and stored according to manufacturer instructions.

Crystallization and diffraction data. For CNL complex crystallization, we mixed the guide and target RNA at a 1:1 molar ratio and annealed them by heating the solution for 10 min at 70 °C followed by slow cooling to room temperature (25 °C). The RNA–protein complex was formed at a 1:1.2 molar ratio with a total concentration of 23 mg ml^{-1}. We carried out crystallization using vapor diffusion methods in hanging drops against a reservoir of 0.8 M KCl, 0.15 M magnesium acetate, 0.05 M sodium cacodylate, pH 6.5, 8% (w/v) PEG 6000. Crystals, which grew to full size (0.3 mm × 0.3 mm × 0.4 mm) at 30 °C within 1 week, were soaked briefly in a solution containing 0.8M KCl, 0.15M magnesium acetate, 0.05M sodium cacodylate, pH 6.5, 8% (w/v) PEG 6000 and 5% (v/v) glycerol, followed by the same solution with 10% (v/v) glycerol, before being flash cooled in a liquid nitrogen stream for data collection. The crystals (cell parameters in **Table 1**) of the CNL complex contained one RNP in each asymmetric unit, with a solvent content of 61.7%.

For CN complex crystallization, the two guide strands and the target RNA were annealed at a 1:1:1 molar ratio. After mixing proteins and RNAs at a 1:1.2 ratio with a total concentration of 18 mg ml^{-1}, the full-size crystals (0.1 mm × 0.2 mm × 0.3 mm) were obtained by vapor diffusion against a reservoir solution of 0.05M sodium cacodylate, pH 6.5, 2.0 mM $CoCl_2$, 30 mM $CaCl_2$, 2.0 mM spermine, 2.0 M LiCl at 30 °C within 45 d. The crystals were soaked in a mother liquor plus 2.5 M LiCl for 5–8 h before being mounted in nylon loops and flash cooled in a liquid nitrogen stream.

Diffraction data were collected at beamlines 22ID and 22BM of the South Eastern Consortium Access Team (SER-CAT) at the Advanced Photon Source (APS) and were processed using HKL2000 (ref. 53).

Phase determination and structure refinement. We determined both structures by molecular replacement methods using Molrep[54]. For the CNL complex structure, the coordinates of Cbf5, Nop10 and L7Ae from the substrate-minus structure (PDB 2HVY)[38] were used as a search model. A single and outstanding solution was found in space group $P2_12_12$. The initial solution was subjected to successive rigid body, energy minimization and simulated annealing refinement using CNS[55]. Electron density computed using the protein coordinates was improved by solvent flattening. A molecular mask generated using a manually constructed RNA-protein complex was used to perform the density modification. At this stage, most RNA nucleotides for the entire guide RNA and partial substrate RNA could be built unambiguously based into the electron density map. We carried out further refinement using CNS[55], REFMAC5 (ref. 56) and manual model building by O[57] and COOT[58] iteratively until the complete model for the RNA-protein complex could no longer be improved. At the final stage of refinement, ten cycles of translation-libration-screw motion (TLS) refinement[59] were performed in addition to restrained refinement in REFMAC5. Each individual protein and RNA molecule was treated as a single 'rigid-body' group, and the final TLS parameters are listed in **Supplementary Table 2**. The real-space correlation coefficient between the final model and the composite omit $3F_o - 2F_c$ map are 0.888, 0.881 and 0.776 for proteins, RNA and water/ions, respectively. The final model was refined to R_{free} 24.8% and R_{work} 21.7% and has an r.m.s. deviation of 0.010 Å and 1.397° from ideal bond lengths and angles, respectively. Of the protein residues, 92.5% lie in the most favored regions of the Ramachandran plot, 7.0% in additional allowed regions, 1 residue (Lys40) in the generously allowed region, and one residue in disallowed regions (Glu97), as similarly observed in the high-resolution structure of the Cbf5–Nop10–Gar1 protein complex[37].

For the CN subcomplex structure, the coordinates of Cbf5 and Nop10 (PDB 2EY4)[37] were used as search models. A single outstanding solution was obtained in the $P6_122$ space group. The RNA nucleotides were built from the lower stem gradually until clear density was available for the remaining structures. Simulated annealing was carried out using only the torsion angle refinement option in CNS[55]. Similarly to what was done for the CNL complex structure, at the final stage of the refinement, we performed ten cycles of TLS refinement in addition to restrained refinement in REFMAC5 (ref. 59). Individual protein and RNA molecules were treated as single 'rigid-body' groups, and the final TLS parameters are listed in **Supplementary Table 2**. The final model was refined to R_{free} 30.6% and R_{work} 27.5%.

Fluorescence studies. The fluorescence assay and data fitting have been described[41]. The wild-type and mutant guide RNAs were transcribed and purified similarly to those used for the crystallographic studies. The 2AP-labeled and 2′-substituted substrate RNAs were purchased from Dharmacon (Thermo Fisher Scientific). Titration curves were obtained in triplicate, from which the s.d. of the K_d was computed.

Pseudouridylation assay. The Pf9_Pf6 composite wild-type and mutant guide RNAs and substrate RNAs were transcribed *in vitro* with T7 RNA polymerase. The same conditions were used for synthesis of uniformly labeled substrate RNAs, except that 6 μCi of [α-^{32}P]UTP (3,000 Ci $mmol^{-1}$) (MP Biomedicals) was added to label all six uridines in the substrate RNA. We performed pseudouridylation assays in a similar manner to those described[31]. Briefly, we incubated 0.2 nM [α-^{32}P]-labeled substrate RNA, 1.2 μM guide RNA and 3 μM of indicated protein components in the reaction buffer containing 100 mM Tris-HCl, pH 8.0, 100 mM ammonium acetate, 5 mM $MgCl_2$, 2 mM DTT, 0.1 mM EDTA for 1 h at 70 °C. The RNAs were extracted by phenol/chloroform/isoamyl alcohol (pH4.5), purified by ethanol precipitation and digested with nuclease P1 (1 unit, United States Biological). The resulting 5′-mononucleotides were separated via thin-layer chromatography, as described[31].

52. McKenna, S.A. *et al.* Purification and characterization of transcribed RNAs using gel filtration chromatography. *Nat. Protocols* **2**, 3270–3277 (2007).
53. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
54. Vagin, A. & Teplyakov, A. MOLREP: an automated program for molecular replacement. *J. Appl. Crystallogr.* **30**, 1022–1025 (1997).
55. Brünger, A.T. *et al.* Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**, 905–921 (1998).
56. Murshudov, G.N., Vagin, A.A., Lebedev, A., Wilson, K.S. & Dodson, E.J. Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallogr. D Biol. Crystallogr.* **55**, 247–255 (1999).
57. Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard, M. Improved methods for binding protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).
58. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
59. Winn, M.D., Isupov, M.N. & Murshudov, G.N. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr. D Biol. Crystallogr.* **57**, 122–133 (2001).

Supplementary Information

"Structure of a functional ribonucleoprotein pseudouridine synthase bound with substrate RNA"

Bo Liang[1], Jing Zhou[2], Elliot Kahen[2], Rebecca M. Terns[3], Michael P. Terns[3], & Hong Li[1,2]

[1]*Institute of Molecular Biophysics,* [2]*Department of Chemistry and Biochemistry, Florida State University, Tallahassee, Florida.*

[3]*Department of Biochemistry and Molecular Biology, University of Georgia at Athens, Athens, Georgia*

Correspondence and requests for materials should be addressed to H.L.

Email: hong.li@fsu.edu Phone: (850) 644-6785 Fax: (850) 644-7244

Nature Structural & Molecular Biology: doi:10.1038/nsmb.1624

Table S1. Dissociation Constants of Substrate Placement

substrate	proteins*	K_d (nM)
amPf9s	*L* + CGN	63 ± 10
amPf9s	*L* + CN	43 ± 2
amPf9s	*L* + C_{loop}GN	221 ± 47
amPf9s	*L* + C_{loop}N	2.8 ± 0.4
inGPf9s	*CGN* + L	380 ± 21

*Proteins in italic fonts are present at a constant 10:1 stoichiometric concentration to RNA while the other protein or protein complex is titrated. Nomenclatures follow those in Figure 4.

Figure S1.

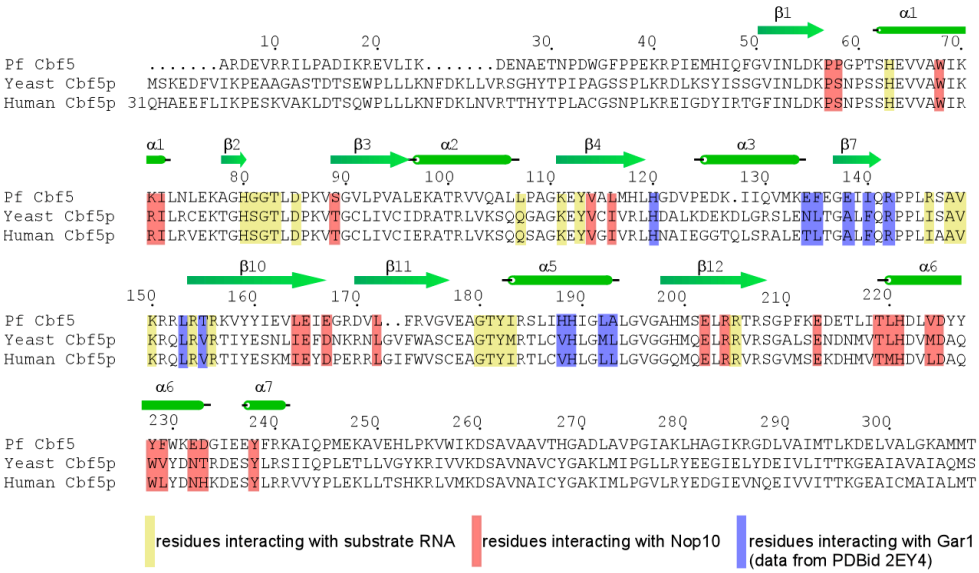

Figure S1 legend. Summary of Cbf5 residues involved in protein-protein and protein-RNA interactions shown with aligned Cbf5 sequences. Note regions where Cbf5 interacts with both substrate RNA and Gar1 (β10) and where Cbf5 interacts with both Nop10 and the substrate RNA (β4 and β12).

Figure S2

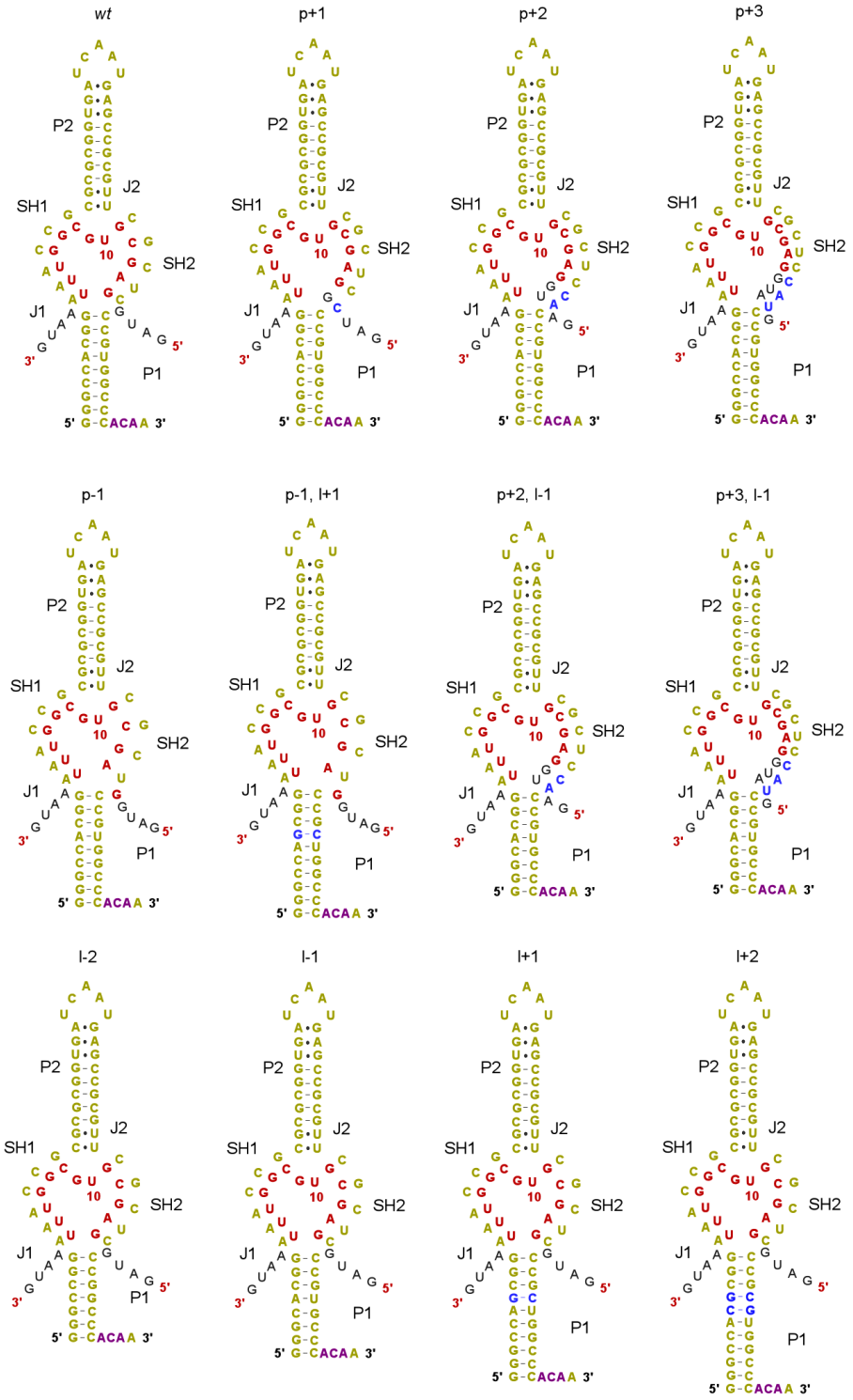

Figure S2 legend. Secondary structures of the composition Pf9_Pf6 guide RNA and its mutants used in pseudouridylation assay.

Table S2. Refined TLS parameters

CNL+g+s crystals

Group 1. Cbf5	T(Å^2)	0.0113	-0.067	-0.064	0.0241	0.022	-0.002		
	L(°^{2})	0.6756	0.5139	0.2299	-0.312	-0.339	0.1892		
	S(Å°)	0.054	0.0356	-0.066	0.0397	0.0078	-0.052	0.0044	0.0131
Group 2. Nop10	T(Å^2)	-0.056	0.0256	-0.053	0.0682	-0.02	0.0069		
	L(°^{2})	1.5331	1.1212	5.3779	-0.625	-1.274	1.6375		
	S(Å°)	0.2887	-0.077	-0.388	-0.085	-0.231	-0.032	0.0197	0.272
Group 3. L7Ae	T(Å^2)	-0.165	0.252	-0.126	0.0887	-0.016	0.1426		
	L(°^{2})	4.5663	3.1417	2.9337	-0.804	0.4893	0.6912		
	S(Å°)	0.1277	-0.236	-0.273	-0.138	-0.373	0.0891	-0.014	0.5612
Group 4. guide RNA	T(Å^2)	-0.039	0.0639	-0.015	0.0887	0.0555	0.0826		
	L(°^{2})	0.8938	0.8352	0.0444	-0.426	0.0772	-0.191		
	S(Å°)	0.0326	0.041	-0.24	-0.214	0.2809	0.0749	-0.065	0.0564
Group 5. substrate RNA	T(Å^2)	-0.009	-0.005	-0.007	-0.002	0.0479	-3E-04		
	L(°^{2})	9.2047	2.9819	2.3742	0.8398	-0.964	-2.658		
	S(Å°)	-0.244	0.0814	-0.85	-0.514	0.0782	-0.138	-0.373	0.0385
Group 6. ions	T(Å^2)	0.0002	0.0002	0	0.0009	-3E-04	-3E-04		
	L(°^{2})	2.8094	2.0492	2.4596	1.5741	-0.203	1.5754		
	S(Å°)	0.258	0.114	-0.058	-0.221	0.6495	-0.344	-0.924	0.0686
Group 7. water	T(Å^2)	0.0729	-0.017	-0.022	0.0187	0.0632	0.0291		
	L(°^{2})	1.3678	0.8984	0.6643	-0.551	-0.122	0.1254		
	S(Å°)	0.0535	0.0629	-0.164	0.0066	0.0467	-0.056	-0.099	0.0203

CN +g+s crystals

Group 1. Cbf5	T(Å^2)	0.1911	-0.111	-0.016	-0.07	0.0264	0.1036		
	L(°^{2})	1.0264	3.8365	1.5393	-1.254	0.0114	1.743		
	S(Å°)	-0.402	-6E-04	-0.34	-0.404	0.2813	0.3159	0.4491	-0.318
Group 2. Nop10	T(Å^2)	-0.05	-0.034	-0.228	0.0361	0.0452	-0.056		
	L(°^{2})	1.4062	2.7992	10.271	0.3171	-3.75	0.0132		
	S(Å°)	0.0764	-0.635	-0.26	-0.086	-0.075	-0.231	-0.129	0.1354
Group 3. 5' guide RNA	T(Å^2)	0.2518	-0.42	-0.018	0.2554	-0.12	0.1627		
	L(°^{2})	4.7505	12.725	2.9772	-1.604	1.0614	-2.035		
	S(Å°)	-1.038	0.5399	1.2291	0.1199	0.1109	-1.788	0.5789	-0.572
Group 4. 3' guide RNA	T(Å^2)	0.3491	-0.508	0.2996	0.1187	-0.137	-0.095		
	L(°^{2})	1.0503	5.2839	1.6471	1.5111	-0.562	1.2368		
	S(Å°)	-1.382	0.109	0.5706	-0.56	1.3033	-1.11	0.4562	-0.487
Group 5. substrate RNA	T(Å^2)	0.0184	0.0449	0.5705	0.1582	-0.433	-0.279		
	L(°^{2})	5.0715	23.933	5.7953	-2.944	4.0126	-9.961		
	S(Å°)	-0.776	1.9514	1.1251	-0.72	2.6641	-2.257	0.4488	0.4963